IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 17, 2018, accepted September 29, 2018, date of publication October 30, 2018,
date of current version December 27, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2878636

Probabilistic Load Flow With Wind Farms Using a Frequency and Duration Method

JINZHOU ZHU[1], YAN ZHANG[1], AND HAIBO CHEN[2,3]

[1]Department of Electrical Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2]State Grid Shanghai Municipal Power Company, Shanghai 200122, China
[3]Shanghai University of Electric Power, Shanghai 200090, China

Corresponding author: Jinzhou Zhu (ZJZ13595608186@sjtu.edu.cn)

This work was supported in part by the National High-Tech R&D Program (the 863 Program) under Grant 2015AA050203 and in part by the State Grid Science and Technology Project of China under Grant 520900150037.

ABSTRACT Probabilistic load flow (PLF) is an important tool in power system planning and operation. One limitation of conventional PLF is that only the probability information of random variables is obtained as a reference for related analyses. Frequency and duration information often plays an important role in power system assessment. In this paper, a frequency and duration method for PLF with wind farms (WFs) is proposed based on Markov chains by using an improved probability-frequency distribution function (PFDF) method. Random input variables, including intermittent loads, conventional generator (CG) power outputs associated with CG failures, and WF power outputs associated with both wind speed uncertainties and wind turbine failures, are modeled using corresponding PFDFs. With the proposed method, not only probability information but also frequency and duration information of random PLF outputs are efficiently and analytically computed through the operations of PFDFs of random inputs. Moreover, an optimal decision-making model for determining the clustering number of random states is proposed to improve the credibility of stochastic process modeling of Markov-chain-based random variables. The performance of the proposed method is verified and compared with that of a sequential Monte Carlo simulation technique using two modified IEEE test systems.

INDEX TERMS Frequency and duration method, Markov chain, probabilistic load flow (PLF), probability-frequency distribution function (PFDF), wind farm.

I. INTRODUCTION

Power flow computation is one of the major tasks facing power system planners and operators. Traditional deterministic load flow (DLF) finds nodal voltages and branch flows only under specified operating conditions. Random factors, such as intermittent loads, conventional generator (CG) failures (or deratings) and power outputs of renewable energy sources, bring great challenges to power system planning and operation. This is why the importance of probabilistic tools for power system analysis is growing. Compared with DLF, probabilistic load flow (PLF) [1]–[12] can provide more comprehensive information regarding a power system with random factors. For example, PLF can be used to assess adequacy indexes such as the probability of a nodal voltage being outside acceptable levels and the probability of a branch flow being greater than its thermal rating, which are extremely useful in the planning and operation of power systems. From the first proposal in the 1970s [1], numerous methodologies [2]–[12] have been presented for PLF to determine the probability information (such as probability distributions, expected values (EVs), and standard deviations) of desired variables (such as node voltages and branch flows) in this problem.

Many PLF formulations use linear load flow equations with which the output variables can be represented as a linear combination of input variables [1]–[8]. With a linearization method, non-sequential Monte Carlo simulation (NSMCS) methods [2] and analytic methods, including the convolution method [1], fast Fourier transform method [3], and cumulant-based method [4], are commonly used to solve PLF problems. In [5] and [6], certain enhancements to the linearization method for PLF were introduced to accommodate the spatial correlation between random input variables. Reference [7] extended the linearization method for PLF to address the nonlinear effects of load flow equations through multilinearization. In [8], random branch outages were incorporated

into the PLF. Moreover, the point estimate method [9], [10] and Latin hypercube sampling method [11], [12] also deserve special mention in the context of PLF studies. Furthermore, some methods proposed for stochastic load flow [13], [14], which is a subject related to PLF, can also be applied to PLF problems.

However, one limitation of most existing PLF methods [1]–[14] is that only the probability information of desired variables is obtained as a reference for related analyses. Compared with probability information, frequency and duration information can provide more insight into the sequential behavior of random variables. Frequency and duration concepts were generally used in the field of power system reliability evaluation [15]. In these concepts, the term ''frequency'' is used to denote the rate of encountering a system state or subset of states, while the term ''duration'' is used to denote the mean residence time in a state or subset of states in one cycle. In practice, frequency and duration concepts can be used in a wide range of power system applications. For example, whether a variable violation (such as violation of voltage or branch flow) occurs once every day (frequency information) with a ten-cycle duration (duration information) or occurs ten times a day with a one-cycle duration makes a difference [16], [17] even though the calculated probabilities of the violation are identical. In addition, the direction of power flow may change when considering random factors. The change frequency of power flow direction and the average duration of each change are intuitively available to both system planners and operators. Although the frequency and duration information of random variables in a power system has showed practical value in many aspects, only a few PLF studies [17] for voltage stability assessments have considered it. Meanwhile, wind energy has been the fastest growing source of energy among renewable energy sources in the past decade. Since the great expansion of wind power in power networks has increased the uncertainty in power systems, many aspects of wind power penetration in PLF considering the intermittent nature of wind speed have been addressed recently [9]–[12], [17]–[21]. However, the power output of a wind farm (WF) is related not only to the intermittent wind speed but also to the probabilistic behavior of wind turbine (WT) failures (or deratings) [22]–[25]. This condition increases the complexity of integrating wind power into power systems in term of randomness modeling of WF power outputs, particularly when the frequency and duration information is considered.

A method used to calculate frequency and duration quantities is generally referred to as a frequency and duration method. In addition to [17], there were some other frequency and duration methods for the assessments of voltage stability [16] and the reliability evaluation of power systems [22]–[28], which may be applied to the PLF with frequency and duration quantities. Among these methods, sequential Monte Carlo simulation (SMCS) [26]–[28] is a traditional straightforward method to solve the frequency and duration problem by simulating random factors chronologically, but this method requires large computational effort to achieve accurate results. Moreover, SMCS based on a time-series method [26], [27], such as the autoregressive moving average (ARMA) model, requires actual hourly data (such as wind speed data and load data) collected over many years for the particular geographic location to construct a simulation model for the specific WF or load; however, long-history time-series data are typically unavailable [27]. NSMCS is an alternative simulation method for PLF. The detailed chronological nature needed for SMCS to describe random factors is not required in NSMCS, and each random factor can be represented by a probability distribution function (PDF). For example, wind speed probability distributions are generally represented using Weibull distributions. Therefore, NSMCS proves useful for cases lacking adequate historical data. However, NSMCS is not suitable for frequency and duration problems since it does not consider the sequential characteristics of random factors.

The other technique widely used to solve frequency and duration problems is the Markov chain. A Markov chain can be used to model the variations of a stochastic process as Markov state transitions in which sequential characteristics are contained. Moreover, a Markov chain requires far less data for acceptable randomness modeling than the time-series-based simulation method [24]. With consideration of frequency and duration characteristics, various analytic methods based on Markov chains have been proposed for voltage stability assessments [16], [17] and power outputs modeling of WFs [22]–[25]. In [16], the intermittent nature of load was modeled using a two-state Markov chain, and the frequency and duration quantities of voltage drops in plants with a large number of intermittent loads were calculated by combining Markov chains of various loads. In [17], a similar approach was proposed for PLF with WFs. In addition to the load, a multistate Markov chain was used to model the WF power output considering the intermittent nature of wind speed. By combining Markov chains of all loads and all WFs, the probability and the frequency and duration of the power system operation states with voltage violations were obtained to conduct the planning study. As mentioned above, the randomness modeling of WF power outputs associated with both wind speed uncertainties and WT failures deserves attention in the context of PLF with WFs. Certain Markov models for WFs have been previously proposed in [22]–[25]. In [22]–[24], a multistate Markov chain was used to obtain the probability and the frequency and duration characteristics of wind speed based on past wind speed observations, and WT failure was represented using a two-state Markov chain. The Markov chain of WF power output was constructed by combining these two types of Markov chains. However, if the number of random variables (such as loads, WFs, and WTs in large-scale WFs) becomes large and the derated states of units are considered [25], the randomness problem in [16], [17], and [22]–[24] presents challenges when using a full Markov model because the dimension of the state space becomes extremely large. Consequently, equivalent Markov

models were further developed in [16], [17], [22], and [24] to improve the efficiency of analysis. In these equivalent models, the number of Markov states was reduced by clustering exact states into various approximate states based on engineering Markov analysis. However, these state-clustering methods are conducted by keeping track of all Markov states and the corresponding transition rates, which can create significant computational complexity when the number of Markov states is large. In [25], multistate Markov chains for WFs and CGs were presented with the consideration of unit deratings. These models were integrated into the universal generating function, which provides a practical approach to determine the probability information of a power system. However, the frequency and duration quantities were not considered.

On the basis of Markov chains, the concepts of state probability and incremental frequency were introduced in [29] to characterize the cumulative probability and frequency of a conventional generation system. With these concepts, the probability-frequency distribution function (PFDF) was presented in [30] and [31] to represent the random model of the CG, in addition to the load, which provides an efficient method to calculate the probability and the frequency and duration reliability indexes for a generation system. However, the traditional PFDF method is not directly suitable for the PLF problem. Certain improvements for the traditional PFDF method were introduced in [32] to solve PLF problems with wind power considering wind speed uncertainties. However, the probabilistic behavior of WT failures (or deratings) was not considered in [32], mainly due to the challenges associated with randomness modeling of large WFs. Additionally, Markov chain modeling is based on the hypothesis that the distribution of residence time in any state is exponential [22]–[25], [32]. However, in the past, the credibility of the exponential hypothesis has rarely been considered when modeling sequential events of Markov-chain-based random variables. The Kolmogorov-Smirnov (KS) test was used in [33] to evaluate the credibility of Markov-chain-based modeling of CG deratings. That test, however, is a post test; thus, the credibility of the Markov chain is challenging to guarantee or actively improve.

The main contributions of the present study for addressing the issues mentioned above are identified as follows.

1) A frequency and duration method for PLF with WFs is proposed based on Markov chains by using an improved PFDF method. Random models of different input variables in PLF problems are developed and represented as corresponding PFDFs. With the proposed method, not only the probability information but also the frequency and duration information of random output variables in PLF problems can be efficiently and analytically computed through the operations of PFDFs of random input variables.

2) A random model for large WFs is developed and represented as PFDF in which both wind speed uncertainties and WT deratings are considered. With the proposed method, the randomness modeling of large WFs becomes simple to implement, and the PFDFs of corresponding WFs are integrated with the PFDFs of conventional elements (including CG and load) to obtain the probability and the frequency and duration quantities in PLF problems.

3) An optimal decision-making model for determining the clustering number of random states is presented to improve the credibility of stochastic process modeling of Markov-chain-based random variables.

The rest of the paper is organized as follows. Section II introduces the PLF problem. Section III presents the optimal decision-making model for the modeling of Markov-chain-based random variables and proposes the improved PFDF method. The randomness modeling of WF power output is described in Section IV. The application of the proposed PFDF method to the PLF problem is presented in Section V. In Section VI, the proposed method is tested on the modified IEEE-RTS79 and IEEE-300 node test grid with WFs, and the results are compared to those of SMCS. Conclusions are presented in Section VII. An appendix is included to provide details regarding WF modeling.

II. PLF PROBLEM

A. LOAD FLOW EQUATIONS

The AC load flow equations for a power system can be written as

$$\begin{cases} \boldsymbol{I}_{AR} = g(\boldsymbol{U}) \\ \boldsymbol{T} = h(\boldsymbol{U}) \end{cases} \tag{1}$$

where $\boldsymbol{I}_{AR}$ is the input vector of active and reactive power injections (such as load demands and power generations of CGs and WFs). $\boldsymbol{U}$ is the output vector of the nodal voltage magnitude U_M and angle. $\boldsymbol{T}$ is the output vector of the active branch flow T_P, and reactive branch flow T_Q. $g(\boldsymbol{U})$ and $h(\boldsymbol{U})$ are nonlinear functions.

In PLF studies, power injections $\boldsymbol{I}_{AR}$ are treated as random input variables. The task of PLF analysis is to characterize the random behavior of output variables $\boldsymbol{U}$ and $\boldsymbol{T}$ from the random information of $\boldsymbol{I}_{AR}$. It is preferable to apply linear approximation to (1) so that the output variables can be solved as a linear combination of input variables, which allows us to solve PLF through fast methods [1]–[8] (such as the convolution method or cumulant-based method). By linearizing (1) around the EV region, the output variables can be represented as [18], [19]

$$\boldsymbol{U}-\boldsymbol{U}_0=\boldsymbol{C}(\boldsymbol{A}-\boldsymbol{A}_0)\Rightarrow\boldsymbol{U}=\boldsymbol{C}\boldsymbol{A}+(\boldsymbol{U}_0-\boldsymbol{C}\boldsymbol{A}_0)=\boldsymbol{U}_L+\boldsymbol{U}_C \tag{2}$$

$$\boldsymbol{T}-\boldsymbol{T}_0=\boldsymbol{Z}(\boldsymbol{A}-\boldsymbol{A}_0)\Rightarrow\boldsymbol{T}=\boldsymbol{Z}\boldsymbol{A}+(\boldsymbol{T}_0-\boldsymbol{Z}\boldsymbol{A}_0)=\boldsymbol{T}_L+\boldsymbol{T}_C \tag{3}$$

where $\boldsymbol{A}$ is the input vector of nodal active power injections. $\boldsymbol{A}_0$ is the EV of $\boldsymbol{A}$. $\boldsymbol{U}_0$ and $\boldsymbol{T}_0$ are the values of $\boldsymbol{U}$ and $\boldsymbol{T}$, respectively, provided by substituting $\boldsymbol{A}_0$ into (1). $\boldsymbol{C}=\boldsymbol{J}_{\boldsymbol{P}}^{-1}$, and $\boldsymbol{Z}=\boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{J}_{\boldsymbol{P}}^{-1}$. For AC load flow equations, $\boldsymbol{J}_{\boldsymbol{h}}$ is the Jacobian matrix that relates $\boldsymbol{T}$ to $\boldsymbol{U}$, and $\boldsymbol{J}_{\boldsymbol{P}}$ is the Jacobian matrix that relates $\boldsymbol{U}$ to $\boldsymbol{A}$. $\boldsymbol{U}_L=\boldsymbol{C}\boldsymbol{A}$ and $\boldsymbol{T}_L=\boldsymbol{Z}\boldsymbol{A}$ are vectors represented as the linear combination of $\boldsymbol{A}$. $\boldsymbol{U}_C=\boldsymbol{U}_0-\boldsymbol{C}\boldsymbol{A}_0$

and $\boldsymbol{T}_C = \boldsymbol{T}_0 - \boldsymbol{Z}\boldsymbol{A}_0$ are vectors of constants corresponding to EV. Reactive power injections are considered to be linearly related to active power injections, since most modern WTs can control the power factor of the WF according to economic incentives and the power factor of loads can be considered constant [18], [19]. Alternatively, if the WF has a variable output power factor that tracks the connection node, then the node with the WF can be considered a PV node with the assumption that the WF has its own voltage control system, and new sensitivity matrixes $\boldsymbol{C}$ and $\boldsymbol{Z}$ can be given accordingly.

In (2) and (3), the sensitivity matrixes $\boldsymbol{C}$ and $\boldsymbol{Z}$ are obtained with the assumption that the swing bus absorbs all changes due to uncertainties of nodal power injections. If dispatch strategies for combined operation of several generators are considered, the linear dispatching model proposed in [18] and [19] can be used. Then, taking $\boldsymbol{C}$ as an example, the new sensitivities can be easily calculated as

$$C'_{mn} = C_{mn} - \sum_{i=1}^{M} k_{in} C_{mi} \tag{4}$$

where C_{mn} is the term (m, n) of the sensitivity matrix $\boldsymbol{C}$. M is the number of nodal injections. k_{in} is the participation factor of the regulating generator in node i to compensate the power injection of node n. Thus, we obtain

$$\boldsymbol{U}_L = \boldsymbol{C}'\boldsymbol{A}, \quad \boldsymbol{U}_C = \left(\boldsymbol{U}_0 - \boldsymbol{C}'\boldsymbol{A}_0\right) \tag{5}$$

where $\boldsymbol{C}'$ is the new sensitivity matrix with elements C'_{mn}.

Similarly, new sensitivities for $\boldsymbol{Z}'$ can also be obtained.

B. SOLUTION METHOD

In most existing PLF studies, random input variables are typically modeled using PDFs (e.g., load demands are commonly modeled as normal variables [5]–[8], and the Weibull distribution is generally used to describe the probabilistic behavior of wind speed [9]–[11], [19], [20]), and multiple calculation methods [1]–[14], [18]–[20] have been proposed for PLF to determine the probability information of desired variables based on these known PDFs. However, the frequency and duration information of variables has rarely been considered. Relative to probability quantities, frequency and duration quantities can provide more information regarding system conditions. Since it is inadequate to represent the sequential characteristics (such as frequency and duration characteristics) of random variables by PDFs alone, suitable random models for associated variables and a corresponding calculation method are desired to solve the PLF problem with frequency and duration quantities.

A Markov chain can perfectly replicate the probability characteristics and the frequency and duration characteristics found in the samples [22]–[25], [29]–[33] and requires far less data for acceptable randomness modeling than the time-series-based simulation method [24]. As a result, stochastic processes of different random variables, such as the load [16], [17], [32], WF output [22]–[25], and CG output [29]–[33], are commonly modeled using corresponding Markov chains. On the basis of Markov chains, an improved PFDF method is introduced to solve the PLF problem with frequency and duration quantities. In the proposed method, the random models for WFs, CGs, and loads are developed and represented as corresponding PFDFs. In addition, high-credibility modeling of Markov-chain-based random variables is desired for the application of the PFDF method, which is also discussed in this study.

III. MARKOV CHAIN AND PFDF

A. RANDOMNESS MODELING USING MARKOV CHAIN

1) BASIC PARAMETERS OF A MARKOV CHAIN

In the real world, every physical event that advances continuously and randomly in time and space can be modeled approximately as a process with a continuous parameter (time) space and a discrete state space [24]. A Markov chain can be used to model the variations of a stochastic process as transitions between Markov states, with each state representing a discrete value. Modeling a stochastic process by a Markov chain requires that the state residence time follows an exponential distribution [22]–[25]. With the hypothesis that the state residence time of random variable X follows an exponential distribution, the transition rate between states i and j of X, namely, λ_{Xij}, can be given based on past observations [24]:

$$\lambda_{Xij} = \frac{N_{Xij}}{D_{Xi}} \tag{6}$$

where N_{Xij} is the number of observed transitions from state i to state j of X, and D_{Xi} is the duration for state i of X calculated during the entire period.

Let λ_{Xi} denote the departure rate for state i of X:

$$\lambda_{Xi} = \sum_{j \neq i, j \in \boldsymbol{B}_X} \lambda_{Xij} \tag{7}$$

where $\boldsymbol{B}_X$ is the state set of X.

The probability of occurrence for state i of X, namely, p_{Xi}, is given by

$$p_{Xi} = \frac{D_{Xi}}{\sum\limits_{k \in \boldsymbol{B}_X} D_{Xk}} \tag{8}$$

2) OPTIMIZATION FOR CLUSTERING NUMBER

Because of the large number of variable states presented in the observations, the representation of all of them in a model can become unfeasible. Therefore, in the Markov modeling of a random variable, it is necessary to answer the following questions: (a) *how are the Markov states chosen* and (b) *how many states do we need in each model?*

To choose the Markov states with discrete values, a technique for clustering observations into a smaller number of states shall be adopted. The aim is to allocate elements of certain common characteristics in clusters, where the step length of discretization may be any desirable or available

unit [22]–[25], [29]–[33]. The *k-means* clustering algorithm is an effective method for splitting data [23]. With the *k-means* method, the clustering number can be adjusted conveniently. In this study, for random variable X, the *k-means* clustering algorithm is used to find C_X centroids based on observations of X. These centroids are the required C_X states of the Markov chain of X, and their values act as quantization levels for X and provide a quantized data set. This quantized data set is then used to derive the transition rate and state probability through (6)–(8). In most of the previous studies [22]–[25], [29]–[33], the number of Markov states was arbitrary or depended only on the required accuracy of the model. In general, *k-means* uses squared Euclidean distances. With the state probability considered, the clustering accuracy index J_X for X is given by

$$J_X = \sum_{i\in \boldsymbol{B}_X} \sum_{x_l \in X_i} p_{Xi} \left(x_l - X_i\right)^2 \tag{9}$$

where $\boldsymbol{B}_X = \{1, 2, \ldots, C_X\}$ is the state set of X and C_X is the clustering number of random states of X. X_i is the centroid of the i-th cluster of X and also denotes the value of X in state i. x_l is the l-th sample in the observations of X, and $x_l \in X_i$ means that x_l belongs to the i-th cluster of X. According to (9), the smaller J_X is, the smaller the average within-cluster sum of sample-to-centroid distance for X is.

Markov-chain-based modeling of a stochastic process is based on the hypothesis that the state residence time follows an exponential distribution. In previous studies, however, the credibility of the exponential hypothesis was either entirely ignored or simply checked using a post test [33]. Therefore, the credibility of stochastic process modeling of Markov-chain-based random variables is challenging to guarantee or actively improve. With the concept of significance level in the KS test [34], we propose an index H_X for measuring the credibility of the hypothesis that the state residence time of X follows an exponential distribution. H_X is defined as

$$H_X = \frac{1}{C_X} \sum_{i\in \boldsymbol{B}_X} (1 - \alpha_i) \tag{10}$$

where α_i is the significance level for the exponential distribution hypothesis of residence time in state i of X. The detailed calculation method for α_i can be found in [34]. Let $N_{Xi} = \sum_{j\neq i, j\in \boldsymbol{B}_X} N_{Xij}$ denote the number of samples used to construct the probability distribution of the residence time in state i of X. When N_{Xi} is sufficiently large, then α_i can be approximately calculated as

$$\alpha_i = \sum_{j=1}^{+\infty} 2(-1)^{j-1} \exp\left(-2N_{Xi} j^2 K_i^2\right) \tag{11}$$

$$K_i = \max_t \left(\left|I_{Xi\_1}(t) - I_{Xi\_2}(t)\right|\right) \tag{12}$$

where $I_{Xi\_1}(t) = \text{Prob}(t_{Xi} \leq t)$ is the empirical probability that the residence time t_{Xi} in state i of X is less than or equal to t, which is calculated using a statistical technique based on the observations of X. $I_{Xi\_2}(t) = 1 - \exp(-\lambda_{Xi} \cdot t)$ is an exponential distribution function with parameter λ_{Xi} given by (7). According to (10)–(12) and [34], the smaller H_X is, the higher the credibility of exponential hypothesis for X is.

To improve the accuracy and credibility of the stochastic process modeling of any Markov-chain-based random variable X, an optimal decision-making model aiming the minimization of J_X and H_X is proposed to determine the number of Markov states and to answer question (b) mentioned at the beginning of this section. Because J_X and H_X have different dimensions, the two-objective optimal decision-making model for determining the clustering number C_X of X is formulated as (13) based on entropy weight [35].

$$\min_{C_X\in \boldsymbol{C}_{FX}} \beta_J \frac{J_X(C_X)}{\max\limits_{C_X\in \boldsymbol{C}_{FX}} (J_X(C_X))} + \beta_H \frac{H_X(C_X)}{\max\limits_{C_X\in \boldsymbol{C}_{FX}} (H_X(C_X))} \tag{13}$$

where $\boldsymbol{C}_{FX}$, representing the range of C_X, can be pre-set according to the observations of X. β_J and β_H are weighting factors for J_X and H_X, respectively. In this study, β_J and β_H are given by the entropy weight method [35].

B. PFDF

1) TRADITIONAL PFDF

On the basis of Markov chains, the incremental frequency of state i of X, namely, f_{Xi_IN}, can be given by [29], [30]

$$\begin{cases} f_{Xi\_IN} = \sum\limits_{j\in \boldsymbol{B}_{Xi}^U} f_{Xij} - \sum\limits_{k\in \boldsymbol{B}_{Xi}^L} f_{Xki}, \quad \forall i \in \boldsymbol{B}_X \\ f_{Xij} = p_{Xi}\lambda_{Xij}, \quad f_{Xki} = p_{Xk}\lambda_{Xki} \end{cases} \tag{14}$$

where $\boldsymbol{B}_{Xi}^U$ and $\boldsymbol{B}_{Xi}^L$ ($\boldsymbol{B}_{Xi}^U$, $\boldsymbol{B}_{Xi}^L \subseteq \boldsymbol{B}_X$) are the upper and lower state set for state i of X, respectively, which means that $X_j > X_i$ if $j \in \boldsymbol{B}_{Xi}^U$ and that $X_k < X_i$ if $k \in \boldsymbol{B}_{Xi}^L$.

The traditional PFDF of X is defined as [30]

$$\boldsymbol{S}_X^0 = \left\{\boldsymbol{X}, \quad \boldsymbol{p}_X, \quad \boldsymbol{f}_{X\_IN}\right\} \tag{15}$$

where $\boldsymbol{X}$ is the vector of X_i ($i \in \boldsymbol{B}_X$). $\boldsymbol{p}_X$ is the vector of p_{Xi} and can be regarded as the PDF of X. $\boldsymbol{f}_{X\_IN}$ is the vector of f_{Xi_IN} and can be regarded as a frequency distribution function (FDF) of X. X_i, p_{Xi}, and f_{Xi_IN} are of one-to-one correspondence.

If random variables X and Y are independent of each other and $Z = X + Y$, then $\boldsymbol{S}_Z^0$ can be determined by integrating $\boldsymbol{S}_X^0$ with $\boldsymbol{S}_Y^0$ as follows:

$$\boldsymbol{S}_Z^0 = \boldsymbol{S}_X^0 * \boldsymbol{S}_Y^0 = \left\{\boldsymbol{Z}, \quad \boldsymbol{p}_Z, \quad \boldsymbol{f}_{Z\_IN}\right\} \tag{16}$$

$$\boldsymbol{p}_Z = \boldsymbol{p}_X * \boldsymbol{p}_Y \tag{17}$$

$$\boldsymbol{f}_{Z\_IN} = \left(\boldsymbol{p}_X * \boldsymbol{f}_{Y\_IN}\right) \oplus \left(\boldsymbol{p}_Y * \boldsymbol{f}_{X\_IN}\right) \tag{18}$$

where $*$ denotes the convolution operation. $\oplus$ denotes the addition operation of functions, which means the addition of corresponding items of functions rather than the addition of vectors. $\boldsymbol{Z}$ is given through the convolution with the consideration of one-to-one correspondence among $\boldsymbol{Z}$, $\boldsymbol{p}_Z$, and $\boldsymbol{f}_{Z\_IN}$.

2) IMPROVED PFDF

In PLF using a linearization method, the associated linear combinations (i.e., (2) and (3)) are more complex than that in (16). Therefore, certain improvements for the traditional PFDF method are essential for solving PLF problems [32].

First, the decremental frequency of state i of X, namely, f_{Xi_DE}, is defined as

$$f_{Xi\_DE} = \sum_{j \in \boldsymbol{B}_{Xi}^{L}} f_{Xij} - \sum_{k \in \boldsymbol{B}_{Xi}^{U}} f_{Xki}, \quad \forall i \in \boldsymbol{B}_X \tag{19}$$

For X, the improved PFDF, which is referred to simply as PFDF in the rest of this study, is defined as

$$\boldsymbol{S}_X = \left\{ \boldsymbol{X}, \quad \boldsymbol{p}_X, \quad \boldsymbol{f}_{X\_IN}, \quad \boldsymbol{f}_{X\_DE} \right\} \tag{20}$$

where $\boldsymbol{f}_{X\_DE}$, i.e., the vector of f_{Xi_DE} corresponding to X_i, can be regarded as the other FDF of X.

The rounding technique in [30] is used in this study to process the PFDF of each random variable after the discrete step length is determined; consequently, $\boldsymbol{X}$, $\boldsymbol{p}_X$, $\boldsymbol{f}_{X\_IN}$, and $\boldsymbol{f}_{X\_DE}$ can be approximated by the rounded discrete values. Furthermore, certain operations are defined. If a random variable $Z = aX$ where a is any real number, then the PFDF of Z can be given by

$$\boldsymbol{S}_Z = \boldsymbol{S}_{aX} = \left\{ a\boldsymbol{X}, \quad \boldsymbol{p}_X, \quad \boldsymbol{f}_X(a), \quad \boldsymbol{f}_X(-a) \right\} \tag{21}$$

where $\boldsymbol{f}_X(n) = \boldsymbol{f}_{X\_IN}$ if $n \geq 0$ and $\boldsymbol{f}_X(n) = \boldsymbol{f}_{X\_DE}$ if $n < 0$.

If $Z = X + a$, then $\boldsymbol{S}_Z$ can be given by

$$\boldsymbol{S}_Z = \boldsymbol{S}_{X+a} = \left\{ \boldsymbol{X} + a\boldsymbol{E}, \quad \boldsymbol{p}_X, \quad \boldsymbol{f}_{X\_IN}, \quad \boldsymbol{f}_{X\_DE} \right\} \tag{22}$$

where $\boldsymbol{E}$ is a vector of ones.

If variables X and Y are independent of each other and $Z = aX + bY$, where b is another real number, then $\boldsymbol{S}_Z$ is given by

$$\boldsymbol{S}_Z = \boldsymbol{S}_{aX} * \boldsymbol{S}_{bY} = \left\{ \boldsymbol{Z}, \quad \boldsymbol{p}_Z, \quad \boldsymbol{f}_{Z\_IN}, \quad \boldsymbol{f}_{Z\_DE} \right\} \tag{23}$$

where $\boldsymbol{p}_Z$ is given by (17), while $\boldsymbol{f}_{Z\_IN}$ and $\boldsymbol{f}_{X\_DE}$ are given as follows:

$$\boldsymbol{f}_{Z\_IN} = \left[\boldsymbol{p}_X * \boldsymbol{f}_Y(b)\right] \oplus \left[\boldsymbol{p}_Y * \boldsymbol{f}_X(a)\right] \tag{24}$$

$$\boldsymbol{f}_{Z\_DE} = \left[\boldsymbol{p}_X * \boldsymbol{f}_Y(-b)\right] \oplus \left[\boldsymbol{p}_Y * \boldsymbol{f}_X(-a)\right] \tag{25}$$

Equations (23)–(25) indicate that the computation for $\boldsymbol{S}_Z$ involves only convolution and addition operations of corresponding distribution functions, which is straightforward to implement. With the fast Fourier transform method [3], the computation burden of convolution can be greatly relieved.

From $\boldsymbol{S}_X$, the probability PR_X and frequency FR_X of specified events of X that occur during the assessment period can be efficiently obtained as follows:

$$PR_X^{up}(x) = \text{Prob}(X \geq x) = \sum_{X_i \geq x} p_{Xi} \tag{26}$$

$$PR_X^{lb}(x) = \text{Prob}(X \leq x) = \sum_{X_i \leq x} p_{Xi} \tag{27}$$

$$FR_X^{up}(x) = \text{Freq}(X \geq x) = \sum_{X_i \geq x} f_{Xi\_DE} \tag{28}$$

$$FR_X^{lb}(x) = \text{Freq}(X \leq x) = \sum_{X_i \leq x} f_{Xi\_IN} \tag{29}$$

where the superscript *up* (or *lb*) denotes the situations in which X is greater (or less) than a certain value. For example, $PR_X^{up}(x)$ denotes the probability of X being greater than or equal to x, and $FR_X^{lb}(x)$ denotes the frequency of the events causing X to be less than or equal to x.

Then, the average duration of corresponding events of X, namely, DU_X, can be given by

$$DU_X^{lb(up)}(x) = \frac{PR_X^{lb(up)}(x)}{FR_X^{lb(up)}(x)} \tag{30}$$

Notably, the sum of $\boldsymbol{p}_X$ should equal unity, and the sum of either $\boldsymbol{f}_{X\_IN}$ or $\boldsymbol{f}_{X\_DE}$ should equal zero. Since the events $X \leq x$ and $X \geq x$ are nearly complementary and frequency balanced, we note that $PR_X^{up}(x) + PR_X^{lb}(x) \approx 1$ and $FR_X^{up}(x) \approx FR_X^{lb}(x)$.

C. RANDOMNESS MODELING OF TRADITIONAL INPUT VARIABLES

1) CG POWER OUTPUT

Random CG failures (or deratings) are generally considered to be power injection uncertainties in PLF studies [1]–[8]. A CG can be modeled using a two-state or multistate Markov chain depending on whether CG deratings are taken into account [29]–[33]. Let G_{mn} denote the active power output of the n-th CG at node m. Based on the generating capacity process observations associated with the availability of the corresponding CG, the PFDF of G_{mn}, namely, $\boldsymbol{S}_{G_{mn}}$, can be given by the method described in Sections III-A and B. Alternatively, this PFDF can be given based on the reliability data of the corresponding CG when observations are unavailable. In general, the failures of different CGs are considered to be independent. Let G_m denote the total active power output of CGs at node m. The PFDF of G_m can be given by (31) based on (23).

$$\boldsymbol{S}_{G_m} = \left(\left(\boldsymbol{S}_{G_{m1}} * \boldsymbol{S}_{G_{m2}} \right) * \cdots \right) * \boldsymbol{S}_{G_{mN_{Gm}}} \tag{31}$$

where N_{Gm} is the number of CGs at node m.

2) LOAD DEMAND

In this study, the load series of the IEEE-RTS79 grid [36] are used as load observations for each load bus independently. The measurement interval is one hour, with a registration of one year. The number of years is not sufficient to use time-series-based simulation methods [26], [27] such as the ARMA model. A Markov chain is commonly used to model the stochastic processes of load demand [16], [17], [32]. Based on load observations, the Markov chain of the load series is constructed through the method proposed in Section III-A. Let L_m denote the load demand at node m. With the method presented in Section III-B, the PFDF

of L_m, namely, $\boldsymbol{S}_{L_m}$, can be given based on the corresponding Markov parameters.

IV. PFDF-BASED MODELING OF A WF

The randomness modeling of a WF is crucial in the context of PLF with WFs. The active power output of a WF depends on three factors: failures (or deratings) of WTs, wind speed, and the relationship between the WT power output and wind speed (i.e., the power curve). Therefore, the randomness modeling of the WF power output is different from that of a CG. To solve the PLF problem with frequency and duration quantities, a random model for large WFs is developed and represented as PFDF.

A. INFLUENCE FACTORS OF A WF POWER OUTPUT

1) WT FAILURES

A WF consists of multiple WTs that are subject to the same wind regime. The failures of WTs affect the power output of the WF. In most research works [22]–[24], a two-state Markov chain is used to represent the failure behavior of a WT. However, certain minor failures of the WT may lead to situations in which the WT continues to operate but at reduced performance (i.e., deratings). Multistate Markov chains of WTs are considered in recent research [25].

On the condition that wind speed is constant and rated, a single WT is similar to a CG with derated states associated with failures [25]. Let W_{mrj} denote the failure-associated active power output of the j-th WT of type r in the WF connected at node m. The PFDF of W_{mrj}, namely, $\boldsymbol{S}_{W_{mrj}}$, can be given by the same approach as that outlined in the modeling of $\boldsymbol{S}_{G_{mn}}$. In general, the failures of different WTs are considered to be independent. Let W_{mr} denote the total active power output of type-r WTs in the WF at node m. The PFDF of W_{mr} can be given by (32) based on (23).

$$\boldsymbol{S}_{W_{mr}} = \left(\left(\boldsymbol{S}_{W_{mr1}} * \boldsymbol{S}_{W_{mr2}}\right) * \cdots\right) * \boldsymbol{S}_{W_{mrN_r}} \tag{32}$$

where N_r is the number of type-r WTs in this WF.

2) WIND SPEED

In this study, the hourly wind speed series from Washington during 2010 [37] are used as wind speed observations for each WF independently. Similar to the load series mentioned above, the time span of the wind speed observations is insufficiently lengthy to use time-series-based simulation methods. A Markov chain is commonly used to model the stochastic processes of wind speed [22]–[25]. To solve the PLF problem with frequency and duration quantities, the Markov chain of wind speed V_m for the WF at node m is constructed through the method presented in Section III-A based on the wind speed observations. Since the wind speed itself is not directly involved in the PLF analysis, the PFDF of wind speed is not essential.

3) POWER CURVE

The power curve can vary depending on the type of WT. The power curve for a WT of type r can be characterized by

$$P_{W\_r}(V) = \begin{cases} 0, & 0 \le V < V_{CI\_r},\ V \ge V_{CO\_r} \\ \left(A_r + B_r V + C_r V^2\right) P_{WN\_r}, & V_{CI\_r} \le V < V_{N\_r} \\ P_{WN\_r}, & V_{N\_r} \le V < V_{CO\_r} \end{cases} \tag{33}$$

where $P_{W\_r}(V)$ is the active power output of a type-r WT under wind speed V. P_{WN_r}, V_{CI_r}, V_{N_r}, and V_{CO_r} are the rated active power output, cut-in wind speed, rated wind speed, and cut-out wind speed for the WT of type r, respectively. The constants A_r, B_r, and C_r are calculated from V_{CI_r}, V_{N_r}, and V_{CO_r} [38].

B. PFDF OF THE WF POWER OUTPUT

1) WITHOUT WT FAILURES CONSIDERED

For the WF at node m, let V_{mi} stand for the i-th state value of the Markov chain of wind speed V_m modeled in Section IV-A-2), and let F_m^0 denote the total active power output of this WF without WT failures considered. If R types of WTs exist in this WF, then the state value of F_m^0 under V_{mi}, namely, F_{mi}^0, is given by

$$F_{mi}^0 = \sum_{r=1}^{R} N_r P_{W\_r}(V_{mi}), \quad i \in \boldsymbol{B}_{V_m} \tag{34}$$

where $\boldsymbol{B}_{V_m}$ is the state set of V_m.

In practice, R and N_r ($r = 1, 2, \ldots, R$) in (34) are constants. Therefore, F_{mi}^0 and V_{mi} are of one-to-one correspondence. By integrating (34) with the Markov chain of V_m, the PFDF of F_m^0 can be expressed as

$$\boldsymbol{S}_{F_m^0} = \left\{ \boldsymbol{F}_m^0, \quad \boldsymbol{p}_{F_m^0}, \quad \boldsymbol{f}_{F_{m\_IN}^0}, \quad \boldsymbol{f}_{F_{m\_DE}^0} \right\} \tag{35}$$

2) WITH WT FAILURES CONSIDERED

On the condition that V_m is fixed at V_{mi} ($\forall i \in \boldsymbol{B}_{V_m}$) and only the failures of WTs are considered, let F_m^i denote the total active power output of the WF at node m. Since a WF consists of multiple types of WTs that are subject to the same wind regime, F_m^i ($\forall i \in \boldsymbol{B}_{V_m}$) is the parallel composition of failure-associated active power output of each type of WT in the corresponding WF. Consequently, the PFDF of F_m^i can be given by (36) based on (23), (32), and (33).

$$\begin{aligned} \boldsymbol{S}_{F_m^i} &= \left(\left(\boldsymbol{S}_{b_{1i}W_{m1}} * \boldsymbol{S}_{b_{2i}W_{m2}}\right) * \cdots\right) * \boldsymbol{S}_{b_{Ri}W_{mR}} \\ &= \left\{ \boldsymbol{F}_m^i, \quad \boldsymbol{p}_{F_m^i}, \quad \boldsymbol{f}_{F_{m\_IN}^i}, \quad \boldsymbol{f}_{F_{m\_DE}^i} \right\}, \quad i \in \boldsymbol{B}_{V_m} \end{aligned} \tag{36}$$

where $\boldsymbol{S}_{W_{mr}}$ is given by (32), and $b_{ri} = P_{W\_r}(V_{mi})/P_{WN\_r}$ which can be obtained from (33).

Equation (36) is formulated by assuming that if the failure-associated Markov chain of a WT under V_{N_r} is given as Fig. 1(a), then the failure-associated Markov chain of this WT

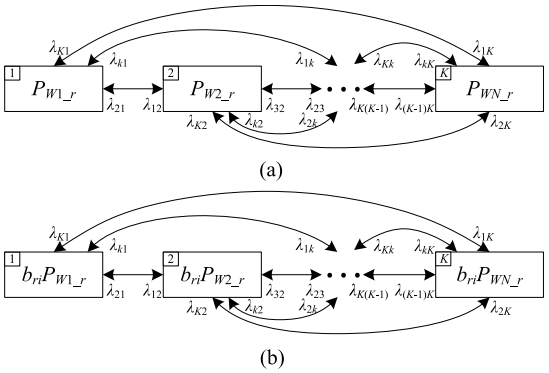

FIGURE 1. Failure-associated Markov chain of a type-*r* WT under different wind speeds. (a) Failure-associated Markov chain of a type-*r* WT under V_{N_r}. (b) Failure-associated Markov chain of a type-*r* WT under V_{mi}.

under V_{mi} can be given as Fig. 1(b) [25]. In Fig. 1, there are K Markov states for the WT, and $P_{W1\_r}, P_{W2\_r}, \ldots, P_{W(K-1)\_r}$ are the derated outputs of this WT in various extents of failures (the K-th state is the rated state).

For F_m^i ($i \in \boldsymbol{B}_{V_m}$), a group of PFDFs exists, each representing the failure-associated WF active power output and corresponding probability-frequency distribution under the selected wind speed V_{mi}. In contrast, there is only one PFDF for F_m^0 that represents the wind-speed-associated WF power output without WT failures considered. Let F_m denote the total active power output of the WF at node m, with all influence factors described in Section IV-A considered. By integrating $\boldsymbol{S}_{F_m^0}$ with $\boldsymbol{S}_{F_m^i}$ ($i \in \boldsymbol{B}_{V_m}$), the PFDF of F_m can be given by (37)–(40) based on (35) and (36).

$$\boldsymbol{S}_{F_m} = \left\{ \boldsymbol{F}_m, \quad \boldsymbol{p}_{F_m}, \quad \boldsymbol{f}_{F_m\_IN}, \quad \boldsymbol{f}_{F_m\_DE} \right\} \tag{37}$$

$$\boldsymbol{p}_{F_m} = \sum_{i \in \boldsymbol{B}_{V_m}} \oplus \left(p_{F_{mi}^0} \boldsymbol{p}_{F_m^i} \right) \tag{38}$$

$$\boldsymbol{f}_{F_m\_IN} = \sum_{i \in \boldsymbol{B}_{V_m}} \oplus \left(p_{F_{mi}^0} \boldsymbol{f}_{F_m^i\_IN} \oplus f_{F_{mi}^0\_IN} \boldsymbol{p}_{F_m^i} \right) \tag{39}$$

$$\boldsymbol{f}_{F_m\_DE} = \sum_{i \in \boldsymbol{B}_{V_m}} \oplus \left(p_{F_{mi}^0} \boldsymbol{f}_{F_m^i\_DE} \oplus f_{F_{mi}^0\_DE} \boldsymbol{p}_{F_m^i} \right) \tag{40}$$

where $p_{F_{mi}^0}$, $f_{F_{mi}^0\_IN}$, and $f_{F_{mi}^0\_DE}$ are the elements of $\boldsymbol{p}_{F_m^0}$, $\boldsymbol{f}_{F_{m}^0\_IN}$, and $\boldsymbol{f}_{F_{m}^0\_DE}$ corresponding to V_{mi}, respectively. $\sum\oplus$ denotes the multiple addition operations of functions. A detailed proof procedure for (37)–(40) is presented in the Appendix.

Equations (37)–(40) combine the $\boldsymbol{S}_{F_m^0}$ and $\boldsymbol{S}_{F_m^i}$ ($i \in \boldsymbol{B}_{V_m}$) into one PFDF (i.e., $\boldsymbol{S}_{F_m}$) in which the uncertainties of wind speed and WT failure behavior are both considered. Notably, the computation for $\boldsymbol{S}_{F_m}$ in (38)–(40) involves only addition and scalar multiplication operations of associated PDFs and FDFs, which simplifies the randomness modeling of the WF power output with frequency and duration characteristics considered.

V. PLF ANALYSIS

A. APPLICATION OF THE PFDF METHOD TO PLF

Uncertainties of nodal active power injections considered in this study include random loads and active power outputs of CGs and WFs. The total active power injection at node m is

$$A_m = G_m + F_m - L_m, \quad \forall m \tag{41}$$

According to (23), the PFDF of A_m can be expressed as

$$\boldsymbol{S}_{A_m} = \left(\boldsymbol{S}_{G_m} * \boldsymbol{S}_{F_m} \right) * \boldsymbol{S}_{(-1)L_m}, \quad \forall m \tag{42}$$

Then, the EV of $\boldsymbol{A}$ (i.e., $\boldsymbol{A}_0$ in (2) and (3)) can be calculated using $\boldsymbol{S}_{A_m}$, and the linear load flow equations (2)–(5) can be obtained by linearizing (1) around the EV region.

Equations (2)–(5) state that $\boldsymbol{U}_L = \boldsymbol{C}\boldsymbol{A}$ (or $\boldsymbol{U}_L = \boldsymbol{C}'\boldsymbol{A}$) and $\boldsymbol{T}_L = \boldsymbol{Z}\boldsymbol{A}$ (or $\boldsymbol{T}_L = \boldsymbol{Z}'\boldsymbol{A}$). Let U_{Lm} denote the m-th element of $\boldsymbol{U}_L$, and let T_{Ll} stand for the l-th element of $\boldsymbol{T}_L$. According to (23), the PFDFs of U_{Lm} and T_{Ll} can be calculated using $\boldsymbol{S}_{A_m}$ as follows:

$$\boldsymbol{S}_{U_{Lm}} = \left(\left(\boldsymbol{S}_{C_{m1}A_1} * \boldsymbol{S}_{C_{m2}A_2} \right) * \cdots \right) * \boldsymbol{S}_{C_{mM}A_M}, \quad \forall m \tag{43}$$

$$\boldsymbol{S}_{T_{Ll}} = \left(\left(\boldsymbol{S}_{Z_{l1}A_1} * \boldsymbol{S}_{Z_{l2}A_2} \right) * \cdots \right) * \boldsymbol{S}_{Z_{lM}A_M}, \quad \forall l \tag{44}$$

where C_{mn} ($n = 1, 2, \ldots, M$) are the m-th row elements of $\boldsymbol{C}$ (or $\boldsymbol{C}'$). Z_{ln} ($n = 1, 2, \ldots, M$) are the l-th row elements of $\boldsymbol{Z}$ (or $\boldsymbol{Z}'$).

Equations (2) and (3) state $U_m = U_{Lm} + U_{Cm}$ and $T_l = T_{Ll} + T_{Cl}$, where U_{Cm} and T_{Cl} are the elements of $\boldsymbol{U}_C$ and $\boldsymbol{T}_C$ corresponding to U_m and T_l, respectively. According to (22), $\boldsymbol{S}_{U_m}$ and $\boldsymbol{S}_{T_l}$ can be given as

$$\boldsymbol{S}_{U_m} = \boldsymbol{S}_{U_{Lm}+U_{Cm}} \tag{45}$$

$$\boldsymbol{S}_{T_l} = \boldsymbol{S}_{T_{Ll}+T_{Cl}} \tag{46}$$

Consequently, the probability and the frequency and duration information of specified events of nodal voltages $\boldsymbol{U}$ and branch flows $\boldsymbol{T}$ can be obtained from $\boldsymbol{S}_{U_m}$ and $\boldsymbol{S}_{T_l}$ according to (26)–(30).

The flowchart of the proposed PLF method in this study is shown in Fig. 2.

B. SPATIAL CORRELATION AND NONLINEAR EFFECT

The calculations in (43) and (44) assume spatial independence between components of $\boldsymbol{A}$. In practice, both the correlation of the random variables among one another for a given time instant (spatial correlation) and the correlation over time (temporal correlation) may exist at the same time. In this study, the temporal correlation, which is essential for solving the PLF problem with frequency and duration quantities, is taken into account using different Markov chains to model corresponding random variables. If only linear spatial correlations among power injections $\boldsymbol{A}$ are considered, then the method proposed in [5] can be directly incorporated into the PFDF method to address the correlation problem. In many studies, various loads and WFs are either close or far away. Therefore, when the study period is short (e.g., in the

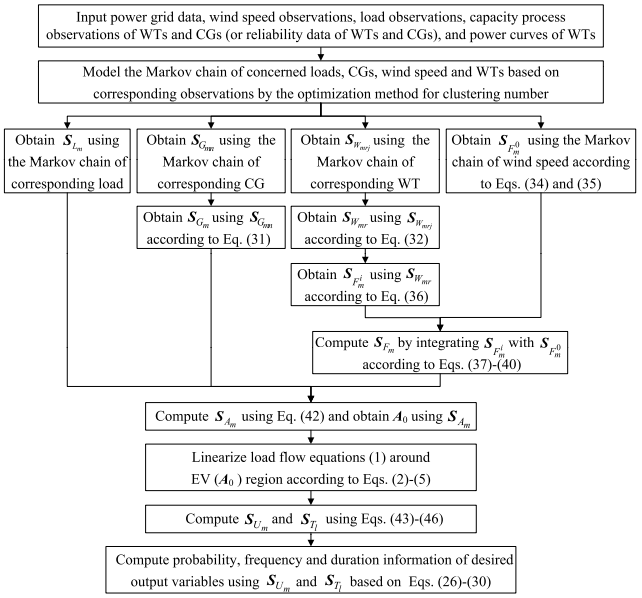

FIGURE 2. Flowchart of the proposed PLF method.

study of system operation), the assumption of spatial independence or linear spatial correlation among $\boldsymbol{A}$ is acceptable. Similar to the initial cumulant-based method [4], one drawback of the PFDF method is that it cannot directly incorporate variables with general (not only linear) spatial correlation. To overcome this drawback, an extension to the cumulant-based PLF problem is presented based on orthogonal transformation [6]. This extension can also be incorporated into the PFDF-based method to handle the correlation problem. First, let $F_m = F_m^0 – (F_{mi}^0 – F_m^i)$ (i.e., $A_m = G_m + F_m^0 – (F_{mi}^0 – F_m^i) – L_m$), where $(F_{mi}^0 – F_m^i)$ approximately represents the effect of WT failures and can be regarded as an independent injection with i (wind speed state) fixed. For example, $\boldsymbol{S}_{F_m^i}$ can be modeled with V_{mi} fixed at the EV of wind speed, and then the PFDF of $(F_{mi}^0 – F_m^i)$ can be given by (21) and (22) since F_{mi}^0 is a constant. In addition, the observations of F_m^0 can be mathematically obtained using (34) to transform associated wind speed observations. Then, the fictitious independent observations can be generated by an orthogonal transformation [6] for original correlated observations of injections (including L_m and F_m^0) at various nodes. In this manner, the PFDFs of fictitious independent random variables can be modeled using the method presented in Sections III-A and B. Finally, $\boldsymbol{S}_{U_m}$ and $\boldsymbol{S}_{T_l}$ can be obtained by the calculation of the PFDFs of fictitious independent random variables based on modified linear load flow equations [6] reflecting the impact of spatial correlations. Essentially, this orthogonalization approach [6] and the method in [5] model spatially correlated variables as a function of several independent ones. The spatial correlation is not included as part of the PFDF method itself; rather, spatial correlation is addressed in the variable modeling, which indicates the practicality of the proposed method itself.

The multilinearization method proposed in [7] can be incorporated into the proposed method to accommodate the nonlinear effects of the load flow equations. With the multilinearization method, various points of linearization can be determined, and the same linearization concept can be applied to such points in addition to EVs. Then, a combination of several PFDF solutions around various linearization points can ensure satisfactory performance in terms of computation time and accuracy.

Moreover, if the branch outages are considered, then the method proposed in [8] can be tailored to solve this problem by simulating branch outages as fictitious power injections. More details of the above extension methods can be found in [5]–[8]. Since the present study focuses on solving the frequency and duration problem in PLF and because space is limited, these extension methods are not discussed further.

VI. NUMERICAL TEST RESULTS

The performance of the proposed method was tested on the IEEE-RTS79 grid and IEEE-300 node test system. Random input variables, including intermittent loads, CG power outputs associated with CG failures, and WF power outputs associated with both wind speed uncertainties and WT failures, were considered in these tests. In all cases, the computations were performed on the MATLAB 7.1 platform using a 3.10-GHz Intel(R) Core(TM) i5-4440/4 GB RAM PC.

A. IEEE-RTS79 GRID

1) SYSTEM DATA

The proposed method was first tested on an IEEE-RTS79 grid [36] that had been modified to include WFs. In this test system, a swing bus absorbs all injection changes, i.e., sensitivity matrixes $\boldsymbol{C}$ and $\boldsymbol{Z}$ are used in the calculations of (43) and (44). A dispatching strategy was considered in the other test (see Section VI-B).

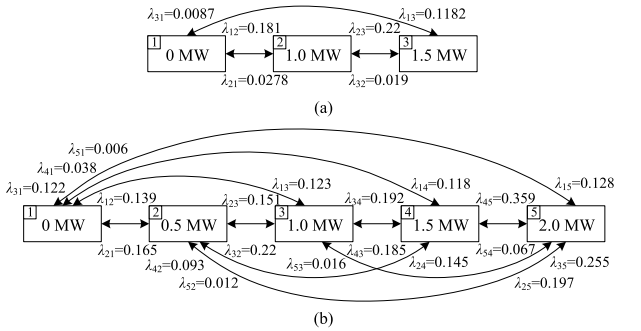

FIGURE 3. Failure-associated Markov chain of a single WT.
(a) Failure-associated Markov chain of a single WT of type 1.
(b) Failure-associated Markov chain of a single WT of type 2.

In the test, two types of WTs are considered. For the WT of type 1, $P_{WN\_1} = 1.5$ WM, $V_{CI\_1} = 4$ m/s, $V_{N\_1} = 15$ m/s, and $V_{CO\_1} = 25$ m/s. $P_{WN\_2} = 2$ WM, $V_{CI\_2} = 3$ m/s, $V_{N\_2} = 12$ m/s, and $V_{CO\_2} = 30$ m/s for the WT of type 2. Due to a lack of WT capacity process observations, the failure-associated Markov chains of a single WT of each type under V_{N_r} $(r = 1, 2)$ are assumed to be known as

shown in Fig. 3(a) and (b), in which the unit of transition rate is occurrence/hour. Two 100 MW CGs at bus 7 and one 350 MW CG at bus 23 are replaced by two WFs. These two WFs are connected at buses 3 and 8. The WF at bus 3 contains 200 WTs of type 1. The WF at bus 8 contains 100 WTs of each type (a total of 200 WTs). WT failures are assumed to be independent. To simplify the analysis, the power factor of each WF is assumed to be the same as that of the load at the corresponding bus. If the power factor of the WF is different from that of the corresponding load, a fictitious node can be added to take the power injection of this WF, and new sensitivity matrixes (i.e., $\boldsymbol{C}$ and $\boldsymbol{Z}$) can be obtained accordingly. Hourly wind speed observations measured in the State of Washington in the U.S. in 2010 [37] are used to model the power output of each WF independently.

Due to a lack of CG observations, a two-state model for each CG is used. The transition rates (occurrence/hour) between two states are assumed to be known as outage and repair rates, which are calculated by the mean time to failure (MTTF) and the mean time to repair (MTTR) [36], respectively. The system is assumed to be operating at the annual peak loading level in the base case, and hourly peak loads with 8736 sampling points expressed in terms of a percent of the annual peak load [36] are used as load observations for each load bus independently.

The actual value, instead of the per-unit value, is used in the PFDF calculation of random variables. The discrete step length in the PFDF calculation is set as $\Delta X = 0.1$ kV for the voltage magnitude and $\Delta X = 0.1$ MW (or Mvar) for the active (or reactive) power.

2) CLUSTERING RESULTS OF OPTIMAL DECISION-MAKING

For random variable X, the index J_X in (9) and the index H_X in (10) may vary with the clustering number C_X. For wind speed V (i.e., $X = V$), Fig. 4 shows the normalized J_V and H_V as a function of the clustering number C_V.

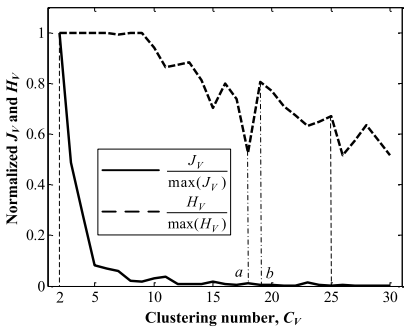

FIGURE 4. Normalized J_V and H_V with variations in C_V.

J_V decreases almost monotonically with increasing C_V, and the decrease in J_V becomes negligible when C_V increases to a certain value (e.g., the decrease of J_V can be ignored when $C_V > 20$). However, prominent fluctuations appear in H_V as C_V increases. Both values of J_V, corresponding to $C_V = a$ and $C_V = b$, are almost the same. However, H_V corresponding to $C_V = a$ is smaller than that corresponding to $C_V = b$, i.e., the value of C_V has a detectable influence on H_V even if its influence on J_V is slight. In the analysis of load L (i.e., $X = L$), a similar influence of C_L can also be found. The clustering number of random states has a considerable influence on the credibility of the Markov-chain-based modeling, and it is necessary to optimize this parameter.

Since the wind speed model is designed to be as simple as possible to improve computing efficiency, $C_V = 18$ can be given by solving (13) when setting $\boldsymbol{C}_{FV} = [2, 25]$. Similarly, $C_L = 168$ can be given for a load ($X = L$) when setting $\boldsymbol{C}_{FL} = [10, 200]$ in (13).

3) RESULTS OF PLF ANALYSIS

The PLF problem of (1)–(3) was solved using the proposed PFDF method. A SMCS with 10 years (10×8760 hours) of simulation time was used as a reference to indicate the accuracy and effectiveness of the proposed method. In the SMCS, the time series data of random input variables are generated through the random residence time sampling technique based on Markov chains [28] modeled above, and DLF is run for each sample.

Equations (26)–(30) indicate that $PR_X^{lb}(x)$, $PR_X^{up}(x)$, $FR_X^{lb}(x)$, $FR_X^{up}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ of random variable X may vary with the parameter x. For the active flow in branch 9–11 (i.e., $X = T_{P9-11}$), Figs. 5(a)–(d) depict the associated $PR_X^{lb}(x)$, $FR_X^{lb}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ as functions of parameter x, respectively. The dashed line represents the PFDF results, and the solid line represents the results obtained from SMCS. As stated in (30), the average duration is the ratio of the probability to the frequency. Therefore, the calculation error for $DU_X^{lb(up)}(x)$ may be amplified when the corresponding $FR_X^{lb(up)}(x)$ in the denominator is excessively small. Moreover, in practice, events with a sufficiently low frequency can be ignored. To limit the error amplification due to the division operation in (30), a lower limit of $R_F = 0.02$ is set for frequency. Only those $DU_X^{lb}(x)$ and $DU_X^{up}(x)$, corresponding to $FR_X^{lb}(x) \geq R_F$ and $FR_X^{up}(x) \geq R_F$ for parameter x, respectively, are calculated. Since the events $X \leq x$ and $X \geq x$ are nearly complementary and frequency balanced as mentioned in Section III-B-2), $PR_X^{up}(x)$ and $FR_X^{up}(x)$ can be derived from $PR_X^{lb}(x)$ and $FR_X^{lb}(x)$, respectively, i.e., $PR_X^{up}(x) \approx 1 - PR_X^{lb}(x)$, and $FR_X^{up}(x) \approx FR_X^{lb}(x)$. For the sake of clarity, $PR_X^{up}(x)$ and $FR_X^{up}(x)$ are not shown in the calculation results. Fig. 5 shows that the results obtained from the proposed PFDF method match closely with the corresponding SMCS results.

The randomness of loads and WF power outputs may cause the power flow to change direction. T_{P9-11} is in the positive direction when $T_{P9-11} \geq 0$, i.e., the active flow in branch 9–11 is from node 9 to node 11. Then, the probability, frequency, and average duration of the events leading to a positive T_{P9-11} can be given as $PR_X^{up}(0)$, $FR_X^{up}(0)$, and $DU_X^{up}(0)$ ($X = T_{P9-11}$), respectively. Accordingly, $PR_X^{lb}(0)$, $FR_X^{lb}(0)$, and $DU_X^{lb}(0)$ can provide associated information of the events causing $T_{P9-11} \leq 0$, which indicates that T_{P9-11} lies in the

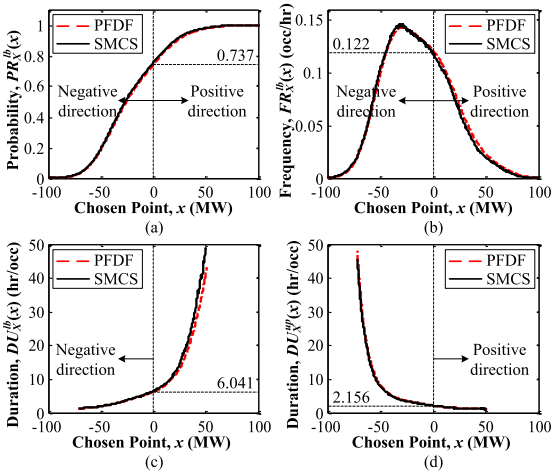

FIGURE 5. $PR_X^{lb}(x)$, $FR_X^{lb}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ of T_{P9-11} in the IEEE-RTS79 grid. (a) $PR_X^{lb}(x)$ of T_{P9-11}. (b) $FR_X^{lb}(x)$ of T_{P9-11}. (c) $DU_X^{lb}(x)$ of T_{P9-11}. (d) $DU_X^{up}(x)$ of T_{P9-11}.

negative direction. For example, the PFDF result $PR_X^{lb}(0) = 0.737$ shown in Fig. 5(a) indicates that the probability of T_{P9-11} lying in the negative direction is 0.737. Moreover, the PFDF result $FR_X^{lb}(0) = 0.122$ hr/occ shown in Fig. 5(b) indicates that the events leading to a negative T_{P9-11} may occur 0.122 times an hour on average. Furthermore, the PFDF result $DU_X^{lb}(0) = 6.041$ hr/occ shown in Fig. 5(c) indicates that the average duration of T_{P9-11} lying in the negative direction is 6.041 hours. Additionally, the PFDF result $DU_X^{up}(0) = 2.156$ hr/occ shown in Fig. 5(d) indicates that the average duration of T_{P9-11} lying in the positive direction is 2.156 hours.

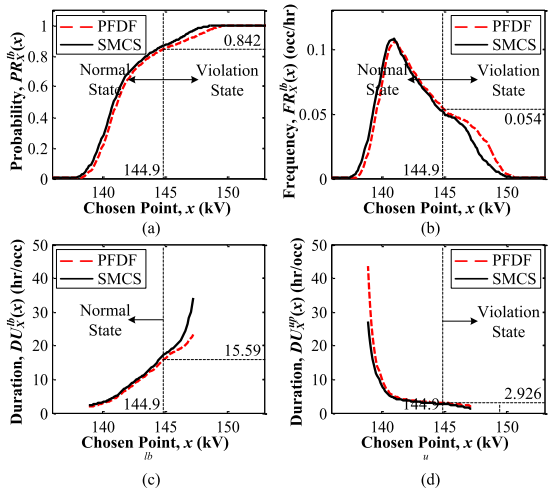

FIGURE 6. $PR_X^{lb}(x)$, $FR_X^{lb}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ of U_{M3} in the IEEE-RTS79 grid. (a) $PR_X^{lb}(x)$ of U_{M3}. (b) $FR_X^{lb}(x)$ of U_{M3}. (c) $DU_X^{lb}(x)$ of U_{M3}. (d) $DU_X^{up}(x)$ of U_{M3}.

The randomness of loads and WF power outputs may also cause the nodal voltages to be outside their acceptable levels. Therefore, we provide accurate frequency and duration assessments for these violation states to reflect the actual operating behavior of a power system. Fig. 6 shows the calculation results of the voltage magnitude at bus 3 (i.e., $X = U_{M3}$). It is known from [36] that the allowable voltage range of U_{M3} is (131.1, 144.9) kV. Then, the probability, frequency, and average duration of U_{M3} (i.e., $X = U_{M3}$) violating its upper bound can be given as $PR_X^{up}(144.9)$, $FR_X^{up}(144.9)$, and $DU_X^{up}(144.9)$, respectively, where $PR_X^{up}(144.9) \approx 1 - PR_X^{lb}(144.9)$ and $FR_X^{up}(144.9) \approx FR_X^{lb}(144.9)$. For example, the PFDF result $PR_X^{up}(144.9) = 1 - PR_X^{lb}(144.9) = 0.158$ (where $PR_X^{lb}(144.9) = 0.842$, as shown in Fig. 6(a)) indicates that the probability of U_{M3} violating its upper bound is 0.158. Moreover, according to the PFDF result $FR_X^{up}(144.9) \approx FR_X^{lb}(144.9) = 0.054$ occ/hr shown in Fig. 6(b), the violation of U_{M3} may occur 0.054 times an hour on average. Furthermore, the PFDF result $DU_X^{up}(144.9) = 2.926$ hr/occ shown in Fig. 6(d) indicates that the average duration for which U_{M3} is in the violation state is 2.926 hours. Additionally, the PFDF result $DU_X^{lb}(144.9) = 15.59$ hr/occ shown in Fig. 6(c) indicates that the average duration of U_{M3} being in the normal state is 15.59 hours. Note that unlike $PR_X^{lb(up)}(x)$ or $FR_X^{lb(up)}(x)$, $DU_X^{lb}(x)$ and $DU_X^{up}(x)$ cannot be derived from each other directly.

Since the linear load flow equations (i.e., (2) and (3)) are used in the proposed PFDF method, the distortion between the SMCS and PFDF results may appear in the region away from the linearization point, whereas the region near the point of linearization shows minimal difference. Some accuracy may be lost, but Figs. 5 and 6 verify that the shapes of lines obtained from SMCS and PFDF remain similar. Moreover, the following test establishes that the calculation efficiency is significantly improved by the proposed method relative to SMCS. Furthermore, the distortion caused by linearization might be further reduced by the multilinearization method proposed in [7]. With the multilinearization method, various points of linearization can be determined, and the same linearization concept can be applied to such points in addition to EVs. Then, a combination of several PFDF solutions around various linearization points can ensure satisfactory performance in terms of computation time and accuracy.

4) ERROR MEASUREMENT

To prove the effectiveness of the proposed method, the normalized sum of square error (NSSE) and average NSSE (ANSSE) used in [39] are calculated to provide error information using the result of SMCS as a reference.

$$\text{NSSE} = \frac{\sum_{x \in \boldsymbol{CP}} (SMCS(x) - PFDF(x))^2}{\sum_{x \in \boldsymbol{CP}} SMCS(x)^2} \tag{47}$$

$$\text{ANSSE} = \frac{\sum_{i=1}^{N_C} \text{NSSE}_i}{N_C} \tag{48}$$

where $PFDF(x)$ denotes the result value of probability, frequency or average duration (hour/occurrence) calculated by the proposed method, e.g., $PFDF(x)$ can be replaced by $PR_X^{lb}(x)$, $PR_X^{up}(x)$, $FR_X^{lb}(x)$, $FR_X^{up}(x)$, $DU_X^{lb}(x)$ or $DU_X^{up}(x)$.

$SMCS(x)$ represents the value given by SMCS corresponding to $PFDF(x)$. $\boldsymbol{CP}$ is the set of x used for comparison, which is evenly chosen from the range of x obtained by SMCS. The interval between each two neighbouring chosen points is set as ΔX. N_C varies with the object of study. When voltage magnitude results are considered, N_C denotes the number of system nodes. When active or reactive branch flow results are considered, N_C denotes the number of system branches.

TABLE 1. NSSE for certain output variables in the IEEE-RTS79 grid.

Variable	NSSE for $PFDF(x)$			
	$PR_X^{lb}(x)$	$FR_X^{lb}(x)$	$DU_X^{lb}(x)$	$DU_X^{up}(x)$
U_{M3}	0.001398	0.047278	0.080087	0.157041
U_{M6}	0.001523	0.046834	0.109742	0.096848
U_{M9}	0.001004	0.040850	0.109191	0.089719
U_{M12}	0.002081	0.058432	0.050403	0.120762
U_{M20}	0.000702	0.050643	0.024073	0.082410
U_{M24}	0.000770	0.041115	0.103601	0.174796
T_{P2-4}	0.000063	0.000317	0.000623	0.000905
T_{P3-24}	0.000004	0.000481	0.002233	0.003365
T_{P9-11}	0.000090	0.001603	0.014877	0.003349
T_{P12-23}	0.000104	0.000527	0.000848	0.000577
T_{P16-17}	0.000112	0.002606	0.005908	0.001861
T_{P19-20}	0.000032	0.001446	0.002877	0.005426
T_{Q2-6}	0.001724	0.043077	0.081366	0.200154
T_{Q3-24}	0.001338	0.047119	0.081370	0.186361
T_{Q9-12}	0.001871	0.033550	0.163412	0.050428
T_{Q14-16}	0.000079	0.003451	0.003782	0.004437
T_{Q16-19}	0.000964	0.016915	0.050713	0.161658
T_{Q21-22}	0.000085	0.018514	0.040563	0.022698

T_{Pa-b} and T_{Qa-b} denote the active and reactive flows in branch a–b, respectively. U_{Ma} denotes the voltage magnitude at node (bus) a.

Table 1 shows the NSSE for various output variables (including voltage magnitudes at some nodes and active/reactive flows in some branches). Table 2 compares the ANSSE for various output variables. The differences between the results obtained by the proposed method and SMCS are small and arise mainly from linearization. Because of the higher nonlinearity of the involved equations, the ANSSE for voltage magnitude or reactive branch flow is greater than that for active branch flow. We take U_{M3} as a reference. Although the NSSE for U_{M3} is relatively large in Table 1, Fig. 6 indicates that the $PFDF(x)$ and $SMCS(x)$ of U_{M3} are nevertheless close.

In each row of Table 1 or 2, the NSSE or ANSSE for $DU_X^{lb}(x)$ and $DU_X^{up}(x)$ are larger than in the corresponding $PR_X^{lb}(x)$ or $FR_X^{lb}(x)$, mainly because the division operation in (30) may amplify calculation errors for $DU_X^{lb}(x)$ (or $DU_X^{up}(x)$), particularly when $PR_X^{lb}(x)$ (or $PR_X^{up}(x)$) becomes large and $FR_X^{lb}(x)$ (or $FR_X^{up}(x)$) becomes small. Such errors will decrease with increasing R_F. Notably, the left side of $DU_X^{lb}(x)$ and the right side of $DU_X^{up}(x)$, where errors are generally small, as shown in Fig. 5 and 6, are the primary concern for the assessment in the PLF problem. Therefore, the effectiveness of the proposed method is confirmed.

TABLE 2. ANSSE for output variables in the IEEE-RTS79 grid.

Variable	ANSSE for $PFDF(x)$			
	$PR_X^{lb}(x)$	$FR_X^{lb}(x)$	$DU_X^{lb}(x)$	$DU_X^{up}(x)$
Voltage magnitude	0.00142	0.04082	0.07394	0.13855
Active branch flow	0.00009	0.0018	0.00481	0.00397
Reactive branch flow	0.00304	0.03975	0.12154	0.15451

B. IEEE-300 NODE GRID

1) SYSTEM DATA

To test the effects of system size on the performance of the proposed method, another test using the IEEE-300 node system was conducted. The IEEE-300 node test system was also modified to include WFs. Conventional generation (10,000 MW) has been reduced proportionally and replaced by wind generation with the same capacity. In this test, all CGs are assumed to have the same reliability data as the 100 MW CG in the IEEE-RTS79 grid. Three WFs are added to share wind power and are connected at buses 37, 126, and 211. In the WF at bus 37, there are 200 WTs of type 1. The WF at bus 126 has 100 WTs of each type (200 WTs in total). The WF at bus 211 also has 100 WTs of each type (200 WTs in total). To provide sufficient wind generation, the rated active power output of each type of WT is set as 10 times the original, i.e., $P_{WN\_1} = 15$ WM, and $P_{WN\_2} = 20$ WM. Since the failures of WTs are simulated as random factors, the computation time of SMCS becomes unacceptable when the number of WTs becomes large. Therefore, P_{WN_r}, rather than the number of WTs, is increased to ensure the feasibility of SMCS, which also indicates the deficiency of SMCS.

The hourly wind speed and load observations described in Section VI-A are still used for each bus independently. Because large fluctuations of power injections are caused in this test by the variations of WF power outputs and loads, a linear dispatching model shown in (4) and (5) is used to obtain new sensitivity matrixes $\boldsymbol{C}'$ and $\boldsymbol{Z}'$ for the calculations of (43) and (44). In the calculation of new sensitivities, the fluctuation of a nodal power injection is assumed to be compensated by all CGs except those connected at this node, and the participation factor of a regulating generator is assumed to be proportional to its reserve capacity.

2) RESULTS OF PLF ANALYSIS

The results of the proposed method are compared with those from the SMCS with 10 years (10×8760 hours) of simulation time for computational accuracy verification. Table 3 presents the ANSSE for various output variables.

TABLE 3. ANSSE for output variables in the IEEE-300 node grid.

Variable	ANSSE for $PFDF(x)$			
	$PR_X^{lb}(x)$	$FR_X^{lb}(x)$	$DU_X^{lb}(x)$	$DU_X^{up}(x)$
Voltage magnitude	0.00647	0.08882	0.15027	0.13822
Active branch flow	0.00146	0.01349	0.02142	0.02525
Reactive branch flow	0.01399	0.09865	0.17035	0.16913

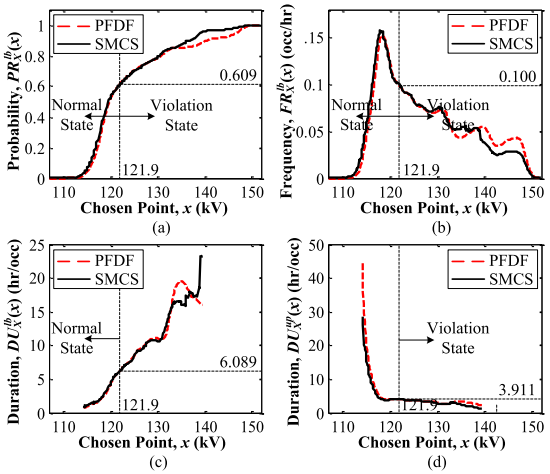

FIGURE 7. $PR_X^{lb}(x)$, $FR_X^{lb}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ of U_{M37} in the IEEE-300 node grid. (a) $PR_X^{lb}(x)$ of U_{M37}. (b) $FR_X^{lb}(x)$ of U_{M37}. (c) $DU_X^{lb}(x)$ of U_{M37}. (d) $DU_X^{up}(x)$ of U_{M37}.

As an example, Fig. 7 depicts the associated $PFDF(x)$ and $SMCS(x)$ of the voltage magnitude at bus 37 (i.e., $X = U_{M37}$), of which the upper bound is 121.9 kV. The PFDF result $PR_X^{up}(121.9) = 1 - PR_X^{lb}(121.9) = 0.391$ (where $PR_X^{lb}(121.9) = 0.609$, as shown in Fig. 7(a)) indicates that the probability of U_{M37} violating its upper bound is 0.391. Moreover, according to the PFDF result $FR_X^{up}(121.9) \approx FR_X^{lb}(121.9) = 0.1$ occ/hr shown in Fig. 7(b), the violation of U_{M37} may occur 0.1 times an hour on average. Furthermore, the PFDF result $DU_X^{up}(121.9) = 3.911$ hr/occ shown in Fig. 7(d) indicates that the average duration for which U_{M37} is in the violation state is 3.911 hours. Additionally, the PFDF result $DU_X^{lb}(121.9) = 6.089$ hr/occ shown in Fig. 7(c) indicates that the average duration of U_{M37} being in the normal state is 6.089 hours.

For the active flow in branch 127–134 (i.e., $X = T_{P127-134}$), Fig. 8 depicts the associated $PFDF(x)$ and $SMCS(x)$. $T_{P127-134}$ is in the positive direction when $T_{P127-134} \geq 0$, i.e., the active flow in branch 127–134 is from node 127 to node 134. Accordingly, $T_{P127-134} \leq 0$ indicates that $T_{P127-134}$ lies in the negative direction.

The PFDF result $PR_X^{lb}(0) = 0.61$ (where $X = T_{P127-134}$) shown in Fig. 8(a) indicates that the probability of $T_{P127-134}$ lying in the negative direction is 0.61. Moreover, the PFDF result $FR_X^{lb}(0) = 0.145$ hr/occ shown in Fig. 8(b) indicates that the events leading to a negative $T_{P127-134}$ may occur 0.145 times an hour on average. Furthermore, the PFDF

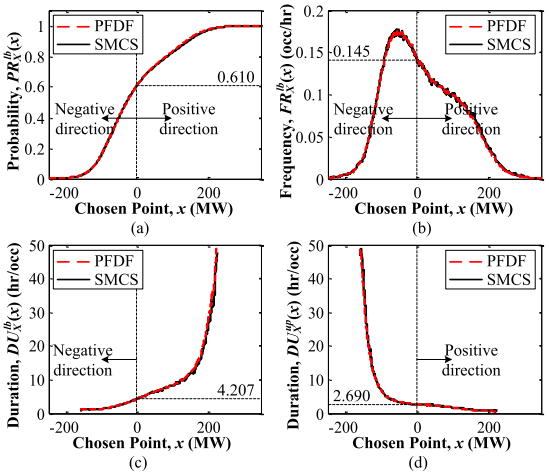

FIGURE 8. $PR_X^{lb}(x)$, $FR_X^{lb}(x)$, $DU_X^{lb}(x)$, and $DU_X^{up}(x)$ of $T_{P127-134}$ in the IEEE-300 node grid. (a) $PR_X^{lb}(x)$ of $T_{P127-134}$. (b) $FR_X^{lb}(x)$ of $T_{P127-134}$. (c) $DU_X^{lb}(x)$ of $T_{P127-134}$. (d) $DU_X^{up}(x)$ of $T_{P127-134}$.

result $DU_X^{lb}(0) = 4.207$ hr/occ shown in Fig. 8(c) indicates that the average duration of $T_{P127-134}$ lying in the negative direction is 4.207 hours. Additionally, the PFDF result $DU_X^{up}(0) = 2.69$ hr/occ shown in Fig. 8(d) indicates that the average duration of $T_{P127-134}$ lying in the positive direction is 2.69 hours.

Figs. 7 and 8 show that the results obtained from the proposed PFDF method match closely with the corresponding SMCS results. The results shown in Table 3 and Figs. 7 and 8 indicate that the proposed method satisfactorily estimates the load flow solution information (including the probability, frequency, and duration information), even for a large system.

Let T_A represent the total active flows of all branches of the test system. The second termination criterion of SMCS is the standard deviation of the expectation estimate of T_A less than 1 MW. For the proposed method and SMCS based on tow termination criteria, Table 4 lists the average computation time on different test systems. Since the number of Markov states for random variables should be optimized before applying the Markov-chain-based method, the time spent on optimizing the state number for the proposed method is the same as that

TABLE 4. Computation time comparison.

Indexes \ IEEE system		IEEE-RTS79	IEEE-300 node
Computation time (s)	TM_1 (for SMCS with 10 years)	200,806.53	6,638,253.08
	TM_2 (for SMCS with $T_A \leq 1$ MW)	37,133.15	9,547,285.46
	TM_3 (for the proposed method)	2.04	163.56
Ratio	TM_1/TM_3	98435	40586
	TM_2/TM_3	18203	58372

for the Markov-chain-based SMCS and is not included in the computation time listed in Table 4.

The proposed method requires much less computational effort than the SMCS method when both methods have similar computational accuracy, and the computational performance of the proposed method is not degraded with increasing system size. Moreover, the computation time of the proposed method using the convolution operation decreases exponentially with increases of the discrete step length ΔX within acceptable accuracy. For example, if ΔX is set as 0.5, then the TM_3 for the IEEE-300 node system decreases from 163 s to 56 s. Moreover, the computational burden of convolution can be greatly relieved using the fast Fourier transform method [3].

VII. CONCLUSION

Conventional PLF methods do not take the frequency and duration information of random variables into account. Here, we propose a frequency and duration method for PLF with WFs. An improved PFDF method is introduced to solve PLF problems with frequency and duration quantities. The random models of WFs, CGs, and loads are developed and represented as corresponding PFDFs based on Markov chains, of which the credibility is improved using the proposed optimal decision-making model to determine the clustering number. With the proposed method, the randomness modeling of the WF power output with frequency and duration characteristics considered becomes straightforward to implement. In addition, the probability, frequency, and duration information of desired variables, such as voltages and branch flows, are computed simply through convolution and addition operations (also including scalar multiplication) of PFDFs.

Our test results indicate the necessity of optimizing the clustering number of random states, which has a considerable influence on the credibility of the Markov-chain-based modeling. Associated probability, frequency, and duration information results of random output variables using the proposed method are compared against the results from SMCS on two IEEE test systems with WFs. The proposed method simplifies the computation process and yields a substantial reduction in computational expense while maintaining a high level of accuracy.

APPENDIX

Let F represent the total active power output of a WF at node m (omitting subscript m below). Considering the uncertainties of WT failures and wind speed V, the state space for F is shown in Fig. 9.

In Fig. 9, λ_{Vik} is the transition rate between states i and k of V, and λ_{Wjk} denotes the transition rate between various states associated with WT failures. The transitions between nonadjacent states are not shown in Fig. 9 for the sake of clarity. The *state probabilities* $p_{F_j^i}$ and $p_{F_i^0}$ are the probability of the corresponding row (WT-failure-associated) state and the column (wind-speed-associated) state, respectively.

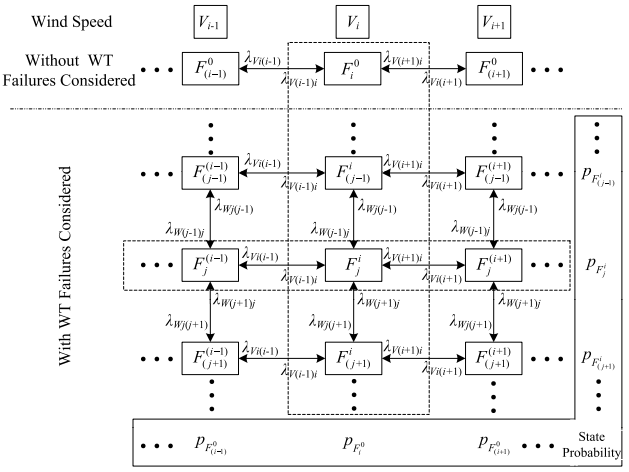

FIGURE 9. State space for the power output of a WF.

Based on the total probability formula, the probability of $F = a$ can be given by

$$p(F=a)=\sum_{i\in \boldsymbol{B}_V} p_{F_i^0} p_{F_j^i}, \quad F_j^i = a, i \in \boldsymbol{B}_V \tag{49}$$

Then, (38) is proved according to the nature of the addition operation of functions.

Assuming that $F_k^0 \le F_i^0$ when $k < i$, the incremental and decremental frequencies corresponding to F_i^0 in (35) are given by

$$f_{F_i^0\_IN} = \sum_{k>i} p_{F_i^0}\lambda_{Vik} - \sum_{k<i} p_{F_k^0}\lambda_{Vki} \tag{50}$$

$$f_{F_i^0\_DE} = \sum_{k<i} p_{F_i^0}\lambda_{Vik} - \sum_{k>i} p_{F_k^0}\lambda_{Vki} \tag{51}$$

Assuming that $F_k^i \le F_j^i$ when $k < j$, the incremental and decremental frequencies corresponding to F_j^i in (36) are given by

$$f_{F_j^i\_IN} = \sum_{k>j} p_{F_j^i}\lambda_{Wjk} - \sum_{k<j} p_{F_k^i}\lambda_{Wkj} \tag{52}$$

$$f_{F_j^i\_DE} = \sum_{k<j} p_{F_j^i}\lambda_{Wjk} - \sum_{k>j} p_{F_k^i}\lambda_{Wkj} \tag{53}$$

Assuming that the order between F_j^k and F_j^i is the same as that between F_k^0 and F_i^0, i.e., $F_j^k \le F_j^i$ when $k < i$ for any j, the incremental and decremental frequencies corresponding to F_j^i, considering both WT failures and the variation of wind speed, are given by

$$\begin{aligned} f'_{F_j^i\_IN} = &\sum_{k>i} p_{F_i^0}p_{F_j^i}\lambda_{Vik} + \sum_{z>j} p_{F_i^0}p_{F_j^i}\lambda_{Wjz} \\ &- \sum_{k<i} p_{F_k^0}p_{F_j^i}\lambda_{Vki} - \sum_{z<j} p_{F_i^0}p_{F_z^i}\lambda_{Wzj} \end{aligned} \tag{54}$$

$$\begin{aligned} f'_{F_j^i\_DE} = &\sum_{k<i} p_{F_i^0}p_{F_j^i}\lambda_{Vik} + \sum_{z<j} p_{F_i^0}p_{F_j^i}\lambda_{Wjz} \\ &- \sum_{k>i} p_{F_k^0}p_{F_j^i}\lambda_{Vki} - \sum_{z>j} p_{F_i^0}p_{F_z^i}\lambda_{Wzj} \end{aligned} \tag{55}$$

Only when the difference between power curves of different types of WTs is significant and the proportion of failure WTs in state j is sufficiently large may the assumption for the order between F_j^k and F_j^i in (54) and (55) produce error due to transitions between the states of $V < V_{CO\_r}$ and $V \geq V_{CO\_r}$ near V_{CO_r}. In practice, the power curves of various types of WTs in the same WF are typically similar, the probability of failure of large-scale WTs is small, and the probability of $V \geq V_{CO\_r}$ should also be small for the case of practical WF planning. Therefore, such error is negligible.

By substituting (50) and (52) into (54) and substituting (51) and (53) into (55), it can be derived that

$$f'_{F_j^i\_IN} = p_{F_i^0} f_{F_j^i\_IN} + f_{F_i^0\_IN} p_{F_j^i} \tag{56}$$

$$f'_{F_j^i\_DE} = p_{F_i^0} f_{F_j^i\_DE} + f_{F_i^0\_DE} p_{F_j^i} \tag{57}$$

Thus, the incremental and decremental frequencies corresponding to the state of $F = a$ are

$$f_{IN}(F = a) = \sum_{i \in \boldsymbol{B}_V} f'_{F_j^i\_IN}, \quad F_j^i = a, i \in \boldsymbol{B}_V \tag{58}$$

$$f_{DE}(F = a) = \sum_{i \in \boldsymbol{B}_V} f'_{F_j^i\_DE}, \quad F_j^i = a, i \in \boldsymbol{B}_V \tag{59}$$

Consequently, (39) and (40) are proved based on the nature of the addition operation of functions.

REFERENCES

[1] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 752–759, May 1974.

[2] A. M. Leite da Silva and V. L. Arienti, "Probabilistic load flow by a multilinear simulation algorithm," *IEE Proc. C, Gener., Transmiss. Distrib.*, vol. 137, no. 4, pp. 276–282, Jul. 1990.

[3] R. N. Allan, A. M. Leite da Silva, and R. C. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 5, pp. 2539–2546, May 1981.

[4] P. Zhang and S. T. Lee, "Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 676–682, Feb. 2004.

[5] A. M. Leite da Silva, V. L. Arienti, and R. N. Allan, "Probabilistic load flow considering dependence between input nodal powers," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 6, pp. 1524–1530, Jun. 1984.

[6] D. Cai, J. Chen, D. Shi, X. Duan, H. Li, and M. Yao, "Enhancements to the cumulant method for probabilistic load flow studies," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, Jul. 2012, pp. 1–8.

[7] R. N. Allan and A. M. Leite da Silva, "Probabilistic load flow using multilinearisations," *IEE Proc. C, Gener., Transmiss. Distrib.*, vol. 128, no. 5, pp. 280–287, Sep. 1981.

[8] Z. Hu and X. Wang, "A probabilistic load flow method considering branch outages," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 507–514, May 2006.

[9] X. Li, J. Cao, and D. Du, "Probabilistic optimal power flow for power systems considering wind uncertainty and load correlation," *Neurocomputing*, vol. 148, pp. 240–247, Jan. 2015.

[10] Q. Xiao, Y. He, K. Chen, Y. Yang, and Y. Lu, "Point estimate method based on univariate dimension reduction model for probabilistic power flow computation," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 14, pp. 3522–3531, Sep. 2017.

[11] X. Xu and Z. Yan, "Probabilistic load flow evaluation considering correlated input random variables," *Int. Trans. Electr. Syst.*, vol. 26, no. 3, pp. 555–572, 2016.

[12] Q. Xu, Y. Yang, Y. Liu, and X. Wang, "An improved Latin hypercube sampling method to enhance numerical stability considering the correlation of input variables," *IEEE Access*, vol. 5, pp. 15197–15205, 2017.

[13] J. F. Dopazo, O. A. Klitin, and A. M. Sasson, "Stochastic load flows," *IEEE Trans. Power App. Syst.*, vol. PAS-94, no. 2, pp. 299–309, Mar./Apr. 1975.

[14] M. Flam and A. M. Sasson, "Stochastic load flow decoupled implementation," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, Mexico City, Mexico, Jul. 1977, pp. 1–8.

[15] C. Singh and R. Billinton, "Frequency and duration concepts in system reliability evaluation," *IEEE Trans. Rel.*, vol. R-24, no. 1, pp. 31–36, Apr. 1975.

[16] F. S. Prabhakara, J. J. Miller, and W. E. Kazibwe, "Frequency and duration of voltage drop in plants with intermittent loads," in *Proc. IEEE Ind. Appl. Soc., Annu. Meeting*, Dearborn, MI, USA, vol. 2, Sep./Oct. 1991, pp. 1584–1590.

[17] Y. Y. Hong and K. L. Pen, "Optimal VAR planning considering intermittent wind power using Markov model and quantum evolutionary algorithm," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2987–2996, Oct. 2010.

[18] J. Usaola, "Probabilistic load flow in systems with wind generation," *IET Gener., Transmiss. Distrib.*, vol. 3, no. 12, pp. 1031–1041, Dec. 2009.

[19] J. Usaola, "Probabilistic load flow with correlated wind power injections," *Elect. Power Syst. Res.*, vol. 80, no. 5, pp. 528–536, 2010.

[20] J. Usaola, "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion," *Int. J. Elect. Power Energy Syst.*, vol. 31, no. 9, pp. 474–481, Oct. 2009.

[21] T. Boehme, A. R. Wallace, and G. P. Harrison, "Applying time series to power flow analysis in networks with high wind penetration," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 951–957, Aug. 2007.

[22] F. C. Sayas and R. N. Allan, "Generation availability assessment of wind farms," *IEE Proc.-Gener., Transmiss. Distrib.*, vol. 143, no. 5, pp. 507–518, Sep. 1996.

[23] A. P. Leite, C. L. T. Borges, and D. M. Falcao, "Probabilistic wind farms generation model for reliability studies applied to brazilian sites," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1493–1501, Nov. 2006.

[24] A. S. Dobakhshari and M. Fotuhi-Firuzabad, "A reliability model of large wind farms for power system adequacy studies," *IEEE Trans. Energy Convers.*, vol. 24, no. 3, pp. 792–801, Sep. 2009.

[25] Y. Ding, C. Singh, L. Goel, J. Østergaard, and P. Wang, "Short-term and medium-term reliability evaluation for power systems with high penetration of wind power," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 896–906, Jul. 2014.

[26] R. Billinton, H. Chen, and R. Ghajar, "Time-series models for reliability evaluation of power systems including wind energy," *Microelectron. Rel.*, vol. 36, no. 9, pp. 1253–1261, 1996.

[27] A. Keane *et al.*, "Capacity value of wind power," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 564–572, May 2011.

[28] R. Billinton and W. Wangdee, "Delivery point reliability indices of a bulk electric system using sequential Monte Carlo simulation," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 345–352, Jan. 2006.

[29] X. Wang and C. Pottle, "A concise frequency and duration approach to generating system reliability studies," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 8, pp. 2521–2530, Aug. 1983.

[30] A. M. Leite da Silva, A. C. G. Melo, and S. H. F. Cunha, "Frequency and duration method for reliability evaluation of large-scale hydrothermal generating systems," *IEE Proc. C, Gener., Transmiss. Distrib.*, vol. 138, no. 1, pp. 94–102, Jan. 1991.

[31] A. M. Leite da Silva, R. A. G. Fernández, and C. Singh, "Generating capacity reliability evaluation based on Monte Carlo simulation and cross-entropy methods," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 129–137, Feb. 2010.

[32] J.-Z. Zhu, T. Zhao, and Y. Zhang, "A Markov chain based method for probabilistic load flow with wind power," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Chicago, IL, USA, Jul. 2017, pp. 1–5.

[33] A. Lisnianski, D. Elmakias, D. Laredo, and H. B. Haim, "A multi-state Markov model for a short-term reliability analysis of a power generating unit," *Rel. Eng. Syst. Saf.*, vol. 98, no. 1, pp. 1–6, 2012.

[34] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric Statistical Methods*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013, pp. 190–200 and 568–578.

[35] W. Li, Y. Shang, and Y. Ji, "Analysis of multiple objective decision methods based on entropy weight," in *Proc. IEEE Pacific–Asia Workshop Comput. Intell. Ind. Appl.*, Wuhan, China, Dec. 2008, pp. 953–956.

[36] IEEE Task Force, "IEEE reliability test system," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 6, pp. 2047–2054, Nov./Dec. 1979.

[37] Iowa Environmental Mesonet. *Wind Speed Data From the Network of AWOS Sensors in the State of Iowa*. [Online]. Available: http://mesonet.agron.iastate.edu/request/awos/1min.php

[38] P. Giorsetto and K. F. Utsurogi, "Development of a new procedure for reliability modeling of wind turbine generators," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 1, pp. 134–143, Jan. 1983.
[39] A. Tamtum, A. Schellenberg, and W. D. Rosehart, "Enhancements to the cumulant method for probabilistic optimal power flow studies," *IEEE Trans. Power. Syst.*, vol. 24, no. 4, pp. 1739–1746, Nov. 2009.

JINZHOU ZHU received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering.

His research interests include power system planning and reliability.

YAN ZHANG received the B.E. degree in power plant and electric power systems from Hefei University of Technology, Hefei, Anhui, China, in 1982, the M.E. degree in high-voltage engineering from China Electric Power Research Institute, Beijing, China, in 1987, and the Ph.D. degree in electric power systems and their automation from Shanghai Jiao Tong University, Shanghai, China, in 1998.

She is currently a Professor with the Department of Electrical Engineering, Shanghai Jiao Tong University. Her research interests include power system stability, planning, and reliability.

HAIBO CHEN received the B.E. degree in electrical engineering and the M.S. degree in business administration from Shanghai Jiao Tong University, Shanghai, China, in 1991 and 1995, respectively.

He is currently the Deputy General Manager of the State Grid Shanghai Municipal Power Company and an Adjunct Professor with Shanghai University of Electric Power. His research interests include power system analysis and power informationization.