TIME-SERIES MODELS FOR RELIABILITY EVALUATION OF POWER SYSTEMS INCLUDING WIND ENERGY

R. BILLINTON
Power System Research Group, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0W0, Canada

HUA CHEN
Department of Electrical Engineering, Chongqing University, Chongqing, 630044 China

R. GHAJAR
Department of Electrical Engineering, Lebanese American University, Byblos, Lebanon

(Received for publication 28 July 1995)

Abstract—An essential step in the reliability evaluation of a power system containing Wind Energy Conversion Systems (WECS) using sequential Monte Carlo analysis is to simulate the hourly wind speed. This paper presents two different time-series models generated using different available wind data. Wind data from Environment Canada and SaskPower are used to illustrate these models. No assumptions or previously estimated factors are included in the models. In order to check the adequacy of the proposed models, the F-criterion and Q-test are used, and the statistical characteristics of the simulated wind speeds are compared with those obtained from the actual wind speeds. The proposed wind models satisfy the basic statistical tests and preserve the high-order auto-correlation, seasonal property and diurnal distributions of the actual wind speed. Copyright © 1996 Elsevier Science Ltd.

1. INTRODUCTION

The development and utilization of renewable energy sources is being given serious consideration due to expressed concerns regarding dwindling conventional energy resources, and enhanced public awareness of the potential impact of conventional energy systems on the environment. Wind is considered by many to be a free, non-depletable and environmentally friendly source of energy. The successful operation of many wind farms throughout the world has illustrated that wind energy can be an encouraging and promising energy option.

At the present time, many utilities are prepared to give an energy credit to a wind facility but are reluctant to assign it a capacity credit. The actual benefits cannot be assigned in the absence of a comprehensive reliability modeling technique for Wind Energy Conversion Systems (WECS). It is therefore both necessary and important to develop reliability evaluation techniques which include WECS. Utilities can then assess the effect of a WECS on the system reliability and choose a suitable or even the optimal wind power penetration level to augment the conventional energy conversion systems.

Most of the reported work on modeling wind power generation and on the use of such models for reliability evaluation is in the analytical domain [1–4]. The most obvious deficiency of analytical methods is that the chronological characteristics of wind velocity and its effects on wind power output cannot be considered. Sequential Monte Carlo simulation, on the other hand, can be used to incorporate these considerations in the adequacy assessment of power systems containing WECS.

One of the crucial steps in reliability evaluation of a power system containing WECS using sequential Monte Carlo simulation is to simulate the hourly wind speed and this has been the subject of several publications [5–8]. In 1981, a method using the Weibull distribution function and a lag-one Auto-Regressive (AR) model which uses 24 previously estimated diurnal cycle factors per month was developed [5]. In this approach, the relatively high order auto-correlation was significantly underestimated. A simplified AR(24) model was subsequently established to mimic both the hourly wind speed and direction [6, 7]. There are many assumptions in this model, one of which is that 'the hourly wind speed must be normally distributed'. In Refs [8] and [9], a series of AR(2) models are presented for simulating the main statistical characteristics of wind speed. There are no indications in these references on whether the proposed models can preserve the diurnal distribution and some of these models do not pass the Chi-square distribution test.

This paper presents two different time-series models generated using different available wind data. Wind data from Environment Canada and SaskPower are used to illustrate these models. No assumptions or previously estimated factors are included in the models. In order to check the adequacy of the models, the F-criterion and Q-test were utilized, and the statistical properties of the simulated wind velocity
are compared with those obtained from the actual wind velocity. The first series of models proposed in this paper can simulate not only the auto-correlation and seasonal property, but also the diurnal distribution. The second series of models proposed in this paper can simulate not only the auto-correlation and seasonal property, but also the diurnal distribution. It is proposed that the first series of models can provide a valid representation of wind velocity for use in reliability studies of power systems including WECS.

2. METHODOLOGY
2.1. General expressions of wind speed models

Let

\begin{align*}
&\text{OW}_t = \text{the observed wind speed at hour } t, \\
&\mu_t = \text{the mean observed wind speed at hour } t, \\
&\sigma_t = \text{the standard deviation of observed wind speed at hour } t, \\
&\mu = \text{the mean wind speed of all the observed data}, \\
&\sigma = \text{the standard deviation of wind speed obtained from all the observed data and} \\
&\text{SW}_t = \text{the simulated wind speed at hour } t.
\end{align*}

Different time series models can be established using different combinations of the above data. Generally, let

\[ y_t = f(\text{OW}_t, \mu_t, \sigma_t, \mu, \sigma, \ldots). \]

The data series set \( y_t \) can be used to build the following Auto-Regressive and Moving Average ARMA\((n, m)\) time series model:

\[ y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_n y_{t-n} + \sigma_1 x_{t-1} - \theta_1 x_{t-1} - \theta_2 x_{t-2} - \cdots - \theta_m x_{t-m}, \]

where \( \phi(i = 1, 2, \ldots, n) \) and \( \theta(j = 1, 2, \ldots, m) \) are the auto-regressive and moving average parameters of the model, respectively. \( [x_t] \) is a normal white noise process with zero mean and a variance of \( \sigma_x^2 \) (i.e. \( x_t \sim \text{NID}(0, \sigma_x^2) \)), where NID denotes Normally Independently Distributed).

A pure AR\((n)\) model can be treated as a special form of Auto-Regressive and Moving Average Model ARMA\((n, m)\) by setting \( m = 0 \). Once the time series model of wind speed is established, the simulated wind speed can be calculated as follows:

\[ \text{SW}_t = f^{-1}(y_t, \mu_t, \sigma_t, \mu, \sigma, \ldots), \]

where \( f^{-1}(\cdot) \) is the inverse function of \( f(\cdot) \).

2.2. Estimation of parameters

The linear least square approach can be used to estimate the parameters \( \phi \) and \( \sigma^2 \) when \( m \neq 0 \). The basic steps used in the least square approach include estimating the initial values and searching for the optimal values based on the initial guesses.

The qualities of the starting values play a very important role in the convergence of the iterations. A systematic method for estimating the initial values was developed in this research work. Three different approaches were used to guess the starting values. The set of guess values that results in the smallest residual sum of squares is chosen as the best set of initial values.

The Gauss–Newton method with the halving mechanism, which is a strategic modification of the classic Gauss–Newton method, is used to minimize the sum of squares. Since this method encounters difficulties in some instances, the Marquart procedure [10] is also used to improve the convergence.

2.3. Determination of the order \((n, m)\)

The question of what are the values of \( n \) and \( m \) before fitting a model is very difficult. Box and Jenkins provided some empirical guidelines for the determination of \( n \) and \( m \) when one of them is zero [11]. It has also been shown that any stationary stochastic system can be approximated as closely as required by an ARMA model of order \((n, n - 1)\) [12]. Consequently, the question of determining \((n, m)\) becomes that of determining \( n \). The basic procedure in [12], which is based on the F-criterion or F-test, is adopted in this paper to determine the value of \( n \) which provides the best fit of the wind speed time series model given by eqn (2). The main steps of this procedure are:

Step 1. Let \( n = 2 \); fit the ARMA\((n, n - 1)\) model using the approach outlined in Section 2.2, calculate the residual sum of squares of the model and designate it as \( \text{RSS}(n, n - 1) \).

Step 2. Fit the ARMA\((n + 1, n)\) model and calculate the residual sum of square \( \text{RSS}(n + 1, n) \) using the same approach as above.

Step 3. Let

\[ F = \frac{\text{RSS}(n, n - 1) - \text{RSS}(n + 1, n)}{\text{RSS}(n + 1, n)} \]

in which \( N \) is the total number of observations and \( r = 2n + 2 \). Perform the following comparisons using the value of \( F \):

1. If \( F > F_p(2, N - r) \), where \( F_p(2, N - r) \) denotes the \( F \)-distribution with 2 and \( N - r \) degrees of freedom at probability level \( p \), then the improvement in the residual sum of squares in going from ARMA\((n, n - 1)\) to ARMA\((n + 1, n)\) is significant at the \((1 - p) \times 100\% \) significance level and therefore there is evidence that the ARMA\((n, n - 1)\) model is inadequate; go to Step 4.

2. If \( F < F_p(2, N - r) \), then the ARMA\((n, n - 1)\) model is adequate at the level of significance, go to Step 5.
Step 4. Set \( n + 1 \rightarrow n \), go to Step 2.

Step 5. Fit a pure AR\((n)\) model and use the
F-criterion to check the adequacy of the model AR\((n)\).
If it is adequate, AR\((n)\) can be used as a possible
substitute model for ARMA\((n, n - 1)\); if it is not
adequate, fit the desired forms of models AR\((n')\) where
\( n' > n \) until an insignificant \( F \) value is reached.
The last AR\((n')\) model can be used as a possible substitute
for ARMA\((n, n - 1)\).

2.4. Diagnostic checking

The procedure for calculating the order \((n, m)\) as
given in Section 2.3 determines the adequacy of the
fitted model from a mathematical point of view.
However, as a precautionary measure, additional
diagnostic checking is needed.

If a fitted ARMA\((n, m)\) model is adequate, \( \{x_t\} \)
should be uncorrelated and normally distributed.
There are different approaches to check the independence
of \( \{x_t\} \) by ensuring that its auto-
correlations are small, say within the permissible band
\( (\pm 2/\sqrt{N}) \) or the more precise Bartlett band.
Another alternative involves using the 'portmanteau' test or
'statistic Q' suggested by Box and Jenkins [11].
If \( Q \) is less than \( x^2(K - n - m) \) at an appropriate
probability level, the \( \{x_t\} \) of the ARMA\((n, m)\) can be
considered as independent. \( K \) should be large enough
so that the Green function \( G_j [12] \) is particularly zero
for \( j \geq K \).

Diagnostic checking from a mathematical point of
view is necessary but not sufficient to determine
whether a wind speed model is feasible or not. A
simulation procedure is further needed to check
whether a wind speed model can retain the main
characteristics of wind speed or not. In order to do
so, the statistical properties of the simulated wind
speed should be compared with those obtained
from the actual wind speed. The main statistical
properties to be compared are the auto-correlation
function, the mean and standard deviation of wind
speed, the seasonal property, the diurnal distribution
and so on.

2.5. Programs

Two computer programs designated as WSERIES
and SWIND have been developed at the University
of Saskatchewan based on the principles outlined in
Sections 2.1–2.4. The first program WSERIES is used
to establish the time series model of wind speed and
check the feasibility of the model from the mathematical point of view. The second program SWIND
is used to simulate the wind speed according to the
established model, thus providing the statistical
properties of the simulated wind speed in order to
compare them with those obtained from the actual
wind speed when determining the feasibility of the
fitted model.

3. WIND SPEED MODELS: TYPE ONE

The actual hourly wind speed for 3 years (from
1 January 1991 to 31 December 1993) and the
hourly mean and standard deviation of wind speeds
from a 37-year database (from 1 January 1953 to 31
December 1989) for a site near North Battleford,
Saskatchewan were obtained from Environment
Canada and used to illustrate this type of wind speed
model.

Let
\[
y_t = (\bar{W}_t - \mu_t) / \sigma_t.
\]

Then \( y_t \) can be used to establish the wind speed model
eqn (2) and the simulated wind speed \( SW_t \) can be
calculated as:
\[
SW_t = \mu_t + \sigma_t \cdot y_t.
\]

Figure 1 presents the auto-correlation functions of
wind speed in years 1991 and 1993. It can be seen

![Fig. 1. Comparison of the auto-correlation functions of wind speed between 1991 and 1993 (site: North Battleford).](image-url)
from this figure that the auto-correlation functions of different years are different, thus the time series models based on one year of actual hourly data may cause somewhat approximate results. The complete three-year record of hourly actual wind speed was therefore adopted in this study.

3.1. ARMA(3,2) model

The program WSERIES was used to establish several models. The first model generated was an ARMA(2,1) model:

\[ y_t = 0.9204y_{t-1} - 0.0081y_{t-2} + x_t - 0.2227x_{t-1} \]
\[ \text{RSS}(2,1) = 5927.595. \quad (6) \]

The second model created was an ARMA(3,2) model, which can be written as:

\[ y_t = 1.7901y_{t-1} - 0.9087y_{t-2} + 0.0948y_{t-3} + x_t - 1.0929x_{t-1} + 0.2892x_{t-2} \]
\[ x_t \sim \text{NID}(0, 0.474762^2) \text{ and } \text{RSS}(3,2) = 5922.804. \quad (7) \]

The residual sum of the squares of the ARMA(3,2) model is smaller than that of the ARMA(2,1) model. The F-criterion shows that:

\[ F = \frac{5927.595 - 5922.804}{5922.804} = 10.63 > F_{0.95}(2, \infty) = 3.00. \]

As the computed F-values is less than the F-distribution value at a 5% level of significance, the ARMA(3,2) model as expressed by eqn (7) can be considered adequate for the given wind speed data. The ARMA(3,2) model was obtained by minimizing the sum of the squares of the difference between the actual wind speed and those given by the model. The distribution of these residuals is given in Fig. 2 which shows how well the residuals satisfy the characteristics of the normal distribution. Figure 3 presents the first 50 auto-correlations for the residuals. It can be seen from this figure that about 96% of the auto-correlations are in the permissible band \( \pm 0.012337 \), thus \( \{x_t\} \) can be roughly taken as independent. An additional test was conducted using the Q-statistic and the results are given in Table 1. It is clear from Table 1 that the Q values are always smaller than the critical values at the 95% probability level for different \( k \). Therefore \( \{x_t\} \) can be considered an independent stochastic variable. The analyses on the normal distribution and independence of \( \{x_t\} \) give the supplementary mathematical support for the suitability of the ARMA(3,2) model.

In addition to the mathematical checking procedure described above, a complete validation of the ARMA(3,2) model should include an analysis of the wind speed generated by the model. The program SWIND was used to generate 38 years of wind data. Several characteristics of the simulated wind speed
are compared with those of the observed wind speed as follows:

(1) The observed average wind speed is 14.62 km h⁻¹, and the simulated value is 14.84 km h⁻¹.

(2) Figure 4 presents a comparison of the auto-correlations of the actual wind speed with those of the simulated wind speed. It can be seen from this figure that the forms of the observed and simulated auto-correlation functions are almost the same including the superimposed sinusoidal damping which reflects the diurnal cycle.

(3) Figure 5 shows the observed and simulated seasonal distributions of wind speed. A comparison of these distributions indicates generally good agreement.

(4) The observed and simulated diurnal distributions of wind speed in August are randomly selected and listed in Fig. 6 which shows a relatively close consistency.

Other properties were also compared and the results did not show any significant difference. It can be concluded from these comparisons that the ARMA(3,2) model proposed in this paper can quite closely reproduce the auto-correlation of hourly wind speed, the seasonal characteristics and the diurnal distribution of wind speed, and therefore can be used as a feasible time series model for reliability evaluation of power systems including WECS at that site.

3.2. Substitute models for ARMA(3,2)

As stated in Section 2, some pure AR models can be established as possible replacements for the fitted
ARMA model. Generally, AR models, because of their simplicity and ease of interpretation, are often used in practical applications.

The first possible substitute model for ARMA(3,2) is AR(3). The residual sum of the squares of the AR(3) model is 5926.74. The F-value resulting from substituting ARMA(3,2) with AR(3) is 8.73, which is much greater than $F_{0.05}(2, \infty)$. This means that the reduction in the residual sum of the squares from AR(3) to ARMA(3,2) is substantial at this significance level. The AR(3) model is therefore rejected and is not an acceptable substitute for the ARMA(3,2) model.

By increasing the order of AR(n), AR(4)–AR(8) can also be formed. The established AR(8) model has almost the same residual sum of the squares as the ARMA(3,2) model. A similar checking procedure was used for the AR(8) model and the results show that it is feasible and can be considered as a reasonable substitute model for ARMA(3,2). The fitted AR(8) is as follows:

$$y_t = 0.6971y_{t-1} + 0.1441y_{t-2} + 0.0464y_{t-3} + 0.01y_{t-4} + 0.0038y_{t-5} - 0.0194y_{t-6} - 0.0068y_{t-7} - 0.0117y_{t-8} + z_t$$

$z_t \sim \text{NID}(0, 0.474801^2)$. (9)

In order to investigate whether there is a much better time-series model than AR(3,2) or AR(8), a complete AR(24) (not the simplified form as presented in [6, 7]) was developed. The results show that the improvement created by the AR(24) model is very slight. The main improvement is in the autocorrelation function of wind speed. The AR(24) model reproduces damping more closely than the AR(3,2) does. As shown in Fig. 4, the AR(3,2) model has almost the same damping as that of the actual wind speed, therefore the more detailed AR(24) model is not
necessary unless more precise simulation procedures are required in some specific applications.

3.3. Comparison of the models for different sites

The wind data for a site near Regina, Saskatchewan which has relatively high wind speed, were also obtained from Environment Canada. A series of wind speed models were established by WSERIES. The ARMA(2,1), ARMA(3,2) were rejected according to the F-criterion. The acceptable adequate model for the Regina site is ARMA(4,3) and can be written as:

\[ y_t = 0.9336y_{t-1} + 0.4506y_{t-2} - 0.5545y_{t-3} + 0.111y_{t-4} + \varepsilon_t - 0.2033\varepsilon_{t-1} - 0.4684\varepsilon_{t-2} + 0.2301\varepsilon_{t-3} \]
\[ \varepsilon_t \sim \text{NID}(0, 0.409442^2). \] (10)

The AR(8) model developed as the substitute model for ARMA(4,3) is as follows:

\[ y_t = 0.7306y_{t-1} + 0.1294y_{t-2} + 0.0475y_{t-3} + 0.0105y_{t-4} - 0.0103y_{t-5} + 0.0047y_{t-6} - 0.0114y_{t-7} - 0.0072y_{t-8} + \varepsilon_t \]
\[ \varepsilon_t \sim \text{NID}(0, 0.409432^2). \] (11)

The above wind speed models for the site near Regina are different from those at the North Battleford; not only in the coefficients, but also in the orders of the models. New time series models should be established using the same procedure for any given location.

4. WIND SPEED MODELS: TYPE TWO

The hourly mean and standard deviation of the wind speed are needed to establish the previously described models. For some sites, however, such records are unavailable or inadequate and therefore a different type of model is required.

Data were provided by SaskPower on a site near Billimun [13]. Only one year of hourly actual wind speed (from 1 August 1993 to 31 July 1994) was available.

Let:

\[ y_t = OW_t - \mu. \] (12)

Then \( y_t \) can be used to establish wind speed model (2), and the simulated wind speed, \( SW_t \), can be calculated as:

\[ SW_t = \mu + y_t. \] (13)

The finally fitted wind speed model for Billimun is ARMA(3,2) as follows:

\[ y_t = 1.4541y_{t-1} - 0.3756y_{t-2} - 0.1116y_{t-3} + \varepsilon_t - 0.3806\varepsilon_{t-1} - 0.1988\varepsilon_{t-2} \]
\[ \varepsilon_t \sim \text{NID}(0, 1.301342^2). \] (14)

The substitute model is AR(4):

\[ y_t = 1.0738y_{t-1} - 0.1688y_{t-2} + 0.041y_{t-3} + 0.0221y_{t-4} + \varepsilon_t \]
\[ \varepsilon_t \sim \text{NID}(0, 1.301547^2). \] (15)

As shown in Fig. 7, underestimation of the auto-correlation function occurs only when the lag is greater than 20 and the fitted ARMA(3,2) model (14) can basically preserve the auto-correlation of the actual wind at Billimun. It should be noted that there is no superimposed sinusoid damping in the auto-correlation function of wind speed at Billimun.

Although the time series model given by (14) passes the statistical tests and fundamentally reproduces the auto-correlation of the actual wind speed, it unfortunately can not retain the seasonal and diurnal distribution of the actual wind speed as shown in Figs 8 and 9, respectively. It can be clearly seen from these figures that the simulated results are significantly different from the observed ones.

Fig. 7. Observed and simulated auto-correlation functions of wind speed at Billimun.
Some improvement can be achieved by dividing the wind data into 12 months and establishing 12 time series models. In this way, the seasonal distribution of wind speed can be retained, but the diurnal distribution cannot be preserved. It can therefore be concluded that the second type of models is not adequate as the wind data used to establish the model are not sufficient. Generally, the more wind data are utilized in developing the model, the more accurate the model is.

5. CONCLUSION

A new procedure for fitting time-series wind speed models is presented in this paper. Two different time-series models are established using different available wind speed data. No assumptions or previously estimated factors are introduced in the models. This approach ensures that there is no inherent distortion in the resulting model. The F-criterion, statistical tests based on the chi-square distribution, and a simulation procedure are used to check the feasibility of the model. The first series of models can pass the statistical tests and reproduce the high-order auto-correlation, the seasonal and diurnal distribution of the actual wind speed and therefore can be used in reliability studies of power system including WECS. The second series of models can not persevere the statistical seasonal characteristics or diurnal distribution as the wind data used to establish them are not sufficient. Availability of actual wind data in sufficient detail is an essential requirement for developing feasible wind speed models.

The studies in this paper also show that the sampling auto-correlation functions of different years at the same site may be significantly different, thus a wind speed model based on only one year of actual wind data should be used with caution.
REFERENCES