



# A Comparison of Different Wind Power Forecasting Models to the Mycielski Approach.

Carsten Croonenboreck

Daniel Ambach

---

European University Viadrina Frankfurt (Oder)  
Department of Business Administration and Economics

Discussion Paper No. 355

May 2014

ISSN 1860 0921

---

# A Comparison of Different Wind Power Forecasting Models to the Mycielski Approach

May 13, 2014

Carsten Croonenbroeck<sup>a</sup> and Daniel Ambach<sup>b</sup>

## Abstract

In the wind power industry, wind speed forecasts are obtained and transformed into wind power forecasts. The Mycielski algorithm has proven to be an accurate predictor for wind speed in short-term scenarios. Moreover, Mycielski has the capability of forecasting wind power directly, instead of wind speed.

This article compares wind power forecasts calculated by the Mycielski algorithm to state-of-the-art forecasters. As such, we use the Wind Power Prediction Tool (WPPT) and the recently developed generalization of it, GWPPT (*Generalized* WPPT). Furthermore, we evaluate statistical time series models such as autoregressive and vector autoregressive models. As an additional benchmark we use the persistence model, which is often used to assess forecasting accuracy. Each model is evaluated and we give a recommendation for the best forecasting model.

**JEL classification:** C35, E27, Q47

**Keywords:** Mycielski algorithm, WPPT, GWPPT, Wind Power, Wind Energy, Forecasting, Prediction

## Addresses:

<sup>a</sup> Corresponding Author: Carsten Croonenbroeck, European University Viadrina, Chair of Economics and Economic Theory (Macroeconomics), Post Box 1786, 15207 Frankfurt (Oder), Germany, Tel. +49 (0)335 5534 2701, Fax +49 (0)335 5534 72701, croonenbroeck@europa-uni.de.

<sup>b</sup> Daniel Ambach, European University Viadrina, Chair of Quantitative Methods and Statistics, Post Box 1786, 15207 Frankfurt (Oder), Germany, Tel. +49 (0)335 5534 2983, Fax +49 (0)335 5534 2233, ambach@europa-uni.de.

# A Comparison of Different Wind Power Forecasting Models to the Mycielski Approach

May 13, 2014

## Abstract

In the wind power industry, wind speed forecasts are obtained and transformed into wind power forecasts. The Mycielski algorithm has proven to be an accurate predictor for wind speed in short-term scenarios. Moreover, Mycielski has the capability of forecasting wind power directly, instead of wind speed.

This article compares wind power forecasts calculated by the Mycielski algorithm to state-of-the-art forecasters. As such, we use the Wind Power Prediction Tool (WPPT) and the recently developed generalization of it, GWPPT (*Generalized* WPPT). Furthermore, we evaluate statistical time series models such as autoregressive and vector autoregressive models. As an additional benchmark we use the persistence model, which is often used to assess forecasting accuracy. Each model is evaluated and we give a recommendation for the best forecasting model.

**JEL classification:** C35, E27, Q47

**Keywords:** Mycielski algorithm, WPPT, GWPPT, Wind Power, Wind Energy, Forecasting, Prediction

# 1 Introduction

1 For the last two decades, energy production has been characterized by renewables.  
2 Conventional energy is deterministic, and as such, easy to predict. Wind energy,  
3 however, possesses stochastic features, so prediction is a tough task. Research  
4 may focus on predicting wind speed and transforming these predictions into wind  
5 power forecasts. Such a mapping from wind speed to wind power is often done  
6 by using the manufacturer’s power curve. However, Wu and Hong (2007) suggest  
7 to use a power curve created by measured wind speed. Alternatively to simple  
8 mapping, wind power can be predicted directly by, e.g., a time series or regression  
9 approach. We compare several wind power forecasting models and evaluate their  
10 accuracy.

11 In the literature, forecasting models are classified into short-term, medium-term  
12 and long-term methods. Although this arrangement is mildly diffuse, short-term  
13 forecasts are usually ones that aim at a forecasting horizon of up to six or twelve  
14 hours. For instance, Lei et al. (2009) make that determination.

15 Wind power data hold a high degree of persistence. Thus, in wind power forecast-  
16 ing, a hard-to-beat benchmark for all newly developed prediction methods is the  
17 persistence model (i.e.  $\hat{x}_{t+k} = x_t$ ). For example, Giebel et al. (2011) emphasize  
18 its importance and point out its surprisingly well prediction performance, at least  
19 for very short forecasting horizons. Another established class of benchmark pre-  
20 diction models are time series approaches. The univariate type of these models is  
21 the autoregressive process. Moreover, we present a vector autoregressive process  
22 that is able to capture linear interdependencies.

23 Wind speed, as the most important explanatory variable for wind power predic-  
24 tion, shows a strong diurnal pattern which can be modeled using a Fourier series.

25 A method based on persistence and daily wind speed patterns is the Wind Power  
26 Prediction Tool (WPPT), as discussed by Nielsen et al. (2007). Croonenbroeck  
27 and Dahl (2014) generalize the idea by adding wind direction as an explanatory  
28 variable and also make use of the a-priori known power range of the turbine, i.e.  
29 the fact that there are predetermined boundaries of wind power forecasts. The  
30 result, named GWPPT, improves forecasting precision by a severe degree.

31 Wind speed prediction is mostly done by univariate time series models. How-  
32 ever, there are alternative ideas such as artificial neural networks or fuzzy logic  
33 models. An impressively simple idea is the one of the Mycielski algorithm. The  
34 algorithm developed by Ehrenfeucht and Mycielski (1992) is originally proposed  
35 as a pseudorandom number generator. An application as a universal prediction  
36 method is refined by Jacquet et al. (2002). Finally, Hocaoglu et al. (2009) use  
37 the method to predict wind speed as measured at three different sites in Turkey.  
38 More recently, Gan et al. (2012) apply the method to three sites whose locations  
39 are not revealed.

40 The Mycielski algorithm has its similarities to Markov chain models. Markov  
41 chain models build up transition probabilities from one condition to another.  
42 Thus, they learn from the history of the whole data set. As a simplification of  
43 that, the Mycielski algorithm directly picks values out of the history as predic-  
44 tors. By using the successor of the longest matching string in the history of data  
45 that is identical to the string in the most recent history, it uses the historic value  
46 that represents the most probable predictor. Although this approach does not  
47 stem from a parametric statistical model, Hocaoglu et al. (2009) as well as Gan  
48 et al. (2012) find good forecasting performance for the algorithm when used as a  
49 wind speed prediction method.

50 Instead of using the Mycielski algorithm as a wind speed forecaster, we evaluate

51 its qualification as a wind power forecaster. The results are compared to WPPT,  
52 GWPPT, persistence and the time series models. Since Mycielski is positioned as  
53 a short-term forecaster, results for horizons of 1 step ahead and 12 hours ahead  
54 are presented.

55 The article is organized in the following way. A description of the analyzed data  
56 is given in Section 2. Thereafter, Section 3 shortly introduces the tested Myciel-  
57 ski algorithm, WPPT, GWPPT and the time series models. In Section 4, results  
58 are discussed. Finally, Section 5 concludes the obtained findings.

## 59 **2 Wind Power Data**

60 The wind power data set used in this study is a high-frequency series collected  
61 in Eastern Germany. The in-sample fit and the out-of-sample forecasts are cal-  
62 culated for four different Fuhrländer FL MD 77 turbines of one wind park. The  
63 observed wind park is situated in a rural plain region. The area has a slight  
64 roughness with fields and some forests. Due to a non-disclosure agreement, the  
65 specific locations cannot be revealed.

66 The turbines, labeled Turbine A to D, exhibit a power range of each  $[0; 1500]$  kW  
67 and write sensor data to log files at a frequency of ten minutes. The observed time  
68 frame spans from Nov. 1, 2010 to Nov. 5, 2012, so there are 105984 observation  
69 per turbine.

## 70 **3 Forecasting Methods**

71 This Section provides an overview of the wind power prediction models investi-  
72 gated in this article. Assume that a forecast for time  $t + k$  is desired, based on

73 historical data for time  $1, 2, \dots, t-1, t$ . Typically, the Mycielski algorithm picks  
74 a short pattern of the most recent history and searches the entire history for that  
75 pattern. Say, the pattern is denoted as  $\mathbf{p} \in \mathbb{R}^d$  and consists of  $d$  elements. In the  
76 simplest case,  $d = 1$ , so  $\mathbf{p} = x_t$ . Every successor of historic observations that are  
77 identical to  $x_t$  are then possible predictors for time  $t + 1$ . In the more general  
78 case of forecasting horizons  $k > 1$ , the  $k$ th successor of that observation would  
79 be the predictor.

80 Whenever multiple candidates are found,  $d$  is increased by 1, so in the second  
81 step,  $d = 2$  and  $\mathbf{p} = (x_{t-1}, x_t)^T$ . That pattern is searched again. The algorithm  
82 continues until the longest matching chain is found. It is assumed that the re-  
83 spective successor of the longest matching chain is the most probable predictor.  
84 Formally, Hocaoglu et al. (2009) note the algorithm as follows. The forecast is  
85 calculated as

$$\hat{x}_{t+k} = f_{t+k}(x_1, \dots, x_t), \quad (1)$$

86 where  $f(\cdot)$  performs the iterative search algorithm: Let  $\mu$  be the index of the  $k$ th  
87 successor of the longest matching string, which itself ends at index  $o$ . Then  $\mu$  is  
88 found by

$$\mu = \arg \max_L (x_t = x_o, x_{t-1} = x_{o-1}, \dots, x_{t-L+1} = x_{o-L+1}), \quad (2)$$

89 where  $L$  denotes chain length. Thus, it follows that  $f_{t+k}(x_1, \dots, x_t) = \hat{x}_{t+k} = x_\mu$ .  
90 As data are usually decimal values, it is unlikely to find long strings that exactly  
91 match the search pattern. To overcome this, Hocaoglu et al. (2009) transform

92 the data. They use rounded integer values during the search process. An alterna-  
93 tive to this is not to search for exact matches in the history, but only for similar  
94 ones. This fuzzy search method can be applied by searching for chains that may  
95 not match exactly, but only lie within a certain tolerance interval. In that case,  
96 tolerance becomes a parameter of the model (denoted as  $\tau$ , interval width is  $2\tau$ ,  
97 then) which can be optimized from the data. One way of optimization would  
98 be to calculate a set of pseudo-out-of-sample predictions using several values of  
99 tolerance and then pick the tolerance that provides the best forecast.

100 WPPT is based on a dynamic regression approach. It takes wind speed as a  
101 regressor and captures diurnal periodicity by a Fourier series of time of day ob-  
102 servations to estimate the parameters of the model

$$\begin{aligned} \hat{x}_{t+k} = & m + a_1 \cdot x_t + a_2 \cdot x_{t-1} + b_1 \cdot w_{t+k|t} + b_2 \cdot w_{t+k|t}^2 + d_1^c \cdot \cos\left(\frac{2\pi d_{t+k}}{144}\right) \\ & + d_2^c \cdot \cos\left(\frac{4\pi d_{t+k}}{144}\right) + d_1^s \cdot \sin\left(\frac{2\pi d_{t+k}}{144}\right) + d_2^s \cdot \sin\left(\frac{4\pi d_{t+k}}{144}\right) + \varepsilon_{t+k}, \quad (3) \end{aligned}$$

103 where  $w_{t+k|t}$  is wind speed at time  $t+k$  given at time  $t$ ,  $d_t$  is time of day for ob-  
104 servation  $t$  and  $\varepsilon_{t+k}$  is assumed white noise. The important wind speed forecasts  
105 may stem from numerical weather predictions (NWP) or from statistical models  
106 themselves.

107 The generalization of this, GWPPT, is based on the same idea, but uses wind di-  
108 rection as an additional explanatory variable. Furthermore, the data is modeled  
109 as both-sided censored: Each wind turbine is produced to operate at a certain  
110 range, the so-called power range. The range is provided by the manufacturer.

111 For example, Fuhrländer FL MD 77 turbines are designed to operate at a max-  
 112 imum load of 1500 kW, so it can be assumed that any forecast that lies outside  
 113 of an interval of  $[0; 1500]$  is wrong. GWPPT makes use of this a-priori known  
 114 information. The model imposes the following structure on wind power:

$$x_t^* = \eta(\mathbf{z}_t) + \varepsilon_t, \quad (4)$$

115 where  $\mathbf{z}_t$  is the vector of explanatory variables,  $\eta$  is a linear function of  $\mathbf{z}_t$ , and  $\varepsilon_t$   
 116 is an assumed Gaussian error term. GWPPT imposes a censored data structure,  
 117 so that

$$x_t = \begin{cases} l, & x_t^* \leq l \\ x_t^*, & x_t^* \in (l, u) \\ u, & x_t^* \geq u, \end{cases} \quad (5)$$

118 where  $l$  and  $u$  are the lower and upper censoring points. Parameters are estimated  
 119 using a generalized probit model. In the end, due to assumed Gaussian errors,  
 120 the forecast is calculated by

$$\hat{x}_{t+k} = (\Phi(f_2) - \Phi(f_1)) \cdot x_{t+k}^* + (\phi(f_1) - \phi(f_2)) \cdot \hat{\sigma} + u \cdot (1 - \Phi(f_2)), \quad (6)$$

121 where

$$f_1 = \frac{l - x_{t+k}^*}{\hat{\sigma}}, \quad (7)$$

$$f_2 = \frac{u - x_{t+k}^*}{\hat{\sigma}}, \quad (8)$$

122 and  $\phi(\cdot)$  and  $\Phi(\cdot)$  denote normal PDF (Probability Density Function) and CDF  
 123 (Cumulative Distribution Function), respectively.

124 Of course it is desirable that a model is able to beat the naïve prediction model  
 125  $\hat{x}_{t+k} = x_t$ . Alternatively, we use a second type of crude benchmark models, au-  
 126 toregressive and vector autoregressive approaches. These models are often used  
 127 in the context of wind speed and wind power prediction. Taylor et al. (2009)  
 128 expand these approaches and introduce periodic regressors and a fractional in-  
 129 tegrated component. Here, we consider AIC-based AR(p) and VAR(p) models  
 130 as benchmarks. These processes are easily estimated by solving the Yule-Walker  
 131 equations. This guarantees a stationary solution and saves computing time. The  
 132 Akaike Information Criterion is used to select the appropriate model order. Con-  
 133 cerning the VAR(p) process, we use wind power, wind speed, wind direction and  
 134 exterior temperature as response variables  $\mathbf{X}_t$ . The autoregressive and vector  
 135 autoregressive models of order  $p$  are given by

$$x_t = \psi(B)x_{t-1} + w_t, \quad \text{where } \{w_t\} \sim WN(0, \sigma^2) \quad (9)$$

136 and

$$\mathbf{x}_t = \Psi(B)\mathbf{x}_{t-1} + \mathbf{w}_t, \quad \text{where } \{\mathbf{w}_t\} \sim WN(0, \sigma^2), \quad (10)$$

137 where  $\mathbf{x}_t$  is an  $m \times 1$  vector. Moreover,  $w_t$  is a white noise error and  $\mathbf{w}_t$  is an  $m \times 1$   
 138 vector white noise error term. The backward shift operator  $B$  is  $B^l X_t = X_{t-l}$ .  
 139  $\psi(B)$  and  $\Psi(B)$  are polynomials defined by  $\psi(B) = 1 - \psi_1 B - \dots - \psi_p B^p$  and  
 140  $\Psi(B) = 1 - \Psi_1 B - \dots - \Psi_p B^p$ . Note that  $\Psi$  is an  $m \times m$  transition matrix which  
 141 maps the dependency of  $\mathbf{x}_t$  on  $\mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-p}$ .

142 Eventually, we investigate three sophisticated models (Mycielski, WPPT and  
143 GWPPT) and three common benchmarks (persistence, AR and VAR).

## 144 4 Results

145 For this article, out-of-sample forecasts are calculated. Initially, roughly 60% of  
146 the entire time frame (approximately 60000 observations) are used as training  
147 data. The considered high frequency observations for the in-sample model cali-  
148 bration start from November 1, 2010 and reach to December 7, 2011. For WPPT,  
149 GWPPT, and the autoregressive models, we calculate the forecasts using a rolling  
150 window of fixed size. To conserve processing time, 4000 forecasts are calculated,  
151 both for one step ahead forecasting horizon (10 minutes) as well as 72 steps ahead  
152 (= 12 hours).

153 To find the optimal tolerance level  $\tau$ , we perform out-of-sample predictions for  
154  $\tau = \{0, 1, 2, 3, 4, 5\}$  as described above. Figure 1 depicts RMSE and MAE at  
155 given tolerance levels for 1 and 72 steps ahead and all turbines. Most curves  
156 suggest a slight decrease in aggregated forecasting errors for  $\tau = 2$ , with a weak  
157 tendency of decreasing values for increasing tolerance. Table 1 shows average per-  
158 forecast chain lengths dependent on  $\tau$ . Lengths increase for increasing tolerance  
159 levels, but the percentage inflation of lengths declines after  $\tau = 2$ . These findings  
160 suggest to use a tolerance level of  $\tau = 2$ . This parametrization provides a trade-  
161 off of well forecasting performance and parsimonious specification. Therefore, all  
162 further results are calculated using  $\tau = 2$ .

163 Figure 2 presents out-of-sample forecasts' behavior in a typical power curve di-  
164 agram. WPPT returns forecasts  $\hat{x}_{t+1} < 0$  for wind speeds below 2  $m/s$ . In the  
165 area of 5  $m/s$ , GWPPT overestimates. Between 8 and 10  $m/s$ , GWPPT and

166 WPPT underestimate. Above 12  $m/s$ , Mycielski and persistence underestimate,  
167 while WPPT strongly overestimates above 13  $m/s$  (several forecasts are even  
168  $\hat{x}_{t+1} > 1500$ ). As expected, GWPPT converges asymptotically against  $l = 1500$ .  
169 AR, VAR and persistence perform similarly. Only in the sparse upper right area,  
170 the spreads between the curves become slightly wider. Results for the other tur-  
171 bines look similar and are presented in the four panels of Figure 3.

172 Figure 4 shows actual values and the forecasts in a time series comparison plot.  
173 Note that the ordinate is differently scaled for WPPT, since as described above,  
174 there is no predetermined limit of domain for this method. Mycielski forecasts  
175 are more volatile than the other ones, possibly due to the fact that there is no  
176 means of capturing persistence of data for this forecaster.

177 Figure 5 presents results for forecasts 72 steps (= 12 hours) ahead. While WPPT  
178 and GWPPT still follow the actual power curve comparably well, Mycielski, AR,  
179 VAR and persistence perform badly. The four panels in Figure 6 show similar  
180 results for all of the four turbines. Figure 7 depicts the time series for 12 hours  
181 forecasts for turbine A. WPPT and GWPPT perform well (again, for WPPT,  
182 the ordinate has different scaling), but Mycielski is even noisier than for the 1  
183 step ahead forecast. Persistence shows good structure at first glance, but as ob-  
184 servations are lagged, forecasts are not well. AR and VAR look good, but only  
185 slightly more plain.

186 In Tables 2 and 3, descriptive statistics of forecasting errors 1 step ahead and 72  
187 steps ahead are shown for all turbines and forecasting methods. Except for per-  
188 sistence, errors are mostly stationary (according to ADF tests). In 1 step ahead  
189 forecasts, errors usually have zero mean ( $E[\hat{\varepsilon}_{t+k}] \neq 0$  is rejected), while in 12  
190 hours ahead forecasts, errors are significantly biased. Few exceptions aside, errors'  
191 distribution is mostly symmetric. For WPPT and GWPPT, errors are mostly

192 leptokurtic, while for Mycielski and persistence, they are usually platykurtic, es-  
193 pecially for the longer forecasting horizon. AR and VAR errors are leptokurtic  
194 in the 1 step scenario, but platykurtic in the longer horizon scheme. One step  
195 ahead, AR and VAR provide the lowest standard deviations and the most nar-  
196 row interval of min/max values. For 12 hours forecasts, GWPPT returns lowest  
197 standard deviation and most narrow min/max ranges.

198 Table 4 presents RMSE and MAE for all stations, all forecasters and both fore-  
199 casting horizons, respectively. The respective best (i.e. lowest) values are bolded.  
200 In all cases, Mycielski is clearly dominated. AR and VAR prove to be strong com-  
201 petitors even to persistence, especially for short-term forecasts. Overall, VAR is  
202 strongest in 1 step forecasts, occasionally outperformed by the simple AR model.  
203 For 72 steps ahead, GWPPT is the overall winner.

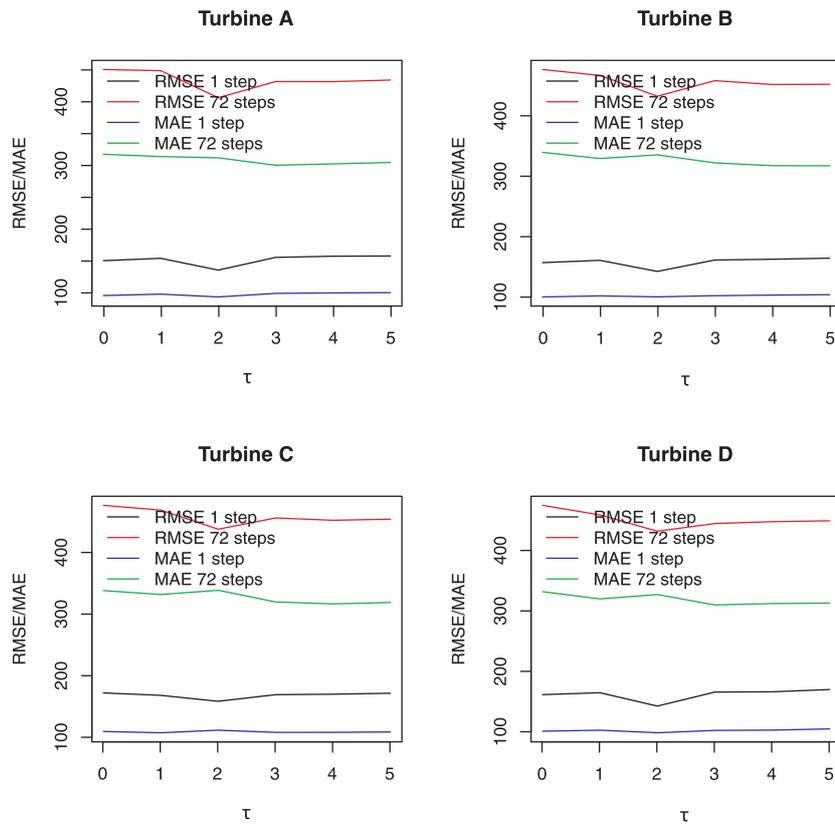


Figure 1. Behavior of RMSE and MAE dependent on  $\tau$ , Turbines A to D, 1 step and 72 steps ahead, time frame December 08, 2011 to January 04, 2012.

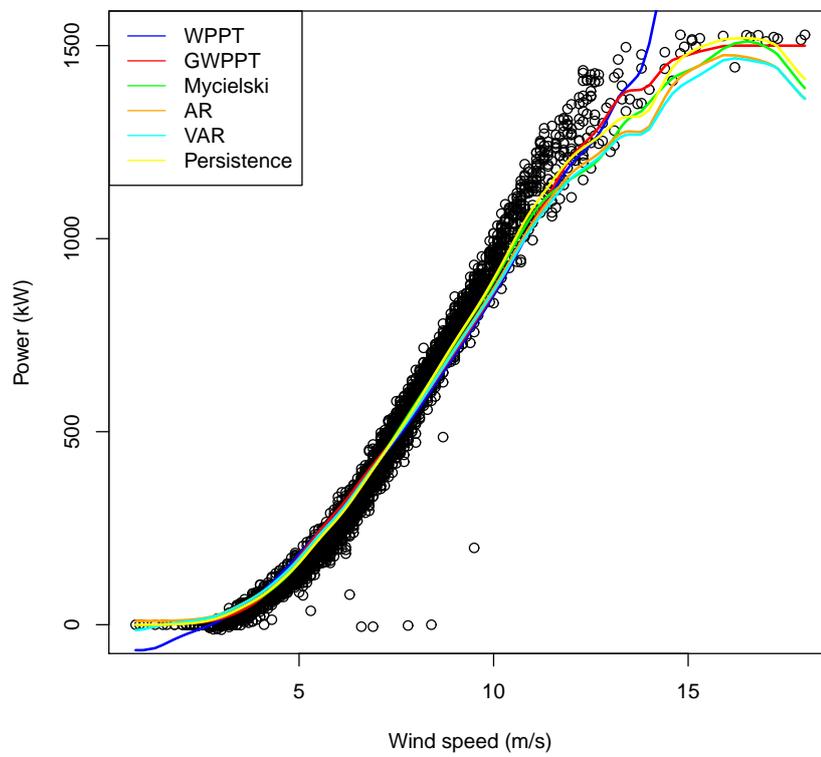


Figure 2. Comparison of the six forecasting methods in a typical power curve diagram. Turbine A, 1 step ahead, time frame December 08, 2011 to January 04, 2012.

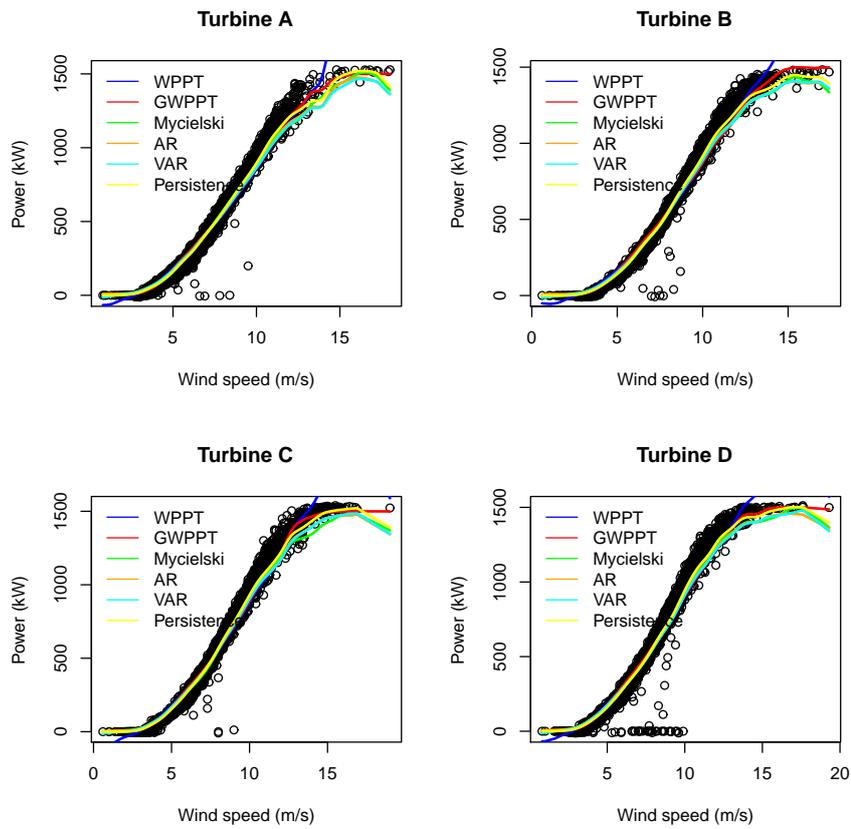


Figure 3. Comparison of the six forecasting methods in a typical power curve diagram. Turbines A to D, 1 step ahead, time frame December 08, 2011 to January 04, 2012.

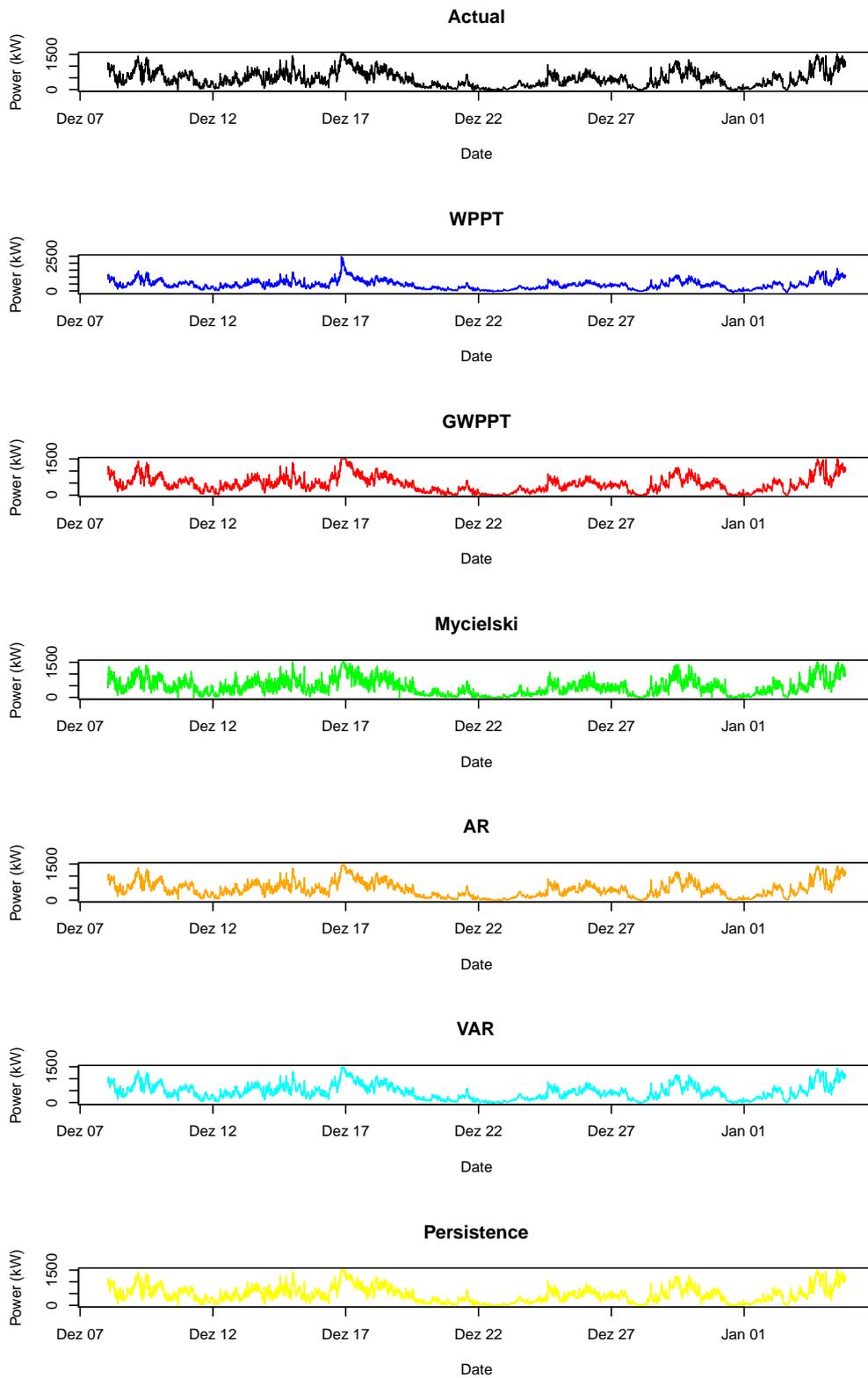


Figure 4. Time series diagram of actual values and the six forecasters. Turbine A, 1 step ahead, time frame December 08, 2011 to January 04, 2012.

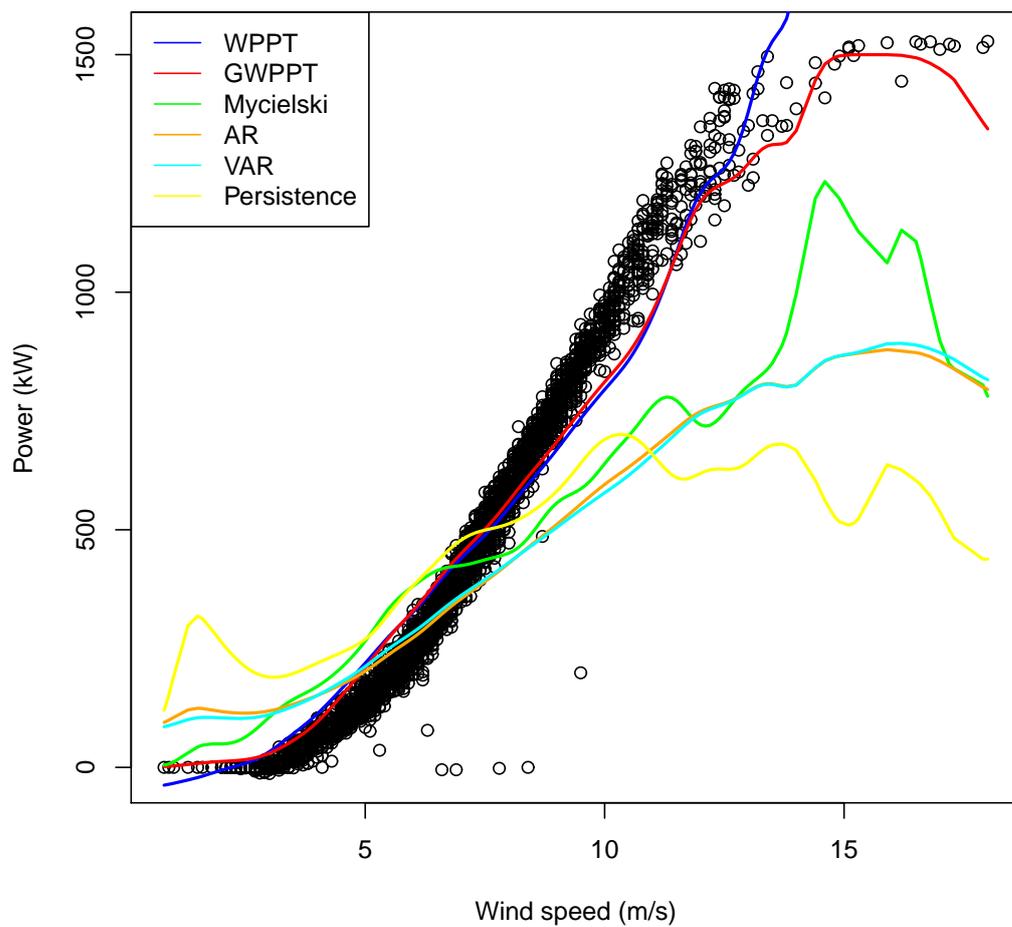


Figure 5. Comparison of the six forecasting methods in a typical power curve diagram. Turbine A, 72 steps ahead, time frame December 09, 2011 to January 04, 2012.

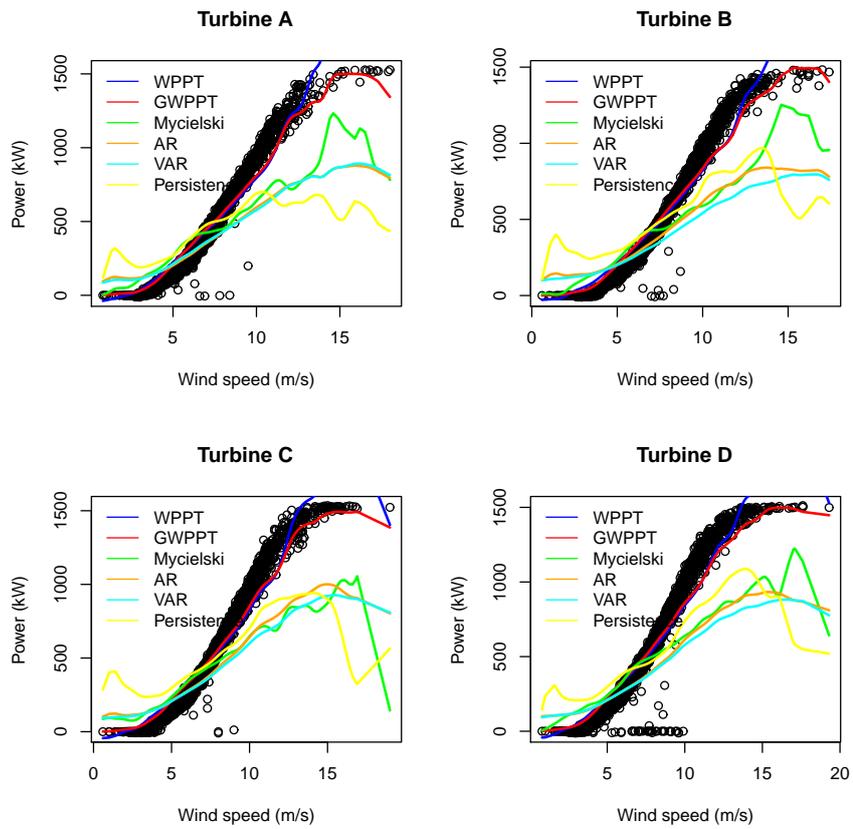


Figure 6. Comparison of the six forecasting methods in a typical power curve diagram. Turbines A to D, 72 steps ahead, time frame December 09, 2011 to January 04, 2012.

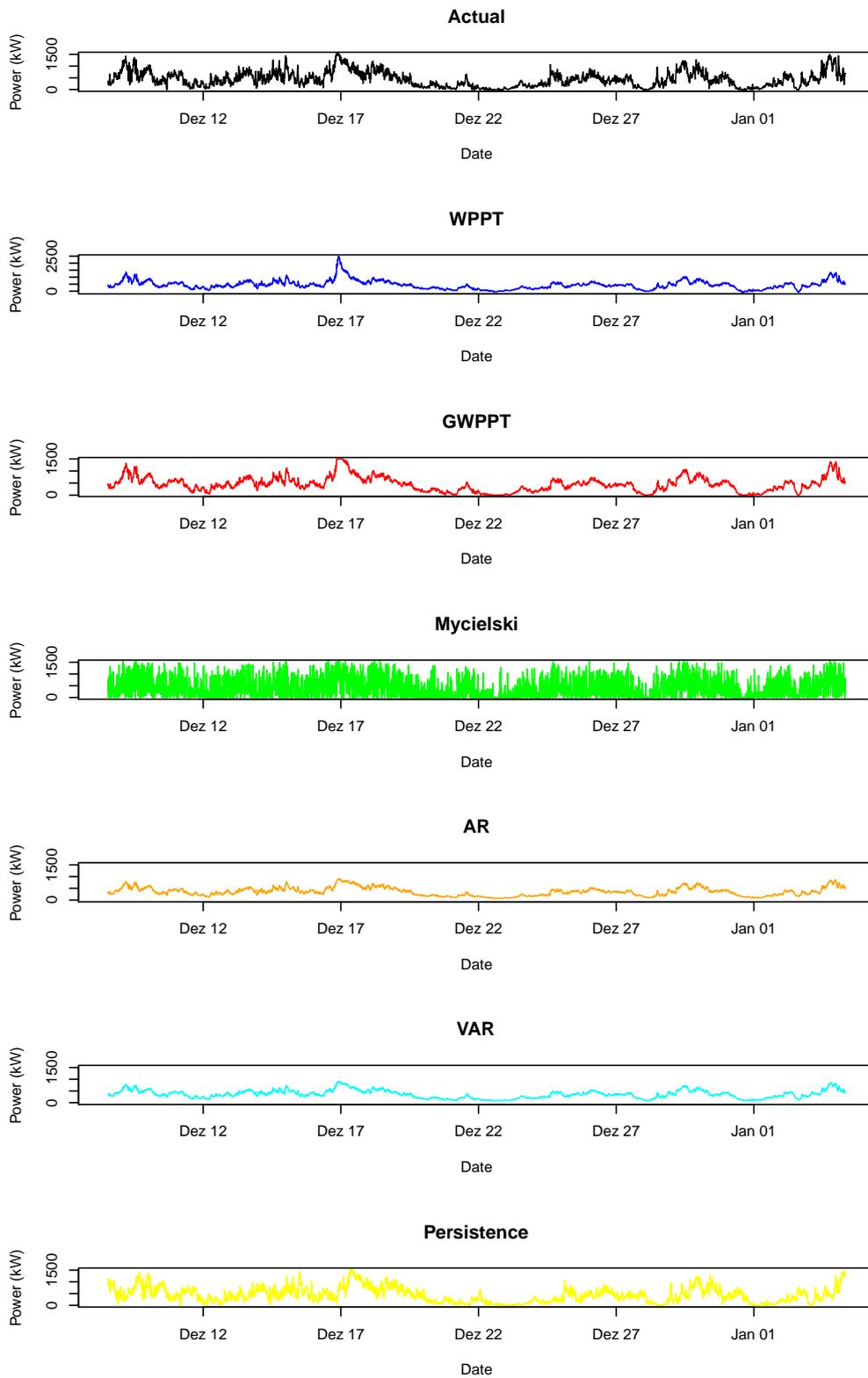


Figure 7. Time series diagram of actual values and the six forecasters. Turbine A, 72 steps ahead, time frame December 09, 2011 to January 04, 2012.

	$\tau = 0$	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$
<b>Turbine A</b>						
1 step	1.00	1.75 (+ 75.31%)	2.38 (+ 35.54%)	2.57 (+ 8.24%)	2.77 (+ 7.61%)	2.91 (+ 5.12%)
72 steps	1.00	1.75 (+ 75.34%)	2.38 (+ 35.54%)	2.60 (+ 9.24%)	2.77 (+ 6.62%)	2.91 (+ 5.15%)
<b>Turbine B</b>						
1 step	1.00	1.73 (+ 73.45%)	2.34 (+ 34.68%)	2.53 (+ 8.41%)	2.73 (+ 7.76%)	2.88 (+ 5.35%)
72 steps	1.00	1.73 (+ 73.25%)	2.33 (+ 34.39%)	2.55 (+ 9.33%)	2.72 (+ 6.90%)	2.88 (+ 5.40%)
<b>Turbine C</b>						
1 step	1.00	1.72 (+ 72.18%)	2.33 (+ 35.27%)	2.52 (+ 8.34%)	2.72 (+ 7.70%)	2.86 (+ 5.15%)
72 steps	1.00	1.72 (+ 72.09%)	2.32 (+ 35.02%)	2.53 (+ 9.03%)	2.71 (+ 7.09%)	2.85 (+ 5.17%)
<b>Turbine D</b>						
1 step	1.00	1.74 (+ 73.53%)	2.39 (+ 37.64%)	2.58 (+ 7.95%)	2.77 (+ 7.37%)	2.91 (+ 5.00%)
72 steps	1.00	1.73 (+ 73.42%)	2.38 (+ 37.31%)	2.59 (+ 8.78%)	2.76 (+ 6.64%)	2.90 (+ 5.03%)

Table 1. Average length of Mycielski chains dependent on  $\tau$  and percentage inflation of chain length in comparison to previous chain length ( $\tau - 1$ ).

	Min	Max	Mean	Median	Std. Dev.	Skewness	Kurtosis	t-value of mean = 0	ADF (p-value)
<b>Turbine A</b>									
WPPT	-976.60	601.77	3.54	0.54	109.24	-0.64	7.25	**2.05	0.37
GWPPPT	-722.09	593.02	1.80	-0.59	102.07	0.05	2.88	1.12	0.38
Mycielski	-770.00	797.00	1.00	1.00	135.78	-0.10	3.12	0.46	*0.07
AR	-686.70	599.41	3.91	-2.53	99.40	0.14	3.33	**2.48	0.37
VAR	-692.26	598.94	5.02	-0.79	99.37	0.17	3.22	***3.19	0.41
Persistence	-685.00	662.00	0.04	0.00	101.25	-0.06	3.57	0.03	**0.04
<b>Turbine B</b>									
WPPT	-966.89	752.40	0.20	-6.35	113.53	-0.33	6.10	0.11	0.16
GWPPPT	-551.69	740.10	-0.23	-3.69	104.30	0.39	2.99	-0.14	0.15
Mycielski	-579.00	737.00	9.65	3.00	142.16	0.18	2.17	***4.29	0.25
AR	-548.73	771.95	4.69	-1.70	102.21	0.40	3.40	***2.90	0.35
VAR	-539.20	768.83	6.13	-0.58	101.98	0.42	3.37	***3.80	0.48
Persistence	-570.00	764.00	0.11	0.00	105.02	0.24	3.47	0.07	**0.02
<b>Turbine C</b>									
WPPT	-1298.30	888.03	0.94	-5.30	120.08	-0.34	7.48	0.49	0.32
GWPPPT	-713.15	870.41	1.14	-2.24	111.09	0.24	3.98	0.65	0.31
Mycielski	-894.00	939.00	8.01	4.00	158.25	0.01	2.53	***3.19	0.98
AR	-789.07	877.74	4.44	-2.38	108.74	0.27	4.22	***2.58	0.48
VAR	-771.58	889.70	5.50	-2.49	108.34	0.31	4.30	***3.21	0.56
Persistence	-839.00	816.00	0.12	0.00	111.13	-0.01	4.33	0.07	**0.03
<b>Turbine D</b>									
WPPT	-1346.12	752.11	-1.99	-4.74	123.44	-0.84	9.26	-1.02	0.38
GWPPPT	-767.27	736.81	-3.22	-3.28	115.48	-0.33	5.43	*-1.76	0.52
Mycielski	-902.00	783.00	4.36	1.00	142.50	-0.13	3.22	*1.93	0.80
AR	-717.93	725.54	4.53	-3.15	107.53	0.27	4.64	***2.67	0.27
VAR	-716.04	717.52	5.47	-1.09	107.69	0.28	4.57	***3.21	0.34
Persistence	-753.00	755.00	-0.15	0.00	109.06	0.04	5.04	-0.09	**0.03

Table 2. Descriptive statistics for 1 step ahead forecasting errors. \*\*\*, \*\* and \* represent significance at 1%, 5% and 10% level.

	Min	Max	Mean	Median	Std. Dev.	Skewness	Kurtosis	t-value of mean = 0	ADF (p-value)
<b>Turbine A</b>									
WPPT	-976.28	798.66	-8.24	-14.27	153.08	-0.40	5.40	***-3.34	0.35
GWPPPT	-736.74	796.50	-9.21	-9.90	139.21	0.37	2.54	***-4.11	0.24
Mycielski	-1276.00	1500.00	30.02	46.00	405.29	-0.17	0.33	***4.59	0.48
AR	-419.23	837.66	83.15	50.26	176.14	0.97	0.89	***29.31	0.50
VAR	-420.45	838.80	81.34	42.57	177.06	1.01	0.84	***28.53	0.50
Persistence	-1318.00	1440.00	2.18	2.00	346.37	0.18	0.55	0.39	**0.02
<b>Turbine B</b>									
WPPT	-1069.81	981.31	-10.86	-20.53	156.83	-0.22	5.03	***-4.30	0.24
GWPPPT	-538.34	963.60	-12.32	-17.57	141.72	0.58	2.45	***-5.40	0.28
Mycielski	-1404.00	1412.00	60.79	83.00	428.23	-0.18	0.30	***8.81	0.62
AR	-339.05	895.27	96.09	61.18	189.39	0.81	0.34	***31.51	0.71
VAR	-286.17	861.92	131.19	89.73	206.17	0.82	0.08	***39.52	0.78
Persistence	-1231.00	1320.00	4.51	14.00	348.36	0.07	0.15	0.80	**0.06
<b>Turbine C</b>									
WPPT	-921.19	1150.19	-9.59	-21.74	168.03	0.05	3.44	***-3.55	0.36
GWPPPT	-690.85	1147.91	-8.29	-14.79	153.27	0.52	2.74	***-3.36	0.25
Mycielski	-1293.00	1422.00	64.28	78.00	433.05	-0.10	0.31	***9.21	0.48
AR	-406.75	1108.43	94.98	57.69	185.98	0.83	0.50	***31.72	0.76
VAR	-311.21	1102.75	112.26	67.28	196.46	0.92	0.40	***35.49	0.83
Persistence	-1307.00	1370.00	11.17	12.00	355.51	0.09	0.39	*1.95	*0.06
<b>Turbine D</b>									
WPPT	-1040.32	871.58	-14.29	-20.04	182.31	-0.55	4.69	***-4.87	0.47
GWPPPT	-983.98	858.66	-15.26	-16.59	167.41	-0.16	3.89	***-5.66	0.57
Mycielski	-1428.00	1478.00	40.95	60.00	429.89	-0.16	0.56	***5.90	0.91
AR	-578.88	929.02	105.37	59.12	199.99	0.89	0.39	***32.72	0.86
VAR	-560.43	923.05	116.47	63.41	214.07	0.94	0.35	***33.79	0.92
Persistence	-1268.00	1328.00	10.47	12.00	357.38	0.04	0.36	*1.82	*0.06

Table 3. Descriptive statistics for 72 steps ahead forecasting errors. \*\*\*, \*\*, \* and \* represent significance at 1%, 5% and 10% level.

## 204 5 Conclusion

205 This article presents a comparison of the forecasting methods Mycielski, WPPT,  
206 GWPPT, AR(p), VAR(p) and the naïve predictor. Out-of-sample forecasts for  
207 four sites in Germany, for both 1 step and 12 hours ahead, show that GWPPT  
208 provides the most accurate forecasts in most cases. Although known as quite well  
209 wind speed forecaster, the Mycielski algorithm does not provide convincing wind  
210 power predictions as measured in RMSE and MAE aggregated forecasting error  
211 measures.

212 Furthermore, the results in this article can be seen as a cross check of Hocaoglu  
213 et al. (2009), Gan et al. (2012) and Croonenbroeck and Dahl (2014). They show  
214 that WPPT and GWPPT are robust and universally applicable for turbines in  
215 different locations and of different types (Croonenbroeck and Dahl, 2014, use  
216 Vestas V90 type turbines). Mycielski, while useful for wind speed forecasts, is  
217 strictly outperformed when used as a wind power predictor. Methods initially  
218 dedicated to wind power forecasting, such as GWPPT, are to be preferred.

	10 min RMSE	10 min MAE	12 hours RMSE	12 hours MAE
<b>Turbine A</b>				
WPPT	109.28	75.68	153.28	107.50
GWPPT	102.07	72.80	<b>139.50</b>	<b>101.26</b>
Mycielski	135.77	93.63	406.34	312.06
AR	<b>99.46</b>	<b>70.25</b>	194.76	142.10
VAR	99.49	70.29	194.83	141.61
Persistence	101.23	70.42	346.33	266.67
<b>Turbine B</b>				
WPPT	113.52	78.72	157.18	112.28
GWPPT	104.28	74.30	<b>142.24</b>	<b>104.89</b>
Mycielski	142.47	100.40	432.47	335.35
AR	102.31	72.20	212.35	156.98
VAR	<b>102.15</b>	<b>71.84</b>	244.34	179.76
Persistence	105.01	73.03	348.34	274.30
<b>Turbine C</b>				
WPPT	120.07	82.46	168.28	122.42
GWPPT	111.09	77.17	<b>153.47</b>	<b>112.59</b>
Mycielski	158.43	111.53	437.74	338.59
AR	108.82	75.87	208.81	153.56
VAR	<b>108.47</b>	<b>75.32</b>	226.25	163.75
Persistence	111.12	76.45	355.64	278.16
<b>Turbine D</b>				
WPPT	123.44	82.66	182.85	126.41
GWPPT	115.51	78.15	<b>168.08</b>	<b>117.86</b>
Mycielski	142.54	98.37	431.77	327.05
AR	<b>107.61</b>	<b>74.11</b>	226.02	162.68
VAR	107.82	74.29	243.68	173.34
Persistence	109.04	73.86	357.49	277.90

Table 4. RMSE and MAE for 1 step ahead and 12 hours ahead forecasts.

## References

- Croonenbroeck, C. and Dahl, C.M. (2014), Accurate Medium-Term Wind Power Forecasting in a Censored Classification Framework, *Viadrina Discussion Papers*, 351, pp. 1–31.
- Ehrenfeucht, A. and Mycielski, J. (1992), A Pseudorandom Sequence – How Random Is It?, *The American Mathematical Monthly*, 99(4), pp. 373–375.
- Gan, M., Ding, M., Huang, Y.z., and Dong, X.p. (2012), The Effect of Different State Sizes on Mycielski Approach for Wind Speed Prediction, *Journal of Wind Engineering and Industrial Aerodynamics*, 109, pp. 89–93.
- 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, Risø DTU, Wind Energy Division.
- Hocaoglu, F., Fidan, M., and Gerek, m. (2009), Mycielski Approach for Wind Speed Prediction, *Energy Conversion and Management*, 50, pp. 1436–1443.
- Jacquet, P., Szpankowski, W., and Apostol, L. (2002), A Universal Predictor Based on Pattern Matching, *IEEE Transactions on Information Theory*, 48(6), pp. 1462–1472.
- 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, pp. 915–920.
- Nielsen, H.A., Pinson, P., Christiansen, L.E., Nielsen, T.S., Madsen, H., Badger, J., Giebel, G., and Ravn, H.F. (2007), Improvement and Automation of Tools for Short Term Wind Power Forecasting, Tech. rep., Scientific Proceedings of the European Wind Energy Conference & Exhibition, Milan, Italy.
- Taylor, J.W., McSharry, P.E., and Buizza, R. (2009), Wind power density forecasting using ensemble predictions and time series models, *Energy Conversion, IEEE Transactions on*, 24(3), pp. 775–782.
- Wu, Y.K. and Hong, J.S. (2007), A literature review of wind forecasting technology in the world, *Power Tech, 2007 IEEE Lausanne*, IEEE, pp. 504–509.