SDE-based Wind Speed Models with Weibull Distribution and Exponential Autocorrelation

Rafael Zárate-Miñano
Escuela de Ingeniería Minera
e Industrial de Almadén
Universidad de Castilla - La Mancha
Almadén, Spain
Email: rafael.zarate@uclm.es

Francesca Madia Mele
School of Electrical
and Electronic Engineering
University College Dublin
Dublin, Ireland
Email: francesca.mele@ucdconnect.ie

Federico Milano
School of Electrical
and Electronic Engineering
University College Dublin
Dublin, Ireland
Email: federico.milano@ucd.ie

***Abstract*—This paper discusses three approaches to construct wind speed models based on Stochastic Differential Equations (SDEs). The methods are applied to construct models able to simulate wind speed trajectories that are statistically described by means of the Weibull distribution and the exponential autocorrelation. The ability of the three models to reproduce stochastic processes with the above indicated statistical properties is duly studied and compared. With this aim, wind speed measurements recorded in a weather station located in Ireland are analyzed. The parameters obtained in this analysis are used to set up the developed models. Finally, the statistical properties of the trajectories generated by the three models are compared with the statistical properties of the considered wind speed data set.**

I. Introduction

The modeling of the wind speed behaviour is an essential aspect in many studies related to power systems. As wind speed is a stochastic phenomenon, the use of stochastic models appears to be a natural choice. Time series, four-component composite model, and models based on Kalman filters, have been traditionally applied in different research fields for the wind speed modeling task [1].

From a statistical point of view, the wind speed behaviour is described by its probability distribution and its autocorrelation. These properties are obtained by means of statistical analyses of historical data recorded for a particular location. Although a variety of probability distributions have been proposed to fit the empirical probability density of the wind speed, the Weibull distribution is generally considered a good fit in many locations around the world [2]. With respect to the autocorrelation, this property has been usually approximated by an exponential decaying function in the time frame of hours, [3], although approximations of the power-law type have been also suggested in [4].

In recent years, Stochastic Differential Equations (SDEs) are gaining popularity as a tool to model stochastic phenomena in power systems (e.g, [5]–[7]). These types of equations have been applied to wind speed modeling in [8]–[10]. The three references above propose three different construction methods to obtain the final wind speed model. The approaches in [8] and [10] are based on the solution of the stationary Fokker-Planck equation while the approach in [9] is based on the theory of translation processes.

Among the three aforementioned approaches, only [10] theoretically guarantee that the constructed model is able to generate wind speed trajectories with the given probability distribution and exponential autocorrelation. The method in [9] leads to models that generate stochastic processes with the given distribution, but an exact exponential autocorrelation with a given decay rate is not guaranteed. Finally, the method proposed in [8] is applied to construct models for wind speed fluctuations with exponential autocorrelation, but based on Gaussian-related probability distributions.

In this paper, the three approaches above are utilized to formulate three different wind speed models on the basis of a given Weibull distribution and autocorrelation functions. The parameters of the Weibull distribution and the decay rate of the exponential autocorrelation are defined according to the results of the statistical analysis of a real-world wind speed data set. The ability of the resulting models to reproduce the statistical properties for what they are designed is duly compared throughout simulations.

II. Construction Methods of SDE-based Models

One-dimensional stochastic differential equations of the Itô type have the following general form:

$$dx(t) = a(x(t), t) \cdot dt + b(x(t), t) \cdot dW(t) \qquad (1)$$

where functions $a(x(t), t)$ and $b(x(t), t)$ are so-called drift and diffusion terms, respectively, and $W(t)$ is a standard Wiener process [11], [12]. The drift and diffusion terms of (1) determines the statistical properties of the variable $x(t)$. In our case, $x(t)$ represents the wind speed and our objective is to define $a(x(t), t)$ and $b(x(t), t)$ such that the stationary probability distribution of $x(t)$ is Weibull and its autocorrelation show an exponential decaying behavior. The following Subsections provide three different approaches that pursue the aforementioned goals. Method I is based on the application of a memoryless transformation, while Methods II and III are based on the solution of the stationary Fokker-Planck equation.

A. Method I

The starting point of this method is the stochastic process represented by the following well-known SDE:

$$dx(t) = -\alpha \cdot (x(t) - \mu) \cdot dt + \sqrt{2 \cdot \alpha} \cdot \sigma \cdot dW(t) \quad (2)$$

The solution of (2) is a stationary stochastic process so-called Ornstein-Uhlenbeck process. This process is characterized by a Normal distribution of mean μ and standard deviation σ, and by an exponential autocorrelation whose decay rate is governed by the coefficient α, i.e.,

$$r(\tau) = e^{-\alpha \cdot \tau} \quad (3)$$

where τ is the time lag. Therefore, (2) fulfills the objective related to the autocorrelation of the process. In order to obtain a given probability distribution, the following memoryless transformation is applied to (2) [9]:

$$y(t) = F_D^{-1}\left(\Phi(x(t))\right) \quad (4)$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) of the Normal distribution $N(\mu, \sigma)$, and $F_D(\cdot)$ is the CDF of the desired probability distribution D. As a result of the memoryless transformation (4), a translation process $y(t)$ is obtained with the desired probability distribution. Note, however, that as a consequence of the memoryless transformation, the autocorrelation of $y(t)$ is not guaranteed to be the exponential autocorrelation of the original process $x(t)$ defined by (2), [9]. Theoretical foundations of this fact can be found in [13].

B. Methods II and III

Both methods are based on the Fokker-Planck equation. This equation is a partial differential equation whose solution provides the time evolution of the Probability Density Function (PDF) of a stochastic process represented by the SDE (1). In the stationary case, the Fokker-Planck equation has the following form:

$$0 = -a(x(t)) \cdot p(x(t)) + \frac{1}{2} \cdot \frac{\partial}{\partial x(t)} \left[b^2(x(t)) \cdot p(x(t))\right] \quad (5)$$

where functions $a(x(t))$ and $b(x(t))$ are the drift and diffusion terms of the SDE, respectively, and $p(x(t))$ is the stationary PDF of the process $x(t)$.

Equation (5) can be solved for the drift term, as follows:

$$a(x(t)) = b(x(t)) \cdot \frac{\partial b(x(t))}{\partial x(t)} + \frac{1}{2} \cdot b^2(x(t)) \cdot \frac{\partial \ln p(x(t))}{\partial x(t)} \quad (6)$$

or it can be also solved for the (squared) diffusion term, as follows:

$$b^2(x(t)) = \frac{2}{p(x(t))} \cdot \int_{-\infty}^{x(t)} a(z(t)) \cdot p(z(t)) \cdot dz(t) \quad (7)$$

for $p(x(t)) \neq 0$, and $b(x(t)) = 0$ if $p(x(t)) = 0$. Therefore, for a given PDF $p(x(t))$, if the diffusion term is known, the drift term is obtained from (6). In a similar way, if the drift term is known, the diffusion term is obtained from (7). The idea is then to proceed in two steps: (i) first, find the expression of one term (the drift or the diffusion) in order to obtain an exponentially autocorrelated process, and (ii) solve the corresponding equation to obtain the form of the other term.

1) Method II: In this method, the drift term is first defined and (7) is solved to obtain the diffusion term. The expression of the drift term is as follows:

$$a(x(t)) = -\alpha \cdot (x(t) - \mu_D) \quad (8)$$

where μ_D is the mean of the desired probability distribution, and α is the autocorrelation coefficient. Defining a drift term as in (8) is a sufficient condition to obtain a stochastic process with exponential autocorrelation, as demonstrated in [10]. Observe that the expression of the drift term (8) is similar to the expression of the drift term of the Ornstein-Uhlenbeck process (2).

2) Method III: In this method, the diffusion term is first defined and (6) is solved to obtain the drift term. The expression of the drift term is as follows:

$$b(x(t)) = \sqrt{2 \cdot \alpha} \cdot \sigma_D \quad (9)$$

where σ_D is the standard deviation of the desired probability distribution, and α is the autocorrelation coefficient. Observe that the expression of the diffusion term (9) is similar to the expression of the diffusion term of the Ornstein-Uhlenbeck process (2). From the best of the authors' knowledge, there is no any theoretical reason that justifies the fact that defining a diffusion term as in (9) lead to a stochastic process with exponential autocorrelation. In [8], this method is applied for Gaussian-related probability distributions (namely, the Gram-Charlier expansion of a Gaussian distribution and a bimodal distribution composed of Gaussians) with good results.

III. Wind Speed Models

In this section, the methods described above are applied to obtain SDE-based models able to generate stochastic processes with Weibull distribution and exponential autocorrelation. The following functions and moments are relevant to develop such models:

$$F_{\mathrm{W}}(u) = \begin{cases} 1 - e^{(u/\lambda)^k} & \text{if } u \geq 0 \\ 0 & \text{if } u < 0 \end{cases} \quad (10)$$

$$p_{\mathrm{W}}(u) = \begin{cases} \frac{k}{\lambda} \cdot \left(\frac{u}{\lambda}\right)^{k-1} \cdot e^{-(u/\lambda)^k} & \text{if } u \geq 0 \\ 0 & \text{if } u < 0 \end{cases} \quad (11)$$

$$\mu_{\mathrm{W}} = \lambda \cdot \Gamma\left(1 + \frac{1}{k}\right) \quad (12)$$

$$\sigma_{\mathrm{W}}^2 = \lambda^2 \cdot \Gamma\left(1 + \frac{2}{k}\right) - \mu_{\mathrm{W}}^2 \quad (13)$$

where k is the shape parameter, λ is the scale parameter, $F_{\mathrm{W}}(\cdot)$ is the CDF, $p_{\mathrm{W}}(\cdot)$ is the PDF, μ_{W} is the mean, and σ_{W} is the standard deviation of the Weibull distribution, respectively. In (12) and (13), $\Gamma(\cdot)$ represents the Gamma function.

A. Model I

To obtain Model I, Method I is applied. For simplicity, the Ornstein-Uhlenbeck process (2) is adapted to a standard Normal distribution, i.e., $\mu = 0$ and $\sigma = 1$. Therefore, the first equation of Model I is

$$dx(t) = -\alpha \cdot x(t) \cdot dt + \sqrt{2 \cdot \alpha} \cdot dW(t) \quad (14)$$

A memoryless transformation of the type of (4) particularized for the Weibull distribution is applied to the solution of (14), i.e,

$$y(t) = F_{\mathrm{W}}^{-1}\left(\Phi(x(t))\right) \quad (15)$$

where $\Phi(\cdot)$ is the CDF of the standard Normal distribution $N(0,1)$, and $F_{\mathrm{W}}(\cdot)$ is the CDF of the Weibull distribution (10). Therefore, Model I is a two-equation model composed of one SDE and one algebraic equation, where variable $y(t)$ represents the wind speed.

B. Model II

This is a single-SDE model that results from applying Method II for the particular case of a Weibull distribution. The drift term is set to

$$a(x(t)) = -\alpha \cdot (x(t) - \mu_{\mathrm{W}}) \quad (16)$$

where μ_{W} is the mean of the Weibull distribution (12). By substituting $p(x(t))$ in (7) by the PDF of the Weibull distribution (11), and solving the equation, the following diffusion term is obtained:

$$b(x(t)) = \sqrt{b_1(x(t)) \cdot b_2(x(t))} \quad (17)$$

with

$$b_1(x(t)) = \frac{2 \cdot \alpha}{p_{\mathrm{W}}(x(t))} \quad (18)$$

and

$$b_2(x(t)) = \lambda \cdot \Gamma\left(1 + \frac{1}{k}, \left(\frac{x(t)}{\lambda}\right)^k\right) - \mu_{\mathrm{W}} \cdot e^{-(x(t)/\lambda)^k} \quad (19)$$

where $\Gamma(\cdot,\cdot)$ is the Incomplete Gamma function.

C. Model III

This model is obtained applying Method III and, similarly to Model II, it is a single-SDE model. The diffusion term is defined as

$$b(x(t)) = \sqrt{2 \cdot \alpha} \cdot \sigma_{\mathrm{W}} \quad (20)$$

where σ_{W} is the standard deviation of the Weibull distribution (13). After setting $p(x(t))$ to the PDF (11) and substituting the diffusion term (20) in (6), the resulting equation is solved and the following drift term is obtained:

$$a(x(t)) = \frac{\alpha \cdot \sigma_{\mathrm{W}}^2 \cdot k}{x(t)} \cdot \left(\frac{k-1}{k} - \left(\frac{x(t)}{\lambda}\right)^k\right) \quad (21)$$

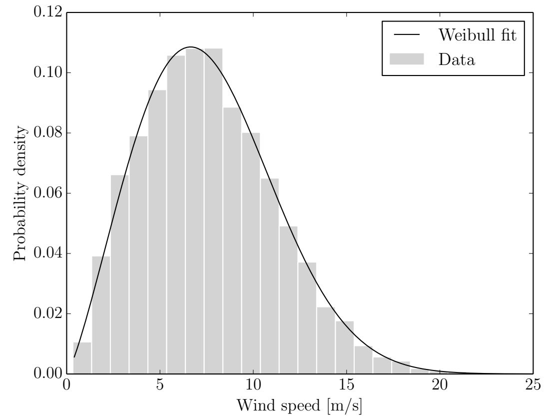

Fig. 1. Probability density of the hourly mean wind speed data.

IV. Case Study

In this section, the three SDE-based models developed above are utilized to model the wind speed of a real-world location. Such a location is Malin Head, head land situated at the most northerly point of Ireland. The data set was provided by Met Éireann, the Irish national meteorological service and consists of wind speed measurements recorded by a synoptic weather center located 22 meters above the mean sea level that includes a 3-cup anemometer. Raw wind speed data are adquired every 0.25 seconds and running averages of 240 values are calculated every minute [14].

A. Data analysis

The original data set is composed of one-minute mean wind speed values expressed in knots for the whole year 2012. For practical reasons, the units of the wind speed measurements are transformed to meters per second, and average is taken such that a data set with hourly mean wind speed values is obtained. As 2012 was a leap year, the new data set contains 8784 values.

Figure 1 depicts the normalized histogram of the considered data set and the PDF fit considering a Weibull distribution. It can be observed that the Weibull PDF fits quite well the empirical probability density of the data. The parameters of the Weibull distribution obtained in the fitting process are $k = 2.2596$ and $\lambda = 8.6055$.

Figure 2 shows the autocorrelation of the considered data set for time lags up to 120 hours (5 days). The grey line is the autocorrelation computed from data, while the black line corresponds to the exponential fit according to expression (3). As it can be observed the exponential function is a good approximation to the autocorrelation for time lags up to approximately 40 hours. The obtained autocorrelation coefficient of the exponential fit is $\alpha = 0.0504$.

As the Weibull distribution and the exponential function are good approximations for the normalized histogram and the autocorrelation of the wind speed data set, respectively, the three models previously developed are suitable to represent

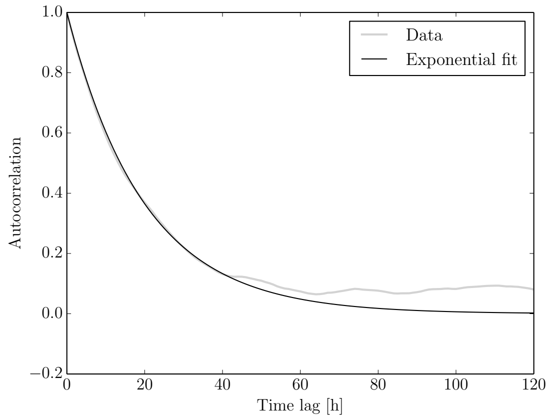

Fig. 2. Autocorrelation of the hourly mean wind speed data.

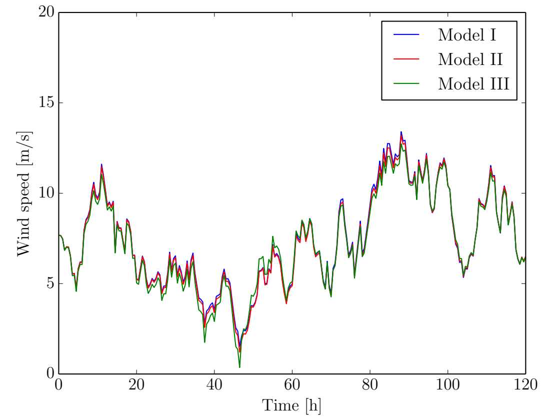

Fig. 3. Wind speed trajectories generated by Models I to III.

the wind speed for the considered location. However, since the wind speed is exponentially autocorrelated only for time lags up to 40 hours, the ability of the developed models to represent the wind speed behavior of the considered location is expected to be acceptable only for that time frame.

B. Simulations

The results obtained with the data analysis above is used to set up the three wind speed models. To obtain wind speed trajectories from the developed models is necessary to use numerical integration techniques for SDEs. For that, we used the multiprocessor stochastic integration tools available in the software Dome [15]. In particular, we applied the following implicit Milstein integration scheme:

$$x(t+h) = x(t) + \frac{h}{2}\left(a(x(t+h)) + a(x(t))\right) + b(x(t))\Delta W + \frac{1}{2}b(x(t))\frac{\partial b(x(t))}{\partial x(t)}\left((\Delta W)^2 - h\right) \qquad (22)$$

where $a(\cdot)$ and $b(\cdot)$ are, respectively, the drift and diffusion terms of the corresponding SDE, h is the integration time step, and $\Delta W \sim \mathcal{N}(0, \sqrt{h})$ are random increments of the Wiener process. The integration scheme (22) reduces to the implicit Euler scheme if the diffusion term $b(\cdot)$ is constant, as it is the case of Model I and Model III. The interested reader can find other integration schemes in [16].

The initial value of the integration scheme (22), i.e, the initial value of all the simulations carried out, is obtained by sampling a Weibull distribution with the parameters given is Subsection IV-A. The sampling of the Weibull distribution, together with the sampling of the normal variates that approximates the Wiener process ΔW, is performed with the pseudo-random number generator provided by the GNU Scientific Library (GSL) [17]. To perform fair comparisons, the simulations of each model are carried out using same seed in the pseudo-random number generator.

Figure 3 depicts three wind speed trajectories, where each trajectory was generated with one model. The trajectories are generated using Models I to III. The three trajectories share the initial value and the sequence of increments of the Wiener process. It can be observed that the trajectories are very close to each other but are not identical. A similar behaviour is observed when using other seed values. The question is whether the small differences observed in the single trajectories translate to noticeable differences in what respect to the statistical properties of the generated wind speed trajectories.

The statistical properties of the processes generated by each SDE-based model were obtained on the basis of multiple simulations considering a time frame of 120 hours. This time frame is appropriate to illustrate the exponential decay of the autocorrelation function, as it can be deduced from Figure 2. A set of 10000 wind speed trajectories was generated by means of each model. It is worth pointing out that in the case of Model III, 9.57% of the simulations were dismissed due to the appearance of wind speed negative values and/or numerical problems. None of these two circumstances were observed for Models I and II. Therefore, the following statistical analysis was performed based on 10000 simulations for Models I and II, and 9430 simulations for Model III.

Figure 4 compares the probability density of the trajectories generated by the Models I to III with the PDF of the Weibull distribution used to construct those same models. These probability densities were computed taking into account the wind speed values generated at the end of each simulation. It is observed that the differences between the curves are not significant from a statistical point of view, so it can be said that the models are able to reproduce this probability density.

Figure 5 depicts the comparison between the exponential function fitted to the autocorrelation of the wind speed data set and the autocorrelation of the trajectories generated by each model. As in the case of the probability density, all models are able to reproduce the autocorrelation exponential decay without significant deviations. This is an unexpected result in what refers to Model III, since it is not theoretically desmonstrated. Regarding Model I, althought obtaining an

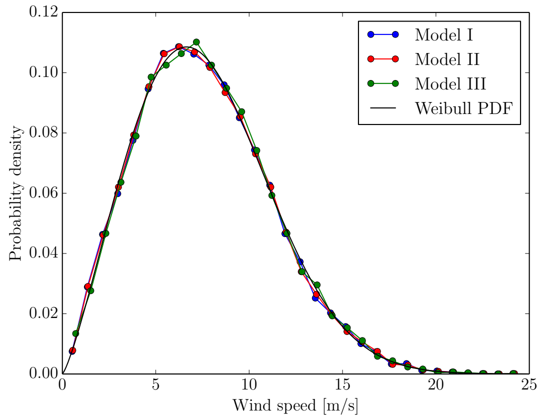

Fig. 4. Probability density of the wind speed generated by Models I to III.

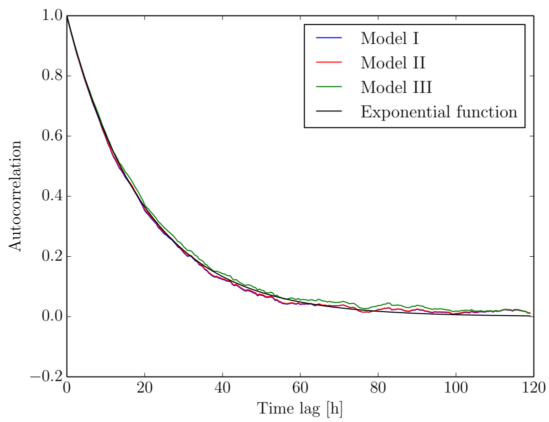

Fig. 5. Autocorrelation of the wind speed generated by Models I to III.

exponential autocorrelation with the given decay rate is not guaranteed, it can be observed that the model is able to aproximate this property with accuracy.

V. Conclusions

This paper describes three methods to construct wind speed models based on SDEs. Such three methods are applied leading to models able to generate stochastic processes with Weibull distribution and exponential autocorrelation. The statistical analysis of a wind speed data set collected in a weather station in Ireland has served to set up the three models according to real-world parameters. Simulation results of the three models reveal that they provide wind speed trajectories with the statistical properties for which they were designed. Although the results obtained with each model are not exactly identical in what refers to trajectories, it can be said that the three models are statistically equivalent. This is not an obvious result, since the formulation of each model is substantially different. Finally, caution should be taken when using Model III since it can produce trajectories with numerical inconsistencies as, for example, negative wind speed values. Future work will focus on the application of the presented approached to develop wind speed models for probability distributions other than the Weibull distribution.

Acknowledgment

The authors would like to thank Met Éireann for kindly providing the wind speed data used in the paper.

The second and third authors are partly funded by the Science Foundation Ireland, under Grant No. SFI/09/SRC/E1780. The opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Science Foundation Ireland. The third author has also benefit from the financial support of EC Marie Skłodowska-Curie Career Integration Grant No. PCIG14-GA-2013-630811.

References

[1] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, May 2009.

[2] J. A. Carta, P. Ramírez, and S. Velázquez, "A review of wind speed probability distributions used in wind energy analysis. Case studies in the Canary Islands," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 5, pp. 933–955, Jun. 2009.

[3] A. C. Brett and S. E. Tuller, "The autocorrelation of hourly wind speed observations," *Journal of Applied Meteorology*, vol. 30, no. 6, pp. 823–833, Jun. 1991.

[4] R. Calif and F. G. Schmitt, "Multiscaling and joint multiscaling description of the atmospheric wind speed and the aggregate power output from a wind farm," *Nonlinear Processes in Geophysics*, vol. 21, no. 2, pp. 379–392, 2014.

[5] Z. Y. Dong, J. H. Zhao, and D. J. Hill, "Numerical simulation for stochastic transient stability assessment," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1741–1749, Nov. 2012.

[6] F. Milano and R. Zárate-Miñano, "A systematic method to model power systems as stochastic differential algebraic equations," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4537–4544, Nov. 2013.

[7] S. V. Dhople, Y. C. Chen, L. DeVille, and A. D. Dominguez-Garcia, "Analysis of power system dynamics subject to stochastic power injections," *IEEE Transactions on Circuits and Systems - I: Regular Papers*, vol. 60, no. 12, pp. 3341–3353, Dec. 2013.

[8] R. Calif, "PDF models and synthetic model for the wind speed fluctuations based on the resolution of the Langevin equation," *Applied Energy*, vol. 99, pp. 173–182, Nov. 2012.

[9] R. Zárate-Miñano, M. Anghel, and F. Milano, "Continuous wind speed models based on stochastic differential equations," *Applied Energy*, vol. 104, pp. 42–49, Apr. 2013.

[10] R. Zárate-Miñano and F. Milano, "Construction of SDE-based wind speed models with exponential autocorrelation," 2015, available at http://arxiv.org/.

[11] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*. New York, NY, 4th edition: Springer-Verlag, 2009.

[12] M. Grigoriu, *Stochastic Calculus. Applications in Science and Engineering*. Boston, MA: Birkhäuser, 2002.

[13] ——, "Simulation of stationary non-Gaussian translation processes," *Journal of Engineering Mechanics*, vol. 124, no. 2, pp. 121–126, Feb. 1998.

[14] Met Éireann, "Met Éireann, The Irish National Meteorological Service," http://www.met.ie/, Accessed November-2015.

[15] F. Milano, "A Python-based software tool for power system analysis," in *IEEE PES General Meeting*, Vancouver, Canada, Jul. 2013, pp. 1–5.

[16] E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. New York, NY, 3rd edition: Springer, 1999.

[17] GNU Scientific Library - Reference Manual, available at http://www.gnu.org/software/gsl/.