Forecasting Wind Power Quantiles Using Conditional Kernel Estimation

James W. Taylor*[a]
Saïd Business School, University of Oxford

Jooyoung Jeon[b]
School of Management, University of Bath

Accepted to Renewable Energy

[a] Corresponding author. Saïd Business School, University of Oxford, Park End Street, Oxford, OX1 1HP, UK
Email: james.taylor@sbs.ox.ac.uk

[b] School of Management, University of Bath, Bath, BA2 7AY, UK
Email: j.jeon@bath.ac.uk

Abstract

The efficient management of wind farms and electricity systems benefits greatly from accurate wind power quantile forecasts. For example, when a wind power producer offers power to the market for a future period, the optimal bid is a quantile of the wind power density. An approach based on conditional kernel density (CKD) estimation has previously been used to produce wind power density forecasts. The approach is appealing because: it makes no distributional assumption for wind power; it captures the uncertainty in forecasts of wind velocity; it imposes no assumption for the relationship between wind power and wind velocity; and it allows more weight to be put on more recent observations. In this paper, we adapt this approach. As we do not require an estimate of the entire wind power density, our new proposal is to optimise the CKD-based approach specifically towards estimation of the desired quantile, using the quantile regression objective function. Using data from three European wind farms, we obtained encouraging results for this new approach. We also achieved good results with a previously proposed method of constructing a wind power quantile as the sum of a point forecast and a forecast error quantile estimated using quantile regression.

Keywords: Wind power; Quantiles; Conditional kernel estimation; Quantile regression

1. Introduction

For many countries, the proportion of electricity consumption generated from renewable sources is rapidly increasing, with ambitious targets aimed at reducing carbon emissions. Wind power generation is a prominent feature of this development in sustainable energy. The high variability and low predictability of the wind present a significant challenge for its integration into electricity power systems [1]. The efficient operation of a power system requires an accurate estimate of the uncertainty in the predicted power output from a wind farm. Quantifying this uncertainty is also important for wind farm operators. A common purpose of wind power forecasting is to set the bid for sales of future production that a wind power producer will make to an energy market. Pinson et al. [2] show that if, as is likely to be the case, the unit cost of surplus and shortage wind power production are different, the optimal bid is not the expectation of future production, but it is instead a quantile. It is, therefore, a prediction of the quantile that is needed, and not a point forecast. The forecasting of wind power quantiles is the focus of this paper.

One possible approach to wind power forecasting is to fit a univariate time series model to wind power time series (e.g. [3]). However, this is very challenging due to the bounded and discontinuous nature of wind power time series. It is more straightforward to fit a time series model to wind speed and direction data, converted to Cartesian coordinates to represent wind velocity variables (e.g. [4]). Forecasts of these variables can then be used as the basis for wind power prediction. This is the approach taken by Jeon and Taylor [5] who use conditional kernel density (CKD) estimation to produce a forecast of the wind power probability density function (i.e. a density forecast). Their methodology incorporates (a) wind speed and direction forecast uncertainty, and (b) the stochastic nature of the dependency of wind power on wind speed and direction. We are not aware of other wind power density forecasting methods that aim to capture these two fundamental sources of uncertainty. The method would, therefore, seem to have strong potential. Although the resultant wind power

density forecasts can be used to provide quantile forecasts, it is our assertion in this paper that superior quantile forecasts can be produced by an adaptation of this CKD-based methodology.

Jeon and Taylor [5] optimise method parameters using an objective function that measures density forecast accuracy. In this paper, we replace this by the objective function of quantile regression, and hence calibrate the approach towards estimation of a particular quantile of interest. The result of this is that the parameters are able to differ across the different quantiles. This is appealing, because different quantiles are likely to have different features and dynamics. For example, the left tail of the wind power distribution may evolve at a faster rate than the right tail.

In this paper, we focus on hourly data from three European wind farms, and we forecast wind power quantiles for lead times ranging from 1 hour up to 3 days ahead. Foley et al. [6] describe how such short lead times are important for power system operational planning and electricity trading. We base the estimation on density forecasts for wind speed and direction, produced by a time series model. It is worth noting that these wind speed and direction density forecasts can be replaced by ensemble predictions from an atmospheric model [7,8,9]. We use density forecasts from a time series model, because this approach has appeal in terms of cost, and the forecasts are likely to compare well with ensemble predictions for short lead times. Also, by contrast with ensemble predictions, time series model predictions can be conveniently produced from any forecast origin, for any lead time, and for any wind farm location for which a history of observations is available.

As we have explained, our proposal is to use the quantile regression objective function within a CKD-based approach. It is worth noting that quantile regression has previously been used for wind power quantile prediction. Bremnes [10] proposes forms of locally weighted quantile regression with wind speed and direction as explanatory variables. The adaptive quantile regression procedure of Møller et al. [11], and the linear model with

spline basis functions of Nielsen et al. [12], both involve the application of quantile regression to the errors of wind power point forecasts produced in a separate procedure. As in the study of Taylor and Bunn [13], Nielsen et al. simultaneously estimate quantiles for a range of lead times by using the forecast error from different lead times as dependent variable, and the forecast lead time as one of the explanatory variables. In this paper, we implement a form of this approach, and compare quantile forecast accuracy with our proposed adaptation of the CKD-based approach of Jeon and Taylor [5]. The CKD-based approaches explicitly try to capture the uncertainties underlying wind power, while the quantile regression method is a pragmatic approach. It is an interesting empirical question as to which is more accurate.

Section 2 discusses the features of wind power, speed and direction data from three wind farms. Section 3 reviews CKD-based wind power density forecasting, and Section 4 describes how the method can be adapted for the prediction of a particular quantile. Section 5 provides an empirical evaluation of the accuracy of our proposed CKD-based quantile forecasting approach, and a quantile regression model based on the approach of Nielsen et al. Section 6 provides a brief summary and conclusion.

2. Wind data and the power curve

2.1. The characteristics of wind data

The data used in this paper consists of hourly observations for wind speed, direction and power, recorded at the following three wind farms: Sotavento, which is in Galicia in Spain, and Rokas and Aeolos, which are on the Greek island of Crete. Our data for Sotavento is for the 23,616 hourly periods from 1 July 2004 to 11 March 2007. For Rokas and Aeolos, the data is for the 8,760 hourly observations from the year 2006. The wind power data corresponds to the total power generated from the whole wind farm. On the final day of each

dataset, the capacities of Sotavento, Rokas and Aeolos were 17.6 MW, 16.3 MW and 11.6 MW, respectively. The data from the two Crete wind farms was used in [5].

Fig. 1 presents the wind speed, direction and power time series for the Sotavento wind farm. The series exhibit substantial volatility, which suggests that point forecasting is likely to be very challenging, and this motivates the development of methods for quantile and density forecasting. The plots also suggest that fluctuations in wind power coincide, to some extent, with variations in wind speed and direction. It is interesting to note that the volatility in the series varies over time. It is this that has prompted the use of generalised autoregressive conditional heteroskedastic (GARCH) models for wind speed data (e.g. [4,8,14]). Fitting such models to wind power time series is not appealing, because the power output from a wind farm is bounded above by its capacity, and this creates discontinuities, as well as distributional properties that are non-Gaussian and time-varying.

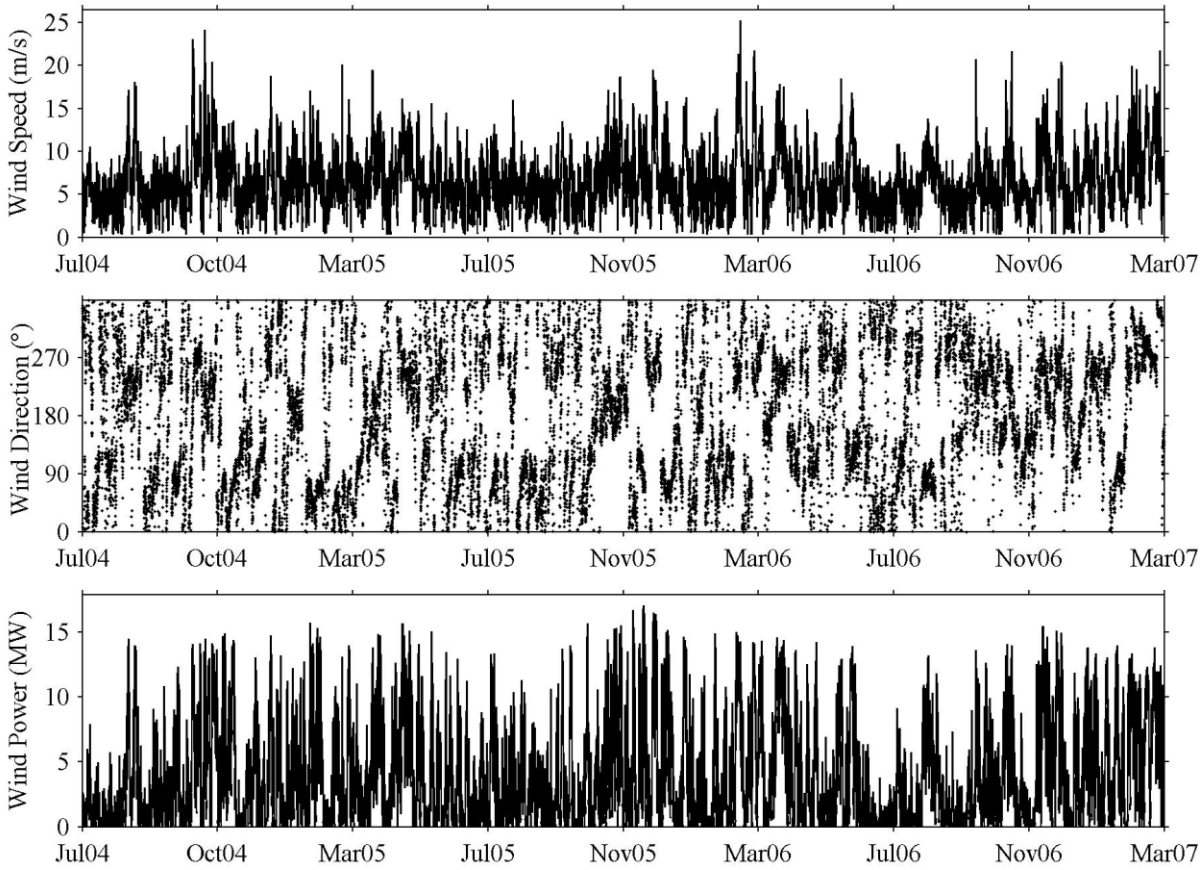

Fig. 1. Wind speed, direction and power time series for Sotavento.

For Sotavento and Rokas, Fig. 2 shows Cartesian plots of wind speed and direction, where the distance of each observation from the origin represents the wind speed. The plot for Rokas shows that north-westerly wind is particularly common at this wind farm.

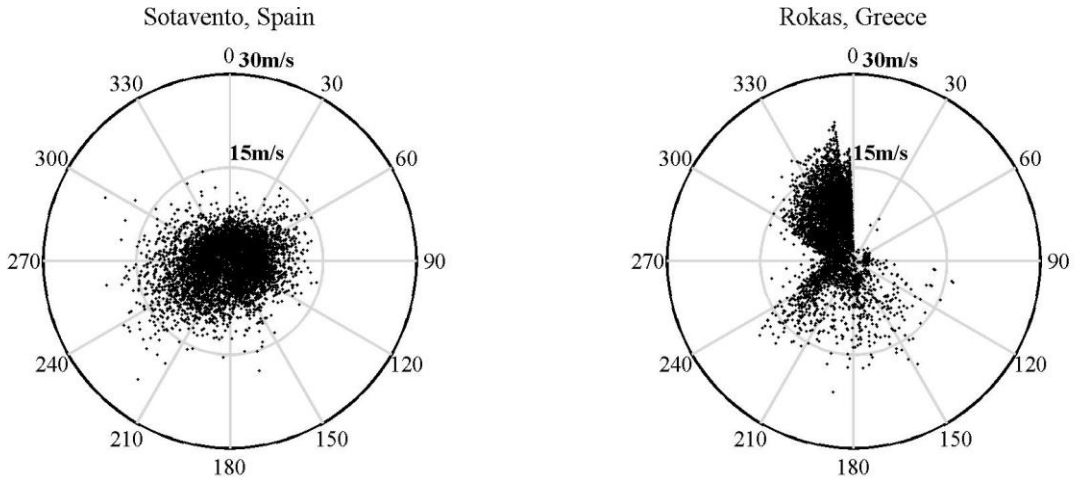

Fig. 2. For Sotavento and Rokas, Cartesian plots of wind speed and direction, where the distance of each observation from the origin is the strength of the wind speed.

2.2. Power curves

The theoretical relationship between the wind power generated and the wind speed is described by the machine power curve, which can be provided by the turbine manufacturer [15]. This curve is deterministic and nonlinear, with the following features: a minimum 'connection speed' below which no power can be generated; as speed rises from this minimum, the power output increases; this continues until a 'nominal speed', which is the lowest speed at which the turbine is producing at its maximum power output; and finally there is a 'disconnection speed' at which the turbine must be shut down to avoid damage.

Fig. 3 plots the empirical power curves, using historical observations, for Sotavento and Rokas. Although the figures show the essential features that we have just described for the machine power curve, it can be seen that, in reality, the power curve for a wind farm is stochastic. Sanchez [15] attributes this to the effect of other atmospheric variables, such as air temperature and pressure, as well as other factors, such as the relationship differing for rising

and falling wind speed, complexities caused by the aggregated effect of different types of turbines in the one wind farm, and the capacity of the wind farm varying over time.

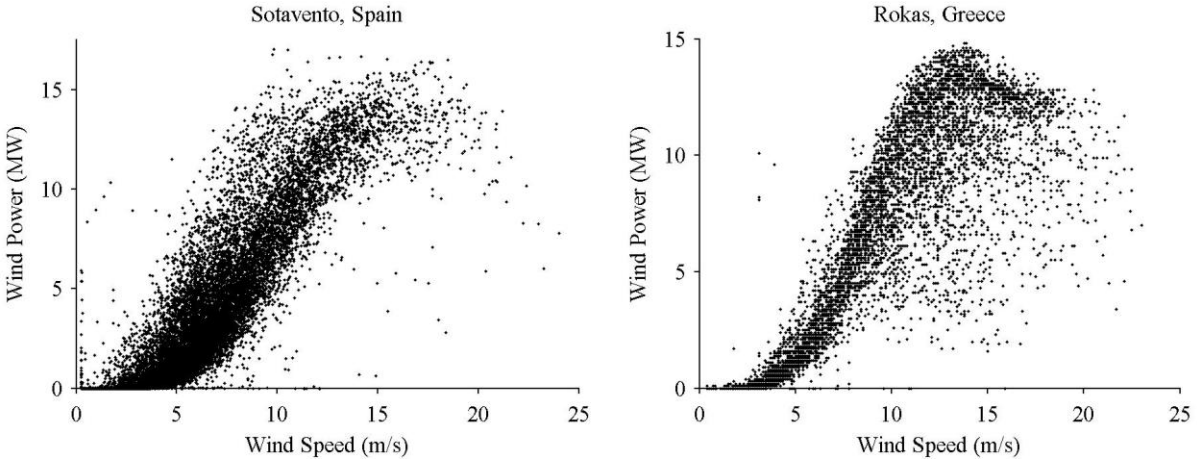

Fig. 3. Empirical power curves for Sotavento and Rokas.

The empirical power curves in Fig. 3 indicate that the dispersion and distributional shape of the variability in wind power depends on the value of wind speed. For example, for Rokas, if wind speed is between about 10 and 15 m/s, the wind power density is skewed to the left with relatively high variability, while for wind speed below about 5 m/s, the wind power density would seem to be skewed to the right with relatively low variability. Therefore, the estimation of the wind power density or quantiles should be conditional on the value of wind speed.

For Sotavento and Rokas, Fig. 4 shows the empirical power curves plotted for two different months. The Sotavento plot shows considerably more variation in November 2005 than in April 2006. Curiously, for the higher values of wind speed, more wind power tended to be generated in November 2005 than April 2006. A similar comment can be made regarding the Rokas empirical power curve, which shows greater efficiency in the conversion of strong values of wind speed to power in January 2006 than September 2006. In essence, the plots suggest that the power curves are time-varying. This can be due to changing weather patterns, and changes in the capacity of the wind farm due, for example, to maintenance or

expansion. Time-variation in the power curve suggests that, when modelling, it may be useful to put more weight on more recent information.

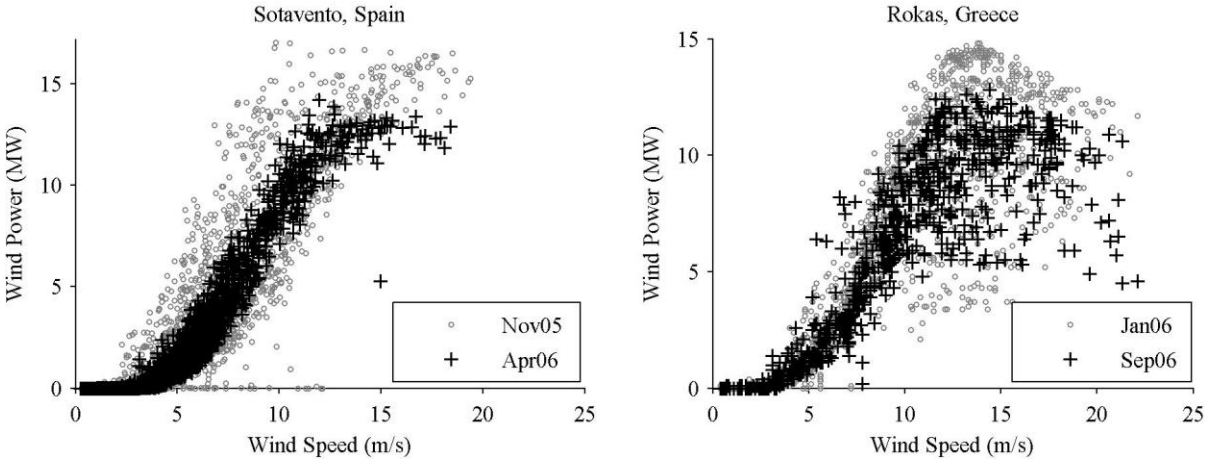

Fig. 4. Empirical power curves for Sotavento and Rokas. Each based on two selected months.

The plots of this section indicate that, when forecasting wind power based on a model relating power to speed, it is important to acknowledge two issues. First, the relationship between wind power and speed is nonlinear and stochastic, and it may be time-varying and dependent on wind direction and other atmospheric variables [15]. Second, the stochastic nature of wind speed will affect the uncertainty in wind power predictions [16], and so should be accommodated in the modelling approach. In the next section, we present a methodology for wind power forecasting that addresses the first of these issues through the use of a nonparametric approach that makes no distributional assumption for wind power, imposes no parametric assumption for the relationship between wind power and speed, and puts more weight on more recent observations. The methodology addresses the second issue by incorporating Monte Carlo sampling from wind velocity density forecasts. These density forecasts could be produced from a time series model or from weather ensemble predictions from an atmospheric model.

3. Conditional kernel estimation for wind power density forecasting

3.1. Conditional kernel density estimation

Kernel density estimation is a nonparametric approach to the estimation of the density of a target variable Y_t. It can be viewed as smoothing the empirical distribution of historical observations. The unconditional kernel density (UKD) estimator (see [17]) is expressed as:

$$\hat{f}(y) = \sum_{t=1}^{n} K_{h_y}(Y_t - y), \tag{1}$$

where n is the sample size, and $K_h(\bullet)=K(\bullet/h)/h$ is a kernel function with bandwidth h. The kernel function is a function that integrates to 1. A common choice is the standard Gaussian probability density function, and we use this for all kernel functions in this paper. The bandwidth is a parameter that controls the degree of smoothing.

In its simplest form, conditional kernel density (CKD) estimation enables the nonparametric estimation of the density of a target variable Y_t, conditional on the value of an explanatory variable X_t. It is nonparametric in two senses: it requires no parametric assumptions for either the distribution of Y_t or the form of the functional relationship between Y_t and X_t. These features make the method particularly attractive for the wind power context, because the wind power distribution is non-Gaussian and unknown, and the form of the relationship between wind power and speed is nonlinear and unknown. The CKD estimator of the conditional density function of Y_t, given $X_t = x$ (see [18]), is expressed as:

$$\hat{f}(y \mid x) = \frac{\sum_{t=1}^{n} K_{h_x}(X_t - x)\, K_{h_y}(Y_t - y)}{\sum_{t=1}^{n} K_{h_x}(X_t - x)}.$$

The kernel K_{h_y} enables kernel density estimation in the y-axis direction, with the observations weighted in accordance to the kernel K_{h_x}, which relates to kernel smoothing in the x-axis direction, enabling a larger weight to be put on historical observations for which X_t is closer to x. For the two kernels, the bandwidths, h_x and h_y, control the degree of smoothing.

3.2. Conditional kernel density estimation for wind power density forecasting

With wind power specified as the target variable Y_t, Jeon and Taylor [5] use CKD estimation with conditioning on wind velocity variables, U_t and V_t, which are the result of transforming wind speed and direction to Cartesian coordinates. They incorporate a decay parameter λ to enable more weight to be put on more recent observations. This is appealing because the shape of the wind power density, and its relationship to wind velocity, can vary over time. We noted this in Section 2.2, in relation to the two plots of Fig. 4, which each show the empirical power curve differing for two separate months of the year. A lower value of λ leads to faster decay. The CKD estimator is presented in the following expression:

$$\hat{f}(y \mid u,v) = \frac{\sum_{t=1}^{n} \lambda^{n-t} K_{h_{uv}}(U_t - u) K_{h_{uv}}(V_t - v) K_{h_y}(Y_t - y)}{\sum_{t=1}^{n} \lambda^{n-t} K_{h_{uv}}(U_t - u) K_{h_{uv}}(V_t - v)} \tag{2}$$

Cross-validation can be used to optimise λ along with the bandwidths h_{uv} and h_y, and we discuss this issue further in the next section. The exponential decay can be viewed as a kernel function, defined to be one-sided with exponentially declining weight [19]. We can, therefore, view the CKD estimator of expression (2) as having three bandwidths, λ, h_{uv} and h_y. The CKD estimator provides an estimate of the density at $Y_t=y$. To estimate the full density, the CKD estimation can be performed for values of y from zero to the wind farm's capacity with small increments. In our implementations of CKD in this paper, we used increments equal to 1% of the capacity, and assumed equal probability within each of the corresponding 100 wind power intervals to deliver an estimate of the full density.

To produce a wind power density forecast, it seems natural to perform the CKD estimation conditional on forecasts of U_t and V_t. This is essentially the approach taken by Juban et al. [20], who condition on point forecasts of wind speed and direction from an atmospheric model. The problem with conditioning on point forecasts is that the resulting wind power density estimate will not capture the potentially significant uncertainty in U_t and

V_t. To address this, Jeon and Taylor [5] provide the following three-stage methodology that effectively enables CKD estimation to be performed conditional on density forecasts for U_t and V_t, which they produced using a time series model:

Stage 1 - The CKD estimator of expression (2) is used to produce an estimate of the full wind power density conditional on each pair of values of U_t and V_t, on a grid from -30 m/s to 30 m/s with an increment of 0.5 m/s. The result is 121×121=14,641 pairs, and, for each, a corresponding conditional wind power density estimate. These are stored for use in Stage 2.

Stage 2 - Monte Carlo simulation of a time series model is performed to deliver 1,000 realisations of pairs of values for U_t and V_t, for a selected lead time. Each value is rounded to the nearest 0.5 m/s, and then for each of the 1,000 pairs, the corresponding conditional wind power density estimate is obtained from those stored in Stage 1.

Stage 3 - The 1,000 wind power density estimates from Stage 2 are averaged to give a single wind power density forecast.

It is worth noting that the methodology relies on density forecasts for the wind velocities, U_t and V_t, and that these could be produced by a time series model or atmospheric model, which would be expected to capture the autocorrelation properties of the wind.

3.3. Optimising conditional kernel density estimation for wind power density forecasting

Fan and Yim [21] and Hall et al. [22] provide support for the use of cross-validation to optimise the bandwidths in kernel density estimation. In our implementation of kernel density estimation in this paper, we followed Jeon and Taylor [5] by using a rolling window of 6 months to produce density estimates, and by selecting the values of λ, h_{uv} and h_y that led to the most accurate wind power density estimates calculated over a cross-validation evaluation period for 1 hour-ahead prediction. They measured accuracy using the mean of the continuous ranked probability score (CRPS), which is described by Gneiting et al. [23] as an

appealing measure of accuracy, capturing the properties of calibration and sharpness in the estimate of the probability density function.

As we explained in the previous section, in our implementation of the kernel density methods, we estimated the density for values of wind power at increments equal to 1% of the capacity, and assumed equal probability within each of the 100 wind power intervals. As this delivers a discrete density and distribution, we evaluated accuracy using the RPS, which is the discrete version of the CRPS (see [24]). We used a three-step cascaded optimisation approach to find the parameter values that minimise the RPS for the cross-validation period. The first step involved a grid search of 100 values for λ, h_{uv} and h_y, log-equally spaced between the following intervals: $0.98 \leq \lambda \leq 1$; $0.0001 \leq h_{uv} \leq 5$; and $0.001 \leq h_y \leq 0.5$. With regard to the interval for h_y, note that, instead of working with wind power measured in MW, we used the capacity factor, which is wind power as a proportion of the wind farm's capacity. The second step of the cascaded optimisation approach used a trust-region-reflective algorithm employed in the 'fmincon' function of Matlab 2012b. The trust-region-reflect algorithm described by [35] further minimised the RPS for the cross-validation period subject to the bound constraints given above. The algorithm uses finite difference approximation and trust regions to ensure the robustness of the iteration. A genetic algorithm was chosen as the final step of the cascaded optimisation. The best individuals from the previous optimisations were used as the population required to operate the genetic algorithm. We did not choose the genetic algorithm in the earlier stage of our cascading optimisation for global optimisation because deceptive problems, which mislead the search to some local optima rather than the global optimum, have been recognised as challenges in genetic algorithms [36]. Increasing the mutation rate or maintaining a diverse population might help to relieve this issue, but in the expense of an exponential increase of the search space. We use the notation CKDλ to refer to the three-stage CKD-based approach of Section 3.2, optimised using the RPS.

4. Conditional kernel estimation for wind power quantile forecasting

In this section, we introduce our proposed approach to wind power quantile forecasting. It is a relatively simple adaptation of the CKD approach of the previous section.

4.1. A limitation of the CKD approach for wind power quantile forecasting

Although CKDλ can certainly be used to deliver quantile forecasts, we would suggest that this has the disadvantage that CKDλ involves the use of the same parameters across different wind power quantiles. With regard to the bandwidth in the wind power direction, h_y, one might imagine that a larger value would be needed for more extreme quantiles, because there are fewer observations in the tails of the density. With regard to the bandwidth in the wind velocity directions, h_{uv}, it seems likely that the optimal value will depend on the value of h_y, as well as the characteristics of the empirical power curve around the quantile under consideration. For example, if that part of the empirical power curve has a relatively high gradient, then a relatively small value of h_{uv} may be needed to avoid over-smoothing. As for the decay parameter, λ, it seems reasonable to assume that different parts of the wind power density will evolve at different rates, and also that the conditionality on the wind velocities may evolve differently for different quantiles. Hence, different values of λ are likely to be optimal for different quantiles. Therefore, the assumption of using the same parameters for different quantiles would seem to hamper accurate quantile estimation.

4.2. Optimising conditional kernel density estimation for wind power quantile forecasting

In this paper, we use the three-stage CKD-based approach, described in Section 3.2, to deliver a wind power density forecast, which we convert into a cumulative distribution function from which we obtain the required θ quantile estimate. However, as our interest is not in the accurate estimation of the entire wind power density, we optimise the approach

specifically towards estimation of the desired θ quantile of interest. More specifically, in the cross-validation approach used to optimise λ, h_{uv} and h_y, we replace the RPS with the following measure, which is the objective function minimised in quantile regression [25]:

$$\frac{1}{n}\sum_{t=1}^{n}\left(Y_t-\hat{Q}_{yt}\right)\left(\theta-I\left(Y_t\le\hat{Q}_{yt}\right)\right) \tag{3}$$

where $\hat{Q}_{yt}$ is an estimate of the θ quantile of a variable Y_t. We refer to expression (3) as the mean quantile regression error (MQRE). It has been proposed as a measure of quantile forecast accuracy, both in the context of wind power [11,26] and in other applications [27,28,29]. We discuss this further in Section 5.3.

Our proposal is, therefore, to produce wind power quantile forecasts using the three-stage CKD-based approach of Section 3.2, with values of λ, h_{uv} and h_y selected to deliver the most accurate quantile estimates, where accuracy is measured using the MQRE, calculated over a cross-validation evaluation period for 1 hour-ahead prediction. We refer to this method as CKQλ. In our empirical work, to minimise the MQRE for the cross-validation period, we used the three-step cascaded optimisation approach that we described in Section 3.3.

5. Empirical study

In this section, we use the hourly data from the three wind farms, described in Section 2, to evaluate forecast accuracy for the 1%, 5%, 25%, 75%, 95%, and 99% conditional quantiles for lead times from 1 to 72 hours ahead. For each wind farm, we used the final 25% of data for post-sample evaluation, and the penultimate 25% for cross-validation.

5.1. Kernel density methods for quantile forecasting

In addition to the CKQλ method, described in Section 4.2, we also implemented, as a sophisticated benchmark, the CKDλ method, described in Sections 3.2 and 3.3. These two

methods differ in that CKDλ uses the RPS as the basis for estimating the parameters, λ, h_{uv} and h_y, while CKQλ uses the quantile regression cost function of expression (3) for estimation. For a given lead time, each of these methods delivers a wind power density forecast, from which the required θ quantile forecast is obtained. We used a 6-month moving window in the CKD estimation, with CKD estimation performed afresh every 24 hours.

Density forecasts of the wind velocity variables, U_t and V_t, were produced using a time series model of the form used by Jeon and Taylor [5], with parameters estimated using the first 75% of the data. This is a bivariate model with vector autoregressive moving average components for the levels, and GARCH components for the variances. Interesting alternative time series models for wind speed and direction include the multivariate kernel density estimation approach of Zhang et al. [30], and the Bayesian approach of Jiang et al. [31].

As a relatively simple benchmark method, we applied the unconditional kernel density (UKD) estimator of expression (1) to a moving window of the most recent historical wind power observations. We optimised the one bandwidth using cross-validation. The resulting density estimate provided quantile estimates that we used as the wind power quantile forecasts for all future periods. We considered moving windows of lengths 24 hours, 10 days and 6 months. The best results were produced with moving windows of 24 hours, and so for simplicity we report only these results in the remainder of this paper. We refer to this method as UKD24.

Table 1
Parameters optimised using cross-validation for Sotavento.

Method	Bandwidth h_{uv} (m/s)	Bandwidth h_y	λ (half-life)
UKD24		0.267	
CKDλ	0.56	0.021	0.999 (28.9 days)
CKQλ-1%	2.55	0.012	0.990 (2.9 days)
CKQλ-5%	0.87	0.015	0.999 (28.9 days)
CKQλ-25%	0.40	0.013	0.999 (28.9 days)
CKQλ-50%	0.50	0.021	0.999 (28.9 days)

CKQλ-75%	0.58	0.010	0.999 (28.9 days)
CKQλ-95%	0.53	0.065	0.999 (28.9 days)
CKQλ-99%	0.46	0.090	0.999 (28.9 days)

NOTE: h_y has no units, because y is the capacity factor.

Table 1 presents the parameters optimised for the three methods using cross-validation, and averaged over the three wind farms. Note that the bandwidth in the y-direction, h_y, has no units because, as we stated in Section 3.3, in our computations, we worked with capacity factor, which is wind power as a proportion of the capacity of the wind farm. In Table 1, each value of the decay parameter λ is accompanied by the corresponding half-life, and these indicate that, although the values of λ may seem rather high, they do imply notable decreasing weight over the 6-month rolling window of hourly observations used for CKDλ and CKQλ. For the CKQλ method, it is interesting to note that the bandwidth in the y-direction, h_y, is larger for more extreme quantiles in the upper tail of the density. This bandwidth relates to kernel density estimation for the wind power density. The need for larger values of h_y for more extreme upper quantiles seems intuitive, because there are fewer observations in the upper tail of the wind power distribution, and hence more kernel smoothing is beneficial. With regard to the values of h_{uv} for CKQλ, it is interesting to note that the values for the 1% quantile and 5% quantile, are notably larger than for the other quantiles. This implies a relatively large degree of smoothing of the empirical power curve, and this seems reasonable as the curve is relatively flat for low values of wind speed.

In Fig. 5, we present the wind power observations and the 6 hour-ahead forecasts for the 5% and 95% quantiles from the CKQλ method for the final 4 weeks of the post-sample period for Sotavento. It is reassuring to see that the quantile forecasts move with the wind power time series. However, the purpose of Fig. 5 is to provide just an informal visual check on the method. A more thorough assessment of quantile forecast accuracy is provided in Section 5.3.

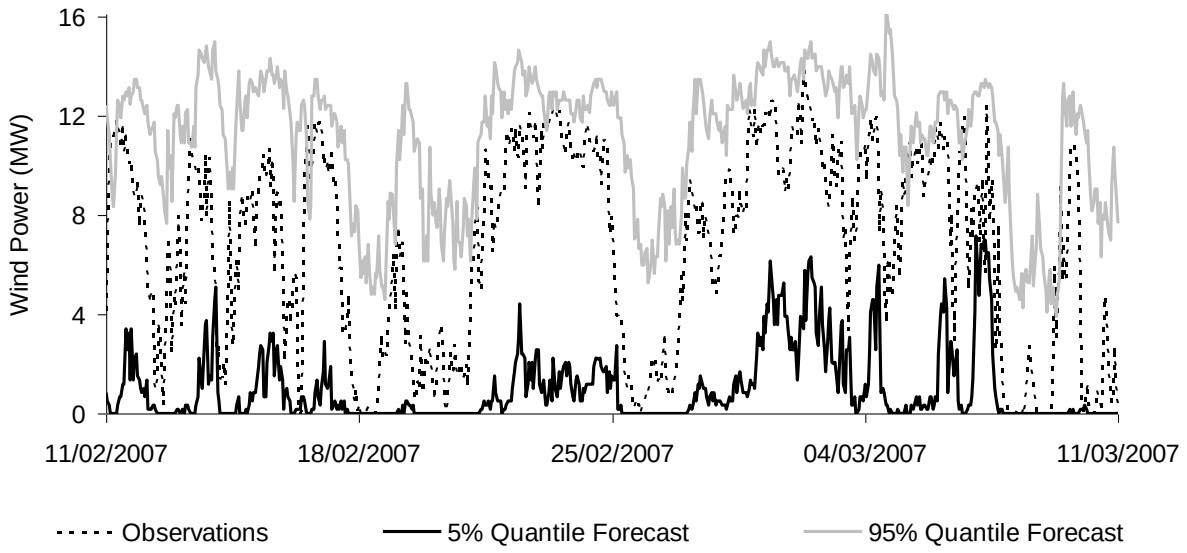

Fig. 5. Time series plots of 6 hour-ahead quantile forecasts from CKQλ for the final 4 weeks of the post-sample period for Sotavento.

5.2. A quantile regression method for quantile forecasting

In addition to the methods, described in the previous section, we also generated quantile forecasts from a quantile regression model. The approach that we used was based on the work of Nielsen et al. [12]. This involves first producing point forecasts, and then using quantile regression to estimate quantile models for the forecast error distribution. For simplicity, as point forecasts, we used the median of the density forecasts of the UKD24 method, which we described in the previous section.

Following the approach taken by Nielsen et al., we chose the quantile regression dependent variable to be a vector constructed by concatenating vectors of (n-72) in-sample forecast errors for each of the 72 lead times of interest, where n is the number of in-sample periods. We included an intercept (C) in the quantile regression, and the following explanatory variables: the lead time (L); the square of the lead time (L^2); the value of the wind power capacity factor at the forecast origin (P); the value of the capacity factor at the forecast origin multiplied by the lead time ($P \times L$); the value of the capacity factor at the forecast origin multiplied by the square of the lead time ($P \times L^2$); the value of wind speed at the forecast origin (S); the value of wind speed at the forecast origin multiplied by the lead

time ($S \times L$); the value of wind speed at the forecast origin multiplied by the square of the lead time ($S \times L^2$); and the point forecast for wind power ($\hat{P}$).

We performed the quantile regression for each of the six quantile probability levels (1%, 5%, 25%, 75%, 95%, and 99%), and for each of the three wind farms. Each wind power quantile forecast was then produced as the sum of the point forecast and the forecast error quantile estimated by a quantile regression model. In Table 2, we provide the parameters estimated for the 5% and 95% quantile regression models for Sotavento. Given that the lead time variable L takes values up to 72, the coefficients of L^2, $P \times L$ and $P \times L^2$ are sufficiently large to imply that the wind power uncertainty is nonlinearly dependent on the lead time and wind power capacity factor at the forecast origin.

Table 2
Parameters of the 5% and 95% quantile regression models for Sotavento.

	C	L	L^2	P	$P \times L$	$P \times L^2$	S	$S \times L$	$S \times L^2$	$\hat{P}$
5%	0.0254	-0.054	-0.00072	1.56	0.0024	-0.00074	-0.0176	0.00200	0.000045	-7.56
95%	0.0161	0.026	-0.00025	1.79	-0.0477	0.00026	-0.0180	0.00008	0.000006	-0.92

5.3. Comparison of post-sample quantile forecast accuracy

To assess the post-sample performance of the wind power quantile forecasting methods, we used the hit percentage and the MQRE of expression (3). In the context of probabilistic wind power forecasting, Pinson et al. [26] describe how quantile forecasts should be assessed in terms of *reliability* and *sharpness*. Reliability is the degree to which the quantile forecast is, on average, correct. Sharpness is the extent to which the quantile forecast varies with the quantile over time. The hit percentage, that we consider here, is a standard measure of reliability (see, e.g., [26] and [32]). Pinson et al. [26] explain that, having assessed reliability, sharpness can be evaluated through the use of the MQRE, which is an overall skill score, measuring both reliability and sharpness.

The hit percentage is the percentage of the post-sample wind power observations that fall below the corresponding quantile forecasts. For estimation of the θ quantile, the ideal value for the hit percentage is θ. For each method and forecast lead time, we calculated the weighted average of the hit percentage across the three wind farms, where the weights were in proportion to the capacities of the wind farms. We present this average hit percentage in Table 3. For clarity of presentation, in Table 3, we group some of the forecast horizons together, with more detailed results shown for the early lead times, as we feel all of the methods have greatest potential for shorter lead times, as they are based in this paper on time series models, rather than on predictions from an atmospheric model. The final column of the table provides the average performance across all lead times. Table 3 shows the simple benchmark method, UKD24, performing relatively poorly, except for estimation of the 75% quantiles. Looking at the final column of Table 3, we see that, overall, CKQλ performed the best for four of the six quantile probability levels, and was poorer than CKDλ for just the 25% probability level. The quantile regression method was relatively poor for the lower three probability levels, but the best overall for estimation of the 95% quantiles, and competitive for estimation of the 75% and 99% quantiles.

The hit percentage is a measure of the unconditional coverage of a quantile estimator. It assesses the average number of times that an observation falls below the estimator. To also assess the degree to which each quantile estimator varies with the wind power series, tests have been proposed for conditional coverage (e.g. [33]). These tests focus on the level of autocorrelation in the series of hits. Unfortunately, these tests are not of use for multi-step-ahead prediction, because the hit variable will naturally tend to be autocorrelated, regardless of the quality of the quantile forecasts [26]. To assess both conditional and unconditional coverage, we use the MQRE, presented in expression (3). As we discussed at the start of this section, the MQRE can also be viewed as an overall skill score measuring both reliability and sharpness. Its use for evaluating quantile forecasts is natural, in view of the common use of

the mean squared error (MSE) for evaluating point forecasts. Table 4 presents the weighted average of the MQRE across the three wind farms, where the weighting was in proportion to the capacities of the wind farms. In this table, the results for the UKD24 method are not competitive for any of the quantiles. The results for the two conditional kernel methods are the same for the lower three probability levels. For the upper quantiles, CKQλ was more accurate than CKDλ, but the results are quite similar. For the 75% probability level, the results for the quantile regression approach are notably the best, while for the other five probability levels, the results for this approach are similar to those for the two conditional kernel methods.

Table 5 investigates how the relative performances of the methods differ across the three wind farms. For each wind farm, the table presents each of the two measures, averaged across the 72 lead times, for each method. The results are reasonably consistent across the three wind farms. An exception to this is that the CKDλ method performed relatively poorly for Aeolos. Another exception is that the UKD24 benchmark method was relatively accurate for the Sotavento wind farm for 95% and 99% quantile estimation.

Table 3
Evaluation of post-sample quantile forecasts using the hit percentage measure of reliability, averaged over the three wind farms with weights in proportion to their capacities.

Horizon (hours)	1	2	3-4	5-6	7-8	9-12	13-24	25-48	49-60	61-72	1-72
1%											
UKD24	22.4	22.5	22.4	22.6	22.6	22.8	22.9	23.6	23.9	24.1	23.5
QuReg	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CKDλ	0.5	0.2	0.1	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0
CKQλ	0.3	0.2	0.1	0.0	0.0	0.1	0.2	0.2	0.0	0.0	0.1
5%											
UKD24	28.1	28.2	28.2	28.5	28.7	28.8	28.8	29.7	30.2	30.8	29.6
QuReg	0.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CKDλ	1.6	1.3	1.3	1.4	1.5	1.6	1.9	2.6	2.9	3.0	2.4
CKQλ	1.5	1.3	1.2	1.3	1.6	1.6	2.1	3.4	4.1	4.6	3.2
25%											
UKD24	53.8	53.8	53.8	54.0	54.2	54.0	54.1	54.9	54.4	53.6	54.3
QuReg	25.6	19.8	15.3	11.9	9.2	5.9	1.7	1.0	1.1	1.9	3.1
CKDλ	18.1	19.4	20.5	21.6	22.6	23.6	23.4	22.1	21.2	20.3	21.8
CKQλ	22.1	23.6	24.9	26.1	26.7	28.3	30.4	30.2	30.3	30.2	29.6
50%											
UKD24	68.9	69.0	68.8	68.8	68.6	68.5	68.2	67.9	67.5	66.6	67.8
QuReg	44.3	46.5	47.9	49.5	51.1	52.6	55.3	60.6	62.8	63.5	58.8
CKDλ	53.2	53.0	51.6	50.4	49.7	48.5	46.9	45.8	44.6	44.3	46.3
CKQλ	50.4	49.7	48.4	47.3	46.7	45.8	45.0	44.5	43.9	43.6	44.8
75%											
UKD24	82.8	82.4	82.3	81.9	81.5	81.3	80.8	79.3	78.5	77.4	79.5
QuReg	81.4	78.3	76.2	75.0	74.9	75.3	77.0	78.4	78.9	78.4	77.9
CKDλ	79.5	77.4	75.9	74.5	73.8	73.0	70.8	68.1	66.3	65.5	68.9
CKQλ	79.5	77.7	76.4	75.3	75.0	74.8	73.6	71.9	71.0	70.6	72.5
95%											
UKD24	92.6	92.2	92.0	91.9	91.6	91.5	91.4	90.5	89.6	89.0	90.5
QuReg	87.8	90.0	90.3	90.2	90.5	91.1	93.7	97.3	97.5	95.0	95.2
CKDλ	96.6	96.0	95.8	95.3	94.6	94.2	93.4	92.1	91.1	90.6	92.4
CKQλ	97.7	97.5	97.2	97.1	96.8	96.6	96.2	95.4	94.7	94.5	95.5
99%											
UKD24	96.5	96.4	96.2	96.0	95.8	95.9	95.8	95.3	94.7	94.4	95.2
QuReg	96.4	96.2	95.9	96.3	96.4	96.6	97.8	99.3	99.7	98.8	98.5
CKDλ	99.0	99.0	99.1	99.2	99.0	98.9	98.7	98.0	97.6	97.5	98.1
CKQλ	99.4	99.3	99.4	99.5	99.4	99.3	99.2	98.9	98.9	98.9	99.0

NOTE: For the θ quantile, the ideal value is θ. The best performing method at each horizon is underlined.

Table 4
Evaluation of post-sample quantile forecasts using the MQRE (×1,000) skill score, averaged over the three wind farms with weights in proportion to their capacities.

Horizon (hours)	1	2	3-4	5-6	7-8	9-12	13-24	25-48	49-60	61-72	1-72
1%											
UKD24	8	8	8	9	9	9	10	12	13	14	12
QuReg	3	3	3	3	3	3	3	3	3	3	3
CKDλ	3	3	3	3	3	3	3	3	3	3	3
CKQλ	3	3	3	3	3	3	3	3	3	3	3
5%											
UKD24	28	28	28	29	30	31	32	35	38	39	35
QuReg	13	13	13	13	13	13	13	13	13	13	13
CKDλ	10	11	12	12	12	13	13	13	13	14	13
CKQλ	11	11	12	12	13	13	13	14	14	14	13
25%											
UKD24	92	93	94	95	96	98	99	104	108	111	104
QuReg	23	32	41	49	55	60	65	67	69	71	64
CKDλ	35	39	43	47	51	55	61	66	67	67	63
CKQλ	34	38	42	47	50	55	61	67	68	68	63
50%											
UKD24	119	120	122	124	125	126	127	133	137	140	132
QuReg	65	70	75	81	86	90	100	115	129	136	113
CKDλ	44	51	60	68	75	83	97	113	120	121	105
CKQλ	40	46	52	59	65	71	83	97	104	106	91
75%											
UKD24	100	100	102	103	104	105	106	110	114	116	110
QuReg	28	38	48	58	65	72	82	98	104	104	91
CKDλ	36	44	51	60	67	76	92	111	122	124	104
CKQλ	36	44	51	59	65	74	90	109	119	121	101
95%											
UKD24	36	37	37	38	39	39	39	40	42	43	40
QuReg	13	16	20	22	25	26	27	29	30	29	28
CKDλ	14	16	18	21	23	26	29	34	37	39	33
CKQλ	15	17	19	21	22	24	27	29	30	30	28
99%											
UKD24	11	11	12	12	12	12	12	13	14	14	13
QuReg	6	6	7	7	8	8	7	7	7	8	7
CKDλ	4	5	5	5	6	6	7	7	7	7	7
CKQλ	4	5	5	5	6	6	6	7	7	7	6

NOTE: Smaller values are better. The best performing method at each horizon is underlined.

Table 5
Evaluation of post-sample quantile forecasts using the hit percentage reliability measure and the MQRE (×1,000) skill score. Values shown are averages across the 72 lead times. The weighted averages use weights in proportion to the capacities of the wind farms.

	Hit percentage				MQRE (×1,000)			
	Aeolos	Rokas	Sotavento	Wtd. Avg.	Aeolos	Rokas	Sotavento	Wtd. Avg.
1%								
UKD24	29.6	22.1	18.7	23.5	17	13	5	12
QuReg	0.0	0.0	0.0	0.0	3	3	2	3
CKDλ	0.0	0.1	0.0	0.0	3	3	2	3
CKQλ	0.0	0.3	0.0	0.1	3	3	2	3
5%								
UKD24	35.4	28.9	24.7	29.6	47	39	18	35
QuReg	0.0	0.0	0.0	0.0	14	14	11	13
CKDλ	0.1	7.0	0.2	2.4	14	14	11	13
CKQλ	0.1	9.1	0.2	3.2	14	14	11	13
25%								
UKD24	57.2	52.3	53.5	54.3	129	106	76	104
QuReg	2.1	5.6	1.6	3.1	69	70	54	64
CKDλ	7.1	36.4	21.7	21.8	68	67	52	63
CKQλ	30.1	36.7	22.0	29.6	69	67	52	63
50%								
UKD24	67.1	66.1	70.2	67.8	159	131	107	132
QuReg	50.0	65.2	61.2	58.8	80	120	139	113
CKDλ	35.7	59.0	44.1	46.3	126	107	83	105
CKQλ	38.5	61.9	34.1	44.8	123	109	40	91
75%								
UKD24	78.3	76.9	83.4	79.5	130	107	91	110
QuReg	79.8	81.9	71.8	77.9	107	94	71	91
CKDλ	54.3	80.5	72.0	68.9	143	92	76	104
CKQλ	56.8	81.2	78.3	72.5	136	93	75	101
95%								
UKD24	89.1	88.0	94.3	90.5	48	43	29	40
QuReg	96.5	94.2	94.9	95.2	31	28	24	28
CKDλ	85.4	95.3	96.5	92.4	46	28	24	33
CKQλ	93.0	95.2	98.3	95.5	30	28	25	28
99%								
UKD24	94.1	92.4	99.2	95.2	16	16	7	13
QuReg	99.0	97.6	98.9	98.5	7	8	7	7
CKDλ	96.3	98.5	99.6	98.1	8	6	6	7
CKQλ	99.1	98.2	99.9	99.0	6	6	6	6

NOTE: For the θ quantile, the ideal value is θ. The best performing method in each column is underlined.

6. Summary and concluding comments

In many parts of the world, the move towards more sustainable power generation has led to a rapid increase in installed wind power capacity. The assessment of the uncertainty in the future power output from a wind farm is of great importance for the efficient management of power systems and wind power plants. The accuracy of the forecasts of a specific quantile of the wind power density is often of more relevance than the overall accuracy of an estimate of the full density. For example, when wind power producers are offering power to the market for a future period, the optimal bid is a quantile of the wind power density.

This paper has focused on a previously proposed CKD-based approach to wind power density forecasting, which captures the uncertainty in wind velocity, and the uncertainty in the power curve. It is appealing because it involves a nonparametric approach that makes no distributional assumption for wind power, it imposes no parametric assumption for the relationship between wind power and wind velocity, and it allows more weight to be put on more recent observations. As we do not require an accurate estimate of the entire wind power density, our new proposal in this paper is to optimise the CKD-based approach specifically towards estimation of the desired quantile, using the quantile regression objective function.

Using data from three wind farms, we found that overall this approach delivered more accurate quantile predictions than quantile forecasts derived from the density forecasts produced by the original CKD-based method and by an unconditional kernel density estimator. We also implemented a quantile regression approach based on the proposals of Nielsen et al. [12]. We also implemented a method, based on the work of Nielsen et al. [12], who construct a wind power quantile as the sum of a point forecast and a forecast error quantile estimated using quantile regression. Interestingly, the results of this method were competitive with the conditional kernel approaches, especially in terms of the MQRE skill score. A disadvantage of the quantile regression approach is that it is not clear how to constrain the wind power quantile to be between zero and the capacity of the wind farm.

Furthermore, we suspect that quantile crossing (see Section 2.5 of [25]) will be more likely from a pair of quantile regression models than from a CKD-based approach.

In future work, it would be interesting to evaluate empirically the conditional kernel methods for wind velocity density forecasts based on weather ensemble predictions. It would also be interesting to consider the possible incorporation of a copula in the CKD-based approach, which would provide a representation of the interdependency between wind power and the wind velocities (see [34]).

Acknowledgements

We are grateful to the Sotavento Galacia Foundation for making the Sotavento wind farm data available. We would also like to thank George Sideratos of the National Technical University of Athens and the EU SafeWind Project for providing the Greek wind farm data. We are also grateful to two anonymous referees for their useful comments on the paper.

References

[1] W. Katzenstein, E. Fertig, J. Apt, The variability of interconnected wind plants, Energy Policy. 38 (2010) 4400-4410.

[2] P. Pinson, C. Chevallier, G.N. Kariniotakis, Trading wind generation from short-term probabilistic forecasts of wind power, IEEE Transactions on Power Systems. 22 (2007) 1148-1156.

[3] Y. Wang, Y. Xia, X. Liu, Establishing robust short-term distributions of load extremes of offshore wind turbines, Renewable Energy. 57 (2013) 606-619.

[4] E. Cripps, W.T.M. Dunsmuir, Modelling the variability of Sydney Harbour wind measurements, Journal of Applied Meteorology, 42 (2003) 1131-1138.

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

[6] A.M. Foley, P.G. Leahy, A. Marvuglia, E.J. McKeogh, Current methods and advances in forecasting of wind power generation, Renewable Energy. 37 (2012) 1-8.

[7] M.S. Roulston, D.T. Kaplan, J. Hardenberg, L.A. Smith, Using medium-range weather forecasts to improve the value of wind energy production, Renewable Energy. 28 (2003) 585-602.

[8] J.W. Taylor, P.E. McSharry, R. Buizza, Wind power density forecasting using wind ensemble predictions and time series models, IEEE Transactions on Energy Conversion. 24 (2009) 775-782.

[9] J.M. Sloughter, T. Gneiting, A.E. Raftery, Probabilistic wind speed forecasting using ensembles and Bayesian model averaging, Journal of the American Statistical Association. 105 (2010) 25-35.

[10] J.B. Bremnes, A comparison of a few statistical models for making quantile wind power forecasts, Wind Energy. 9 (2006) 3-11.

[11] J.L. Møller, H.A. Nielsen, H. Madsen, Time-adaptive quantile regression, Computational Statistics and Data Analysis. 52 (2008) 1292-1303.

[12] H.A. Nielsen, H. Madsen, T.S. Nielsen, Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts, Wind Energy. 9 (2006) 95-108.

[13] J.W. Taylor, D.W. Bunn, A quantile regression approach to generating prediction intervals, Management Science. 45 (1999) 225-237.

[14] R.S.J. Tol, Autoregressive conditional heteroscedasticity in daily wind speed measurements. Theoretical Applied Climatology. 56 (1997) 113-122.

[15] I. Sanchez, Short-term prediction of wind energy production, International Journal of Forecasting. 22 (2006) 43-56.

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

[17] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, London, 1986.

[18] M. Rosenblatt, Conditional probability density and regression estimators, in: P.R. Krishnaiah (Ed.), Multivariate Analysis II, Academic Press, New York, 1969, pp. 25-31.

[19] I. Gijbels, A. Pope, M.P. Wand, Understanding exponential smoothing via kernel regression, Journal of the Royal Statistical Society Series B. 61 (1999) 39-50.

[20] J. Juban, L. Fugon, G. Kariniotakis, Probabilistic short-term wind power forecasting based on kernel density estimators, European Wind Energy Conference, Milan, Italy, 2007.

[21] J. Fan, T.H. Yim, A crossvalidation method for estimating conditional densities, Biometrika. 91 (2004) 819-834.

[22] P. Hall, J. Racine, Q. Li, Cross-validation and the estimation of conditional probability densities, Journal of the American Statistical Association. 99 (2004) 1015-1026.

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

[24] E.S. Epstein, A scoring system for probability forecasts of ranked categories, Journal of Applied Meteorology. 8 (1969) 985-987

[25] R.W. Koenker, Quantile Regression, Cambridge University Press, Cambridge, UK, 2005.

[26] P. Pinson, H.A. Nielsen, J.K. Moller, H. Madsen, G.N. Kariniotakis. Non-parametric probabilistic forecasts of wind power: Required properties and evaluation. Wind Energy. 10 (2007) 497-516.

[27] R. Koenker, J.A.F. Machado. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association. 94 (1999) 1296-1310.

[28] J.W. Taylor. Evaluating volatility and interval forecasts. Journal of Forecasting 18 (1999) 111-128.

[29] R. Giacomini, I. Komunjer. Evaluation and combination of conditional quantile forecasts. Journal of Business and Economic Statistics. 23 (2005) 416-431.

[30] J. Zhang, S. Chowdhurya, A. Messac, L. Castillo, A multivariate and multimodal wind distribution model, Renewable Energy. 51 (2013) 436-447.

[31] Y. Jiang, Z. Song, A. Kusiak, Very short-term wind speed forecasting with Bayesian structural break model, Renewable Energy. 50 (2013) 637-647.

[32] P. Pinson, G.N. Kariniotakis. Conditional prediction intervals. IEEE Transactions on Power Systems. 25 (2010) 1845-1856.

[33] R.F. Engle, S. Manganelli, CAViaR: Conditional autoregressive value at risk by regression quantiles, Journal of Business and Economic Statistics. 22 (2004) 367-382.

[34] R.J. Bessa, V. Miranda, A. Botterud, Z. Zhou, J. Wang, Time-adaptive quantile-copula for wind power probabilistic forecasting, Renewable Energy. 40 (2012) 29-39.

[35] T.F. Coleman, Y. Li, An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds, SIAM Journal on Optimization, 6 (1996) 418–445.

[36] D. E. Goldberg, Simple genetic algorithm and the minimal deceptive problem, Genetic Algorithms and Simulated Annealing, Morgan Kaufmann, San Mateo, CA, (1987) 74–88.