

On the Uncertainty of Wind Power Forecasts

Waleed Alhaddad

King Abdullah University of Science (KAUST)

waleed.alhaddad@kaust.edu.sa

Abstract

Reliable wind power generation forecasting is crucial to meet energy demand, to trade and invest. We propose a model to simulate and quantify uncertainties in such forecasts. This model is based on Stochastic Differential Equations whose time-dependent parameters are inferred using continuous optimization of an approximate Likelihood function. The result is a skew stochastic process that simulates the uncertainty of wind power forecasts accounting for maximum power production limit and other temporal effects. We apply the model to historical Uruguayan data and forecasts.

1. Introduction

We propose a model to simulate and quantify uncertainties in wind power generation forecasts. This model is based on Stochastic Differential Equations (SDEs) whose time-dependent parameters are inferred using optimization techniques of an associated approximate Likelihood function. Through continuous optimization, we find the optimal parameters of an unbounded convex problem with convex constraints.

We are able to simulate and quantify uncertainties in a variety of wind power generation forecasts while taking into account the skewness of the errors. This method is non-intrusive and is independent of forecasting technology or future developments. Through optimization, we update and tune the parameters of our SDE as we receive new sets of observations and their associated forecasts. Additionally, we are able to compare in a quantitative manner the different forecast technologies and how they behave in different real-world scenarios. This model is to be extended to the uncertainty quantification of other power generation sources such as the uncertainties in the power generation of solar power plants. This introduces new challenges in terms of optimization and modeling. Finally, we apply our model to synthetic and real wind power generation data for benchmarking. We apply the model to Uruguayan wind power production as determined by historical data and correspond-

ing numerical forecasts for the year 2016-2017.

2. Model

2.1. Physical Model

Let X_t be the wind power generation forecasts stochastic process defined by the following parameterized stochastic differential equation (SDE),

$$\begin{aligned} dX_t &= a(X_t; \theta)dt + b(X_t; \theta)dW_t \quad t > 0 \\ X_0 &= X_0 \end{aligned} \quad (1)$$

- $a(X_t; \theta)$: Drift function
- $b(X_t; \theta)$: Diffusion function
- θ : a vector of parameters
- dW_t : Standard Wiener random process.

2.2. Physical Restrictions

It is necessary to keep the power generation forecast plus its errors inside the range $[0, 1]$ after normalizing the power generated to the capacity of the power plant. To ensure that, we choose our model to have zero diffusion near the boundaries. Hence, we formulate the following diffusion term,

$$\begin{aligned} b(V_t; \theta) &= \sqrt{2\theta\alpha p_t(1-p_t)X_t(1-X_t)} \\ &= \sqrt{2\theta\alpha p_t(1-p_t)(V_t + p_t)(1-V_t - p_t)} \end{aligned}$$

It remains to control the drift term. As a first requirement, we need the drift term to be mean reverting to the power generation forecast. Thus, we formulate the following drift,

$$a(V_t; \theta) = -\theta(X_t - p_t) \quad (2)$$

However, this does not track the rate of change of the forecast \dot{p} which results in X_t being out of sync with the forecast as can be seen in Figure (1)

For clarity of exposition, we switch to the point of view of the forecast prediction process $X_t = V_t + p_t$ and we include the derivative tracking in the base model,

$$\begin{aligned} dX_t &= \hat{a}(X_t; \theta)dt + \hat{b}(X_t; \theta)dW_t \quad t > 0 \\ X_0 &= x_0 \end{aligned} \quad (3)$$

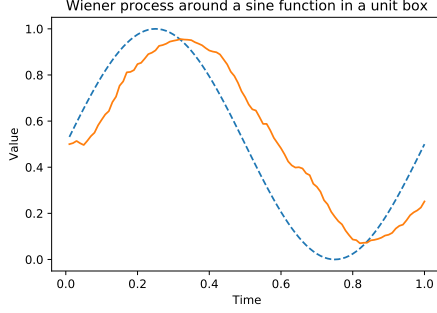


Figure 1. Example of the process X_t following a sine wave without derivative tracking.

where

$$\begin{aligned}\hat{a}(X_t; \boldsymbol{\theta}) &= \dot{p}_t + a(X_t; \boldsymbol{\theta}) = \dot{p}_t - \theta (X_t - p_t) \\ \hat{b}(X_t; \boldsymbol{\theta}) &= b(X_t; \boldsymbol{\theta}) = \sqrt{2\theta\alpha p_t(1-p_t)X_t(1-X_t)}\end{aligned}\quad (4)$$

Here the term \dot{p}_t is not controlled to maintain that X_t stays inside the range $[0, 1]$. In other words, the zero drift line defined by setting $\hat{a}(X_t; \boldsymbol{\theta}) = 0$ must be contained inside the range $[0, 1]$. This line is given by $\frac{\dot{p}_t}{\theta} + p_t$, we call it the line of zero drift or the line of mean stationarity. Thus, must have that,

$$0 \leq \frac{\dot{p}_t}{\theta} + p_t \leq 1 \quad (5)$$

to keep the process X_t inside the interval $[0, 1]$. The condition can be rewritten as,

$$\frac{-|\dot{p}_t|}{p_t} \leq \theta \leq \frac{|\dot{p}_t|}{1-p_t} \quad (6)$$

To enforce the condition, we define an adjusted drift $\theta = \theta_0 f(p_t, \dot{p}_t)$ which satisfies the condition when

$$\theta = \theta_0 f(p_t, \dot{p}_t) \leq \frac{|\dot{p}_t|}{\min(p_t, 1-p_t)} \quad (7)$$

Hence, we choose θ such that,

$$\theta = \max\left(\theta_0, \frac{|\dot{p}_t|}{\min(p_t, 1-p_t)}\right) \quad (8)$$

2.3. Model prescription

We conclude that in order to respect the physical restrictions discussed above, we prescribe the following specifications and the final model becomes,

$$\begin{aligned}dV_t &= a(V_t; \boldsymbol{\theta})dt + b(V_t; \boldsymbol{\theta})dW_t \quad t > 0 \\ V_0 &= v_0\end{aligned}\quad (9)$$

$$\begin{aligned}a(V_t; \boldsymbol{\theta}) &= -\theta V_t \\ b(V_t; \boldsymbol{\theta}) &= \sqrt{2\theta\alpha p_t(1-p_t)(V_t + p_t)(1-V_t - p_t)}\end{aligned}\quad (10)$$

with

$$\theta = \max\left(\theta_0, \frac{|\dot{p}_t|}{\min(p_t, 1-p_t)}\right) \quad (11)$$

Where

- p_t : Numerical wind power generation forecast.
- $\theta > 0$: Mean reversion parameter.
- $\alpha > 0$: Variability parameter.

The SDE above defines the stochastic process V_t which can be characterized by its transition density. Consider a set of M paths with N observations each, $V^{M,N} = \{V_{t_1}^{M,N}, V_{t_2}^{M,N}, \dots, V_{t_N}^{M,N}\}$ observed in intervals of Δ_N . Then, the transitions of V_t are given by the likelihood function,

$$\mathcal{L}(\boldsymbol{\theta}; V) = \prod_{j=1}^M \prod_{i=1}^N \rho(V_{j,i+1} | V_{j,i}, \boldsymbol{\theta}) \rho(V_{j,0}) \quad (12)$$

The standard approach is approximating (12) by Gaussian transitions. However, this is inadequate in our case as we expect a skew symmetric stochastic process due to physical restrictions. Thus, we approximate the transitions by a Beta distribution. This is possible by matching the moments of the Beta distributions with the moments of the SDE defined above. The moments are given by,

$$\begin{aligned}\mathbb{E}[V_t^k] &= \mathbb{E}\left[\left(\int_0^t a(V_i; \boldsymbol{\theta}_i)dt + \int_0^t b(V_i; \boldsymbol{\theta}_i)dW_i\right)^k\right] \quad t > 0 \\ V_0 &= v_0\end{aligned}\quad (13)$$

The above moments define a system of Ordinary Differential Equations,

$$\begin{aligned}\frac{dm_1(t)}{dt} &= -\theta_t m_1(t) \\ \frac{dm_2(t)}{dt} &= -2m_2(t)[\theta_t + \alpha_t \theta_t p_t(1-p_t)] \\ &\quad + 2m_1(t)[\alpha_t \theta_t p_t(1-p_t)(1-2p_t)] \\ &\quad + 2\alpha_t \theta_t p_t^2(1-p_t)^2\end{aligned}\quad (14)$$

With initial conditions,

$$\begin{cases} m_1(0) = \mathbb{E}[V_0] = v_0 \\ m_2(0) = \mathbb{E}[V_0^2] = v_0^2 \end{cases} \quad (15)$$

where $m_1(t)$ and $m_2(t)$ are the first and second moments of the stochastic process V_t . As we are dealing with numerical wind power generation forecasts p_t , we do not have a closed form solution to this system of ODEs. Thus, this system is iteratively solved by means of numerical integration.

Once the moments $m_1(t)$ and $m_2(t)$ are obtained, we match them with the shape parameters of the Beta distribution. To allow for negative and positive values of the process V_t , we formulate a modified Beta distribution whose support is defined on $[a, b] = [-1, 1]$. We use the linear transformation $y = g(x) = a + (b - a)x$ and we obtain the probability density function,

$$\begin{aligned} f_Y(y; \alpha, \beta) &= f_X(g^{-1}(y)) \left| \frac{\partial g^{-1}(y)}{\partial y} \right| \\ &= \frac{1}{B(\alpha, \beta)(b-a)} \left(\frac{a-y}{b-a} \right)^{\alpha-1} \left(1 - \frac{a-y}{b-a} \right)^{\beta-1} \end{aligned} \quad (16)$$

where $B(\cdot, \cdot)$ is the Beta function. Also, we find that

$$\mathbb{E}[Y] = a + (b-a) \frac{\alpha}{\alpha + \beta} \quad (17)$$

$$\mathbb{V}[Y] = (b-a)^2 \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (18)$$

And by matching the moments, we find that

$$\alpha = \frac{(\tilde{\mu}_t - a)(\tilde{\mu}_t^2 - \tilde{\mu}_t(a+b) + \tilde{\sigma}_t^2 + ab)}{\tilde{\sigma}_t^2(a-b)} \quad (19)$$

$$\beta = \frac{(\tilde{\mu}_t - b)(\tilde{\mu}_t^2 - \tilde{\mu}_t(a+b) + \tilde{\sigma}_t^2 + ab)}{\tilde{\sigma}_t^2(b-a)} \quad (20)$$

where $\tilde{\mu}_t = m_1(t)$ and $\tilde{\sigma}_t^2 = m_2(t) - m_1^2(t)$

2.4. Likelihood Optimization

Combining the previous results and (12), we can express the log-likelihood function for M paths,

$$\begin{aligned} \ell(\alpha, \beta | V) &= \sum_{j=1}^M \sum_{i=1}^N (\alpha - 1) \log \left(\frac{V_{j,i+1} + 1}{2} \right) \\ &\quad + (\beta - 1) \log \left(1 - \frac{V_{j,i+1} + 1}{2} \right) - \log(2B(\alpha, \beta)) \end{aligned} \quad (21)$$

where α and β are given by,

$$\alpha = -\frac{(1 + \tilde{\mu}_t)(\tilde{\mu}_t^2 + \tilde{\sigma}_t^2 - 1)}{2\tilde{\sigma}_t^2} \quad \beta = \frac{(\tilde{\mu}_t - 1)(\tilde{\mu}_t^2 + \tilde{\sigma}_t^2 - 1)}{2\tilde{\sigma}_t^2} \quad (22)$$

We will show that the log-likelihood function in (21) is concave. Note that it is sufficient to check the log-likelihood of the standard Beta distribution for concavity since the modified Beta distribution log-likelihood is obtained through non-negative weighted summations and a

linear transformation $g(x)$ in the given data that does not affect the parameters of interest. Thus, the concavity is preserved.

We study the concavity of log-likelihood of the standard Beta,

$$\ell^*(\alpha, \beta | x) = -\log(B(\alpha, \beta)) + (\alpha - 1)\log(x) + (\beta - 1)\log(1 - x)$$

It has been shown that the Beta function $B(\alpha, \beta)$ is log-concave in both its arguments on $(0, \infty)^2$ (see [1]) and the remaining terms are linear in α and β . Then, we conclude that the likelihood in (12) is indeed log-concave.

In order to infer the shape parameters α and β , we perform the following constrained maximum likelihood optimization,

$$(\alpha^*, \beta^*) = \arg \min_{\alpha > 0, \beta > 0} -\ell(\alpha, \beta | V^{M,N}) \quad (23)$$

We solve the above optimization problem using L-BFGS to avoid full numerical computation of the Hessian matrix.

2.5. Data

We apply the aforementioned method to Uruguayan Windpower Generation data and their corresponding numerical wind power production forecast.

The wind power generation data set contains hourly samples of aggregated wind power production throughout the country. The samples are normalized according to the full installation power which varies across the data set. Each sample path contains 72 hourly observations and we have a total of 1217 sample paths spanning the year 2016 to 2017 (87,624 data points).

The numerical wind power generation forecast set contains 1217 forecasts of the wind power generation paths corresponding to the wind power generation data set mentioned above.

2.6. Results

We are able to obtain the parameters based on the complete data sets mentioned earlier. The parameters are given by $(\theta^*, \alpha^*) \approx (8, 1)$. We show in Figure(??) the contours of the log-likelihood in (21) using the full data set.

In Figure(4) and (3), we simulate sample forecast errors using the optimal parameters $(\theta^*, \alpha^*) \approx (8, 1)$ and obtain the associated confidence bands empirically using 10,000 sample paths for each forecast. Note that the confidence bands obtained are non-trivial and reflect the complexity of the uncertainties in the numerical wind power generation forecasts.

2.7. Conclusion

In this project, we have proposed a model to quantify the uncertainties in wind power generation forecasts. It is important to note that this approach is agnostic of forecasting

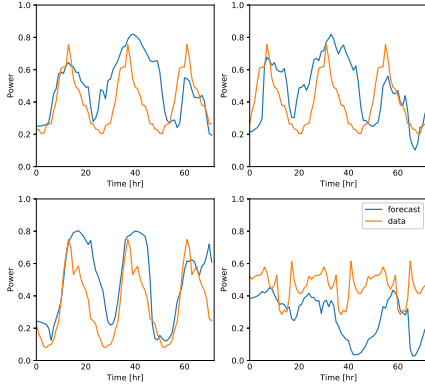


Figure 2. Examples of wind power generation from the data set along with the wind power generation forecast.

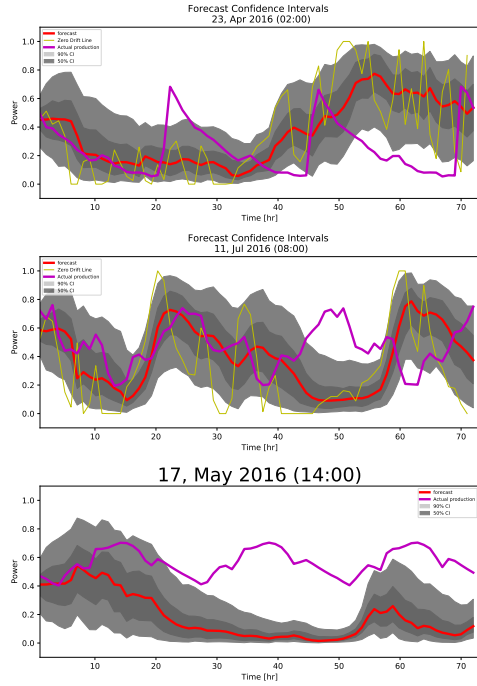


Figure 3. Examples of confidence bands obtained for the full 72 hour forecasts.

technology. We have also taken into account the physical limits of the model by respecting the full installation power of the wind farm. Additionally, we have taken into account skew-symmetric errors for which the Beta distribution is an appropriate choice. The skewness of the forecast errors is necessary to correctly model the errors when operating near full installation power production or in cases of extremely low wind power production.

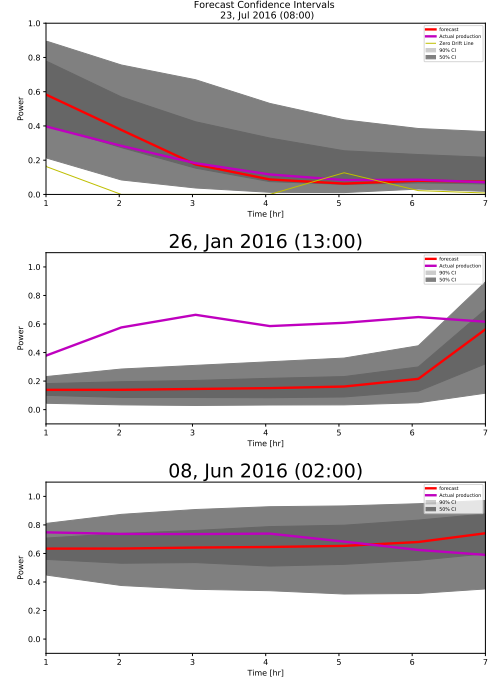


Figure 4. Examples of confidence bands obtained for the first 6 hours of the forecasts. This is important as this specific forecasting company recomputes a new forecast every 6 hours.

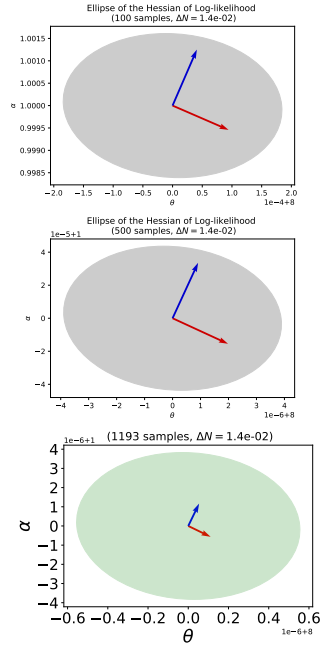


Figure 5. Shrinkage of the ellipse determined by the Hessian matrix of the log-likelihood around the point of optimality $(\theta, \alpha) \approx (8, 1)$

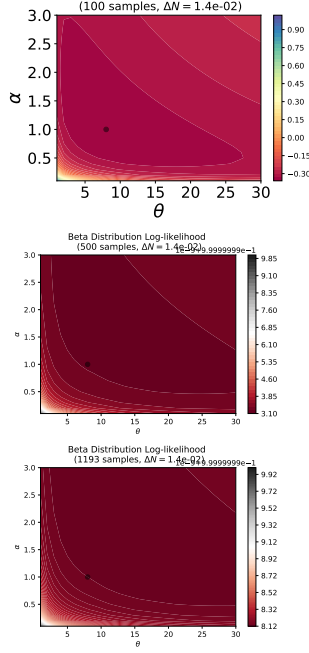


Figure 6. Contour plot of the log-likelihood as more forecasts and data is added, point of optimality $(\theta, \alpha) \approx (8, 1)$ indicated in black.

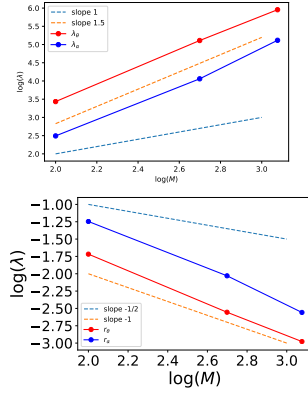


Figure 7. Convergence of the major and semi-axis of the ellipse of the Hessian of the log-likelihood at the point optimality $(\theta, \alpha) \approx (8, 1)$. Note that it is slightly faster than the expected rate of $1/\sqrt{M}$. This due to the fact that segments of some of the M are uncorrelated, thus acting as an additional uncorrelated sample to the rest of the path.

2.8. References

- [1] Dragomir, S., Agarwal, R., & Barnett, N. (2000). Inequalities for Beta and Gamma functions via some classical and new integral inequalities. *Journal of Inequalities and Applications*, 2000(2), 504054. doi:10.1155/s1025583400000084