Probabilistic Forecasts of Wind Power Generation by Stochastic Differential Equation Models.

Jan Kloppenborg Møller[1,3], Pierre Pinson[2], and Henrik Madsen[1]
[1] DTU Compute, Technical University of Denmark, Denmark.
[2] DTU Electrical Engineering, Technical University of Denmark, Denmark.
[3] Corresponding author: Jan Kloppenborg Møller, e-mail: `jkmo@dtu.dk`

Abstract

The increasing production of renewable energy, and in particular wind energy, introduces highly volatile sources of energy in the total production. This implies that methods for reliable probabilistic forecasts of future wind power production are essential.

Today there exist numerous methods and tools for providing point forecasts of wind power generation. However, for efficient and safe regulation, and for harvesting optimal trading strategies reliable information on the uncertainty is also needed. In this paper we focus on forecasts on the 1-48 hour horizon. It is well-known that the form of the conditional density for the wind power production is highly dependent on the level of predicted wind power in addition to the prediction horizon. This paper describes a new approach for wind power forecasting based on state dependent stochastic differential equations (SDEs). Specifically we will use a logistic type stochastic differential equation to account for the natural restrictions (wind power cannot exceed installed capacity and cannot be below zero). The SDE is driven by a widely used point predictor for wind power forecast, and the SDE formulation allows us to calculate both state dependent conditional uncertainties as well as correlation structures. Evaluation and optimization of the model is obtained by evaluating the likelihood of a 48-dimensional random vector when accounting for the correlation structure defined by the SDE-formulation. We explore the correlation parameters and skewness of the model and input-variables (prediction horizon and point predictions), by a non-parametric (spline based) model for the parameters.

Keywords: nonlinear forecasting, state space model, stochastic differential equations, wind power prediction.

1 Introduction

In 2012 wind power accounted for about 30 % of the total demand for electricity in Denmark. Integrating this world leading amount of wind power into the power systems calls for state-of-the-art methods for wind power forecasting. Methods for generating point forecasts (e.g. Giebel et.al., 2011) of wind power for operational decision-making have been used since 1995. Today, however, decision-making and operational optimization problems can be characterized by assymmetric and time-varying cost functions. This implies that state-of-the-art tools must be able to provide a reliable description of full predictive densities, with lead times ranging from minutes to days.

Still, even after about 20 years of research, wind power forecasting remains a challenge from a statistical point of view. This is firstly due to the nonlinear and double-bounded nature of the stochastic process describing the variation in time of wind power production (Pinson,

2012; Trombe, 2012). In a probabilistic framework, conditional predictive densities may be generated in a parametric framework (e.g. Pinson, 2012; Trombe, 2012), or alternatively in a nonparametric framework (e.g. Møller, 2008; Nielsen, 2006). Trajectories of wind power generation may be based on Gaussian copula (Pinson, 2009) or by nonlinear transformation of ensemble forecasts (Nielsen, 2004) of relevant meteorological variables.

In this paper a new method based on stochastic differential equations (SDEs) is suggested. The advantage of this method is that by selecting a proper specification and parameterization of the SDE, basically all the required properties, like nonlinearities, double bounded variations, and varying quantiles, can be described. The multivariate distribution of wind power data will (after normalised with installed capacity) be contained in $[0,1]^n$, where n is the dimension of the forecast (in this paper we condider $n = 48$).

The scope is to demonstrate the improvements that can be achieved by considering SDE's, and we assume that the general distribution can be approximated by a multivariate normal distribution, describing the variations of prediction erreros around point predictions by the widely used Wind Power Prediction Tool (WPPT, see (Madsen, 1998; WPPT, 2012)).

2 Data

The data studied here consist of hourly averages of wind-power production from Klim wind power plant located in the northern part of Denmark. Production data are normalised with the installed capacity (21MW). Every 6th hour a new forecast based on a 48 hour weather forecast is issued, one observation in our setting consist of a 48 hour forecast and the corresponding measurements.The data set cover the period March 2001 through April 2003 (2593 observations). The SDE estimation is quite time consuming and it was therefore decided to base the estimation on 150 observations (chosen to span the situations in data) only. The observations not in the training set is used as our test set.

3 SDE-models

The continuous-discrete time stochastic state space model consist of a SDE describing the dynamics, and an observation equation. We want the deterministic part of the model to have the stationary solution $x_{i,t} = \hat{p}_{i,t}$ (with $\hat{p}_{i,t}$ being the WPPT forecast). The diffusion is allowed to depend on the state and the input, and since we want the diffusion to die off at the boundaries a minimal requirement is $\sigma(1, \hat{p}_{i,t}, t) = \sigma(0, \hat{p}_{i,t}, t) = 0$. Which is fulfilled for logistic type SDEs (e.g. Schurz, 2007). Here we select the special case

$$dx_{i,t} = -\theta_{i,t} \cdot (x_{i,t} - \hat{p}_{i,t})dt + \sqrt{2\theta_{i,t}a_{i,t}\hat{p}_{i,t}(1 - \hat{p}_{i,t})x_{i,t} \cdot (1 - x_{i,t})}dw_t. \tag{1}$$

For each fixed value of $\theta_{i,t}$ and $\hat{p}_{i,t}$ this have a stationary solution for $a_{i,t} \in [0,1)$ (see e.g. Forman and Sørensen, 2008). We want to transform (1) in such a way that the transformed system is independent of the state. In order to do this we use the Lamperti transformation (e.g. Baadsgaard et. al., 1997; Iacus, 2008), $z_{i,t} = \arcsin(2x_{i,t} - 1)$, with $z_{i,t} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The first main result of the paper is summarised in the following theorem.

Theorem 1 *The Lamperti transformed process corresponding to* (1)*, is given by*

$$dz_{i,t} = \tilde{f}(z_{i,t}, \hat{p}_{i,t}, t)dt + \sqrt{2 \cdot \theta_{i,t}a_{i,t}\hat{p}_{i,t}(1 - \hat{p}_{i,t})}dw_t. \tag{2}$$

with

$$\tilde{f}(z_{i,t},\hat{p}_{i,t},t) = \frac{2\left(-\theta_{i,t}\cdot(\frac{1}{2}(1+\sin(z_{i,t}))-\hat{p}_{i,t})+\frac{1}{2}\sin(z_{i,t})\cdot a_{i,t}\cdot\theta_{i,t}\hat{p}_{i,t}(1-\hat{p}_{i,t})\right)}{\cos(z_{i,t})}. \tag{3}$$

Moreover the second order moment representation of the linearised process can be solved for $a_{i,t} \in [0,2)$.

PROOF. Omitted

The restriction on a is therefore relaxed and we consider the SDE-formulation

$$dx_{t,i} = -\theta\cdot(x_{t,i}-\hat{p}_{t|0,i})dt + 2\sqrt{\theta\alpha\hat{p}_{t|0,i}(1-\hat{p}_{t|0,i})x_{t,i}\cdot(1-x_{t,i})}dw_{t,i} \tag{4}$$

with $\alpha \in [0,1)$. The "bias" of the Lamperti-transformed SDE $(\alpha\hat{p}_{t|0}(1-\hat{p}_{t|0})(1-2x_t))$ points in the "wrong" direction, and a compensator is therefore introduced

$$dx_{t,i} = -\theta\cdot(x_{t,i}-\hat{p}_{t|0,i}-c\hat{p}_{t|0,i}(1-\hat{p}_{t|0,i})(1-2x_{t,i}))dt+$$
$$2\sqrt{\theta\alpha\hat{p}_{t|0,i}(1-\hat{p}_{t|0,i})x_{t,i}\cdot(1-x_{t,i})}dw_{t,i} \tag{5}$$

where $c \geq 0$ is a constant. Again the results are are given in a theorem.

Theorem 2 *The second order moment representation of the linearised Lamperti transformed process corresponding to (5) exist for* $\alpha \in [0,1)$ *and* $c \geq 0$.

The complete solution (second order moment representation of the Lamperti transformed process) is given by (for given $z_{0|0}$*)*

$$\hat{z}_t = \arcsin\left\{\left(\int_0^t \theta(2\hat{p}_{s|0}-1)e^{\int_0^s g_u du}ds + \sin(z_{0|0})\right)e^{-\int_0^t g_s ds}\right\} \tag{6}$$

where $g_t = \theta_t[1-2(1-c_t)\alpha_t\hat{p}_{t|0}\cdot(1-\hat{p}_{t|0})]$ *and the covariance is given by (with* $P_0 = 0$*)*

$$P_t = \int_0^t \sigma_s^2 e^{-2\int_0^s A_u du}ds\cdot e^{2\int_0^t A_s ds}; \quad with \quad A_t = \left.\frac{\partial f_z(z)}{\partial z}\right|_{z=\hat{z}_t} \tag{7}$$

PROOF. Omitted

In addition to the theorem above we use $\rho_{t,t+s} = e^{\int_t^{t+s} A_u du}$ to approximate the correlation structure. These equation are solved by numerical integration to obtain the mean $\hat{z}_i \in \mathbb{R}^{48}$ and a covariance matrix $P_i \in \mathbb{R}^{48\times48}$ for the i'th observation.

For all of the models the observation variance is constant (different for each model). The parameters α, θ and c are modelled by

$$\gamma(\hat{p}_{t|0},t) = \frac{K_\gamma}{1+exp(-\gamma_0-f_\gamma(\hat{p}_{t|0})-g_\gamma(t))}, \tag{8}$$

with $\gamma = \alpha, \theta$ or c where the functions in f_γ and g_γ are modelled by natural cubic splines. K_γ is used to control the range γ (we use $K_\alpha = 1$ and $K_\theta = K_c = 20$).

We further assume that the observations are given by

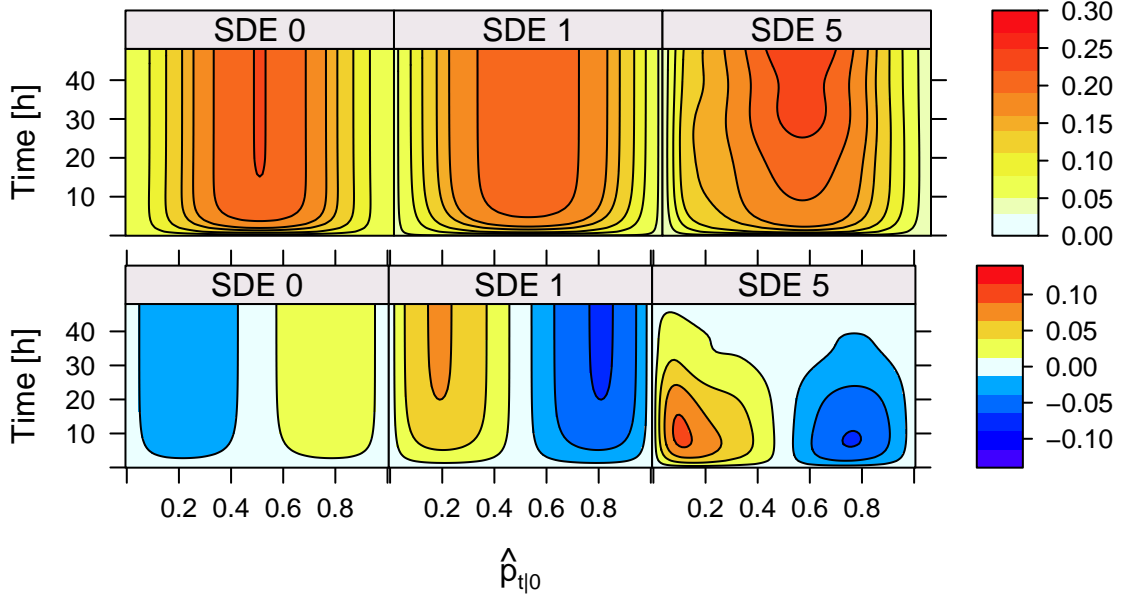$$p_i = \hat{x}_i + e_i = h(\hat{z}_i) + e_i, \tag{9}$$

Figure 1: Standard deviation (top row) and bias (bottom row) as a function of $\hat{p}_{t|0}$ and look ahead time (for $\hat{p}_{t|0}$ in time) for SDE models 0, 1, and 5.

where x_i is found by the integration described in the above and $e_i \sim N(0, \Sigma^e)$. Using the error propagation law and assuming that p_i is Gaussian we get

$$p_i \sim N\left(\frac{1}{2}(1 + \sin(\hat{z}_i)), C_i P_i C_i + \Sigma_i^e\right), \tag{10}$$

where $C_i = \frac{\partial h(\hat{z}_i)}{\partial \hat{z}_i} = \frac{1}{2}\mathrm{diag}(\cos(\hat{z}_i))$. The covariance function of e_i is modelled by $cov(e_{t_1,i}, e_{t_2,i}) = s^2 \rho(e_{t_1,i}, e_{t_2,i})$, where s^2 is a constant to be estimated and $\rho(e_{t_1,i}, e_{t_2,i})$ is determined from the correlation of the covariance matrix $C_i P_i C_i$.

4   Results

The models are found by including more and more structure in the parameters starting from the simple model as indicated in Table 1 (only a selection of the tested models are shown). In each step all possible inclusions are tested and the best model in terms of the likelihood is chosen. The simplest model (SDE 0) contain only a constant observation variance (s), a time constant (θ), and a diffusion parameter (α). Further inclusion of time varying parameters (either directly on time or by including function of $p_{t|0}$) gives large significant improvements in all steps (Table 1).

SDE-models 0 and 1 are both time-homogeneous, the standard deviation grow very fast in the first part of the series and find a stationary level after a few hours (Figure 1, top row), the effect of the compensator (c) is clear from Figure 1 (bottom row) where the bias is reversed, and the complicated structure introduced by time-varying parameter is illustrated by SDE-model 5.

Table 1: Value of the log-likelihood function for the training set and the test set.

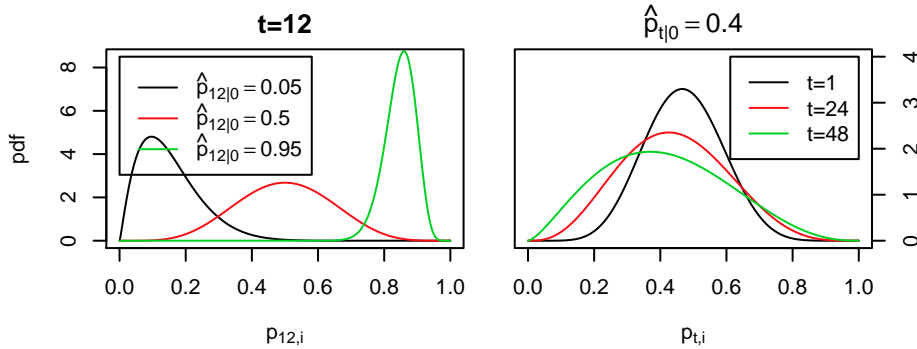|	df	α	θ	c	l(train)	p-value	l(test)
SDE 0	3	0.307	0.206	0	8408.4		161598
SDE 1	4	0.472	0.129	2.879	8532.4	< 0.0001	165011
SDE 5	24	f_α	$f_\theta + g_\theta$	g_c	8728.9	< 0.0001	**166296**

Figure 2: Stationary pdf for look ahead time (t) equal 12 hours for different predicted values of power (left column), and for predicted power ($\hat{p}_{t|0}$) equal 0.4 and different look ahead times (right column).

Even though the presented SDE's are always in the transient, it is illustrative to look at the stationary distribution. If the (time-varying) parameters are fixed the stationary distribution of the underlying SDE is a Beta distribution with parameters $\frac{1+c_t \cdot (1-\hat{p}_{t|0})}{2\alpha_t \cdot (1-\hat{p}_{t|0})}$ and $\frac{1+c_t \cdot \hat{p}_{t|0}}{2\alpha_t \cdot \hat{p}_{t|0}}$ (see Figure 2). The figure clearly illustrates the skewness obtained by the state dependent formulation of the diffusion.

5 Conclusion and Discussion

The starting point of the modelling is a quite simple SDE driven by a state of the art forecasting algorithm for prediction of the mean value of wind power. By considering natural restriction on the diffusion process we were able to formulate a set of ODEs to govern the second order moment representation of the wind power. The considered SDE models perform far better than complex benchmark models with larger number of parameter (not discussed here).

The SDE's offer a nice split between parameter that govern bias, variance and correlation. The focus in this paper is how to use SDE for scenario forecast when correlation is taken into account. In general it might be advisable to use some kind of transformation, e.g. using quantiles, before the SDE formulation is applied, but in order to keep the formulation and reasoning simple we have chosen not to consider transformations.

The SDE based state space formulation offers a simple and flexible framework for constructing very complex correlation or inter-dependence structures that are difficult to formulate explicitly in a direct parametric set up. The introduced framework for forecasting is able to provide an intergrated description of the inter-dependence structure as well as the sequence

of bounded and asymmetrical state-dependent predictive densities for the future wind power production.

The mapping from wind speed to wind power is very complex, and depends for instance on how dirty the blades of the turbines are. Hence, state-of-the-art methods for wind power forecasting of today, includes some methods for adaptive estimation of the parameters (e.g. Madsen, 2008). Such techniques must be integrated with the SDE approach before any operational state-of-the-art performance can be expected.

References

Baadsgaard, M., Nielsen, J. N., Spliid, H., Madsen, H., and Preisel, M. (1997) Estimation in stochastic differential equations with state dependent diffusion term. *SYSID '97 - 11th IFAC symposium on system identification, IFAC*

Forman, J., and Sørensen, M. (2008) The Pearson Diffusion: A Class of Statistically Tractable Diffusion Processes. *Scandinavian Journal of Statics*, **35**, 438-465

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M. and Draxl, C. (2011) *The State-of-the-Art in Short-Term Prediction of Wind Power: A Literature Overview*, The ANEMOS.plus Project, Paris, France

Iacus, S. M. (2008) Simulation and Inference for Stochastic Differential Equations - With R Examples. *Springer Series in Statistics, Springer New York.*

Madsen, H., Nielsen, T.S. and Tøfting, J. (1998) WPPT - A tool for online wind power prediction EPRI-DOE-NREL Wind Power Forecasting Workshop, San Francisco.

Madsen, H. Time Series Analysis *Chapman and Hall, Texts in Statistical Science, Boca Raton*

Møller, J.K., Nielsen, H.Aa., and Madsen, H. (2008) Time-adaptive quantile regression *Computational Statistics and Data Analysis*, **52**(3), 1292-1449

Nielsen, H.Aa., Madsen, H., Nielsen, T.S., Badger, J., Giebel, G., Landberg, Sattler, K., and Feddersen, H. (2004) Wind power ensemble forecasting Proceedings of the 2004 Global Wind Power Conference, Chicago.

Nielsen, H.Aa., Madsen, H. and Nielsen, T.S. (2006) Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts *Wind Energy*, **9**(2), 95-108

Pinson, P., Madsen, H., Nielsen, H.Aa., Papaefthymiou, G. and B. Klöckl (2009) From probabilistic forecasts to statistical scenarios of short-term wind power production *Wind Energy*, **12**(1), 51-62

Pinson, P. (2012) Very short-term probabilistic forecasting of wind power with generalized logit-Normal distributions *Journal of the Royal Statistical Society, Series C*, **61**(4), 555-576

Schurz H. (2007) Modelling, analysis and discretization of stochastic logistic equations *International journal of numerical analysis and modelling*, **4**(2), 178-197

Trombe, P-J, Pinson, P., and Madsen, H. (2012) A general probabilistic forecasting framework for offshore wind power fluctuations *Energies*, **5**, 621-657

Øksendal, B. (2003) Stochastic Differential Equations - An Introduction with Applications, Sixth edition. *Springer-verlag, Berlin*

WPPT (2012), http://enfor.eu/wind_power_prediction_tool_wppt.php