ARIMA-Based Frequency-Decomposed Modeling of Wind Speed Time Series

Kalid Yunus, *Student Member, IEEE*, Torbjörn Thiringer, *Member, IEEE*, and Peiyuan Chen, *Member, IEEE*

***Abstract*—In this paper, a modified auto regressive integrated moving average (ARIMA) modeling procedure that can capture time correlation and probability distribution of observed wind-speed time-series data is presented. The procedure introduces frequency decomposition (splitting the wind-speed data into high frequency (HF) and low-frequency (LF) components), shifting, and limiting in addition to differencing and power transformation which are used in the standard ARIMA modeling procedure. The modified modeling procedure is applied to model 10 minute average measured wind-speed data from three locations in the Baltic Sea area and the results show that the procedure can capture time correlation and probability distribution of the data. In addition, it is shown that, for 10-min average wind-speed data in the Baltic Sea area, it could be sufficient to use ARIMA(6,0,0) and ARIMA(0,1,6) to model the HF and LF components of the data, respectively. It is also shown that, in the Baltic Sea area, a model developed for an observed wind-speed data at one location could be used to simulate wind-speed data at a nearby location where only the average wind-speed is known.**

***Index Terms*—Auto correlation coefficient (ACC), auto regressive integrated moving average (ARIMA), CDF, partial auto correlation coefficient (PACC), PDF, Q-Q plot, time-series model, wind power, wind speed.**

I. Introduction

WIND power is getting increased public support and government incentives because of its socio-environmental benefits. As a result, it is the most developed and commercially utilized nonhydro renewable energy resource. It is also among the fastest growing producer of electric energy in the world. As a result, the cumulative global installed wind power capacity by the end of 2012 was 283 GW and has continued to grow at a rate of approximately 20% annually [1]. On the other hand, the impacts of the growing integration of wind power into the existing power systems are also significant, which in some areas have resulted in limiting the growth of wind power integration [2], [3]. In order to analyze the impacts of stochastic wind power generation on power system reliability and long-term planning, a probabilistic approach based on sequential Monte Carlo simulation is usually adopted [4], [5]. In a sequential Monte Carlo simulation, it is important to capture the probability distribution and chronological characteristics of wind power generation, load profiles, and transition states of all of the system components [4], [5]. Thus, for this purpose, it is of great importance to capture the probability distribution and temporal correlation of a wind power time series accurately in a wind power model.

Developing a wind power model requires measured wind power data which can only be obtained if wind power plants (WPP) are already installed in the area of interest, which is not always the case. Even in areas where there are WPPs, availability of the measured data for public use are limited since the data are often confidential for the WPPs operators. Hence, the most suitable way to obtain wind power data is by using a wind-speed model that can be used to simulate wind-speed data. Wind power can be simulated by transforming the wind-speed data through a power curve [4], [6].

Wind-speed time-series data are stochastic in nature, and some of their important characteristics are strong time correlation, probability distribution, and diurnal and seasonal periodic properties. Hence, a good wind-speed model should be able to capture these properties. There are different ways of modeling wind-speed data. Wind speed data can be modeled by using the Weibull distribution [7]–[10]. This method captures only probability distribution, not temporal correlation of the wind-speed data. Wind speed can also be modeled by reverse transformation of the power spectral density (PSD) [11], [12]. The PSD method is used to capture different frequency components of the wind-speed data and does not consider the temporal correlation in the wind-speed data. Wind speed can also be modeled using the discrete Markov model which can capture probability distribution and time correlation of the data [13]–[15]. However, Markov's model has three major problems. First, it has quantization error. Second, it requires many model parameters. Third, it requires a large set of time-series data [16]. Another method to simulate wind-speed data, which is used in this article, is the ARIMA modeling procedure [17]–[21]. Similar to Markov's model, the ARIMA modeling can capture both probability distribution and time correlation of wind-speed data. Unlike Markov's model, the ARIMA modeling does not have quantization error. ARIMA modeling results in a lower number of model parameters, and it can be used with a smaller set of data [16].

The ARIMA modeling procedure is quite classical and is extensively studied in the literature [4], [6], [22]–[24]. As a result, a standard ARIMA-based modeling procedure was recently published [25]. However, an important finding in this article is that the standard modeling procedure can hardly be used directly to model 10-min average wind-speed time-series data. This is

Manuscript received June 30, 2014; revised December 08, 2014, March 28, 2015, June 30, 2015, and August 05, 2015; accepted August 08, 2015. This work was supported by the Swedish Energy Agency. Paper no. TPWRS-00887-2014.

The authors are with the Department of Electric Power Engineering, Chalmers University of Technology, Gothenburg 412 96, Sweden (e-mail: kalid.yunus@chalmers.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2015.2468586

because the standard modeling procedure is designed to capture the first- and second-order moments of a linear stationary time series. The procedure is not designed to capture the whole joint probability distribution of a nonlinear, nonstationary time series such as wind-speed data. In order to analyze the impacts of stochastic wind power generation on power system, an accurate modeling of the probability distribution and temporal correlation of the wind-speed data is essential. Hence, the main purpose and contribution of this paper is threefold. First, the purpose is to present to the modified ARIMA modeling procedure that is able to simulate 10-min average wind. The main novelty compared with previous efforts is the use of decomposition to split the wind-speed data into a HF and LF parts. Second, the purpose is to present a generic ARIMA model structure that can be used to simulate 10-min average wind-speed data in the Baltic Sea area which can avoid the effort needed to determine model structure of a similar wind-speed data the area. Finally, the purpose is also to present the model structure and the corresponding model parameters that can be used to simulate wind-speed data in the Baltic Sea area with a specific average value and reasonable characteristics in case where measured wind-speed data is not available.

II. Standard ARIMA Modeling Procedure

Here, a summary of the standard ARIMA modeling procedure is presented. Interested readers are referred to [25] and [26] for more detailed information.

A. Standard ARIMA Model

Let $y(1), y(2), \ldots y(N)$ be an observed time-series data which is modeled by a random process $Y(t)$. The standard ARIMA model of the process $Y(t)$, which is often represented as ARIMA (p, d, q), is expressed as

$$\left(1-\sum_{i=1}^{p}\phi_i B^i\right)(1-B)^d Y(t) = \theta_0 + \left(1-\sum_{i=1}^{q}\theta_i B^i\right)a(t) \tag{1}$$

where p is the order and ϕ_i is the coefficient of the auto regressive (AR) component of ARIMA, B is a back shift operator where $B^i Y(t) = Y(t-i)$; d is a degree of differencing which is used to transform data with time dependent mean to data with time independent mean; $Y(t)$ is the time-series model; θ_0 is referred to as deterministic trend term; q is the order and θ_i is the coefficient of the MA(Moving Average) component of ARIMA and $a(t)$ is a white Gaussian noise with zero mean and variance of σ_a^2 [25], [26].

For a time-series data with time dependent variance, the Box-Cox's power transformation can be used to stabilize the variance [25], [26]. The criteria for the selection of the transformation factor (v) is to find the value of v that minimizes the residual sum of squares which is calculated as

$$SS(v) = \sum_{t=1}^{N}\left(T\left(y(t)\right) - \widehat{m}_y\right)^2 \tag{2}$$

where $T(y(t)) = (y^v(t)-1)/vG^{v-1}$, $G = (\prod_{t=1}^{N} y(t))^{1/N}$

B. Model Identification and Diagnostics Checking

In a time-series analysis, the most crucial step is to identify an appropriate structure of the ARIMA model for a given time-series data. Model identification can be done in three steps [25], [26]. The first step is to analyze the observed time-series data and determine the required type of transformation (differencing and/or power transformation, if necessary). The second step is to identify the corresponding ARIMA model structure by comparing the estimated auto correlation coefficient (ACC) and partial auto correlation coefficient (PACC) of the transformed data with the theoretical ACC and PACC (available in [25] and [26]). For the given time-series data $y(t)$, the estimated ACC at time lag k, $\hat{\rho}_y(k)$ can be calculated as

$$\hat{\rho}_y(k) = \frac{\frac{1}{N-k}\sum_{t=1}^{N-k}\left(y(t)-\overline{y}\right)\left(y(t+k)-\overline{y}\right)}{\sigma_y^2} \tag{3}$$

where $\overline{y}$ and σ_y^2 are the mean and variance of the data respectively and N is the length of the data [25], [26]. Similarly, given the estimated sample ACC $(\hat{\rho}_k)$, the estimated PACC of the data can be calculated as

$$\hat{\Phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k}\hat{\Phi}_{kj}\hat{\rho}_{k+1-j}}{1-\sum_{j=1}^{k}\hat{\Phi}_{kj}\hat{\rho}_j} \tag{4}$$

where $\hat{\Phi}_{k+1,j} = \hat{\Phi}_{kj} - \hat{\Phi}_{k+1,j+1}\hat{\Phi}_{k,k+1-j}$. The third and last step is to determine the deterministic trend term θ_0 such that the mean of the simulated time-series data is equal to the mean of the observed time-series data.

Once the model structure is identified, the model parameters can be estimated by using the maximum likely hood estimator. A diagnostic check is then made on the resulting model. This can be done by using the ACC and PACC of the residual[1] where both have to be within the critical limits for time lags different from zero. The values $\pm 2\sqrt{\mathrm{var}[\hat{\rho}_k]}$ and $\pm 2\sqrt{\mathrm{var}[\hat{\Phi}_{kk}]}$ can be used as a critical limits on $\hat{\rho}_k$ and $\hat{\Phi}_{kk}$ of the residual to test the hypothesis that the residual is a white noise [25], [26]

C. Verification and Final Diagnostic Checking

To justify that the identified model is the right model, it has to pass the final diagnostic check. The final diagnostic check involves comparing the simulated time-series data against measurement. The relevant parameters to be compared at this stage are the mean, variance, time correlation, and probability distribution of the simulated and the measured time-series data. If the simulated and the measured data are not the same, the deterministic trend term, transformation factors, and the identified model structure need to be updated until the comparison of the simulated time-series data against the measurement is acceptable. Observe that the main difference between the diagnostic checking discussed in Section II-B and the final diagnostic checking described in this section is that, while the first

[1]The residual is the value of $a(t)$ in (1) for a given time-series data $y(t)$ and the estimated model parameters

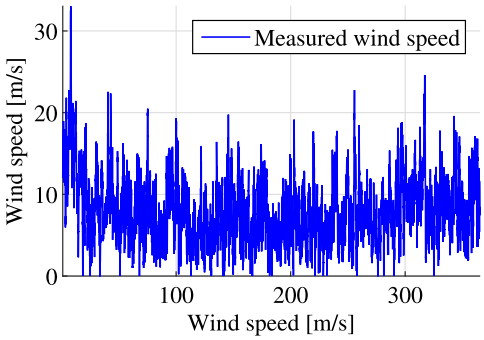


Fig. 1. Measured wind-speed data from Näsudden over a year.

TABLE I
BOX-COX POWER TRANSFORMATION AND LOGARITHM OF THE RESIDUAL SUM OF SQUARES OF THE TIME-SERIES WIND-SPEED DATA

Transformation factor (v)	Transformation	$\lg(SS(v))$
-2	$y(t)^{-2}$	16.1604
-1	$y(t)^{-1}$	14.1791
-0.5	$\left(\sqrt{y(t)}\right)^{-1}$	13.7081
0	$\lg(y(t))$	13.4937
0.5	$\sqrt{y(t)}$	13.4682
1	$y(t)$	13.5948
2	$y(t)^2$	14.2430

deals with the transformed time-series data, the latter deals with the inverse transformed time-series data. Observe also that both the diagnostic checking can be automated. The flowchart and the description of the procedures described above along with the modification made to the standard ARIMA modeling procedure is presented in the following section.

III. ARIMA-BASED WIND-SPEED MODEL

Here, the inadequacy of the standard ARIMA modeling procedure is demonstrated by using a 10-min average measured wind-speed data shown in Fig. 1. The data are obtained from the Näsudden peninsula in Gotland, Sweden, where the measurement was made during the period between the 1st of January to 31st of December 2005 at a height of 100 m. As can be seen from the figure (as can also be seen the ACC and PACC of the data not shown), the measured wind-speed data is highly non-stationary (both in mean and variance), which requires proper transformation before modeling.

Table I shows some values of the transformation factors v, the corresponding transformation, and the residual sum of squares of the time-series data. Observe that the transformation factor $v = 0.5$results in the minimum residual sum of squares. After transforming the wind-speed data using $d = 1$ and $v = 0.5$, ARIMA(2,1,2) emerged from model identification process. Note that the transformed wind-speed data is assumed to be stationary and linear.

The results in Fig. 2 shows that, although the transformation factor $v = 0.5$ results in the minimum residual sum of squares; sensitivity analysis of the simulated wind-speed data to different values of the transformation factors indicates that

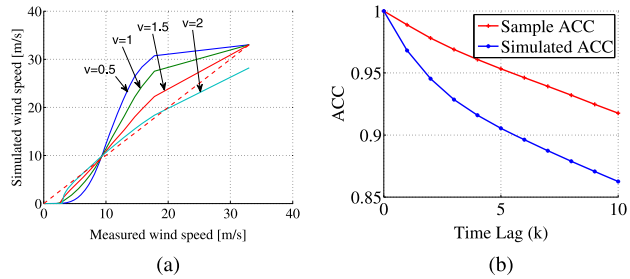


Fig. 2. ARIMA(2,1,2). (a) Q-Q plot a function of v. (b) ACC plot for $v = 2.1$

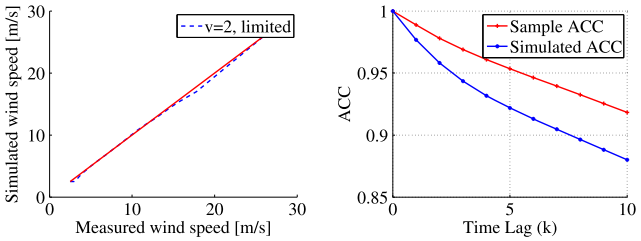
the factor $v = 2$ produces relatively the best simulation result. Fig. 2(a) shows the quantile-quantile (Q-Q)[2] plot of the simulated wind speed against the measurement for different transformation factors. Fig. 2(b) shows the ACC of the measured and the simulated wind-speed data for a transformation factor of $v = 2$. The results in Fig. 2 indicate that, for the 10-min average wind-speed data considered in this article, the standard ARIMA modeling procedure does not produce an acceptable match in time correlation (ACC) and probability distribution (Q-Q plot) compared with the measurement. This requires a modification of the standard ARIMA modeling procedure to produce a good simulation result, which is the main contribution of this paper and is discussed in the following sections.

IV. MODIFIED ARIMA MODELING PROCEDURE

A. Limiting the Wind Speed Before Modeling

The mismatch in the Q-Q plots of the measured and the simulated wind speed, shown in Fig. 2(a) at lower and higher wind speeds for $v = 2$, might be due the fact that the transformation factor does not produce a proportional weight at all wind-speed ranges. Note that the maximum wind speed is about 33 m/s and that wind speeds above 26 m/s occur only for about 0.1% of the year. Note also that wind speeds below 2 m/s occur for about 1.9% of the year. In addition, observe that there is no power production when wind speed is above 26 m/s and below 2 m/s. On the other hand, the wind speed between 2 to 26 m/s occur for about 98% of the time. A sensitivity analysis of the model to the range of the wind-speed data showed that by limiting the wind speed in the range of [2.5, 26] m/s before modeling (i.e. the wind-speed data above and below the limits are set to the maximum and minimum respectively), better result can be achieved, as can be seen Fig. 3. Fig. 3(a) shows the Q-Q plot of the measured and the simulated wind speed. Note that the variance of the simulated wind speed in this case is 14.19 m^2/s^2 compared with 14.72 m^2/s^2 of the limited measured wind speed, while the mean of the simulated and the measured wind speed are the same (8.01 m/s). Fig. 3(b) shows the ACC of the measured $(+)$ and the simulated $(*)$ wind speed. Although limiting the wind speed within the indicated range and then modeling it produces a better match in the probability distribution, the match in the time correlation is still poor, only a small improvement is achieved,

[2]The Q-Q plot is a graphical method of comparing two probability distributions. If the two distribution being compared are similar, the Q-Q plot lies on the straight line (y = x).



(a) (b)

Fig. 3. ARIMA(2,1,2) with limited wind-speed data between [2.5, 26] m/s. (a) Q-Q plot. (b) ACC plot of the measured and simulated wind-speed data.

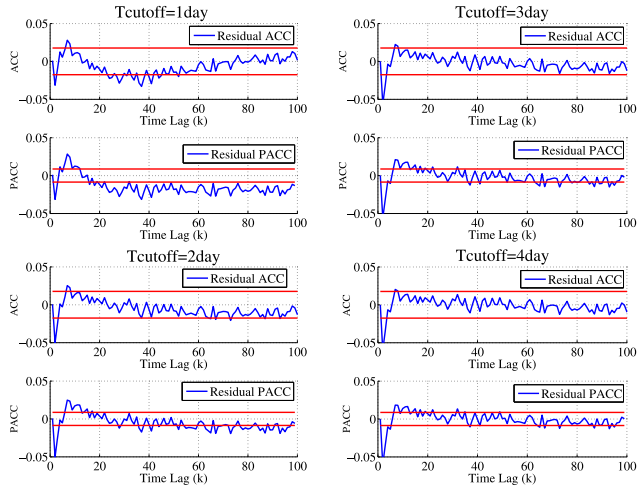
as can be seen in Fig. 3(b). Note also ACC is a measure of the periodic characteristics of wind-speed time-series data. Another way of looking at the periodic characteristics is by using Power Spectral Density (PSD) of the data. According to [25], [26], there is a direct relationship between PSD and ACC of the time-series data. Hence, a good match in the ACC of the time-series data imply that there is a good match in PSD and hence a good match in periodicity of the measured and the simulated time-series data. Although not evident from the result in Fig. 3(b) due to the small window of time lag (k $=$ 10 corresponding to 100 minutes), it is observed that the ARIMA modeling procedures considered so far poorly capture time correlation (chronological characteristics) and periodic characteristics of the measured wind-speed data. In order to address these issues, the concept of frequency decomposition is introduced into the standard ARIMA modeling procedure in the following subsection.

B. Frequency Decomposition

The proposed frequency decomposition-based approach involves first splitting the wind-speed data into high-frequency (HF) and low-frequency (LF) components. The process of frequency decomposition involves mainly four steps. The first step is determining the cutoff frequency (F_{cutoff}) which is used to split the observed wind-speed data into HF and LF components. The second and the third steps are modeling the HF and LF components of the wind-speed data. The final and the fourth step is combining the simulation results from the HF and LF models and comparing it with the observed wind-speed data. Each of these steps are discussed in the following subsections.

1) Determining the Cut-Off Frequency (F_{cutoff}): As stated before, wind-speed time-series data has nonstationary characteristics. According to the study in [27], the major contributor to the nonstationary characteristics of the wind-speed data is the LF component of the wind speed. Based on this fact, the HF component of a wind-speed dataset can be considered stationary. Hence, the F_{cutoff} can then be chosen in such a way that the HF component is stationary. Note that $F_{\text{cutoff}} = 1/T_{\text{cutoff}}$, where T_{cutoff} is the cutoff period.

For a time-series data to be stationary, the ACC and PACC of the residual data need to have the characteristics of a white noise, i.e., the ACC and PACC of the residual has to be within a critical limits for a large time lag. Fig. 4 shows the ACC and PACC of the residual as a function of T_{cutoff} when ARIMA(2,0,0) model structure is used to model the HF component of the observed wind-speed data. Observe that the model structure (ARIMA(2,0,0)) for HF component is obtained by using the initial model identification procedure discussed in Section II-B. As the T_{cutoff} is increased from 1 day to 4 days, the ACC and PACC of the residual became closer and closer to the ACC and PACC of the white noise at a larger time lag. At $T_{\text{cutoff}} =$ 4 day, the ACC and PACC of the residual are within the limiting values for a larger time lags which indicates that the HF component at this T_{cutoff} can be assumed stationary. It is important to notice at this point that the cutoff time ($T_{\text{cutoff}} =$ 4 day) that emerged from this analysis is consistent with the result from the Van der Hoven spectrum [27].



(a) (b)

Fig. 4. Residual ACC and PACC as a function of T_{cutoff} when an ARIMA(2,0,0) model structure is used to model the HF component.

2) Modeling the HF Component: Note that, at $T_{\text{cutoff}} =$ 4 day, the residual ACC and PACC in Fig. 4(b) at a lower time lag are not bounded within the limit. This indicates that the initially identified ARIMA(2,0,0) model for the HF component needs to be modified. Following the diagnostic checking procedure described in Section II-B and II-C, the updated model structure for the HF components of the wind speed is found to be ARIMA(6,0,0). Note that the residual, when ARIMA(6,0,0) is used to model the HF component, is a white noise. The result in Fig. 5 shows the ACC and PACC of the residual when ARIMA(6,0,0) is used to model the HF component of the wind-speed data. Note that the residuals are within the critical limits indicating the sufficiency of the model.

3) Modeling the LF Component: Once the HF component of the wind speed is determined using $T_{\text{cutoff}} =$ 4 day, the LF component can be determined by subtracting the HF component from the observed wind-speed data. As expected, the resulting LF component is non-stationary and the corresponding ARIMA model structure can be developed by following the model identification and diagnostic checking procedures described in Sections II-B and II-C.

Before transforming and then modeling the LF component of the wind-speed data, it is important to resample the data. Note that the sampling frequency (FS) of the measured wind speed is 144 samples per day ($FS =$ 144 cycle/day), which is too high compared with the maximum frequency ($F_{\max}$) in the LF

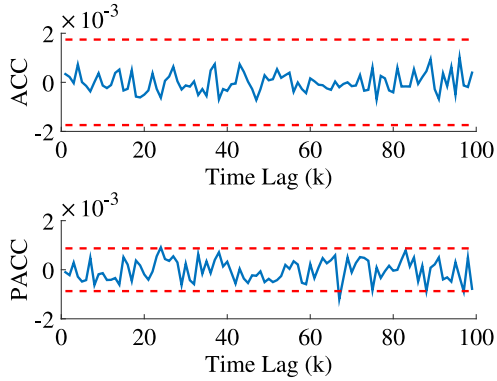


Fig. 5. Residual ACC and PACC of the HF component of the wind-speed data when modeled by using ARIMA(6,0,0).

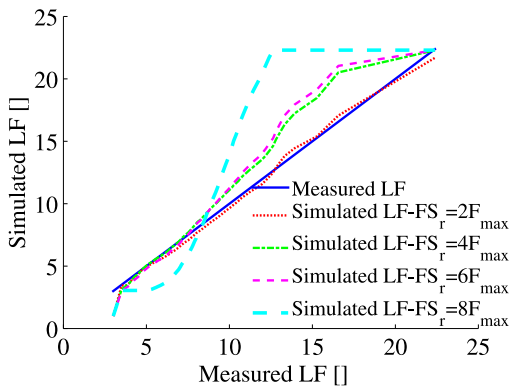


Fig. 6. Effect of the choice of resampling frequency (FS_r) on the results of the LF component of the wind-speed data.

data which is 0.25 cycle per day. Note that $F_{\max}$ is the maximum frequency component in the LF data which is given by $F_{\max} = F_{\text{cutoff}} = 1/T_{\text{cutoff}}$. This makes it difficult to see the temporal correlation of the data and hence the ACC and PACC of the data from which the initial model identification can be made. According to the Nyquist sampling criteria [28], reconstruction of a signal is possible when the sampling frequency is greater than twice the maximum frequency of the signal being sampled. In order to choose the right value of the frequency to resample the LF component of the wind-speed data, the effects of different re-sampling frequencies on the characteristics of the LF component is analysed. Hence, re-sampling frequencies of $2 * F_{\max}$, $4 * F_{\max}$, $6 * F_{\max}$, $8 * F_{\max}$ are considered in this case, which corresponds to $FS_r = 0.5$ cycle/day, $FS_r = 1$ cycle/day, $FS_r = 1.5$ cycle/day, $FS_r = 2$ cycle/day, respectively. Fig. 6 shows how different resampling frequencies alter the CDF (Cumulative Distribution Function) of the LF component of the wind-speed data. The results in the figure show that the re-sampled LF component with the resampling rate of $FS_r = 2 * F_{\max}$ better matches the original LF component of the wind-speed data.

The resampled LF component of the observed wind-speed data is then transformed by using a transformation factor of $v = 0$ (logarithmic transformation) followed by one degree of differencing ($d = 1$). Using the modeling procedure described in Section II, the resulting time-series data are fit to an

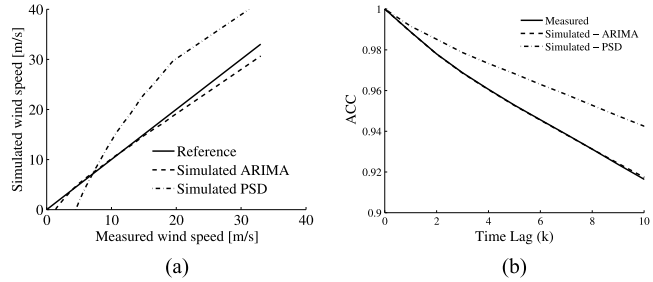


Fig. 7. Measured and simulation results of the wind-speed time-series data. (a) Q-Q plots. (b) ACC plots.

ARIMA(0,1,6) model structure. A lower order of the model introduces more error in the residual at a lower time lag (as shown in Fig. 4 for HF modeling) and a higher order model introduces more parameters which are not desired.

4) Combining the Results From HF and LF Models to Get the Simulated Wind-Speed Data: Once the HF and LF components of the wind-speed time-series data are modeled separately, the last step is to combine the results from the two models to get the complete representation of the simulated wind-speed data. Fig. 7 shows a Q-Q plot and ACC of the observed and the simulated wind-speed time-series data. Observe that the results are for the measured wind-speed data over one year and multiple realization ($\times 100$) of the simulated wind speed. Hence, the simulation results are like an average results. Fig. 7(a) shows Q-Q plots and Fig. 7(b) shows ACC of the observed and the simulated wind-speed data. The results show that the proposed ARIMA modeling procedure produces simulation result with a tight match in ACC and Q-Q plot compared with the measurement. Note that the observed wind-speed data is not limited before modeling; all the data are used in the model to produce the results presented in Fig. 7. Note from Fig. 7(a) that the match in the Q-Q plot (from ARIMA simulation) is not tight at higher wind speed. This is because of the fact that there is little information regarding wind speeds higher than 20 m/s, as can be seen from Fig. 1 to be captured by the model. Extreme wind conditions can be handled separately when needed [29]. The results in Fig. 7 also show that the proposed modeling procedure produces much better results that the inverse PSD method, the claim which is also supported by study in the [12].

Note from Fig. 7(b) that ACC up to ten time lag is presented, which corresponds to approximately 1.7 h. Recall also that the ACC is closely related to periodic characteristics [25], [26]. Hence, from Fig. 7(b), it is difficult to see if the proposed modeling procedure can capture the daily, monthly, and seasonal characteristics of the observed wind-speed data in some way. In order to show the characteristics of the ACC of the observed and the simulated wind-speed data at larger time lags, the results are presented in a better way, as can be seen in Fig. 8. The figure shows the ACC of the observed and the simulated wind speed over a time lag of a month as a function of T_{cutoff}. As cab be seen from the figure, the model does capture the pattern of the observed ACC up to 500 time lags which corresponds to approximately three days. The following values are less significant since the value ACC for high time lag is less than 0.2. Fig. 8 also shows that the choice T_{cutoff} affect the results, and,

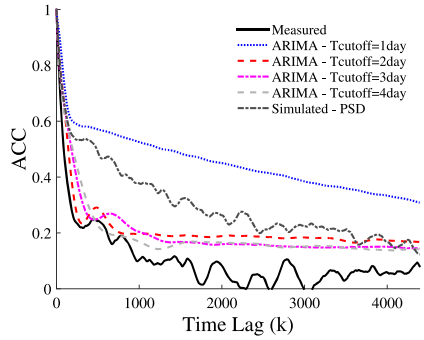


Fig. 8. ACC of measured and simulated wind speed over a month.

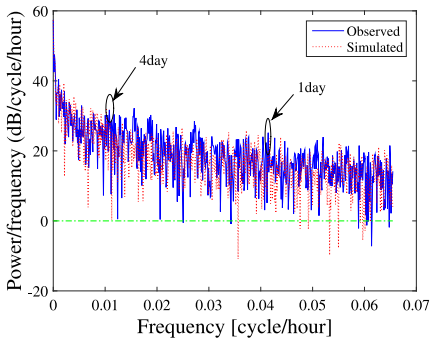


Fig. 9. Periodograms of the measured and the simulated wind-speed data when Tcuff = 4 day is used.

similar to the conclusion made above, the proposed method produce much better results at higher time lag compared with the spectral method.

Although there is relationship between the ACC and the periodicity of the time-series data; Fig. 8 shows it is difficult to see the significant periodic (daily, monthly, seasonal and so on) characteristics the time-series data from the ACC alone. The best way to look at the general trend of the periodic characteristics of a time-series data is to analyze the periodogram of the data. The periodogram is an estimate of the spectral density of a signal. Fig. 9 shows the periodogram of the observed wind-speed data and the simulated wind-speed data when $T_{\text{cutoff}} = 4$ day. Fig. 9 shows that the proposed modeling procedure can capture the general periodic characteristics of the observed wind-speed data. Note from the figure that the dominant periodic characteristics of the wind-speed data descends from higher to lower in a smooth way. Note also that periodicity up to 0.08 cycles per hour (which corresponds to 12.5 h) is presented in the figure. The patter is similar at higher frequency except that the magnitudes of the spectrum are lower.

C. Shifting the Wind-Speed Data

Although not used in the modeling of the observed wind-speed data presented in the preceding sections, shifting (adding a constant offset value to the signal) the observed time-series data before transformation can sometimes improve results, as can be seen in Table IV.

D. Flowchart of the Modified ARIMA-Based Wind-Speed Modeling Procedure

Fig. 10 shows the flowchart of the modified ARIMA-based modeling procedure which consists of three stages referred to as first stage, second stage, and third stage. In the first stage, the procedure starts by taking the time-series data as an input (which can be wind speed) followed by splitting the time-series data into HF and LF components. A nonlinear transformation followed by a linear transformation is applied on each frequency component (when necessary) to transform the resulting data into stationary time-series data.

In the second stage, the initial model identification of the transformed data is made followed by a model parameter estimation. By using the estimated model parameters, diagnostic checking is made on the identified model. If the model does not satisfy the diagnostic checking, the model structure is updated using the information from the diagnostic checking and a new parameter estimation is made. The same process is repeated until the model is justified by the diagnostic checking. If the model satisfies the diagnostic checking, it is used to simulate time-series data, which is the third stage of the modified modeling procedure.

The third stage starts from the transformed simulated time-series data (indicated as $TTZ_{\text{sim}}(t)$). The main task in this stage is the inverse transformation of $TTZ_{\text{sim}}(t)$. In order to make sure that the simulated wind-speed time-series data does not have maximum and minimum values that are not found in the observed wind-speed data, limiters are used during the inverse transformation. Especially when there is an integrator during the inverse transformation (originated from differencing), it is likely that the simulated wind speed could go out of the upper and lower bounds. Hence, the upper and lower bounds of the limiters are set based on the maximum and minimum values in the observed wind-speed data. Note also that the main difference between the limiter used in the first stage and the limiters used in the third stage is that, while the limiter in the first stage is applied on the observed wind-speed data (where the upper and the lower limits of the limiter is decided by the user), the limiters in the third stage are applied on the simulated time-series data (where the upper and lower limits are set based on the measured data). The resulting ARIMA model with a limiter is then termed as Limited-ARIMA, the idea of which is proposed in [16]. Once the simulated wind-speed time-series data are ready, the final diagnostic checking is made on the simulated data by comparing the simulated time-series data against the measurement as described in Section II-C. If the comparison does not result in a good match, the process is repeated with updated values of transformation parameters and model structures until acceptable results are achieved. It is important to point out that the main modifications made on the standard modeling procedure are the introduction of frequency decomposition, shifting, and limitation (before modeling and during simulation), as can be seen in Fig. 10.

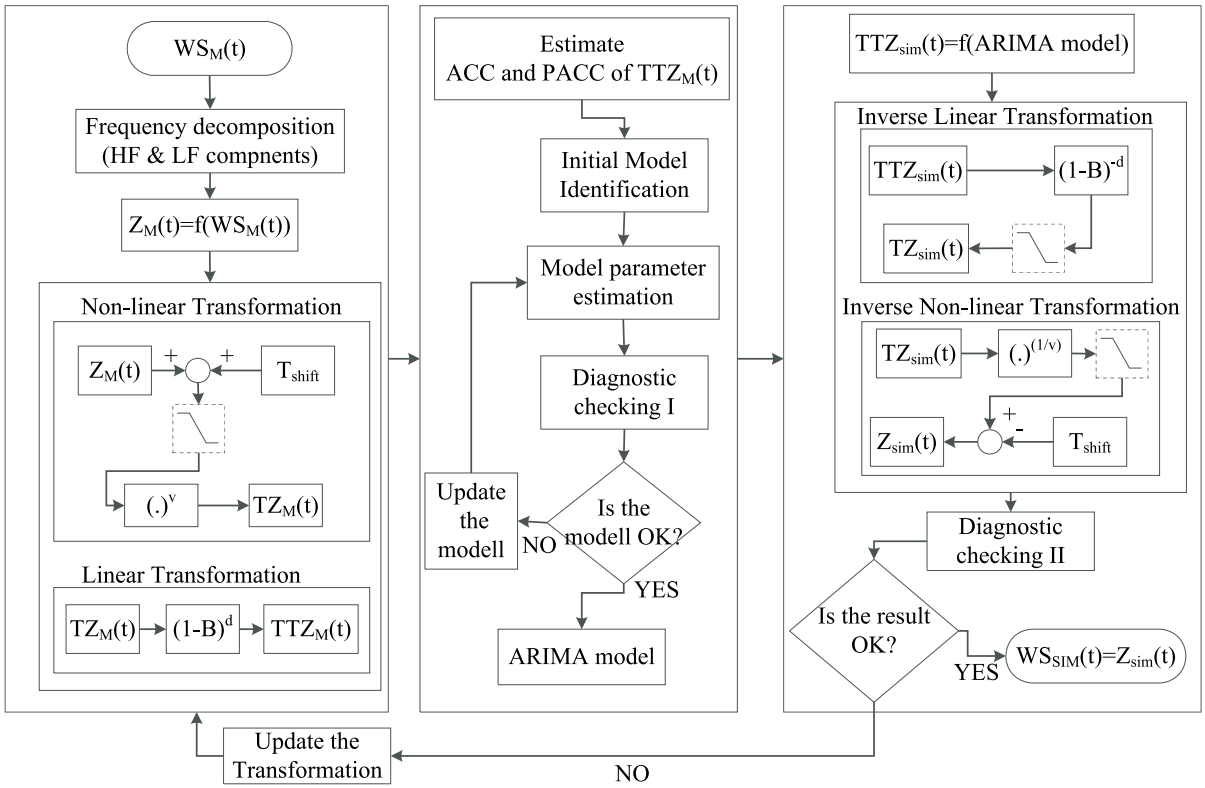


Fig. 10. Flowchart of the modified ARIMA-based modeling procedure.

Fig. 11. Geographical locations of measured wind-speed sites.

V. Application of the Proposed Modeling Procedure

A. ARIMA(6,0,0) for HF and ARIMA(0,1,6) for LF Component of Observed Wind Speed Data in the Baltic Sea Area

As discussed in Section II-B, getting an appropriate model structure for a given time-series data is not an easy task. Hence, it is of great importance to find a way that can be used to avoid the step to determine the model structure for a give time-series data in a particular area. For this purpose, a detailed analysis of wind-speed modeling at different locations over many years is made by using the proposed modeling procedure. In this analysis, some areas in the Baltic Sea region are considered.

Fig. 11 shows the geographical locations of the measured/observed wind-speed sites considered. The distance between the Utgrunden and the Näsudden wind site is about 150 km and the distance between Utgrunden and Tjaereborg is about 500 km.

TABLE II
Average Wind Speed at Different Locations and in Different Years

Average wind speed [m/s]	Näsudden	Utgrunden	Tjaereborg
All year (average)	7.3116	7.4634	8.7548
Year 1	6.6068	7.2836	8.5854
Year 2	7.1334	7.787	8.8471
Year 3	7.3203	7.4938	8.8319
Year 4	7.9113		
Year 5	7.6855		

From each sites, different lengths of 10-min average wind-speed time-series data are available. A five-year time-series data, measured at a height of 100 m was available at Näsudden. A set of three-year data, measured at a height of 38 m, was available from Utgrunden. A set of three-year data, measured at a height of 90 m, was available from Tjaereborg to make the analysis.[3] The average values of the wind-speed data at each locations, over the years, are given in Table II.

In order to show the possibility of using a generic model structure for a given wind-speed time-series data at a given area, the wind-speed data in Year 1 from the three locations are investigated by using the proposed modeling procedure. The results from the investigation showed that the model structures ARIMA(6,0,0) and ARIMA(0,1,6) were sufficient to model the 10-min average wind-speed data from both locations. Table III presents the model parameters of the HF components of the observed wind-speed data at each locations, and Table IV presents the model parameters of the LF components of the observed wind-speed data at each locations. The residual

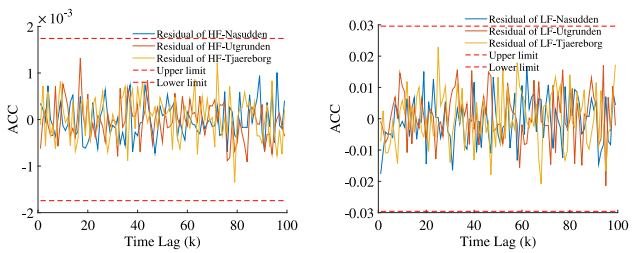
[3][Online]. Available: http.winddata.com

TABLE III
MODEL PARAMETERS OF THE HF COMPONENTS OF WIND-SPEED DATA FROM DIFFERENT LOCATIONS. THE MODEL STRUCTURE IN EACH CASE IS ARIMA(6,0,0)

Data source	Model Parameters
Näsudden HF	$\theta_0 = -0.0001$, $T_{shift} = 0$ $\phi_1 = 0.9662$, $\phi_2 = -0.0571$, $\phi_3 = 0.0232$ $\phi_4 = 0.0245$, $\phi_5 = -0.0030$, $\phi_6 = 0.0210$
Utgrunden HF	$\theta_0 = -0.003$, $T_{shift} = 0$ $\phi_1 = 1.0350$, $\phi_2 = -0.1177$, $\phi_3 = 0.0352$ $\phi_4 = 0.0251$, $\phi_5 = 0.0058$, $\phi_6 = -0.0110$
Tjaereborg HF	$\theta_0 = -0.001$, $T_{shift} = 0$ $\phi_1 = 0.9227$, $\phi_2 = -0.0648$, $\phi_3 = 0.0659$ $\phi_4 = 0.0143$, $\phi_5 = 0.0149$, $\phi_6 = 0.0159$

TABLE IV
MODEL PARAMETERS OF THE LF COMPONENTS OF WIND-SPEED DATA FROM DIFFERENT LOCATIONS. THE MODEL STRUCTURE IN EACH CASE IS ARIMA(0,1,6)

Data source	Model Parameters
Näsudden LF	$\theta_0 = -0.0025$, $T_{shift} = 5$ $\theta_1 = 0.3258$, $\theta_2 = 0.5157$, $\theta_3 = -0.0890$ $\theta_4 = 0.1951$, $\theta_5 = -0.0132$,$\theta_6 = -0.0520$
Utgrunden LF	$\theta_0 = 0.0003$, $T_{shift} = 10$ $\theta_1 = 0.4008$, $\theta_2 = 0.5747$, $\phi_3 = 0.0352$ $\phi_4 = 0.0251$, $\phi_5 = 0.0058$, $\phi_6 = -0.0110$
Tjaereborg LF	$\theta_0 = 0.00025$, $T_{shift} = 5$ $\theta_1 = 0.2029$,$\theta_2 = 0.7413$, $\theta_3 = -0.0986$ $\theta_4 = 0.2570$, $\theta_5 = -0.0433$,$\theta_6 = -0.2016$



(a) (b)

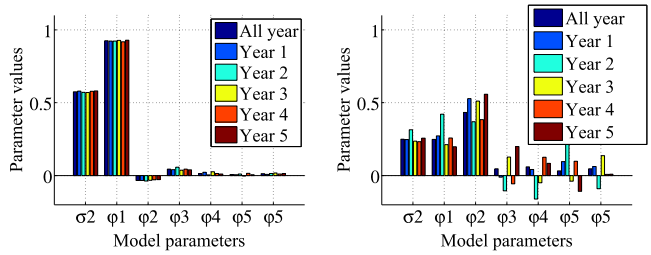
Fig. 12. Residual ACC of the models presented in (a) Table III (b) Table IV.

ACC of the models presented in the tables are presented in Fig. 12. In the process of modeling the wind-speed data from the three locations, it is observed that the use of shifting factors proved to be useful in improving the results and is considered as an added degree of freedom to be exploited in future works. Note that only positive values of the shift factors are considered in this study and the choice is made by trial and error.

Based on the results presented in Section III, Section V-A and the results presented in the following section (Section V-B), it is concluded that the proposed model structures (ARIMA(6,0,0) for HF components and ARIMA(0,1,6) for LF components) could be sufficient to model 10 minute average wind-speed data in the Baltic Sea area. This means that, if one has a 10-min average wind-speed data from the Baltic Sea area (different from the ones considered above), it could be sufficient to use the proposed model structures.

B. Common Model to Simulate Wind Speed at Adjacent Multiple Locations

The wind-speed models developed so far involves measurements. However, measured wind-speed data are not available at all places of interest. Hence, it is of great importance to develop a model that can be used to simulate wind speed data with *acceptable* properties at places where there are no measured wind-speed time-series data. The key to this is to define the term *acceptable* and, in order to define it, it is important to study the variability of model parameters and relevant properties of measured wind-speed data from place to place and from time to time.



(a) (b)

Fig. 13. Annual variability of model parameters when measured wind speed from different year are considered at Näsudden. (a) HF and (b) LF components.

In order to study the variability of the model parameters (both HF and LF model parameters) as a function of location and time, the wind-speed time-series data from Näsudden, Utgrunden and Tjaereborg are investigated, which are all located close to the Baltic Sea region. Note that the wind-speed models developed so far are based on measured wind-speed data over one year. To see the variability of the model parameters from year to year, wind-speed data over many years are considered. As described in Section V, the measurement data from Näsudden are 10 minute average wind speed over five year; the measurement data from Utgrunden are 10-min average wind-speed data over 3 years and the measurement data from Tjaereborg are 10-min average wind-speed data over three years. The average wind speed at each location from year to year is given in Table II.

Fig. 13 shows the model parameters of the wind-speed data from Näsudden over different years. Fig. 13(a) shows the parameters of the HF components, and Fig. 13(b) shows the parameters of the LF components. Note that there are similar patterns in the both HF and LF model parameters. For the HF components, the model parameters from year to year are almost the same. However, there is a slight variation in model parameters of the LF components from year to year.

Fig. 14 shows the model parameters of the wind-speed data from Utgrunden over different years. The pattern and the values of the model parameters in this case are similar to the one presented in Fig. 13. Note that the model parameters of the LF component of the data from Utgrunden is more stable or uniform compared with the one from Näsudden.

Although they are hundreds of kilometers apart, similar results are observed when analyzing wind-speed time-series data at Tjaereborg, which are shown in Fig. 15.

It is worth pointing out two important points from the above results. The first point is that the model parameters, especially those of the HF components, are not greatly affected when measurement data changes from year to year at a given location. The second point is that, by comparing the model parameters

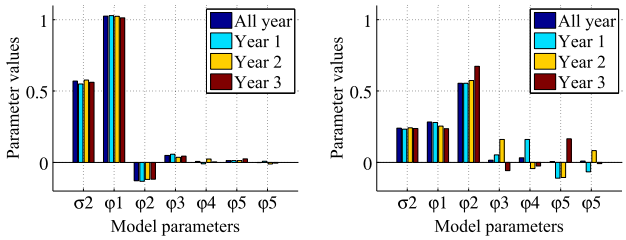


Fig. 14. Annual variability of model parameters when wind-speed data from different year are considered at Utgrunden. (a) HF and (b) LF components.

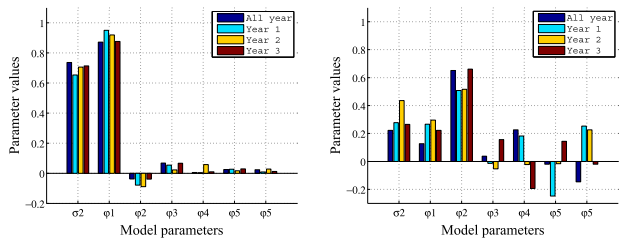


Fig. 15. Annual variability of model parameters when wind-speed measurements from different year are considered at Tjaereborg. (a) HF and (b) LF components.

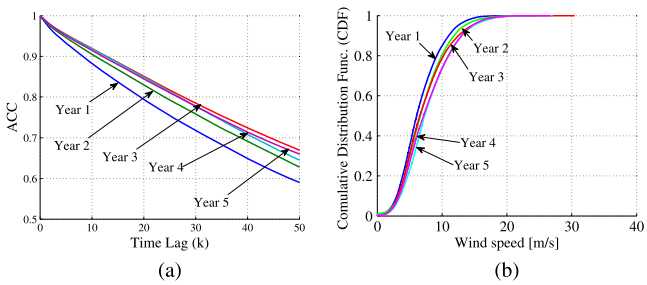


Fig. 16. Annual variability of relevant characteristics of wind-speed time-series data. (a) ACC. (b) CDF.

at Näsudden and Utgrunden, it can be seen that there is a similar pattern in model parameters with a slight difference in the magnitude of the model parameters. This is an interesting observation since it can be used as a basis to use a model developed at one location to simulate wind-speed time-series data at a nearby location. To develop the observation further, it is important to study the variability of wind-speed characteristics from year to year. It is true that average values, variance, time correlation, and probability distribution of wind-speed time-series data varies from year to year [30]. Hence, if a wind-speed model developed at a given location can produce a simulated wind-speed data at a nearby location with the expected annual variation at the nearby location, then it is fair to say that it is acceptable to use the wind-speed model to simulate wind-speed data at the nearby location.

Fig. 16 shows the variation of the relevant characteristics of the measured wind-speed data from year to year at Näsudden. Fig. 16(a) shows the annual variation in ACC of the measured wind-speed time-series data, and Fig. 16(b) shows the annual variation in CDF of the measured wind-speed time-series data. This is the expected annual variation of the characteristics in this

TABLE V
HF MODEL PARAMETERS

	M1	M2	M3	M4
σ^2	0.5697	0.5493	0.5776	0.5617
ϕ_1	1.0247	1.0288	1.0232	1.0132
ϕ_2	-0.1291	-0.1316	-0.1188	-0.1182
ϕ_3	0.0491	0.0576	0.0357	0.0448
ϕ_4	0.0069	-0.0098	0.0235	0.0029
ϕ_5	0.0136	0.0126	0.0128	0.0246
ϕ_6	-0.0013	0.0079	-0.0106	-0.0047

TABLE VI
LF MODEL PARAMETERS

	M1	M2	M3	M4
σ^2	0.2395	0.2325	0.2425	0.2376
θ_1	0.2832	0.2795	0.2539	0.2362
θ_2	0.5549	0.5547	0.5736	0.6733
θ_3	0.0156	0.0525	0.1602	-0.0582
θ_4	0.0322	0.1603	-0.0429	-0.0254
θ_5	0.0064	-0.1111	-0.1061	0.1653
θ_6	0.0097	-0.0662	0.083	-0.0081

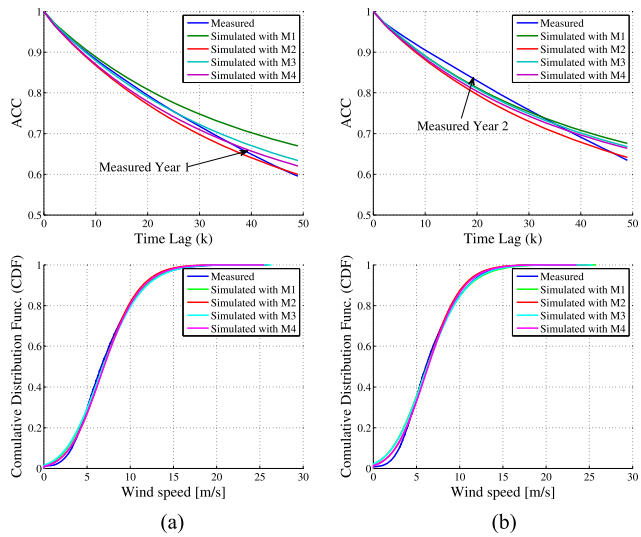


Fig. 17. Deviation of the ACC and CDF of the simulated wind-speed data from the measured wind speed. (a) Näsudden Year 1. (b) Näsudden Year 2.

area and can be used as a basis to check if the simulated wind-speed time-series data at this location, by using the wind-speed model at another location, is acceptable or not.

In order to show the possibility of simulating wind-speed time-series data at different locations in the Baltic Sea region using a model developed at a given location; the model and model parameters developed at Utgrunden, which are presented in Fig. 14, are used to simulate the wind-speed data at Näsudden. The actual values of the model parameters are given in Tables V and VI. Model M1 stands for the model developed by using all year (three years) data at Utgrunden, similarly model M2 is associated with Year 1 data, M3 with Year 2 data and M4 with year 3 data. Table V shows the model parameters of the HF components of the wind-speed data and Table VI shows the model parameters of the LF components of the wind-speed data at Utgrunden.

In order to simulate wind-speed data at Näsudden by using the models developed at Utgrunden, it is assumed that the average wind speed at Näsudden is known or given. Fig. 17 shows

the properties of the simulated wind-speed data at Näsudden produced by using different models developed at Utgrunden. Note that, to produce the results presented in Fig. 17, T_{shift} and θ_0 are set to zero to simulate the HF components of the wind speed. Similarly, while simulating the LF components of the wind-speed data, θ_0 is adjusted in such a way that the average wind speeds at Näsudden, presented in Table II, are achieved. Fig. 17(a) shows the comparison of the measured wind-speed data during Year 1 at Näsudden with the simulated data produced using different models developed at Utgrunden in terms of ACC and CDF. Similarly, Fig. 17(b) shows the comparison of the ACC and CDF of the measured wind-speed data during Year 2 at Näsudden and the ACC and CDF of the simulated wind-speed data generated with different models from Utgrunden.

As can be seen from the results in Fig. 17, the variation of the characteristics of the simulated wind-speed data around the measurement is within the expected variation compared with the expected variation presented in Fig. 16. Compared with the results in Fig. 16, the deviation of the simulated wind speed from the measured wind-speed data is relatively smaller. Similar result is observed when analyzing Year 3, Year 4 and Year 5 wind-speed data at Näsudden. From this, it can be concluded that, in places where there are no measured wind-speed time-series data in the Baltic Sea region, one can consider using the model parameters presented in Tables V and VI to simulate wind-speed time-series data with the only knowledge of the average wind-speed value.

VI. Conclusion

Three main points and contributions are discussed. The fist contribution is the presentation of the proposed modified ARIMA modeling procedure that can be used to model wind-speed time-series data. The modified modeling procedure introduces frequency decomposition, shifting and limiting in addition to differencing and power transformation which are used in the standard ARIMA modeling procedure. The application of the modified modeling procedure is demonstrated by using 10 minute average wind-speed time-series data from multiple locations in the Baltic Sea region over many years. The results show that the proposed modeling procedure works quite well with the studied wind-speed data. The second contribution is that, based on the analysis of the wind-speed data from multiple locations in the Baltic Sea area over many years, it is concluded that it could be sufficient to use ARIMA(6,0,0) and ARIMA(0,1,6) model structure to model the HF and LF components of 10 minute average observed wind-speed data in the Baltic Sea region. An important implication of this is that the effort needed to determine model structure (model identification process), which is the most important step in the modeling procedure, can be avoided. The third and final contribution is that, in areas where there are no measured wind-speed data in the Baltic Sea region, the model parameters presented in Table V and VI can be used to simulate wind-speed time-series data with a desired average value. It is shown that the variation of the relevant characteristics of the simulated wind-speed data can be considered acceptable compared with the expected annual variation of the wind-speed characteristics from year to year at a given area.

One of the limitations in the proposed modeling procedure is that it is assumed that the transformed time-series data is stationary, which might not be entirely valid as there is no guarantee that the transformed time series is strictly stationary. However, as the purpose is to capture the temporal correlation and probability distribution of the wind-speed time-series data, and as long as these two properties are accurately modeled, we consider the transformation of the time series adequate. In addition, the modeling procedure also assumes that the HF component of the wind-speed data is stationary, which might not be necessarily true for the considered cut-off period. Furthermore, the data used in this study is 10 minute average wind-speed data over a year and is of high-quality without any gaps. However, often the data contains gaps and could be of different sampling frequency and length. The application of the procedure to data with different sampling frequency and length and to quantify how much gaps that can be allowed and still obtain models with good validity would accordingly be interesting to study. Apart from this, it is shown that the properties of the measured wind data changed from year to year. It is also shown that although the general pattern of the periodic characteristics is captured, it is might not be enough. It would thus be interesting to incorporate these features into the modeling procedure, perhaps by decomposing the LF components of the wind-speed data further. In addition, to apply proposed methodology to other sea areas would definitely be of very high interest. A final remark is that the application of the proposed modeling procedure in wind-speed forecasting is a potential area that is interesting investigate in a future work.

References

[1] F. R. Pazheri, M. F. Othman, and N. H. Malik, "A review on global renewable electricity scenario," *Renewable Sustain. Energy Rev.*, vol. 31, pp. 835–845, Mar. 2014.

[2] P. S. Georgilakis, "Technical challenges associated with the integration of wind power into power systems," *Renewable Sustain. Energy Rev.*, vol. 12, no. 3, pp. 852–863, Apr. 2008.

[3] M. H. Albadi and E. F. El-Saadany, "Overview of wind power intermittency impacts on power systems," *Electric Power Syst. Res.*, vol. 80, no. 6, pp. 627–632, June 2010.

[4] R. Billinton and W. Wangdee, "Reliability-based transmission reinforcement planning associated with large-scale wind farms," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 34–41, Feb. 2007.

[5] R. Billinton, R. Karki, Y. Gao, D. Huang, P. Hu, and W. Wangdee, "Adequacy assessment considerations in wind integrated power systems," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2297–2305, Nov. 2012.

[6] R. Billinton, H. Chen, and R. Ghajar, "Time-series models for reliability evaluation of power systems including wind energy," *Microelectron. Reliabil.*, vol. 36, no. 9, pp. 1253–1261, Sep. 1996.

[7] A. Bensoussan, P. Bertrand, and A. Brouste, "A generalized linear model approach to seasonal aspects of wind speed modeling," *J. Appl. Statistics*, vol. 41, no. 8, pp. 1694–1707, Aug. 2014.

[8] C. G. Justus, W. R. Hargraves, A. Mikhail, and D. Graber, "Methods for estimating wind speed frequency distributions," *J. Appl. Meteorol.*, vol. 17, no. 3, pp. 350–353, Mar. 1978.

[9] K. Conradsen, L. B. Nielsen, and L. P. Prahm, "Review of Weibull statistics for estimation of wind speed distributions," *J. Climate Appl. Meteorol.*, vol. 23, no. 8, pp. 1173–1183, Aug. 1984.

[10] J. P. Hennessey, "Some aspects of wind power statistics," *J. Appl. Meteorol.*, vol. 16, no. 2, pp. 119–128, Feb. 1977.

[11] C. Nichita, D. Luca, B. Dakyo, and E. Ceanga, "Large band simulation of the wind speed for real time wind turbine simulators," *IEEE Trans. Energy Conversion*, vol. 17, no. 4, pp. 523–529, Dec. 2002.

[12] C. Gavriluta, S. Spataru, I. Mosincat, C. Citro, I. Candela, and P. Rodriguez, "Complete methodology on generating realistic wind speed profiles based on measurements," in *Proc. Int. Conf. Renewable Energies and Power Quality*, 2012, pp. 828–829.
[13] G. D'Amico, F. Petroni, and F. Prattico, "Wind speed modeled as an indexed semi-Markov process," *Environmetrics*, vol. 24, no. 6, pp. 367–376, Sep. 2013.
[14] A. Shamshad, M. A. Bawadi, W. M. A. Wan Hussin, T. A. Majid, and S. A. M. Sanusi, "First and second order Markov chain models for synthetic generation of wind speed time series," *Energy*, vol. 30, no. 5, pp. 693–708, Apr. 2005.
[15] P. Ailliot, V. Monbet, and M. Prevosto, "An autoregressive model with time-varying coefficients for wind fields," *Environmetrics*, vol. 17, no. 2, pp. 107–117, Mar. 2006.
[16] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2010.
[17] S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *Proc. North Amer. Power Symp.*, Sep. 2010, pp. 1–8.
[18] D. Hill, D. McMillan, K. Bell, and D. Infield, "Application of autoregressive models to U.K. wind speed data for power system impact studies," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 134–141, Jan. 2012.
[19] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable Sustain. Energy Rev.*, vol. 13, no. 4, pp. 915–920, May 2009.
[20] I. Colak, S. Sagiroglu, and M. Yesilbudak, "Data mining and wind power prediction: A literature review," *Renewable Energy*, vol. 46, pp. 241–247, Oct. 2012.
[21] J. A. Benth and F. E. Benth, "Analysis and modelling of wind speed in New York," *J. Appl. Statistics*, vol. 37, no. 6, pp. 893–909, Jun. 2010.
[22] J. Haslett and A. E. Raftery, "Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource," *J. Royal Stat. Soc. C, Appl. Stat.*, vol. 38, no. 1, pp. 1–50, Jan. 1989.
[23] B. G. Brown, R. W. Katz, and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power," *J. Climate Appl. Meteorol.*, vol. 23, pp. 1184–1195, Aug. 1984.
[24] W. Wangdee and R. Billinton, "Considering load-carrying capability and wind speed correlation of WECS in generation adequacy assessment," *IEEE Trans. Energy Conversion*, vol. 21, no. 3, pp. 734–741, Sep. 2006.
[25] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, May 2013.
[26] A. Chuang, "Time series analysis: Univariate and multivariate methods," *Technometrics*, vol. 33, no. 1, pp. 108–109, Feb. 1991.
[27] I. Van der Hoven, "Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour," *J. Meteorol.*, vol. 14, no. 2, pp. 160–164, Apr. 1957.
[28] W. G. Ruesink, "Introduction to sampling theory," in *Sampling Methods in Soybean Entomology*, ser. Springer Series in Experimental Entomology, M. Kogan and D. C. Herzog, Eds. New York, NY, USA: Springer, Jan. 1980, pp. 61–78.
[29] L. Fawcett and D. Walshaw, "Markov chain models for extreme wind speeds," *Environmetrics*, vol. 17, no. 8, pp. 795–809, Dec. 2006.
[30] N. B. Negra, O. Holmstrøm, B. Bak-Jensen, and P. Sørensen, "Model of a synthetic wind speed time series generator," *Wind Energy*, vol. 11, no. 2, pp. 193–209, Mar. 2008.

Kalid Yunus (S'13) was born in Mechara, Ethiopia, on July 6, 1983. He received the B.Sc. degree in electrical engineering from Arba-Minch University in 2006, and the M.Sc. degree in electric power engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2008, where he is currently working toward the Ph.D. degree.

He was an Assistant Lecturer with Haramaya University from 2006 to 2008. His field of interest is in power systems, power electronics for power systems, and HVDC transmission system.

Torbjörn Thiringer (M'96) received the M.Sc. and Ph.D. degrees from the Chalmers University of Technology, Gothenborg, Sweden, in 1989 and 1996, respectively.

He is a Professor in applied power electronics with the Chalmers University of Technology. His research interests include the modeling, control and grid integration of wind energy converters into power grids as well as power electronics and drives for other types of applications, such as electrified vehicles, buildings, and industrial applications.

Peiyuan Chen (S'07–M'10) received the B.Eng. degree in electrical engineering from Zhejiang University, China, in 2004, the M.Sc. degree in electric power engineering from Chalmers University of Technology, Gothenborg, Sweden, in 2006, and the Ph.D. degree in stochastic modeling and analysis of power system with renewable generation from Aalborg University, Aalborg, Denmark, in 2010.

Currently, he is an Assistant Professor with both Aalborg University and Chalmers University of Technology. His main research interests are optimal operation and planning of power system with integration of wind power.