DOI: 10.48011/asba.v2i1.1078

Probabilistic Power Flow Analysis Based on Point-Estimate Method for High Penetration of Photovoltaic Generation in Electrical Distribution Systems

Cordero B. Luis*. Franco B. John*

**São Paulo State University, Brazil, CEP 1538500*
**** Electrical Engineering Department,*
e-mail: luis.gc.bautista@ unesp.br.
**e -mail: fredy.franco@unesp.br*

Abstract: Environmental awareness and energy policies led to decarbonization targets, fostering the adoption of distributed energy resource in the distribution network. Particularly, photovoltaic systems have been gaining momentum due to cost-competitive option and financial benefits. However, traditional distribution networks were not designed for intermittency in power generation. This poses technical issues such as reverse power flow, overvoltage, and thermal overloading. Furthermore, the growth in intermittency and variability of distributed energy resources increases the uncertainty, hence, it brings challenges for the operation, planning, and investment decisions. In this context, probabilistic methods to cater for these uncertainties are essential to address this issue. This paper presents a probabilistic power flow method based on point estimate method combined Edgeworth expansion for high penetration of photovoltaic generation in distribution networks. Normal distribution and Beta distribution are considered for load and solar irradiation modelling, respectively. The method is assessed for different cases using the IEEE 33-bus distribution test system with photovoltaic systems installation. The point estimate method combined Edgeworth expansion provided satisfactory results with lower computational effort and high fitting accuracy of statistical information compared to Monte Carlo simulation.

Resumo: Políticas ambientais e de energia têm como meta a redução de gases de efeito estufa, incentivando a adoção de recursos energéticos distribuídos. Em particular, sistemas de geração fotovoltaica tem aumentado sua presença nas redes elétricas de distribuição devido a preços competitivos e benefícios financeiros. Entretanto, essas redes não foram projetadas para uma alta injeção de geração intermitente, causando desafios técnicos como fluxo reverso, sobretensões e sobrecarga. Além disso, o crescimento da intermitência e variabilidade dos recursos energéticos distribuídos aumenta as incertezas, portanto, traz desafios na operação, planejamento e a decisão de investimento. Nesse contexto, os métodos probabilísticos que atendem as incertezas são essenciais para abordar esse assunto. Este artigo apresenta o fluxo probabilístico de potência baseado no método de estimação por pontos junto a expansão Edgeworth para um sistema com alta geração fotovoltaica. A modelagem estocástica da carga e a irradiação solar seguem a distribuição normal e Beta, respectivamente. O método foi avaliado para diferentes casos usando o sistema teste IEEE-33 barras com instalação de sistemas fotovoltaicos. Foi verificado que o método de estimação por pontos, junto a expansão Edgeworth fornece resultados satisfatórios com baixo esforço computacional, alta precisão das informações estatísticas quando comparados à simulação de Monte Carlo.

Keywords: Probabilistic power flow, photovoltaic systems, point estimate method, Monte Carlo simulation, reverse power flow, Edgeworth expansion.

Palavras-chaves: Fluxo probabilístico de potência, sistemas fotovoltaicos, método de estimação por pontos, simulação de Monte Carlo, fluxo reverso de potência, a expansão Edgeworth.

1. INTRODUCTION

The adoption of distributed energy resources (DERs) into the traditional distribution network increased significantly over the recent years. That transition of the distribution network took place not only as a measure of environmental impact reduction and energy policies to counteract carbon emission from traditional fossil fuels but also to harness the potential of low-carbon technologies (Carstens and da Cunha, 2018).

Among the DERs, photovoltaic (PV) systems have been gaining momentum due to its cost-competitive investments, more efficient production capacity and financial incentives. This technology not only brought systems support benefits but also energy bills reduction and increases the independence from traditional fossil fuels from small to utility-scale projects (REN21, 2020).

However, traditional distribution networks were not designed to operate for the intermittency in generation, which also brought technical challenges, for instance, reverse power flow , overvoltage, and thermal overloading (de Lima Vianna *et al.*, 2018). Furthermore, the growth in intermittency and variability of distributed energy resources (e.g. wind and solar power plant, energy storage systems, electric vehicles) increases the uncertainty in operation and investment decision.

In order to understand the effects of this non-dispatchable source in the operation and planning of the distribution network, the traditional approach such as the deterministic power flow is not enough because it was designed to solve problem for one specific point in time without taking into consideration uncertainties; which may lead to unrealistic outcomes. Therefore, a probabilistic power flow (PPF) is essential to cater for the intermittency and variability of distributed resources, leading to a more realistic representation of the characteristics of the power system (Chen, Chen and Bak-Jensen, 2008).

A well-known method to deal with the PPF is the Monte Carlo simulation (MCS). It was used to investigate the effects of reverse power flow in the presence of wind and photovoltaic systems for medium voltage (MV) distribution network (Constante-Flores and Illindala, 2019). Although the MCS has the ability to represent complex system behaviour by evaluating thousands of deterministic analyses, the major drawback has been its high computational effort.

Another alternative for the PPF is the *2m+1* point estimate method (PEM). That approach was used to solve the optimal PPF with the objective to minimize cost grid operation for unbalanced three phase MV networks (Giraldo *et al.*, 2019). This approximate method is a strong technique that accurately approximates the expected value and variance of a variable, furthermore, the computational burden is reduced (Plattner, Farah Semlali and Kong, 2017).

In order to obtain more statistical information of a variable of interest, approximate curves of the probability density function (PDF) and cumulative distribution function (CDF), the Gram-Charlier expansion were proposed by Zhang and Lee (2004) to estimate the PDF of branch power flow for transmission systems. However, the Gram-Charlier expansion is divergent for higher orders. On the other hand, the Edgeworth expansion for nearly Gaussian distribution provides better results (Blinnikov and Moessner, 1998; Fan *et al.*, 2012).

For the probabilistic approach, historical measurements, theoretical modelling or synthetic profiles can be considered to calculate the probability density function of input variables (Soroudi, 2014).

According to (Constante-Flores and Illindala, 2019), the load demand can be represented by a normal probability density function. For the photovoltaic (PV) generation, the solar irradiation and temperature intermittency influences the PV power output, hence, the PV characteristic and location are important. For estimating the solar irradiation, the Beta probability density function was adopted by authors (Dhayalini and Mukesh, 2019; Fernandes *et al.*, 2019).

This paper addresses the lack of significant statistic information on the behaviour of main feeder utilization, voltage levels and power flow variables. The knowledge of probability distribution of state variables is of special interest to meet technical and economic optimization of the distribution network. Therefore, the PPF based on $2m+1$ PEM combined with the Edgeworth expansion approximation is used to calculate PDF and CDF of state variables. This method solves the PPF with high penetration of photovoltaic systems as well as their effects by providing statistical information regarding confidence interval for a variable of interest

This paper is organized as follow. Section 2, the MCS and PEM methods used in this paper with their respective implementation steps, and the Edgeworth expansion for PDF and CDF. Second, uncertain variables modelling such as load and photovoltaic generation. Section 3, the application of the PPF based on *2m+1* PEM. Section 4, the assessment of the PEM is carried out for different cases using a modified IEEE 33-bus radial distribution test system with photovoltaic systems installation. Section 5, some relevant conclusions are summarized.

2. PROBABILISTIC TECHNIQUES

Probabilistic power flow techniques are developed for impact assessment of the uncertainties such as the random behaviour of energy consumers and the introduction of DERs in the power systems. The MCS is used for purpose comparison as benchmark, and this section describes general steps for its implementation. The *2m+1* PEM method is also presented with a generalized steps for implementation.

2.1 Monte Carlo Simulation

The MCS method is a well-known method to obtain accurate statistical results, though the computational burden to reach an acceptable result is unattractive (Zhang and S.T. Lee, 2004). This method usually involves the calculation of thousands deterministic analysis. For each deterministic case, a scenario is created containing a set of points randomly selected from each uncertain variable (Constante-Flores and Illindala, 2019).

The MCS can be implemented through the following general steps:
Step 1: Compute the deterministic analysis for each set of scenarios;
Step 2: Store results for the variable of interest;
Step 3: Verify the stopping criterion (the maximum number of iterations or the convergence of the expected value);
Step 4: Calculate the statistical information (e.g, expected value, standard deviation) for the variable of interest.

2.2 2m+1 Point Estimate Method

This approximate technique performs as many deterministic analysis as two times uncertain variables are considered plus one deterministic analysis for the expected value of all uncertain variables (Giraldo *et al.*, 2019). Let $\mathbf{Z}_{\boldsymbol{r}}$ represents a set of points for the uncertain variable $\boldsymbol{r}$, where $\boldsymbol{r} \in \{\mathbf{1}, \mathbf{2}, \dots, \boldsymbol{m}\}$. From $\mathbf{Z}_{\boldsymbol{r}}$, the method selects two points based on mean $\mu_{\boldsymbol{r}}$, standard deviation $\sigma_{\boldsymbol{r}}$, skewness $\lambda_{\boldsymbol{r},\mathbf{3}}$ and kurtosis $\lambda_{\boldsymbol{r},\mathbf{4}}$ coefficients. The generalized following steps are applied:

Step 1: Calculate concentrations $\{y_{r,i}, \omega_{r,i}\}$, for its location $y_{r,i}$ of all the uncertain variables $\boldsymbol{r}$ using (1)-(2), and let $\boldsymbol{i} \in \{\mathbf{1}, \mathbf{2}\}$;

$$\xi_{r,i} = \lambda_{r,3}/2 + (-1)^{3-i}\sqrt{\lambda_{r,4} + 3(\lambda_{r,3})^2/4} \tag{1}$$

$$y_{r,i} = \mu_r + \xi_{r,i}\sigma_{r,i} \tag{2}$$

Step 2: Perform a deterministic analysis for each location $y_{r,i}$, while fixing the remaining variables to their expected values;
Step 3: Perform an extra deterministic analysis using the expected value (μ_r) of the uncertain variables;
Step 4: Define $\boldsymbol{Y_x} = \{x_{r,,i} \cup x_0\}$, whose length is *2m+1*. This vector stores results $x_{r,,i}$ from step 2 and x_0 from step 3 for a variable of interest x*;*
Step 5: Calculate the weighting factor $\omega_{r,i}$ for each location $y_{r,i}$ using (3);

$$\omega_{r,i} = (-1)^{3-i}\left(\xi_{r,i}(\xi_{r,1} - \xi_{r,2})\right)^{-1} \tag{3}$$

Step 6: Calculate an extra weighting factor for Step 3 using (4);

$$\omega_0 = 1 - \sum_{r=1}^{m}\left(\lambda_{r,i}(\lambda_{r,1} - \lambda_{r,2})\right)^{-1} \tag{4}$$

Step 7: Define $\boldsymbol{W_x} = \{\omega_{r,i} \cup \omega_0\}$, with length *2m+1*.
Step 8: Calculate the *j*-th moment $\boldsymbol{E}[\boldsymbol{x^j}]$ for a variable of interest $\boldsymbol{x}$ using (5);

$$E[x^j] = \boldsymbol{W_x}(\boldsymbol{Y_x})^{\boldsymbol{j}} = \sum_{r=1}^{m}\left(\omega_{r,i}(x_{r,,i})^{\boldsymbol{j}}\right) + \omega_0(x_0)^{\boldsymbol{j}} \tag{5}$$

Step 9: Calculate the statistical information in terms of the raw moments $E[x^j]$. Calculate the mean and standard deviation of a variable of interest $\boldsymbol{x}$ using (6) and (7).

$$\mu_x = E[x^1] \tag{6}$$

$$\sigma_x = \sqrt{E[x^2] - (E[x^1])^2} \tag{7}$$

This paper uses the Edgeworth expansion because of its asymptotic expansion property to estimate the PDF and the CDF for a variable of interest x, *w*hose υ^{th} cumulant can be calculated in terms of the raw moments α_υ(e.g. $\alpha_1 = E[x^1]$) using (8) (Fan *et al.*, 2012; Pender, 2014).

$$\kappa_\upsilon = \alpha_\upsilon - \sum_{i=1}^{\upsilon-1}\binom{\upsilon-1}{i-1}\kappa_\upsilon\alpha_{\upsilon-i} \tag{8}$$

The reference function $\varphi(z)$ represents a PDF, and $\phi(z)$ represents a CDF for normal distribution written in (9)-(10).

$$\varphi(z) = \exp(-z^2/2)/\sqrt{2\pi} \tag{9}$$

$$\phi(z) = \int_{-\infty}^{z}\varphi(z)dz \tag{10}$$

The PDF of x with a reference function $\varphi(z)$ up to the 5^{th} cumulant is written in (11), in which z is a normalized variable of x with mean μ_x and standard deviation σ_x. For the CDF of x, the $F(x)$ is calculated using $\phi(z)$ instead of $\varphi(z)$ in $h(x) = \sigma f(x)$.

$$\begin{aligned} f(x) = (\varphi(z) &- \kappa_3\varphi^{(3)}(z)/3\,! + \kappa_4\varphi^{(4)}(z)/4\,! \\ &- \kappa_5\varphi^{(5)}(z)/5\,! \\ &+ 10{\kappa_3}^2\varphi^{(6)}(z)/6\,!)/\sigma \end{aligned} \tag{11}$$

3. UNCERTAIN LOAD AND PHOTOVOLTAIC GENERATION MODELLING

3.1 Probabilistic Load Modelling

The normal distribution function is considered to represent the active and reactive power of the load buses (Constante-Flores and Illindala, 2019). Buses with load are treated as PQ buses and the main feeder is the reference bus. Let L represent the load, the normal PDF for L is written in (12).

$$f(L) = \exp(-(L - \mu_L)/(2\sigma_L^2))/\sqrt{2\pi\sigma^2} \tag{12}$$

in which μ_L is the mean, and σ_L is the standard deviation for the variable L.

2.2 Probabilistic Photovoltaic Generation Modelling

For the buses where PV generations is installed, this non-dispatchable form of generation can be considered as PQ buses with a unitary power factor (Constante-Flores and Illindala, 2019). The PV power injected is intermittent due to the stochastic nature of solar irradiation. To calculate the PV injected power, the solar irradiance s $[kW/m^2]$ is modelled using the Beta distribution (Dhayalini and Mukesh, 2019).

$$f_\beta(s) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}s^{(\alpha-1)}(1-s)^{\beta-1} \tag{13}$$

In which, $f_\beta(s)$ represents the Beta PDF of s in (13), $\Gamma(\cdot)$ represents the gamma function, and the parameter α, β are determined in terms of the mean and standard deviation of s calculated using (14) and (15).

$$\alpha = \mu^2(1-\mu)/\sigma^2 - \mu \tag{14}$$

$$\beta = \alpha(1-\mu)/\mu \tag{15}$$

The PV injected power P_{output} [W] of N modules is calculated using (16)-(20).

$$P_{output}(s) = \mathrm{N} * \mathrm{FF} * \mathrm{V}_s * I_s \tag{16}$$

$$\mathrm{FF} = (\mathrm{V}_{MPP} * \mathrm{I}_{MPP})/(\mathrm{V}_{OC} * I_{SC}) \tag{17}$$

$$T_C = T_A + \mathrm{s} * (N_{OCT} - 20)/0.8 \tag{18}$$

$$\mathrm{V}_s = \mathrm{V}_{OC} + K_V * T_C \tag{19}$$

$$\mathrm{I}_s = \mathrm{s} * (I_{SC} + K_I * (T_C - 25)) \tag{20}$$

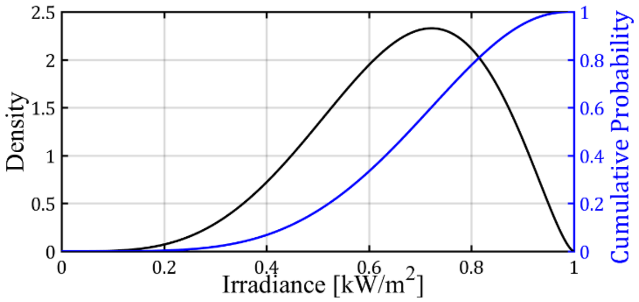

Fig. 1 PDF and CDF of solar irradiance- Summer

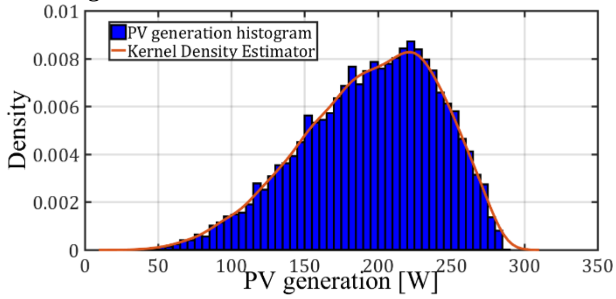

Fig. 2 PDF of HIT module power output

Table 1 Technical specification HIT PV module

V_{MPP}	I_{MPP}	V_{OC}	I_{SC}	N_{OCT}	K_V	K_I
58.0	5.70	69.7	6.07	44	-0.164	0.00334

in which, FF is known as the fill factor, V_s [V] and I_s [A] is the actual output voltage and current. V_{MPP} [V] and I_{MPP} [A] are the maximum power point voltage and current.

$\boldsymbol{V_{OC}}$ [V] and $\boldsymbol{I_{SC}}$ [A] are the open-circuit voltage and short-circuit current. $\boldsymbol{T_C}$, $\boldsymbol{T_A}$ and $\boldsymbol{N_{OCT}}$ are cell, ambient and nominal operating cell temperature [°C]. Finally, $\boldsymbol{K_V}$ [$\mathbf{V}$/°C] and $\boldsymbol{K_I}$ [$\mathbf{A}$/°C] are the voltage and current temperature coefficient.

This paper assumes an expected solar irradiation of 0.663 $\boldsymbol{kW/m^2}$ and a standard deviation of 0.162 $\boldsymbol{kW/m^2}$ for a period between 12:00h and 13:00h of a summer day as in (Dhayalini and Mukesh, 2019). The PDF and CDF of solar irradiance for this period of time is presented in Fig. 1.

To evaluate the injected PV power, the technical specification in Table 1 of HIT module is considered since it outperforms others types of technology in terms of generated energy and area ratio (Mesquita *et al.*, 2019). The PDF of the PV power output for $\mathbf{N = 1}$, and considering $\boldsymbol{T_A} = \mathbf{31}$°C is presented in Fig. 2.

4. APPLICATION OF THE *2m*+1 PEM TO THE POWER FLOW ANALYSIS

Based on the above probabilistic techniques, the procedure to perform the PPF based on the *2m+1* PEM is shown in Fig. 3.

Step 1: Read data for load modelling and data of solar irradiation for PV generation modelling;
Step 2: Compute the *2m+1* PEM method to determine the estimated points taken from the uncertain variables using (2).
Step 3: Execute deterministic power flow analysis with each estimating point in order to obtain the raw moments for a variable of interest (e.g. voltage, current, branch power flow);

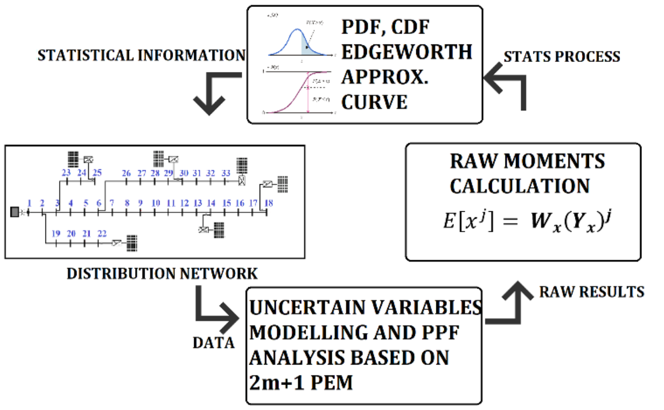

Fig. 3 PPF based on the *2m+1* PEM computational procedure

Step 4: Calculate the approximate curve for PDF and CDF of a variable of interest using (11).

5. CASE STUDY AND RESULTS COMPARISON

The PEM was applied to the IEEE 33-bus radial distribution test system, which consist of 33 buses and 32 lines. The modified distribution is represented using single line diagram with high penetration of PV systems shown in Fig. 6. The system has a rated voltage of 12.66 kV, and a total demand of 3.715 MW and 2.3 MVAr. The original data of the aforementioned systems is taken from the work of Vita (Vita, 2017).

The Monte Carlo simulation performs 10 000 deterministic power flow analysis. This high sampling is intended to guarantee a good accuracy of the estimated values. Thus, the MCS results are taken as reference to verify the accuracy of the *2m+1* PEM. The implementations of the probabilistic methods were done in MATLAB on a computer with an Intel Core i5 2.4GHz processor with 6GB of RAM.

This paper also considers the following data for the analysis of the PPF with high penetration of PV systems. For load modelling, normal distribution is considered standard deviations arbitrarily set as follow: 5% for buses #1-#6, 7% for buses # 19 -#22, 7% for buses #23-#25, 10% for buses #7-#18 and for buses #26 -#33. It is assumed a maximum power of the PV system equal to 1.5 MWp. The PDF of solar irradiance in Fig. 1 is considered for the stochastic PV system power. The current limits for lines #1-#9 are 400, and for lines #10-#32 are 200 A.

The following cases are considered: Case 1, the distribution network without PV systems; Case 2, 2 PV systems are connected at buses #18 and #30; Case 3, 4 PV systems are connected at buses #22, #25, #18, and #30; Case 4, 6 PV systems are connected to buses #25, #22, #30, #18, #14, and #33.

5.1 Index for Accuracy

Variables of interest are generally large. Therefore, the average root mean square index provides a degree of accuracy for the quality of an estimator. The average root mean square (ARMS) index in (21) is defined to quantitatively assess the performance of the *2m+1* PEM compared to the MCS.

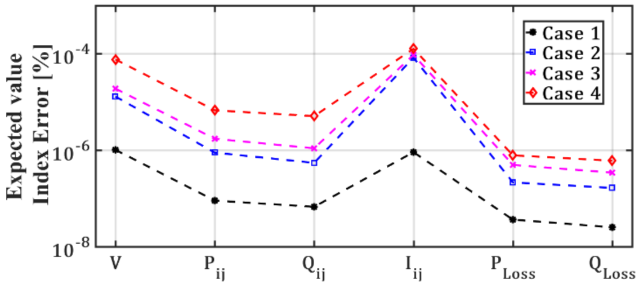

Fig. 4 Variable of interest Index Error - Expected value for Case 1-4

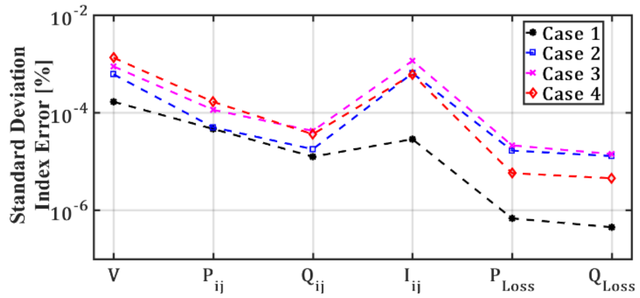

Fig. 5 Variable of interest Index Error – Standard Deviation for Case 1-4

$$\text{ARMS}^{\mu,\sigma} = \frac{1}{\text{n}} * \left(\sqrt{\sum_{i=1}^{n} \left(X_{MCS,i} - X_{PEM,i} \right)^2} \right) \qquad (21)$$

in which X represents the type of variable of interest for mean or standard deviation values, $X_{MCS,i}$, $X_{PEM,i}$ are the *i-th* point of the variable $\boldsymbol{X}$ for the MCS and the PEM method, and **n** is the total sample of a specific variable.

The results for the ARMSs in Fig. 4 and Fig. 5 indicate that the *2m+1* PEM method can accurately estimate the expected and standard deviation values, with index errors is closer to zero for the variables of interest; thus, demonstrating the robustness of the *2m+1* PEM method. The variables of interest show in Fig. 4 and Fig. 5 are the voltage ($\boldsymbol{V}$), active and reactive power flow ($\boldsymbol{P_{ij}}, \boldsymbol{Q_{ij}}$), line current ($\boldsymbol{I_{ij}}$), active and reactive energy loss ($\boldsymbol{P_{Loss}}, \boldsymbol{Q_{Loss}}$).

5.2 Probabilistic Power Flow Analysis

The PPF provides a more realistic characteristic of the distribution network state variables under the stochastic presence of load and photovoltaic generation. The *2m+1* PEM method has a good degree of accuracy for the expected and standard deviation values, as shown in Table 2 for the expected and standard deviation up to three decimal places of accuracy for the main feeder supply and total energy loss.

According to Table 2, the main feeder experieces a reverse power flow for Case 3 and Case 4. In the later, the integration of 6 PV systems increased the total energy loss due to significant reverse power flow through lines. The total energy loss is comparable to Case 1.

Similarly, the integration of 4 PV systems (Case 3) causes reverse power flow in the main feeder, though its value is only 3.92% compared to feeder utilization in Case 1. Furthermore,

Table 2 Main feeder and total energy loss statistical information – expected and standard deviation

			Case 1	Case 2	Case 3	Case 4
Main Feeder supply	P [MW]	μ	3.919	1.795	-0.226	-2.125
		σ	0.074	0.338	0.469	0.541
	Q [MVAr]	μ	2.437	2.380	2.384	2.481
		σ	0.072	0.073	0.073	0.083
Total Energy loss	P [MW]	μ	0.203	0.109	0.107	0.238
		σ	0.009	0.013	0.015	0.056
	Q [MVAr]	μ	0.135	0.080	0.084	0.181
		σ	0.006	0.012	0.014	0.042

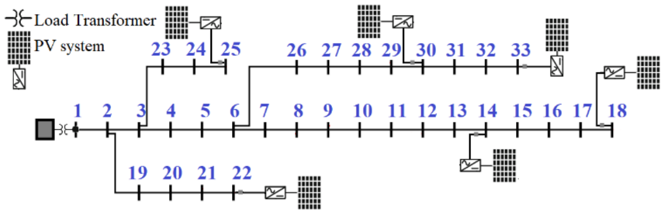

Fig. 6 Single line diagram IEEE 33-bus with PV systems

this case presents greater energy loss reduction compared to the other cases. On the other hand, the main feeder utilization is aliviated in due to integration of 2 PV systems (Case 2) which are located far from the main feeder. Furthermore, the total energy loss is significantly reduced.

The expected and standard deviation for voltage and branch power flow behaviour in the distribtuion network is depicted in Fig. 7 to Fig. 10. It can be seen that the variation of the standard deviation increases when PV systems are installed in the distribution network as shows in Cases 2,3 and 4.

Another important aspect to point out is that buses where PV systems are installed improve voltage levels at the expense of reverse power flow to the adjacent bus. For instance, Case 2 presents power flow reversion which not only relates to line #17 and #29 which are closer to the PV installations but also from lines #7 to #16 and from lines #25 to #26. For Case 3, the installation of 4 PV systems at buses #18, #22, #25, and #30 increases reverse power flow for Line #1, and from line # 7 to #29. Similarly, Case 4 presents reverse power flow from almost all lines towards the main feeder.

Regarding the power flow reversion, the presence of 2 PV systems in Fig. 8 modifies the traditional one direction power flow shown in Fig. 7. It is also observed for Case 2 in Fig. 8 that line #17 presents the highest power flow reversion.

In planning, the confidence interval of power flow through lines is of great importance to prevent overloading. For instance, using the *2m+1* PEM an 80% confidence interval of power flow through line #17 is between [-1.2262; -0.5989] MW for the Case 2. Similarly, for Case 3, it is calculated the 80% confidence interval of the maximum power flow reverse for line #21 and its confidence interval is [-1.2277; -0.5965] MW.

For voltage analysis, it can be observed that PV systems improves the voltage level for all cases. However, Case 4 presents overvoltages in different buses, thus, the probability

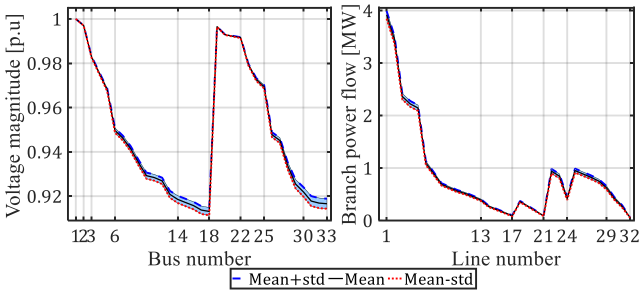

Fig. 7 Voltage and branch power flow for Case 1

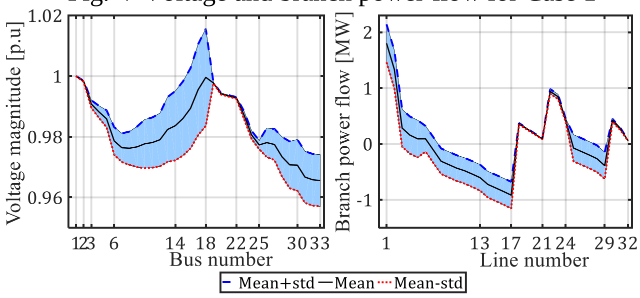

Fig. 8 Voltage and branch power flow for Case 2

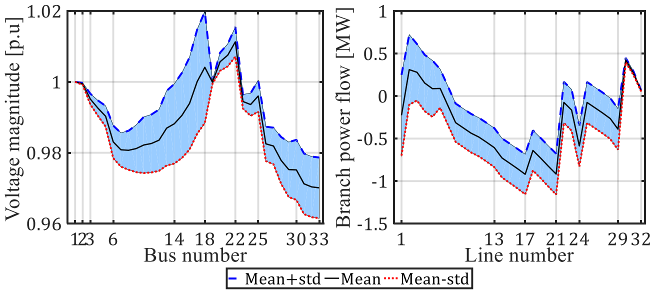

Fig. 9 Voltage and branch power flow for Case 3

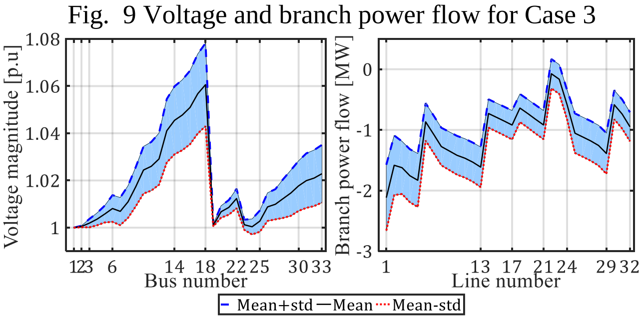

Fig. 10 Voltage and branch power flow for Case 4

of overvoltage ocurrence increased with the integration of 6 PV systems. Furthermore, Case 4 shows a crititcal power flow reversion and this may compromise the assets utilization in the distribution network.

5.3 PDF and CDF analysis

The *2m+1* PEM method is combined with the Edgeworth series approximation in order to obtain the PDF and CDF of a variable of interest.

For the MCS fitting curve of PDF and CDF, the kernel function estimator is used to describe the PDF and CDF of the variables of interest. This technique is used to visualize the shape of MCS data as a reference curve from the obtained histogram. MCS is also approximated using the normal distribution.

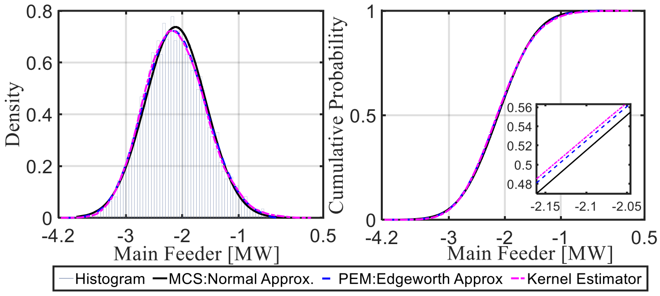

Fig. 11 Density and Cumulative probability function of the main feeder for Case 4

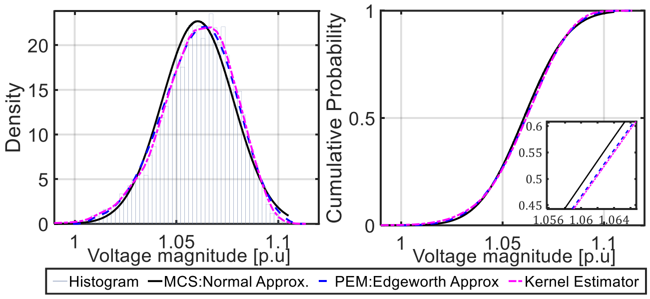

Fig. 12 Density and Cumulative probability function of voltage at bus 18 for Case 4

For the *2m*+1 PEM method, the Edgeworth expansion is used to represent the curve approximation of PDF and CDF of a variable of interest.

The PDF and CDF for the main feeder is shown in Fig. 11 and for the voltage at bus #18 in Fig. 12 considering the Case 4, where 6 PV systems are installed. For the main feeder in Case 4, it is observed that the reverse power flow behaves as a Normal distribution. However, its base distribution has a positive skewness, and the Edgeworth expansion has a better approximation. Furthermore, the CDF of the reverse power flow in the main feeder has a high degree of accuracy as shown in Fig. 11.

For the voltage at bus #18, which presents the maximum voltage in Case 4; it is observed that its PDF is affected by the dominance of the PV generation distribution probability, whose skewness value is negative. The Edgeworth expansion shows a better approximation both the PDF and CDF as shown in Fig. 12 compared to the normal distribution which presents an acceptable approximation.

The curve approximation PDF and CDF using *2m+1* PEM combined with Edgeworth expansion overcomes the lack of perfect knowledge from the stochastic variable probability function as shown in Fig. 11 in Fig. 12. Therefore, this approximation requires few statistical moments to represent with high accuracy the curve approximation.

Table 3 provides the information regarding the shape of the probability distribution in terms of skewness and kurtosis for voltages, branch power flow, current and total energy loss. This information is taken from the MCS to understand how the integration of PV systems affects the behaviour of state variables. For a short description, skewness is the degree of distortion from the symmetrical bell curve. If the skewness is

Table 3 Average values of Skewness and Kurtosis for variables of interest in each considered case

	Case 1		Case 2		Case 3		Case 4	
	$\lambda_{r,3}$	$\lambda_{r,4}$	$\lambda_{r,3}$	$\lambda_{r,4}$	$\lambda_{r,3}$	$\lambda_{r,4}$	$\lambda_{r,3}$	$\lambda_{r,4}$
V	0.03	3.04	-0.41	2.95	-0.40	2.92	-0.38	2.99
P_{ij}	-0.02	3.01	0.32	2.85	0.40	2.81	0.40	2.86
I_{ij}	0.01	3.02	0.21	3.27	0.23	3.33	-0.16	2.72
S_{Loss}	0.05	3.02	0.40	2.62	0.40	2.82	0.17	2.66

positive, more data relies on the right side of the distribution, and vice versa negative skewness means data relies on the left side.

For skewness between -0.5 and 0.5, the data are fairly symmetrical. Kurtosis is a measure of outliers presented by describing the tails of the distribution, a kurtosis value of three is similar to a normal distribution. If kurtosis is greater than 3, tails are fatter (profusion of outliers). If kurtosis is less than 3, tails are thinner (lack of outliers). From the information of Table 3, the variables of interest for Case 1 presents a similar distribution that matches a normal distribution. Case 2, Case 3 and Case 4 differ from a proper Normal distribution; the voltage variable stands out with a negative skewness (e.g. Fig. 12). Similarly, it can be observed for the power flow through lines that presents a positive skewness (e.g. Fig. 11).

5.4 Confidence Interval

The CDF provides information regarding how probable a value is below or above in the distribution data. For the main feeder, the estimated 80% and 90% confidence interval is given in Table 4 and Table 5 by using a normal approximation for MCS, the Edgeworth approximation for *2m*+1 PEM, and the kernel estimator for MCS.

For example, the range that covers 80% of possibilities for the power through the main feeder means that there is an 80% of chance the power through the main feeder falls somewhere between this range. This information is of great interest for operation and planning due to the fact the main feeder should not exceed its capacity. Furthermore, it is observed in Case 3 and Case 4, the presence of power flow reversion towards the main feeder, which is uncommon for distribution networks operation.

5.5 Analysis of Computational Effort

The computation effort of the *2m+1* PEM is of great interest for practical applications. For this reason, the analysis of accuracy and time effort is necessary. The required calculation time for 10000 trials of MCS is 220.5 seconds and for the *2m+1* PEM method is just 4.1 seconds. The approximate method time is less time consuming compared to MCS; and this method is almost 50 times faster.

6. CONCLUSIONS

The probabilistic analysis is essential to cater for the uncertainties introduced by load and non-dispatchable generation in order to control and minimize the risk associated with operation and planning in power systems. This work presents the probabilistic power flow based on the *2m+1* point

Table 4 Main Feeder [MW] - 80% Confidence Interval for Cases 1-4

	Main Feeder [MW] - 80% Confidence Interval					
	MCS: Normal		PEM: Edgeworth		MCS: Kernel	
Cases	10%	90%	10%	90%	10%	90%
Case 1	3.823	4.014	3.821	4.016	3.821	4.015
Case 2	1.378	2.241	1.362	2.242	1.368	2.250
Case 3	-0.819	0.395	-0.831	0.391	-0.826	0.404
Case 4	-2.817	-1.419	-2.816	-1.426	-2.826	-1.405

Table 5 Main Feeder [MW] - 90% Confidence Interval for Cases 1-4

	Main Feeder [MW] - 90% Confidence Interval					
	MCS: Normal		PEM: Edgeworth		MCS: Kernel	
Cases	5%	95%	5%	95%	5%	95%
Case 1	3.794	4.041	3.795	4.042	3.793	4.042
Case 2	1.281	2.389	1.259	2.373	1.272	2.400
Case 3	-0.970	0.577	-0.982	0.565	-0.982	0.588
Case 4	-2.982	-1.223	-2.989	-1.232	-2.996	-1.199

estimate method considering a high penetration of PV systems in the IEEE 33-bus radial distribution tests system.
Reverse power flow, overvoltage and feeder energy supply and energy losses are estimated. Normal and Beta distribution have been used to model the load behaviour and the solar irradiation for a specific period of time in a summer season.

The expected and standard deviation of state variables (e.g. voltage, branch power flow, current, energy loss) are accurately estimated for the distribution network; this calculation is based on average root mean square results, which are closer to zero for all considered cases. Therefore, the results show an acceptable error of expected and standard deviation for variables of interest.

The PV system voltage support benefits is clearly seen among the considered cases, however; it introduces reverse power flow in the system up to a point of increasing energy loss (e.g. Case 4).The PDF and CDF of voltage and branch power flows is mainly affected by the location and generation where PV is installed as shown in Fig.6-10, and Table 3.

For the curve approximation of PDF and CDF was used the Edgeworth expansion which shows a high fitting accuracy. Therefore, it can be used for calculation of required probabilities and confidence intervals.

Finally, the computational burden of the *2m*+1 PEM method is considerable lower which is crucial for practical applications. Further studies can be carried out for the implementation of control methods and risk minimization aiming the DER integration to cope with overvoltage and thermal limits, based on the probabilistic PEM method.

ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

REFERENCES

Blinnikov, S. and Moessner, R. (1998) ‘Expansions for nearly Gaussian distributions’, *Astronomy and Astrophysics Supplement Series*, 130(1), pp. 193–205. doi: 10.1051/aas:1998221.

Carstens, D. D. D. S. and da Cunha, S. K. (2018) 'Solar energy growth in Brazil: Essential dimensions for the technological transition', *International Journal of Energy Economics and Policy*, 8(4), pp. 293–302. Available at: https://econjournals.com/index.php/ijeep/article/view/6567.

Chen, P., Chen, Z. and Bak-Jensen, B. (2008) 'Probabilistic load flow: A review', in *2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*. IEEE, pp. 1586–1591. doi: 10.1109/DRPT.2008.4523658.

Constante-Flores, G. E. and Illindala, M. S. (2019) 'Data-Driven Probabilistic Power Flow Analysis for a Distribution System With Renewable Energy Sources Using Monte Carlo Simulation', *IEEE Transactions on Industry Applications*. IEEE, 55(1), pp. 174–181. doi: 10.1109/TIA.2018.2867332.

Dhayalini, K. and Mukesh, R. (2019) 'Optimal setting and sizing of distributed solar photovoltaic generation in an electrical distribution system', in *2019 Innovations in Power and Advanced Computing Technologies (i-PACT)*. IEEE, pp. 1–5. doi: 10.1109/i-PACT44901.2019.8960225.

Fan, M. *et al.* (2012) 'Probabilistic Power Flow Studies for Transmission Systems With Photovoltaic Generation Using Cumulants', *IEEE Transactions on Power Systems*. IEEE, 27(4), pp. 2251–2261. doi: 10.1109/TPWRS.2012.2190533.

Fernandes, A. T. *et al.* (2019) 'Statistical Modeling of Solar Irradiance for Northeast Brazil', in *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE, pp. 386–391. doi: 10.1109/ICRERA47325.2019.8996714.

Giraldo, J. S. *et al.* (2019) 'Probabilistic OPF Model for Unbalanced Three-Phase Electrical Distribution Systems Considering Robust Constraints', *IEEE Transactions on Power Systems*, 34(5), pp. 3443–3454. doi: 10.1109/TPWRS.2019.2909404.

de Lima Vianna, P. *et al.* (2018) 'A General Analysis of the Distributed Generation Impact on Electrical Energy Planning', in *2018 53rd International Universities Power Engineering Conference (UPEC)*. IEEE, pp. 1–6. doi: 10.1109/UPEC.2018.8542020.

Mesquita, D. D. B. *et al.* (2019) 'A review and analysis of technologies applied in PV modules', in *2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*. IEEE, pp. 1–6. doi: 10.1109/ISGT-LA.2019.8895369.

Pender, J. (2014) 'Gram Charlier Expansion for Time Varying Multiserver Queues with Abandonment', *SIAM Journal on Applied Mathematics*, 74(4), pp. 1238–1265. doi: 10.1137/120896815.

Plattner, G., Farah Semlali, H. and Kong, N. (2017) 'Analysis of probabilistic load flow using point estimation method to evaluate the quantiles of electrical networks state variables', *CIRED - Open Access Proceedings Journal*, 2017(1), pp. 2087–2091. doi: 10.1049/oap-cired.2017.0179.

REN21 (2020) *Renewables 2020 Global Status Report*. Available at: https://www.ren21.net/reports/global-status-report/.

Soroudi, A. (2014) *Taxonomy of Uncertainty Modeling Techniques in Renewable Energy System Studies*, *Green Energy and Technology*. Edited by J. Hossain and A. Mahmud. Singapore: Springer Singapore (Green Energy and Technology). doi: 10.1007/978-981-4585-30-9.

Vita, V. (2017) 'Development of a Decision-Making Algorithm for the Optimum Size and Placement of Distributed Generation Units in Distribution Networks', *Energies*, 10(9), p. 1433. doi: 10.3390/en10091433.

Zhang, P. and Lee, Stephen T. (2004) 'Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion', *IEEE Transactions on Power Systems*, 19(1), pp. 676–682. doi: 10.1109/TPWRS.2003.818743.

Zhang, P. and Lee, S.T. (2004) 'Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion', *IEEE Transactions on Power Systems*, 19(1), pp. 676–682. doi: 10.1109/TPWRS.2003.818743.