ANALYSIS OF PROBABILISTIC LOAD FLOW USING POINT ESTIMATION METHOD TO EVALUATE THE QUANTILES OF ELECTRICAL NETWORKS STATE VARIABLES

Guillaume PLATTNER
EDF - R&D – France
guillaume.plattner@edf.fr

Hicham FARAH SEMLALI
Enedis – France
hicham.farahsemlali@enedis.fr

Nicolas KONG
Enedis - France
nicolas.kong@enedis.fr

ABSTRACT

The objective of the paper is to provide a clear and comprehensive analysis of probabilistic load flow using Point Estimation Method, and its accuracy in the evaluation of the quantiles of the state variables of real electrical distribution networks. Three Points Estimation Method (TPEM) has been implemented to evaluate the first four moments of output variables (voltages, currents and active power flows), and several methods to reconstruct the probability density function from moments and calculate the quantiles have been compared, including Generalized Lambda Distribution and Gram-Charlier Development.

Monte-Carlo PLF has been taken as reference to evaluate the accuracy of aforementioned TPEM results and the scope of application of the method in a real electrical distribution system.

INTRODUCTION

The planning of electrical networks consists of an operational process of decision making for the investment in the development and strengthening of the network. This process involves a techno-economical optimization to minimize the cost of overall system while maintaining a good level of quality and continuity of supply of electricity. Since the development of electrical grid, Deterministic Load Flow (DLF) has been used to evaluate most of the electrical constraints such as voltage drop and line flows, in reference "extreme cases" scenarios. But the recent increase of the part of intermittent renewable energy, and the more active role played by consumers, increases the uncertainties in the power production and consumption. In this scenario, DLF lacks a significant amount of information on the behaviour of load flow's (LF's) solution variable, and especially the quantification of the risk of occurrence of the chosen reference scenario. Considering the random nature of renewable energy, uncertain loads and network configuration, it is possible to build a Probabilistic Load Flow (PLF) model to calculate the probability distribution of output variables. The knowledge of the quantiles of output variables is especially interesting to obtain information about the risk of occurrence of constraints on the network, and thus optimize the dimensioning of network economically as well as technically.

Many PLF methods have been proposed to study load flow uncertainty problem [1]. Those methods can be classified as numerical (MC, PEM, UT), analytical (quadratic PLF) and the combination of both (multi-linearization).

The simplest and easiest-to-implement method for PLF is the Monte Carlo (MC) method [2]. MC PLF method relies on repeated random sampling from the probability distributions of input variables. The approximate probability distribution of LF's output variables can be evaluated with a significant number of iterations. The only drawback of MC PLF method is its computational time, which makes it unsuitable to implement in the real system. Analytical methods [3] can decrease the computation burden significantly, but are only accurate for specific probability distributions. The multi-linearization method [4] has shown accurate results compared to analytical method but the simulation time was not decreased significantly.

In [5] and [6], Oke has studied a different numerical method known as enhanced unscented transform (UT) method. In this method, the continuous input random variables are transformed into discrete probability distributions represented by 3 points called sigma points along with their corresponding weights (probability), using moment equalizing technique. The moments of output random variables are then evaluated. Furthermore, a mathematical modification of UT method called BDR has also been studied to reduce the amount of error in the evaluation of moments.

In [7] and [8], PLF using different variations of PEM has been studied. PEM has been modified since Rosenblueth presented a *2m* variation [9]. The modification have resulted in several variations like: *Km* or *Km+1* scheme from Hong [10], *2m+1* from Delgado [8] and 5PEM from Outcalt [11], where *m* is the number of input random variables. The development of PEM is similar to UT method and is based on the transformation of input variable to discrete probability distributions with the same moments. The results of PEM are the first four moments of its output variables whereas the input variable is represented by its first four moments in the case of *2m+1* variation and by its first 8 moments in the case of 5PEM using *4m+1* variation. The low number of evaluations in the case of *2m+1* variation decreases the computational time to make it a potential candidate to be utilized in the real system's studies. The main limitation of PEM found in the literature is its inability to provide accurate 3rd (Skewness) and 4th (Kurtosis) moments of LF's solution variables.

This paper is dedicated to the study of the Three Points Estimation Method (TPEM). The 1st (Mean) and 2nd (Standard deviation) moments of PLF's output variables

have already been evaluated accurately in the aforementioned literature on small networks. However, a comprehensive analysis using real networks, and an evaluation of the precision of Skewness and Kurtosis results, as well as the possibility to evaluate the quantiles of LF's output variables, have not been found in the study of PLF using PEM. This paper will provide a comprehensive analysis of the accuracy of TPEM's results in terms of the first four moments of output variables, and a comparison of the different methods available to reconstruct the variables' Probability Density Function (PDF) from their moments and calculate their quantiles.

PROBABILISTIC LOAD-FLOW METHOD

The studied PLF method is divided in two steps: first, the moments of the output variables are calculated using TPEM, then the variables PDF are reconstructed and the quantiles can be calculated.

Three Points Estimation Method

The point estimation method is a modification of Gaussian quadrature to evaluate the probability distribution of a random variable Z which is a function of single or several random variables as follow:

$$Z = f(X_1, X_2, X_3, \dots, X_m)$$

The objective of PEM is to reduce the number of evaluations of f hence the computation time by replacing the continuous input variables X_k by discrete variables X'_k with distributions represented by few concentration points and their corresponding weights (probability). The discrete variables X'_k are obtained by matching their first few moments with those of X_k, the number of moments to match depending on the number of points of X'_k.

The probability distributions of the X'_k variables with three points are:

$$f_k(x) = w_{k,0}.\delta(x - \overline{X_k}) + w_{k,1}.\delta(x - x_{k,1}) + w_{k,2}.\delta(x - x_{k,2})$$

Where $w_{k,i}, i \in \{0,1,2\}$ is the weight of a point, $\overline{X_k}$ is the mean of random variable X_k, $x_{k,i}, i \in [\![1,2]\!]$ is a point and $f_k(x)$ is the discrete PDF. One point should always be at the mean for *2m+1* technique.

The three points can be calculated using the moment matching technique described in [8], which consists in matching the moments of X_k with those of X'_k as follow:

$$1 = w_{k,0} + w_{k,1} + w_{k,2}$$
$$0 = \xi_{k,1}.w_{k,1} + \xi_{k,2}.w_{k,2}$$
$$1 = \xi_{k,1}^2.w_{k,1} + \xi_{k,2}^2.w_{k,2}$$
$$S_k = \xi_{k,1}^3.w_{k,1} + \xi_{k,2}^3.w_{k,2}$$
$$K_k = \xi_{k,1}^4.w_{k,1} + \xi_{k,2}^4.w_{k,2}$$

Where σ_k is the standard deviation, S_k is the skewness (normalized third moment) and K_k is the kurtosis (normalized fourth moment) of X_k, and:

$$x_{k,i} = \overline{X_k} + \xi_{k,i}.\sigma_k, i \in [\![1,2]\!]$$

The solutions of the equations will be:

$$\xi_{k,i} = \frac{S_k}{2} + (-1)^{3-i}.\sqrt{K_k - \frac{3S_k^2}{4}} \quad , i \in [\![1,2]\!]$$
$$w_{k,i} = \frac{(-1)^{3-i}}{\xi_{k,i}.(\xi_{k,1} - \xi_{k,2})} \quad , i \in [\![1,2]\!]$$
$$\xi_{k,0} = 0\,;\; w_{k,0} = 1 - w_{k,1} - w_{k,2}$$

The goal is now to approximate the first four moments of output random variable Z, function of multiple random variables X_k, $k \in [\![1,m]\!]$. The process consists of evaluating Z at each one of the three points $x_{k,i}, i \in [\![0,2]\!]$ of all the input random variable as follow:

$$Z_{k,i} = f(\overline{X_1}, \dots, x_{k,i}, \dots, \overline{X_m}) \quad , i \in [\![1,2]\!], k \in [\![1,m]\!]$$

During each evaluation, a single random variable is held at one of its non-mean points and all other input variables are held at their mean values. This will produce *2m* evaluations of output variable. The (2m+1)th evaluation will be performed at mean of all input variables as follow:

$$Z_0 = f(\overline{X_1}, \overline{X_2}, \dots, \overline{X_m})$$

When TPEM is used for PLF, the variable Z could be any of the output variables of PLF such as the voltage at a node or current flowing in a line.

The jth raw moment of voltage Z_n in the nth node can be evaluated as:

$$E(Z_n^j) = \sum_{k=1}^{m}\sum_{i=1}^{2}(w_{n,k,i}.Z_{n,k,i}^j) + w_{n,0}.Z_{n,0}^j$$

Where $w_{n,0} = 1 - \sum_{k=1}^{m}\sum_{i=1}^{2}(w_{n,k,i})$

The first four normalized moments can then be calculated from the raw moments as follow:

$$\mu(Z_n) = E(Z_n^1)$$
$$\sigma(Z_n) = [E(Z_n^2) - E(Z_n^1)^2]^{1/2}$$
$$S(Z_n) = \frac{E(Z_n^3) - 3.E(Z_n^2).E(Z_n^1) + 2.E(Z_n^1)^3}{\sigma(Z_n)^3}$$
$$K(Z_n) = \frac{E(Z_n^4) - 4.E(Z_n^3).E(Z_n^1) + 6.E(Z_n^2).E(Z_n^1)^2 - E(Z_n^1)^4}{\sigma(Z_n)^4}$$

Where μ, σ, S, K are mean, standard deviation, skewness and kurtosis respectively.

Aforementioned derivation of moments of an output random variable Z is basically derived using Taylor series [12]. During the Taylor series derivation of TPEM only first two terms of the series have been considered to avoid the complexity of the process. The elimination of higher order terms of Taylor series is the root cause of residual error in the higher order moments of variable Z.

Reconstruction of the Probability Density Functions

Several methods exist to reconstruct the PDF of a variable from its moments. Since TPEM provides the first four moments of output variables, with an accuracy decreasing with the order of the moment (see below), we will compare methods using the first two, three or four moments of a variable. The methods considered are listed below:

- **Simplification as a Gaussian Distribution (SGD)**

The simplest method consists in making the assumption

that the variable follows a Gaussian distribution. Only the mean μ and the standard deviation σ calculated with TPEM are thus kept. For a Gaussian distribution, the quantile q_p can be calculated as follow:

$$q_p = \mu + \sigma . K_p \text{ , with } K_p = \sqrt{2}.erf^{-1}(2p-1)$$

- **Third order Gram-Charlier development (3GC)**

The Gram-Charlier development consists in modifying a reference PDF (usually the Gaussian distribution) by multiplying it by a polynomial function, in order that the moments of the resulting PDF fit those of the variable we want to calculate. For a third order development, only the first three moments μ, σ and S are used, and the PDF of the variable is thus given by:

$$f_z(x) = N(\mu,\sigma)(x).\left[1 + \frac{S}{6}.H_3\left(\frac{x-\mu}{\sigma}\right)\right]$$

Where $N(\mu, \sigma)$ is the Gaussian distribution, and $H_3(x) = x^3 - 3x$

The quantiles can then be calculated numerically from the PDF.

- **Fourth order Gram-Charlier development (4GC)**

For a fourth order development, the first four moments μ, σ, S and K are used, and the PDF of the variable is thus given by:

$$f_z(x) = N(\mu,\sigma)(x).\left[1 + \frac{S}{6}.H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{K-3}{24}.H_4\left(\frac{x-\mu}{\sigma}\right)\right]$$

- **Generalized Lambda Distribution**

The GLD is a continuous probability distribution defined by four parameters ($GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$) in terms of its inverse distribution function (Quantile function) as follow:

$$Q(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2}$$

Where $Q(p)$ is the quantile function, inverse of the cumulative function $F(x) = p$, $p \in [0,1]$.

The objective of PDF reconstruction is here to calculate the four parameters of GLD using a moment matching method. The idea behind it is to equalize the moments of a variable calculated using TPEM to the ones of theoretical GLD to calculate its parameters. The moments of GLD are functions of λ_1, λ_2, λ_3 and λ_4, and the resolution of these equations can be done numerically, and is well documented.

In our study, three variations of GLD were considered, depending on the number of moments from TPEM to be kept, the other ones being fixed to the values of a Gaussian distribution:

- **GLD using two Moments (GLD2)**: μ and σ from TPEM are kept, S is fixed at 0 and K at 3.
- **GLD using three Moments (GLD3)**: μ, σ and S from TPEM are kept, K is fixed at 3.
- **GLD using four Moments (GLD4)**: μ, σ, S and K from TPEM are kept.

CASE STUDIES

The TPEM and the different methods to calculate the quantiles have been tested in different cases: three different networks, and four different scenarios for the PDF of input variables.

Networks

The three networks are existing middle-voltage (20 kV) radial distribution networks.

- **Network 1**

Network 1 corresponds to a semi-urban feeder, comprising 314 nodes, 225 lines, 78 loads (2 MV clients and 76 distribution substations), and 1 producer. The total average consumed power is 5.113 MW, and the producer's power is 0.483 MVA.

- **Network 2**

Network is the same that Network 1, except that 3 big producers are added. The total average produced power is now 4.48 MVA.

- **Network 3**

Network 3 corresponds to all feeders departing from a main substation: 5 rural feeders and 1 dedicated feeder to a big consumer, for a total of 891 nodes, 727 lines and 264 loads (14 MV clients and 250 distribution substations). There is no producer on this network. The total average consumed power is 9.705 MW.

Scenarios

Several PDF have been used to simulate the distribution of consumed and produced powers in 4 scenarios. For every load and producer, the mean power is the same in every scenario. Furthermore, the standard deviation of the power distribution is arbitrarily set at 1/5 of the mean power in every scenario, except for Rayleigh distributions (see below). Finally, we have considered a constant $\cos \varphi$ for every load and producer.

- **Scenario 1**

Every load and producer follows a Gaussian distribution. Every input variable has thus a Skewness equal to 0, and a Kurtosis equal to 3.

- **Scenario 2**

Every load and producer follows a log-normal distribution. The location and scale parameters of the distributions are calculated in order to respect the values of the mean and standard deviation defined previously. Every input variable has thus a Skewness equal to 0.608 and a Kurtosis equal to 3.66.

- **Scenario 3**

Every load follows a Gaussian distribution, and every producer follows a Rayleigh distribution. The Rayleigh distribution having only one parameter, the scale parameter, its value is calculated in order to respect the value of the mean (μ) defined previously. The standard deviation is then equal to $\mu.\sqrt{\frac{4-\pi}{\pi}}$, the Skewness is equal to 0.63 and the Kurtosis is approximately equal to 3.24.

- **Scenario 4**

Every load and producer follows a Gaussian distribution, and the loads are correlated with each other. This correlation is modeled by adding a variable, which will be common to every load. The power P of a load is thus equal

to the sum of two random variables, independent between them, one peculiar to this load (X_p), and the other common to all loads (X_c). These two variables follow Gaussian distributions, and their parameters are calculated in order that the total power of the load follow a Gaussian distribution (by additivity) with a mean and a standard deviation equal to those previously defined. The two variables are weighted in order to represent respectively 80% and 20% of the total variance (arbitrary values). The power of a load is thus given by:

$$P = \sqrt{0.2}.\sigma.X_c + X_p \;\; with \;\; X_c = N(0,1)$$
$$and \; X_p = N(\mu, \sqrt{0.8}.\sigma)$$

Thus:

$$P = N\left(0, \sqrt{0.2}.\sigma\right) + N\left(\mu, \sqrt{0.8}.\sigma\right) = N(\mu, \sigma)$$

RESULTS

The MC method was used as a reference to evaluate the precision of TPEM results in terms of moments and quantiles. 100 000 samplings were made for every PLF using the MC method.

Two indicators are used to evaluate the precision of TPEM results: the Mean Absolute Percentage Error (MAPE) and the Maximum Relative Error (MRE), which are calculated as follow:

$$MAPE(\%) = \frac{100}{nb_{variables}} \times \sum_{variables\ Z} \left|\frac{Z_{MC} - Z_{TPEM}}{Z_{MC}}\right|$$

$$MRE(\%) = 100 \times \max_{variables\ Z} \left|\frac{Z_{MC} - Z_{TPEM}}{Z_{MC}}\right|$$

Moments

The first step consists in evaluating the precision of the moments calculated by TPEM. To do that, the mean, standard deviation, Skewness and Kurtosis calculated by TPEM were compared to those calculated by MC for different variables: the current amplitude (*I*) and the active power (*P*) in every line, and the voltage amplitude (*U*) at every node.

Mean values

The evaluation of mean values of the three variables is in most cases precise, resulting in low MAPE, as shown in table 1, where MAPE and MRE are calculated over every network and every scenario.

VARIABLE	MAPE (%)	MRE (%)
I	0.15	12.44
P	0.41	97.38
U	0.001	0.005

Table 1: Precision of TPEM results in terms of mean values.

For current amplitude and active power, MRE show that the evaluation is bad for certain lines. For active power, this is however not dramatic, since the relative error is high only for very low values of *P*, i.e. lines transiting very few power, as shown in figure 2, resulting in low absolute errors.

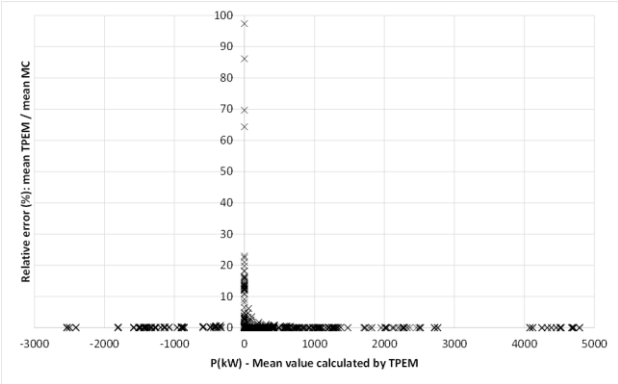

Figure 1: Relative error in the evaluation of mean value of P in every line, for every network and every scenario.

Concerning current amplitude, the most significant errors occur for network 2, and always for lines at the feeder's head, where power flow is sometimes positive and other times negative. This must result in higher non-linearity in the load-flow equations, and thus decrease the precision of the TPEM method.

Standard deviation

The evaluation of standard deviation of the three variables is in most cases also precise, resulting in low MAPE, as shown in table 2, where MAPE and MRE are calculated over every network and every scenario.

VARIABLE	MAPE (%)	MRE (%)
I	1.37	91.77
P	1.01	65.51
U	0.22	0.70

Table 2: Precision of TPEM results in terms of standard deviation.

The high MRE for current amplitude and active power are due to errors happening in the same lines and scenarios that resulted in errors in the evaluation of the mean values.

Skewness

The evaluation of Skewness of the three variables is in most cases not precise, resulting in high MAPE, as shown in table 3, where MAPE and MRE are calculated over every network and every scenario.

VARIABLE	MAPE (%)	MRE (%)
I	99.3	12000
P	126	30000
U	56.2	440

Table 3: Precision of TPEM results in terms of Skewness.

Despite this lack of precision, a good correlation could be observed between the Skewness calculated by TPEM, and the Skewness calculated by MC, for each of the three variables.

It is thus difficult to establish if the calculation of quantiles should use the Skewness calculated by TPEM or not. This issue will be answered below, by comparing the different methods to calculate the quantiles, taking into account the Skewness or not.

Kurtosis

The evaluation of Kurtosis of the three variables is in most cases not precise, resulting in high MAPE, as shown in table 4, where MAPE and MRE are calculated over every network and every scenario.

VARIABLE	MAPE (%)	MRE (%)
I	65.2	40000
P	34.1	1100
U	70.5	23000

Table 4: Precision of TPEM results in terms of Kurtosis.

Unlike for Skewness, there is no correlation between the Kurtosis results of TPEM and of MC, and these results are thus not usable to calculate the quantiles, as will be shown below.

Quantiles

Several methods were described above to calculate the quantiles of the variables from their first few moments. These methods will be used to compute the quantiles 5%, 10%, 90% and 95% of the current amplitude, active power and voltage amplitude. The results will be compared to the quantiles calculated from the MC results, and MAPE and MRE will be calculated to compare the methods.

Due to the imprecision of the Kurtosis results, the 4GC method was in many cases resulting in non-coherent results (probability density functions having negative values), and will thus not be considered.

The results of the different methods are compared in table 5, where MAPE and MRE are calculated over every network and every scenario.

VARIABLE	LN	GC3	GLD2	GLD3	GLD4
I – MAPE (%)	1.79	1.22	1.80	1.17	1.59
I – MRE (%)	69.7	152	173	154	75.3
P – MAPE (%)	1.59	0.40	1.97	0.88	2.10
P – MRE (%)	192	41.1	950	148	908
U – MAPE (%)	0.012	0.004	0.012	0.004	0.034
U – MRE (%)	0.18	0.064	0.18	0.053	0.22

Table 5: Precision of quantiles results for different methods.

From these results, it appears that the GC3 and GLD3 methods are always more precise than the other three methods, meaning that it is beneficial to use the Skewness coefficients calculated by TPEM, even though they are not precise. Since the GC3 method is much faster in terms of computation time compared to GLD, and has better precision for the evaluation of active power quantiles, it may be considered as the best method to calculate quantiles.

CONCLUSION

The results above show that TPEM is precise to evaluate the mean and standard deviations of voltages, currents and power flows, but much less to evaluate their Skewness and Kurtosis coefficients. The overall precision of the quantiles results, compared to those obtained with Monte-Carlo, is very good for voltages (less than 0.1% of relative error), and acceptable for currents and power flows (a significant error being made only for very small values, and thus having no impact for network planning).

This study has thus shown that TPEM is a relevant alternative to the Monte-Carlo method to perform probabilistic load-flows for networks with a high number of nodes and loads. Such a tool enables to calculate the quantiles of networks' state variables, and thus quantify the risk of occurrence of a constraint.

REFERENCES

[1] P. Chen, Z. Chen, B. Bak-Jensen, 2008, "Probabilistic Load Flow: A Review", *DRPT2008.* 6-9, 1586-1591.

[2] P. Jorgensen, J.S. Christensen, J.O. Tande, 1998, "Probabilistic load flow calculation using Monte Carlo techniques for distribution network with wind turbines", *Proceedings Harmonics and Quality of Power conference*, IEEE/PES and NTVA, 1146-1151.

[3] M. Brucoli, F. Torelli, R. Napoli, 1985, "Quadratic probabilistic load flow with linearly modelled dispatch", *Electrical Power and Energy Systems.* vol. 7, 138-146.

[4] A.M. Leite Da Silva, V.L. Arienti, 1990, "Probabilistic load flow by a multilinear simulation algorithm", *Proceedings IEE conference*, vol.137, 276-282.

[5] O.A. Oke, D.W.P. Thomas, G.M. Asher, 2011, "A new probabilistic load flow method for systems with wind penetration", *Proceedings PowerTech IEEE conference*, 2011.

[6] O.A. Oke, 2013, *Enhanced Unscented Transform Method for Probabilistic Load Flow Studies*, Thesis submitted to the University of Nottingham.

[7] G. Verbic, C.A. Canizares, 2006, "Probabilistic Optimal Power Flow in Electricity Markets Based on a Two-Point Estimate Method", *IEEE Transactions on Power Systems* vol. 21, 1883-1893.

[8] C. Delgado, J.A. Dominguez-Navarro, 2014, "Point estimate method for probabilistic load flow of an unbalanced power distribution system with correlated wind and solar sources", *Electrical Power and Energy Systems.* vol. 61, 267-278.

[9] E. Rosenblueth, 1975, "Point estimates for probability moments", *Proc. Nat. Acad. Sci. USA*, vol.72, 3812-3814.

[10] H.P. Hong, 1998, "An efficient point estimate method for probabilistic analysis", *Reliability Engineering & System Safety*, vol.59, 261-267.

[11] D.M. Outcalt, 2009, *Probabilistic load flow for high wind penetrated power systems based on a five point estimation method*, Thesis submitted to the University of Wisconsin-Milwaukee.

[12] C.L. Su, 2005, "Probabilistic Load-Flow Computation Using Point Estimate Method", *IEEE Transactions on power systems*, vol.20, 1843-1851.