Electrical Power and Energy Systems 62 (2014) 273–283

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Statistical characterization of aggregated wind power from small clusters of generators ☆

Guzmán Díaz *, Javier Gómez-Aleixandre, José Coto

Dept. of Electrical Engineering, University of Oviedo, Campus de Viesques, s/n, Gijón 33204, Asturias, Spain

ARTICLE INFO

Article history:
Received 21 June 2013
Received in revised form 24 March 2014
Accepted 21 April 2014
Available online 23 May 2014

Keywords:
Wind power
Correlation
Aggregation
Microgrids
Duration curves
Forced outage rate

ABSTRACT

In this paper we address the problem of aggregating wind power. The purpose of the methodology presented here is to avoid the assumption of extreme values of correlation, meaning perfect dependence or perfect independence of the production. That is, we accept intermediate values of correlation, which we argue is of special interest for small-scale siting analysis, where the fluctuations of wind power production affect the capacity value or the size of energy storage.

We provide a formulation that is based on the integration of the joint probability density function (PDF) of the wind power. We formulate this PDF by means of copula theory in order to cope with the involved representation of the marginal PDFs.

As a result, we characterize the PDF of the aggregated wind power and the associated duration curve. We also present a simple formulation of the joint forced outage rate. These serves us for verifying, through a case analysis based on NREL datasets, that in some cases the assumption of extreme dependence in small-scale sites does not hold.

Introduction

The joint production of wind energy by a cluster of generators is characterized by uncertainty, which stems both from the individual wind speed uncertainty and from the correlation between wind sites. This latter has a fundamental impact on the joint production, because if the generators are not producing in unison, the joint production will be less fluctuating; less affected by individual fluctuations. In small-scale wind power integration—the subject of this paper—the role of the correlation is exacerbated. This is the case of dispersed generation, microgrids, or in general a cluster of wind generators that is considered to be an only producer. In such cases, the correlation may be non-negligible, and we may even find a perfect coherence between pairs of generators. Obviously, this largely affects the capacity credit of the cluster or, in the case of islanded microgrids, the size of the required storage.

Certainly the most direct approach to aggregating generation or loads when they are randomly fluctuating is to take one of the extreme values of dependence. If on the one side the power is assumed to be perfectly coherent in all the generators, namely fluctuating in unison, the aggregated power is obtained as the power in one generator times the number of generators (or if the powers are different in value, as the sum of powers). On the other extreme, if it is assumed a perfect independence between individual powers (or loads), the joint probability of a given value of aggregated power is equal to the product of the probabilities of every individual (marginal) power. In this case if the statistical characterization is done by means of probability density functions (PDFs), then the joint PDF is again the product of the marginal PDFs. And as long as the aggregated power can be generally obtained as the convolution of the joint PDF over the power domains, the decomposition into a product of marginal PDFs facilitates the task of solving the problem.

It can be argued that in some cases this simplified approach may well serve the purpose of the investigation. For instance in [1], Vallée et al. exploited both extreme values of dependence to investigate the reliability of a power system. They claimed that the two extreme values were representative of the worst and best scenarios. In particular, perfect dependence in their four-area power system represented the worst scenario, because the wind power fluctuations were amplified. Conversely, perfect independence produced a "softening" of the wind power fluctuation, which yielded the minimum reserve requirements. For optimization purposes of hybrid generation systems, Tina et al. also exploited the extreme dependence approach [2,3]. In this case they only

☆ This work was supported by the Spanish Ministry of Science under Grant ENE2010-14899.

* Corresponding author. Tel.: +34 985182592.
E-mail addresses: guzman@uniovi.es (G. Díaz), jgomez@uniovi.es (J. Gómez-Aleixandre), jcoto@uniovi.es (J. Coto).

http://dx.doi.org/10.1016/j.ijepes.2014.04.033
0142-0615/

considered perfect independence, arguing that solar and wind energy could be assumed to be perfectly independent, and that otherwise a time scale sufficiently small might be chosen. The other extreme, perfect dependence, can be noted in for instance [4], where the productions of all generators in a wind farm were assumed to be coherent.

An approach based on extreme dependence—either perfect dependence or perfect independence—may not always be adequate, however. An indication of the need for intermediate dependence analysis may be induced for instance from Østergaard's work [5], in which the author reflects on the issue that the larger the geographical extension of the area of Denmark in which wind power is exploited, the lower are the required reserves—an indication of the *progressive* "softening" of fluctuations when the generators are geographically spread. That is, the work of Vallée et al. provides the extreme values of reserves, but Østergaard narrows the reserve requirements and indicates that they depend on the spread of generators. This is particularly more relevant if we review Holttinen's work in [6], where the correlation between wind speeds as a function of distance in Sweden is represented as a decaying exponential $r = e^{-\frac{d}{500}}$; where d is the distance in kilometers. This means that the correlation varies from $r \approx 1$ for nearby sites to as low as 0.6 by a few hundred kilometers. It also means that to achieve perfect independence, we must compare sites in the thousands of kilometers apart. Similar results can be found in [7]. (Ultimately, these results agree with our own results from a survey conducted by using the wind speed year samples of 210 sites from NREL dataset, hosted in http://wind.nrel.gov/, comprising 21,945 pairs of generators with the pair proximity ranging from 1.3 through 390 km. Again, the value of r was between approximately 0.5 and 1.)

The objective of our paper is to introduce a methodology to compute the aggregated power, which does not rely on the assumption of extreme dependence. We will argue in the last part of this paper, when we investigate the power duration curves and the forced outage rates, that the approaches based on extreme dependence give remarkably poor results in wind power scenarios where the values of correlation are intermediate. This is specially relevant, following the above discussion, in the analysis of small areas of wind production where the dependence is expected to be from medium to high; sometimes almost perfect.

The methodology that we propose can be boiled down to three major operations, (i) the statistical characterization of the *univariate* marginals, (ii) the statistical characterization of the *multivariate* dependence structure, and (iii) the wind power aggregation proper. By characterization we denote the parametric representation of the cumulative distribution (CDF) and probability density (PDF) functions. And eventually through aggregation, we will obtain the PDF of the aggregated wind power and the duration curve—a particular interpretation of the wind power CDF.

Importantly, in the statistical characterization we shall describe not only the wind speed random variables, but also wind power. The difference between both variables will be evident and relevant, because the wind power distribution function is non-smooth, unlike the wind speed distribution. Notably, we will exploit this feature found in the wind power distribution to improve the statistical characterization of the aggregated wind power by incorporating an interpretation of the forced outage rate (FOR) based on copula calculations. This FOR index has been employed in reliability analysis [8–10], and it has been computed by means of bin counts in [1] (without considering intermediate dependence) or by intensive Monte Carlo simulations in [11]. Our proposal based on copula theory—a byproduct indeed of the characterization of the aggregated power distribution—will demonstrate to be easier and straightforward to apply, without the need of intensive simulations.

Methodology

Marginals

We begin by presenting the random vectors which form the core of the statistical characterization, concentrating in this Section on the transformation existing between them. In an N-generator analysis, let $\mathbf{W} = (W_1, \ldots, W_N)$ be the vector of random wind speeds. The corresponding vector of wind powers is $\mathbf{P} = (P_1, \ldots, P_N)$. We assume that each i-th component of $\mathbf{W}$ is related with the corresponding i-th entry of $\mathbf{P}$ through a known wind speed–power curve (WSP) of the turbine located at the i-th site.

There is a large deal of research work on modeling wind speed distributions. Preferably Weibull, but also Rayleigh (a special Weibull case in which the shape is 2), inverse Gaussian, lognormal, or even bimodal distributions, have been published to characterize the random distribution of wind speed. In what follows, these wind speed cumulative distribution functions (CDF) and probability distribution functions (PDF) will be denoted by F_{W_i} and f_{W_i}. These are the *marginal* functions of i-th site. (Alternatively when we refer to the multivariate *joint* distribution function of wind speed, we will denote it by $F_{\mathbf{W}}$.)

It is in wind power, however, that we are interested here. More particularly, we aim at finding the value of the total aggregated power produced by N generators, which we will denote by $P_\mathrm{T} = \sum_{i=1}^{N} P_i$. Obviously, P_T is also a random variable—the sum of random variables—with a distribution F_{P_T}. Characterizing this distribution is one of the aims of this paper. We have divided the problem into two subproblems: (i) the characterization of the marginal wind power CDFs and (ii) the aggregation proper of the marginals.

The WSP characteristic of most wind turbines is a piecewise-defined function consisting of four main sections, which are conventionally separated by three characteristic wind speeds; see $h(w_i)$ in Fig. 1. The cut-in speed, w_i, is the lower wind speed at which the turbine can produce power. Rated speed, w_r, is the wind speed above which the output power is maximum. Finally, cut-out speed, w_f, represents a higher value of the wind speed limit, above which pitching the blades cannot reduce the power output to be below the rated power. Above w_f the power output is then null.

This special definition of the WSP function makes it problematic finding F_{P_i} from F_{W_i}. Although F_{W_i} is smooth, h is not, and thus it is necessary to eventually characterize F_{P_i} also as a piecewise function:

$$F_{P_i}(p_i) = \begin{cases} 1 - F_{W_i}(w_{\mathrm{fi}}) + F_{W_i}(w_{\mathrm{ii}}), & \text{if } p_i = 0\,\text{p.u.} \\ 1 - F_{W_i}(w_{\mathrm{fi}}) + F_{W_i}[h^{-1}(p_i)], & \text{if } 0 < p_i < h(w_{\mathrm{ri}})\,\text{p.u.} \\ 1, & \text{if } p_i = 1\,\text{p.u.} \end{cases} \tag{1}$$

The first line of (1) indicates that the probability of null P_i is equal to the probability that the corresponding wind speed is either above or below the cut-out and cut-in wind speeds, respectively. The third line indicates that the probability that the power output is below the rated power is one. Finally the second line includes a straight transformation from F_{W_i} into F_{P_i}, which can be performed within the cut-in and rated wind speeds, where the transforming function h is invertible. ($F_{P_i}(p_i) = \mathrm{pr}\{P_i \leqslant p_i\} = \mathrm{pr}\{h(w_i) \leqslant p_i\} = \mathrm{pr}\left\{w_i \leqslant h^{-1}(p_i)\right\} = F_{W_i}[h^{-1}(p_i)]$.) Note that for proper definition, the probability that the wind speed exceeds w_{fi} must be added to $F_{W_i}[h^{-1}(p_i)]$ in this second line.

The wind power CDF, F_{P_i}, is thus expressed in terms of the wind speed CDF, F_{W_i}, and the WSP curve, $h(w_i)$. Fig. 1 shows the remarkable difference between F_{W_i} in panel (a) and F_{P_i} in panel (b), a consequence of the combination of both functions. Particularly,