Probabilistic load flow in systems with wind generation

J. Usaola

Departamento de Ingeniería Eléctrica, Universidad Carlos III de Madrid, Spain
E-mail: jusaola@ing.uc3m.es

Abstract: A method for solving a probabilistic power flow that deals with the uncertainties of (i) wind generation, (ii) load and (iii) generation availability in power systems is proposed. Dependence between random variables has been considered. The method is based on the properties of cumulants of random variables. Cornish-Fisher expansion series are used to obtain the cumulative distribution function (CDF) of the output variables. Multimodal CDF are obtained by convolutions, whose number has been minimised in order to decrease the computation requirements.

1 Introduction

The great expansion of intermittent generation in power systems has increased their uncertainty. This leads to greater needs of probabilistic analysis tools, both for system planning [1] and for the daily system operation.

Probabilistic power flow is one of the best known probabilistic tools. From the first proposals in the 1970s [2, 3], a great deal of literature can be found about it. The most straightforward method of solving this problem is Monte Carlo simulation [4]. This technique involves repeated simulations with values obtained from the probability density function (PDF) of the considered random variables. For an adequate representation, many simulations must be considered in real systems, which makes sometimes this approach unpractical. One of the alternatives is the convolution of the PDF of the random variables involved, when they are independent of each other, and linearly related. Although this reduces the computational burden, it is costly to obtain the PDF of a line when several random power injections are considered even if fast fourier transform (FFT) [5] techniques are used. Convolution techniques and FFT were also used in [6] for distribution networks with wind energy, where a simplified estimate of the PDF for short-term wind power prediction was made. A multilinear approach was proposed in [7], and fuzzy techniques [8] have been used to solve a probabilistic optimal power flow. Another recent proposal

is the point estimate method [9, 10], which approximates the moments of the system variables of interest. This method has been also applied to the probabilistic optimal power flow problem [11].

The previous approaches assume that the random variables considered are independent. However, dependence between the uncertainties of power injections should be considered for loads and for wind generation. The generalisation of some of these methods for considering the dependence between random variables is very complex, or impossible. There have been proposals that consider this dependence, only between loads, in [12], where it is modelled with a linear relation, or in [13], where the covariance between dependent random variables has been considered.

Probabilistic load flow has mostly included the uncertainty of load. This uncertainty is not usually very high, especially for day-ahead operation, and it can be modelled using Gaussian probabilistic density functions. Wind energy proliferation, however, poses new challenges, since the variability of wind power production is much higher, and the PDF of the uncertainties are not Gaussian.

The use of cumulants and the approximation of a PDF by orthogonal series (Gram-Charlier A expansion series or Laguerre polynomials [14]) have also been recently proposed for the probabilistic power flow [15], with discrete distributions [16], or for the probabilistic optimal

power flow [17]. It has interesting properties, and is computationally inexpensive. For large transmission networks, it seems that it is very adequate because of its low computational requirements. It has the disadvantage of the necessary linearisation but it may be generalised for dependent random variables. However, for non-Gaussian PDF, Gram-Charlier A expansion series have serious convergence problems, and the Cornish-Fisher expansion gives better results without more computational burden [18].

The aim of this paper is to propose an analytical method [called enhanced linear method (ELM)] for the problem of probabilistic load flow. This method improves the approach proposed by the author in [18], by including the dependendent continuous random variables and discrete random variables. The ELM method can be applied to grids where the wind power uncertainty, load uncertainty and generation availability must be considered. Mathematically, this means that the considered random variables may be continuous and/or discrete, and also dependent and/or independent.

The method combines different approaches, but it is based on the cumulant method, generalised for the case of dependent random variables. A refinement is also added to better consider the non-linearity of load flow equations. Cornish-Fisher expansion series are used to evaluate the CDF of the output random variables, and a short number of FFT-based convolutions are made only when the resultant CDF is multimodal. The method has been applied to the daily operation problem, using as inputs the scheduled generation, the forecasted load and the wind prediction for a specific moment of the day ahead. The intended contributions of this paper are the consideration of the dependence between input random variables (for instance the wind power predictions in an area), and to deal with multimodal distributions minimising the required computing time.

The paper is organised as follows. Section 2 gives the statistical background used in the paper. Section 3 gives an overview of short-term wind power prediction. In Section 4, load flow equations are formulated. The computational procedure is explained in Section 5, and Section 6 includes the results of an application example. The Conclusion with the main intended contributions ends the paper.

2 Statistical background

2.1 Moments and cumulants of multivariate distributions

Definitions: For the sake of simplicity, the definitions given here will be for the bivariate case. Generalisation to a higher number of variables is straightforward. More information can be found in [19].

Let x_1 and x_2 be two random variables, $f(x_1, x_2)$ their joint PDF and $F(x_1, x_2)$ their joint bivariate CDF. For these random variables, the joint moment of order $k + r = n$ is

$$m_{kr} = E[x_1^k x_2^r] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^r f(x_1, x_2)\, dx_1\, dx_2 \quad (1)$$

The values $\eta_1 = m_{10}$ and $\eta_2 = m_{01}$ are the means of the marginal distributions. The joint central moments of x_1 and x_2 are the moments of $(x_1 - \eta_1)$ and $(x_2 - \eta_2)$

$$\mu_{kr} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \eta_1)^k (x_2 - \eta_2)^r f(x_1, x_2)\, dx_1\, dx_2 \quad (2)$$

The joint moment generation function of x_1 and x_2 is defined as

$$\phi(s_1, s_2) = E[e^{s_1 x_1 + s_2 x_2}]$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{s_1 x_1 + s_2 x_2} f(x_1, x_2)\, dx_1\, dx_2 \quad (3)$$

And it can be shown that

$$\frac{\partial^k \partial^r}{\partial s_1^k \partial s_2^r} \phi(0, 0) = m_{kr}$$

The joint cumulant generating function is defined as $\psi(s_1, s_2) = \ln \phi(s_1, s_2)$. The values of the derivatives of this function

$$\frac{\partial^k \partial^r}{\partial s_1^k \partial s_2^r} \psi(0, 0) = \kappa_{kr}$$

are by definition the joint cumulants κ_{kr} of order k and r of random variables x_1 and x_2. The mathematical relations between multivariate moments and cumulants are rather complex [20].

Linear combination of random variables: The following equations will also be formulated for two variables. Let z be the random variable $z = a_1 x_1 + a_2 x_2$ where x_1 and x_2 are random variables and a_1 and a_2 are real constants. Then, the moment of order n of variable z, $m_{z,n}$ is

$$m_{z,n} = E[z^n] = E[(a_1 x_1 + a_2 x_2)^n] \quad (4)$$

$$= E\left[\sum_{j=0}^{n} C_n^j (a_1 x_1)^{n-j} (a_2 x_2)^j\right] \quad (5)$$

where

$$C_n^j = \frac{n!}{j!(n-j)!}$$

and thanks to the linearity of the expression

$$m_{z,n} = \sum_{j=0}^{n} C_n^j a_1^{n-j} a_2^j m_{(n-j)j} \quad (6)$$

where $m_{(n-j)j} = E[x_1^{n-j} x_2^j]$. It could be also demonstrated (see [20]) that

$$\mu_{z,n} = \sum_{j=0}^{n} C_n^j a_1^{n-j} a_2^j \mu_{(n-j)j} \tag{7}$$

$$\kappa_{z,n} = \sum_{j=0}^{n} C_n^j a_1^{n-j} a_2^j \kappa_{(n-j)j} \tag{8}$$

where $\mu_{z,n}$ and $\kappa_{z,n}$ are, respectively, the central moment and cumulant of order n of the variable z, and $\mu_{(n-j)j}$ and $\kappa_{(n-j)j}$ are the joint central moment and cumulant of order $(n - j)$ and j of the variables x_1 and x_2. In general, it is better to work with cumulants for the following reasons [20]:

• Most statistical calculations using cumulants are simpler than the corresponding calculation using moments.

• For independent random variables, the cumulants of a sum are the sum of cumulants.

• For independent random variables, the cross-cumulants are zero.

• Series expansion such as the Cornish-Fisher expansion is most conveniently expressed using cumulants.

• Where approximate normality is involved, high-order cumulants can usually be neglected, but not higher order moments.

2.2 Cornish−Fisher expansion series

To approximate the CDF of a random variable from its moments or cumulants is a complex mathematical problem not yet completely solved [21]. In the area of probabilistic power flow, previous works [10, 15] have proposed the Gram-Charlier A series to solve it. However, the convergence of this series is poor when the involved distributions are not Gaussian (see [22] or [23]). Numerical errors may make this convergence worse [24].

The Cornish-Fisher expansion provides a better approximation of a quantile α of a distribution function $F(x)$ in terms of the quantile of a normal $N(0, 1)$ distribution Φ and the cumulants of $F(x)$. Using the first five cumulants, the series expansion can be written as (see [22])

$$x(\alpha) \simeq \xi(\alpha) + \frac{1}{6}(\xi^2(\alpha) - 1)\kappa_3 + \frac{1}{24}(\xi^3(\alpha) - 3\xi(\alpha))\kappa_4$$
$$- \frac{1}{36}(2\xi^3(\alpha) - 5\xi(\alpha))\kappa_3^2 + \frac{1}{120}(\xi^4(\alpha) - 6\xi^2(\alpha) + 3)\kappa_5$$
$$- \frac{1}{24}(\xi^4(\alpha) - 5\xi^2(\alpha) + 2)\kappa_3\kappa_4$$
$$+ \frac{1}{324}(12\xi^4(\alpha) - 53\xi^2(\alpha) + 17)\kappa_3^3 \tag{9}$$

where $x(\alpha) = F^{-1}(\alpha)$ and $\xi(\alpha) = \Phi^{-1}(\alpha)$ and κ_r is the cumulant of order r of the distribution function F. More

information about this series expansion can be found in [21, 25] or [26].

Although the convergence properties of Cornish-Fisher series are difficult to demonstrate [21], their behaviour for non-Gaussian CDF is usually better. A comparison between Gram-Charlier A and Cornish-Fisher expansion series applied to a probabilistic load flow can be found in [18], where this issue is discussed in more detail.

2.3 Multimodal distributions

A multimodal distribution has several modes, and comes from a combination of a continuous and a discrete random variables. In a multimodal distribution, it is not possible to estimate the CDF of the random variable through series expansion from moments, and convolutions must be made. Since these are time consuming, it would be shown here how to minimise their number.

Let z be a linear combination of the discrete random variable x_d and the continuous random variable x_c. Both variables are independent. Then

$$z = x_c + a_d x_d \tag{10}$$

where a_d is a real constant. One approach to the estimate of the PDF of z is the convolution of the PDF of those variables, since they are independent. A property of the Fourier transform is that the transform of a convolution of two functions is the product of their Fourier transforms, that is

$$f_z(z) = f_{x_1}(x_1) * f_{x_2}(x_2) \leftrightarrow \mathcal{F}_z(\omega) = \mathcal{F}_{x_1}(\omega) \cdot \mathcal{F}_{x_2}(\omega)$$

where the asterisk means convolution and $\mathcal{F}$ stands for the Fourier transform of f. An example of a transform of a continuous exponential PDF a discrete binomial PDF and its product is given in Fig. 1. It must be remarked that the Fourier transform of the discrete distribution is a periodic function in the frequency domain. The result of the convolution is the multimodal PDF shown in Fig. 2. Not all linear combinations of discrete and continuous variables render z a multimodal distribution. This depends on the nature of both, and on the coefficient a_d.

Since the convolution is a computationally expensive procedure, if there are many discrete variables and many combinations to be considered, the computational burden increases hugely. Therefore it could be found under which conditions the convolution of a discrete and a continuous variable yields a multimodal PDF in order to reduce the number of convolutions as much as possible. When the distribution is unimodal, the Cornish-Fisher series expansion provides a good approximation.

To obtain the condition for a multimodal distribution, it is better to work in the frequency domain. The condition is that the transform of the continuous variable has a sufficiently low
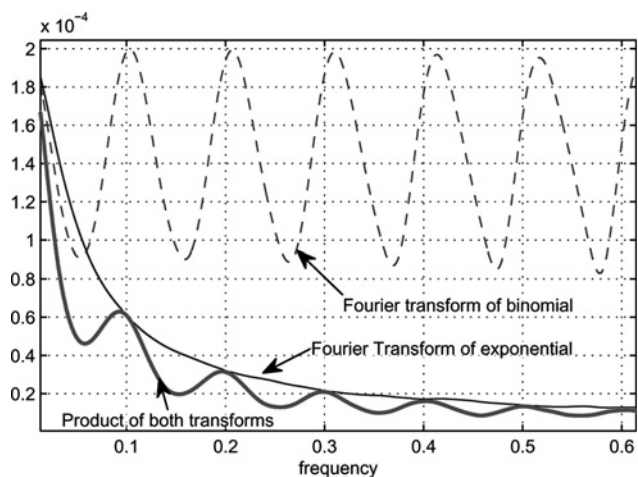
Figure 1 *Transforms of a discrete and continuous PDF, and their product*
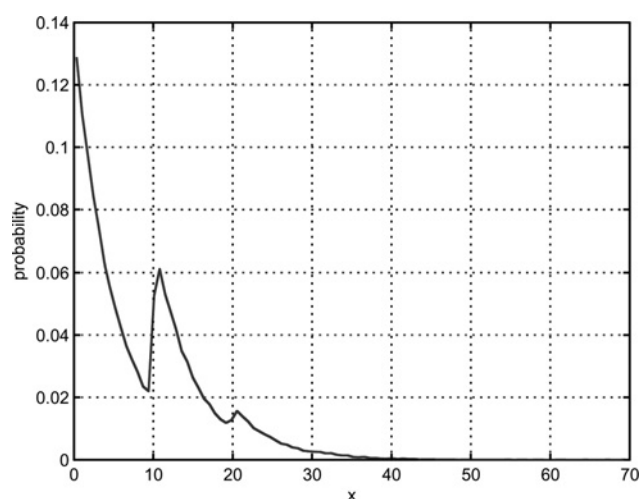
Multimodal case

Figure 2 *Multimodal PDF from a combination of a discrete and continuous functions*

value at the frequency of the first minimum of the Fourier transform of the discrete variable. This minimum is given by

$$f_{\mathrm{lim}} = \frac{1}{2a_{\mathrm{d}}\Delta x_{\mathrm{d}}} \qquad (11)$$

where Δx_{d} is the separation between two non-zero values of the discrete distribution. Hence, the unimodality condition is

$$|\mathcal{F}_{\mathrm{c}}(f_{\mathrm{lim}})| < \varepsilon_{\mathrm{f}} \qquad (12)$$

where ε_{f} is a number sufficiently small and $\mathcal{F}_{\mathrm{c}}$ is the Fourier transform of the continuous function. If this condition is fulfilled, the result of the convolution will be an unimodal function, and a Cornish-Fisher expansion may be applied to estimate the CDF. Otherwise, a convolution should be made.

2.4 Generation of correlated random numbers

The uncertainty of dependent short-term wind power predictions is modelled in this paper as a multivariate Beta

PDF (see Section 3). Random numbers with this distribution are not easily generated, and that is why a method is given here to generate multivariate random numbers with any distribution.

The method is based on the inverse transformation of a uniform distribution. Let x_1 be a random variable with uniform distribution $U(0, 1)$ (the CDF of a $U(0, 1)$ is $H(x) = x$), and let us take n random samples of this variable. Let $x_{1,k}$ be the sample k. Then, to generate random numbers, $x_{2,k}$, of a given distribution with invertible CDF $F(x)$, it is necessary to perform the operation $x_{2,k} = F^{-1}(x_{1,k})$. If we transform this new variable, forming $x_{3,k} = F(x_{2k})$, then x_{3k} will have again a $U(0, 1)$ distribution.

The method begins by generating random numbers of a multivariate normal random variable, with a given correlation matrix, forming the array $\boldsymbol{x}_1 \in \boldsymbol{R}^{m,n}$, where m is the number of variables and n the sample size. Each element is written as $x_{1,ij}$, where i is the variable and j the sample. These numbers can be easily generated by standard programs.

In a second step, a normal transformation is made to these values in order to obtain a multivariate uniform distribution, $\boldsymbol{x}_2$. That is to say, $\boldsymbol{x}_2 = F(\boldsymbol{x}_1)$, where $F(\boldsymbol{x})$ is a multivariate normal CDF.

The third step consists in transforming the obtained multivariate uniform random numbers into series with the wished marginal distributions G. Then, sample j of the new variable i $x_{3,ij}$ will be obtained as $x_{3,ij} = G^{-1}(x_{2,ij})$. The obtained random numbers with the wished distribution will have correlations very close to the original one of the normal multivariate distribution. The difference is due to the non-linearity of the transformation performed.

A more complex approach, such as copula modelling, could be used to preserve the desired correlation among variables, that is in the origin of the transformations (see, for instance, [21] or [27] for wind power), but to go deeper into this problem is beyond the aim of this paper.

3 Uncertainty of short-term wind power prediction

Short-term wind power prediction programs are tools that provide an estimate of future power production of a wind farm for the next hours. They use meteorological forecasts from a Numerical Weather Prediction tool, and sometimes real-time SCADA data from the wind farms. Data of the wind farms, such as rated power, type and availability of wind turbines etc., are also necessary. The output of these programs is the hourly average wind farm production for the next hours. Typically, predictions are issued for the next 48 h, but longer time horizons are possible. The accuracy of these tools are smaller than load prediction programs and decreases with the time horizon. A survey of the accuracy of these tools is given in [28].

The forecasts provided by a short-term wind power prediction program are uncertain. This uncertainty changes with the range of the wind farm power output, since this value is bounded between zero and the rated power. Besides, the power curve of a wind turbine or wind farm is non-linear. If we assume that the wind speed forecasts have Gaussian uncertainty, then the PDFs of the power predictions will not be Gaussian. The shape of these PDFs is also affected by the time lag elapsed between the prediction and the operation times. A sample of an heuristical PDF of the uncertainty of short power prediction is given in Fig. 3 This function shows the uncertainty of a wind power prediction made with a time horizon of 7 h when the forecasted power was 0.2 p.u.

It is not within the purposes of this paper to model this uncertainty, and a reasonable approximation will be used. Owing to the bounded nature of the power produced by a wind farm, a beta distribution will be used, as in [29]. In our case, the mean of the distribution will be the predicted power at the time of interest, whereas the standard deviation σ will depend on the level of power injected, with respect to the wind farm rated power. This dependence has been obtained heuristically for some wind farms, and the results are shown in Fig. 4, where the value of standard deviation is normalised to the rated power of the wind farm. Although there are wide variations, an approximation by a quadratic curve (shown in the picture) may provide realistic results.

The uncertainty of short-term wind power prediction of geographically close wind farms are correlated, since the wind power in all of them are due to similar meteorological conditions. This dependence has not yet been modelled, but studies such as [30] show the dependence between productions in a wide area. These results may be considered as an estimate of actual correlation values, although it is necessary to wait for more specific studies.
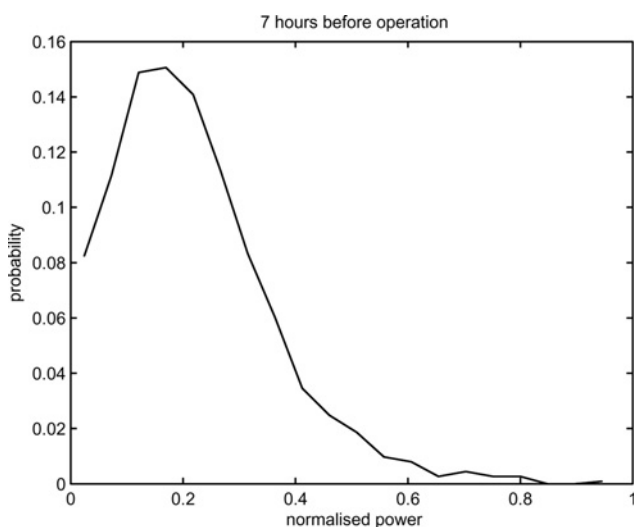

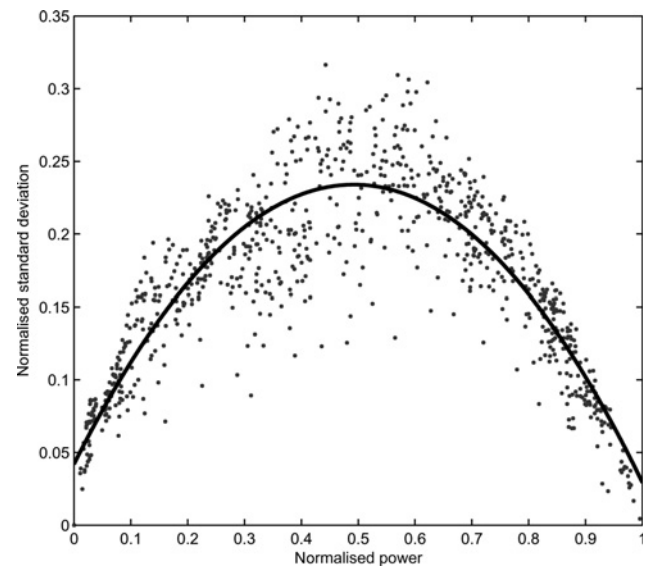**Figure 3** *Sample PDF of the uncertainty of wind power prediction*

Figure 4 *Relation between standard deviation and mean for the uncertainty of predictions*

4 Load flow equations

The load flow equations for a power system may be written as (13)

$$S = g(Z)$$
$$P_f = h(Z) \tag{13}$$

where Z is the vector of nodal voltages and angles, S, the input vector of real and reactive power injections, and P_f, the output vector of line active power flows; g and h are non-linear functions. Linearising these equations around a working point yields, after some calculation

$$\Delta P_f = J_h J_P^{-1} \Delta P = \Lambda_f \Delta P \tag{14}$$

ΔP is the vector of incremental active power injections taken from vector S. J_h is the jacobian matrix of non-linear function h, whereas J_P is the submatrix of the jacobian matrix of function g that relates line active power flows to variables Z. Reactive power injections have been considered as linearly related to active power injections, since most modern wind generation may control the power factor of wind farms according to economic incentives and the power factor of loads will be considered constant. Hence the changes in the reactive power injected or demanded are proportional to the active power.

Matrix Λ_f is a sensitivity matrix whose terms are the system power transfer distribution factors (PTDF). The definition of these PTDFs assumes that the power injections are compensated by opposite power injections at the slack bus. This could be an acceptable assumption when the injections have a small value. However, large fluctuations because of changes of power in wind farms are compensated by the combined operation of several generators. Hence, the

sensitivity coefficients used in this work have been calculated considering the distribution of input power among different generators. Evidently, conventional PTDF are a particular case of these sensitivities. The expression of these new sensitivity coefficients is

$$\lambda'_{qi} = \lambda_{qi} - \sum_{r=1}^{R} k_{ir} \lambda_{qr} \qquad (15)$$

where λ_{qi} is the term (q, i) of the sensitivity matrix Λ_f, that is to say, the PTDF of line q with respect to an injection in node i, k_{ir} is the part of power injection in node i that the regulating generator r assumes, for example $(k_{ir} = 1/R)$. R is the number of generators that compensate the injection in node i. Of course, any other sharing among generators is possible. Hence, it may be written that

$$\Delta \boldsymbol{P}_f = \boldsymbol{\Lambda}'_f \Delta \boldsymbol{P} \qquad (16)$$

where $\Delta \boldsymbol{P}$ includes only the considered power injections, that is, the random power injections in our case.

A similar approach may be made to estimate the reactive power and the voltage sensitivities. The accuracy of these approximations is smaller, because of the higher non-linearity of the involved equations. This approximation can be written, for reactive power, as

$$\Delta \boldsymbol{Q}_f = \boldsymbol{\Gamma}'_f \Delta \boldsymbol{P} \qquad (17)$$

5 Computational procedure

The method proposed in this paper calculates the CDF of the line power flows from the CDF of the power injections through a linear approximation of the load flow equations, taking into account the dependence among power injections. The input random variables are wind power, demand and availability of conventional generation. The proposed method will be called ELM.

The time range considered in the paper is the daily operation, although it could also be applied for other time ranges. The process goes as shown in the next paragraphs. A flowchart of the process is given in Fig. 5.

1. *Base case:* The input data are the grid data, the forecasted load and wind power and the scheduled power of the power plants. The sensitivities of the line power flows to power injections are calculated as shown in Section 4.

2. *Uncertainty of power injections:* Wind power uncertainties are modelled as dependent beta distributions whose means are the forecasted power, and whose standard deviations are estimated as shown in Section 3. Load uncertainties are modelled as dependent or independent normal variables with a given standard deviation and correlation matrix. Power plant uncertainties are modelled as independent

binomial variables. Random numbers with the mentioned distributions are generated as shown in Section 2.4. As it can be seen, the method does not depend on the accuracy of the correlation between the input data.

3. *Moments and crossed moments of input variables:* Central moments of inputs variables are calculated numerically from the random numbers previously generated, using a discretisation of (2). From the central moments, cumulants are found as described in [20].

4. *Estimation of mean values of output variables:* The mean values of the output variables are estimated using a $(2m + 1)$ point estimate (PE) method, without considering the correlation between random variables. In spite of this simplification, the accuracy of the approximation for the mean value is very good.

5. *Higher order cumulants of output variables:* Cumulants of order 2 and higher are calculated using (8), irrespectively of the continuous or discrete character of the random variables. The linear coefficients that relate the output variables (power flows through branches) to the input variables (power injections) are given by (16). The cumulants up to the fifth order are obtained, since they are needed for the Cornish-Fisher expansion series, as shown in (9).

6. *CDF of output variables:* Once the cumulants of distributions of the output variables are known, it is necessary to estimate their CDF. This process follows three steps:

(a) First, the CDF of the output variables, without considering the discrete variables (power plant availabilities), is estimated using Cornish-Fisher expansion series.

(b) Then, the unimodality condition of (12) is checked for each output variable of interest and for each power plant.

(c) If the resulting CDF is not multimodal, then a Cornish-Fisher expansion series is used to approximate the CDF. If the CDF is multimodal, then the necessary convolutions are made in order to obtain the estimate of the CDF. This process is made at the end, in order to minimise the number of necessary convolutions.

6 Study case

In order to show the possibilities of ELM and to quantify its accuracy, a simulation study has been made for a test case.

Data: The IEEE-118 test system [31] has been modified to include wind generation. Wind generation replaces part of the conventional generation, which has been reduced proportionally. The wind power installed is 693 MW, and the total load is 3670 MW. The injected wind power in this situation, 416 MW, is the 60% of the installed wind power. Wind farms are clustered into three groups. Wind farms of the same group are highly correlated, but the groups are
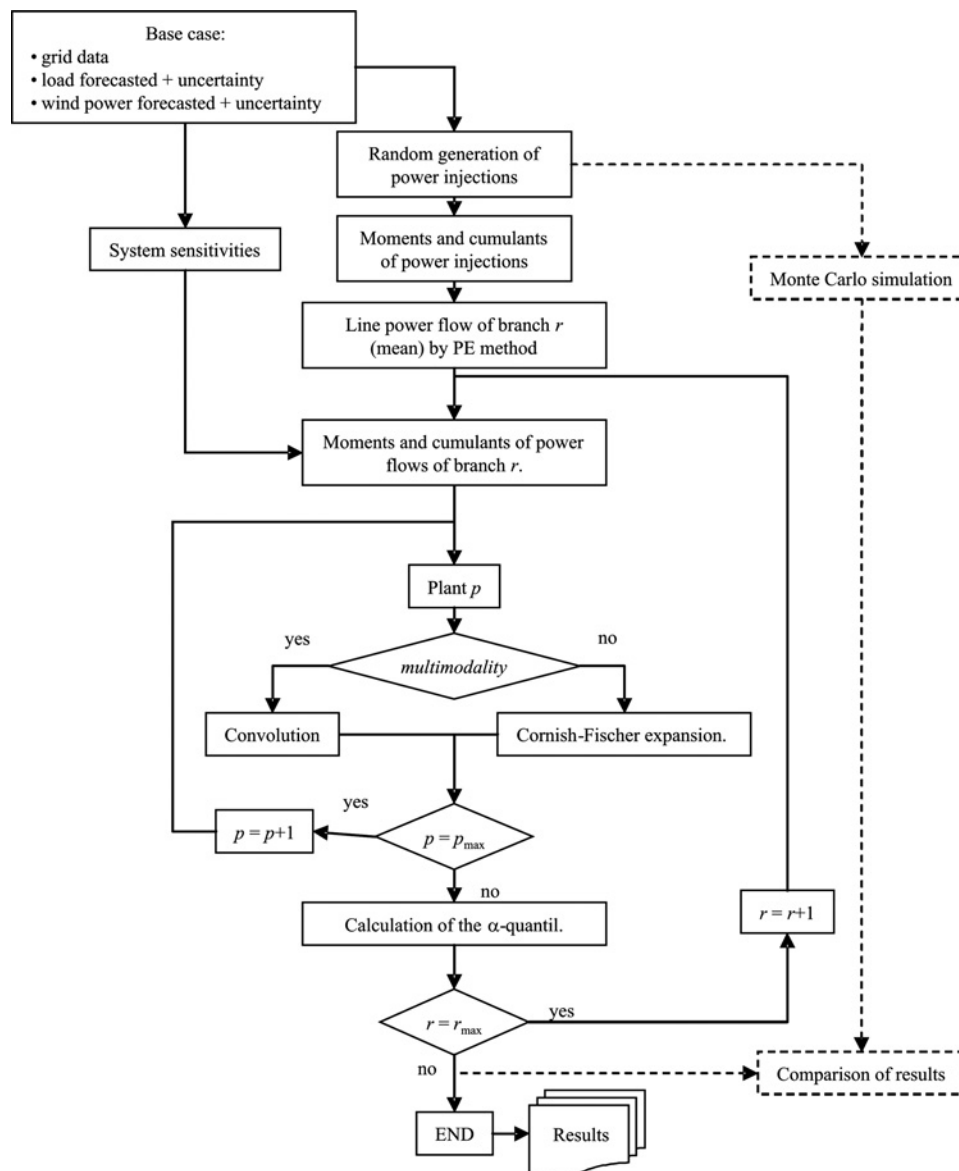
Figure 5 *Flowchart of ELM*

considered to be independent between them. The values of the correlation coefficients were taken from actual correlation coefficients between nearby wind farm productions in peninsular Spain. Although this is not a wholly rigorous approach, it could be enough as a first approximation, since there are not yet studies concerning spatial dependence between short-term wind power predictions, to the knowledge of the author. The data of this system are given in the Appendix (Tables 5 and 6).

The availability of power plants has been considered as a binomial variable. The plants that compensate load and generation changes are those located in nodes 10, 25, 46, 54, 61, 66 and 100. The changes in generation or load power are shared equally among them.

In order to assess the performance of ELM and the importance of considering the dependence between

random variables, its results are compared to the PE method [9, 10]

Computation times: The method has been programmed in MATLAB. Computation times depend on the number of output variables of interest and the number of convolutions performed, but some values are given in Table 1. They have been obtained with a processor Intel Pentium of 2.13 GHz with 1 Gb of RAM. The number of samples in the Monte Carlo simulation is 10 000.

Results: The accuracy of the approximation of moments is shown in Table 2 for the ELM and $(2m + 1)$ PE method, compared with Monte Carlo simulation results. Two cases have been considered, with and without dependence between input variables, in order to show the importance of considering this dependence. This table shows the average error in % between moments of the line power flows,

Table 1 Computation times, in s

Branches	Conv.		ELM	Monte Carlo
28	23		14.52	332.26
172	43	AC	43.98	311.7
		DC	27.23	271.14

obtained analytically, by ELM and the PE method, and those obtained through Monte Carlo simulation. The error is defined by

$$\varepsilon_n = \frac{1}{N_B} \sum_{j=1}^{N_B} \frac{|\mu_{n,j}^{an} - \mu_{n,j}^{MC}|}{|\mu_{n,j}^{MC}|} \cdot 100 \qquad (18)$$

where $\mu_{n,j}^{an}$ is the moment of order n of branch j found analytically (either with ELM or PE), whereas $\mu_{n,j}^{MC}$ is the same moment obtained by the Monte Carlo method. N_B is the number of branches in the grid.

Comparison with Monte Carlo is made in order to check the accuracy of the proposed method, and no attempt has been made to minimise the number of Monte Carlo simulations. In order to assess the accuracy of the results of the Monte Carlo simulations in the given system, the following data are given. The 95% confidence interval for the mean estimate by Monte Carlo have been found for all the lines. Since its size is very variable, because also the power flow is very different among lines, average quantities have been considered. The average power flowing in the lines (absolute value) is 0.4762 p.u., whereas the average size of the 95% confidence intervals is 0.0037 p.u., which is a 5.6% of the average power. The number of Monte Carlo simulations necessary to reduce this value to a 5% of the average absolute power is about 12 600. The average 95% confidence interval of the variance is a 5.55% of the average variance of all the lines. For this last test, normality of the power flow distribution has been assumed, so the result is only approximate.

Two formulations, AC and DC, have been considered for ELM. The DC formulation uses the DC load flow equations for the analytical approach and the Monte Carlo simulation, whereas the AC case uses the sensitivities obtained in Section

4, and the results are compared to those of Monte Carlo simulations with AC load flow.

The PE method yields a good approximation for the independent case up to the second-order moment. Estimates of higher order moments have also higher errors. The method, however, behaves poorly when the input random variables are dependent. ELM, on the other hand gives slightly worse results for the independent case. For the dependent case, the method yields satisfactory results.

The error in the 90% percentile of active power, that is calculated as in (19), has also been found

$$\varepsilon_{90} = \frac{1}{N_B} \sum_{j=1}^{N_B} \frac{|p_{90,j}^{ELM} - p_{90,j}^{MC}|}{|p_{90,j}^{MC}|} \times 100 \qquad (19)$$

where $p_{90,j}^{ELM}$ and $p_{90,j}^{MC}$ are the 90% quantile of branch j for ELM and the Monte Carlo simulation, respectively. The values of ε_{90} are given in Table 3.

The results for reactive power flows in branches are given in Table 4, where the error is defined as in (18). The results of the PE method in the independent case are better than ELM for the second moment, but in the dependent case ELM behaves better than the PE method. In all the cases the error ε_n is high, as could be expected. This is due not only to the higher non-linearity of reactive power with respect to input power, but also to the smaller variability of reactive power flows and the definition of ε_n. For instance, for the case of linearised AC equation with dependent variables, the average variance of branch active powers is 0.019 (p.u.)^2, whereas that of reactive powers is 5.73×10^{-4} (p.u.)^2. The maximum values of these parameters are 0.4749 (p.u.)^2 and 0.0132 (p.u.)^2, respectively. For higher order moments, the differences are even larger. This implies that the importance of considering reactive power variations is smaller.

The PDF resulting from Monte Carlo simulation, the Cornish-Fisher expansion and the normal approximation is also compared for the power flows in two lines. Fig. 6 shows the PDF of the power flow through line 5−11. It

Table 2 Value of ε_n: active power in lines

Moment	Independent			Dependent		
	ELM		PE (2m + 1)	ELM		PE (2m + 1)
	DC	AC		DC	AC	
1	0.31	0.36	0.13	0.35	0.36	0.40
2	0.88	2.00	0.78	0.91	1.82	29.93
3	8.55	13.39	10.07	9.97	12.94	60.56
4	2.24	5.34	59.36	2.72	5.39	77.86
5	11.47	25.97	–	13.65	17.26	–

Table 3 Errors for the 90% quantile: active power in lines

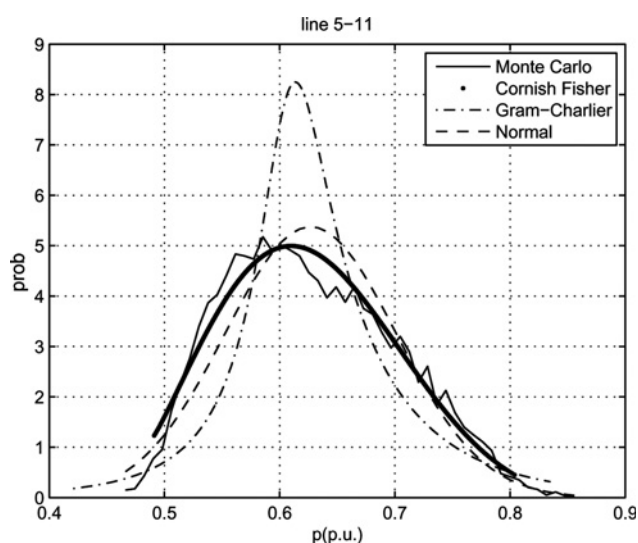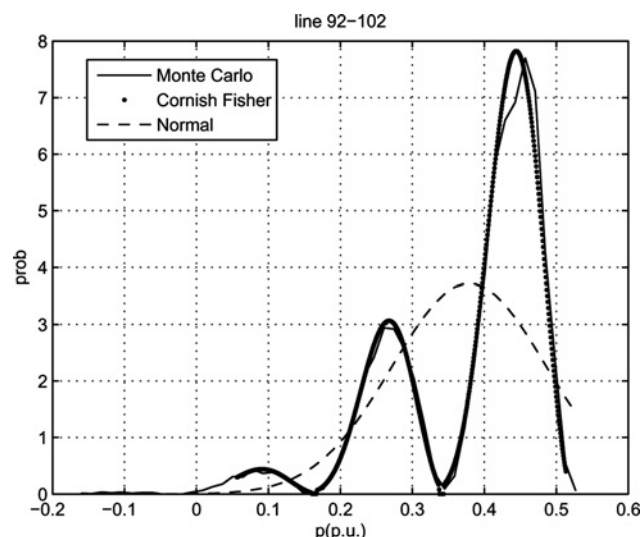|	Independent				Dependent			
	AC		DC		AC		DC	
	C-F	N	C-F	N	C-F	N	C-F	N
ε_{90}	1.93	2.73	1.95	3.06	1.83	2.03	1.81	2.13
max. (p.u.)	0.160	0.089	0.142	0.094	0.077	0.061	0.064	0.066

C-F: Cornish-Fisher expansion; N: Normal distribution

Table 4 Value of ε_n: reactive power in lines, grid IEEE-118

Moment	Independent		Dependent	
	ELM	PE $(2m+1)$	ELM	PE $(2m+1)$
1	1.01	0.95	2.04	1.80
2	25.41	9.40	22.88	30.36
3	102.24	152.46	97.42	155.48
4	58.45	82.26	45.74	86.84
5	142.33	–	132.12	–

can be seen that the Cornish-Fisher result fits better to the Monte Carlo PDF. The PDF obtained through Gram-Charler A series is also included. It can be seen that this series expansion leads to a bad approximation.

When the influence of a discrete variable (the uncertainty of a power plant production) on the line power flows is large, multimodality appears, as shown in Fig. 7. The normality assumption gives a bad approximation for the PDF, whereas the convolution result fits well. It must be remarked that convolutions are only performed whenever they are needed: in the previous case (Fig. 6), no convolution was made.

Figure 7 PDFs of active power in branch 92−102

Although the method has been developed also for voltages, the changes in the voltages, when power variations are compensated by different nodes, are extremely small, and for this reason the results are not presented here.

7 Conclusion

Probabilistic power flow becomes more important in systems with high wind power penetration, because of the high variability of the injected power. For daily system operation, it is necessary to consider the uncertainties of short-term wind power predictions. These uncertainties are correlated for nearby wind farms. In general, they cannot be assumed to be Gaussian, and a beta distribution is a better choice.

The use of cumulants to estimate the moments of line power flows from the uncertainty of random power injections seems to be a good choice, because of the accuracy of the results and the easiness to include dependence between input random variables.

Estimation of CDF from cumulants using Cornish-Fisher expansion series seem to behave reasonably well in unimodal non-Gaussian functions (such as those of the wind power prediction uncertainty). To approximate the output PDF by a normal distribution also gives an approximation of the higher percentiles, even if the fitting of the resulting PDF is poor.

Figure 6 PDFs of active power in branch 5−11

When power plant availability or other discrete input random variables are to be considered, some of the PDF of the output variables may be multimodal, and it is necessary to perform convolutions between the PDF of random variables. This technique is computationally expensive, and it is better to perform it only when necessary. To discriminate between necessary and unnecessary convolutions, a unimodality test can be made in the frequency domain. If convolutions are made only when this test is not passed, the number of necessary convolutions reduces dramatically. Obviously, the normality assumption does not hold for multimodality conditions.

8 References

[1] LEITE DA SILVA A.M., RIBEIRO S.M.P., ARIENTI V.L., ALLAN R.N., DE COUTO FILHO M.B.: 'Probabilistic load flow techinques applied to power system expansion planning', *IEEE Trans. Power Syst.*, 1990, **PWRS-5**, pp. 1047−1053

[2] BORKOWSKA B.: 'Probabilistic load flow', *IEEE Trans. Power Appl. Syst.*, 1974, **PAS-93**, pp. 752−759

[3] DOPAZO J.F.: 'Stochastic load flow', *IEEE Trans. Power Appl. Syst.*, 1975, **PAS-94**, pp. 299−309

[4] GENTLE J.E.: 'Random number generation and Monte Carlo methods' (Springer, New York, 2003, 2nd edn.)

[5] ALLAN R.N., DA SILVA A.M.L., BURCHETT R.C.: 'Evaluation methods and accuracy in probabilistic load flow solutions', *IEEE Trans. Power Appl. Syst.*, 1981, **PAS-100**, pp. 2539−2546

[6] HATZIARGYRIOU N.D., KARAKATSANIS T.S., PAPADOPOULOS M.: 'Probabilistic load flow in distribution systems containing dispersed wind power generation', *IEEE Trans. Power. Syst.*, 1993, **PWRS-8**, pp. 159−165

[7] LEITE DA SILVA A.M., ARIANTI V.L.: 'Probabilistic load flow by a multilinear simulation algorithm', *IEE Proc.*, 1990, **137**, (Pt. C), pp. 276−282

[8] MIRANDA V., SARAIVA J.T.: 'Fuzzy modelling of power system optimal power flow'. Proc. Power Industry Computer Application Conference, Baltimore, USA, 1991

[9] SU C.-L.: 'Probabilistic load-flow computation using point estimate method', *IEEE Trans. Power Syst.*, 2005, **20**, (4), pp. 1843−1851

[10] MORALES J.M., PEREZ-RUIZ J.: 'Point estimate schemes to solve the probabilistic power flow', *IEEE Trans., Power Syst.*, 2007, **22**, (4), pp. 1594−1601

[11] VERBIC G., CAÑIZARES C.A.: 'Probabilistic optimal power flow in electricity markets based on a two point estimate method', *IEEE Trans. Power Syst.*, 2006, **21**, pp. 1883−1893

[12] LEITE DA SILVA A.M., ARIENTI V.L., ALLAN R.N.: 'Probabilistic load flow considering dependence between input nodal powers', *IEEE Trans. Power. Appl. Syst.*, 1984, **PAS-103**, pp. 1524−1530

[13] SANABRIA L.A., DILLON T.S.: 'Power system reliability assessment suitable for a deregulated system via the method of cumulants', *Electric Power Energy Syst.*, 1993, **20**, pp. 203−211

[14] TIAN W.D., SUTANTO D., LEE Y.B., OUTHRED H.R.: 'Cumulant based probabilistic power system simulation using Laguerre polynomials', *IEEE Trans. Energy Convers.*, 1989, **4**, pp. 567−574

[15] ZHANG P., LEE T.: 'Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion', *IEEE Trans. Power Syst.*, 2004, **PWRS-19**, pp. 676−682

[16] HU Z., WANG X.: 'A probabilistic load flow method considering branch outages', *IEEE Trans. Power Syst.*, 2006, **PWRS-21**, pp. 507−514

[17] SCHELLENBERG A., ROSEHART W., AGUADO J.: 'A cumulant-based probabilistic optimal power flow (P-OPF) with gaussian and gamma distributions', *IEEE Trans. Power. Syst.*, 2005, **PWRS-20**, pp. 773−781

[18] USAOLA J.: 'Probabilistic load flow with wind production uncertainty using cumulants and Cornish-Fisher expansion'. Proceedings of the 2008 Power System Computation Conference, Glasgow, UK, July 2008

[19] PAPOULIS A., PILLAI S.U.: 'Probability, random variables and stochastic processes' (McGraw-Hill, Boston, 2002, 4th edn.)

[20] MCCULLAGH P.: 'Tensor methods in statistics' (Chapman & Hall, London, 1987)

[21] JASCHKE S.R.: 'The Cornish-Fisher-expansion in the context of delta-gamma-normal approximations'. http://www.jaschke-net.de/papers/CoFi.pdf. Discussion Paper 54, Sonderforschungsbereich 373, Humboldt-Universitaet zu, Berlin

[22] STUART A., ORD K.: 'Kendall's advanced theory of statistics' (Arnold, 1994, vol. I, 6th edn.)

[23] CRAMER H.: 'Numerical methods of statistics' (Princeton University Press, Princeton, NJ, 1946)

[24] BLINNIKOV S., MOESSNER R.: 'Expansions for nearly Gaussian distributions', *Astron. Astrophys. Suppl. Ser.*, 1998, **130**, pp. 193−205

[25] CORNISH E.A., FISHER R.A.: 'Moments and cumulants in the specification of distributions', *Revue de l'Institut Inter. Statis.*, 1937, **4**, pp. 307−320

[26] HILL G.W., DAVIS A.W.: 'Generalized asymptotic expansions of Cornish-Fisher type', *Ann. Math. Stat.*, 1968, **39**, (4), pp. 1264–1273

[27] PAPAEFTHYMIOU G.: 'Integration of stochastic generation in power systems'. PhD dissertation, University of Delft, 2007

[28] KARINIOTAKIS G., *ET AL*.: 'Evaluation of advanced wind power forecasting models – results of the Anemos project'. Proceedings of the EWEC, Athens, 2006

[29] FABBRI A., GOMEZ SAN ROMAN T., RIVIER J., MENDEZ QUEZADA V.H.: 'Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market', *IEEE Trans. Power Syst.*, 2005, **20**, (3), pp. 1440–1446

[30] FOCKEN U., LANGE M., MOENNICH K., WALDL H.P., BEYER H.G., LUIG A.: 'Short-term prediction of the aggregated power output of wind farms – a statistical analysis of the reduction of the prediction error by spatial smoothing effects', *J. Wind Eng. Ind. Aerodyn.*, 2002, **90**, pp. 231–246

[31] Power Systems Test Case Archive. [Online]. Available at: http://www.ee.washington.edu/research/pstca

9 Appendix: Data of the study cases

Table 5 Wind farms included in the IEEE-118 system

WF	Node	Group	P (MW)	σ (MW)	Rated power (MW)
1	52	1	59.3	25.6	98
2	44	1	31	13.2	51
3	53	1	14.8	6.48	25
4	50	1	8.5	3.66	14
5	84	2	20.1	9.4	36
6	86	2	17	7.2	28
7	83	2	33	15.11	58
8	82	2	50.3	20.84	82
9	2	3	33	14.1	55
10	5	3	20	9.4	36
11	16	3	27	11.2	44
12	13	3	37.5	16.1	62
13	3	3	27	10.5	42
14	14	3	37.5	16.1	62
		total	416		693

Table 6 Correlation coefficients: IEEE-118 case

Group 1					
Wind farm	1	2	3	4	
1	1	0.88	0.87	0.91	
2	0.88	1	0.85	0.87	
3	0.87	0.85	1	0.85	
4	0.91	0.87	0.85	1	
Group 2					
Wind farm	5	6	7	8	
5	1	0.82	0.85	0.9	
6	0.82	1	0.85	0.88	
7	0.85	0.85	1	0.89	
8	0.9	0.88	0.89	1	
Group 3					
Wind farm	9	10	11	12	13	14
9	1	0.85	0.86	0.83	0.82	0.91
10	0.85	1	0.88	0.83	0.89	0.92
11	0.86	0.88	1	0.85	0.95	0.87
12	0.83	0.83	0.85	1	0.89	0.91
13	0.82	0.89	0.95	0.89	1	0.82
14	0.91	0.92	0.87	0.91	0.82	1