Optimum Data Sampling Frequency for Short-Term Analysis of Power Systems with Wind

Guðrún Margrét Jónsdóttir, *Student Member, IEEE*, Brendan Hayes, and Federico Milano, *Fellow, IEEE*
School of Electrical & Electronic Engineering, University College Dublin, Ireland
gudrun.jonsdottir@ucdconnect.ie, {brendan.hayes, federico.milano}@ucd.ie

***Abstract*—The aim of this paper is to determine the optimum data sampling frequency for building wind speed models for dynamic analysis of power systems. Higher sampling frequencies increase the computational burden of the model, while smaller frequencies cannot capture faster wind variations. Another aspect to take into account when selecting the sampling frequency is that the wind turbine blades damp faster wind speed variations. This paper addresses these issues by using wind speed measurements as well as the continuous-time autoregressive moving average modeling approach to study the effect of the sampling frequency and blade damping on the stochastic properties of the proposed wind speed model.**

***Index Terms*—Wind speed modeling, sampling frequency, short-term analysis of power systems, continuous-time ARMA models.**

I. Introduction

Wind energy is one of the fastest growing renewable energy sources for electricity worldwide [1]. This has led to an increase in the volatility and uncertainty present in power systems. Therefore, accurate stochastic models are required to simulate the wind speed fluctuations in power systems. In order to construct these models wind speed data is required. Typically, only minutely or hourly data samples are available as well as average values over the same time frames [2], [3]. In some instances however, such as the dynamic analysis of power systems, time steps of less then one minute are required. Little work has been done to study the effect of wind speed sampling times on short-term power system analysis. This research identifies the effect the data sampling frequency has when building wind speed models based on data for short-term analysis of power systems.

Wind speed is typically modeled as a stochastic process. Since wind speed is site dependent, there is no universal time step, dataset or model for all locations. Each site requires its own data-based model [4], [5]. Traditionally, models used for dynamic analysis of power systems are based on data with a minutely sampling time or are not based on data [6]–[10]. For example, the four-component composite model used in [6] and [7] and the Weibull distribution wind speed model used in [8] are not based on data. In [9] and [10], the modeling approaches intended for short-term analysis, which ranges from seconds to a few minutes, of power systems are presented utilizing minutely wind speed data.

The wind speed models generate processes that are fed to the aerodynamic part of the wind turbine rotor model. In [11] and [7], a simplified aerodynamic model for wind turbines is presented. It models the damping effect of the rotor blades as a low-pass filter. The time constant of the low-pass filter is dependent on the rotor radius and the average wind speed. The effective wind speed on the rotor of the wind turbine is therefore a filtered version of the actual wind speed. This filters out some of the faster wind speed variations and thereby sets a limit on the required sampling frequency.

The focus of this research is to determine the optimum sampling frequency that captures the stochastic properties of wind speed on a minutely to secondly bases. The damping of the rotor complicates the selection of the sampling frequency. A higher sampling frequency increases the computational burden and if the sampling frequency is too low, the models do not capture the faster wind speed variations. With this aim, the continuous-time equivalent of the well-established Autoregressive Moving Average (ARMA) is used to model the wind speed. This enables building models based on data with any sampling rate and simulating them with a time step of seconds. This approach is based on the observation that if an ARMA model is stationary then it has a continuous-time counterpart termed Continuous-Time ARMA (CARMA) [10], [12].

The contributions of this paper are twofold:

- To analyze what resolution is required when sampling the wind speed data to build models intended for use in a secondly to minutely time frame, specifically for dynamic analysis of power systems.
- To study how the effective wind speed on the rotor of the wind turbine depends on the data sampling rate used.

The reminder of this paper is organized as follows. In Section II, the steps to construct the CARMA wind speed models are presented. Section III presents the filter model used to represent the damping effect of the turbine blades. The wind speed data with different sampling times is presented in Section IV. In Section V, CARMA wind speed models are built based on data with different sampling rates and compared. Finally, Section VI draws conclusions and outlines possible areas of future research.

II. Construction of CARMA models

This section presents a procedure to obtain a continuous-time ARMA model from wind speed measurements. The proposed procedure consists of four steps, as shown in Fig. 1. Each step is discussed below. Further information on the construction procedure as well as the accuracy of this construction method can be found in [10].

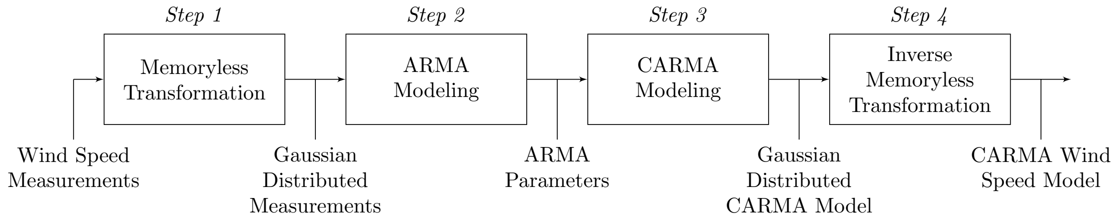

Fig. 1. The four steps to construct the proposed CARMA wind speed models.

Step 1: Memoryless transformation

The memoryless transformation fits non-normally distributed data to the Gaussian distribution while retaining its original stochastic properties. This is achieved by applying the inverse Gaussian Cumulative Distribution Function (CDF) to the CDF of the wind speed data.

$$y(t) = \Phi^{-1}(F(X(t))) \ , \tag{1}$$

where Φ^{-1} is the inverse CDF of the Gaussian distribution and F is the CDF of the probability distribution of the wind speed [13]. The result is wind speed that is normally distributed which can be used to build an ARMA model.

Step 2: ARMA *modeling*

ARMA models can be divided into two components, namely, Autoregressive (AR) and Moving Average (MA):

- *Autoregressive*: relates the current value of the wind speed to past values.
- *Moving Average*: relates the current value of the wind speed to past error values.

The ARMA(p,q) model is given by

$$X_t = \underbrace{\sum_{i=1}^{p} \phi_i X_{t-i}}_{AR} + \underbrace{\sum_{i=1}^{q} \theta_i \varepsilon_{t-i}}_{MA} + \varepsilon_t \ , \tag{2}$$

where ε_t is white noise with a standard deviation σ_a, ϕ_i are the autoregressive parameters and θ_i are the moving average parameters. ϕ_p and θ_q are non-zero.

ARMA models of second order or higher have been widely used to model wind speed [14]–[16]. The ARMA(2,1) model is a special case as it is the lowest order ARMA model that captures the statistical properties of wind speed. It can be written as

$$\phi(B) X_t = \theta(B) \varepsilon_t \ , \tag{3}$$

where B is the backward operator such that $BX_t = X_{t-1}$ and

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 \tag{4}$$

$$\theta(z) = 1 + \theta_1 z \ . \tag{5}$$

In this research, the Maximum Likelihood method is used to estimate the ARMA parameters. The method finds the parameter values of the ARMA model which maximizes the Likelihood Function of the sampled data. The Likelihood Function is based on the Gaussian CDF of the sampled data. The estimated ARMA parameters are used to find the equivalent CARMA parameters.

Step 3: CARMA *modeling*

CARMA models are the continuous-time counterparts of the discrete-time ARMA models. A CARMA(p,q) model denoted by $x(t)$ is a Stochastic Differential Equation (SDE) of the form

$$\begin{aligned} &\frac{d^p x}{dt^p} + c_{p-1}\frac{d^{p-1}x}{dt^{p-1}} + \ldots + c_1 \frac{dx}{dt} + c_0(x(t) - \mu) \\ &\quad = b_0 dW(t) + b_1 \frac{dz}{dt} + \ldots + b_q \frac{d^q z}{dt^q} \ , \end{aligned} \tag{6}$$

where $W(t)$ is the standard Wiener process. c_i are the autoregressive coefficients and b_i are the moving average coefficients. c_i and b_i are real and $b_q \neq 0$ [17]–[19].

Generally, a stationary CARMA(p,q) model sampled regularly can be written as a ARMA(p,p-1) model with $q < p$ [12]. The simplest example is the Ornstein-Uhlenbeck process, i.e., CARMA(1,0)

$$dX_t + c_0 X_t = dW(t) \ , \tag{7}$$

which is equivalent to an ARMA(1,0) model viewed with a fixed time step h

$$X_t = \exp(-c_0 h) X_{t-1} + \varepsilon_t \ . \tag{8}$$

The CARMA(2,1) model is used in the remainder of this paper and can be written as

$$c(D)X(t) = b(D)dW(t) \ , \tag{9}$$

where D is the differential operator and

$$c(z) = z^2 + c_1 z + c_0 \tag{10}$$

$$b(z) = b_0 + b_1 z \ . \tag{11}$$

An equivalent discrete-time ARMA(2,1) model can be found if the CARMA(2,1) model is stationary. The CARMA(2,1) model is stationary if the real parts of the roots of (10), α_1 and α_2, are negative. The autoregressive parameters of the continuous-time model, c_1 and c_0, can be directly connected to the autoregressive parameters of the discrete-time model, ϕ_1 and ϕ_2, using the z-transformation.

$$\phi_1 = e^{\alpha_1 h} + e^{\alpha_2 h} \tag{12}$$

$$\phi_2 = -e^{(\alpha_1 + \alpha_2)h} \ . \tag{13}$$

The theoretical auto-covariance function of a discrete-time ARMA(2,1) model is defined as

$$\gamma_{\text{ARMA}}(k) = \begin{cases} \phi_1\gamma(1) + \phi_2\gamma(2) + \theta_1(\phi_1 + \theta_1)\sigma_a^2 + \sigma_a^2 & \text{if } k = 0 \\ \phi_1\gamma(0) + \phi_2\gamma(1) + \theta_1\sigma_a^2 & \text{if } k = 1 \\ \phi_1\gamma(k-1) + \phi_2\gamma(k-2) & \text{if } k > 1 \ , \end{cases} \tag{14}$$

where k is the time lag and $\gamma = \gamma_{\text{ARMA}}$. The theoretical auto-covariance function of a CARMA(2,1) model is

$$\gamma_{\text{CARMA}}(h) = e^{\alpha_1 h} \frac{b(\alpha_1)b(-\alpha_1)}{c'(\alpha_1)c(\alpha_1)} + e^{\alpha_2 h} \frac{b(\alpha_2)b(-\alpha_2)}{c'(\alpha_2)c(\alpha_2)}. \quad (15)$$

The moving average parameter of the continuous-time model, b_1, is set so that the auto-covariance of the discrete-time model, γ_{ARMA}, is equal to the auto-covariance of the continues time model, γ_{CARMA}.

Step 4: Inverse memoryless transformation

The inverse of the memoryless transformation in *Step 1* is used to impose the true probability distribution of the data. The inverse CDF of the wind speed data is applied to the Gaussian CDF of the CARMA model

$$y(t) = F^{-1}(\Phi(X(t))) \ . \quad (16)$$

This obtains the desired probability distribution of the wind speed data [13]. The memoryless transformation and its inverse enable the use of ARMA and CARMA models to model data with any probability distribution that has a defined CDF and inverse CDF.

III. Equivalent wind speed

The CARMA method, presented in Section II, is used to model wind speed based on the data available in [20]. Wind speed is not uniform across the rotor blade area. For example, wind speed at the tip, center and hub can differ [11]. However, these variations even out over the blade area. This damping effect by the rotor blades is modeled as a low-pass filter, shown in Fig. 2. The input is the CARMA data-based model generating wind speed at hub height of the wind turbine. The output is the equivalent wind speed that produces the same torque as the actual wind field [7].

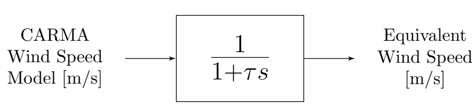

Fig. 2. The low-pass filter that represents the damping effect of the blades of the wind turbine.

The time constant, τ [s], of the low-pass filter is proportional to the rotor radius, R [m], and the average wind speed at the hub height, U [m/s]. The time constant of the low-pass filter is:

$$\tau = \frac{\gamma R}{U} \ , \quad (17)$$

where γ is the decay factor over the disc [11]. The radius of the rotor typically ranges from $20 - 90$ m. Thus, the time constant of the low-pass filter is between $5 - 20$ s. Therefore, the filter has a cut-off frequency of $0.03 - 0.008$ Hz which corresponds to $30 - 120$ s. Hence, the filter is damping wind speed variations in the minutely time frame.

The fastest wind speed variations have frequencies above 10 Hz. The Nyquist-Shannon sampling theorem states that for discrete samples, to capture all the information of a continuous-time signal, a sampling frequency, f_s, has to be chosen which is twice the highest frequency component, $f_{\max}$ of the continuous-time signal. That is $f_s \geq 2f_{\max}$. The sampling frequency required in this case is limited by the low-pass filter. Due to the cut-off frequency of the filter the fast wind speed variations in this paper are in the range $0.001 - 10$ Hz. Furthermore, since the rotor damping is modeled as a first-order filter, the wind speed variations close to cut-off frequency are damped, not eliminated. It is difficult to ascertain the extent to which these higher frequencies pass through the filter. Hence, this research utilizes simulation to identify the optimum sampling frequency of wind speed data for models intended for short-term analysis of power systems.

IV. Wind speed data

The wind speed data used throughout this research were gathered in the *RE<C (Renewable Electricity Less Than Coal) project* and are available online [20]. The dataset consists of one month of wind speed measurements sampled at 7.6 Hz. The data is collected close to Tracy, California in the United States. Two groups of datasets with different sampling times are created in order to study the effect of the sampling time as well as averaging on the stochastic properties of wind speed data. Both groups of datasets are centered around minutely sampled data as the wind turbine blades damp oscillations with higher frequencies.

From a statistical point of view, the wind speed can be characterized by its Probability Density Function (PDF) and autocorrelation. The best fit for the autocorrelation and PDF of wind speed is dependent on both the location and the sampling time [4], [5]. The PDF gives the likelihood of the occurrence of all possible wind speeds at a specific location. The autocorrelation indicates how much the speed is likely to change based on its current value. The autocorrelation measures the relationship between the wind speeds current value and its past and future values.

The CARMA models used in this paper aim to capture both the autocorrelation and PDF of the wind speed data. The two wind speed dataset groups studied in this paper are presented in Sections IV-A and IV-B.

A. Group 1

The datasets in Group 1 contain one month of data and are sampled as follows; fifteen minutely, five minutely, minutely and minutely averages.

The two-parameter Weibull PDF that best fits each dataset in Group 1 is shown in Fig. 3. The lower wind speeds are slightly more prevalent in the fifteen minutely data as well as in the average minutely data. The fifteen minutely data is sampled less frequently and is therefore less likely to capture the higher wind speeds. By taking the average over a minute the extreme wind speeds are averaged out. The difference between the PDF of the minutely and five minutely data is not visible and the overall difference between the four datasets is minimal. Therefore, the selection of the sampling time step is not critical for capturing the PDF.

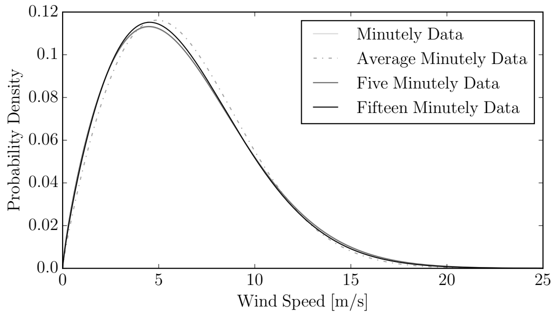

Fig. 3. PDFs of the wind speed datasets in Group 1.

The autocorrelation of the wind speed datasets in Group 1 are shown in Fig. 4. The noticeable drop in the autocorrelation of the four datasets in the first time instants is due to short-term wind speed variations, such as turbulence and gusts. The initial drop in the autocorrelation of the average minutely data is much smaller than for the other three datasets. This is because by taking the average, the effect of the short-term variations is reduced. The minutely, five minutely and fifteen minutely datasets have a steep drop after the first sampling instance and then settle to have similar slope.

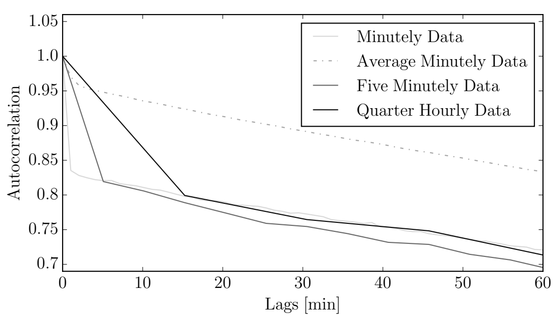

Fig. 4. Autocorrelations of the wind speed datasets in Group 1.

B. Group 2

The datasets in Group 2 contain one week of data and are sampled as follows; minutely, half minutely, twenty secondly, ten secondly and secondly.

The autocorrelation of the wind speed datasets in Group 2 are shown in Fig. 5. When the time between samples is decreased the initial drop in the autocorrelation gets steeper. The autocorrelation of the datasets indicates that the models will have the same long-term behavior but the short-term fluctuations depend on the time step of the data.

V. Simulation results

In this section, the CARMA construction method presented in Section II is used to build wind speed models based on both Group 1 and 2 datasets. The generated wind speed processes are filtered using the low-pass filter, presented in Section III, that represents the damping effect of the turbine blades and the resulting effective wind speed on the turbine is examined.

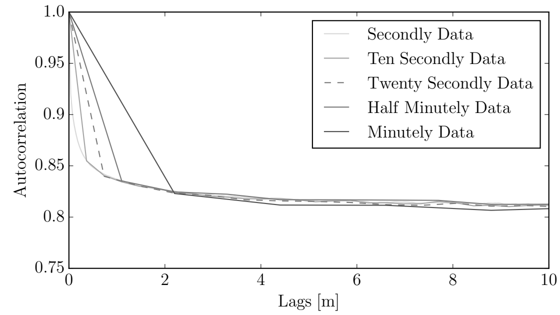

Fig. 5. Autocorrelations of the wind speed datasets in Group 2.

A. Group 1

The CARMA construction method is used to build wind speed models based on the Group 1 datasets. Each wind speed model is simulated $1,000$ times. The model based on average minutely data generates wind processes that vary less then the other three models. This is predicted by the autocorrelation in Fig. 4. The initial variations of the wind speed increase as the sampling frequency increases. However, this difference decreases with time as illustrated by the autocorrelation in Fig. 4. The slope of the autocorrelation defines how fast the wind speed varies over a certain time frame. The variance of the wind speeds is shown in Fig. 6 to better compare the wind speed processes.

The generated wind speed processes are fed to the low-pass filter, presented in Section III, to model the damping effect of the blades of the wind turbine. The resulting filtered wind speed is the effective wind speed on the wind turbine. The filter in this case is set with the turbine radius of both 40 m and 90 m. The average wind speed in both cases is set to 6 m/s and the decay factor of the disc is set to $\gamma = 1.3$. The resulting time constants of the filter are found using (17); $\tau_{40} = 8.67$ s with $R = 40$ m and $\tau_{90} = 19.5$ s with $R = 90$ m. The cut-off frequency of the filter is 0.0184 Hz and 0.0019 Hz, respectively.

In Fig. 6, the variance of the equivalent filtered wind speed processes after 10 s and 100 s are shown. The filter slightly damps the processes based on five and fifteen minutely data but damps the oscillations of the minutely based processes more because the variations have a higher frequency. This indicates that the maximum sampling step can be one minute. Therefore, models based on datasets with a higher sampling frequency are studied.

B. Group 2

The wind speed models based on the Group 2 datasets are simulated $1,000$ times. The variance of the generated wind speed processes after 10 s and 100 s are shown in Fig. 7. The effect of the sampling is visible in the variance after 10 s but not after 100 s. This is because the initial drop in the autocorrelation, shown in Fig. 5, is dependent on the sampling time. After one minute all the models have the same autocorrelation thus, after 100 s all models have the same variance.

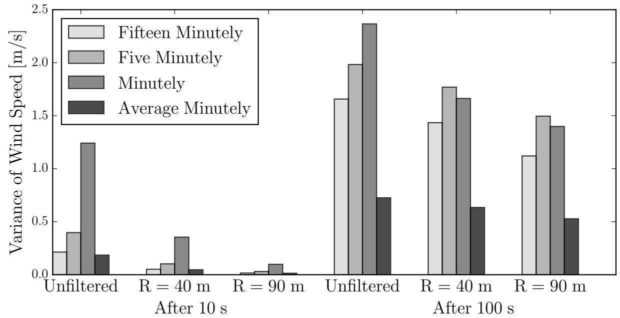

Fig. 6. The variance of the 1,000 processes simulated using wind speed models based on the Group 1 datasets.

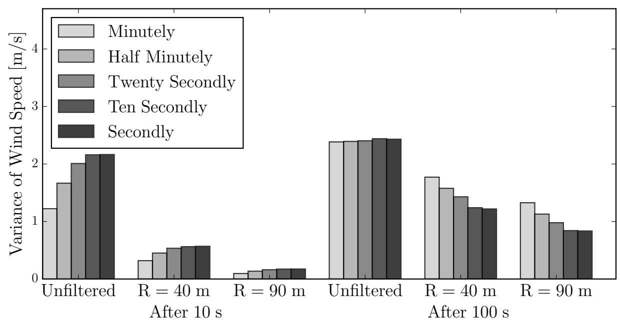

Fig. 7. The variance of the 1,000 processes simulated using wind speed models based on the Group 2 datasets.

The variance of the filtered wind speed processes are shown in Fig. 7. The low-pass filter damps the variance of all five models. As the sampling frequency is increased the newly introduced faster variations are damped more. The variance of the model based on ten secondly and secondly data is almost identical. This indicates that ten secondly sampling is sufficient as faster variations are damped by the low-pass filter.

VI. Conclusions

This work studies the effect of the sampling frequency on the stochastic properties of wind speed data. The paper takes in to account the damping effect of the turbine blades. This damping effect is modeled using a low-pass filter so that wind speed variations close to the cut-off frequency of the filter are damped, not eliminated. The CARMA construction method is used to build wind speed models based on datasets with sampling times from fifteen minutely to secondly.

Wind speed measurements indicate that the PDF of the wind is not dependent on the sampling frequency. However, the autocorrelation depends on the sampling frequency as it defines the slope of the autocorrelation. Results also show that ten secondly sampling is the optimum sampling rate. Higher sampling frequencies will increase the computational burden of the model, while, a lower sampling frequencies will not capture fast wind speed variations.

The proposed CARMA model is tailored for short-term dynamic analysis. Long-term effects, such as daily and seasonal ones, are not taken into account in this study. Future work will attempt to define a stochastic wind speed model that is able to reproduce both the short- and long-term behavior of wind speed.

References

[1] M. Bilgili, A. Ozbek, B. Sahin, and A. Kahraman, “An overview of renewable electric power capacity and progress in new technologies in the world,” *Renewable and Sustainable Energy Reviews*, vol. 49, pp. 323–334, 2015.

[2] Met éireann - The Irish Meterological Service Online, Historical Data (2017). [Online]. Available: http://met.ie/climate-request/

[3] Wind Energy Center, University of Massachusetts, Wind Data (2017). [Online]. Available: https://www.umass.edu/windenergy/

[4] R. Zárate-Miñano and F. Milano, “Construction of SDE-based wind speed models with exponentially decaying autocorrelation,” *Renewable Energy*, vol. 94, pp. 186–196, 2016.

[5] R. Calif, “PDF models and synthetic model for the wind speed fluctuations based on the resolution of langevin equation,” *Applied Energy*, vol. 99, pp. 173–182, 2012.

[6] A. Abrous, R. Wamkeue, and E. M. Berkouk, “Modeling and simulation of a wind model using a spectral representation method,” in *2015 3rd International Renewable and Sustainable Energy Conference (IRSEC)*, Dec 2015, pp. 1–6.

[7] T. Ackermann, *Wind power in power systems*. John Wiley & Sons, 2005.

[8] A. Ortega and F. Milano, “Comparison of different control strategies for energy storage devices,” in *Power Systems Computation Conference (PSCC), Genova, Italy*, 2016, pp. 1–7.

[9] J. Sexauer and S. Mohagheghi, “Hybrid stochastic short-term models for wind and solar energy trajectories,” in *Green Technologies Conference (GreenTech), 2015 Seventh Annual IEEE*. IEEE, 2015, pp. 191–198.

[10] G. M. Jónsdóttir, B. Hayes, and F. Milano, “Continuous-time ARMA models for data-based wind speed models,” in *Power Systems Computation Conference (PSCC), Dublin, Italy, 2018.*, 2018, pp. 1–7. [Online]. Available: http://faraday1.ucd.ie/archive/papers/windcarma.pdf

[11] T. Petru and T. Thiringer, “Modeling of wind turbines for power system studies,” *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1132–1139, 2002.

[12] H. Tómasson, “Some computational aspects of Gaussian CARMA modelling,” *Statistics and Computing*, vol. 25, no. 2, pp. 375–387, 2015.

[13] M. Grigoriu, *Applied non-Gaussian processes: Examples, theory, simulation, linear random vibration, and MATLAB solutions*. Prentice Hall, 1995.

[14] A. Lojowska, “Wind speed modeling,” Master’s thesis, Delft, the Netherlands, 2009.

[15] E. Erdem and J. Shi, “ARMA based approaches for forecasting the tuple of wind speed and direction,” *Applied Energy*, vol. 88, no. 4, pp. 1405–1414, 2011.

[16] K. Yunus, T. Thiringer, and P. Chen, “ARIMA-based frequency-decomposed modeling of wind speed time series,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2546–2556, 2016.

[17] M. Phadke and S. Wu, “Modeling of continuous stochastic processes from discrete observations with application to sunspots data,” *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 325–329, 1974.

[18] F. Lu, K. K. Lin, and A. J. Chorin, “Comparison of continuous and discrete-time data-based modeling for hypoelliptic systems,” *arXiv preprint arXiv:1605.02273*, 2016.

[19] P. Brockwell, “Continuous-time ARMA processes,” *Handbook of Statistics*, vol. 19, pp. 249–276, 2001.

[20] RE<C: Renewable Electricity Less Than Coal - Surface Level Wind Data Collection (2017). [Online]. Available: https://code.google.com/archive/