Electric Power Systems Research 119 (2015) 266–277

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Probabilistic load flow with versatile non-Gaussian power injections

Cristina Carmona-Delgado [1], Esther Romero-Ramos *, Jesús Riquelme-Santos [2]

Department of Electrical Engineering, Camino los Descubrimientos S/N, 41092, University of Sevilla, Spain

ARTICLE INFO

Article history:
Received 8 April 2014
Received in revised form 11 August 2014
Accepted 4 October 2014
Available online 29 October 2014

Keywords:
Probabilistic load flow
Gaussian Mixture Model
Probabilistic density function
Monte Carlo simulation

ABSTRACT

A probabilistic load flow distinguished by the versatility in the way in which input data can be provided is presented. The main contribution of the proposal involves taking advantage of the data available as well as completing any missing information. This enables the proposal to be applied at any voltage level, even in medium-voltage networks where there is a glaring lack of systematic data collection. The use of the Gaussian Mixture Model is also a key feature of the proposed solution, and determinant in the final solution. Not only does the detailed and thorough analysis through numerous tests demonstrate the good performance of the proposed procedure, but it also confirms the accuracy of the results.

1. Introduction

Probabilistic load flow (PLF) is becoming a relevant and essential planning tool since it enables the assessment of all the possible working conditions of a power network. PLF takes into account uncertainties in power systems, such as the random nature of demand and supply, which includes the inherent uncertainty of the increasing distributed generation (DG). By means of the appropriate modelling of random input variables, PLF provides valuable results: for example, it can yield the likelihood that a bus voltage or a power flow falls outside its respective permissible limit. This method therefore overshadows the traditional deterministic load flow (DLF), which fails to consider the uncertainty associated to the problem. Ref. [1] constitutes a good review of the state of the art of PLF techniques.

The first proposals on the PLF problem appeared in the mid-70s [2,3]. Since then numerous methodologies have been proposed, the majority of which focus on the most efficient and accurate way either to model the uncertainty of the loads or to solve the problem. These methodologies can be roughly divided into two main groups according to the way that the PLF is solved. On the one hand, numerical methods are the most straightforward. Within this group, Monte Carlo simulation (MCS) is the most widely used technique to handle the PLF problem. It involves running repeated simulations of an exact deterministic model from the probability density function (PDF) of the random variables considered [4,5], which constitutes the main concern regarding the MCS method due to the need for a large number of simulations. On the other hand, there are analytical methods, which are computationally more efficient compared to MCS. Among these, two of the most popular techniques developed for the reduction of the computational burden are: those that use the discrete frequency domain convolution by applying the Fast Fourier Transformation (FFT) [6] and those that employ a combination of the cumulants and Gram-Charlier expansion series [7] and Cornish–Fisher expansion [8]. The main concerns about analytical methods, apart from the complicated mathematical computation, include the requirement of mathematical assumptions, such as linearization of equations and independence between input variables, which leads to inaccurate results. Apart from numerical and analytical methods, according to certain authors, two additional recent groups of advantageous methodologies could be considered: those known as heuristic procedures, such as fuzzy logic [9]; and those known as approximate techniques, from among which the point estimate method deserves special mention [10]. The point estimate method obtains the mean and standard deviation of any power system variable in a very efficient way, while the MCS method requires the knowledge of the probability distributions. However, the point estimate method needs a highly complex formulation in order to consider the correlation between random variables [11]. Other authors have even ventured to combine the properties of several methods in order to overcome the drawbacks of non-numerical methods [12].

From this brief bibliographic review, it becomes clear that, for the last 40 years, PLF has been widely studied, and that its application to network planning in order to adapt to the uncertainties

* Corresponding author. Tel.: +34 95 448 12 72.
E-mail addresses: cristinacarmona@us.es, cristinacarmonadelgado@gmail.com (C. Carmona-Delgado), eromero@us.es (E. Romero-Ramos), jsantos@us.es (J. Riquelme-Santos).
[1] Tel.: +34 95 448 21 74.
[2] Tel.: +34 95 448 12 74.

http://dx.doi.org/10.1016/j.epsr.2014.10.006
0378-7796/

introduced by the increasing DG is extensively accepted [13]. Even so, the planning of the electric power system is still carried out according to a traditional DLF based on peak-load estimations and their corresponding supply by programmed conventional generation. These estimations are based on the current regulatory framework, which fails to take into account all the effects and consequences of the high and increasing proportion of DG ([14] in the case of Spain). Regarding DG, neither can the amount of the supply, nor can the geographical location of said supply be governed in the same way as for conventional generation On the one hand, the uncertainty associated to DG supply supports the application of a probabilistic tool instead of a deterministic one. On the other hand, the scattered and distant locations of DG from consumption centres tend to increase the power flows through certain critical lines of the system, whose capacity could be put at risk. These risky scenarios warrant verification with a probabilistic load flow tool.

In this work, a PLF procedure is proposed that is able to take into consideration the uncertainty inherent to the modern electric power systems based on MCS. The main contribution of the paper is the versatility of the modelling of the PDFs, regardless of the availability of measurements for loads (none, some, or all measurements available) and of the shape of the PDFs (standard or otherwise). On the one hand, through the use of MCS, simplifying assumptions no longer remain necessary and correlation among bus loads and uncertainty of the DG are easily taken into account. On the other hand, the versatility of the modelling of the PDFs comes from approximating the non-Gaussian distributions by a Gaussian Mixture Model (GMM), thereby enabling any type of PDF given by measurements to be modelled [15], whereas those loads without measurements are modelled as standard distributions defined by parameters. This versatility enables the proposal to be applied even in medium-voltage systems (which often do not have measurements available in most loads).

The proposed procedure features a new encompassing and versatile proposal incorporating a number of isolated contributions that are properly merged and put together in order to create a necessary and practical tool adapted to the latest changing electric system conditions.

The paper is arranged as follows: Section 2 introduces the GMM and its treatment from a set of measurements. Section 3 sets out the formulation of the proposed PLF, which includes DG. Section 4 is dedicated to testing the proposal in the cases of absence of measurements and of availability of measurements. The main conclusions are summarized in Section 5.

2. Gaussian Mixture Model

The vast majority of load PDFs cannot be represented by any standard probabilistic distribution, especially those in medium- and low-voltage networks, where levels of aggregation decrease. The GMM approach can represent any type of load distribution as a combination of several Gaussian components [16], and represents a parameter estimation problem.

As stated in Section 1, the proposal presented in this paper uses the GMM approach to model PDFs associated to those loads which do have measurements or *monitored loads*. The conversion of a set of measurements into the PDF itself is carried out by means of a density histogram. The density histogram is generated from the segments into which the range of the data can be divided (commonly known as *bins*). The probability density of each bin is assessed from the relative frequency of load data falling into each bin, thus computing the discrete PDF. Fig. 1 illustrates the example of a discrete PDF associated to a real load. Note that the shape of the PDF does not follow any known standard distribution function as happens in most loads in real networks. This leads to the concept of GMM, which is discussed below.

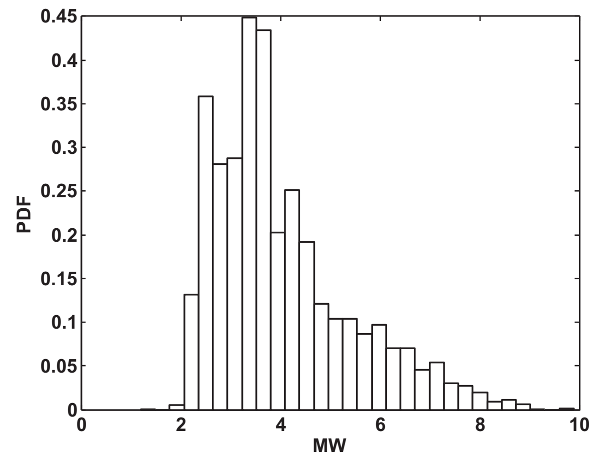

Fig. 1. Discrete probability density function.

The Gaussian PDF of a random variable is widely-known and defined as (1),

$$f_{\mathrm{Gauss}(\mu,\sigma)}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-((x-\mu)^2/2\sigma^2)} \tag{1}$$

where μ is the mean value, and σ the standard deviation. A GMM is the combination of n Gaussian distribution components, whose PDF, for one random variable X, is defined as (2),

$$f_X(x) = \sum_{k=1}^{n} \omega_k \cdot f_{\mathrm{Gauss}(\mu_k,\sigma_k)}(x) \tag{2}$$

where ω_k, μ_k and σ_k are the proportion, mean, and standard deviation of the kth component of the Gaussian combination, respectively [17]. The proportion parameters, in turn, must be subject to the following constraint to ensure the specific conditions of a PDF (3),

$$\omega_k \in (0, 1] \quad \text{and} \quad \sum_{k=1}^{n} \omega_k = 1 \tag{3}$$

Any PDF can be approximated by the GMM, by taking into account that the higher the number of Gaussian components, the better the approximation; however, the number of parameters to estimate are also higher. Therefore, there are two different estimations to be carried out: the number of Gaussian components, and the parameters of those Gaussian components. The way to proceed involves:

a. Assume a minimum number of four Gaussian components (e.g. $n = 4$).
b. Solve the parameter estimation problem, for these n components, by using the expectation maximization (EM) algorithm [18], widely considered one of the most effective methodologies in this regard. The EM algorithm finds the maximum-likelihood estimate of the parameters of a distribution from a given data set.
c. Obtain Akaike's Information Criterion (AIC) for this n-component GMM [19]. AIC_n is introduced below.
d. Save the information and go to point b, and increase the number of Gaussian components by one unit, i.e. $n = n + 1$.
e. Repeat this process until $|\mathrm{AIC}_{n+1} - \mathrm{AIC}_n| < \xi$.

The number of Gaussian components is approximated from the minimization of Akaike's Information Criterion. The AIC provides