ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
"Antonio Ruberti"

CONSIGLIO NAZIONALE DELLE RICERCHE

A. Naimo

NEW SYNTHETIC WIND FORECAST DATA GENERATION MODEL

R. 32   Dicembre 2011

Arianna Naimo – Istituto di Analisi dei Sistemi ed Informatica del CNR, viale Manzoni 30 - 00185 Roma, Italy. Email: arianna.naimo@iasi.cnr.it.

Abstract

Wind energy sources utilization and integration into electrical systems is assuming considerable importance. For this reason, new improved wind synthetic data generation models are needed, in order to properly generate forecasts of wind speed and power. These data are fundamental in simulations carried out to analyze and improve wind generating units performances, individuating wind turbines technical parameters that directly affect power production. In the present study, a new model is developed in order to generate realistic synthetic wind forecast data. Wind speed is modeled as a Weibull distribution, while wind speed forecasting error is simulated using *First-Order Auto-Regressive Moving Average - ARMA(1,1)* time-series models. Operations Research formulations for the *Assignment Problem* are used to model wind speed persistence features, which, as simulations results show in this work, are essential to properly obtain wind speed and power output forecasts.

Key words: Energy generation systems, Wind synthetic data generation, Time-series, Wind power forecasting.

1. Introduction

Wind energy sources utilization and integration into conventional electrical generation systems is rapidly becoming significant on a world scale. Currently, wind energy is produced in more than seventy countries and in some of them wind installed capacity is reaching a considerable percentage over the total electricity production [1]. Wind energy generation reaches more than 20% in Denmark and 15% in Spain [2]. Fig. 1 and Fig. 2 show how wind energy represents a high proportion of world electricity.
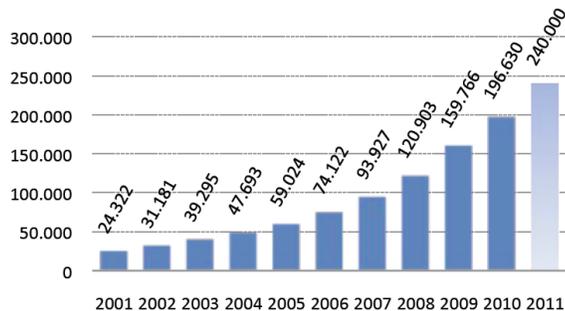

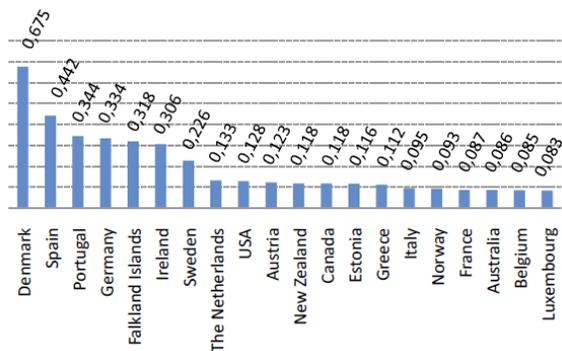Figure 1: World Total Wind Installed Capacity

Figure 2: Wind Capacity per capita

In order to use wind as an energy resource, its main characteristics, such as uncertainty and variability, related to climate variations and terrain physical complexities, should be properly analyzed.

Wind persistence is one of the most important factor to be studied in order to understand wind dynamic and intermittent behavior. Statistically, past values of wind speed time series, recorded in a given place and time horizon, are related to present wins speed values. This concept is quantified with a well known mathematical and statistical function, called autocorrelation. In this sense, Maraun *et al.* propose a definition of persistence: *The time dependency of linear stochastic models is completely captured by the autocorrelation function, describing the average linear relation of two points in time with lag. The sum of the autocorrelation function overall lags is often called memory or persistence of the process* [3]. For this reason, wind persistence properties, based on wind speed autocorrelation, should be known in order to obtain a proper wind speed and power output forecasting.

Many works presented in literature have shown that wind speed is modeled as the well known statistical distribution, called Two-Parameters Weibull distribution [4] - [11].

While studying wind characteristics, a correct estimation of wind speed forecasting error is fundamental, because of the uncertain and intermittent intrinsic nature of wind. *First-Order Auto-Regressive Moving Average - ARMA(1,1)* time-series models are the most applied method to simulate the error associated with wind speed forecasting [12].
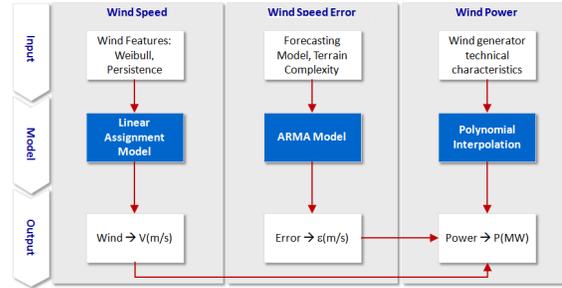
4.

Figure 3: Synthetic Wind Data Generator

The main aim of this study is to define a new model to generate synthetic forecast wind speed and power realistic data, in order to obtain proper wind data forecasting. In this model, wind speed is considered as a Weibull distribution, while ARMA(1,1) models are used to estimate wind speed forecasting error. Wind speed persistence features are modeled taking into account Operations Research formulations for the *Assignment Problem*.

Wind features simulations, carried out varying wind data and wind speed forecasting error characteristics have shown that properly modeling wind speed persistence characteristics and correctly estimating wind speed and power error is fundamental in wind data forecasting.

2. Summary

New synthetic wind data generation model proposed in this work is defined, describing how Operations Research formulations for the *Assignment Problem* could be used to take into account wind speed persistence features.

A brief overview on the characteristics of wind speed Weibull distribution is provided. Mathematical definition and parameters that determine the main features of this statistical distribution are described.

The importance of providing a correct estimation of the error associated with wind speed forecast is then explained. The main aspects related to ARMA(1,1) models are briefly described.

The following sections show how a typical wind turbine works, giving particular attention to the definition of the main aspects related to a typical wind power technical curve.

Finally, simulations results obtained with new generation model are presented and directions for future work are briefly highlighted.

3. New synthetic wind data generation model

In this work, a new synthetic wind data generation approach, based on OR models, is proposed. This generator is used to produce forecast wind speed and wind power realistic data, in order to obtain a proper wind data forecasting.

In particular, the objectives of new wind data generator are the following:

- simulating wind features, varying wind data and wind speed forecasting error characteristics

- simulating wind turbine features, individuating the technical parameters that mainly affect wind generators performance

Fig. 3 describes how new synthetic wind data generator proposed in this study works.

A brief description of the structure and the model of this generator is provided in the following sections.

3.1. Synthetic wind data generator: input parameters

New synthetic wind data generator takes the following input parameters:

- wind speed features: parameters for wind speed Weibull Two-Parameters Distribution (*shape* and *scale* parameters); wind persistence characteristics

- wind speed forecasting error: parameters for ARMA(1,1) forecasting model (*alpha*, *beta* and *sigma* parameters), parameters describing wind site terrain complexity

- wind turbine technical characteristics: main operating features of typical wind generators, like *cut-in* and *cut-out* wind speed, *rated* power and blade surface

3.2. Synthetic wind data generator: models

New synthetic wind data generator is based on three main models:

- *Assignment Model*: main wind speed characteristics such as wind speed persistence and Weibull statistical distribution are modeled with these Operations Research formulations

- *ARMA(1,1) model*: wind speed forecasting error is modeled with ARMA models, taking into account realistic values for alpha, beta and sigma parameters

- *wind turbine technical characteristics* are modeled taking into account a polynomial interpolation algorithm used to perform wind power data polynomial interpolation

3.3. Synthetic wind data generator: output parameters

The output parameters that can be obtained with new synthetic wind data generator are the following:

- forecast *wind speed* (measured in m/s)

- forecast *wind speed error* (measured in m/s)

- forecast *wind power* (measured in MW)

3.4. Forecast wind speed curve: simulation with persistence

Many works presented in literature have shown that wind speed distribution is represented by a *Weibull Two Parameters Distribution* [13].

Weibull distribution is defined as follows

$$f(v) = \left(\frac{k}{\lambda}\right) \left(\frac{v}{\lambda}\right)^{k-1} exp\left[-\left(\frac{v}{\lambda}\right)^{k}\right] \tag{1}$$

where:

- $f(v)$ is the probability to observe wind speed v

- k is the *shape parameter*, which is used to determine the shape of the Weibull distribution and usually varies from 1.2 to 2.75 [4] - [11].

- λ is the *scale parameter*, which represents a scale factor for the Weibull distribution

Weibull distribution has the following properties

1. **first property**:

$$P(v_1 < v < v_2) = exp\left[-\left(\frac{v_1}{\lambda}\right)^{k}\right] - exp\left[-\left(\frac{v_2}{\lambda}\right)^{k}\right]$$

2. **second property**: wind speed expected value is given by

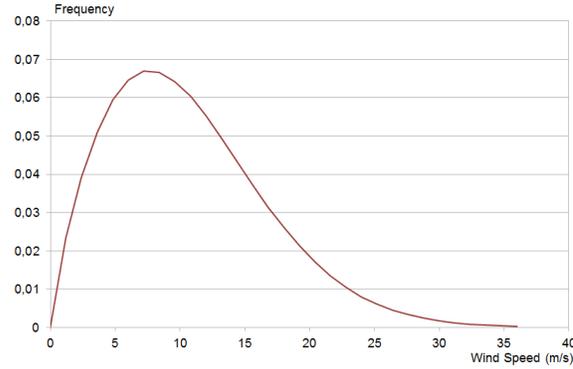$$\mu = \lambda\Gamma\left(1 + \frac{1}{k}\right)$$
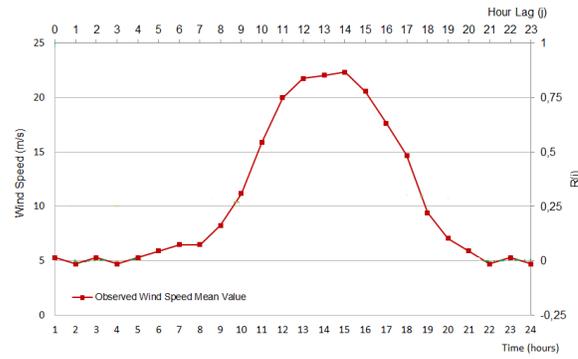
6.

Figure 4: A typical Weibull distribution curve

Figure 5: Hourly observed mean wind speed curve

3. **third property**: wind speed distribution variance σ^2 is given by

$$\sigma^2 = \lambda^2 \left[\Gamma \left(1 + \frac{2}{k} \right) - \Gamma^2 \left(1 + \frac{1}{k} \right) \right]$$

Weibull *cumulative distribution function (cdf)* is defined as follows

$$1 - e^{-(v/\lambda)^k} \tag{2}$$

Weibull distribution can be determined by substituting the input *shape and scale parameters* in the expression (1).

Besides, Weibull *cdf* can be determined by substituting the input *shape and scale parameters* in the expression (2).

A typical Weibull distribution is shown in Fig. 4.

3.5. Wind speed forecast curve simulation

Wind speed forecast curve can be simulated, according to the Two-Parameters Weibull Distribution previously defined, taking into account the observed mean wind speed curve, which is shown in Fig. 5.

In order to simulate wind speed forecast curve, we can consider the expression (1) and we can assume that v is a random value x chosen between 0 and 1. It is thus possible to obtain the expression for wind speed y, as follows

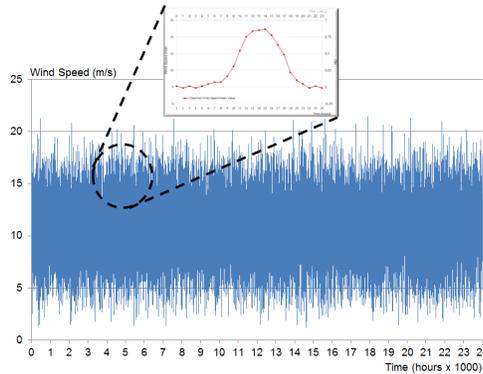$$y = -\lambda^k ln(1-x)^{\frac{1}{k}} \tag{3}$$

Figure 6: Mean observed wind speed simulations over 24000 hours

$x = rand(0, 1)$ is a random value generated according to a normal distribution between 0 and 1.

Wind speed forecast curve, measured in m/s, is thus simulated and it is shown in Fig. 6.

Fig. 6 shows that, even if the statistical wind speed properties of the Weibull distribution are maintained, wind speed seems to have a random behavior. Actually, wind speed does reflect persistence characteristics; in fact, Fig. 6 is the result of the simulations of wind speed over 24000 hours, taking into account observed mean wind speed.

Nevertheless, for this reason, and for the apparent random behavior of wind speed, enhanced models to correctly investigate wind speed characteristics and to properly simulate wind persistence are needed.

3.6. Weibull Distribution parameters estimation

Many works in literature have shown that wind speed distribution reflects a Two-Parameters Weibull Distribution, by properly estimating λ and k Weibull distribution parameters, considering a given mean μ and an assigned variance σ^2.

In literature, the most common methods, used to estimate Weibull distribution parameters, are [14]:

- *graphic method*

- *maximum likelihood method*

- *moment method*

The most applied method, called *moment method*, has been proposed by Justus *et al.* [14]. They showed that, by equaling the following expressions

$$\mu^2 = \lambda^2 \Gamma^2 \left(1 + \frac{1}{k}\right)$$

$$\sigma^2 = \lambda^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right]$$

the following result is obtained

$$\frac{\sigma^2}{\mu^2} = \left(\frac{\sigma}{\mu}\right)^2 = \frac{\lambda^2 \Gamma\left(1 + \frac{2}{k}\right)}{\lambda^2 \Gamma^2\left(1 + \frac{1}{k}\right)} - 1$$

which is equivalent to

$$\lambda = \frac{\mu}{\Gamma\left(1 + \frac{1}{k}\right)}$$

8.

By the iterative application of this method, it is possible to calculate Weibull distribution λ and k parameters, given a mean μ and a variance σ^2 [14].

Besides, it is possible to show that wind speed hourly forecast, determined previously, is a Weibull distribution.

3.7. Modeling wind speed persistence: Assignment Problem formulations

Modeling wind persistence properties is essential in carrying out wind speed forecasting, for the following reasons:

- wind speed is distributed according to the well known Weibull statistical distribution

- wind speed autocorrelation reflects synoptic and seasonal wind speed behavior

- persistence is measured as the mean duration of wind speed in a given time interval for a concrete site

Many models for wind speed persistence have been proposed in literature [1]. The most applied models are:

- Auto-Regressive Moving Average (ARMA) models

- Markov models

- Wavelet models

Auto-Regressive Moving Average (ARMA) models are based on wind time-series over different time scales. In these models, even if wind speed autocorrelation partially reflects synoptic cycles, wind speed persistence and Weibull statistical characteristics are altered.

Markov models are based on Markov chains which are used to give a model of wind speed characteristics. In these models, even if wind speed persistence and Weibull statistical characteristics are partially maintained, wind speed autocorrelation does not reflect synoptic cycles. Several variants of Markov models exist, like auto-regressive and Montecarlo Markov models, which represent a combination of ARMA and Montecarlo models with Markov models [16].

Wavelet models are based on wavelet approximating functions which are decomposed and recomposed in order to generate wind speed sequences. Wavelet models are currently the most applied approaches even if wind speed autocorrelation and persistence are not completely modeled.

In this work, in order to overcome all the drawbacks of the models presented in literature, wind speed persistence features are modeled taking into account Operations Research formulations for the **Assignment Problem**.

The *Assignment Problem* is one of the most important Combinatorial Optimization problems [17]. Given two sets, containing the same number of elements, each of one differs from the others, it is required to assign exactly one element belonging to the first set to one element belonging to the second set, so that all the elements belonging to the sets have been assigned and the total cost of the assignment is minimized.

In our study:

- the elements belonging to the first set are represented by mean wind speed values, which are individuated on a given temporal horizon (time intervals are usually equal to one hour)

- the elements belonging to the second set are represented by wind speed values generated in a random way, according to a Weibull distribution, and that have to be reassigned, in order to consider also wind speed persistence features

- assignment costs are represented by distances between mean observed wind speed values and random wind speed values
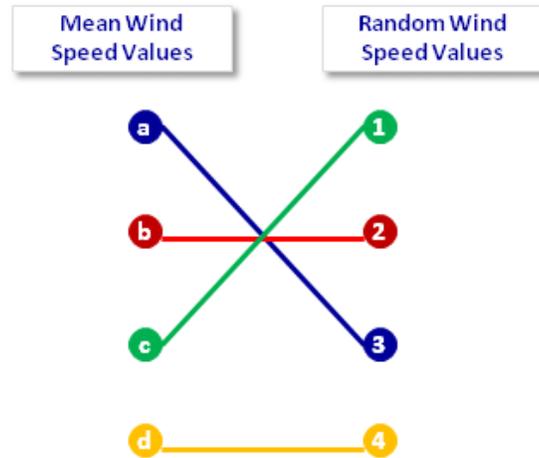
Figure 7: An example for the Assignment Problem, used to model wind speed persistence characteristics

In other terms, the *Assignment Problem* consists in finding an assignment of the elements as shown in Fig. 7, where the first set is represented in the left side of the picture, and the second set is shown in the right side of the figure.

In this example:

- random wind speed value **1** has been assigned to mean wind speed value **c**

- random wind speed value **2** has been assigned to mean wind speed value **b**

- random wind speed value **3** has been assigned to mean wind speed value **a**

- random wind speed value **4** has been assigned to mean wind speed value **d**

In this study, the *Assignment Problem* has been solved considering the Hungarian Algorithm, which is a well-known Operations Research algorithm used to solve this type of problem. The elements of cost matrix are represented by the distances between the mean wind speed values and the wind speed values generated in a random way without considering persistence features.

These distances have been calculated with the following formula:

$$d_t = \frac{|\bar{v}_t - v_{rand,t}|}{2} \tag{4}$$

where

- d_t represents the distance between mean and random wind speed values at time t

- $\bar{v}_t$ represents mean wind speed value at time t

- $v_{rand,t}$ represents random wind speed value at time t

Solving the *Assignment Problem* with the Hungarian Algorithm, the correct re-assignment of the random wind speed values to the mean wind speed values is thus obtained, minimizing the total distance between the random points and the mean points.
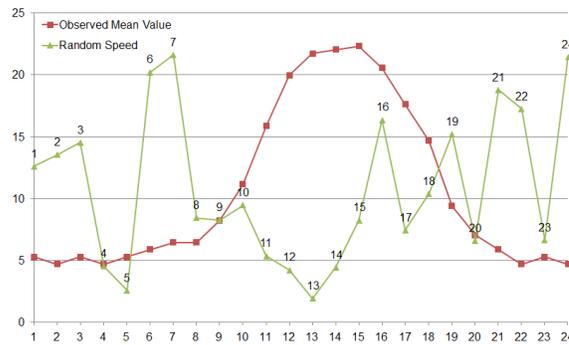
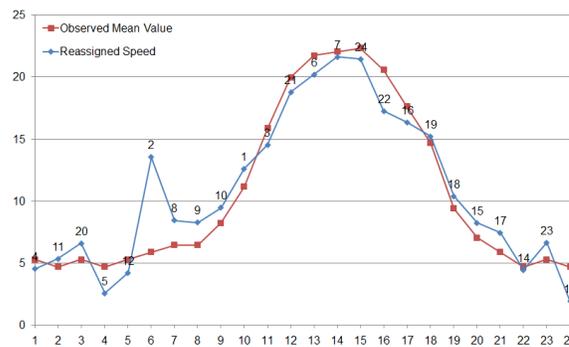Figure 8: Wind speed generated according to a random distribution

Figure 9: Wind speed generated according to an Assignment Problem formulation

3.8. Assignment Problem: simulations results

Operations Research formulations for the *Assignment Problem* have been considered in order to model wind speed persistence characteristics. Observed mean wind speed is depicted in Fig. 8 (red curve).

If wind speed values are generated in a random way, taking into account only Weibull distribution features, but without considering wind speed persistence, the main result that is obtained in simulation phase is that this random wind speed curve does not reflect the observed wind speed at all. These results can be easily verified by analyzing green curve depicted in Fig. 8.

If wind speed values are generated according to a solution of the *Assignment Problem*, taking into account both Weibull distribution features and wind speed persistence, the main result that is obtained in simulation phase is that this re-assigned wind speed curve reflects the observed wind speed in a realistic way. These results can be easily verified by analyzing blue curve depicted in Fig. 9.

If wind speed values are generated in a random way, random wind speed mean assumes a constant value and random wind speed auto-correlation is next to zero. This means that random wind speed does not reflect synoptic wind speed cycles and persistence features.

If wind speed values are generated according to a solution of the *Assignment Problem*, re-assigned wind speed mean assumes values next to the observed wind speed mean values and re-assigned wind speed auto-correlation differs from zero. This means that re-assigned wind speed reflects synoptic wind speed cycles and persistence features. Thus, simulations results show that formulations for the *Assignment Problem* represent a useful instrument to correctly model wind speed persistence features, which are fundamental characteristics to perform a proper wind speed forecasting.

Fig. 12 compares the following wind speed distributions, obtained in simulation phase:

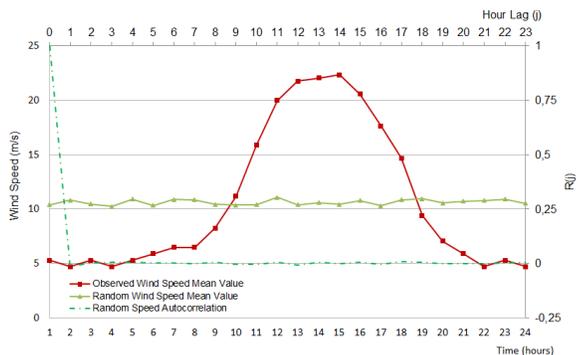- nominal wind speed Weibull distribution

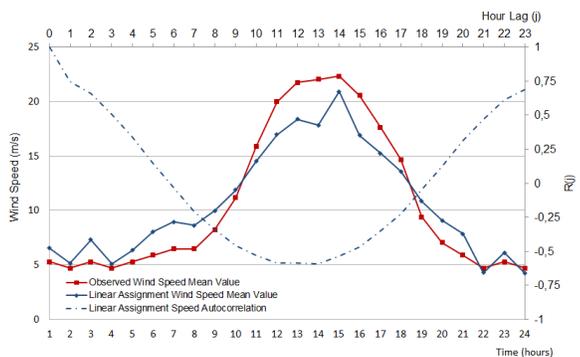Figure 10: Simulation results obtained with random generation

Figure 11: Simulation results obtained with an Assignment Problem formulation
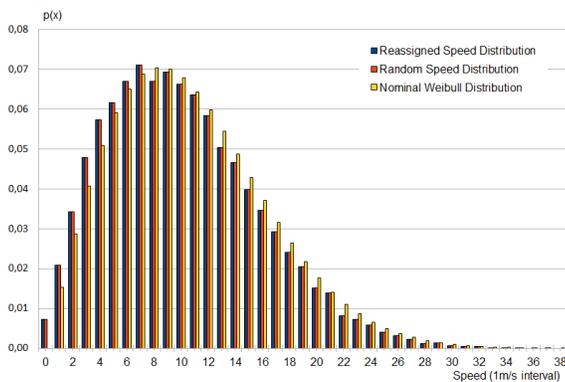

Figure 12: Weibull Distribution obtained in simulation phase

- random wind speed Weibull distribution

- wind speed Weibull distribution, re-assigned according to the solution of the *Assignment Problem*

Fig. 12 shows that all the wind speed distributions mentioned above realistically reflect Weibull distribution. This means that solving the re-assignment with the *Assignment Problem* does not alter the statistical features of wind speed, but, moreover, takes also in account wind speed persistence features.

4. Forecast wind speed error curve simulation

Given a certain wind speed hourly forecast, it is necessary to simulate the associated forecast wind speed error. A realistic measure of wind speed forecasting error can be obtained with ARMA(1,1) models, which was proposed by Soder [12] [15].

The $ARMA(1,1)$ model - *1st-order Auto-Regressive Moving Average Model* simulates the forecast wind speed error. These models are based on *time-series*, which are sequences of observations; each observation is recorded in a specified time interval. Examples of *time-series* are actually represented by forecast wind speed errors.

Different types of *time-series* exist:

- *discrete time-series*: observations are recorded in discrete time instants, with a constant time interval between these instants

- *parametric time-series*: *time-series* characteristics can be described by constant parameters

- *uni-variate time-series*: it is composed by observations of a single variable

- *multi-variate time-series*: it is composed by observations of different variables.

$ARMA(1,1)$ model creates discrete, parametric and multi-variate *time-series*. These *time-series* are defined as follows

$$X(0) = 0 \tag{5}$$

$$Z(0) = 0 \tag{6}$$

$$X(t) = \alpha X(t-1) + \beta Z(t-1) + Z(t) \tag{7}$$

where

- $X(t)$ is forecast wind speed error at time t

- α is an auto-regressive constant parameter which determines at which grade the previous value in *time-series* influences the current value

- β is a *moving-average* constant parameter which determines at which grade the gaussian random variable of the previous parameter in *time-series* influences the current value

- $Z(t)$ is a gaussian random variable with

  - mean zero
  - standard deviation σ_z at time t.

An $ARMA(1,1)$ *time-series* is described by the following 3 parameters

- α

- β

- σ_z

We assume that $ARMA(1,1)$ parameters depend on 3 factors:

1. precision of forecast wind speed estimation (this is related to prediction system simplicity and accuracy)

2. terrain complexity of the given site where wind turbines are placed

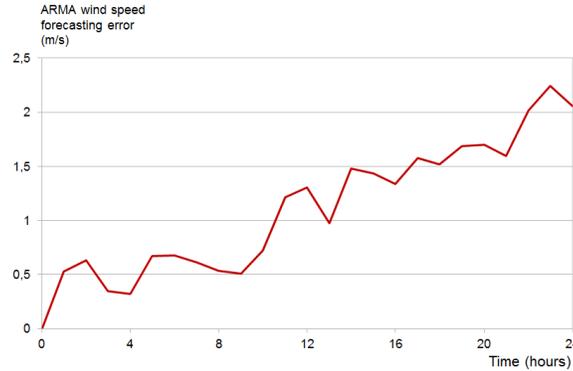|		α	β	σ_z
S.T.	**S.P.S.**	0.98	-0.7	$1.0 \div 2.5 \ f(\overline{w})$
S.T.	**A.P.S.**	> 0.98	$-0.8 \div -0.4 \ f(\overline{w})$	$0.75 \div 1.5 \ f(\overline{w})$
C.T.	**S.P.S.**	0.8	-0.1	$2.5 \div 3 \ f(\overline{w})$
C.T.	**A.P.S.**	0.8	-0.1	$0.50 \div 2.0 \ f(\overline{w})$

Table 1: Possible values that α, β and σ_z parameters can assume

Figure 13: Wind speed forecast error curve

3. medium wind speed level at the given site where wind turbines are placed

The following table shows the possible values that α, β and σ_z parameters can assume (S.T. means *Simple Terrain*, C.T. means *Complex Terrain*, S.P.S. means *Simple Prediction System*, A.P.S. means *Advanced Prediction System*, $f(\overline{w})$ means that parameters depends on mean wind speed $\overline{w}$).

Many works in literature [12] have shown that the most realistic values that $ARMA(1,1)$ parameters can assume are $\alpha = 0.95$, $\beta = -0.6$, $\sigma_z = 0.5$.

Wind speed forecast error curve determined with the $ARMA(1,1)$ model is shown in Fig. 13.

5. Forecast wind turbine power curve

Analyzing wind power technical curve is fundamental to obtain correct wind power forecasting. Given a wind power plant, the forecast power curve, for an assigned forecast wind speed, is shown in the wind turbine technical *datasheet*.

A typical wind turbine usually works in four modalities which differ according to the input wind speed, as depicted in Fig. 14:

- area 1: forecast wind speed values are lower than cut-in wind speed

- area 2: forecast wind speed values are comprised between cut-in wind speed and rated wind speed

- area 3: forecast wind speed values are comprised between rated wind speed and cut-out wind speed

- area 4: forecast wind speed values are greater than cut-out wind speed

As reported in wind turbines technical datasheets, three types of characteristic wind speeds exist:

- cut-in speed: minimum speed which guarantees that wind turbine works properly

- rated speed: minimum speed which guarantees that wind turbine works at nominal power

- cut-out speed: wind turbine must be blocked if wind speed assumes values greater than the cut-out speed, in order to avoid structural damages
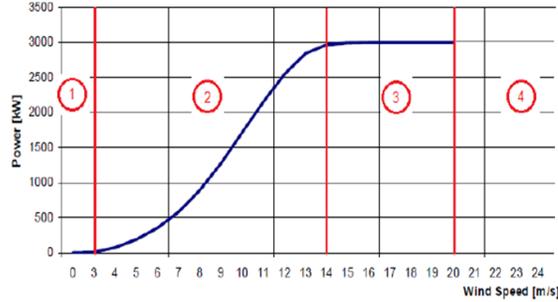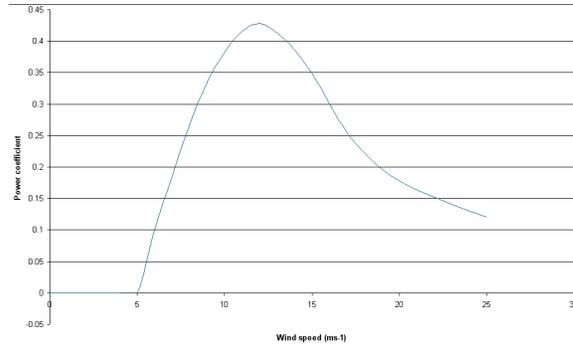
14.

Figure 14: Wind turbine technical curve

Figure 15: C_P power coefficient curve for a given wind turbine

Power curve is calculated by the following expression

$$P = 0.5\rho C_P \pi R^2 (v_j^i)^3 \qquad (8)$$

where

- ρ is the air density factor

- C_P is the wind turbine *power coefficient*, which is useful to measure its efficiency

- $2R$ is the wind turbine diameter, given in technical *datasheet*

- v_j^i is the medium wind speed for wind power plant i at time j

Formula (8) presents a difficulty in the calculation of C_P *power coefficient*.

In fact, C_P term cannot be determined in a closed form, but we can calculate only its maximum value, which is about 0.59 according to *Betz's Theory*.

C_P *power coefficient* depends on the technical characteristics of the given wind turbine. A typical C_P *power coefficient* curve is shown in Fig. 15.

Power curve is determined according to power coefficient curve. For these reasons, power curve cannot be easily reproduced.

These issues motivate the necessity to use the technical *datasheet* power curve as an input for our wind data generator.

5.1. Wind power data interpolation

Given the forecast wind speed and the wind turbine *datasheet* power curve, it is possible to determine the wind turbine forecast power associated with a given wind speed. Polynomial data interpolation, based
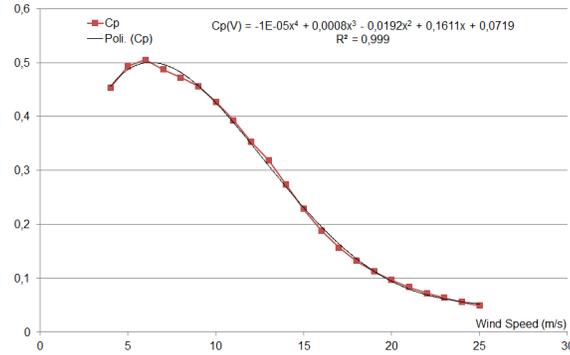
Figure 16: C_P power coefficient curve polynomial interpolation

on real technical data, is performed to determine wind turbine forecast power associated with a given wind speed.

Technical *datasheets* present a discrete power curve, it means that there exist a correspondence between each wind speed value and a punctual power value.

Nevertheless, the input forecast wind speed value could not be present in the technical *datasheet*; for this reason, it is necessary to make a data interpolation, according to the following steps:

- C_P is determined for all the given *datasheet* wind speeds and an assigned air density factor

- C_P is approximated with a polynomial curve

- C_P is re-calculated according to the interpolation equation

- power associated with input wind speed is determined, by substituting C_P value, calculated at step 3, in the power equation

6. New synthetic wind data generator: simulation results

New synthetic wind data generator, subject of this work, has been developed in order to analyze wind forecast realistic data, according to different scenarios.

The most important simulations results that have been obtained are briefly commented in the following sections.

6.1. Input parameters for simulations

Simulations have been performed not altering the wind turbine technical input parameters, while wind speed Weibull distributions parameters have been changed.

In particular, a block consisting of 1000 simulations has been performed over a time horizon of 24 hours, considering the following parameters for the Weibull distribution

- $\lambda = 12$

- $k = 3,9$

All the other input parameters, like wind turbine technical parameters, described in the previous sections, have not been altered.
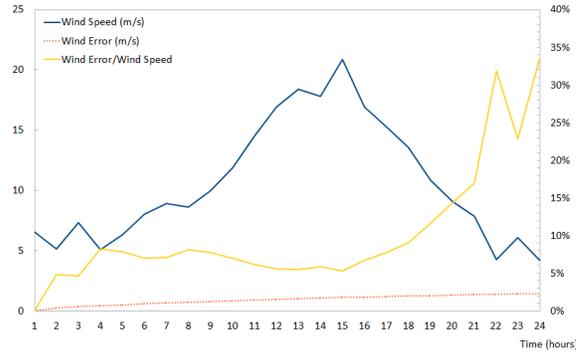
16.

Figure 17: Wind speed hourly distribution

6.2. Simulations Results: wind speed analysis

Simulations results obtained with new synthetic wind data generator show important aspects related to wind speed. In particular, by analyzing Fig. 17, it is possible to observe the following curves:

- wind speed hourly distribution curve (blue curve)

- wind speed forecasting error curve (red curve)

- relationship between wind speed forecasting error curve and wind speed itself (yellow curve)

Fig. 17 shows that wind speed hourly distribution curve assumes a smoother profile if the value of Weibull distribution k parameter increases.

6.3. Simulations Results: wind power analysis

Simulations results obtained with new synthetic wind data generator show important aspects related to wind power. Simulations have been performed considering the following wind turbine technical parameters:

- *cut-in* wind speed

- *cut-out* wind speed

- *rated wind speed*

- air density factor ρ

- wind turbine blade area S

- *rated wind power*

- coefficients of polynomial curve which approximates C_P power coefficient curve

Given these parameters, it has been possible to determine:

- an *upper bound* on forecast wind speed $v + e$ for each time interval t (where e is the error on forecast wind speed and v is forecast wind speed)

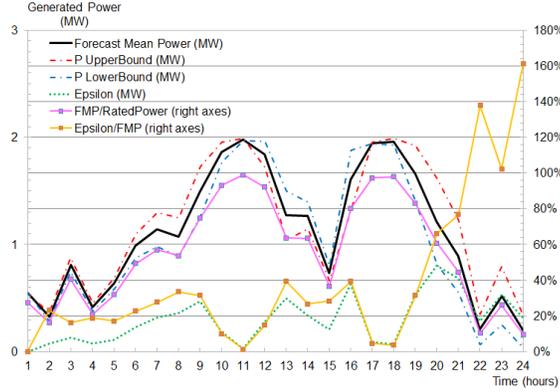- a *lower bound* on forecast wind speed $v - e$ for each time interval t

Figure 18: Simulation results obtained with an Assignment Problem formulation

- an *upper bound* on forecast wind power P for each time interval t.

  This *upper bound* is determined considering the following expression

  $$P = 0.5\rho C_P \pi R^2 (v_j^i)^3 \tag{9}$$

  where C_P is determined by the polynomial equation which approximates its curve, while wind speed is given by the *upper bound* $v + e$

- a *lower bound* on forecast wind power P for each time interval t.

  This *lower bound* is determined considering the expression 9, where wind speed is given by *lower bound* $v - e$

- forecast wind power error ϵ, which is determined as follows

  $$\epsilon = P_{UB} - P_{LB} \tag{10}$$

- medium forecast power for the assigned wind power plant, which is determined as follows

  $$P_{med} = \frac{P_{UB} + P_{LB}}{2} \tag{11}$$

In particular, by analyzing Fig. 18, it is possible to observe the following curves:

- mean forecast wind power (black curve)

- upper bound (UB) on forecast wind power (red dot curve)

- lower bound (LB) on forecast wind power (blue dot curve)

- wind speed forecasting error (green dot curve)

- relationship between forecast mean wind power (FMP) and rated power (RP) (pink curve)

- relationship between wind power forecasting error and forecast mean wind power (FMP) (yellow curve)

Fig. 18 shows that the peaks in the wind speed curve (figure 17) correspond to a decreasing in wind power. This behavior is reasonable, since wind turbines are arrested and they do not produce power anymore if wind speed values are greater than a given maximum value; this is made in order to prevent technical damages to wind turbines.
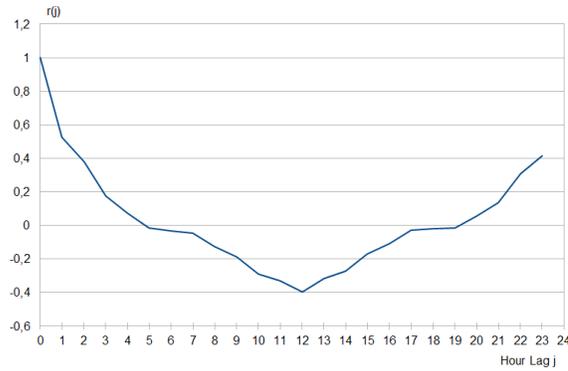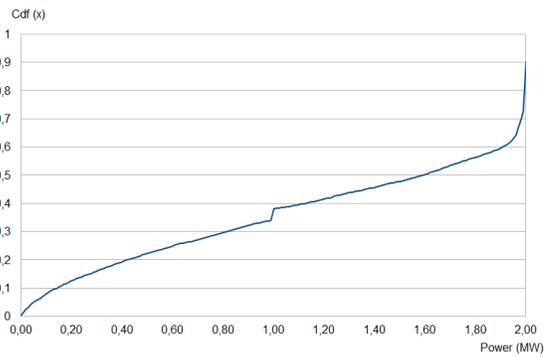
Figure 19: Wind power auto-correlation

Figure 20: Wind power cdf

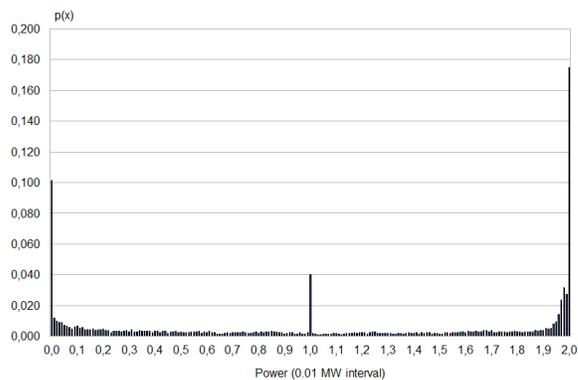Figure 21: Wind power pdf

Analyzing Fig. 19, Fig. 20 and Fig. 21 it could be easily verified that power statistical properties are completely different from wind speed ones.

This aspect is fundamental in wind speed forecasting, since wind power cdf and pdf cannot be taken into account to perform a correct wind forecasting, but wind speed main features, such as persistence, have to be considered.
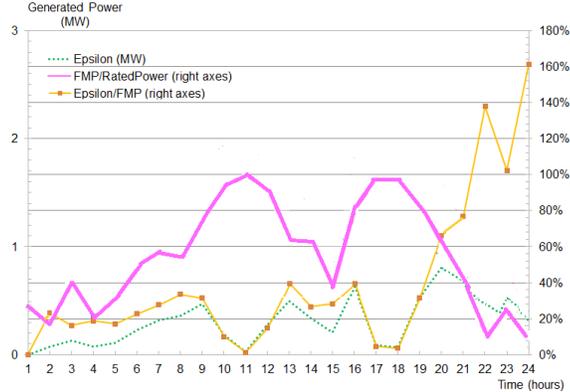
Figure 22: Wind speed forecasting error is comparable to power in long forecast time horizons

6.4. Simulations Results: wind speed forecasting error analysis

If we analyze the forecast wind power curve from a qualitative point of view, we can observe two main behaviors of the measure of ϵ. This is a measure of the error related to medium forecast wind turbine power and it is determined as the difference between the *lower bound* on power (related to wind forecast in an overestimation case, which leads to a negative error ϵ) and the *upper bound* on power (related to wind forecast in an underestimation case, which leads to a positive error ϵ). In particular, the error:

- assumes values comparable to power, if the forecast time horizon increases. For this reason, considering long forecast time horizons for solving Wind UC problem could not be effective, from a practical point of view, as shown in Fig. 22. This means that Unit Commitment Problem, in presence of wind energy sources, could not be solved over a long-term time horizon, because, in this case, wind forecasting error will assume a value that should not be neglected; in order to solve the UC problem in a more efficient way, it should be preferable to adopt the Rolling-UC approach, which consists in solve the UC over a short-term time horizon (which is normally equal to six hours);

- it is strongly greater than forecast wind speed error, because it depends on v^3 and on C_P power coefficient.

Besides, ϵ depends also on wind turbine technical parameters (for instance, *rated wind speed*) and Weibull distribution parameters (for instance λ).

7. Conclusions and directions for future work

In this study, a new model has been developed to generate synthetic wind forecast data. Wind speed has been modeled as a Weibull distribution, while wind speed forecasting error has been simulated using *First-Order Auto-Regressive Moving Average - ARMA(1,1)* time-series models. Operations Research formulations for the *Assignment Problem* have been used to model wind speed persistence features. Simulations results obtained in this work have shown that wind persistence features are essential in carrying out a complete wind speed and power output forecasting.

In the future, work will be addressed in developing a refined model for synthetic wind data generator, considering more accurated ARMA models, which could take into account different terrain complexities and weather conditions.

Assignment Problem formulations for random wind speed values re-assignment will be enhanced, in particular studying more accurate mode to define distances between wind speed values. In fact, varying the way in which distances are calculated, different results for wind speed values are obtained.

By performing a large-scale simulations plan, it could be possible to individuate the most meaningful patterns that could be useful to evaluate the influence of wind features and wind turbines technical characteristics on wind power generation.

20.

In particular, wind forecast synthetic data, generated with the model proposed in this work, could be used to carry out simulations studies in order to individuate wind turbines operational parameters that mainly affect wind generators performances, like cut-in, cut-out and rated wind speed, wind turbine blade surface, rated power and air density.

Wind data generator could be also used to analyze forecast wind speeds and errors related to these forecasts. In particular, for instance, we could:

- study the forecast wind speed curve, according to different values that Weibull distribution *shape and scale parameters* could assume

- study the forecast wind speed error curve, according to different values that $ARMA(1,1)$ model α, β and σ_z parameters could assume

- study the *upper bound* and the *lower bound* forecast wind power curves and the associated error ϵ, according to different values that wind turbine technical parameters could assume (for instance, Weibull and $ARMA(1,1)$ parameters). In particular, *upper bound* and *lower bound* on forecast wind power could be used to define a Robust formulation for Wind Unit Commitment.

References

[1] Cancino-Solrzano Y., Gutirrez-Trashorras A.J., Xiberta-Bernat J., *Analytical methods for wind persistence: Their application in assessing the best site for a wind farm in the State of Veracruz, Mexico.*, Renewable Energy, 35:2844-2852, 2010.

[2] *World Wind Energy Report 2010*, WWEC 2011, 10th World Wind Energy Conference and Renewable Energy Exhibition, Greening Energy: Converting Deserts into Powerhouses, Cairo, Egypt, 31 October - 2 November 2011

[3] Maraun D, Rust H, Timmer J., *Tempting long-memory - on the interpretation of DFA results.*, Nonlinear Processes in Geophysics 2004;11:495-503.

[4] Corotis R.B., Sigl A.B.. Klein J., *Probability models of wind velocity magnitude.*

[5] Hennessey Jr. J.P., *Some aspects of wind power statistics.*, Journal of Applied Meteorology, 1977;16:119-28.

[6] Justus C.G., Hargraves W.R., Yalcin A., *Nation wide assessment of potential output from wind powered generators.*, Journal of Applied Meteorology, 1976;15:673-8.

[7] Stevens M.J.M., Smulders P.T., *The estimation of the parameters of the Weibull wind speed distribution for wind energy utilization purposes.*, Wind Engineering, 1979;3:132-45.

[8] Rohatgi J.S., de Meiros A.L.R., Cavalcanti R.A., *Statistical Testing of Weibull and other frequency distributions of wind speed variations for different sites.*, In: EWEC '89, Proceedings of European Wind Energy Conference and Exhibition, Brazil: Peter Peregrino Ltd; 1989,880-883.

[9] Rehman S., Halawani T.O., Husain T., *Weibull parameters for wind speed distribution in Saudi Arabia.*, Solar Energy 1994;53:473-9.

[10] Garcia A., Torres J.L., Prieto E., De Francisco A., *Fitting wind speed distributions: a case study.*, Solar Energy 1998;62:139-44.

[11] Gupta B.K., *Weibull parameters for annual and monthly wind speed distributions for five locations in India.*, Solar Energy 1986;37:469-71.

[12] Boone A., *Simulation of short-term wind speed forecast errors using a multi-variate ARMA(1,1) time-series model.*, Dept of Electrical Engineering, Electric Power Systems Stockholm, Sweden Royal Institute of Technology, 2005.

[13] Bivona S., Burlon R., Leone C., *Hourly wind speed analysis in Sicily.*, Renewable Energy, 28:1371-1385, October 2002.

[14] Justus C.G., Hargraves W.R., Graber A.M., *Methods for estimating wind speed frequency distributions.*, Journal of Applied Meteorology, 17:350-353,1978.

[15] Soder L., *Simulation of wind speed forecast errors for operation planning of multi-area power systems.*, In 8th International Conference on Probabilistic Methods Applied to Power Systems, 12 - 16 September 2004, Iowa State University, United States, 2004.

[16] Brokish K., Kirtley J., *Pitfalls of Modeling Wind Power Using Markov Chains*, IEEE Transactions on Power Systems, 2009.

[17] Burkard R., Dell'Amico M., Martello S., *Assignment Problems.*, SIAM, 2009.