A Hybrid Model for Probabilistic Analysis of Modern Power Systems with Integration of Renewable Energy Resources

Nguyen Thi Ai Nhi, Le Dinh Duong*, Ngo Van Duong, Huynh Van Ky

Abstract—Modern power systems faces various uncertainties both from conventional sources, due to stochastic nature of both the load and the availability of generation resources and transmission assets, and from renewable resources. The increasing penetration of wind and solar power generation introduces additional uncertainty, causing more difficulties in power system analysis. To deal with uncertainties, Probabilistic Power Flow (PPF) has been introduced as an efficient tool. In this paper, we develop a hybrid model that combines scenario analysis technique and cumulant based PPF approach. It can take into account various sources of uncertainty in the power system and their correlations. The proposed approach is performed on IEEE-118 bus test system, indicating good performance in comparison with Monte Carlo approach.

Index Terms—Renewable energy, Uncertainty, Probabilistic Power Flow, Cumulant, Scenario analysis.

1. Introduction

During the operation of the power system, the parameters of operating mode such as power flow or current flowing through branches, voltage of nodes, etc., must be regularly calculated and then compared with their permissible limits to evaluate as well as to propose suitable solutions to handle and ensure the security of the system in case of a risk of unsafety. The traditional Deterministic Power Flow (DPF) method [1] is the tool used to determine the mode parameters; however, in the calculation process, the DPF method only uses power injections from nodes (from loads, generation resources, etc.) which are fixed values and the grid structure that is known in advance, so the uncertainty factors (random factors) from loads, power generation, grid structure (such as random failures of lines and equipment) are not taken into account. This is the large drawback of the DPF calculation method.

In order to overcome the above disadvantages, Probabilistic Power Flow (PPF) method was introduced and become a very effective calculation tool to account for uncertainties in the system. Power injected by loads and power sources, working status of elements such as lines, transformers... can follow certain probabilistic laws [2]. Especially, for today’s power systems, integrating more and more renewable energy sources into the system such as wind power, solar power, etc., introduces large amounts of uncertainties that increase the level of uncertainty in the system. PPF must fully take into account all such the uncertainties in the computation process and give outputs such as nodal voltage, current and power through the lines... in the form of probability distributions. Therefore, the PPF calculation method allows to evaluate the probability of the existence of dangerous modes (voltage, current, transmission power exceeding the allowable values) in the system, from that the operator can propose an appropriate solution to improve safety for the system.

The PPF method was first proposed by Borkowska in 1974 [3] and since then many research works on this field have been published in the world. In general, PPF methods can be divided into three main groups: analytical methods, approximation methods and numerical methods.

Analytical methods use algorithms or techniques of analysis such as convolution [4-6] and semi-invariant (cumulant) [7-11]. Applying analytical techniques combined with the input-output relationship of the problem allows to determine the distribution functions of the output random variables (such as nodal voltage, current and power transmitted through branches) according to system parameters (line impedance, transformer impedance...) and probability distributions from input random variables from load, power generation of conventional generators and renewable energy sources, etc.

The input-output relationship of the power flow calculation problem is nonlinear. However, the analytical method only works with the linear input-output relationship. Therefore, before using analytical techniques, the input-output relationship is linearized using expansion methods such as McLaren, Taylor expansion.
The general advantage of the analytical method is that the calculation results are quickly obtained. Among two methods using convolution and cumulant techniques, the convolution method is more computationally intensive, requires more memory, and gives slower results than the cumulant method. Therefore, the cumulant method is now more commonly used than the convolution method. The cumulant method is often used in conjunction with series expansion techniques, e.g., Gram-Charlier [7, 8], Cornish-Fisher expansions [11], to obtain a distribution function for the output random variables. Thanks to the advantages of fast calculation, the analytical method can be used to calculate for large power systems in practice. However, the analytical method has the following main disadvantages: 1) Due to the use of techniques to linearize the input-output relationship, the accuracy of the analytical method is greatly affected when the input variable changes in a wide range such as the output of wind sources; 2) Analytical method using expansion techniques such as Gram-Charlier, Cornish-Fisher and these techniques give high accuracy if the distribution functions of the input variables are normal distribution (Gaussian distribution) or close to the normal distribution. However, for real power systems, the distributions of the input random variables often follows the distributions different from the Gaussian distribution, so the results obtained when applied in practice are not very accurate.

The typical approximation method in PPF is the point estimate method [12-15]. In this method the input random variables are parsed out into a series of values and corresponding weights, then the output random variable’s moment is calculated as a function of the input random variables. From that, the distribution function of the output random variable is established based on the calculated moment.

The advantage of the point estimate method is that it gives relatively fast results, so it can also be applied to calculations for large power systems. In addition, this method uses the nonlinear input-output relationship of the power flow calculation problem, so the calculation results do not depend on the linearization process like the analytical method. However, the point estimation method has the disadvantage that the accuracy decreases with increasing the order of the moment, so the distribution function of the output variable has a decrease in accuracy. Another limitation of the approximation method is that when applying calculations to large-scale power systems with an increased number of input random variables, the computational burden increases, making the total calculation time increase significantly.

Monte Carlo simulation (MCS) based PPF is a typical method of the group of simulation methods [16-18]. In the Monte-Carlo method, the input random variables (representing processes, random events) will be sampled and then the power flow calculation process (using the same methods as the traditional DPF method) will be performed for all those samples. The accuracy of this method depends greatly on the sampling technique and on the number of samples taken (number of samples is often very large). This method uses the nonlinear input-output relationship of the problem like the traditional power flow calculation methods. To increase the efficiency of the computation, the techniques of Latin hypercube [19], Latin supercube [20] and importance sampling [21], etc., are used.

One of the most advantage of the Monte-Carlo simulation method is that the results obtained are very accurate and reliable. The probability distributions of the input random variables are generally easier to implement than analytical and approximation methods. However, the largest disadvantage of Monte Carlo simulation is that it requires too much computation time, so it is difficult to apply power system calculations, especially large power systems in practice.

From the above overview and analysis, it is shown that each PPF method has its own characteristics, so depending on the actual application, we choose the most appropriate PPF method.

In order to overcome a number of the above-mentioned difficulties, in this paper we develop a methodology for PPF based on the combination of cumulant and scenario analysis techniques that allows to integrate various types of input probability distribution functions, to take into account correlation of input random variables and to reduce the computation time while ensuring to a certain degree of accuracy.

2. Uncertainty modeling

2.1. Using probability distributions to model uncertainty

Modelling of uncertainties from input random variables is necessary for PPF computation. In the paper, the uncertainty sources from solar, wind, and load demand are considered and modelled.

- Wind generation

The Weibull distribution is widely used in wind speed analysis [22]. The probability density function (PDF) of Weibull distribution is

\[ f(v) = \frac{k}{c} \left( \frac{v}{c} \right)^{k-1} \exp \left( -\left( \frac{v}{c} \right)^k \right) \]

where \( v \) is wind speed; \( k \) is the shape parameter; \( c \) is the scale parameter.

The characteristic curve of a wind turbine can be represented as follows

\[ P_{wp}(v) = \begin{cases} 0 & v \leq v_{c_i} \text{ or } v > v_{c_0} \\ P_w & v_{r_i} - v_{c_i} \leq v < v_{c_i} \\ v_{c_i} < v < v_r \\ P_w & v_r < v < v_{c_0} \end{cases} \]

Where \( v_r \), \( v_{c_i} \), and \( v_{c_0} \) are the rated, cut-in, and cut-out wind speed, respectively; \( P_w \) and \( P_{wp} \) are the rated and output power of wind generation, respectively.

- Solar generation
Beta distribution [23] is usually used to model solar radiation:
\[
f(r) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \left( \frac{r}{r_{\text{max}}} \right)^{\alpha-1} \left( 1 - \frac{r}{r_{\text{max}}} \right)^{\beta-1}
\]
where \( r_{\text{max}} \) and \( r \) are the maximum and actual solar radiations, respectively; \( \alpha \) and \( \beta \) are distribution parameters; \( \Gamma(\cdot) \) is the well-known Gamma function.

\[P_{\text{pv}}(r) = \begin{cases} 
\frac{P_S}{r_{\text{std}}} r^2 & r < r_c \\
\frac{P_w}{r_{\text{std}}} r & r_c \leq r \leq r_{\text{std}} \\
P_S & r > r_{\text{std}} 
\end{cases}
\]

where \( r_{\text{std}} \) and \( r_c \) are the solar radiation in the standard condition of environment and a certain radiation point, respectively; \( P_S \) is the rated power of the photovoltaic (PV) unit; \( P_{\text{pv}} \) is the output power. Normally, solar generation operates in the unity power factor.

- **Load**

  The uncertainty of each load is represented by a Gaussian distribution [9, 10] in which the mean (expected value) is its base power and the standard deviation is assumed to be equal to a certain percentage, e.g., 8%, of the expected value.

2.2. Scenario based uncertainty modelling

In order to solve problems under uncertainty, scenario analysis is one of the most useful tools. The uncertainties as well as the correlations of input random variables can be characterized by using the scenario analysis.

- **Scenarios generation**

  Scenario generation is the process of representing uncertainty with a large number of scenarios. Sampling techniques are commonly used to generate scenarios based on the distributions and correlations of uncertainty factors. In this paper, the comprehensive and realistic method in [24] is exploited to generate scenarios. The model makes use of time series analysis and Principal Component Analysis (PCA) along with data preprocessing techniques to explicitly capture the salient characteristics of uncertainty factors such as distinct seasonal and diurnal patterns, spatial and temporal correlations, and non-Gaussianity. Moreover, the model is able to reduce the dimensions of data sets, so it is helpful for working with high-dimensional data.

- **Scenarios reduction**

  Generally, to capture the salient features of uncertainties, a scenarios generation technique needs to generate a very large number of scenarios leading to high computational burden. A scenario reduction technique is then used to reduce the computational burden and obtain a small number of representative scenarios. For this task, we apply fast forward selection (FFS) method [25].

3. Methodology

3.1. Cumulant based PPF method

The method of calculating PPF based on cumulant [9, 10] is performed as follows:

The equations for calculating power flow in matrix form are as

\[w = g(x)\]
\[z = h(x)\]

where, \( w \) is the vector of power injected into nodes; \( x \) is the vector of the state variable (voltage, phase angle); \( z \) is the vector of the transmitted power on the branches; \( g(x) \) are power flow equations; \( h(x) \) is the function to calculate power flow.

Initially, DPF is performed to calculate the power flow for the system. After that, Taylor expansion can be used to linearize the equations around their solution \( \bar{x} \). The results are as follows:

\[\Delta x = G|_{\bar{x}} \Delta w\]
\[\Delta z = H|_{\bar{x}} \Delta w\]

where, \( G|_{\bar{x}} \) and \( H|_{\bar{x}} \) are the inverse of the Jacobian matrix and the sensitivity matrix of the power flow according to the power injected into the nodes, respectively; \( G|_{\bar{x}} \) and \( H|_{\bar{x}} \) are calculated for \( \bar{x} \).

In PPF, each component of the vectors \( x, w, \) and \( z \) is treated as a realization of the random variable corresponding to the state variable, the power injected into the nodes, and the power on the branches. Based on the linear relationship (7) and (8), the cumulant-based PPF calculation method is performed in the following steps:

- Calculate the DPF for the system and get the solution \( \bar{x} \) and \( G|_{\bar{x}} \) and \( H|_{\bar{x}} \);
- Compute the cumulant of the state variables and the power transmitted on the branches according to the cumulant of the input random variables based on the linear relationship (7) and (8);
- Use expansion techniques to estimate the probability density function (PDF) and/or cumulative distribution function (CDF) functions of the output random variables.

3.2. Scenario based cumulant PPF method

The procedure of the developed scenario based cumulant PPF (denoted as SBCPPF) is:

- **Step 1:** Input the required data for PPF (loads, renewable energy resources, network topologies, correlation information);
- **Step 2:** Form structure \( \omega = \{\omega_1, \omega_2, \ldots, \omega_1, \ldots, \omega_N\} \) for \( N \) non-Gaussian input random variables and generate scenarios as presented in Section 2.2;
- **Step 3:** Carry out the scenario selection technique FFS to obtain a reduced number \( N_s \) of representative scenarios \( \omega^* = \{\omega_1^*, \omega_2^*, \ldots, \omega_{N_s}^*\} \);
- **Step 4:** Perform the cumulant method for each representative scenario \( \omega_i^* \) to obtain the corresponding
means and standard deviations of output random variables;

- Step 5: Form the CDFs and PDFs for output random variables by the weighted sum of Gaussian functions series based on (9) and (10):

$$f(x|\omega^*) = \pi^{(w_1)} f(x|\omega^*_1) + \cdots + \pi^{(w_{ns})} f(x|\omega^*_n)$$

$$= \sum_{i=1}^{ns} \pi^{(w_i)} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x - \mu_i)^2}{2\sigma_i^2}}$$  \hspace{1cm} (9)$$

$$F(x|\omega^*) = \pi^{(w_1)} F(x|\omega^*_1) + \cdots + \pi^{(w_{ns})} F(x|\omega^*_n)$$

$$= \sum_{i=1}^{ns} \pi^{(w_i)} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(\gamma - \mu_i)^2}{2\sigma_i^2}} d\gamma$$  \hspace{1cm} (10)$$

where $\pi^{(w_i)}$ is the occurrence probability of scenario $\omega^*_i$.

- Step 6: Perform probabilistic analysis and assess the security of the system (e.g., probability of over-/under-voltage, probability of line overloading).

4. Tests and discussion

To illustrate and compare the above methods, a case study related to IEEE 118-bus test system is used. The tests are implemented in Matlab (on a PC using 2.53 GHz CPU and 4 GB RAM).

The required data for DPF of the IEEE 118-bus test system are provided in [26]. To test the developed method, the system is modified, i.e., four wind farms and six solar parks are connected to ten buses as given in Table 1 and Table 2, respectively. The information relevant to wind farms, solar parks, loads and their uncertainties are assumed to be known. For the sake of simplicity, load at each bus is modelled by a Gaussian distribution. The expected value of the distribution is equal to its base value while its standard deviation is assumed, for example, to be equal to 8% of its mean. Similarly, the uncertainties from wind farms and solar parks are assumed to follow Weibull distributions and Beta distributions, respectively, with the parameters as given in Table 1 and Table 2. In addition, the correlation coefficient for wind outputs at different buses is equal to 0.5 while that information for solar outputs at different buses is 0.6.

**TABLE 1: Information for wind resources**

<table>
<thead>
<tr>
<th>Node</th>
<th>$v_c$ (m/s)</th>
<th>$v_r$ (m/s)</th>
<th>$v_{co}$ (m/s)</th>
<th>$c$</th>
<th>$k$</th>
<th>$P_{in}$ (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>12</td>
<td>25</td>
<td>20</td>
<td>2</td>
<td>114</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>12</td>
<td>25</td>
<td>15</td>
<td>1</td>
<td>132</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>12</td>
<td>25</td>
<td>20</td>
<td>1.8</td>
<td>70</td>
</tr>
<tr>
<td>50</td>
<td>3</td>
<td>12</td>
<td>25</td>
<td>25</td>
<td>2.1</td>
<td>72</td>
</tr>
</tbody>
</table>

After obtaining the necessary input information for the problem, scenario generation technique, discussed in Section 2.2, is used to generate 1000 ($N = 1000$) scenarios that characterize non-Gaussian input random variables, i.e., wind and solar resources. Next, 50 ($N_s = 50$) representative scenarios are obtained by using FFS method.

**TABLE 2: Information for solar resources**

<table>
<thead>
<tr>
<th>Node</th>
<th>$a$ (kW/m^2)</th>
<th>$b$ (kW/m^2)</th>
<th>$r_{min}$ (W/m^2)</th>
<th>$r_{max}$ (W/m^2)</th>
<th>$r_e$ (W/m^2)</th>
<th>$r_{std}$ (W/m^2)</th>
<th>$P_v$ (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.2</td>
<td>4.1</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>90</td>
</tr>
<tr>
<td>7</td>
<td>2.6</td>
<td>4.2</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>92</td>
</tr>
<tr>
<td>14</td>
<td>2.7</td>
<td>3.8</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>92</td>
</tr>
<tr>
<td>31</td>
<td>3.1</td>
<td>4.0</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>82</td>
</tr>
<tr>
<td>38</td>
<td>3.2</td>
<td>3.9</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>30</td>
</tr>
<tr>
<td>86</td>
<td>2.4</td>
<td>3.7</td>
<td>0</td>
<td>1000</td>
<td>150</td>
<td>900</td>
<td>102</td>
</tr>
</tbody>
</table>

In order to assess the accuracy and efficiency of the developed PPF method (i.e., SBCPPF), a MCS with 10,000 samples has been performed and taken as reference.

For evaluating the accuracy of the considered method, the Average Root Mean Square (ARMS) error is also computed using the results from MCS as in [9]. ARMS is calculated as [9]:

$$ASRM = \frac{1}{NP} \sum_{i=1}^{NP} (MCS_i - SBCPPF_i)^2$$  \hspace{1cm} (11)$$

where $MCS_i$ and $SBCPPF_i$ are the $i$-th value on CDF curves obtained by MCS and by the SBCPPF method, respectively. $NP$ is the number of samples considered (in this case, an interval of p.u. is used). The base power of 100 MVA is used.

Besides MSC and SBCPPF, CMPPF is also implemented. Probability distributions of all output random variables can be estimated. Nevertheless, for illustration purposes, the CDFs of some selected output random variables are depicted. Fig. 1 and Fig. 2, for example, plot CDFs of real power flow through line 69-75 (i.e., $P_{69-75}$) and of voltage at bus 58 (i.e., $V_{58}$), respectively. As can be seen from the figures, the results from the developed approach SCBPPF is better than the results obtained by CMPPF in comparison to MCS results.

![CDFs of real power flow through line 69-75 ($P_{69-75}$)](image)

For evaluating the accuracy according to ARMS, the smaller values of ARMS errors obtained by SCBPPF in comparison to that of CMPPF show the higher accuracy of SCBPPF in estimation of probability distributions. To demonstrate the accuracy, the mean (Mean)
and maximum (Max) values of ARMS errors of all output random variables of the system (i.e., active power flows, reactive power flows, nodal voltages and angles) are also shown as follows: $\text{ARMS}_{\text{SBCPPF}}^{\text{Max}} = 0.12\%$, $\text{ARMS}_{\text{CMPPF}}^{\text{Max}} = 0.54\%$ and $\text{ARMS}_{\text{SBCPPF}}^{\text{Mean}} = 0.07\%$, $\text{ARMS}_{\text{CMPPF}}^{\text{Mean}} = 0.36\%$.

![Fig. 2: CDFs of real power flow through line 69-75 (P_{69-75})](image)

Regarding computational burden, Table 3 clearly shows that all the cumulant methods considered give results in a few seconds, compared to a hundred of seconds from MCS. SBCPPF requires more computation time, compared to CMPPF; however, the amount of increase is not significantly. Hence, SBCPPF is suitable for calculation and analysis of a practical large scale system.

<table>
<thead>
<tr>
<th>Method</th>
<th>Time (s)</th>
<th>SBCPPF</th>
<th>CMPPF</th>
<th>MCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (s)</td>
<td>253</td>
<td>4.85</td>
<td>2.38</td>
<td></td>
</tr>
</tbody>
</table>

PPF provides results in terms of probability distributions for output variables that are useful for assessing power systems taking into account the uncertainties from input variables.

For instance, the real power flow limit of line 69-75 is assumed to be equal to 115 MW (the vertical dotted line in Fig. 1), the probability so that power flow through line 69-75 is over its limit can be computed as

$$P\{P_{69-75} > 115\} = 2.1\%$$ (12)

With the above result calculated, the operators of the system can identify the possible overload level of the branch so that depending on the importance of the branch, they can propose appropriate solutions to avoid overloading such as reducing the load or using devices to regulate the power transmitted through that branch.

Probability so that voltage at a considered bus is out of its operating range (i.e., assumed to be $[0.95, 1.05]$ p.u. in this test) can also be calculated. It can be seen from Fig. 2 that voltage at node 58 is within its operating range.

Therefore, the system operator can use the results in probabilistic form from PPF to evaluate the operating states of the system. In case a risk is detected, the operator can suggest reasonable solutions to deal with the issue.

5. Conclusion

In this paper, a scenario based cumulant PPF approach is developed. It can account for various sources of uncertainty in power systems and their correlations.

The developed approach together with the traditional cumulant method and MCS are carried out on IEEE-118 bus system. The results obtained are compared showing that the proposed approach can achieves high accuracy, while requiring less computation time in comparison with MCS. The proposed methodology application results are useful in power system analysis and in assessing power system security, such as probability of line overloading, probability of over-/under-voltage.

Acknowledgment

This research was funded by the Ministry of Education and Training, Vietnam under project number CT2022.07.DNA.03.

References


Nguyen Thi Ai Nhi received her Ph.D. degree in Electrical Engineering from Polytechnic University of Milan (Politecnico di Milano), Italy in 2016. Her research interests include renewable energy, power systems, and energy storage system.

Le Dinh Duong received his Ph.D. degree in Electrical Engineering from Polytechnic University of Milan (Politecnico di Milano), Italy in 2014. He worked as a visiting research scholar at the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign (UIUC), USA from February to August 2012. His research interests include probabilistic approaches applied to power systems, integration of renewable energy resources into the grid, power system security assessment, power system operation and planning, and energy storage system.

Ngo Van Duong received his Ph.D. degree in network and power system from Hanoi University of Science and Technology, Vietnam in 2002. His research interests include renewable energy and power systems.

Huynh Van Ky received his Ph.D. degree in network and power system from The University of Danang, Vietnam in 2020. His research interests include renewable energy and power systems.