Obtaining Superior Wind Power Predictions from a Periodic and Heteroscedastic Wind Power Prediction Tool

Daniel Ambach

Carsten Croonenbroeck

European University Viadrina Frankfurt (Oder)

Department of Business Administration and Economics

Discussion Paper No. 361

September 2014

ISSN 1860 0921

Obtaining Superior Wind Power Predictions from a Periodic and Heteroscedastic Wind Power Prediction Tool

Daniel Ambach and Carsten Croonenbroeck

Abstract The Wind Power Prediction Tool (WPPT) has successfully been used for accurate wind power forecasts in the short to medium term scenario (up to 12 hours ahead). Since its development about a decade ago, a lot of additional stochastic modeling has been applied to the interdependency of wind power and wind speed. We improve the model in three ways: First, we replace the rather simple Fourier series of the basic model by more general and flexible periodic Basis splines (B-splines). Second, we model conditional heteroscedasticity by a threshold-GARCH (TGARCH) model, one aspect that is entirely left out by the underlying model. Third, we evaluate several distributional forms of the model's error term. While the original WPPT assumes gaussian errors only, we also investigate whether the errors may follow a Student's t-distribution as well as a skew t-distribution. In this article we show that our periodic WPPT-CH model is able to improve forecasts' accuracy significantly, when compared to the plain WPPT model.

1 Introduction

Compensation systems for renewable energy like wind energy are pluralistic through several countries. Many differences aside, accurate wind power forecasts are essential to the energy producer.
Research on wind power forecasting has been manifold. Lei et al. (2009) and, more recently, Giebel et al. (2011) provide an overview. There are models based on the physics of wind speed and power, models based on machine learning, wavelet mod-

Daniel Ambach
European University Viadrina, Große Scharrnstraße 59, 15230 Frankfurt (Oder), GERMANY, e-mail: ambach@europa-uni.de

Carsten Croonenbroeck
European University Viadrina, Große Scharrnstraße 59, 15230 Frankfurt (Oder), GERMANY e-mail: croonenbroeck@europa-uni.de

els and others. Simple, yet accurate stochastic models like the Wind Power Prediction Tool (WPPT), as introduced by Nielsen et al. (2007), are quite successful. WPPT is put to wide usage, especially in Denmark, the world leader in wind energy harvesting, as Giebel et al. (2011) point out. However, WPPT disregards several characteristics of the wind speed against wind power relationship. First, WPPT uses a Fourier series to capture diurnal periodicity. This is a straightforward way to model periodic effects, but replacing the Fourier terms by periodic B-spline functions introduces more flexibility. The idea to use periodic B-splines is inspired by Harvey and Koopman (1993), who use these functions to forecast the hourly electricity demand. These functions enable the model to follow the diurnal periodic structure independently from seasonal or yearly periodicity that may be present. Second, the residuals of the WPPT show a strongly heteroscedastic behavior, which also seems to be askew. We capture the skew (or: leveraged) heteroscedastic variance by modeling the error term as a TGARCH process, as introduced by Rabemananjara and Zakoïan (1993). Finally, the WPPT model's assumption of gaussian errors may be violated.[1] As the residual's density exhibits fat tails (particularly on its left-hand side), we investigate alternative distributional assumptions. After all, we show that our periodic WPPT-CH model generates forecasts that perform significantly more accurate than those obtained by the plain WPPT model.
This article is organized as follows. In Section 2, we describe the analyzed data. Section 3 introduces the underlying WPPT model and our new periodic wind power prediction model with TGARCH effects. The results of the in-sample fit and out-of-sample predictions are presented in 4 and section 5 provides a short conclusion.

2 Description of the Wind Power Data

The data used in this study are collected from a Fuhrländer FL MD 77 Turbine in Germany. Due to a non-disclosure agreement, the specific location cannot be revealed. Wind speed, wind direction and wind power is recorded at a frequency of 10 minutes. This Turbine exhibits a power range of $[0;1500]$ kW. The observed time frame for the training data set spans from October 31, 2010 to August 19, 2011 (40000 observations).

3 A New Wind Power Forecasting Method

WPPT, as given by Nielsen et al. (2007), models wind power P_t as a dynamic regression approach. It includes lagged wind speed and diurnal periodicity as regressors. The periodic behavior is captured by a Fourier series. Clearly, the important wind speed forecasts are not deterministic. They may stem from numerical weather pre-

[1] Shapiro-Wilk-Tests generally reject the hypothesis of gaussian WPPT errors.

dictions (NWP) or could be predicted from statistical model approaches. In this article, we use the recently developed predictions from a periodic ARFIMA-APARCH model with time varying regressors, as discussed by Ambach and Schmid (2014). As WPPT includes only two lags of wind speed in addition to periodicity as explanatory variables, it is not flexible enough to capture important features of wind power. The high-frequency data investigated here clearly show a strong presence of autocorrelation, see Figure 1. The autoregressive model order should be extended and include multiple lags. Furthermore, the variance structure shows heteroscedastic disturbance, see Figure 2. Therefore, we model the conditional standard deviation by a TGARCH model.

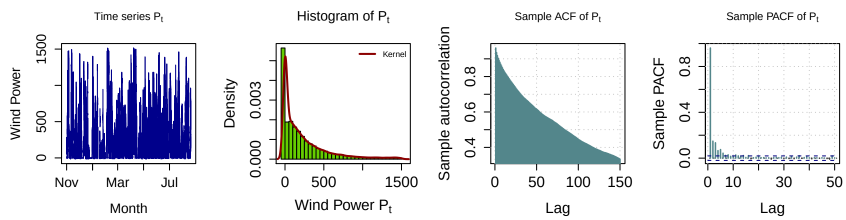

Fig. 1 Time series of wind power P_t (first panel), histogram of wind power (second panel), ACF and PACF for wind power (third and fourth panel), time frame October 31, 2010 to August 19, 2011.

Furthermore, we include wind direction as an additional explanatory variable, as due to the Turbine's uneven surroundings, local wind speed may depend on wind direction. Wind power shows a correlation structure that suggests that it may very well be dependent on wind direction. Figure 3 provides evidence on that assumption.

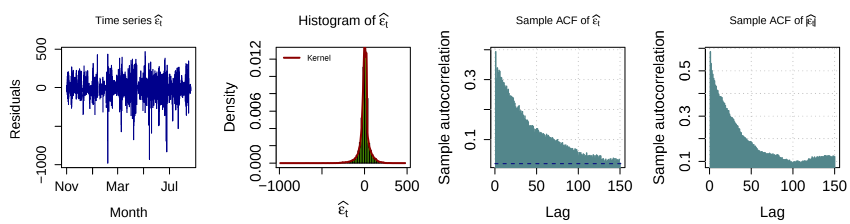

Fig. 2 WPPT residuals $\widehat{\varepsilon_t}$ (first panel), histogram of $\widehat{\varepsilon_t}$ (second panel), ACF of $\widehat{\varepsilon_t}$ (third panel) and ACF of $\widehat{|\varepsilon_t|}$ (fourth panel), time frame October 31, 2010 to August 19, 2011.

Our new wind power prediction model uses B-splines to model the periodic structure, instead of the Fourier series used in WPPT. The considered periodic basis functions are inspired by Harvey and Koopman (1993). The B-spline approach uses

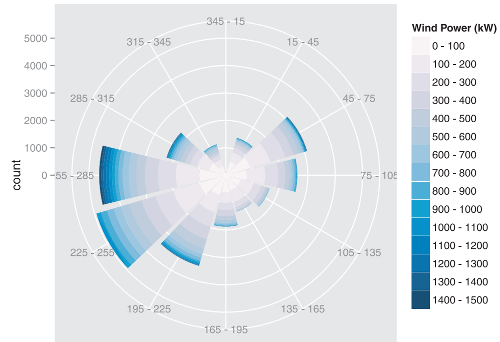

Fig. 3 Absolute number of wind power observations and the perceived wind directions.

local basis functions to provide more flexibility, especially when modeling time series with nonstationary impacts such as wind power. After all, our new model suggestion is

$$P_t = \vartheta_0 + \sum_{j=1}^{m} \theta_i P_{t-j} + \theta_{144} P_{t-144} + b_1 W_t + b_2 (W_t)^2 \tag{1}$$

$$+ \sin(A_t) + \cos(A_t) + \sum_{k=1}^{K} \widetilde{B}_k(t) + \sum_{r=1}^{o} \phi_r \varepsilon_{t-r}, \tag{2}$$

where $j \in 1,\ldots,m \backslash \{144\}$, W_t is the wind speed (in m/s) and A_t is the wind direction (Azimuth). Wind direction is measured in degrees. To avoid numerical problems resulting from that, we split the wind direction information into two components $cos(A_t)$ and $sin(A_t)$. The autoregressive coefficients θ provide a stationary solution if and only if $P(z) = 1 - (\sum_j^m \theta_i z^i + \theta_{144} z^{144}) \neq 0$ for $|z| \leq 1$. ϕ represents the moving average components. We include MA coefficients to reduce the parameters of the model and to capture the strongly persistent behavior of the dependent variable. The essential model enhancements are the periodic B-splines $\widetilde{B}_k(t)$. For our definition of these splines, we follow Ziel and Steinert (2014). The basis functions of the wind power series are given by

$$\widetilde{B}_k(t) = \sum_{k \in \mathbb{Z}} B_{kS,d_\kappa}(t) = \sum_{k \in \mathbb{Z}} B_{kS,d_\kappa}\left(t; \kappa\left(d_\kappa, T, D\right), D\right), \tag{3}$$

where κ is a set of equidistant knots, T depicts the central point of this set of knots, S represents the periodicity and D is the degree of the spline. The distance of the equidistant knots is given by d_κ. As the data frequency is 10 minutes, the diurnal periodicity for the wind power is $S = 144$. It is reasonable to choose the number of included basis functions to be a common denominator of S. Thus, we decide to use $\lambda = 6$ basis functions. Hence, we obtain a distance d_κ of $S/\lambda = 24$. Finally, we are able to iteratively define the complete set of basis functions $\widetilde{B}_k(t) = \widetilde{B}_{k-1}(t-24)$, where $k \in \{1,...,5\}$. Besides, we choose $D = 3$, to get the popular cubic splines, which are twice continuously differentiable. Furthermore, we have to remark that $\sum_{k=1}^{K} \widetilde{B}_k(t)$ is constant. Hence, the last component is omitted to avoid singularities. Henceforth, we have to introduce the TGARCH model described by Rabemananjara and Zakoïan (1993)

$$\varepsilon_t = \sigma_t \eta_t, \tag{4}$$

$$\sigma_t = \alpha_0 + \sum_{l=1}^{q} \alpha_l (|\varepsilon_{t-l}| - \gamma_l \varepsilon_{t-l}) + \sum_{i=1}^{p} \beta_i \sigma_{t-i}, \tag{5}$$

$$= \alpha_0 + \sum_{l=1}^{q} \alpha_l (1-\gamma_l)\varepsilon_t^+ - \sum_{l=1}^{q} \alpha_l (1+\gamma_l)\varepsilon_t^- + \sum_{i=1}^{p} \beta_i \sigma_{t-i}, \tag{6}$$

where $\{\eta_t\}_{t\in\mathbb{Z}} \sim F$ is i.i.d. with $E[\eta_t] = 0$ and $Var[\eta_t] = 1$. Specifically, the error term $\{\eta_t\}$ is assumed to be either standard normally distributed, normalized t-distributed or skewed t-distributed. The TGARCH parameter γ_l are the asymmetry parameters with $|\gamma_l| \leq 1$ for $l = 1,...,q$. This parameter depicts the asymmetry within the conditional variance. Furthermore, $\alpha_0 > 0$, $\alpha_l \geq 0 \ \forall \ l = 1,...,q$ and $\beta_i \geq 0 \ \forall \ i = 1,...,p$ are the classical GARCH parameters. Rabemananjara and Zakoïan (1993) discuss the existence of a stationary solution of the TGARCH process. One definition of the skewed t-distribution is given by Fernandez and Steel (1998). This approach combines two halves of a symmetric base distribution, which are differently scaled. The subsequential equation provides the density function

$$f_x(x) = \frac{2\xi}{(\xi^2+1)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}\sqrt{\nu}\sigma} \left[1 + \frac{\left(\frac{x-\mu}{\sigma}\right)^2}{\nu} \left(\frac{1}{\xi^2} I(x \geq \mu) + \xi^2 I(x < \mu)\right)\right]^{-\frac{\nu+1}{2}} \tag{7}$$

with ν degrees of freedom, $I(\cdot)$ the indicator function, expectation μ and variance $\sigma^2 = \nu/(\nu-2)$. Moreover, ξ is the skewness parameter with $\xi > 0$. It reduces f to the noncentral t-distribution if $\xi = 1$.

4 Comparison of the Forecasting Performance

We want to produce wind power forecasts that are significantly better than those computed by previous models. In contrast to WPPT, which is estimated by a least-squares approach, we use the maximum likelihood method for our periodic WPPT-CH model. We apply our method to three different distributional assumptions. Figure 2 provides the histogram of the WPPT residual process, $\{\widehat{\varepsilon}_t\}$. This histogram and the Shapiro-Wilk-Test reject the assumption of normally distributed residuals. Implying normality, the ACF of $\{\widehat{\varepsilon}_t\}$ and the ACF of $\{\widehat{|\varepsilon_t|}\}$ suggest the presence of

strong autocorrelation.
We use Akaike/Bayesian information criteria to select the best model for the residuals $\{\widehat{\varepsilon}_t\}$. Hence, for the autoregressive order, we choose $m = 5$ and $o = 4$ for the moving average part. Moreover, for the TGARCH model, we choose $q = 2$ and $p = 4$. For a comparison of the underlying distribution, we choose the same model order for each model. Figure 4 provides ACF of $\widehat{\varepsilon}_t$ and $\widehat{|\varepsilon_t|}$. This figure depicts a clear model improvement over normality. A huge amount of autocorrelation vanishes. The Ljung-Box-Test supports the assumption of no remaining autocorrelation for each distribution. Furthermore, we use the Kolmogorow-Smirnow-test to test for the underlying distribution. Only the assumption of the skew t-distribution is not rejected.

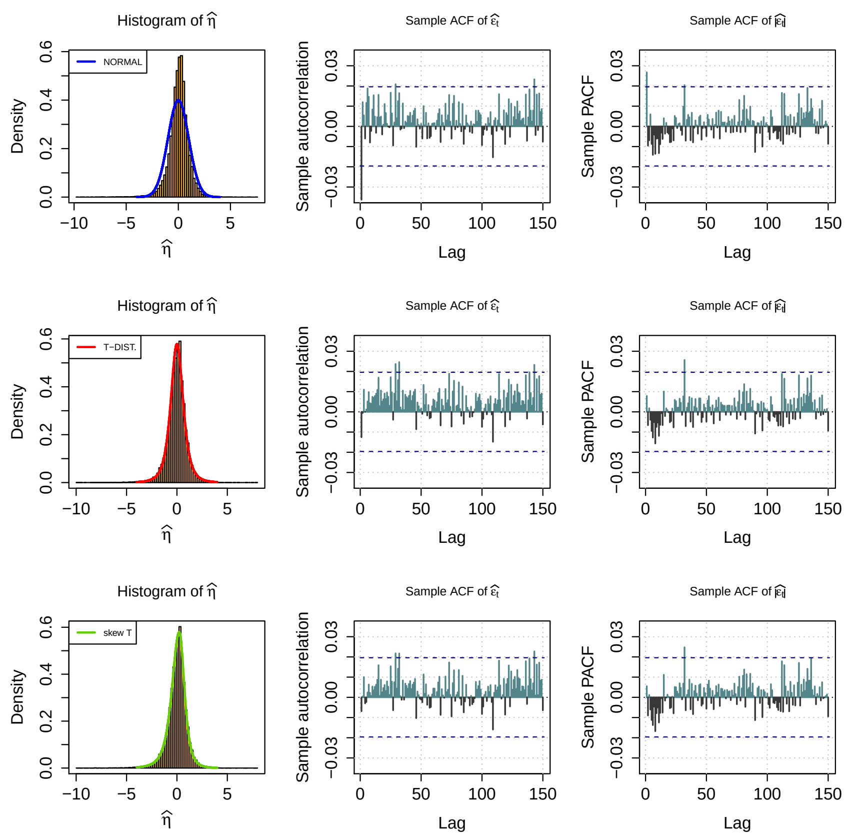

Fig. 4 Histogram of $\widehat{\varepsilon}_t$ (first column), ACF of $\widehat{\varepsilon}_t$ (second column) and ACF of $\widehat{|\varepsilon_t|}$ (third column), time frame October 31, 2010 to August 19, 2011 for the p-WPPT-CH and all distributions.

Regarding the in-sample fit, we conclude that the periodic WPPT-CH with skew t-distributed residuals provides the best fit. Therefore, we expect that this model outperforms the other models concerning the out-of-sample forecasts. We use the first 40000 observations as a training data set. The following week from August 19, 2011 to August 26, 2011 is used for out-of-sample forecasts. Here, we consider forecasts up to a maximum of half a day. Using a rolling window technique, we re-estimate each model for each forecast with a part of the information set available at the period t, namely $P_{t-10000+1},...,P_t$. Subsequently, we derive $\widehat{P}_{t+\tau|t}$, where $\tau \in \{1,...,72\}$. This procedure is repeated 1.000 times. We evaluate the out-of-sample forecasts for the classical WPPT and the periodic WPPT-CH model with normally, t-distributed and skew t-distributed errors. Besides, we take the persistence predictor $\widehat{P}_{t+\tau} = P_t$ and an AR(p) model as benchmark. Table 1 presents aggregated root mean square errors (RMSE)[2] for all models and forecasting horizons of $\tau \in \{1,18,36,72\}$. The best (smallest) values are bolded.

Table 1 RMSE for all models, time frame August 19, 2011 to August 26, 2011, 10 minutes, 3 hours, 6 hours and 12 hours ahead.

	1 step	18 steps	36 steps	72 steps
Persistence	110.77	258.04	281.11	282.45
AR	112.08	207.38	213.11	**204.18**
WPPT	**101.97**	208.48	211.41	205.93
pWPPT-CH-n	102.53	**202.89**	**206.85**	204.60
pWPPT-CH-t	102.85	204.10	208.05	204.58
pWPPT-CH-st	102.04	203.60	207.61	205.06

In almost all cases, the periodic WPPT-CH model provides the lowest forecasting errors. Figure 5 visualizes the RMSE inflation paths of the models by forecasting horizons. The Figure emphasizes the findings presented in Table 1.[3] Thus, we conclude that the best in-sample model also provides the best predictions.

5 Conclusion

We provide a new class of models that replace the Fourier series utilized by WPPT by more appropriate and more flexible periodic B-splines. Beyond, we capture conditional heteroscedasticity by modeling the error term as a TGARCH process. Finally, different error distributions are used. The skew t-distribution seems to capture the residuals' empirical properties quite better than the normal distribution. Our

[2] We also calculate MAE. Results are quite similar and omitted here to conserve space. Tables and Figures are available upon request.

[3] For lucidity, the Figure depicts only the worst (persistence), best (pWPPT-CH-st) and the WPPT benchmark model. All other curves lie inside the spanned range.

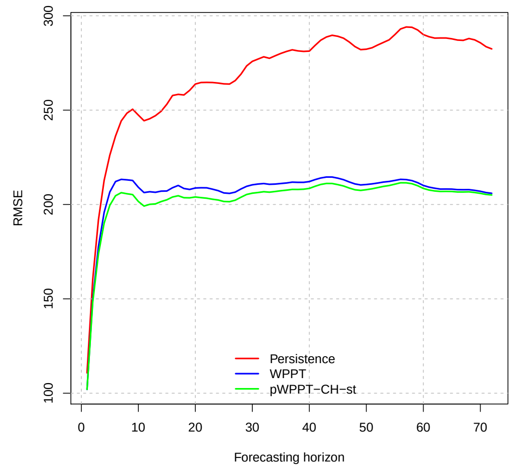

Fig. 5 RMSE for all models and all forecasting horizons, time frame August 19, 2011 to August 26, 2011.

models improve the in-sample features quite well, when compared to the classical WPPT model. Clearly, there are further improvements possible, for example, using a heteroscedastic multivariate periodic model. Nevertheless, the new wind power prediction model derived here does improve forecasts over WPPT already.

References

1. Ambach D and Schmid W (2014) Periodic and Long Range Dependent Models for High Frequency Wind Speed Data, Forthcoming.
2. Fernandez C and Steel M (1998) On Bayesian modelling of fat tails and skewness, Journal of the American Statistical Association, 93:359-371.
3. Giebel G, Brownsword R, Kariniotakis G, Denhard M and Draxl C (2011) The State-Of-The-Art in Short-Term Prediction of Wind Power, Tech. rep., ANEMOS.plus, Riso DTU, Wind Energy Division.
4. Harvey A, and Koopman S J (1993) Forecasting hourly electricity demand using timevarying splines, Journal of the American Statistical Association, 88(424):1228-1236.
5. Lei M, Shiyan L, Chuanwen J, Hongling L and Zhang Y (2009) A Review on the Forecasting of Wind Speed and Generated Power, Renewable and Sustainable Energy Reviews, 13:915-920.
6. Nielsen H, Pinson P, Christiansen L, Nielsen T, Madsen H, Badger J, Giebel G and Ravn H (2007) Improvement and Automation of Tools for Short Term Wind Power Forecasting, Tech. rep., Scientific Proceedings of the European Wind Energy Conference and Exhibition, Milan, Italy.
7. Rabemananjara R and Zakoïan J M (1993) Threshold arch models and asymmetries in volatility, Journal of Applied Econometrics, 8(1):31-49.
8. Ziel F and Steinert R (2014) Efficient Modeling and Forecasting of the Electricity Spot Price, arXiv preprint arXiv:1402.7027.