Applied Energy 162 (2016) 21–30

Contents lists available at ScienceDirect

Applied Energy

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

Wind power scenario generation through state-space specifications for uncertainty analysis of wind power plants

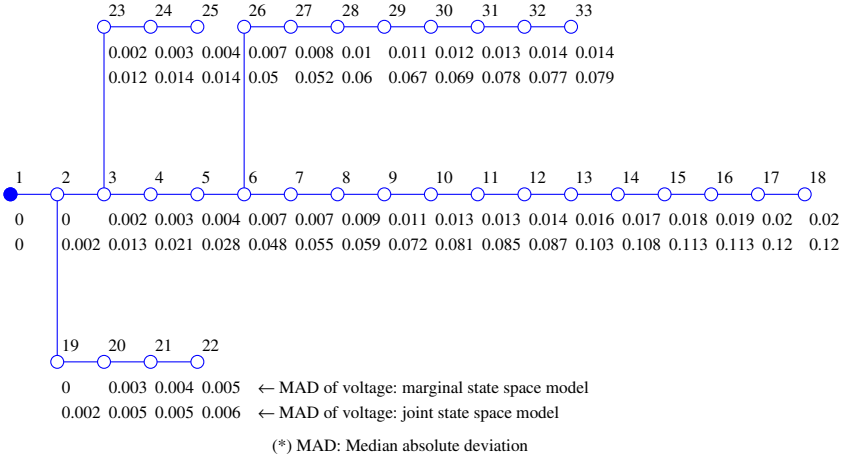
Guzmán Díaz *, Javier Gómez-Aleixandre, José Coto

Dep. of Electrical Engineering, University of Oviedo, Campus de Viesques, s/n, 33204, Spain

HIGHLIGHTS

- State space representations for simulating wind power plant output are proposed.
- The representation of wind speed in state space allows structural analysis.
- The joint model incorporates the temporal and spatial dependence structure.
- The models are easily integrable into a backward/forward sweep algorithm.
- Results evidence the remarkable differences between joint and marginal models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 29 June 2015
Received in revised form 5 October 2015
Accepted 6 October 2015
Available online 11 November 2015

Keywords:
Wind power
Multivariate stochastic processes
Simulation
State space

ABSTRACT

This paper proposes the use of state space models to generate scenarios for the analysis of wind power plant (WPP) generation capabilities. The proposal is rooted on the advantages that state space models present for dealing with stochastic processes; mainly their structural definition and the use of Kalman filter to naturally tackle some involved operations. The specification proposed in this paper comprises a structured representation of individual Box–Jenkins models, with indications about further improvements that can be easily performed. These marginal models are combined to form a joint model in which the dependence structure is easily handled. Indications about the procedure to calibrate and check the model, as well as a validation of its statistical appropriateness, are provided.

Application of the proposed state space models provides insight on the need to properly specify the structural dependence between wind speeds. In this paper the joint and marginal models are smoothly integrated into a backward–forward sweep algorithm to determine the performance indicators (voltages and powers) of a WPP through simulation. As a result, visibly heavy tails emerge in the generated power probability distribution through the use of the joint model—incorporating a detailed description of the dependence structure—in contrast with the normally distributed power yielded by the margin-based model.

* Corresponding author.
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.apenergy.2015.10.052
0306-2619/

1. Introduction

Uncertainty analysis of a wind power plant (WPP) provides knowledge about the reliability of its design parameters, its integration into the power system, and ultimately about decisions resting on its estimated performance [1]. Essentially, these analyses aim at producing probabilistic distributions of selected performance indicators (voltages, powers, etc.) subject to the uncertain variation of independent variables. Wind speed is arguably the most significant of those variables in a WPP. Its random variations—with involved both temporal and spatial dependencies—makes scenario generation through simulation a most valuable tool to facilitate the uncertainty analysis.

Cross-sectional sampling is a first suite of methods for simulating wind speed to investigate WPP performance. They are the basis of Monte Carlo analyses in which time as a variable is of no interest. In these analyses the extraction of samples is not necessarily sequential. Indeed, vector operations are indicated to improve sampling speed [2]. In the wind power literature, several versions appear. The simplest rest on drawing unstratified samples of the probability distribution [3], or stratified through Latin hypercube-sampling (LHS) [4,2] or lattice sampling [5] to improve performance. They are simple to use because they do not necessarily require parameter estimation. If the marginal distribution is obtained through a kernel estimation, the distribution parametric specification can be avoided [2]. Even so, they may accurately model simple spatial dependence between pairs of machines by using a linear transformation based on the Cholesky decomposition of the correlation matrix [3,2]. Alternatively, where the dependence structure is more involved, copula methods have been applied, but following the same time independence [6,2].

At times it is necessary not only to focus on the probabilistic properties of the wind power sample, but also to show the longitudinal dependence structure, which stands for sequential sampling. That is the case when the wind power must be confronted to other stochastic processes—electricity price being the most relevant [1]—or when the evolution of a power system is investigated [7–9]. Box–Jenkins's ARMA models—with the property of resting in past values to regress the actual wind speed—have been favored in such cases. Indeed, Billinton et al. claimed that any individual wind speed process may be modeled by ARMA$(n, n-1)$ models [9]. And Torres et al. after intensive research concluded that other more parsimonious ARMA specifications also represented these processes adequately [10].

However, the ability to incorporate a sequential dependence makes ARMA-based models more complex to employ than their cross-sectional counterparts. The two major problems are the parameter estimation of individual wind speed series and the incorporation of spatial cross-correlation between sources. The first issue requires trial and error procedures as well as expert judgment, and it has been sufficiently covered in the related literature; including the classical work by Box and Jenkins [11]. The second issue, the correlation, has been addressed in the wind power literature in two ways: one resting on forcing the correlation to estimated individual models, and other using compound models covering several wind speed sources simultaneously. An instance of the first approach is reported in [12], where Gao et al. proposed a modification of the random number generation to affect the MA errors in such a way that the correlation was forced. The model was complex because it required a heuristic search of the appropriate seeds. Also following the individual path, Morales et al. proposed in [13] a methodology based on Nataf's method, popularized in [14], to obtain correlated samples of wind speed after having estimated the individual models. The correlation was incorporated by employing a technique of transformation similar to that in [3,4]. Alternatively other authors have recently followed the compound model path by employing vector autoregressive (VAR) models. For instance, in [15] VAR(p) models were employed for simulating wind speeds subject to directional components. Correia et al. restricted their analysis to VAR(1) models [16], and Hill et al. to VAR(2) [17]. The common feature in these studies is that the authors employed VAR, but not VARMA, models. That is, the error regression was not considered, though it has been stated in [9,10] that it is a fundamental component.

A recent addition to the previous specifications of wind speed autoregressive models is that of Chen and Yu in [18]. They proposed the translation of an AR model into state space (SS) form. Indeed, AR models are but a subfamily of the more general SS models. The ensuing advantages of using Chen and Yu's approach, rather than Box–Jenkins's, were detailed by Durbin and Koopman in [19, Section 3.2.1]. First and foremost, the problem can be structurally analyzed. This is in contrast with Box–Jenkins approach, which does not investigate the structure of the problem. This structural analysis makes the SS approach really flexible for incorporating trends and seasonalities. By contrast, Box–Jenkins approach requires a previous deseasonalizing and detrending. In addition, Durbin and Koopman cite other superior features of SS models compared with ARMA specifications, such as for instance the treatment of missing observations, the easy incorporation of explanatory variables, the possibility of time-varying regression coefficients, and the use of Kalman filter to naturally forecast forward in the future (the subject of [18]).

This paper contributes to the literature on wind power scenario generation by proposing a SS representation of the wind speed. The contributions with respect to previous works are the following:

1. First, Section 2.2 generalizes the SS model in [18] to also consider the contribution of previous unobserved errors. The ensuing generalized model exhibits a structure that makes it susceptible to easy expansions.
2. The use of Kalman filtering to estimate those marginal SS model parameters is favored by the transformation of the original dataset into Gaussian random variables. This paper shows that the method proposed in [20] though proving to be useful is incomplete, for it cannot cope with calm wind speeds. A solution to this problem is offered in Section 2.1.
3. This paper shows in Section 2.3 that a joint SS model can be easily built, preserving the structure of the marginal models. The joint model expands the VAR formulations in [15–17], by incorporating the MA terms, as advised in [9,10]. Further, it completes the marginal state space model in [18] by integrating the spatial dependence between sources through the use of a multivariate white noise into the transition equation, with covariances estimated from the original data.
4. Finally, this paper shows how to integrate the model into a backward–forward sweep algorithm to obtain the simulated performance of a WPP (Section 3.2). Moreover, clear evidence about the error of employing non-dependent wind speeds to simulate the aggregated power generation of a WPP is provided: Though there may be no deviations on the mean node voltages and generated power, the extreme and more probable values are visibly different.

2. State space model characterization

This section describes the proposed procedure for building the SS model of wind speed of a WPP. The model is built in a normalized space, starting with an uncorrelated SS model in which the marginal distributions are independent, and ending with the specification of the correlation.

2.1. Normalization

Wind speed distributions are not generally Gaussian. Actually the most popular probability distribution employed for representing wind speed is the two-parameter Weibull (see for instance [21,22] and references therein). This is a visibly skewed distribution that represents the superior frequency of relatively low wind speeds by means of two parameters: scale (which serves as an approximation of the mean wind speed) and shape (which refers to the skewness and kurtosis). As a matter of fact, only when the shape parameter is in the open interval (3.25, 3.61) the Weibull can be assimilated to a Gaussian distribution according to [23]. But this requires a steady provision of wind speeds, which only occurs at few locations. For instance, the trade winds are known to be fairly constant, and they meet this bound condition at some locations [24]. However, the literature reports that their distribution may not be Gaussian either, but for instance bimodal [25].

It is controversial whether the non-normal distributed observations should be subject to a normalization prior to the calibration of a model. Several authors advocate for this preprocessing of wind speed [26,13,27,28,10]. Others on the other hand, as Chatfield in his classic book [29, Section 2.4], advice against this practice on the ground that the transformed variable may lack physical interpretation. Indeed, strictly speaking, in autoregressive models normally distributed observations are not crucial, but instead the lack of correlation and finite variance of the error term.

In this paper we have chosen the normalization option for practical reasons. First, normalization extends the range of possible analysis: AIC (Akaike information criterion) for order selection or the asymptotic prediction of confidence bounds, for instance [30, Section 5.4]. Second, the proposed model primarily depends on Kalman filter for a number of operations (parameter estimation, forecasting, etc.). Employing Gaussian random variables allows using an optimal Kalman filter algorithm; for which well-established implementations exist. Finally, as it is shown below, a modified transformation easily deals with discontinuities in wind speed, making calibration easier.

A normalization method that has proven to be effective is reported in [8,26,13,31] based on the use of cumulative distribution functions (CDFs). In [20] it has been termed a "memoryless" transformation. The rationale behind this transformation is that the probability of not exceeding a value must be the same for transformed and untransformed observations. This is achieved by means of the CDFs in two steps: a first transformation of wind speed into uniform random variables, and a second transformation converting the uniform variable into normally distributed variables. In compact form it reads:

$$w'_t = \Phi(F_W^{-1}(w_t)). \tag{1}$$

$\Phi(\cdot)$ is the standard Normal CDF. There are not restrictions on the specification of the wind speed CDF, $F_W(w_t)$. It can be parametric [8,26,13], which permits obtaining closed formulas for the transformation [31]. Or alternatively it can be non-parametric, based on kernel density estimation [2]. We have opted for this second alternative, because it enables calm wind to be taken into account.

Calm winds introduce a jump at the origin in the cumulative distribution of wind speed. In the probability density function, this is shown as a disruptive infinite value at the origin. The height of the jump indicates the frequency of such winds. This discontinuity makes it difficult to produce a simple SS model that yields pure zeros in a proportion statistically similar to that of the original series. In cross-sectional analysis, Takle and Brown proposed the use of hybrid Weibull distributions [32], where a delta operator indicates the existence of null wind speed. However, we did not find a parallel, simple solution for the SS model.

The solution proposed in this paper to account also for calm winds is based on an expanded inverse CDF of wind speed, defined as

$$F_W^{*-1}(w_t) = \begin{cases} \mathcal{U}(0, \mathrm{pr}W = 0), & \text{if } w_t = 0 \\ F_W^{-1}(w_t), & \text{otherwise} \end{cases} \tag{2}$$

This is congruent with the original Nataf transformation, in which the first step is the transformation into uniform random variables. In (2) that first transformation is preserved for non-calm winds. But in the case of calm winds the transformation is obtained directly through a uniform sampling, bounded between zero and the probability of wind speed being zero. This probability represents the height of the "jump" of the CDF.

2.2. Marginal model

The conventional Box–Jenkins ARMA(p,q) model for wind speed, w_t, is:

$$w_t = c + \sum_{i=1}^{p} \phi_i w_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{3}$$

where the ϕ_i are the parameters of the autoregressive component, the θ_i are the parameters of the moving average part, and the $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the error term.

In state space form, with state vector $\mathbf{x}_t$ and observation vector $\mathbf{y}_t$, the wind power model is:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}u_t \tag{4a}$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}v_t, \tag{4b}$$

where u_t and v_t are uncorrelated, unit-variance white noise vector processes. The first equation is referred to as the state equation and the second as the observation equation. The matrices defining the model, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, are the transition, disturbance loading, measurement sensitivity, and observation innovation matrices, respectively.

To model the ARMA form (3) in SS form, first we let the wind speed w_t be specified by the state variable $x_{1t} = w_t$. Further, the p AR elements are represented by hidden states, so that recursively $x_{it} = w_{t-(i-1)}$, with $i = 2, \ldots, p$. That is, each hidden state represents a past observation of wind speed. Similarly for the MA terms, a state variable directly represents the disturbance of the state equation as $x_{(p+1)t} = \epsilon_t$; and recursively the rest of disturbance regressions are defined through the use of hidden states, $x_{jt} = \epsilon_{t-(j-1)}$, with $j = p+2, \ldots, p+q$. Eventually by following this approach and comparing (3) and (4a), we can establish a state space representation of wind speed through Eq. (5), where $u_t \sim \mathcal{N}(0,1)$.

$$\underbrace{\begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ \vdots \\ x_{pt} \\ \hline x_{(p+1)t} \\ x_{(p+2)t} \\ \vdots \\ x_{(p+q)t} \\ \hline x_{(p+q+1)t} \end{pmatrix}}_{\mathbf{x}_t} = \begin{pmatrix} w_t \\ w_{t-1} \\ w_{t-2} \\ \vdots \\ w_{t-p+1} \\ \hline \epsilon_t \\ \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-q+1} \\ \hline x_{(p+q+1)t} \end{pmatrix} = \underbrace{\left(\begin{array}{cccc|cccc|c} \phi_1 & \phi_2 & \cdots & \phi_p & \theta_1 & \theta_2 & \cdots & \theta_q & c \\ \hline 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right)}_{\mathbf{A}_t} \begin{pmatrix} w_{t-1} \\ w_{t-2} \\ w_{t-3} \\ \vdots \\ w_{t-p} \\ \hline \epsilon_{t-1} \\ \epsilon_{t-2} \\ \vdots \\ \epsilon_{t-q} \\ \hline x_{(p+q+1)(t-1)} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma_\epsilon \\ 0 \\ 0 \\ \vdots \\ 0 \\ \hline \sigma_\epsilon \\ 0 \\ \vdots \\ 0 \\ \hline 0 \end{pmatrix}}_{\mathbf{B}_t} u_t, \tag{5}$$

By inspection it can be seen that this representation includes all the features displayed in the Box–Jenkins model in (3) through a highly sparse matrix of size $(p+q+1)\times(p+q+1)$. That matrix is **A** in (4a). Except for the first row, the matrix is filled exclusively with zeros and ones, and importantly it can be easily and efficiently built in a structured way if it is divided into the component blocks shown in (5).

This representation of the transition equation has some particular features. First, if needed the constant term c can be represented through the last row of (5), in which it is simply stated that $x_{(p+q+1)t} = x_{(p+q+1)(t-1)}$; which means that the last state does not vary. In addition, expansion or contraction to different orders is immediate. It can represent any ARMA(p,q) model—including for instance mean-reverting processes employed in specifying electricity spot prices, which are but ARMA(1,0) [33]. However, when p or q are zero, care must be taken in avoiding eliminating the corresponding term. For instance, an ARMA(0,1) without constant term process must not be represented by a $(p+q)\times(p+q) = 1\times 1$ matrix. The non-regressed state space component, w_t, and the disturbance component, ϵ_t, must always be present (though with the corresponding parameter set to zero). More specifically, the ARMA(0,1) process would be:

$$\begin{pmatrix} x_t \\ \epsilon_t \end{pmatrix} = \begin{pmatrix} 0 & \theta_1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ \epsilon_{t-1} \end{pmatrix} + \begin{pmatrix} \sigma_\epsilon \\ \sigma_\epsilon \end{pmatrix} u_t \tag{6}$$

which yields

$$x_t = \theta_1 x_{t-1} + \underbrace{\sigma_\epsilon \overbrace{u_t}^{u_t\sim\mathcal{N}(0,1)}}_{\epsilon_t\sim\mathcal{N}(0,\sigma_\epsilon^2)} \tag{7}$$

Importantly, the representation in (5) is a translation of an autoregressive specification of a wind speed process, which can be readily enriched. For instance, seasonal patterns can be easily incorporated. Laine et al. show in [34] how by inserting blocks of the form $\begin{pmatrix} \cos(2k\pi/12) & \sin(2k\pi/12) \\ -\sin(2k\pi/12) & \cos(2k\pi/12) \end{pmatrix}$ in **A**, where k is the harmonic order, and correspondingly expanding the state vector and disturbances in pairs, cyclic components can be considered. That is, the model can intrinsically separate seasonal from purely autoregressive components.

The disturbance matrix **B** also has a structured representation, resting on the component blocks shown in (5). Its size is $(p+q+1)\times 1$, and it provides the information about the variance of the process. Its relation with the innovation terms in (3) is emphasized in (7).

Ultimately only the wind power is of interest, while the hidden states are not particularly relevant. To reveal only this state, it suffices to employ the observation Eq. (4b) to obtain the current wind speed:

$$w_t = (1 \quad 0 \quad \cdots \quad 0)\mathbf{x}_t. \tag{8}$$

It follows that wind power is obtained as:

$$P_t = g(w_t), \tag{9}$$

where $g(w_t)$ is the mapping from wind speed to wind power, according to the characteristic power transformation curve of the wind turbine.

The observation innovation matrix **D** is here arbitrarily set to zero, so that the model is exactly an equivalent to the Box–Jenkins model. However, were there measurement errors or uncertainties in the wind speed measurement, these would be accounted for in SS formulation by explicitly stating a non-zero matrix **D**.

2.3. Joint model

The previous model describes the autoregressive form of the wind speed at one only site. One of the advantages of employing the state space representation proposed in this paper is that it is readily expandable to account for correlated wind speeds at different sites. The proposed structure is as follows:

$$\mathbf{A} = diag(\mathbf{A}_i) \tag{10a}$$

$$\mathbf{B} = diag(\mathbf{B}_i) \tag{10b}$$

$$\mathbf{C} = diag(\mathbf{C}_i) \tag{10c}$$

That is, initially the model is simply expanded by diagonally stacking the components of each i-th marginal model. This procedure, however, provides an uncorrelated model of wind speed.

The correlation can be introduced either in the transition matrix **A** or, more simply, in the second term of (4a). Both formulations would be equivalent. However, operating on the second term of (4a) makes it more straightforward the specification of the correlation, because it can be directly imported from the analysis of correlation of the original sample. Besides, the general matrix **A** can be in this way built through stacking of the marginal components, without introducing off-diagonal terms which would complicate the characterization of **A**. When the system is expanded to represent n_s generation sites, the disturbance u_t also expands to a vector $\mathbf{u}_t$ of size n_s; each entry representing the stochastic deviation of each marginal wind speed. For simulation of correlated wind speeds, the key is not in producing Normal i.i.d. for each component of the vector $\mathbf{u}_t$, but doing it jointly from a multivariate normal distribution. That is, the sampling should not be done as $u_{it} \sim \mathcal{MVN}(0,1)$, but following $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0},\mathbf{Q})$. Hence matrix **Q**—the covariance matrix obtained from the original wind speed series—helps model the correlation among the process deviations through the product $\mathbf{Bu}_t$. The variance of the marginal models was already included in **B** through the specification of the marginal models in (5). This accordingly requires the different wind speeds to follow deviations that obey the original correlation between series. It is noticeable that if the correlation were instead specified in matrix **B** as a sum of one process own deviation plus the deviation proposed by adjacent processes through correlation (with the samples of deviations being i.i.d. from a standard normal distribution) the maximum values of wind speeds would reach values several times higher than those of the original series. This would have been a consequence of the lack of control about the amount of the deviations, by considering the errors additively.

The procedure is summarized in the following steps.

- Normalize the k input time series by (a) employing an empirical or kernel estimate-based CDF of each sample to produce uniform samples and (b) obtaining the Gaussian values of those uniform samples. Time sequence of the observations must be preserved.
- Decide equivalent ARIMA order p and q, and build the corresponding *marginal* state space models.
- Individually fit each marginal model to the normalized series by a combination of Kalman filtering and minimum loglikelihood. (Follow Durbin and Koopman guidelines for guessing the initial values [19].)
- Check the validity of the fitness: whiteness and uncorrelation of the standardized residuals must be ensured. If not, suggest other p and q values and repeat.
- Build the joint model from the marginal models by diagonal stacking. As an instance of the result, the following are the joint model for sites 25,225 and 25,228 of NREL dataset:

$$\mathbf{x}_t = \left(\begin{array}{cccc|cccc} 0.951 & 0.278 & -0.022 & -0.018 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ \hline 0.000 & 0.000 & 0.000 & 0.000 & 0.936 & 0.255 & 0.159 & 0.041 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & 0.000 \end{array}\right) \mathbf{x}_{t-1} + \left(\begin{array}{c|c} 0.069 & 0.000 \\ 1.000 & 0.000 \\ 0.000 & 0.000 \\ 0.000 & 0.000 \\ \hline 0.000 & -0.107 \\ 0.000 & 1.000 \\ 0.000 & 0.000 \\ 0.000 & 0.000 \end{array}\right) \mathbf{u}_t.$$

$$\mathbf{C} = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right).$$

The marginal models and the stacked structure are readily seen ($p = 1$ and $q = 3$).

- Compute the correlation between the original wind speed samples. For sites 25,225 and 25,228 it is 0.896. Consequently the covariance matrix is

$$\mathbf{Q} = \begin{pmatrix} 1 & 0.896 \\ 0.896 & 1 \end{pmatrix}.$$

Therefore $\mathbf{u} \sim \mathcal{MVN}(\mathbf{0}, \mathbf{Q})$.

- Simulate normalized wind speeds by drawing a sample of $\mathbf{u}_t$ at each time step t, to get the value of $\mathbf{x}$ at $t+1 : \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t$. See Durbin and Koopman [19] for guidelines about the first sample guess.
- The simulated samples of *normalized* wind speed are $\mathbf{y} = \mathbf{C}\mathbf{x}$.
- Finally, "untransform" by reverting step 1: (a) the simulated sample is transformed to uniform by the Gaussian inverse CDF, and (b) those uniform samples are transformed through the estimate (kernel-based or empirical) CDF of the original sample.

The entire procedure for characterizing the SS model of the wind speed of the WPP, including correlation, is summarized in Figs. 1 and 2.

3. Case analysis

In what follows two case analyses are presented. One employs a reduced number of generators. Its objective is to illustrate the most relevant features of the proposed model. The second analysis expands the study to a full 32-generator WPP to further demonstrate the ability of the model to smoothly integrate with WPP load flows, and to show the errors when the marginal models are employed rather than the joint versions.

3.1. Six-generator model

In this section we illustrate the procedure for scenario generation by using a small scale WPP composed of six generators. In order to focus on the procedure, what follows does not take into account the load power flow (i.e. the line losses) in the WPP.

3.1.1. Data set

To illustrate the procedure for scenario generation, we first selected a six-generator sample from the NREL data set in The data employed to illustrate the SS specification can be found in [35]. The site identifications and the main features about wind power production are listed in Table 1. We carefully selected the data set, all corresponding to sites close in proximity, to have a varied representation of wind profiles and correlation levels. The data set contains measurements for three years (we employed 2006) with observations of wind speed every ten minutes.

3.1.2. Marginal model estimate

As with Box–Jenkins models, the first step is to calibrate the model. The calibration comprises both a trial and error procedure

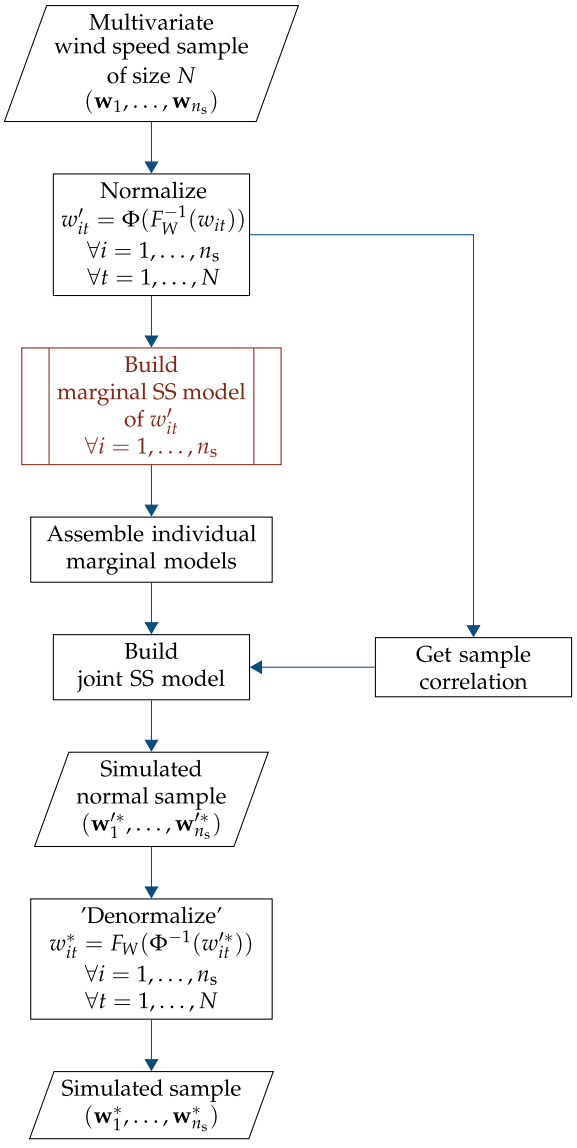


Fig. 1. General procedure for simulating correlated wind samples.

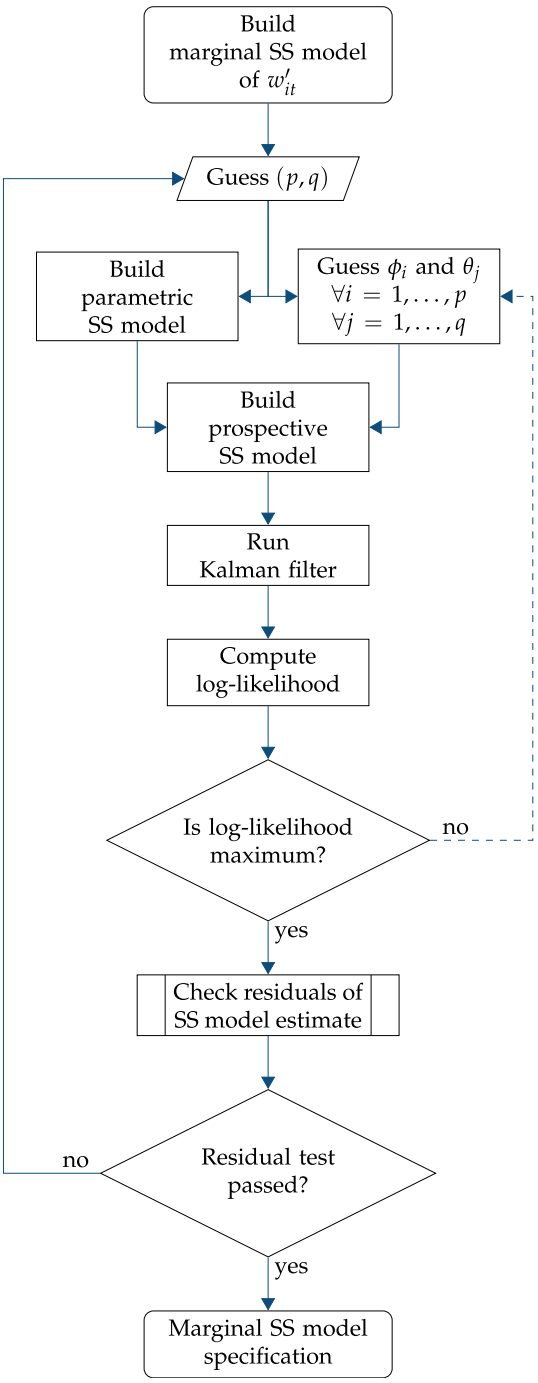


Fig. 2. Detail of the procedure for building the individual marginal models in Fig. 1. The dashed line loop indicates that the procedure of finding the ϕ_i and θ_j is undertaken by an optimization routine.

and expert judgment. The process must test candidate models of different orders and, subsequently, assess their accuracy. In this respect, we based on the information provided by Torres et al. in [10]. After investigating 54 wind datasets employing autoregressive models in the form of ARMA(p,q), they concluded that the most accurate models were those described by a first order autoregression, with $q \in [2,4]$ q an integer in the closed interval $[2,4]$. Therefore, for each site we assumed that the processes were governed by a generic system of five state variables, through the transition Eq. (5), having as unknown parameters $\phi_1, \theta_1, \ldots, \theta_4$, and σ_ϵ.

For normalization of wind speed, we employed the kernel density estimation algorithm in [36] to non-parametrically estimate

Table 1
Main features of the investigated data set.

NREL id.	Cap. factor (%)	Power dens. (W/m^2)	Mean speed (m/s)
25,181	38.8	849.2	8.7
25,228	23.6	600.9	6.7
25,331	35.0	759.0	8.2
25,377	26.4	389.2	6.9
25,194	43.1	1162.9	10.2
25,293	31.8	756.0	7.9

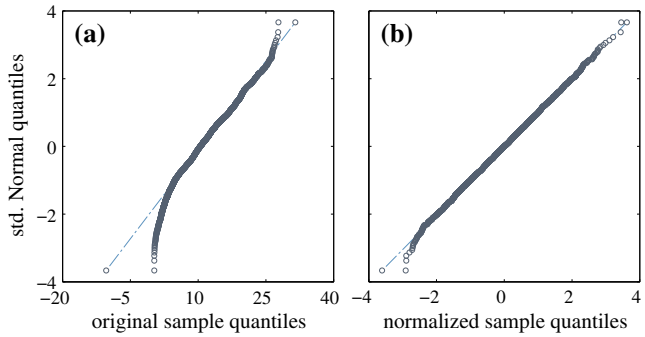


Fig. 3. Normal probability plots of NREL id. 25,377. (a) Original series. (b) Normalized series after employing the first part of Nataf transformation using an intermediate kernel density estimation.

the marginal wind speed CDFs. It is a fast algorithm that provides the uniform version of wind speed detailed in (2). The difference between distributions (pre- and post-transformed) is shown in Fig. 3.

Because of the normalization of variables the model is now a linear Gaussian SSM, and the optimal Kalman filter can be advantageously employed to construct the log-likelihood function, through which the unknown parameters can be estimated [37, Section 5.3]. The procedure is quite standard, thanks to the normalization of data: The Kalman filter recursively estimates the unobserved components at each observation, based on the previous available information, which permits a dynamic estimation of the log-likelihood function. This process of parameter estimation is individually repeated for each wind speed series in the data set.

The calibration procedure also encompasses adequacy checks, to corroborate whether the residuals of the fit have the appropriate specification (normality and independence). Durbin and Koopman showed in [19] that the analysis of autocorrelation plots of standardized prediction errors provide a quick measure of the effectiveness of the model in reproducing the characteristics of the original series. The procedure comprises the following steps. First, the Kalman filter is employed to filter the estimated model against the original wind speed series. Then using the filter output, the prediction error and its covariance are obtained by comparing the original and forecast series. The standardized error is the quotient between the mean and the square root of the covariance. This is the error that eventually is fed to the autocorrelation plot. As for normality compliance, it can be assessed by a number of specialized tests. Particularly, we employed the well-known Augmented Dickey–Fuller test.

Fig. 4 illustrates this point. The standardized prediction error is shown overlapping the original normalized series in the top panel. The normalization makes the original series oscillate around the zero mean, presenting negative speeds. The prediction error shows also the same oscillating behavior, with larger deviations where the original series presents more oscillations. Eventually, with the parameter values given in the caption, the prediction shows a normal distribution with autocorrelation at different lags fairly withing the standard deviation margins.

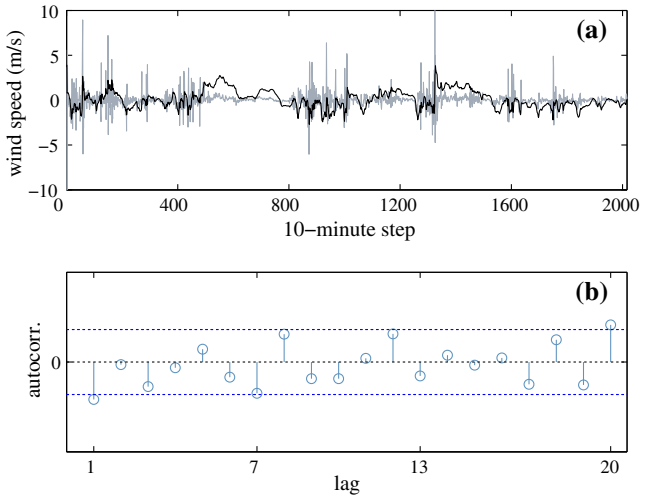


Fig. 4. Standardized error analysis of NREL site id. 25,377 for the following parametric specification: $\phi_1 = 0.9599$, $\theta_1 = 0.0741$, $\theta_2 = 0.0147$, $\theta_3 = \theta_4 = 0$, and $\sigma_\epsilon = 0.9443$. (a) Two weeks of original normalized series (black line) and standardized prediction error (gray line). (b) Autocorrelation plot.

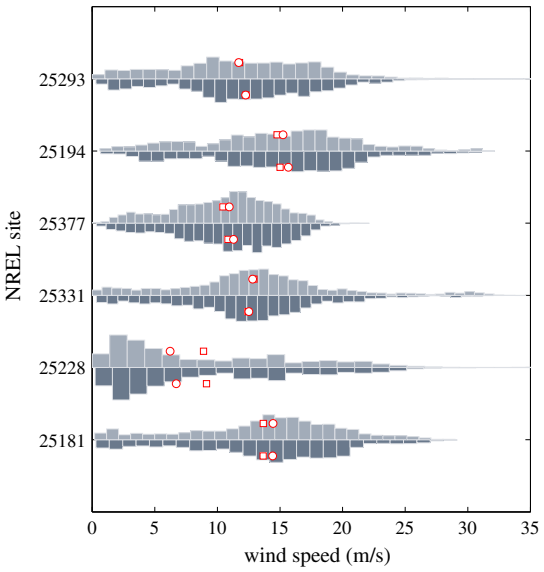


Fig. 5. Histograms comparing the original (top, light shade) and the simulated (bottom, dark shade) series. The markers are the sample mean (□) and median (○).

Table 2
Wind speed correlation of the original series.

NREL id.	25,181	25,228	25,331	25,377	25,194	25,293
25,181	1.0000	0.1099	0.6233	0.6645	0.8823	0.6131
25,228	0.1099	1.0000	0.4532	0.4323	0.2929	0.5339
25,331	0.6233	0.4532	1.0000	0.7024	0.6972	0.7933
25,377	0.6645	0.4323	0.7024	1.0000	0.6715	0.7521
25,194	0.8823	0.2929	0.6972	0.6715	1.0000	0.6764
25,293	0.6131	0.5339	0.7933	0.7521	0.6764	1.0000

We repeated these steps until we were satisfied with the results. If one marginal model parameterization did not comply with the residual requirements, we would check another fit, but without changing the model structure. It suffices to introduce constraints to the values of the parameters at the stage of maximization of the log-likelihood function. For instance, an ARMA(1,2) model can be imposed in the general structure by setting $\theta_3 = \theta_4 = 0$. Eventually this trial and error procedure yields the marginal, uncorrelated SS models of wind speed at the six sites.

3.1.3. Joint model

After the estimation of the marginal models, we proceeded with the specification of the joint model. As explained above, this specification consists of a stacking of the marginal model matrices $\mathbf{A}_i$, $\mathbf{B}_i$, and $\mathbf{C}_i$, $i = 1, \ldots, 6$; with the provision of a multivariate error term. The ensuing model has 30 state variables, inherited from marginal models of 5 states. Simulations can then be obtained through the recursive extraction of error samples from a multivariate normal distribution of zero mean, employing the original correlation matrix to compute the off-diagonal entries of the covariance matrix of the distribution (Table 2).

3.1.4. Model performance assessment

The accuracy of the model in representing the statistical distribution of the wind speed is corroborated in Fig. 5. The selected data set comprised wind speed records of difficult specification through parametric distributions (Weibull for instance). This is the case of site id. 25,293 or 25,194, where the fitting of a unimodal distribution seems difficult. However, the SS model estimate produces simulated data of characteristics similar to those of the original data.

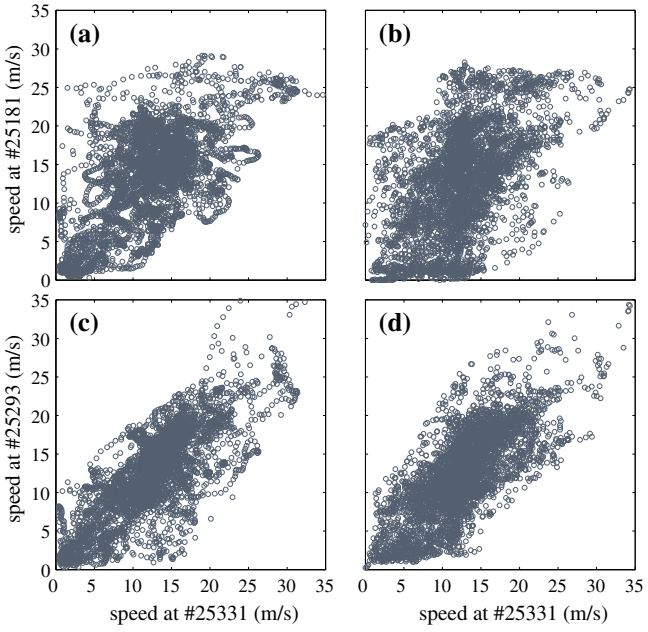


Fig. 6. Correlation plot comparing the original data set (left) to the simulated data (right).

Table 3
Wind speed correlation of the simulated series.

NREL id.	25,181	25,228	25,331	25,377	25,194	25,293
25,181	1.0000	0.1144	0.5183	0.6708	0.8536	0.5761
25,228	0.1144	1.0000	0.3752	0.1807	0.2143	0.4341
25,331	0.5183	0.3752	1.0000	0.5733	0.4728	0.7402
25,377	0.6708	0.1807	0.5733	1.0000	0.5695	0.6994
25,194	0.8536	0.2143	0.4728	0.5695	1.0000	0.5717
25,293	0.5761	0.4341	0.7402	0.6994	0.5717	1.0000

Not only the shapes (modes) of the distribution are well replicated, but also the mean and median statistics are coherent.

The correlation among processes is investigated next. Fig. 6 shows again original and simulated data, and Table 3 provides

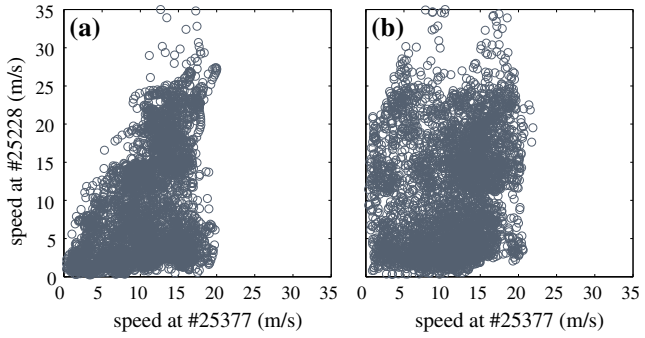


Fig. 7. (a) Original data. (b) Simulated data.

the correlation values obtained from the simulated series. Panels (a) and (b) in Fig. 6 compare the correlation between two wind speed series of relatively low correlation (0.62 in the original data set and 0.52 in the simulated series); and (c) and (d) show the match of correlation between simulated and original data of two series of higher correlation (about 0.79). In both cases, high and low correlation, the model is fairly accurate in representing the statistical dependence between wind speed series. This accuracy is repeated throughout all the data set, with few exceptions, such as the case of the site pairs 25,377/25,228 and 25,194/25,331. We investigated these cases and found that the lack of accuracy occurs because of the difficulty in modeling complex correlation patterns through a unique parameter. The off-diagonal elements of the correlation matrix indicate the strength of a linear relationship between the two involved variables. Fig. 7 shows that the simulated data correctly complies with the approximate limits of the wind speed shown in the original data (around 20 m/s for site 25,377 and 35 m/s for site 25,228) because it depends on the individual fitting of each time series. However, the "distortion" of the original correlation plot cannot be reproduced by the simulated data, where a more linear, even normalized, data is observed.

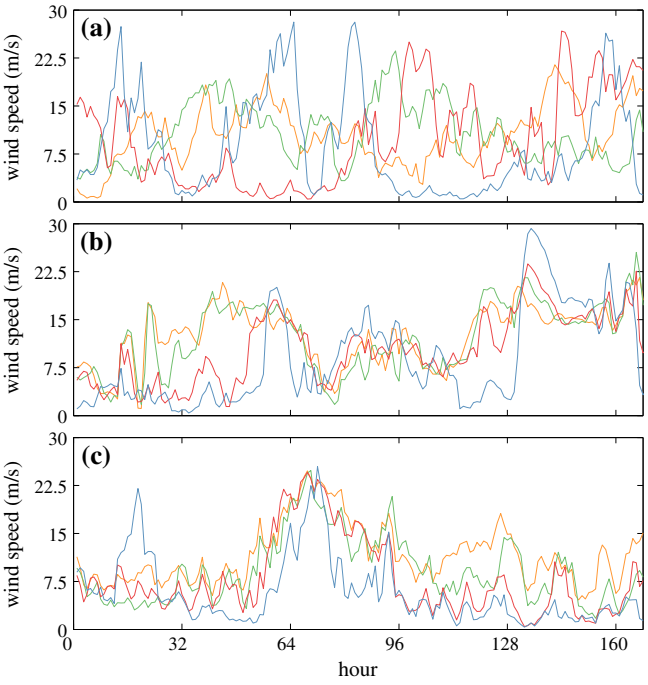


Fig. 8. One-sample simulation of the six time series detailed in Table 1. (a) Simulation without considering the correlation—i.e., simulation before building the joint model. (b) Original data set. (c) Simulation of the joint model.

Nonetheless, because the model is built so that the correlation is enforced in the error sampling, possibly a more complex sampling procedure—maybe based on vine copula theory—should facilitate a more exact specification of the dependence.

Finally, the need for a correlation adjustment upon the building of the complete model is further emphasized in Fig. 8. The underlying variability an reversion obtained from the simulation of the models—either before compounding or after compounding—is similar to that of the original data set: see for instance the speed limits and the reversion speed of original and simulated data. However, the dependence observed in Fig. 8b is only reproduced in Fig. 8c, where the correlation is enforced at the time of recursively generating the observations.

3.2. 32-generator model

This section broadens the previous case analysis by integrating a larger number of processes into a simulated WPP. The main purpose is to asses the error of not modeling a compound system when the aggregated response is analyzed, and ultimately corroborate that the problem is tractable in a power flow framework even with such a large number of state variables.

3.2.1. Test system and load power flow

The selected distribution system representing the generator interconnection in the WPP was the 33-bus test system described in [38]. We defined a generator at each node except at the slack bus.

We integrated the wind power SS model into a backward–forward sweep (BFS) algorithm in order to compute node voltages, branch currents, and total produced power in the test system. To achieve the integration, we modified the direct approach presented in [39] by introducing a node incidence matrix, $\mathbf{\Gamma}$, for representing the test system. This is an $N_b \times N$ matrix, where N_b is the number of branches (32 in the test system analyzed here), such that its element γ_{ij} is [40]:

- $\gamma_{kl} = +1$ when current in branch k leaves node l;
- $\gamma_{kl} = -1$ when current in branch k flows towards node l; and
- $\gamma_{kl} = 0$ when no connection exists.

This matrix makes it possible to relate all the branch currents to the drawn currents at every node at step t by stating

$$\mathbf{I}_{branch,t} = \mathbf{\Gamma}^{-T}\mathbf{I}_{node,t}, \tag{11}$$

where the superscript $-T$ indicates inverse transpose [40]. And because the drawn currents are related to the drawn power, all the currents of the network can be computed in the backward sweeps as

$$\mathbf{I}_{branch,t} = \mathbf{\Gamma}^{-T}(\mathbf{P}_t \text{ø} \mathbf{U}_t), \tag{12}$$

where $\mathbf{U}_t$ is the vector of node voltages—assumed known in the backward sweep—and ø is the element-wise division of vectors.

The forward sweep assumes that the voltage at the slack bus is 1.0 p.u and that the currents computed in the backward sweep are correct. From that point and again employing matrix $\mathbf{\Gamma}$, the new node voltages can be computed in just one operation as:

$$\begin{aligned}\mathbf{U}'_t &= 1.0 - \mathbf{\Gamma}^{-1}\, diag(\mathbf{Z}_{branch})\,\mathbf{I}_{branch,t} \\ &= 1.0 - \mathbf{\Gamma}^{-1}\, diag(\mathbf{Z}_{branch})\,\mathbf{\Gamma}^{-T}(\mathbf{P}_t \text{ø} \mathbf{U}_t),\end{aligned} \tag{13}$$

where $diag(\mathbf{Z}_{branch})$ is the diagonal matrix of ordered line impedances.

The procedure is derivative free and efficient. It computes the value of voltages from the simulated injected powers in vector $\mathbf{P}_t$

in few seeps; hence providing a quick simulation of the network state at each extraction of a sample in the transition equation.

3.2.2. Data set

The data set employed to model the wind speed of the 32 generators has the following codes in the NREL data base [35]: 24,483, 24,484, 24,485, 24,486, 24,487, 24,488, 24,489, 24,490, 24,491, 24,492, 24,493, 24,494, 24,495, 24,496, 24,497, 24,498, 24,499, 24,430, 24,431, 24,432, 24,433, 24,680, 24,681, 24,682, 24,563, 24,564, 24,565, 24,566, 24,567, 24,568, 24,569, and 24,570. These codes are sorted here in the same order as the numbering of the nodes in the test system of [38]. Moreover, if the map in [35] is consulted, it can be observed that we intentionally selected the branches of the system following actual rows of generators, so that the dependencies in every branch were representing an actual WPP.

The simulated wind speed was mapped into wind power by means of the characteristic curve of the Enercon E40 600 kW wind turbine (see [4] for details). We set the system base power to 500 kW. In order to simulate a WPP, we eliminated all the loads in the nodes, and to be consistent with the convention in [38], we considered the power *injected* in the nodes as negative. We also set the reactive power injections to zero.

The joint wind speed model consisting of 32 individual processes had 121 active state variables. We performed a 10,000-sample Monte Carlo experiment employing 10-min scale observations—as in the NREL data set—which amounts to approximately a one-week simulation. The approximate running time for the Monte Carlo experiment—including sampling and BFS solutions—was around 10 s in an four-processor Intel(R) Core(TM) i7-3770, 3.4 GHz, PC with 8 GB RAM. This shows that the relative high order of the system does not produce a relevant computational burden, because of the sparsity of the involved matrices.

The difference in the results when considering the generator dependencies is substantial. First, the voltage profile has remarkable discrepancies. Both models produce similar mean voltages (Fig. 9a). However, the use of independent models (Fig. 9a) consistently produces results of narrower interquartile ranges. That is, the dispersion of the data is more reduced. Moreover, the marginal models yield normal distributions. On the whole, the combination of these distribution features makes the aggregation of marginal models fail to reveal the possibility of a flat voltage profile in any

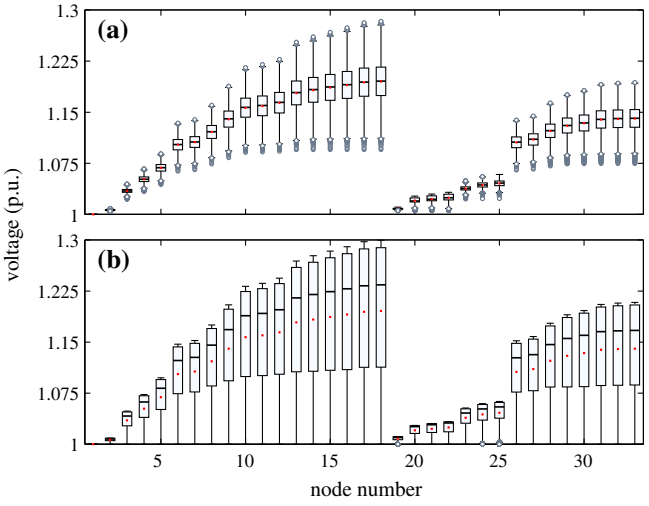


Fig. 9. Tukey's box plots showing the voltage profile. The red dot inside the box plot refers to the distribution mean. (a) Computation from marginal models. (b) Computation from the joint model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

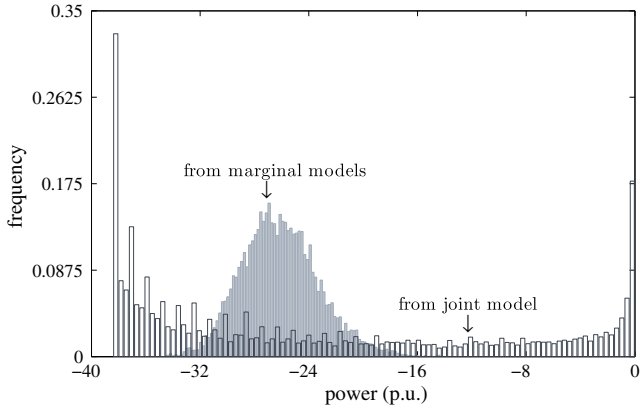


Fig. 10. Injected power through the slack bus into the main grid.

branch of the WPP. This would be a consequence of a possible lack of wind speeds above the cut-in speed in all the generators of the WPP branch simultaneously.

The error in the aggregated production estimate can neither be neglected when the dependence structure is not considered. The maximum aggregated power in per unit (p.u.) system produced by the independent models is 34.6 p.u., whereas the minimal is 17.3 p.u. This is truly inaccurate when compared with the results of the joint model: 38.4 and 0.0 p.u. Moreover, as shown in Fig. 10, the distribution of probabilities is very different. Again the independent models produce a normal distribution, indicating a mean production equal to 25.7 p.u.; which is practically equal to that of the joint model. However, the dependent, joint model shows that the extremes are the most frequent event. Most of the time the WPP will be producing at the rated power because of the prevalent simultaneous wind speeds occurring in the range between the rated and the cut-off speed of the Enercon E40 (with rated power equal to 1.2 p.u.). Second in frequency of occurrence is the null production event. Other frequent events appear in the histogram of Fig. 10 in the form of spikes. These reflect the concurrence of various generators producing at rated when the rest are stopped. These spikes can be seen at $1.2 \times 32 = 38.4$, $1.2 \times 31 = 37.2$, $1.2 \times 32 = 36.0, \ldots$ p.u. On the whole, the independent scenario generation fails to reproduce all these extreme events.

4. Conclusions

This paper provides a comprehensive procedure to represent wind speed in a WPP by means of joint state space models comprising all the wind speed information of the WPP into one only expression. The representation is of interest to researchers and practitioners working on estimation, forecasting, or decision making, concerning the performance of a WPP where an analytical solution is too complex and sampling of multiple possible scenarios is required. Importantly, these scenarios—statistically equivalent to the actual conditions in the WPP—incorporate not only the marginal statistical information of each individual power production, but also the dependence structure among the generators in the WPP. Based on an autoregressive specification, the proposed state space representation allows that dependence be not only cross-sectional (site-dependent), but also longitudinal (time-dependent).

There are several advantages that advocate for the use of state space models to generate wind speed scenarios. The most important, as emphasized by Durbin and Koopman, is the structural specification. State space models give the ability to study the internal structure of wind speed series. In this paper the focus is on the

specification of Box–Jenkins marginal models through joint state space models that include the dependence structure between those marginals. However, more detailed structures may be approached, including trends and seasonalities, to further investigate the structure of wind speed in a WPP. Additionally, state space models provide several advantages over Box–Jenkins models in dealing with missing or irregular-spaced data, or time-varying parameters. Resting on the Kalman filter, state space models provide a convenient way of dealing with such irregularities, as well as they prove useful in facilitating forecasting and smoothing of the underlying processes.

This paper presents a procedure that is simple and—as proved by a number of through several numerical simulations—effective. It comprises a direct transformation into Gaussian variables so that the proposed structured state space model can be handled by means of the optimal Kalman filter. The structure dependence is also easily handled inside the transition equation by defining a multivariate normally distributed innovation; yielding a joint model where wind processes are dependent, in contrast to the individual margin-based model. Through the integration of the wind speed state space model into a backward–forward sweep algorithm, this paper shows that marginal scenario generation fails to reproduce important details of the aggregated power in a WPP. Mean values of power and voltage may be correctly obtained. But extreme values are not correctly shown through those marginal specifications.

References

[1] Conejo AJ, Carrion M, Morales JM. Decision making under uncertainty in electricity markets. New York: Springer; 2010.
[2] Díaz G, Casielles PG, Coto J. Simulation of spatially correlated wind power in small geographic areas – sampling methods and evaluation. Int J Electr Power Energy Syst 2014;63:513–22.
[3] Feijoo A, Cidras J, Dornelas J. Wind speed simulation in wind farms for steady-state security assessment of electrical power systems. IEEE Trans Energy Convers 1999;14:1582–8.
[4] Diaz G, Abd-el Motaleb AM, Mier V. On the capacity factor of distributed wind generation in droop-regulated microgrids. IEEE Trans Power Syst 2012;28. 1–1.
[5] Niknam T, Azizipanah-Abarghooee R, Narimani M. An efficient scenario-based stochastic programming framework for multi-objective optimal micro-grid operation. Appl Energy 2012;99:455–70.
[6] Hagspiel S, Papaemannouil A, Schmid M, Andersson G. Copula-based modeling of stochastic wind power in Europe and implications for the Swiss power grid. Appl Energy 2012;96:33–44.
[7] Morales J, Baringo L, Conejo A, Minguez R. Probabilistic power flow with correlated wind sources. IET Gener Transm Distrib 2010;4:641.
[8] Delgado C, Domínguez-Navarro J. Point estimate method for probabilistic load flow of an unbalanced power distribution system with correlated wind and solar sources. Int J Electr Power Energy Syst 2014;61:267–78.
[9] Billinton R, Chen H, Ghajar R. Time-series models for reliability evaluation of power systems including wind energy. Microelectron Reliab 1996;36:1253–61.
[10] Torres J, García A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Solar Energy 2005;79:65–77.
[11] Box GEP, Jenkins GM. Time series analysis: forecasting and control. rev. ed. San Francisco: Holden-Day; 1976.
[12] Gao Y, Billinton R. Adequacy assessment of generating systems containing wind power considering wind speed correlation; 2009.
[13] Morales J, Mínguez R, Conejo A. A methodology to generate statistically dependent wind speed scenarios. Appl Energy 2010;87:843–55.
[14] Liu P-L, Der Kiureghian A. Multivariate distribution models with prescribed marginals and covariances. Probab Eng Mech 1986;1:105–12.
[15] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405–14.
[16] Correia P, Ferreira de Jesus J. Simulation of correlated wind speed and power variates in wind parks. Electric Power Syst Res 2010;80:592–8.
[17] Hill DC, McMillan D, Bell KRW, Infield D. Application of auto-regressive models to U.K. wind speed data for power system impact studies. IEEE Trans Sust Energy 2012;3:134–41.
[18] Chen K, Yu J. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl Energy 2014;113:690–705.
[19] Durbin J, Koopman SJ. Time series analysis by state space methods. 2nd ed. Oxford University Press; 2012.
[20] Zárate-Miñano R, Anghel M, Milano F. Continuous wind speed models based on stochastic differential equations. Appl Energy 2013;104:42–9.
[21] Usta I, Kantar YM. Analysis of some flexible families of distributions for estimation of wind speed distributions. Appl Energy 2012;89:355–67.
[22] Ouarda T, Charron C, Shin J-Y, Marpu P, Al-Mandoos A, Al-Tamimi M, et al. Probability distributions of wind speed in the UAE. Energy Convers Manage 2015;93:414–34.
[23] Dubey SYD. Normal and Weibull distributions. Naval Res Log Quart 1967;14:69–79.
[24] Chadee XT, Clarke RM. Large-scale wind energy potential of the Caribbean region using near-surface reanalysis data. Renew Sust Energy Rev 2014;30:45–58.
[25] Anderson L, Stephen B, Galloway S, McMillan D, Ault G. Statistical profiling of site wind resource speed and directional characteristics. IET Renew Power Gener 2013;7:583–92.
[26] Papavasiliou A, Oren SS, O'Neill RP. Reserve requirements for wind power integration: a scenario-based stochastic programming framework. IEEE Trans Power Syst 2011;26:2197–206.
[27] Daniel A, Chen A. Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica. Solar Energy 1991;46:1–11.
[28] Nfaoui H, Buret J, Sayigh A. Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco). Solar Energy 1996;56:301–14.
[29] Chatfield C. The analysis of time series: an introduction, 5th ed., vol. 2013. CRC Press; 2013.
[30] Brockwell PJ, Davis RA. Introduction to time series and forecasting. 2nd ed. New York: Springer; 2002.
[31] Qin Z, Li W, Xiong X. Incorporating multiple correlations among wind speeds, photovoltaic powers and bus loads in composite system reliability evaluation. Appl Energy 2013;110:285–94.
[32] Takle E, Brown J. Note on the use of Weibull statistics to characterize wind-speed data. J Appl Meteorol 1978;17:556–9.
[33] Lucia JJ, Schwartz ES. Electricity prices and power derivatives: evidence from the nordic power exchange. Rev Derivatives Res 2002;5:5–50.
[34] Laine M, Latva-Pukkila N, Kyrölä E. Analysing time-varying trends in stratospheric ozone time series using the state space approach. Atmos Chem Phys 2014;14:9707–25.
[35] National Renewable Energy Laboratory. NREL wind dataset. <http://wind.nrel.gov/Web_nrel/> (accessed 23 August 2015).
[36] Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. Ann Stat 2010;38:2916–57.
[37] Lemke W. Term structure modeling and estimation in a state space framework. Springer Science & Business Media; 2006.
[38] Baran M, Wu F. Network reconfiguration in distribution systems for loss reduction and load balancing. IEEE Trans Power Deliv 1989;4:1401–7.
[39] Teng Jen-Hao. A direct approach for distribution system load flow solutions. IEEE Trans Power Deliv 2003;18:882–7.
[40] Diaz G, Gonzalez-Moran C, Gomez-Aleixandre J, Diez A. Complex-valued state matrices for simple representation of large autonomous microgrids supplied by PQ and Vf generation. IEEE Trans Power Syst 2009;24:1720–30.