Probabilistic Forecasting of Wind Power Ramp Events using Autoregressive Logit Models

James W. Taylor
Saïd Business School, University of Oxford, Park End Street, Oxford, OX1 1HP, UK.
james.taylor@sbs.ox.ac.uk

European Journal of Operational Research, 2017, Vol. 259, pp. 703-712.

Abstract

A challenge for the efficient operation of power systems and wind farms is the occurrence of wind power ramps, which are sudden large changes in the power output from a wind farm. This paper considers the probabilistic forecasting of a ramp event, defined as exceedance beyond a specified threshold. We directly model the exceedance probability using autoregressive logit models fitted to the change in wind power. These models can be estimated by maximising a Bernoulli likelihood. We introduce a model that simultaneously estimates the ramp event probabilities for different thresholds using a multinomial logit structure and categorical distribution. To model jointly the probability of ramp events at more than one wind farm, we develop a multinomial logit formulation, with parameters estimated using a bivariate Bernoulli distribution. We use a similar approach in a model for jointly predicting one and two steps-ahead. We evaluate post-sample probability forecast accuracy using hourly wind power data from four wind farms.

Key words: OR in energy; wind power ramps; probability forecasting; autoregressive logit models

1. Introduction

Wind power is the world's fastest growing source of renewable energy. For electricity systems with high wind power penetration, forecasts of wind power are needed to support the efficient dispatch of generation units and to maintain the reliability of the system. Due to the uncertainty in the weather, wind power is inherently stochastic, and so, in order to convey a realistic representation of future power generation, probabilistic forecasts are required. Probabilistic wind power forecasting enables system operators to optimise levels of reserve, and allows a wind farm to improve operational efficiency and refine the risk-return trade-off in bidding strategies (Pinson and Kariniotakis, 2010).

In forecasting wind power, a significant challenge is the occurrence of wind power ramps, which are sudden large changes in the power generated from a wind farm, resulting from variations in wind conditions. An upward ramp needs to be swiftly balanced by a reduction in other sources of generation, and a downward ramp requires the use of reserve load, or a request for power, at short notice, from other sources (see, for example, Wang et al., 2015). This increases costs, and reduces the benefit in terms of carbon emissions. In terms of immediate financial impact, a downward ramp is sometimes considered more serious than an upward ramp of equal size (see, for example, Sherry and Rival, 2015). However, there is no consensus on this, as it will depend on their respective costs in the power system of interest. Research in the area of ramp forecasting is still in its infancy (Gallego et al., 2013). In this paper, we develop models for predicting the probability of a ramp event. Such probabilistic forecasts have been used by system operators to improve situational awareness through the use of graphical displays (see, for example, Zack et al., 2010). As operators gain confidence in the quality of the forecasts, they could feed directly into an automated system, which would incorporate tolerances regarding ramp size, timing and likelihood (Cui et al., 2015). This system could initiate the use of reserve, prompt the lowering of output from other sources, and terminate power output from a wind farm. Potter et al. (2009) show how reserve requirements can be optimized by incorporating forecasts of the probability of a ramp event in an analysis of system operating costs.

Wind power probability forecasts can be based on probabilistic forecasts of wind speed and other meteorological variables, generated from a statistical time series model (see, for example, Hering and Genton, 2010) or a physical model, such as a Numerical Weather Prediction (NWP) system (see, for example, Pinson and Madsen, 2009). Probabilistic prediction from an NWP model relies on ensemble predictions, which are generated by running the model with multiple different initial conditions. This limits the geographical

coverage and frequency of updating of the resultant forecasts. For short lead times, such as one or two hours, statistical models can be very competitive, because recent meteorological observations can better describe the state of the atmosphere than a physical model (Yoder et al., 2014). However, basing wind power prediction on meteorological variables is complicated by the relationship being stochastic (Jeon and Taylor, 2012). In view of this, and our interest in short lead times, we use statistical models of wind power based on only historical wind power observations.

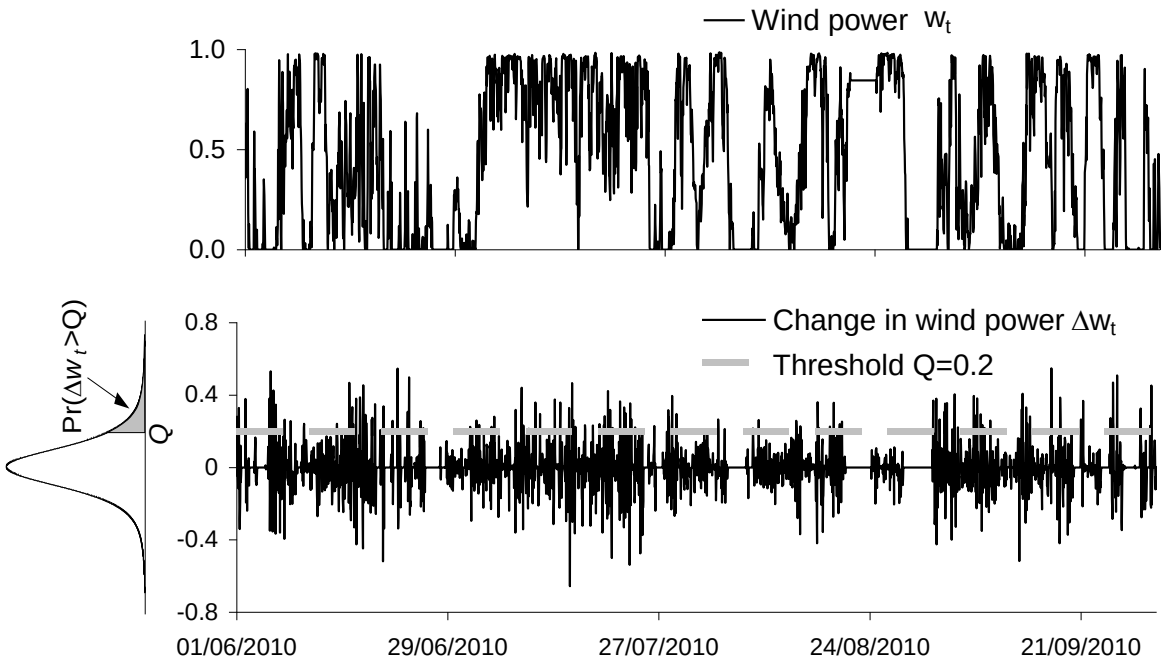
Of the methods designed specifically for predicting ramp events, the great majority do not deliver probabilistic forecasts (see, for example, Cutler et al., 2007; Zheng and Kusiak, 2009; Gallego et al., 2011). Exceptions to this are the methods of Bossavy et al. (2010, 2013) and Cui et al. (2015). Bossavy et al. (2010) use quantile regression based on point forecasts of wind power, speed and direction, as well as information on the magnitude and timing of the most recent ramp. Bossavy et al. (2010, 2013) present methods based on weather ensemble predictions, and consider forecasting up to several days. Cui et al. (2015) derive a probabilistic forecast of a ramp event from a wind power density forecast. They use a neural network with lagged wind power values as inputs for predicting ramp event probabilities for each hour of the next day. By contrast, our interest is in lead times of just one or two hours.

There is no consensus regarding the precise definition of a ramp (Gallego et al., 2013; Bossavy et al., 2015). They are usually considered to be large changes in power occurring over several minutes or hours, and their magnitude is typically measured as a percentage of the capacity of the wind farm. Interest can be in the change in power between two periods, the cumulative change over several periods, or the difference between the maximum and minimum power during a number of periods (see Ferreira et al., 2010). In this paper, we use hourly data, and define a ramp simply as a change in successive wind power observations that exceeds a specified constant threshold Q. Our aim is to forecast the probability of exceedance, as illustrated in Fig. 1 for data from the Aeolos wind farm in Crete. We express wind power as the capacity factor w_t, which is calculated as generated power divided by the wind farm's capacity. Fig. 1 shows wind power changes Δw_t, a threshold of 0.2, and a rotated probability density function showing the ramp event probability.

In principle, a ramp event probability forecast could be obtained from the predictive distribution of a statistical model fitted to wind power, w_t, or wind power changes, Δw_t. However, this requires a distributional assumption, which is not obvious, and so we directly model the probability using discrete choice models (see, for example, de Jong and Woutersen, 2011). More specifically, we adapt the conditional autoregressive logit (CARL)

models of Taylor and Yu (2016), who developed the models for financial returns series. Such data has similarities to our series of Δw_t in that it is high-frequency, and possesses apparent autocorrelation in the variance.

Fig. 1. Wind power w_t (upper panel) and power changes Δw_t (lower panel) for hourly Aeolos data. Rotated density indicates ramp event probability for threshold, Q=0.2.



In addition to adapting the CARL models for ramp event forecasting, we develop three new models, which we term *conditional autoregressive multinomial logit* (CARML). The first simultaneously estimates the probabilities of exceeding different thresholds, using a categorical distribution. The second relates to the literature on the spatial modelling of wind power (see, for example, Gneiting et al., 2006; Hering and Genton, 2010; Elberg and Hagspiel, 2015), as it involves jointly modelling the probability of a ramp event at two wind farms, using a bivariate Bernoulli distribution. The third CARML model uses a similar approach for jointly predicting one and two steps-ahead.

Section 2 describes the data used in this paper. Section 3 introduces the CARL and CARML models. Section 4 presents an empirical study comparing forecast accuracy. Section 5 concludes the paper.

2. The Wind Data

In this paper, we use hourly wind power data, recorded for 2010 at the following four wind farms on the Greek island of Crete: Aeolos, Iweco, Plastika and Rokas. Each hourly observation is the average of the six wind power readings recorded in the previous hour at 10 minute intervals. For all four wind power series, observations were missing for the same 367 of the 8760 hourly periods. These 367 periods were excluded from our analysis. As shown in Fig. 2, all four wind farms are in the east of the island, which is about 160 miles in length. In 2010, the capacities of Aeolos, Iweco, Plastika and Rokas were 8.3 MW, 4.5 MW, 11.8 MW and 11.5 MW, respectively. As stated in Section 1, we work with the capacity factor w_t, which is measured on a scale of 0 to 1. Crete is an interesting focus for our study, because, as noted by Sørensen et al. (2007), ramps can have particularly serious implications for the operation and security of a small island's power system, especially when there is no interconnector to the mainland, which is the case for Crete.

Fig. 2. Locations of the four wind farms.

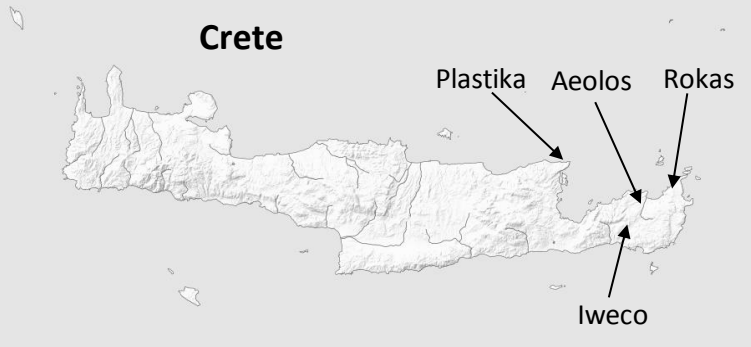


Table 1 provides summary statistics for each series of wind power changes Δw_t, calculated using the full year of data. For a Gaussian distribution, skewness is 0 and kurtosis is 3, and so it is interesting to see the high kurtosis values in Table 1, indicating that the series are far from Gaussian. It is also interesting to note that the autocorrelation is significant for Δw_t and Δw_t^2. Autocorrelation in Δw_t^2 suggests that autoregressive modelling of the variance in Δw_t may be useful within a model for the probability of a ramp event. This is supported by the clustering of periods of relatively high volatility and relative tranquillity in Fig. 1.

Table 2 presents the correlations between the four wind power series, and Table 3 provides the correlations between the four series of wind power changes. The full year of data was used to calculate these values. The correlations between Aeolos, Plastika and Rokas are higher than between any of these three and Iweco. The reason for this is not clear, although we note that Iweco is somewhat further from the north coast than the other three wind farms.

Table 1. Summary statistics for the four time series of changes in wind power Δw_t.

	Aeolos	Iweco	Plastika	Rokas
Mean	0.000	0.000	0.000	0.000
Standard deviation	0.126	0.125	0.110	0.097
Skewness	0.144	0.072	0.260	0.060
Kurtosis	7.215	8.015	8.504	9.260
Autocorrelation at lag 1 in Δw_t	0.049**	0.071**	0.114**	0.133**
Autocorrelation at lag 1 in Δw_t^2	0.187**	0.170**	0.190**	0.132**

* and ** indicate significant autocorrelation at 5% and 1% levels, respectively.

Table 2. Correlation between the four time series of wind power w_t.

	Aeolos	Iweco	Plastika
Iweco	0.55**		
Plastika	0.71**	0.45**	
Rokas	0.76**	0.51**	0.73**

* and ** indicate significance at 5% and 1% levels, respectively.

Table 3. Correlation between the four time series of changes in wind power Δw_t.

	Aeolos	Iweco	Plastika
Iweco	0.07**		
Plastika	0.13**	0.06**	
Rokas	0.26**	0.03**	0.17**

* and ** indicate significance at 5% and 1% levels, respectively.

Table 4. Percentage of periods with wind power change Δw_t exceeding each threshold.

	Threshold					
	-0.3	-0.2	-0.1	0.1	0.2	0.3
Aeolos	2.2	5.6	13.8	14.0	5.8	2.3
Iweco	2.1	5.4	13.0	13.1	5.6	2.3
Plastika	1.2	3.5	12.3	12.0	4.2	1.4
Rokas	1.0	2.9	9.9	10.8	3.0	0.9

In our empirical work, we consider six thresholds: -0.3, -0.2, -0.1, 0.1, 0.2 and 0.3. These span the range of commonly considered thresholds for wind power ramps (see Ferreira et al., 2010; Gallego et al., 2013). Table 4 shows the percentages of wind power changes exceeding each threshold in 2010. Each percentage is an estimate of the *unconditional* expectation of the probability of the wind power change exceeding the threshold. Our aim is to model the *conditional* expectation of this probability. As we explained in Section 1, to

model this probability, we use statistical models based on only historical wind power observations. As we are modelling the probability of a ramp, it seems natural to use, as a covariate, a variable that indicates whether or not a ramp occurred in the previous period. In addition, in view of the autocorrelations shown in Table 1, we also consider, as covariates, lagged wind power changes and measures of the volatility of wind power changes.

For each month, Figs. 3 and 4 show the percentage of periods for which the wind power change Δw_t exceeded the thresholds 0.2 and -0.2, respectively. For both thresholds, the total number of ramps for the four locations was the highest for February and the lowest for August. This was also the case for the other four thresholds, except 0.1 and -0.3, for which January had slightly more ramps than February. However, as we are using just one year of data, we cannot make confident inference regarding seasonality. In our empirical work, we re-estimate model parameters for each month using just the previous four months of data, which enables the parameter values to change across the year. If we had a time series extending over several years, our modelling could also consider covariates aimed at capturing the annual seasonality.

Fig. 3. For each month, percentage of periods with wind power change Δw_t exceeding the threshold Q=0.2.

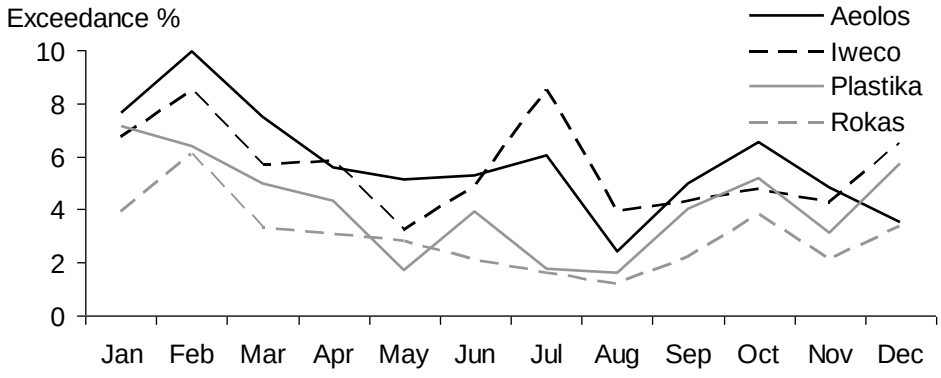
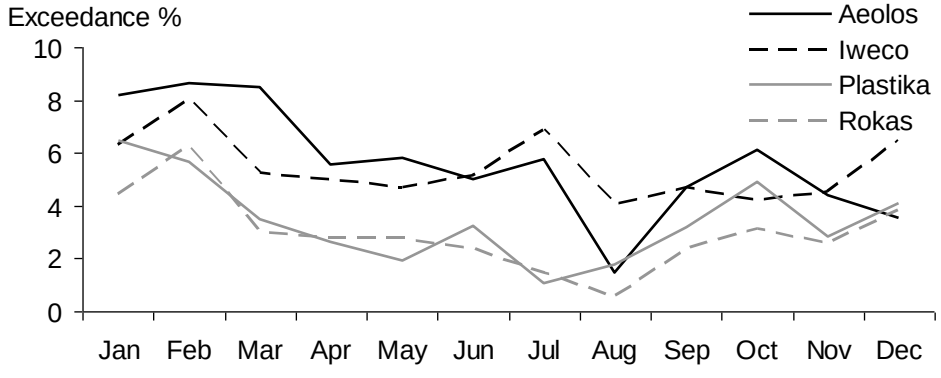


Fig. 4. For each month, percentage of periods with wind power change Δw_t exceeding the threshold Q=-0.2.



3. CARL and CARML Models

3.1. CARL Models for a Single Threshold and a Single Wind Power Series

In this section, we adapt the conditional autoregressive logit (CARL) models of Taylor and Yu (2016) for the wind power ramp context, where we model the probability of Δw_t falling beyond a threshold Q. For simplicity, throughout Section 3.1, we present models for only $Q>0$. We use the logistic function of expression (1), with five alternative formulations for the logit x_t. The parameters in each are estimated separately for different values of Q. This amounts to a rather flexible modelling of the conditional distribution of Δw_t with dynamics that potentially differ across the distribution.

$$\Pr(\Delta w_t > Q) = \frac{1}{1+\exp(-x_t)} \tag{1}$$

Our first formulation for the logit is the *CARL-Indicator* model of expressions (2) and (3). The term $I(\Delta w_{t-1} \le Q)$ allows the probability to change according to whether or not there was a ramp in the previous period, and the inclusion of this term, along with the term $I(\Delta w_{t-1} \ge -Q)$, helps the model adjust to the volatility in the data.

$$x_t = \begin{cases} h_t & \text{if} \quad w_{t-1} \le 1-Q \\ L & \text{otherwise} \end{cases} \tag{2}$$

$$h_t = \alpha_0 + \alpha_1 I(\Delta w_{t-1} \le Q) + \alpha_2 I(\Delta w_{t-1} \ge -Q) + \beta_1 h_{t-1} \tag{3}$$

where α_i and β_1 are constant parameters. Throughout this paper, L is chosen as a large negative number. Recall that $w_{t-1} \in [0,1]$, and that we are considering $Q>0$. If $w_{t-1}>1-Q$, then Δw_t cannot take a value above Q. This is captured in expression (2), because if $w_{t-1}>1-Q$, the logit is set to L, which forces the probability to be close to 0. This feature was not needed in the CARL models of Taylor and Yu (2016), which were developed for financial returns. The CARL-Indicator model for a negative threshold is presented in Appendix A.

In view of the autocorrelation in the volatility of Δw_t, which we noted in Section 2, the following *CARL-Absolute* model is specified to adapt to changes in the magnitude of Δw_t:

$$x_t = \begin{cases} h_t & \text{if} \quad w_{t-1} \le 1-Q \\ L & \text{otherwise} \end{cases}$$

$$h_t = \alpha_0 + \alpha_1 |\Delta w_{t-1}| + \beta_1 h_{t-1}$$

where α_i and β_1 are constant parameters.

The *CARL-GARCH* model of expressions (4) and (5) uses a GARCH formulation to capture the autocorrelation in the volatility of Δw_t.

$$x_t = \begin{cases} \gamma_0 + \gamma_1 h_t^{-\frac{1}{2}} & \text{if} \quad w_{t-1} \le 1-Q \\ L & \text{otherwise} \end{cases} \tag{4}$$

$$h_t = \alpha_0 + \alpha_1 \Delta w_{t-1}^2 + \beta_1 h_{t-1} \tag{5}$$

where α_i, β_1 and γ_i are constant parameters. This model can be motivated by considering the assumption of a constant distribution F_1 for standardised values of Δw_t. This would imply $\Pr(\Delta w_t \le Q) = F_1((Q-\mu)/h_t^{\frac{1}{2}})$, where h_t is the variance, and μ is the mean, which we assume to be zero. This can be rewritten as $\Pr(\Delta w_t > Q) = F_2((\mu - Q)/h_t^{\frac{1}{2}})$, where F_2 is a constant distribution function. This suggests that the probability can be modelled using a logistic function with the logit term a linear function of $h_t^{-\frac{1}{2}}$, and this is the form of expression (4). We impose the constraints $\alpha_1, \beta_1 \ge 0$, and, given that the variance of Δw_t is stationary, we set $\alpha_0=(1-\alpha_1-\beta_1)h$, where h is the variance of the in-sample values of Δw_t. This approach to reducing the number of parameters in a variance model is known as *variance targeting*.

To capture the autocorrelation in Δw_t, which was evident in Table 1, we include Δw_{t-1} in the logit expression, as shown in expression (6). We term this *CARL-GARCH with* Δw_{t-1}.

$$x_t = \begin{cases} \gamma_0 + \gamma_1 h_t^{-\frac{1}{2}} + \phi_1 \Delta w_{t-1} & \text{if} \quad w_{t-1} \le 1-Q \\ L & \text{otherwise} \end{cases} \tag{6}$$

where α_i, β_1, γ_i and ϕ_1 are constant parameters, and h_t is modelled as in expression (5). We also considered w_{t-1} in the logit expression, but this did not improve the empirical results. We present the model for a negative threshold in Appendix B.

Following the standard approach for binary choice models, we estimated the CARL models using the Bernoulli likelihood of expression (7), where T is the sample size.

$$\prod_{t=1}^{T} \Pr(\Delta w_t \le Q)^{I(\Delta w_t \le Q)} (1 - \Pr(\Delta w_t \le Q))^{1-I(\Delta w_t \le Q)} \tag{7}$$

The CARL models can be adapted for other definitions of a ramp, such as a ramp being defined as a large change between periods that are not successive (see Ferreira et al., 2010). For example, if interest is in the change over three periods, this could be modelled by replacing Δw_t with $(w_t - w_{t-3})$ in the likelihood of expression (7).

Table 5 presents the parameters for CARL-GARCH with Δw_{t-1}, estimated using four-month windows of the Aeolos data for the threshold Q=0.2. As is common with GARCH models of daily financial returns, the sum of α_1 and β_1 is just below 1 for each model. The negative values for γ_1 are intuitive because they imply that, as the volatility of Δw_t rises, there is an increase in the probability of w_t rising by at least 0.2. The positive values for ϕ_1 imply

that, if w_t increased in the previous period, the probability of it then rising by at least 0.2 is more than if w_t had decreased in the previous period. This seems reasonable, given the autocorrelation in Δw_{t-1} that we noted in Table 1. Although there are some sizeable differences in the values of parameters when estimated using the different four-month windows, it is not possible to make confident inference regarding seasonality, as we are using just one year of data in this paper.

Table 5. Parameters of CARL-GARCH with Δw_{t-1}, estimated using four-month windows of Aeolos data for threshold Q=0.2.

In-Sample	γ_0	γ_1	ϕ_1	α_0	α_1	β_1
Jan-Apr	-1.628**	-0.032*	0.307	0.000227	0.970**	0.020
Feb-May	1.604**	-0.040	0.404	0.000140	0.782**	0.211*
Mar-Jun	-0.946**	-0.144**	0.851*	0.000105	0.582**	0.351**
Apr-Jul	-0.787*	-0.161**	0.977*	0.001101	0.560**	0.365**
May-Aug	-1.069**	-0.118**	1.380**	0.000616	0.748**	0.198*
Jun-Sep	-0.950**	-0.133**	1.115*	0.000721	0.693**	0.250**
Jul-Oct	-0.949**	-0.138**	0.948	0.000676	0.545**	0.405**
Aug-Nov	-1.483**	-0.090**	1.282*	0.000270	0.531**	0.447**

* and ** indicate significance at 5% and 1% levels, respectively.

For Aeolos, the upper and middle panels of Fig. 5 show the October observations for w_t and Δw_t, respectively, and the threshold Q=0.2. The lower panel shows the one hour-ahead post-sample probability forecasts $\Pr(\Delta w_t > 0.2)$, produced from the CARL-Indicator model and CARL-GARCH with Δw_{t-1}, estimated using the four-month window of data from June to September, inclusive. When w_t is very high, it cannot rise by 0.2, and so, as expected, the probability forecasts are close to 0 in Fig. 5.

For the same periods of the Aeolos data, Fig. 6 shows the Δw_t observations, the threshold Q=-0.2, and the corresponding forecasts for the ramp probability $\Pr(\Delta w_t \le -0.2)$. There is some indication in the probability forecasts of Figs. 5 and 6 that, when there is high volatility in w_t (and hence in Δw_t), a ramp event is more likely. Comparing the probability forecasts from the two models, we see that the CARL-GARCH forecasts are more responsive to changes in Δw_t. This is understandable, because the CARL-GARCH forecasts react to the magnitude of Δw_t, while the CARL-Indicator forecasts adjust by fixed amounts depending on whether or not Δw_t exceeds Q or -Q.

Fig. 5. Wind power w_t (upper panel), change in wind power Δw_t (middle panel), and post-sample forecasts (lower panel) of $\Pr(\Delta w_t>0.2)$ from two CARL models, estimated using Aeolos data for June to September, inclusive.

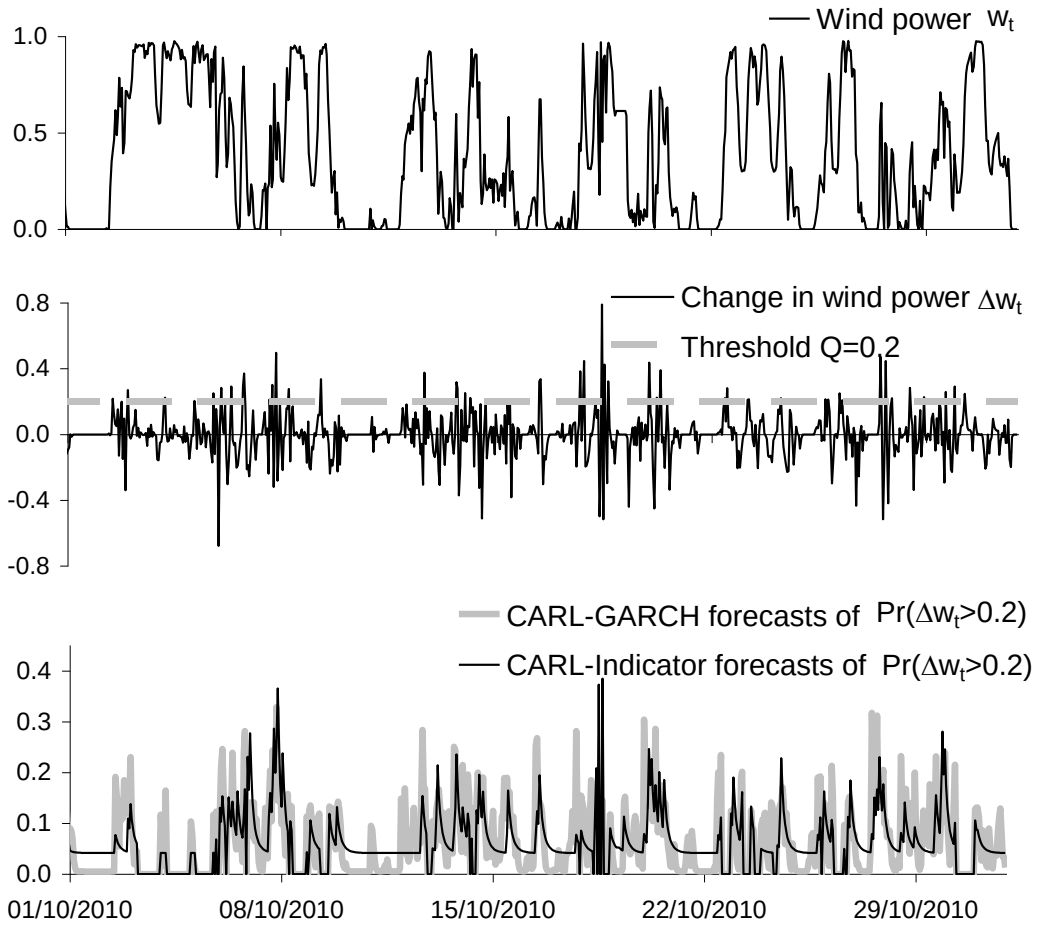
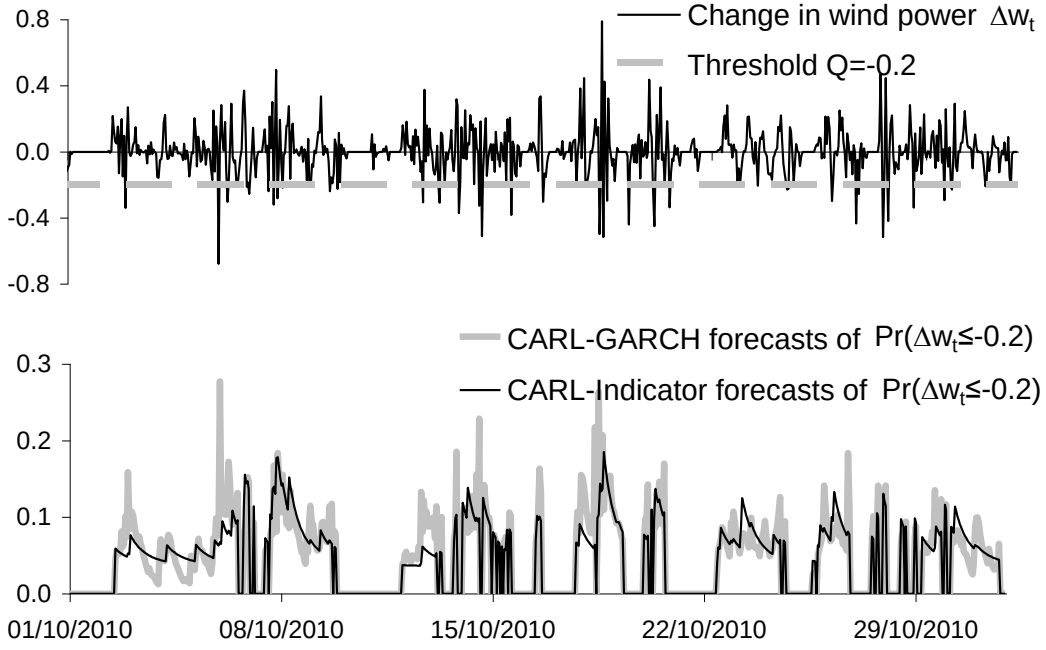


Fig. 6. Change in wind power Δw_t (upper panel) and post-sample forecasts (lower panel) of $\Pr(\Delta w_t\leq-0.2)$ from two CARL models, estimated using Aeolos data for June to September, inclusive.



3.2. A Multi-Threshold CARML Model for a Single Wind Power Series

If there is interest in predicting the probabilities of ramp events for ramps of different sizes in a single wind power time series, it would seem efficient to model the probabilities together. Furthermore, such simultaneous estimation enables us to avoid a ramp event probability estimate being higher for the more extreme of two thresholds of the same sign. We term this problem *probability crossing*; it is analogous to quantile crossing in quantile estimation (see Section 2.5, Koenker, 2005).

In expressions (8)-(12), we present a multi-threshold conditional autoregressive multinomial (CARML) model for K_1 negative and K_2 positive thresholds. We use the notation Q_1 to Q_{K_1} for the negative thresholds, and Q_{K_1+1} to $Q_{K_1+K_2}$ for the positive thresholds.

$$\Pr\left(Q_{i-1} < \Delta w_t \le Q_i\right) = \frac{\exp\left(x_{it}\right)}{1+\sum_{j=1}^{K_1+K_2}\exp\left(x_{jt}\right)} \qquad i = 1 \text{ to } K_1 \tag{8}$$

$$\Pr\left(Q_i < \Delta w_t \le Q_{i+1}\right) = \frac{\exp\left(x_{it}\right)}{1+\sum_{j=1}^{K_1+K_2}\exp\left(x_{jt}\right)} \qquad i = (K_1+1) \text{ to } (K_1+K_2) \tag{9}$$

$$\Pr\left(Q_{K_1} < \Delta w_t \le Q_{K_1+1}\right) = \frac{1}{1+\sum_{j=1}^{K_1+K_2}\exp\left(x_{jt}\right)} \tag{10}$$

$$x_{it} = \begin{cases} \gamma_{0i} + \gamma_{1i} h_t^{-\frac{1}{2}} + \phi_{1i}\Delta w_{t-1} & \text{if } w_{t-1} \ge -Q_i \text{ and } w_{t-1} \le 1-Q_i \\ L & \text{otherwise} \end{cases} \qquad i = 1 \text{ to } (K_1+K_2) \tag{11}$$

$$h_t = \alpha_0 + \alpha_1 \Delta w_{t-1}^2 + \beta_1 h_{t-1} \tag{12}$$

where α_i, β_1, γ_{ji} and ϕ_{1i} are constant parameters. We order the thresholds so that $Q_i<Q_{i+1}$, and we define extreme thresholds as Q_0=-1.01 and $Q_{K_1+K_2+1}$=1. Expressions (8)-(10) show a multinomial logit structure. For each of the K_1 negative thresholds (i=1 to K_1), expression (8) shows a logit x_{it} corresponding to the probability of Δw_t falling in the interval $(Q_{i-1},Q_i]$. For each of the K_2 positive thresholds (i=(K_1+1) to (K_1+K_2)), expression (9) shows a logit x_{it} corresponding to the probability of Δw_t falling in the interval $(Q_i,Q_{i+1}]$. Each logit x_{it}, therefore, relates to the probability of exceeding Q_i, but not exceeding the adjacent threshold that has larger magnitude than Q_i. Defining the logits like this allows us to impose, in expression (11), that the exceedance probability is close to 0 if $w_t<-Q_i$ for a negative threshold, and if $w_t>1-Q_i$ for a positive threshold.

We estimate the model by maximising the likelihood of expression (13). This is based on a categorical distribution, which is a generalisation of the Bernoulli distribution for a random variable with more than two possible outcomes.

$$\prod_{t=1}^{T}\prod_{i=1}^{K_1+K_2+1}\Pr\left(Q_{i-1}<\Delta w_t\le Q_i\right)^{I\left(Q_{i-1}<\Delta w_t\le Q_i\right)} \tag{13}$$

Table 6 presents the parameters of the model with all six thresholds, estimated using the Aeolos wind farm data from June to September, inclusive. For this model, we have K_1=3 negative thresholds (Q_1=-0.3, Q_2=-0.2 and Q_3=-0.1) and K_2=3 positive thresholds (Q_4=0.1, Q_5=0.2 and Q_6=0.3). The negative values for γ_{1i} are intuitive because they imply that, as the volatility in Δw_t increases, all six logits increase, and hence there is an increase in the probability of exceeding each threshold. ϕ_{11} and ϕ_{16} are not significant, suggesting that the sign of the change in wind power does not affect the probability of exceeding -0.3 or 0.3 in the next period. For the other thresholds, the signs of ϕ_{1i} indicate that a wind power change increases the probability of a change of the same sign in the next period.

Table 6. Parameters of multi-threshold CARML with thresholds -0.3, -0.2, -0.1, 0.1, 0.2 and 0.3, estimated using four months of Aeolos data.

i	γ_{0i}	γ_{1i}	ϕ_{1i}	α_0	α_1	β_1
				0.031	0.686**	0.284**
1	-1.906**	-0.104**	0.274			
2	-1.673**	-0.071**	-2.351**			
3	-0.969**	-0.050**	-1.749**			
4	-0.647**	-0.085**	1.324**			
5	-0.841**	-0.143**	1.499*			
6	-1.345**	-0.126**	1.078			

* and ** indicate significance at 5% and 1% levels, respectively.

3.3. A Multi-Location CARML Model for a Single Threshold and Two Wind Power Series

If an estimate is required for the probability of a ramp event at two wind farms, that are located relatively close to each other, joint modelling of the two ramp event probabilities may lead to improved estimation. For example, using the wind power observations from the two locations may improve the modelling of their volatilities. Also, the wind conditions at one location may provide a leading indicator of conditions at a downwind location. We inspected wind direction data from our four wind farm locations, but there was no clear visual

evidence that any one of our locations was generally downwind from another of the locations. Nevertheless, in light of the significant correlations, shown in Table 3, between the four series of wind power changes, we feel it is interesting to investigate further the empirical benefits of joint modelling for two locations.

We present here a CARML model for two wind farm locations and one positive threshold Q. In expressions (14)-(17), we use a multinomial logit structure for the probabilities of the four possible joint events regarding whether the changes in wind power at the two locations, Δw_{1t} and Δw_{2t}, are above or below the threshold. Expression (18) allows each logit to depend on wind power changes at the two wind farms, and on the corresponding volatilities, which we model in expression (19).

$$\Pr\left(\Delta w_{1t} \le Q, \Delta w_{2t} \le Q\right) = \frac{1}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{14}$$

$$\Pr\left(\Delta w_{1t} \le Q, \Delta w_{2t} > Q\right) = \frac{\exp\left(x_{1t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{15}$$

$$\Pr\left(\Delta w_{1t} > Q, \Delta w_{2t} \le Q\right) = \frac{\exp\left(x_{2t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{16}$$

$$\Pr\left(\Delta w_{1t} > Q, \Delta w_{2t} > Q\right) = \frac{\exp\left(x_{3t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{17}$$

$$x_{it} = \begin{cases} \gamma_{0i} + \sum_{j=1}^{2}\gamma_{ji} h_{jt}^{-\frac{1}{2}} + \sum_{j=1}^{2}\phi_{ji}\Delta w_{j,t-1} & \text{if} \quad w_{1,t-1} \le 1 - QI(i \ne 1) \quad \text{and} \quad w_{2,t-1} \le 1 - QI(i \ne 2) \\ L & \text{otherwise} \end{cases}$$

$$i=1 \text{ to } 3 \tag{18}$$

$$h_{jt} = \alpha_{0j} + \alpha_{1j}\Delta w_{j,t-1}^{2} + \beta_{1j} h_{j,t-1} \qquad j=1 \text{ to } 2 \tag{19}$$

where α_{ij}, β_{1j}, γ_{ji} and ϕ_{ji} are constant parameters.

As we are modelling the joint events regarding two dichotomous variables, to estimate model parameters, we maximise the likelihood of the bivariate Bernoulli distribution presented by Teugels (1990). This likelihood is shown in expression (20). A similar multinomial logit model structure to ours is considered by Dai et al. (2013), who also use a bivariate Bernoulli likelihood, although they do not consider an application to wind energy. In principle, our model could be extended for more than two wind farm locations, using a multivariate Bernoulli distribution for estimation (see Teugels, 1990). However, parameter

optimisation is likely to be challenging, because, of all the models in this paper, the multi-location model of (14)-(19) was the most difficult to optimise. We found that the optimisation was sensitive to the initial values of the parameters. We discuss our approach to optimisation in Section 4.1. The optimisation problem is fundamentally different to that for the other models in this paper, because the model is fitted to two series.

$$\prod_{t=1}^{T} \Pr\left(\Delta w_{1t} \le Q, \Delta w_{2t} \le Q\right)^{I\left(\Delta w_{1t} \le Q, \Delta w_{2t} \le Q\right)} \Pr\left(\Delta w_{1t} \le Q, \Delta w_{2t} > Q\right)^{I\left(\Delta w_{1t} \le Q, \Delta w_{2t} > Q\right)} \times \Pr\left(\Delta w_{1t} > Q, \Delta w_{2t} \le Q\right)^{I\left(\Delta w_{1t} > Q, \Delta w_{2t} \le Q\right)} \Pr\left(\Delta w_{1t} > Q, \Delta w_{2t} > Q\right)^{I\left(\Delta w_{1t} > Q, \Delta w_{2t} > Q\right)} \tag{20}$$

Table 7 presents the parameters of the model for a threshold of 0.2, and estimation based on the Aeolos and Plastika wind power time series from June to September, inclusive. In the model, w_{1t} and w_{2t} correspond to Aeolos and Plastika, respectively. It is interesting to see that the values of γ_{13} and γ_{23} are not significant, which suggests that the volatilities of Δw_{1t} and Δw_{2t} are not useful for modelling the probability that there will be a ramp at both locations. The value of ϕ_{13} is significant, implying that the sign of the change in wind power at Aeolos has an impact on this probability.

Table 7. Parameters of multi-location CARML with a single threshold Q=0.2, estimated using four months of Aeolos and Plastika data.

i	γ_{0i}	γ_{1i}	γ_{2i}	ϕ_{1i}	ϕ_{2i}	α_{0i}	α_{1i}	β_{1i}
1	-1.557**	0.002	-0.149*	-0.535	0.821	0.0004	0.710**	0.257**
2	-1.422**	-0.128**	0.023	0.401	3.700**	0.0005	0.302**	0.645**
3	-3.615**	-0.023	-0.068	5.511**	2.617			

* and ** indicate significance at 5% and 1% levels, respectively.

3.4. A Multi-Step-Ahead CARML Model for a Single Threshold and a Single Power Series

The models that we have proposed so far are suitable only for one step-ahead prediction. Gneiting et al. (2006) describe how a two-hour lead time, for wind power, is important for transmission, resource allocation and generation dispatch. In expressions (21)-(26), we present a CARML model for jointly modelling the ramp event probabilities for one and two step-ahead prediction, for the case of one wind power series and a single positive threshold. The model uses a similar structure and estimation approach to the multi-location model of the previous section.

$$\Pr\left(\Delta w_t \le Q, \Delta w_{t+1} \le Q\right) = \frac{1}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{21}$$

$$\Pr\left(\Delta w_t \le Q, \Delta w_{t+1} > Q\right) = \frac{\exp\left(x_{1t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{22}$$

$$\Pr\left(\Delta w_t > Q, \Delta w_{t+1} \le Q\right) = \frac{\exp\left(x_{2t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{23}$$

$$\Pr\left(\Delta w_t > Q, \Delta w_{t+1} > Q\right) = \frac{\exp\left(x_{3t}\right)}{1+\sum_{i=1}^{3}\exp\left(x_{it}\right)} \tag{24}$$

$$x_{it} = \begin{cases} \gamma_{0i} + \gamma_{1i} h_t^{-\frac{1}{2}} + \phi_{1i}\Delta w_{t-1} & \text{if } w_{t-1} \le 1 - Q(1+I(i=3)) \\ L & \text{otherwise} \end{cases} \qquad i=1 \text{ to } 3 \tag{25}$$

$$h_t = \alpha_0 + \alpha_1 \Delta w_{t-1}^2 + \beta_1 h_{t-1} \tag{26}$$

where α_i, β_1, γ_{ji} and ϕ_{1i} are constant parameters.

The parameters are estimated by maximising the bivariate Bernoulli likelihood of expression (27). Table 8 presents the parameters of the model with Q=0.2, estimated using the Aeolos data from June to September, inclusive. The values of γ_{1i} indicate that the volatility has an impact on the probability of a ramp event at each lead time. Indeed, the volatility seems to be of more importance in the model than the lag of the wind power changes. The values of ϕ_{12} and ϕ_{13} are not significant, implying that the lag of the change in wind power does not have a significant impact on the probability of a ramp event at the first lead time.

$$\begin{aligned} \prod_{t=1}^{T} & \Pr\left(\Delta w_t \le Q, \Delta w_{t+1} \le Q\right)^{I\left(\Delta w_t \le Q, \Delta w_{t+1} \le Q\right)} \Pr\left(\Delta w_t \le Q, \Delta w_{t+1} > Q\right)^{I\left(\Delta w_t \le Q, \Delta w_{t+1} > Q\right)} \\ & \times \Pr\left(\Delta w_t > Q, \Delta w_{t+1} \le Q\right)^{I\left(\Delta w_t > Q, \Delta w_{t+1} \le Q\right)} \Pr\left(\Delta w_t > Q, \Delta w_{t+1} > Q\right)^{I\left(\Delta w_t > Q, \Delta w_{t+1} > Q\right)} \end{aligned} \tag{27}$$

Table 8. Parameters of multi-step-ahead CARML with a single threshold Q=0.2, estimated using four months of Aeolos data.

i	γ_{0i}	γ_{1i}	ϕ_{1i}	α_0	α_1	β_1
1	-1.860**	-0.091**	-2.751**	0.001	0.651**	0.288**
2	-0.776**	-0.163**	0.940			
3	-2.939**	-0.101*	0.361			

* and ** indicate significance at 5% and 1% levels, respectively.

In Section 4, we evaluate the multi-step-ahead CARML model in terms of its ability to predict, from a forecast origin at period *t*, the probability of a ramp in period *t*+1 and the probability of a ramp in period *t*+2. It is worth noting that the model can also be used to predict the probability of a ramp occurring in either period *t*+1 or period *t*+2, and indeed a ramp event could be defined as a large change occurring in any one period within a given time interval.

4. Empirical Evaluation of Probability Forecasts

Our analysis proceeded by using the first four calendar months of 2010 to estimate model parameters. We then produced probability forecasts for all periods in the next month. We moved the window of four calendar months forward by one month, re-estimated the parameters, and again produced forecasts for the next month. We did this six more times to deliver, in total, forecasts for the final eight months of the year. These are post-sample forecasts, because each was generated using only data on or before the forecast origin. Our main focus was one hour-ahead prediction, but we also evaluated two hours ahead from the multi-step-ahead CARML model.

4.1. Probability Forecasting Methods

As a simple benchmark, we produced forecasts for the probability $\Pr(\Delta w_t \le Q)$ as the proportion of the previous four calendar months that the observed Δw_t were less than the threshold Q. This would be termed by meteorologists the *climatology forecast*, updated each month. As more sophisticated benchmarks, we implemented an exponentially weighted moving average (EWMA) and a GARCH(1,1) model to estimate the standard deviation σ_t of Δw_t. We optimised the EWMA smoothing parameter by minimising the sum of in-sample variance forecast errors, and we fitted the GARCH model using a Student-*t* distribution. From the empirical distribution of the standardised changes $\Delta w_t/\sigma_t$, we derived the probability of a standardised change being less than or equal to Q/σ_t for the forecast period. We used this as the estimate of $\Pr(\Delta w_t \le Q)$. We also implemented a quantile regression approach, motivated by the work of Bossavy et al., who produce a distributional forecast based on quantile models corresponding to probability levels 5% to 95%, with 5% increments. In their models, they use wind speed and direction forecasts, and consider lead times much longer than ours. In our study, we used autoregressive quantile models. More specifically, we modelled the quantiles of Δw_t using the symmetric absolute value conditional autoregressive value at risk (CAViaR) model of Engle and Manganelli (2004), which is estimated using

quantile regression. We derived each forecast of $\Pr(\Delta w_t \le Q)$ from a distributional forecast based on quantile models corresponding to probability levels 1% to 99%, with 1% increments. To produce the full distributional forecast, we used linear interpolation between quantile forecasts, and bounded the estimated distribution by the historical lowest and highest values of Δw_t. If $w_{t-1}<-Q$ or $w_{t-1}>1-Q$, a ramp was not possible for period t, and so in this case we set the ramp probability forecasts from all methods to be zero for period t.

We implemented the CARL and CARML models of Section 3. For each, we initialised h_t using the first 100 observations. For example, in the CARL-GARCH models, h_1 was calculated as the variance of the first 100 observations for Δw_t. We considered two versions of the multi-threshold model. The first included all six thresholds. The second involved one model estimated for the three negative thresholds, and a separate model for the three positive thresholds. To maximise the likelihoods, we used a similar optimisation approach to that used by Engle and Manganelli (2004) for their CAViaR quantile models. It proceeded by sampling J vectors of parameters, using a uniform random number generator to sample values between a lower and upper bound, which were set for each parameter based on initial experimentation. Of the J sampled vectors, the three that gave the highest likelihood values were used, in turn, as the initial vector in a quasi-Newton algorithm. The resulting vector, corresponding to the highest log likelihood, was chosen as the final parameter vector. We set $J=10^5$ for the multi-location model and the multi-threshold model with six thresholds, while for all other models, we used $J=10^4$.

4.2. Post-Sample Results

The Brier score is commonly used to evaluate probability forecasts. It is presented in expression (28) for a forecast $\hat{p}_t$ of $\Pr(y_t \le Q)$, and N post-sample periods. It is a proper scoring rule, meaning that the expected value of the score is lowest for the true probability (see, for example, Gneiting et al., 2007). We calculated the Brier score for each method, along with the logarithmic score, which is another popular evaluation measure (see, for example, Wilks, 2011). As the ranking of the methods for these two measures was very similar, we consider only the Brier score in the remainder of this paper.

$$\frac{1}{N}\sum_{t=T+1}^{T+N}\left(I(y_t \le Q) - \hat{p}_t\right)^2 \tag{28}$$

For each method, we calculated the Brier skill score, which is shown in expression (29) (see, for example, Wilks, 2011). This measure compares the Brier score to a reference

method, which we chose as the simple benchmark approach. The measure was used by Bossavy et al. (2010, 2013) to evaluate ramp event probability forecasts.

$$\text{Brier skill score} = \left(1 - \frac{\sum_{t=n+1}^{T+N} \left(I\left(y_t \le Q\right) - \hat{p}_t \right)^2}{\sum_{t=n+1}^{T+N} \left(I\left(y_t \le Q\right) - \hat{p}_{reference,t} \right)^2} \right) \times 100 \qquad (29)$$

Table 9 presents the results for the Aeolos wind farm. Higher values indicate superior accuracy, and positive values imply greater accuracy than the reference method. The final column summarises performance across the six thresholds. To obtain the values in this column, we calculated the geometric mean of the ratios of the Brier score for each method to the Brier score for the reference method, we then subtracted 1, and multiplied the result by 100. The same form of calculation was used to produce Table 10, which averages the Brier score results across the four wind farms. In the tables, for each of the six thresholds, and for the summary column, bold indicates the best three results.

Table 9. For one hour-ahead ramp event probability forecasts, Brier skill scores for Aeolos.

	Threshold						
	-0.3	-0.2	-0.1	0.1	0.2	0.3	Geometric mean
Benchmarks							
EWMA for variance	**1.3**	**2.5**	3.8	4.5	3.0	1.0	2.7
GARCH for variance	-0.5	1.7	3.5	6.0	3.6	0.4	2.4
CAViaR	-0.1	1.5	2.1	7.0	3.5	1.1	2.5
CARL (Section 3.1)							
CARL-Indicator	-0.4	0.4	3.0	4.8	1.6	0.6	1.7
CARL-Absolute	0.0	1.4	2.8	5.0	3.0	1.7	2.3
CARL-GARCH	**1.0**	1.8	3.4	7.7	4.6	**2.3**	3.4
CARL-GARCH with Δw_{t-1}	0.7	**1.9**	4.9	8.5	4.7	**2.3**	3.8
Multi-threshold CARML (Section 3.2)							
6 thresholds	0.3	1.2	5.1	**9.6**	**4.9**	2.0	3.8
3 thresholds (Q_i<0 separate from Q_i>0)	0.5	1.4	5.2	**9.8**	**4.9**	1.8	**3.9**
Multi-location CARML (Section 3.3)							
Aeolos & Iweco	0.9	1.1	4.4	8.4	4.6	1.9	3.5
Aeolos & Plastika	**2.1**	1.6	**5.8**	**10.0**	**5.8**	0.2	**4.2**
Aeolos & Rokas	0.4	1.2	**5.5**	8.5	3.9	1.8	3.5
Multi-step-ahead CARML (Section 3.4)							
Model for 1 & 2 hours ahead	0.8	**2.2**	**5.3**	9.3	**5.0**	**2.3**	**4.1**

Notes. Higher values are better. Bold indicates best three methods in each column.

Table 10. For one hour-ahead ramp event probability forecasts, Brier skill scores averaged across the four wind farms.

	Threshold						
	-0.3	-0.2	-0.1	0.1	0.2	0.3	Geometric mean
Benchmarks							
EWMA for variance	2.2	3.3	4.4	2.6	1.5	0.3	2.4
GARCH for variance	1.2	3.0	5.4	4.3	1.4	-1.9	2.2
CAViaR	**2.4**	3.5	4.5	4.6	2.2	0.7	3.0
CARL							
CARL-Indicator	0.9	2.2	4.5	3.7	1.0	0.0	2.0
CARL-Absolute	1.3	3.6	4.7	3.2	1.8	0.8	2.5
CARL-GARCH	**2.4**	3.6	5.3	5.8	3.0	**1.3**	3.6
CARL-GARCH with Δw_{t-1}	2.3	**3.9**	6.6	6.5	3.0	1.2	3.9
Multi-threshold CARML							
6 thresholds	1.8	3.4	**7.0**	**7.5**	**3.4**	**1.3**	**4.0**
3 thresholds (Q_i<0 separate from Q_i>0)	2.0	**3.7**	**7.2**	**7.8**	**3.4**	1.2	**4.2**
Multi-step-ahead CARML							
Model for 1 & 2 hours ahead	**2.6**	**4.0**	**7.1**	**7.3**	**3.2**	**1.3**	**4.2**

Notes. Higher values are better. Bold indicates best three methods in each column.

Almost all of the values in Tables 9 and 10 are positive indicating outperformance of the simple benchmark. Comparing the four different CARL models, the final summary column in both tables shows that the two forms of CARL-GARCH were more accurate than CARL-Indicator, CARL-Absolute, and the EWMA, GARCH and CAViaR benchmarks. Of the four CARL models, CARL-GARCH with Δw_{t-1} delivered the greatest accuracy, and it is this that led us to use this form of model within the CARML models. The results are encouraging for the multi-threshold CARML models, with the best results produced by estimating the three negative thresholds in one model, and estimating the three positive thresholds in a separate model. The results of Table 9 for the multi-location CARML models are mixed, with two of the three not being particularly competitive, while the joint model for Aeolos and Plastika performed well. The other two wind farms were notably closer to Aeolos than Plastika, and this may explain why, for modelling power generation at Aeolos, those wind farms did not provide useful information in addition to that provided by the Aeolos wind power series. As we explained in Section 3.3, we inspected wind direction data at the four locations, but there was no clear evidence that any one was generally downwind from another.

Tables 9 and 10 show that the multi-step-ahead CARML model performed very well for one step-ahead prediction. This model was estimated for a single threshold, and so it is interesting to see that it outperformed the CARL models estimated for a single threshold and single lead time. It seems that the information for two step-ahead estimation, captured by the multi-step-ahead model, enhanced the modelling for one step-ahead estimation. Table 11 reports the Brier skill score for two step-ahead estimation for Aeolos. The multi-step-ahead CARML model clearly outperformed the EWMA, GARCH and simple benchmarks. Table 12 shows that this was also the case when averaging over the results for the four wind farms.

Table 11. For two hour-ahead ramp event probability forecasts, Brier skill scores for Aeolos.

	Threshold						
	-0.3	-0.2	-0.1	0.1	0.2	0.3	Geometric mean
Benchmarks							
EWMA for variance	-1.9	-0.1	2.0	2.6	0.4	-2.6	0.0
GARCH for variance	-6.8	-2.7	1.4	3.2	-0.1	-5.0	-1.7
Multi-step-ahead CARML							
Model for 1 & 2 hours ahead	1.4	3.0	6.3	5.4	3.2	2.3	3.6

Notes. Higher values are better.

Table 12. For two hour-ahead ramp event probability forecasts, Brier skill scores averaged across the four wind farms.

	Threshold						
	-0.3	-0.2	-0.1	0.1	0.2	0.3	Geometric mean
Benchmarks							
EWMA for variance	-1.6	-0.8	0.8	1.6	-0.1	-1.9	-0.4
GARCH for variance	-6.0	-2.6	2.1	2.0	-1.7	-6.2	-2.1
Multi-step-ahead CARML							
Model for 1 & 2 hours ahead	2.3	3.8	7.0	4.7	2.4	1.4	3.6

Notes. Higher values are better.

5. Concluding Comments

This paper has presented a set of autoregressive logit models for the short-term probabilistic forecasting of a ramp event, which we defined as the occurrence of a change in hourly wind power exceeding a given threshold. For a single threshold, we adapted CARL models previously proposed by Taylor and Yu (2016) for daily financial data. We introduced a new multi-threshold CARML model for simultaneously estimating ramp event probabilities

for more than one threshold. This model has a multinomial logit formulation, and is estimated using a categorical distribution. It has the appeal of avoiding probability crossing. We found it to be more accurate than the CARL models. We also introduced a CARML formulation for modelling the probability of a ramp event at two wind farm locations. The results were mixed, which was perhaps due to estimation being challenging for this model. The model has a multinomial logit structure, and is estimated using a bivariate Bernoulli distribution. We employed a similar formulation, and the same estimation approach, in a new CARML model for jointly predicting one and two steps-ahead. The one step-ahead results were very competitive with the other models, and the two step-ahead results were also very promising.

In a recent review of the literature on ramp event forecasting, Gallego-Castillo et al. (2015) describe how this is a rapidly growing area of research. Further work with the CARL and CARML models could involve diurnal and annual seasonal factors, predictions of wind speed and direction, probit models, forecast combinations (see, for example, Lessmann et al., 2012), or higher frequency data. For example, the forecasting of the probability of a wind power ramp could perhaps be improved by knowledge of the occurrence of a ramp in wind speed at a downwind location. The models could be adapted for other definitions of a ramp, such as a large change between periods that are not successive. Drawing on the work of Yoder et al. (2014), the models could be adapted for predicting the probability of wind power increasing, decreasing or remaining the same.

Acknowledgements

We are grateful to two referees for providing useful comments on the paper. We would like to thank George Sideratos of the National Technical University of Athens and the EU SafeWind Project for providing the data. We are also grateful for the insightful comments of James Mitchell and participants at the workshop on Advances in Economics and Business Forecasting, held by the Economic Modelling and Forecasting Group, Warwick Business School, 2014, as well as participants in the London Mathematical Society workshop on Statistics for Risk Analysis, held at Brunel University, 2015.

Appendix A

The CARL-Indicator model, of Section 3.1, for a negative threshold Q:

$$\Pr\left(\Delta w_t \le Q\right) = \frac{1}{1+\exp\left(-x_t\right)}$$

$$x_t = \begin{cases} h_t & \text{if} \quad w_{t-1} \ge -Q \\ L & \text{otherwise} \end{cases}$$

$$h_t = \alpha_0 + \alpha_1 I\left(\Delta w_{t-1} \le Q\right) + \alpha_2 I\left(\Delta w_{t-1} \ge -Q\right) + \beta_1 h_{t-1}$$

Appendix B

The *CARL-GARCH with* Δw_{t-1} model, of Section 3.1, for a negative threshold Q:

$$\Pr\left(\Delta w_t \le Q\right) = \frac{1}{1+\exp\left(-x_t\right)}$$

$$x_t = \begin{cases} \gamma_0 + \gamma_1 h_t^{-\frac{1}{2}} + \phi_1 \Delta w_{t-1} & \text{if} \quad w_{t-1} \ge -Q \\ L & \text{otherwise} \end{cases}$$

$$h_t = \alpha_0 + \alpha_1 \Delta w_{t-1}^2 + \beta_1 h_{t-1}$$

References

Bossavy, A., Girard, R., & Kariniotakis, G. (2010). Forecasting uncertainty related to ramps of wind power production. *European Wind Energy Conference and Exhibition 2010, EWEC 2010*. Vol. 2, European Wind Energy Association.

Bossavy, A., Girard, R., & Kariniotakis, G. (2013). Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy*, 16, 51-63.

Bossavy, A., Girard, R., & Kariniotakis, G. (2015). An edge model for the evaluation of wind power ramps characterization approaches. *Wind Energy*, 18, 1169-1184.

Cui, M., Ke, D., Sun, Y., Gan, D., Zhang, J., Hodge, B. (2015). Wind power ramp event forecasting using a stochastic scenario generation method. *IEEE Transactions on Sustainable Energy*, 6, 422-433.

Cutler, N., Kay, M., Jacka, K., & Nielsen, T.S. (2007). Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT. *Wind Energy*, 10, 453-470.

Dai, B., Ding, S., & Wahba, G. (2013). Multivariate Bernoulli distribution. *Bernoulli*, 19, 1465-1483.

de Jong, R.M., & Woutersen, T. (2011). Dynamic time series binary choice. *Econometric Theory*, 27, 673-702.

Elberg, C., & Hagspiel, S. (2015). Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241, 260-272.

Engle, R.F., & Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22, 367-381.

Ferreira, C., Gama, J., Matias, L., Botterud, A., & Wang, J. (2010). A survey on wind power ramp forecasting. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

Gallego, C., Costa, A., & Cuerva, Á. (2011). Improving short-term forecasting during ramp events by means of regime-switching artificial neural networks. *Advances in Science and Research*, 6, 55-58.

Gallego, C., Costa, A., Cuerva, Á., Landberg, L., Greaves, B., & Collins, J. (2013). A wavelet-based approach for large wind power ramp characterization. *Wind Energy*, 16, 257-278.

Gallego-Castillo, C., Cuerva-Tejero, A., & Lopez-Garcia, O. (2015). A review on the recent history of wind power ramp forecasting. *Renewable and Sustainable Energy Reviews*, 52, 1148-1157.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society, Series B*, 69, 243-268.

Gneiting, T., Larson, K., Westrick, K., Genton, M. G., & Aldrich, E. (2006). Calibrated probabilistic forecasting at the Stateline Wind Energy Center: the regime-switching space-time method. *Journal of the American Statistical Association*, 101, 968-979.

Hering, A. S., & Genton, M. G. (2010). Powering up with space-time wind forecasting. *Journal of the American Statistical Association*, 105, 92-104.

Jeon, J., & Taylor, J.W. (2012). Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association*, 107, 66-79.

Koenker, R.W. (2005). *Quantile regression*. Cambridge, UK: Cambridge University Press.

Lessmann, S., Sung, M.-C., Johnson, J.E.V., & Ma, T. (2012). A new methodology for generating and combining statistical forecasting models to enhance competitive event prediction. *European Journal of Operational Research*, 218, 163-174.

Pinson, P., & Kariniotakis, G. (2010). Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems*, 25, 1845-1856.

Pinson, P., & Madsen, H. (2009). Ensemble-based probabilistic forecasting at Horns Rev. *Wind Energy*, 12, 137-155.

Potter, C.W., Grimit, E. & Nijssen, B. (2009). Potential benefits of a dedicated probabilistic rapid ramp event forecast tool. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pp. 1-5. IEEE.

Sherry, M. & Rival, D. (2015). Meteorological phenomena associated with wind-power ramps downwind of mountainous terrain. *Journal of Renewable and Sustainable Energy*, 7, 033101.

Sørensen, P., Cutululis, N.A., Vigueras-Rodríguez, A., Jensen, L.E., Hjerrild, J., Donovan, M.H., & Madsen, H. (2007). Power fluctuations from large wind farms. *IEEE Transactions on Power Systems*, 22, 958-965.

Taylor, J.W., & Yu, K. (2016). Forecasting the exceedance probability for financial asset Returns using conditional autoregressive logit models. *Journal of the Royal Statistical Society, Series A*, forthcoming.

Teugels, J.L. (1990). Some representations of the multivariate Bernoulli and binomial distributions. *Journal of Multivariate Analysis*, 32, 256-268.

Wang, S., Yu, D., & Yu, J. (2015). A coordinated dispatching strategy for wind power rapid ramp events in power systems with high wind power penetration. *Electrical Power and Energy Systems*, 64, 986–995.

Wilks, D.S. (2011). *Statistical methods in the atmospheric sciences*. Oxford, UK: Academic Press.

Yoder, M., Hering, A.S., Navidi, W.C., & Larson, K. (2014). Short-term forecasting of categorical changes in wind power with Markov chain models. *Wind Energy*, 17, 1425-1439.

Zack, J.W., Young, S., Nocera, J., Aymami, J., & Vidal, J. (2010). Development and testing of an innovative short-term large wind ramp forecasting system. *Proceedings of the European Wind Energy Conference & Exhibition*, Warsaw, Poland, April, 20-23.

Zheng, H., & Kusiak, A. (2009). Prediction of wind farm power ramp rates: a data-mining approach. *Journal of Solar Energy Engineering*, 131, 031011-1–31011-8.