Probabilistic Modeling of Multi-site Wind Farm Production for Scenario-based Applications

Duong D. Le, *Member, IEEE,* George Gross, *Life Fellow, IEEE,* and Alberto Berizzi, *Member, IEEE*

***Abstract*—The deepening penetration of wind resources introduces major challenges into power system planning and operation activities. This is due to the need to appropriately represent salient features of wind power generation from multiple wind farm sites such as non-stationarity with distinct diurnal and seasonal patterns, spatial and temporal correlations, and non-Gaussianity. Hence, an appropriate model of multi-site wind power production in systems with integrated wind resources represents a major challenge to meet a critical need. In this paper, we aim at defining a new methodology to improve the quality of generated scenarios by means of historical multi-site wind data and effective deployment of time series and principal component techniques.**

Scenario-based methodologies are already available in power systems, but sometimes lack in accuracy: the paper proposes a methodology able to capture the main features of wind: it can both characterize spatio-temporal properties and be used to reduce size of data sets in practical applications without using any simplifying assumption. Extensive testing indicates good performance in effectively capturing the salient wind characteristics to provide useful models for various problems related to multi-site wind production, including security assessment, operational planning, environmental analysis and system planning. An application to security assessment is presented.

***Index Terms*—Correlation, stochastic processes, wind power generation, time series, modeling.**

I. Introduction

THE deepening penetration of wind resources with their various climatological and geographic sources of uncertainty provides new challenges in the planning and operation of such systems. Wind speed is a highly uncertain, time-varying and intermittent phenomenon and, therefore, so is wind generation. The effective representation of wind generation is fraught with major difficulties since there is no analytic characterization for wind speed and the output is a nonlinear function of wind speed. Wind speed is a non-Gaussian and non-stationary stochastic process with distinct diurnal and seasonal patterns; wind speed at a given location is temporally correlated. Moreover, when many wind farms are installed at many sites in the power system, building a multi-site model brings additional complexity. Such model must capture the correlation among wind speeds at different locations as well as their time correlation. In light of these requirements, the modeling must rely heavily on the collection of appropriate data sets, whose analysis provides the basis for the modeling of multi-site wind installations.

The early contribution in wind modeling area was presented in [1], in which auto-correlation in wind speed is characterized but its non-Gaussianity is not considered. To deal with non-Gaussianity and non-stationarity, the authors in [2], [3] apply transformation and standardization to hourly wind speed time series to obtain approximate Gaussian and stationary data: while Auto-Regressive (AR) model is used for fitting the resulting data in [2], the authors in [3] adopt a more general model, i.e., Auto-Regressive Moving Average (ARMA) model. These schemes are suitable for a single time series data. To take into account spatial correlation of wind speed between different zones in UK, a multivariate AR model is used in [4]. However, building a multivariate time series model for real wind speed data from multiple sites is complicated, especially with a large number of wind sites. In a different way, reference [5] proposes a wind regime model to capture both the seasonal and the diurnal variations of wind resources and their correlation with the load seasonal and diurnal changes but this model is only suitable for planning studies.

To represent a stochastic process, a set of scenarios, constructed as a set of realizations – so-called sample paths or trajectories – over the predefined time horizon, can be considered: in [6], wind power scenarios are generated from non-parametric probabilistic forecasts, while in [7] time series analysis and Monte Carlo Simulation (MCS) are used to generate wind power scenarios. Reference [8] presents a method to characterize forecast error via empirical distributions of a number of forecast bins and based on statistical uncertainty and variability to generate a large number of wind power scenarios. These approaches have been applied to a single wind farm or an aggregate wind power data set and some of them may be extended to apply to a multi-site wind data set for capturing spatial correlation.

For generating space-time wind scenarios, in order to characterize interdependence structure of multivariate stochastic processes, Gaussian copula method [9] is widely used. Among existing approaches, references [10] and [11] work on a similar topic. In [10], the authors build the model for multi-site wind speed using a noise vector that drives a vector AR process. In order to deal with non-Gaussianity and preserve the marginal distribution associated with observed data at each site in the wind speed scenarios generated, we transform the time series of historical values of each site into a Gaussian time series, similar to that in [10] and [11]. For capturing

D. D. Le is with the Department of Energy, Politecnico di Milano, Milan, 20156 Italy, and also with the Department of Electrical Engineering, Danang University of Science and Technology, Danang, Vietnam (e-mail: dinhduong.le@polimi.it).

G. Gross is with the Electrical and Computer Engineering Department, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: gross@illinois.edu).

A. Berizzi is with the Department of Energy, Politecnico di Milano, Milan, 20156 Italy (e-mail: alberto.berizzi@polimi.it).

temporal correlation and variability, we make use of time series model as in [10], [11]. Nevertheless, to deal with spatial correlation of wind speed at different sites, [10] and [11] simplify the problem by introducing an assumption that the matrix of time series coefficients is diagonal: this implies that spatial correlation is modeled fully by the underlying noise vector. By doing so, the multivariate time series model can be decoupled into different univariate time series models. On the contrary, in this paper we do not use any simplifying assumption. Instead, we solve it explicitly by making use of Principal Component Analysis (PCA) to transform correlated multivariate time series of wind data at multiple sites into different univariate time series, i.e., Principal Component (PC) time series, which are not cross-correlated with other time series. In terms of handling non-stationarity, while [10] uses stationary assumption or suggests using seasonal ARMA model, we adopt pre-processing techniques, similarly to [11].

A comprehensive modeling methodology of multi-site wind generation that captures all its salient features is crucial for power system planning and operation studies. To account for spatial-temporal information, full knowledge of distribution of multivariate stochastic processes of wind, e.g., joint probability density functions ($p.d.f.$s) and cumulative distribution functions ($c.d.f.$s), is necessary. It is, however, difficult to use such functions in power system planning and operation; to this issue, a spatial-temporal scenario set is an alternative and efficient way to characterize multivariate stochastic processes. In addition, the model needs to be realistic, i.e., simplifying assumptions, if necessary, must be reasonable.

In this paper, a new framework of multi-site wind modeling is proposed. Starting from time series relevant to wind speed (or wind power) data coming from multiple different sites, we first build a model capturing the salient features of wind and use this model to generate a set of accurate scenarios by using PCA [12], [13] and time series analysis [14]. Each scenario is a set of generated time series of wind speed (or power) reproducing the time/space features of the input data used. Scenarios can be used for any scenario-based power system application, as it will be discussed in the paper. The resulting wind speed scenarios for each wind site are then transformed into wind power scenarios via a suitable aggregate power curve. The PCA-time series combination is effective in obtaining the analytical characterization of the statistical features of the spatio-temporal model of the wind output at the multi-site farms. The proposed model is able to reduce the size, without losing significant information, of a data set; this is very useful in cases of high-dimensional data, such as the wind data from a large number of wind farm locations. Moreover, for properly working of PCA and time series model, we propose some techniques to obtain approximately stationary and Gaussian data from observed wind data (that are typically non-stationary and non-Gaussian), thus removing any limiting simplifying assumption. The proposed methodology is, therefore, comprehensive and realistic so that it is applicable to wind data in real power systems. The model results provide a wide range of applications in power systems with integrated multi-site wind energy resources, including security assessment, operational planning, planning and environmental analysis. These applications provide valuable insights into the impacts of the wind contributions.

In Section II, we present the fundamental background of PCA and time series analysis. The proposed methodology is described in Section III, while in Section IV the results obtained on wind data from multiple wind farms in Sicily and on security assessment of Sicilian power system are discussed; discussions on assessing the resulting scenarios are also given. In Section V, further discussion on applicability of scenario-based methods is presented. Concluding remarks are provided in Section VI.

II. Fundamental background

A. Principal Component Analysis

PCA performs an orthogonal transformation on data in order to transform a correlated data set into an uncorrelated one. The underlying technique is the eigenanalysis, applied to a symmetrical matrix such as either the correlation or the covariance matrix.

PCA applies to a matrix **W** whose elements w_s^h are the data of wind speed or power available at site $s \in \{1, 2, ..., S\}$ and at time $h \in \{1, 2, ..., N\}$. We assume that all considered wind sites have the same number of observations N spanning, for example, one year, and that they are synchronized and equally spaced in time (e.g., one hour).

$$\mathbf{W} = \begin{bmatrix} w_1^1 & w_1^2 & \cdots & w_1^N \\ w_2^1 & w_2^2 & \cdots & w_2^N \\ \vdots & \vdots & \ddots & \vdots \\ w_S^1 & w_S^2 & \cdots & w_S^N \end{bmatrix} \tag{1}$$

Each element w_s^h can be interpreted as a realization of corresponding random variable $\widetilde{W}_s^h$ of the random process at site s and time h.

At first, data are centered [13] in matrix $\boldsymbol{W_c}$ by subtracting the mean μ_s of each time series at each site:

$$\mu_s = \frac{1}{N}\sum_{h=1}^{N} w_s^h \tag{2}$$

$$\boldsymbol{W_c} = \boldsymbol{W} - \boldsymbol{\mu} \tag{3}$$

where $\boldsymbol{\mu} = diag\{\mu_1, \mu_2, \cdots, \mu_S\}\boldsymbol{J}$ in which $\boldsymbol{J}$ is a $S \times N$ matrix of ones.

Next, correlation or covariance matrix of the centered data is calculated. PCA can use either correlation or covariance matrix. Correlation matrix must be adopted when the considered variables are not comparable [13], for example when considering the real power outputs of wind parks of different rating. In this case, another option is to normalize values and adopt normalized covariance matrix. On the contrary, covariance matrix can be directly used for wind speed data. In the following, the covariance matrix is considered, for the sake of simplicity.

Covariance matrix $\boldsymbol{\Sigma}$ is a symmetric $S \times S$ matrix:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,S}^2 \\ \sigma_{2,1}^2 & \sigma_2^2 & \cdots & \sigma_{2,S}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{S,1}^2 & \sigma_{S,2}^2 & \cdots & \sigma_S^2 \end{bmatrix} \tag{4}$$

where, σ_i^2 is the variance (σ: standard deviation) and $\sigma_{i,j}^2$ is the covariance between the time series at site *i* and the time series at site *j*:

$$\sigma_{i,j}^2 = \frac{1}{N}\sum_{h=1}^{N}(w_i^h - \mu_i)(w_j^h - \mu_j) \tag{5}$$

The covariance matrix $\boldsymbol{\Sigma}$ is a symmetric positive semi-definite matrix and all its eigenvalues $\lambda_i, \ i = 1, 2, ..., S$, are positive and are the roots of (6):

$$\det(\boldsymbol{\Sigma} - \lambda_i \boldsymbol{I}) = 0 \tag{6}$$

where, $\boldsymbol{I}$ is the $S \times S$ identity matrix, $\det(\cdot)$ is the determinant.

Eigenvalues are then ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_S$. There exists a vector $\mathbf{u}_i$ corresponding to λ_i such that:

$$\boldsymbol{\Sigma}\mathbf{u}_i = \lambda_i \mathbf{u}_i \tag{7}$$

Vector $\mathbf{u}_i$ is called the eigenvector of $\boldsymbol{\Sigma}$ associated with the eigenvalue λ_i. Matrix $\mathbf{U}$ is formed by the corresponding columns $\mathbf{u}_i, \ i = 1, 2, ..., S$:

$$\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & | & \mathbf{u}_2 & | & \cdots & | & \mathbf{u}_S\end{bmatrix} \tag{8}$$

Its elements are also known as *PC coefficients*. Finally, PCs are derived [13] as

$$\mathbf{Z} = \mathbf{U}^T \boldsymbol{W_c} \tag{9}$$

where $\mathbf{Z}$ is a $S \times N$ matrix. The i-th row of matrix $\mathbf{Z}$, $\mathbf{z}_i$, is the i-th PC, that is, a time series univariate and uncorrelated with other PCs. The techniques to characterize such a time series are reported in [14].

The reconstruction of wind data from PCs is implemented inversely:

$$\begin{aligned}\mathbf{W} = \boldsymbol{\mu} + \mathbf{UZ} = & \boldsymbol{\mu} + \begin{bmatrix}\mathbf{u}_1 & | & \mathbf{u}_2 & | & \cdots & | & \mathbf{u}_S\end{bmatrix} \\ & . \begin{bmatrix}\mathbf{z}_1^T & | & \mathbf{z}_2^T & | & \cdots & | & \mathbf{z}_S^T\end{bmatrix}^T\end{aligned} \tag{10}$$

It is worth noting that if the distribution considered is multivariate Gaussian, resulting PCs will be independent. Otherwise, PCs will be uncorrelated but still dependent (the diagonal covariance matrix of PCs only implies that they are uncorrelated). In fact, for multivariate non-Gaussian distribution, the first and second statistical moments do not characterize totally the distribution. In the present paper, we adopt pre-processing and transformation techniques to obtain approximately stationary and Gaussian data sets (see Section III). This is a significant improvement in using PCA as confirmed by the results in Section IV.

Another very interesting application of PCA is that it provides an excellent tool to approximate a large data set by reducing its dimension [12]. This function makes PCA a powerful tool for high-dimensional data analysis. As it derives from eigenanalysis, each PC $\mathbf{z}_l$ can be seen as a mode, whose variance is weighted by the relevant eigenvalue λ_l. It should be noted that the variance of each PC time series is equal to the eigenvalue associated with that PC. Therefore, the contribution of the l-th PC to total variance of the data [12] can be computed as:

$$\gamma_l = \frac{\lambda_l}{\sum\limits_{i=1}^{S} \lambda_i} \times 100\% \tag{11}$$

and the cumulative contribution of the first l PCs is:

$$\Gamma_l = \sum_{i=1}^{l} \gamma_i \tag{12}$$

Hence, the first row vector $\mathbf{z}_1$ corresponding to the largest eigenvalue λ_1 and eigenvector $\mathbf{u}_1$ is the most important component (dominant component) which contains most of the variance in the data set, followed by the second component $\mathbf{z}_2$, and so on. If only the first K ($K < S$) components are considered, $\mathbf{W}$ will be approximated by

$$\begin{aligned}\hat{\mathbf{W}}_K = \boldsymbol{\mu} + & \begin{bmatrix}\mathbf{u}_1 & | & \mathbf{u}_2 & | & \cdots & | & \mathbf{u}_K\end{bmatrix} \\ & . \begin{bmatrix}\mathbf{z}_1^T & | & \mathbf{z}_2^T & | & \cdots & | & \mathbf{z}_K^T\end{bmatrix}^T\end{aligned} \tag{13}$$

The choice of the most suitable K for application of PCA to dimensional approximation is dependent on the comparison of Γ_K to a threshold (e.g., $\Gamma_K \geq 90\%$). This is also a very useful application of PCA, from the practical point of view: it makes it possible to describe most of the features of the data set by a reduced number of variables.

B. Time series analysis

A time series is a sequence of observations ordered in time, usually at equally-spaced intervals. There are several methods for fitting to a time series, if it is stationary. For a stationary stochastic process, the joint probability distribution, and therefore the mean, variance, and autocorrelation structure, etc., do not change over time. The typical linear model for a stationary time series is ARMA [14], which can be used to characterize a stationary process and for prediction as well. The ARMA model consists of two parts: AR and MA. The model is usually referred to as the ARMA(p, q) model where p is the order of the AR part and q is the order of the MA part. An ARMA(p, q) model of a stochastic process can be mathematically represented as:

$$w_s^h = \sum_{j=1}^{p} \alpha_j w_s^{h-j} + \varepsilon_s^h - \sum_{l=1}^{q} \beta_l \varepsilon_s^{h-l} \tag{14}$$

where, $\alpha_1, \alpha_2, ..., \alpha_p$ and $\beta_1, \beta_2, ..., \beta_q$ are the parameters of AR and MA, respectively. The stochastic process $\{\varepsilon_s^h\}$ is referred to as a white noise [14].

If q = 0, then the ARMA(p, q) model becomes an AR(p) model. On the other hand, when p = 0, the process becomes a MA(q) model. An AR model expresses a time series as a linear combination of its past values. The order of p tells how many lagged past values are included in the model. The MA model includes lagged terms on the noise process.

To build a time series model, we follow the procedure proposed by Box-Jenkins, clearly described in [14].

It should be noted that stationarity is a necessary condition in building an ARMA model. However, this condition may not always hold with real time series data. In such a case, data must be pre-processed before building an ARMA model. In this paper, we carry out various pre-processing and transformation techniques, presented in the next sections, and apply them to wind speed time series data.

III. Methodology

In this section, we discuss in detail the proposed approach to capture main characteristics of wind data from multiple sites and to build a spatio-temporal model.

The input is observed wind speed or wind power data (in time series) for each site, in the form (1). The process is implemented step by step as follows:

Step 1: The first requirement to be fulfilled is stationarity of the process described by (1). This is achieved first by removing diurnal and seasonal effects [2], [3], [11]:

$$w'^{h}_{s} = (w^{h}_{s} - \mu^{h,m}_{s})/\sigma^{h,m}_{s} \tag{15}$$

where $\mu^{h,m}_{s}$ and $\sigma^{h,m}_{s}$ are the mean and standard deviation at site s and time h for epoch m such as month, season, etc., which is selected based on the periodic features of the data. The resulting stationarity must be assessed by a statistical test on $\{w'^{h}_{s}\}$. In this paper, we used Augmented Dickey-Fuller (ADF) test [15] and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test [16]. If the modified $\{w'^{h}_{s}\}$ does not pass the test, the pre-processing adopted is not sufficient and a different pre-processing must be carried out, still based on (15). This formula depends on the proper identification of epochs, each one identified by m. Therefore, if the first application of (15) is not satisfying, it is necessary to revise the partitioning of the considered time interval in epochs, always using (15), until a suitable partition is found out, stationarity is obtained and the test is passed. If it is impossible to find epochs so as to obtain stationarity and pass the test, the only option is to check if there is any trend in data and remove it [17]. In our case, that has been proven not to be necessary.

Step 2: As the obtained stationary data set could still be non-Gaussian, $\{w'^{h}_{s}\}$ is then transformed into Gaussian data set [10], [11] by:

$$w''^{h}_{s} = \Phi^{-1}[\hat{F}_s(w'^{h}_{s})] \tag{16}$$

where $\hat{F}_s(\cdot)$ is the estimated $c.d.f.$ of the stationary process associated with $\{w'^{h}_{s}\}$ and $\Phi^{-1}(\cdot)$ is the inverse of the $c.d.f.$ of the standard normal distribution. Statistical tests to assess if the resulting distribution is Gaussian, such as Lilliefors goodness-of-fit test [18], Jarque-Bera hypothesis test [15], etc., are carried out.

Step 3: PCA is adopted according to Section II-A. Each PC is a univariate time series and not cross-correlated with other PCs. The data, however, still contain temporally correlated information.

Step 4: Each PC time series is fitted by a time series model according to Section II-B, as PC time series data herein satisfy all necessary conditions for building a time series model like ARMA or simpler forms such as AR or MA.

Step 5: The obtained time series model for each PC is then used to generate an adequate number of time series for future time; for example, time frames of operation (e.g., 6 hours, 24 hours ahead, etc.) and/or planning (e.g., weeks, months, years ahead, etc.). Thanks to the use of ARMA methodology, variability of input data is implicitly considered.

Step 6: The generated time series in terms of PCs are reconstructed by using (10). If dimensional approximation is desired, (13) can be used.

Step 7: The obtained data from *Step 6* are back-transformed into non-Gaussian data [10], [11] by:

$$w'^{h}_{s} = \hat{F}_s^{-1}[\Phi(w''^{h}_{s})] \tag{17}$$

and then the items removed in the pre-processing step are added back to obtain scenarios obeying all the characteristics of the observed wind data for each site.

The outputs of the procedure are time series, i.e., scenarios or trajectory sets, of wind data over the predefined time horizon for each site. The novelty of the proposed approach is that it can explicitly capture the main features of stochastic processes of multi-site wind data: marginal distribution, spatial correlation, temporal correlation, diurnal and seasonal non-stationarity and non-Gaussianity.

As discussed in [8], in operation, it is important to deal with both uncertainty and variability, i.e., forecast errors and fluctuations. In the proposed methodology, the focus is on uncertainty; however, variability is taken into account implicitly by the methodology thanks to the use of time series methods to generate PC time series (*Step 4* and *Step 5*). This approach is already present in the technical literature. For example, in order to capture variability of wind and also of other resources in power system analysis and security assessment, time series-based methods have been adopted: in [19], hourly time series data of wind and demand are used to determine overload conditions or to specify non-firm connection agreements for new generators; in [20], the authors use time series data of load and variable resources such as solar photovoltaic and gas-fired micro-CHP to quantify the technical impact of high penetration of such resources on the operation of distribution systems; a development of time series power flow-based analysis to assess the impact of wind generation on the voltage stability of power systems is presented in [21]. While these studies use historical time series data in the analysis, we use the proposed methodology to characterize variability of wind to provide wind power time series (i.e., scenarios or trajectory sets) as input for security assessment (Section IV-C).

The above procedure can be applied to wind speed as well as to wind power data. However, often wind power data are neither available nor reliable. When wind power data are not available, the only chance is to model wind speed and then to derive wind power data. Moreover, it would be very difficult to model non-stationarity and non-Gaussianity of wind power and, in the end, to use time series and PCA. In case of wind speed data, an aggregate power curve for each entire wind site is needed for mapping wind speed scenarios into wind power scenarios. In this paper, we make use of the method of bins

[22]. To estimate power curve for a site, measurement data of wind power-wind speed pairs of the site are used. Before adopting the method, some techniques are applied to reject erroneous data to improve the estimation of power curve.

Resulting scenarios from the proposed model and the estimated power curve for each site are assessed and discussed in detail in the next section.

IV. Results

In this section, we apply the proposed multi-site wind model to observed wind speed from different sites in Sicily, Italy. As a possible application, we present its exploitation for the security assessment of the Sicilian power system for highlighting the attractive features of the proposed approach.

Further discussions on the wide range of applications of the multi-site wind modeling results are given in the next section.

A. Wind sites in Sicily

Sicily is the largest Italian island. We use hourly wind speed and wind power data from September 1, 2011 to August 31, 2012 measured at ten sites in Sicily: wind speed data are used for multi-site wind speed modeling, while wind power-wind speed pairs are used for estimating power injections. The resulting wind power scenarios are then used as input for assessing security of Sicilian power system, including MV, HV, and EHV levels; it consists of 539 buses, 664 branches, and 261 generating units.

For the sake of simplicity, the sites are denoted as S_1, S_2, ...,S_{10}. Correlation coefficients calculated from wind speed at ten sites range from 0.21 to 0.75, indicating that they are more or less correlated, depending on their geographical features, e.g., their positions, distances, etc.

B. Multi-site wind modeling

1) Wind speed modeling: Observed wind speed at ten wind farm sites in Sicily (see Fig. 1) is used as input for wind speed modeling.

As discussed, PCA works properly when the data used are Gaussian; furthermore, a time series model like ARMA requires stationary data. To this goal, the data were partitioned by month, i.e., m denotes month in (15), and the initial wind speed was pre-processed according to *Step 1*. The resulting data passed the stationarity tests; eventually, transformation (16) was applied and an approximately stationary and Gaussian set was obtained and passed the relevant tests. The $c.d.f.$s of the stochastic processes associated with the data before (i.e., non-Gaussian) and after (i.e., Gaussian) using (16), for example, for location S_1 are depicted in Fig. 2.

It is worth noticing that, so far, we have used techniques in statistics and obtained the data set associated with a stationary Gaussian process for each site without any assumptions. This is a particularly attractive feature of the proposed methodology: the method can be used for real wind data in power systems.

The refined data from all sites are then transformed into PCs by using (9). All eigenvalues are sorted in descending order

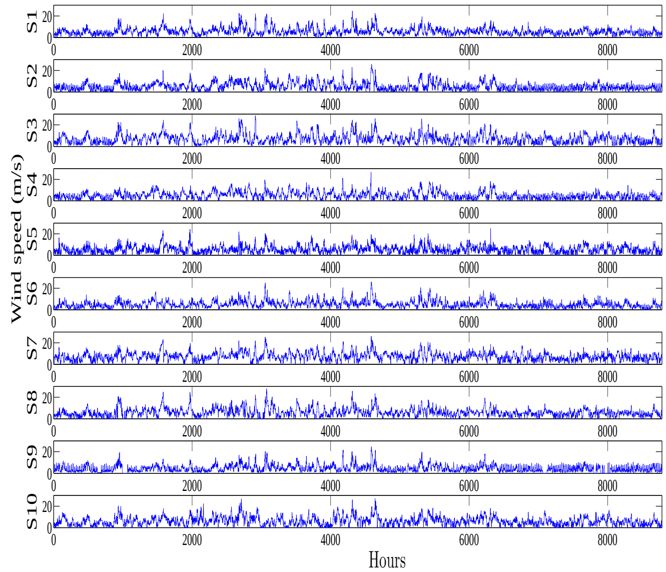

Fig. 1: Observed wind speed at ten wind sites in Sicily

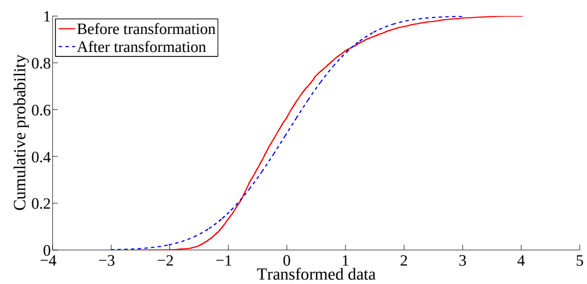

Fig. 2: $c.d.f.$s before and after conversion to stationary and Gaussian for location S_1

TABLE I: Contribution of PCs

PC - l	1	2	3	4	5	6	7	8	9	10
$\gamma_l(\%)$	54.15	16.28	8.69	5.19	4.30	3.36	2.44	2.27	1.89	1.43
$\Gamma_l(\%)$	54.15	70.44	79.12	84.31	88.60	91.97	94.41	96.68	98.57	100.00

and plotted in Fig. 3. The plots of the time series relevant to each PC are shown in Fig. 4. The contribution of each PC and the cumulative contribution of the first PCs are calculated by (11) and (12), respectively, and presented in Table I. As can be seen from Fig. 4, PCs are quite different in terms of magnitudes. The variance of each PC time series is equal to the eigenvalue associated with its PC. The first PC ($\mathbf{z}_1$) contains the largest percentage of variance in the data set (54.15%); the second PC ($\mathbf{z}_2$) the second largest percentage (16.28%) and so on. The first few PCs cover large amount of variance and can be used as an approximation. This is also a very important aspect of the proposed model so as to make it a very effective tool to reduce the size of high-dimensional data set.

After applying PCA, the obtained PCs are uncorrelated. In this case, PCs are also independent because transformation (16) was adopted. If the proper transformation had not been applied (i.e., data were still non-Gaussian), dependency would still have existed between the PCs. Independence can also be deduced from Fig. 5 which depicts the scatter plot of $\mathbf{z}_1$ and

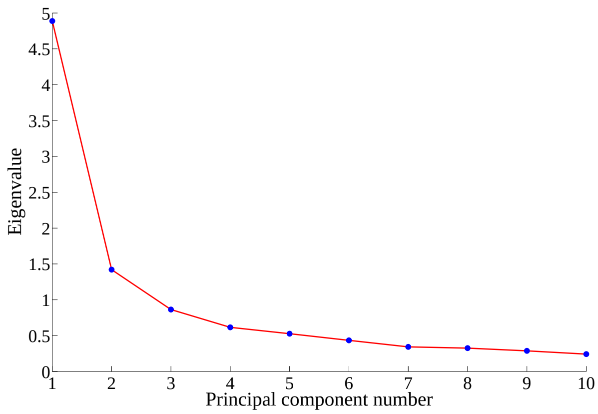

Fig. 3: Ordered eigenvalues associated with PCs

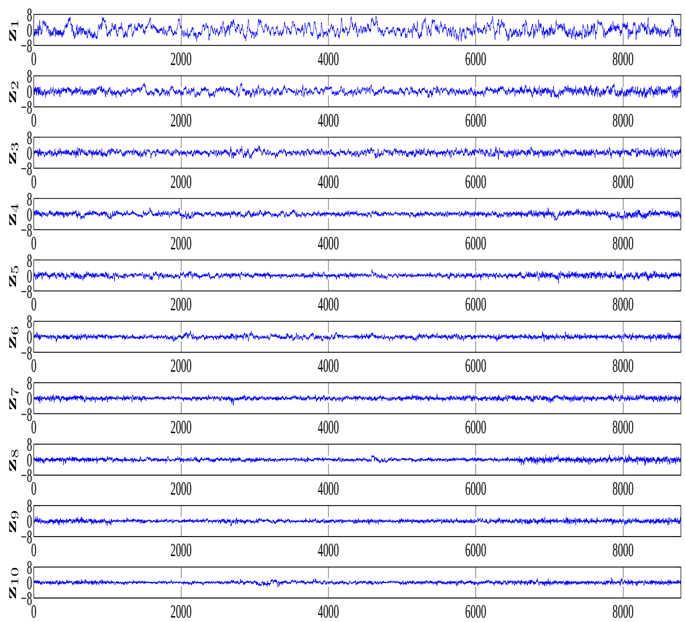

Fig. 4: Features of PCs

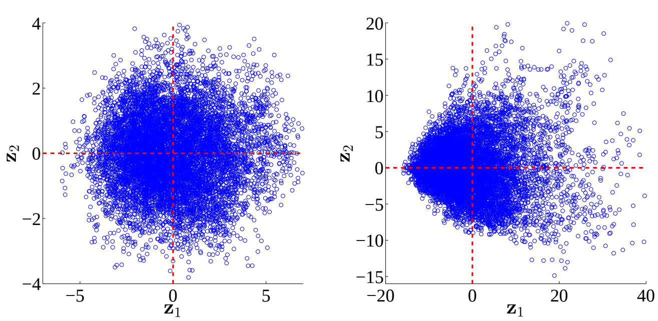

Fig. 5: Scatter plot of $\mathbf{z}_1$ and $\mathbf{z}_2$

$\mathbf{z}_2$ for the case of applying transformation (16) on the left of the figure and the case without using transformation on the right.

The time series built on the PCs fulfill the assumptions for a successful use of stationary time series models. We follow the procedure proposed in [14], and obtain a time

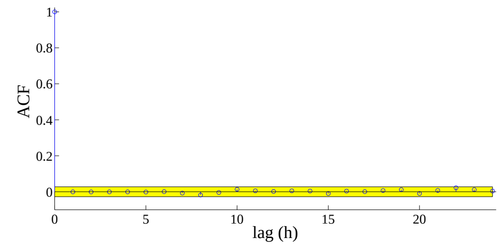

Fig. 6: Residual test for time series model of $\mathbf{z}_1$

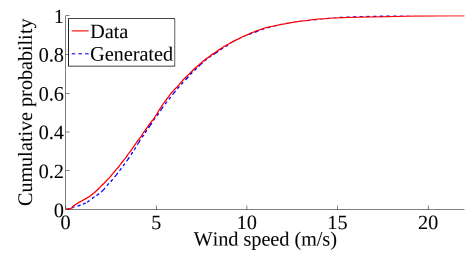

Fig. 7: $c.d.f.$s of observed wind speed data and generated data for site S_5

series model for each PC time series, for example, $\mathbf{z}_1$: AR(4) with corresponding coefficients [1.144;-0.164;0.012;-0.031] and $\mathbf{z}_2$: AR(7) with corresponding coefficients [0.983;-0.131;0.021;-0.008;0.008;0.003;0.036]. The resulting models are then assessed by the residual (error) test. As an example, Fig. 6 illustrates the resulting residual test for $\mathbf{z}_1$ model (Auto-Correlation Function – ACF is plotted). As the residual is a white noise, the AR(4) model of $\mathbf{z}_1$ is valid. The same process was carried out for the other PCs.

The models of PCs are then used to generate a large number of time series for future time instants. After that, the inverse process is used to obtain wind speed scenarios which are the output of the wind speed modeling.

2) Assessing the quality of wind speed scenarios: The main goal of the proposed wind modeling is to explicitly capture main statistical features of wind speed stochastic processes at multiple sites, i.e., marginal distribution, spatial correlation, temporal correlation, diurnal and seasonal non-stationarity and non-Gaussianity, and to generate scenarios retaining these features. In order to properly assess statistical properties of the generated processes in comparison to properties explored from the observed wind speed data, time span should be sufficiently long. For this purpose, we generate scenarios spanning 8000 hours ahead.

Figure 7 compares the $c.d.f.$ relevant to the input data at site S_5 and the $c.d.f.$ computed on one of the generated scenarios at the same site: it is clear that the methodology preserves the marginal distribution of the observed data.

The non-stationarity of the observed data is explored and dealt with by pre-processing techniques (*Step 1*) and preserved by the inverse process in *Step 7*. Similarly, the non-Gaussianity is treated by transformation (16) and back-transformation (17). Temporal correlation existing in the observed data is captured by time series models (*Step 4*), which are validated by the residual test, illustrated in Fig. 6. When generating scenarios for multivariate stochastic processes, consistency between

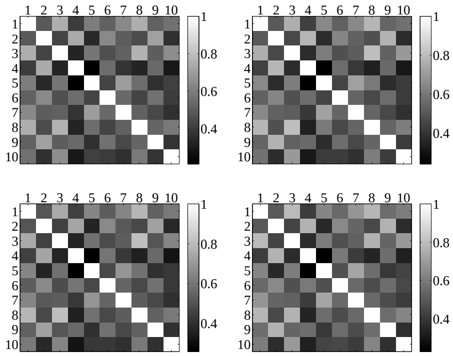

Fig. 8: Visualization of cross-correlation matrices of scenarios and observed data

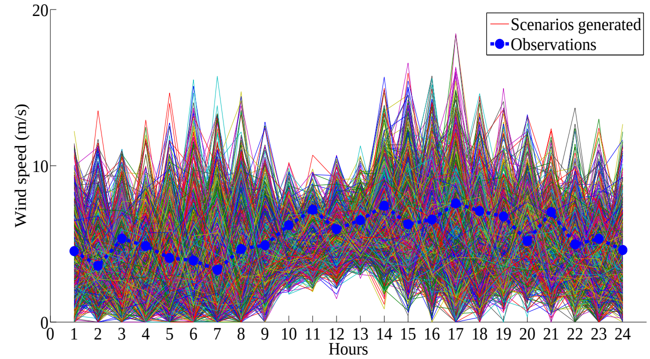

Fig. 9: Wind speed scenarios and observations at site S_5

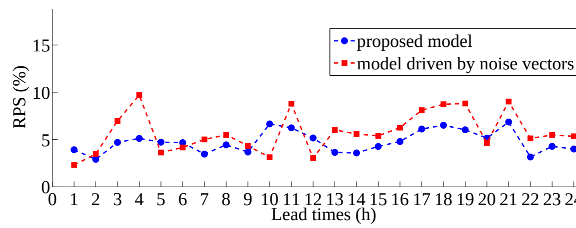

Fig. 10: RPS computation for site S_5

processes should be considered, i.e., cross-correlation between processes (here, spatial correlation between wind speed at different sites) should be maintained. In this modeling, PCA ensures cross-correlation: it is captured by PCA (*Step 3*) and the correlation structure is reproduced by reconstruction in (10). It should be noted that when generating one scenario, we actually generate (one for each site) ten parallel time series and we can generate as many scenarios as desired. For instance, Fig. 8 shows the visualization of cross-correlation matrices of three randomly picked scenarios, compared to the one observed (upper-left sub-figure): they are very similar, showing that spatial correlation is retained.

For probability and ensemble forecasts, the obtained results can be evaluated by adopting some scoring criteria such as Brier Score (BS), Ranked Probability Score (RPS) [23], [24], and so on. While traditional assessment tools [23], [24] have their own merits, reference [25] developed an event-based verification framework to assess a set of scenarios generated that is expected to capture the probability of a certain event. In this paper, wind speed forecasts are represented by a set of discrete scenarios which can be treated as multi-categorical forecasts by partitioning the range of values into exclusive intervals (bins), then event-based verification approach can be adopted. It should be noted that the verification tool chosen should account for the ordering of categories [23]. While BS is widely used for binary events, RPS is an extension of BS to ordinal multi-categorical forecasts [23]. Therefore, RPS is suitable for evaluating wind speed scenarios generated in this paper. RPS is negatively oriented and its values range from 0 to 1: it assigns lower values to better forecasts and a score of 0 indicates that the forecast is perfect.

Assume that the range of wind speed scenario forecasts is divided into different categories by using C thresholds $\zeta_1 < \zeta_2 < \cdots < \zeta_C$. The events A_c $(c = 1, 2, ..., C)$ are defined for categories c as: $A_c = \{\widetilde{W}^h \leq \zeta_c\}$, where $\widetilde{W}^h$ is wind speed random variable at time h. The RPS can be made horizon-dependent as a function of the lead time h [23]:

$$RPS^h = \frac{1}{C}\sum_{c=1}^{C}(F^h_{f,c} - F^h_{o,c})^2 \qquad (18)$$

where, $F^h_{f,c}$ and $F^h_{o,c}$ are $c.d.f.$s of scenario forecasts and observation belonging to bin c at time h, respectively.

Figure 9 depicts 10000 wind speed scenarios generated by the proposed model and observations, for instance, at site S_5 for 24 hours ahead.

Figure 10 shows RPS values computed by using (18) with splitting the range of possible forecasts into 10 bins: they are small (about 5%), indicating that the model proposed gives a good performance for the considered horizon.

In particular, in order to provide a comparison with other methods to generate scenarios, in Fig. 10 the results obtained by the method proposed are compared to the results obtained by the model driven by noise vectors, introduced in [10] and improved in [11]: RPS values of the proposed model are generally lower than RPS values corresponding to the noise vector model, showing quality improvement of scenarios generated by the proposed model.

3) From wind speed to wind power: For mapping wind speed scenarios into wind power scenarios, we make use of the method of bins [22], [26] to estimate an aggregate power curve for each site. Due to erroneous values existing in the measurement data of wind power-wind speed pairs, data should be filtered. Some criteria are proposed to eliminate spurious data points; data points falling into the following cases (i.e., data not representing the normal operating conditions or caused by wrong measurement and other effects) must be neglected: data points that do not match the number of turbines available for generation (e.g., power output measured is greater than $\sum_{i=1}^{n_t} P_{r,i}$, where $P_{r,i}$ is the rated power output of turbine i and n_t is the total number of turbines available); data points at wind speed higher than the cut-out speed with the corresponding power outputs different from zero; data points corresponding to very low wind speed (with respect to cut-in speed) and non-zero power output; data points corresponding

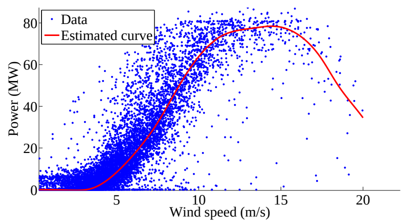

Fig. 11: Estimated power curve for site S_6

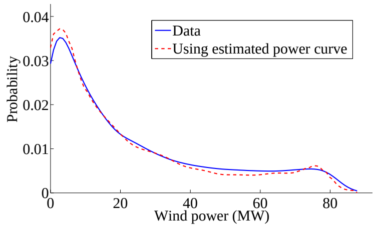

Fig. 12: Comparison between $p.d.f.$s of observed wind power and wind power obtained by using estimated power curve for site S_6

to zero power output and wind speed within normal operation region; data points with constant wind speed over a too long period (e.g., two hours), etc. The remaining pairs are then used for the method of bins, in which wind speed is divided into bins for the range from 0 m/s to cut-out speed (it should be noted that when all turbines in a site are not identical, the maximum cut-out speed is used) and for each bin, the average wind speed and wind power are computed and used as a point for estimation. In this paper, the span of each bin is 1 m/s. Figure 11 depicts power curve estimated for site S_6 (consisting of 113 turbines, each rated 850 kW), while Fig. 12 compares $p.d.f.$s of observed wind power data and the one of wind power obtained by mapping from observed wind speed at the same site via the estimated curve: it shows a good performance of the estimated curve with a small deviation.

C. Application to security assessment

In the following, the use of the model proposed applied to security assessment of Sicilian power system is shown as an example. We adopt MCS with 10000 samples to carry out power flow computation for a time frame up to 24 hours ahead. This test aims at showing as a possible application the usefulness of the proposed model results in security assessment [27], especially highlighting attractive features of the model. About the sources uncertainty, the following holds: for loads, their uncertainty distributions are assumed to be stationary and normally distributed, with expected values equal to their base case data and standard deviations equal to 9% of the expected values; random outages of 170 lines are also considered with the probability of failure equal to 0.1%. Distributed slack bus formulation [28] is exploited so as to possibly include the steady-state behaviour of the frequency regulation of conventional generation in the calculation.

First, dimensional reduction is considered in the proposed application. Figure 13 shows $p.d.f.$ of current, for example, between buses 176 and 227 at time 8h ahead for different

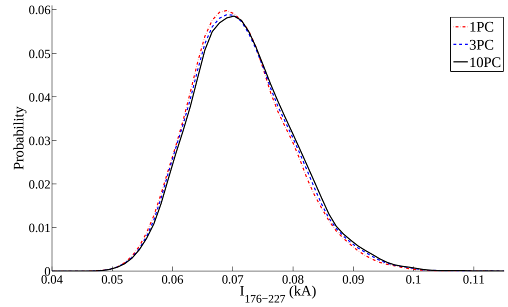

Fig. 13: $p.d.f.$ of current between buses 176 and 227 at 8h

number of PCs used. When all PCs are used, total variance in wind resources is taken into account; otherwise, part of the variance is neglected due to using first PCs to approximate the dimension of wind data. The figure indicates that, in spite of using the first few PCs, the curve is very close to the case of all PCs used. If the first three PCs are used (covering 79.12% of the total variance), one dimension of wind data set, i.e., space, (in total of two dimensions, i.e., time and space) will be reduced from ten to three; in this case, computation time for obtaining wind speed scenarios decreases from 46.1 to 36.5 s: this means that the information lost is negligible, while the gain in computation time is 21%. This is a very attractive feature of the proposed method to deal with high-dimensional data. From $p.d.f.$s of currents as well as voltages, we can evaluate probability of line overloading and probability of over-/under-voltage [28]; however, in this test, there is no violation of voltages and flows.

Second, the proposed modeling explicitly captures both temporal and spatial correlations in wind resources and provides valuable results, especially for applications where correlation information is not negligible. It is worth noticing that MCS based Probabilistic Power Flow (PPF) [28] using samples obtained by sampling non-temporal probability distributions of input random variables at different time-steps (e.g., wind power distributions provided by a probabilistic forecast technique [29]) can provide output (voltages, currents, power flows) in terms of probability distributions at each time-step which may be sufficient for assessing probability of violation for these quantities.

When security issues relevant to variability of wind are considered, the ramping capability of generators is involved and temporal correlation can not be neglected. This is particularly important in some electricity markets where links among different market periods are considered. Conventional generators connected to 18 buses are distributed slack, so that any mismatch, and/or any uncertainty, in the system is shared by the relevant generators with corresponding participation factors. Conventional generator at bus 468 (g_{468}), for instance, is assigned in real power allocation process with the participation factor 0.15 [28], resulting in its power output $p_{g_{468}}^{h,\xi}$ at time h for wind output scenario ξ ($\xi = 1, 2, ..., \Omega$, where Ω is total number of generated wind scenarios) as in Fig. 14. To evaluate the probability of rampability violation from time h to time $h+1$, its ramping r must be calculated

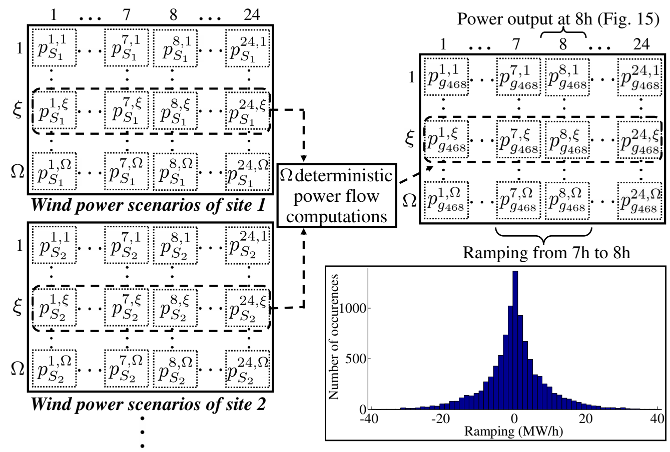

Fig. 14: Illustration of accounting for spatial and temporal correlations of wind resources in the computation

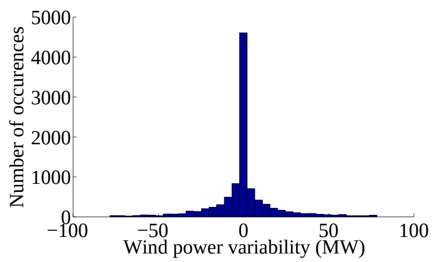

Fig. 15: Histogram of wind power variability from 7h to 8h at site S_6

for each power output trajectory ξ from h to $h+1$ so that temporal correlation is explicitly taken into account: $\{r_{g468}^{h,h+1,\xi}\} = \{p_{g468}^{h+1,\xi} - p_{g468}^{h,\xi}\}$. Generally, the probability of rampability violation of conventional generators in power systems is affected by several factors such as uncertainties and variations in the forecasts of loads and production from non-controllable generation (e.g., wind, photovoltaic solar). Variability of wind power can also be managed by the proposed model and the evaluation of the probability of rampability violation provides an example: the ramping of conventional generator g_{468} from 7h to 8h in Fig. 14 is partly contributed by variability of wind power rw_s from 7h to 8h at each site s, i.e., $\{rw_s^{h,h+1,\xi}\} = \{p_s^{h+1,\xi} - p_s^{h,\xi}\}$, $h = 7$ [8]; Fig. 15 depicts histogram of wind power variability from 7h to 8h, for example, at site S_6.

Each wind power scenario of a site, e.g., the ξ^{th} scenario at site 1 in Fig. 14, is simultaneously generated with the ξ^{th} scenario of sites from 2 to 10 in the proposed space-time correlation structure, and then they are used at once in the computation. By this way, spatio-temporal correlation of wind resources are accounted for. This is a significant improvement in capturing correlation, in comparison with either point or probabilistic forecast techniques [29].

Figure 16 compares the $p.d.f.$ of random variable $\widetilde{P}_{g468}^{8}$ of power output at 8h for different number of PCs used. Assume that the regulation band of the generator is $\pm 6\%$ of its rated power (equal to 250 MW), the probability of violation of over-/under-regulation limits (P_{g468}^{up} and P_{g468}^{low}, vertical lines in Fig. 16) of the generator at time 8h are given in Table II.

TABLE II: Probability of violation of over-/under-regulation limits of g_{468}

First PCs used	1	3	10
$\mathbb{P}\{\widetilde{P}_{g468}^{8} > P_{g468}^{up}\}(\%)$	0.89	0.72	0.58
$\mathbb{P}\{\widetilde{P}_{g468}^{8} < P_{g468}^{low}\}(\%)$	3.73	4.89	6.77

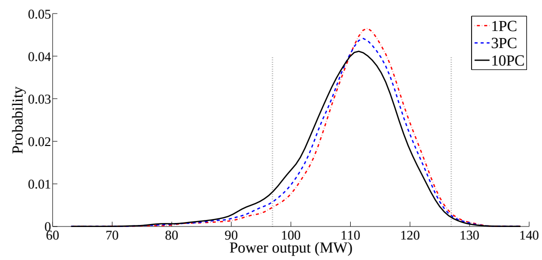

Fig. 16: $p.d.f.$ of power output of g_{468} at time 8h

V. Discussion on most promising scenario-based applications

In this paper, a powerful tool to improve the quality of generated scenarios has been proposed. In this section, a discussion on possible applications of the proposed methodology is presented. Scenario-based analysis is currently used for several applications in power systems [30], [31], from operation, to operational planning, to long-term planning.

In operation framework, the information about the uncertainty of expected wind generation is the focus, because this information is of great value for forecast users such as TSOs and market participants that need to decide their strategies. The model proposed provides valuable information about forecast uncertainty of wind generation at multiple sites: the temporal correlation of forecast uncertainty between different time instances as well as the spatial correlation between different sites are embedded in the scenarios and can be used as in [30] for reserve assessment. Another possible use of the proposed method is to provide a very suitable input for solving decision-making problems under uncertainty in electricity markets [31]. For dealing with decision-making problems, future wind power related information can be provided by forecast techniques such as point forecast and probabilistic forecast [29]; however, such techniques do not help decision makers, because temporal correlation between forecast errors (i.e., uncertainty) at different time-steps is not captured, different from the proposed model. Additionally, the methodology presented in this paper provides a set of discrete scenarios for each wind site, which is suitable for stochastic programming, like in [30], [31]. Another useful application in operational planning framework is security assessment, as the application presented in the previous section demonstrated. Similarly, other possible applications of the proposed method are: stochastic optimal power flow, stochastic unit commitment, and so on. In power flow analysis, there is a possible way to reduce computational burden: wind power scenarios generated from the proposed model can be used to estimate a $p.d.f.$ of wind power for each time-step which can be directly used in PPF adopting either an analytical technique like cumulant [28] or an approximation technique like point estimate [32].

In long-term planning [5], the scenarios generated by the proposed model can provide not only the possible range of wind production for each instance but also its dynamics over a long-term horizon. Moreover, the model can be easily extended to other sources of uncertainty such as photovoltaic solar power at multiple sites and loads (their modeling is usually expected less complicated than that for wind resources). Consequently, both the seasonal and the diurnal variations of all power injections of these resources as well as their temporal characteristics can be explicitly assessed. The model results are also applicable to other planning problems such as transmission expansion, planning reserve requirement, transmission planning, var planning, etc. [10]. Also environmental analysis in planning domain can take advantage of accurate scenarios, for quantifying the impacts of wind resources on the emission outputs [5].

VI. Conclusions

In this paper, a comprehensive modeling methodology for multi-site wind power generation scenarios is presented. The model exploits time series and Principal Component techniques together with data pre-processing techniques to explicitly capture the salient wind characteristics from multiple sites such as distinct diurnal and seasonal patterns, non-Gaussianity, spatial and temporal correlations.

Moreover, the proposed model is able to reduce the dimensions of necessary data sets, so it is very useful for working with high-dimensional data, such as wind data from a large number of sites; it is realistic, because it can be used for real wind data in practice without using any simplifying assumption.

The proposed methodology can be used for solving a wide range of problems related to multi-site wind power production: for operational planning like in stochastic power flow, stochastic optimal power flow, stochastic unit commitment, operating reserve requirement and so on; for planning studies such as transmission expansion with multi-site wind production, planning reserve requirements, transmission planning, etc.; as well as for environmental analysis to quantify the impacts of wind resources on the emission outputs. Furthermore, the proposed methodology can be easily extended to other resources such as photovoltaic solar power at multiple locations and loads.

Applications of the model to study multiple wind farms integrated into Sicilian network in Italy and extensive testing indicate good performance in effectively capturing the salient features of multi-site wind power production and providing useful insights into the impacts of the wind contributions.

Acknowledgment

We are grateful to TERNA (Italian TSO) for the data provided for the testing of the proposed approach.

References

[1] K. C. Chou and R. B. Corotis, ”Simulation of hourly wind speed and array wind power,” *Solar Energy*, vol. 26, no. 3, pp. 199-212, 1981.

[2] B. G. Brown, R. W. Katz, and A. H. Murhpy, ”Time series models to simulate and forecast wind speed and wind power,” *Journal of Climate and Applied Meteorology*, vol. 23, no. 4, pp. 1184-1195, 1984.

[3] J. L. Torres, A. Garcia, M. D.vBlas, and A. D. Francisco, ”Forecast of hourly wind speed with ARMA models in Navarre (Spain),” *Solar Energy*, vol. 79, no. 1, pp. 65-77, Jul. 2005.

[4] M. S. Miranda and R. W. Dunn, ”Spatially correlated wind speed modelling for generation adequacy studies in the UK,” *Power Engineering Society General Meeting*, Tampa, USA, pp. 1-6, June 24-28, 2007.

[5] N. Maisonneuve and G. Gross, ”A Production Simulation Tool for Systems With Integrated Wind Energy Resources,” *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2285-2292, Nov. 2011.

[6] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klockl, ”From probabilistic forecasts to statistical scenarios of short-term wind power production,” *Wind Energy*, vol. 12, no. 1, pp. 51-62, 2009.

[7] P. Meibom, R. Barth, B. Hasche, H. Brand, C. Weber, and M. O'Malley, ”Stochastic optimization model to study the operational impacts of high wind penetrationsin ireland,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1367-1379, Aug. 2011.

[8] X. Y. Ma, Y. Z. Sun, and H. L. Fang, ”Scenario Generation of Wind Power Based on Statistical Uncertainty and Variability,” *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 894-904, Oct. 2013.

[9] J. Tastu, P. Pinson, and H. Madsen, ”Space-time scenarios of wind power generation produced using a Gaussian copula with parametrized precision matrix,” *Technical Report*, Technical University of Denmark, 2013.

[10] J. M. Morales, R. Minguez, and A. J. Conejo, ”A methodology to generate statistically dependent wind speed scenarios,” *Applied Energy*, vol. 87, pp. 843-855, 2010.

[11] A. Papavasiliou and S. S. Oren, ”Stochastic Modeling of Multi-Area Wind Power Production”, *12th International Conference on Probabilistic Methods Applied to Power Systems*, Istanbul, Turkey - June 10-14, 2012.

[12] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer, 2002.

[13] J. E. Jackson, *A Users Guide to Principal Component Analysis*. New York: Wiley, 1991.

[14] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1976.

[15] J. B. Cromwell, W. C. Labys, and M. Terraza, *Univariate Tests for Time Series Models*. Thousand Oaks, California: Sage, 1994.

[16] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, ”Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root,” *Journal of Econometrics*, vol. 54, pp. 159-178, 1992.

[17] R. Chandler and M. Scott, *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*. John Wiley and Sons, 2011.

[18] H. W. Lilliefors, ”On the Kolmogorov-Smirnov test for normality with mean and variance unknown,” *Journal of the American Statistical Association*, vol. 62, pp. 399-402, 1967.

[19] T. Boehme, A. R. Wallace, and G. P. Harrison, ”Applying times series to power flow analysis in networks with high wind penetration,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 951-957, Aug. 2007.

[20] M. Thomson and D. G. Infield, ”Network power-flow analysis for a high penetration of distributed generation,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1157-1162, Aug. 2007.

[21] E. Vittal, M. O'Malley, and A. Keane, ”A steady-state voltage stability analysis of power systems with high penetrations of wind,” *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 433-442, Feb. 2010.

[22] IEC - International Electrotechnical Commission, ”Wind Turbines - Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines,” *IEC-61400-12*, 2005.

[23] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification - A Practitioner's Guide in Atmospheric Science*, Chap. 7. New York: Wiley, 2003.

[24] D. S. Wilks, *Statistical methods in the atmospheric sciences*, Chap. 8. 3rd ed. London: Academic Press, 2011.

[25] P. Pinson and R. Girard, ”Evaluating the Quality of Scenarios of Short-Term Wind Power Generation,” *Applied Energy*, vol. 96, no. 1, pp. 12-20, 2012.

[26] Y. H. Wan, E. Ela, and K. Orwig, ”Probabilistic security assessment for power system operations”, *WindPower*, Dallas, Texas, May 23-26, 2010.

[27] J. D. McCalley, S. Asgarpoor, L. Bertling, R. Billinion, H. Chao, J. Chen, J. Endrenyi, R. Fletcher, A. Ford, C. Grigg, G. Hamoud, D. Logan, A. P. Meliopoulos, M. Ni, N. Rau, L. Salvaderi, M. Schilling, Y. Schlumberger, A. Schneider, and C. Singh, ”Probabilistic security assessment for power system operations”, *Power Engineering Society General Meeting*, pp. 212-220, 2004.

[28] D. D. Le, A. Berizzi, C. Bovo, E. Ciapessoni, D. Cirio, A. Pitto, and G. Gross, ”A probabilistic approach to power system security assessment under uncertainty”, *Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid, 2013 IREP Symposium*, pp. 1-7, Greece, Aug. 2013.

[29] A. Botterud, Z. Zhou, J. Wang, R. J. Bessa, H. Keko, J. Mendes, J. Sumaili, and V. Miranda, ”Use of wind power forecasting in operations decisions”, *Technical report*, Argonne National Laboratory, Sep. 2011.

[30] A. Papavasiliou, S. S. Oren, and R. P. O’Neill, ”Reserve requirements for wind power integration: A scenario-based stochastic programming framework,” *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2197-2206, Nov. 2011.

[31] A. J. Conejo, M. Carrion, and J. M. Morales, *Decision Making Under Uncertainty in Electricity Markets*. New York: Springer, 2010.

[32] C. L. Su, ”Probabilistic load-flow computation using point estimate method,” *IEEE Trans. on Power Systems*, vol. 20, no. 4, pp. 1843-1851, 2005.

Duong D. Le (S’12–M’14) received his Ph.D. degree in Electrical Engineering from Politecnico di Milano, Milan, Italy in 2014.

He worked as a visiting research scholar at the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, USA from February to August 2012. During Feb. 2014 - Feb. 2015, Dr. Le was at the Department of Energy, Politecnico di Milano as a post-doctoral researcher.

He is currently a Lecturer at the Department of Electrical Engineering, Danang University of Science and Technology, Vietnam. His current research interests include probabilistic approaches applied to power systems, integration of renewable energy resources into power grids, power system security assessment, power system planning, operations and analysis.

George Gross (F’86–LF’11) is Professor of Electrical and Computer Engineering and Professor, Institute of Government and Public Affairs, at the University of Illinois at Urbana-Champaign. His research and teaching activities are in the areas of power system analysis, economics and operations, utility regulatory policy and industry restructuring. He was formerly with the Pacific Gas and Electric Company, where, Dr. Gross founded the company’s Management Science Department and held other key management, technical and policy positions. During 1992-93, Dr. Gross was at the Electric Research Power Institute to develop research directions on open access transmission. George Gross is a co-founder of POWERWORLD and served on its Board of Directors from 1996 - 2001. A Fellow of IEEE, Dr. Gross received the Franz Edelman Management Science Achievement Award by the Institute of Management Science. Dr. Gross is the author of a large number of publications and book chapters. He was a Visiting Professor at the Politecnico di Milano, University of Pavia and the Politecnico di Torino during the academic year 1999 - 2000. George Gross earned his undergraduate degree at McGill University in Montreal, Canada, and he did his graduate studies at the University of California, Berkeley.

Dr. Gross has consulted on electricity issues with utilities, government organizations and research institutions in North America, Europe, South America, Australia and Asia. He has lectured widely and has given numerous invited presentations at leading universities and research institutions throughout the world. His numerous publications have appeared in the leading journals in the field and his research results have been presented at a wide array of international conferences. He has made a broad range of contributions in various areas of power system planning, operations, analysis and control.

His work on smart grid issues has focused on both the technical and the regulatory aspects. The principal areas of involvement include the design of AMI architectures to ensure cyber security, the deployment of AMR for demand response, the integration of demand-side response, renewable and storage resources into the grid and the economics of smart grid implementation. He has co-organized one of the first workshops on the public policy issues in Cyber-Security and Privacy for Smart Grid Technology. He was a co-author of the paper entitled ”An Integrated Architecture for Demand Response Communications and Control,” which received the Best Paper Award in the Engineering and Economics Interactions Minitrack of the Electric Power Systems Track at the 2008 Hawaii International Conference on System Sciences HICSS-41, January 2008.

Alberto Berizzi (M’93) received his Ph.D. degree (1994) in Electrical Engineering. He is now Professor of Electric Power Systems at Politecnico di Milano, where he is also responsible for the PhD programme in Electrical Engineering at the same institution.

His areas of research include power system analysis, security and optimization.