Wind Power Density Forecasting Using Ensemble Predictions and Time Series Models

J. W. Taylor, P. E. McSharry, *Senior Member, IEEE*, and R. Buizza

Abstract-- **Wind power is an increasingly used form of renewable energy. The uncertainty in wind generation is very largely due to the inherent variability in wind speed, and this needs to be understood by operators of power systems and wind farms. To assist with the management of this risk, this paper investigates methods for predicting the probability density function of generated wind power from one to 10 days ahead at five UK wind farm locations. These density forecasts provide a description of the expected future value and the associated uncertainty. We construct density forecasts from weather ensemble predictions, which are a relatively new type of weather forecast generated from atmospheric models. We also consider density forecasting from statistical time series models. The best results for wind power density prediction and point forecasting were produced by an approach that involves calibration and smoothing of the ensemble-based wind power density.**

Index Terms— **Density forecasting; GARCH models; weather ensemble predictions; wind power; wind speed.**

I. INTRODUCTION

WIND generation is the fastest growing source of renewable energy. However, due to the erratic nature of the earth's atmosphere, there is great variability in wind generated power, and this poses a number of complexities that act as a limiting factor for this source of energy. Fluctuations in wind speed cause the amount of wind power generation to vary with time and location. This variability generates uncertainty that needs to be understood by power system operators in order to ensure that supply and demand are balanced. Reserve electricity supplies are needed whenever wind generators fail to produce anticipated amounts of energy [1]. For wind farm operators, an understanding of the uncertainty is important for both operational and financial reasons. The availability of wind power forecasts and their stochastic properties enables the optimization of the operational strategy for wind turbines [2]. Costs due to wind energy prediction error have been shown to be as much as 10% of total generator energy income, implying a strong need to manage the risks of unexpected levels of generation [3]. This risk clearly relates to the supply side of the power system, but weather risk also has an impact on the demand side [4, 5].

To assist with the management of risk, this paper develops methods to predict the probability density function of wind power generation for lead times from one to 10 days ahead. These density forecasts provide a description of the expected future value and the associated uncertainty. Our empirical analysis uses data for five different UK wind farm locations.

We consider a relatively new type of weather forecast supplied by the European Centre for Medium-Range Weather Forecasts (ECMWF). These forecasts, known as weather ensemble predictions, are generated from atmospheric models and consist of multiple scenarios for the future value of a weather variable. They, therefore, provide an understanding of the uncertainty in the variable. Using an appropriate power curve, the wind speed scenarios can be converted into scenarios for wind power generation. The distribution of these scenarios can be used as a density forecast, with the mean of the scenarios providing a point forecast for wind power.

We also implement a number of statistical time series techniques for generating density forecasts of wind energy. These include the application to daily wind speed data of AR-GARCH and long memory time series models designed to deal specifically with the stochastic and seasonal patterns of the data. The resulting wind speed density forecasts are converted, via a power curve, into wind power density predictions. An alternative approach is to use neural networks, which enable nonlinear modeling [6, 7, 8]. Indeed, these models have also been used for calibration of atmospheric model predictions [9].

Both the atmospheric and time series model approaches involve the generation of wind speed density forecasts, which are converted to wind power predictions. We could have just evaluated the quality of the wind speed density forecasts. However, because of the nonlinear relationship between wind speed and wind power, the suitability of a method for wind speed forecasting is no guarantee of that method's usefulness for wind power prediction [10]. Therefore, we convert all wind speed density predictions into wind power density forecasts in order to evaluate their relative worth.

In Section II, we introduce our five wind farm locations and define the wind power curve that we employ. Section III presents statistical time series models for wind speed. Weather ensemble predictions are described in Section IV. Section V compares point forecasts produced by the different approaches, and in Section VI we compare density forecasts. The final section provides a summary and conclusion.

J. W. Taylor is with the Saïd Business School, University of Oxford, Park End Street, Oxford OX1 1HP, UK (e-mail: james.taylor@sbs.ac.uk).

P. E. McSharry is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK (corresponding author; tel: 44-1865-273095; fax: 44-1865-273905; e-mail: patrick@mcsharry.net).

R. Buizza is with European Centre for Medium-range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, UK (e-mail: r.buizza@ecmwf.int).

II. Wind Speed Data and Power Curve

We used data for wind speed at a height of 10 m, recorded at the five wind farms summarized in Table I. The farms vary in size and capacity, with the newer farms having the more powerful turbines. Our dataset contained daily midday wind speed observations from 1 January 1995 to 30 June 2004, and weather ensemble predictions corresponding to forecast origins of 1 January 1997 to 30 June 2004. We estimate method parameters using data up to the end of 2002, with the final 18 months used for post-sample forecast evaluation.

TABLE I
WIND FARM LOCATIONS USED IN THIS STUDY

Online	Wind farm	Turbine capacity (MW)	No. of turbines	Farm capacity (MW)	Annual homes equivalent
Dec-1992	Blood Hill, Norfolk	0.225	10	2.25	1258
Oct-1997	Llyn Alaw, Anglesey	0.6	34	20.4	11407
Sep-2001	Bears Down, Cornwall	0.6	16	9.6	5368
Mar-2002	Bu Farm, Orkney	0.9	3	2.7	1510
Mar-2002	Cemmaes, Powys	0.85	18	15.3	8555

Fig. 1 shows the Bears Down wind speed series. Although the series is dominated by variation, the peaks every January suggest that there is seasonality within each year. Skewness is also evident in the series, with occasional large values not matched by extremes in the lower direction because of the lower bound of zero. In order to try to gain further insight into the seasonality, in Fig. 2, we plot the wind speed observations against the day of the year for the eight-year estimation period 1995 to 2002. This plot shows that there is seasonality in both the level and variance of the series.
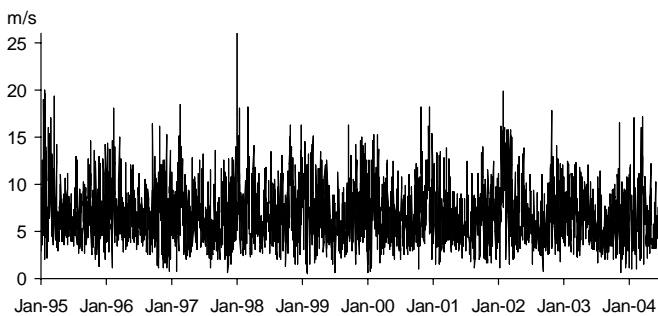

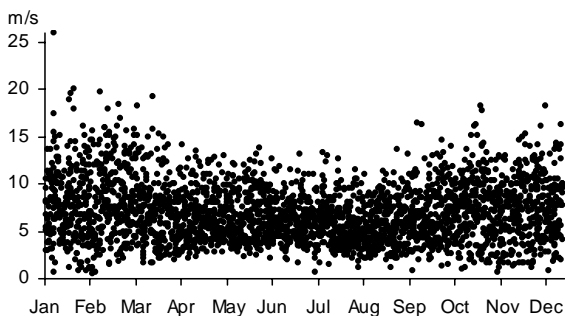Fig. 1. Daily midday wind speed at Bears Down wind farm.

Fig. 2. Daily midday wind speed observations at Bears Down wind farm plotted against the day of the year for the estimation period 1995 to 2002.
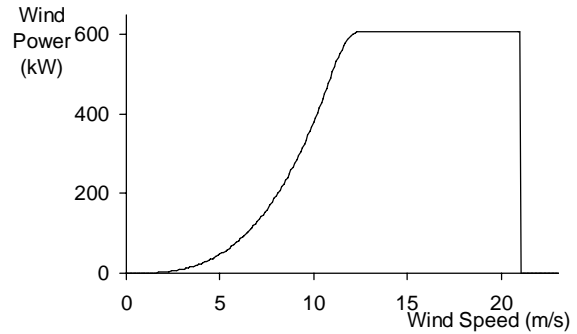

Fig. 3. Wind power curve used in this study for all five locations.

Although a variety of wind turbines are used at the five locations, in this paper, for all five we used a common power curve to convert wind speed to wind power. We did this so that we could average wind power forecasting performance in order to compare methods more easily. The power curve used in this study is shown in Fig. 3. The nonlinear relationship is similar to that used in other studies (e.g. [11]). The curve is essentially cubic up to what is known as the 'nominal speed' beyond which the power generated is limited by the capacity of the turbine. At the 'disconnection speed', the turbine is shut down in order to prevent damage from excessively strong wind. We specified the capacity of the curve to be approximately the average capacity of the five turbine types used at the different wind farms. In practice, the power curve can vary depending on several factors, including other meteorological variables, the power control system of the specific turbine used [12], and on the location and topographical conditions of the wind farm. However, in this paper, for simplicity, we use a single deterministic power curve that is realistic and can be easily reproduced. The same approach is employed in [11]. An alternative approach is to consider the use of a stochastic power curve, and this has been investigated in [12]. In our future work, we intend to model empirically the power curves for a number of wind farms, and use these as a basis for wind power forecasting.

III. Time Series Models For Wind Speed Density Forecasting

A. ARMA-GARCH Models

Recent studies of wind speed time series have used generalized autoregressive conditional heteroskedasticity (GARCH) for the conditional variance. GARCH components allow the variance to evolve in an autoregressive manner over time. In [13] an AR-GARCH model is fitted to daily Canadian mean wind speed data. In order to accommodate the asymmetric nature of the data, rather than the common Gaussian assumption, a Gamma distribution was used. For data recorded at one-minute time intervals, in [14] a bivariate GARCH model is used to model the east-west and north-south wind velocities in Sydney Harbour. In an analysis of wind speed data collected in Texas with a 15-minute interval [15], an ARMA-GARCH-in-mean model is considered, which allow the conditional standard deviation to be included in the ARMA equation for the mean.

As discussed in Section II, our daily wind speed series possess a yearly seasonal pattern. None of the studies mentioned so far in this Section has included seasonal terms. For a seasonal version of ARMA-GARCH, we turn to the model presented in [16] for weekly Dutch temperature data. This model is presented in expression (1):

$$y_t = S(\boldsymbol{\mu}, t) + \phi_1\, y_{t-1} + \varepsilon_t$$
$$\varepsilon_t = \sigma_t\, \eta_t$$
$$\sigma_t^2 = S(\boldsymbol{\omega}, t) + \alpha\left(\varepsilon_{t-1} - S(\boldsymbol{\gamma}, t)\right)^2 + \beta\, \sigma_{t-1}^2$$

(1)

where y_t is the temperature variable, ε_t is an error term, η_t is an i.i.d. error term, σ_t is the conditional standard deviation (volatility), α and β are parameters, and $\boldsymbol{\mu}$, $\boldsymbol{\omega}$ and $\boldsymbol{\gamma}$ are vectors of parameters. The seasonality term, $S(\boldsymbol{\mu}, t)$, appears in the equation for the mean along with a first order autoregressive term. A similar term, $S(\boldsymbol{\omega}, t)$, is used to model the seasonality in the volatility, as well as the asymmetric seasonal impact of temperatures lower and higher than expected on conditional volatility, $S(\boldsymbol{\gamma}, t)$. The seasonality is modeled as a quadratic function. Similar models for daily mean US temperature are used in [17], with the difference that Fourier terms are used to model the seasonality.

In [18], this type of seasonal AR-GARCH model is used for daily wind speed data recorded at five UK cities. In this paper, we also use these models, but our application differs in that we use data for actual wind farm locations and, furthermore, we apply the models to transformed wind speed. In order to stabilize the variance and make the marginal distribution of the data closer to Gaussian, a square root transformation is taken in [19]. The effect that this has on our Bears Down series can be seen from Figs. 4 and 5, which are analogous to the untransformed wind speed plots of Figs. 1 and 2. The transformed series still possesses seasonality, but the transformation has substantially reduced the skewness in the data. An assumption of Gaussian errors would seem to be considerably more appropriate for the transformed series.
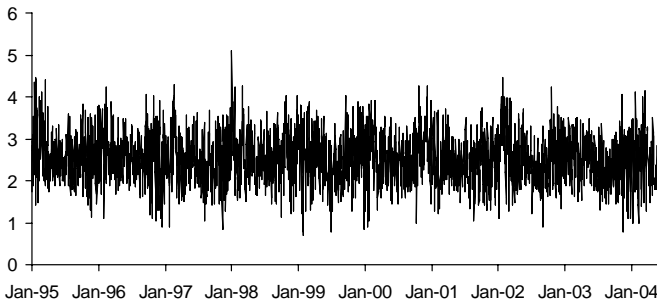

Fig. 4. Square root of daily midday wind speed at Bears Down wind farm.
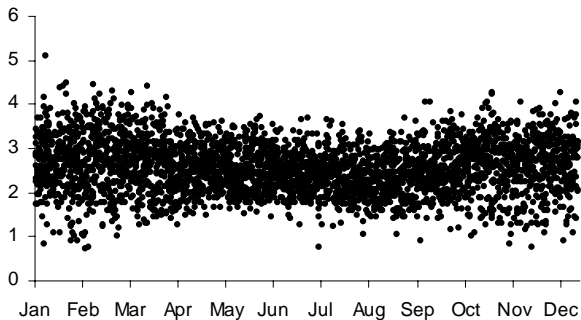

Fig. 5. Square root of daily midday wind speed observations at Bears Down wind farm plotted against the day of the year for the estimation period 1995 to 2002.

TABLE II
PARAMETERS AND DIAGNOSTICS FOR THE AR-GARCH MODEL IN EXPRESSION (2) FITTED TO THE SQUARE ROOT OF WIND SPEED

	Blood Hill	Llyn Alaw	Bears Down	Bu Farm	Cemmaes
Equation for Mean					
μ_0	1.44	1.10	1.58	1.42	1.03
	(17.66)	(15.61)	(18.75)	(16.85)	(15.78)
μ_1	0.03		0.03	0.05	0.03
	(2.60)		(2.46)	(3.19)	(2.07)
μ_2	0.06	0.07	0.10	0.10	0.04
	(4.34)	(4.66)	(7.07)	(6.06)	(3.44)
μ_4	-0.03			-0.04	
	(-2.38)			(-2.82)	
ϕ_1	0.34	0.38	0.38	0.32	0.41
	(17.81)	(19.69)	(20.22)	(17.26)	(20.71)
ϕ_2		0.05		0.07	
		(2.46)		(3.59)	
ϕ_3					0.05
					(2.24)
ϕ_4	0.06	0.05			0.04
	(3.05)	(2.42)			(2.07)
ϕ_5				0.05	
				(2.32)	
ϕ_7		0.06			0.05
		(3.10)			(2.82)
Equation for Variance					
ω	0.0035	0.0057	0.0050	0.0054	0.0041
	(1.16)	(1.46)	(1.40)	(1.41)	(1.46)
α	0.03	0.03	0.04	0.03	0.03
	(3.19)	(3.68)	(4.79)	(3.69)	(3.73)
β	0.96	0.95	0.94	0.96	0.95
	(53.01)	(47.02)	(56.15)	(58.94)	(51.09)
Diagnostics					
$\hat{\eta}_t$ LB Q(7)	4.66	4.88	9.97	1.76	6.14
$\hat{\eta}_t^2$ LB Q(7)	13.72*	10.51	10.20	11.48	4.94
JB for $\hat{\eta}_t$	10.86*	30.72*	0.11	21.74*	3.16
SBC	1.34	1.48	1.55	1.76	1.32

We estimated models similar to those in expression (1), for the square root of our wind speed series using maximum likelihood under the assumption that η_t was Gaussian. We selected models using the Schwarz Bayesian Criterion (SBC) to judge fit. We considered AR terms up to order seven in the equation for the mean. Inclusion of MA terms did not lead to improvements. Fourier modeling of seasonality gave better fit than quadratic modeling. We did not find significant parameters in the asymmetric seasonal volatility function, $S(\boldsymbol{\gamma}, t)$. For two series, we found significant Fourier terms in the variance equation. However, their inclusion resulted in an insignificant GARCH parameter, α. Removal of the GARCH terms resulted in autocorrelation in the squared standardized residuals, indicating an inadequate model of the conditional variance. We, therefore, opted to remove the Fourier terms from the variance equation. Expression (2) shows the model.

$$y_t = \mu_0 + \mu_1 \sin\left(2\pi\, \tfrac{d(t)}{365}\right) + \mu_2 \cos\left(2\pi\, \tfrac{d(t)}{365}\right)$$
$$+ \mu_3 \sin\left(4\pi\, \tfrac{d(t)}{365}\right) + \mu_4 \cos\left(4\pi\, \tfrac{d(t)}{365}\right) + \sum_{i=1}^{7} \phi_i\, y_{t-i} + \varepsilon_t$$
$$\varepsilon_t = \sigma_t\, \eta_t$$
$$\sigma_t^2 = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2$$

(2)

where y_t is the square root of wind speed and $d(t)$ is a repeating step function that numbers the days from 1 to 365 within each year. We removed 29 February from each leap year in our sample in order to maintain 365 days in each year.

In Table II, for each of our five series, we present our preferred model. The table presents the following diagnostics: each parameter with its t-statistic in parentheses; the Ljung-Box Q-statistic to test for autocorrelation in standardized residuals ($\hat{\eta}_t = \hat{\varepsilon}_t / \hat{\sigma}_t$) and squared standardized residuals; the test statistic for the Jarque-Bera test for normality of standardized residuals; and the SBC. The asterisks indicate significance at the 5% level, indicating that the model assumptions can be rejected. Although, three of the models have significant diagnostics, we could not find better simple alternative models, so we decided to use these models.

The AR-GARCH models enable predictions to be made for the mean and variance. A density forecast can then be constructed using either a Gaussian assumption or the empirical distribution of standardized residuals.

B. ARFIMA-GARCH Models

We also implemented the long memory models presented in [19]. Spatio-temporal models are fitted to the square root of daily wind speed recorded at several locations in Ireland, but, as our focus is purely on temporal modeling, we omitted the spatial element in our implementation of the approach. In this study, the square root of wind speed series is deseasonalized, and then an autoregressive fractionally integrated moving average (ARFIMA) model is fitted. This type of model is a generalization of the ARIMA class of models in which the order of differencing, d, is allowed to take non-integer values. If it is found that $0<d<0.5$, the series is stationary, but will possess significant autocorrelation for many lags, which is why they are termed long-memory models.

We deseasonalized each series by subtracting a fitted quadratic function of the day of the year counter, $d(t)$. We were unable to use Fourier terms for this purpose because their use resulted in zero values in the empirical periodogram, which led to problems in identifying the order of fractional differencing. We estimated the degree of fractional integration, d, using the log-periodogram regression estimator [20]. The estimated values of d are shown at the top of Table III. For all series, d was significantly different from zero and 0.5, indicating stationarity and long memory. After applying a long-memory filter to the square root of wind speed, we fitted ARMA-GARCH models to the filtered data. Inclusion of MA terms did not lead to improvements, and so the models in Table III can be described as ARFI-GARCH models. The parameters and diagnostics for these models are presented in the table using the same notation as for the model in expression (2) and Table II. Our use of GARCH terms is an extension of the Haslett and Raftery approach because they did not consider such terms, which is understandable because at the time GARCH was not well established. Using the mean and variance forecasts from these models we constructed separate sets of density forecasts using a Gaussian assumption and the empirical distribution of standardized residuals.

TABLE III
PARAMETERS AND DIAGNOSTICS FOR THE ARFI-GARCH MODEL FITTED TO THE LONG-MEMORY FILTERED SQUARE ROOT OF WIND SPEED

	Blood Hill	Llyn Alaw	Bears Down	Bu Farm	Cemmaes
Order of Fractional Differencing					
d	0.18	0.26	0.24	0.19	0.26
$H_0: d = 0$	(5.74)	(8.05)	(7.39)	(5.93)	(8.23)
$H_0: d = 0.5$	(10.63)	(7.71)	(8.04)	(9.73)	(7.85)
Equation for Mean					
μ_0	0.05	0.00	0.03	0.03	0.00
	(-0.43)	(-0.47)	(-0.34)	(-0.63)	(-0.46)
ϕ_1	0.16	0.13	0.13	0.14	0.14
	(8.42)	(6.60)	(6.93)	(7.37)	(7.36)
ϕ_3	0.05		0.04		
	(2.54)		(2.30)		
Equation for Variance					
ω	0.0034	0.0055	0.0049	0.0055	0.0040
	(2.02)	(2.88)	(2.59)	(2.59)	(2.60)
α	0.03	0.03	0.04	0.03	0.03
	(3.75)	(4.48)	(5.24)	(4.65)	(4.45)
β	0.96	0.95	0.95	0.96	0.95
	(78.24)	(70.66)	(81.58)	(89.77)	(72.62)
Diagnostics					
$\hat{\eta}_t$ LB Q(7)	4.81	5.87	4.00	7.69	4.91
$\hat{\eta}_t^2$ LB Q(7)	15.16*	10.75	8.73	12.34	15.37*
JB for $\hat{\eta}_t$	8.77*	20.71*	9.61*	29.72*	4.82
SBC	1.34	1.48	1.55	1.74	1.31

IV. ATMOSPHERIC MODEL WEATHER ENSEMBLE PREDICTION FOR WIND SPEED DENSITY FORECASTING

A. Ensemble Prediction

Traditional point forecasts are generated from atmospheric models by running the model once at high resolution (horizontal grid spacing) with best estimates for the initial conditions. However, it is important to acknowledge the existence of two sources of uncertainty. First, because the weather is a chaotic system, small errors in the initial conditions of a forecast grow rapidly, and affect predictability. Second, predictability is limited by model errors due to the approximate simulation of atmospheric processes in a numerical model.

Ensemble prediction systems aim to provide an assessment of weather uncertainty by deriving a sophisticated estimate of the probability density function for the weather variables. They involve the generation of multiple realizations of numerical predictions by using a range of different initial conditions in the numerical model of the atmosphere run at a lower resolution than for traditional point prediction. An estimate of the density function is provided by the frequency distribution of the different realizations, which are known as ensemble members. The initial conditions are designed to sample directions of maximum possible growth [21]. They are not sampled as in a statistical simulation because this is not practical for the complex, high-dimensional weather model.

Following the leading examples of the US National Center for Environmental Predictions and the European Centre for Medium-range Weather Forecasts (ECMWF), global ensemble forecasts are now produced daily at the

major meteorological centers. In all these ensemble systems, the number of ensemble members is limited by the necessity to produce forecasts in a reasonable amount of time with the available computer power.

In our work, ensemble forecasts generated at ECMWF have been used. At the time of writing (January 2008), the ECMWF Ensemble Prediction System include 51 members, consisting of one forecast started from the unperturbed, best estimate of the atmosphere initial state plus 50 others generated by varying the initial conditions. The ensemble systems also incorporate stochastic physics, which aims to simulate model uncertainties due to random model error. Since 25 March 2003, the ECMWF ensemble forecasts have been produced twice a-day, with midnight and midday being the two forecast origins, for lead times coinciding with midday and midnight on each of the next 10 days. In the empirical work of this paper, we used only ensemble predictions for the period 1997-2002, and we considered only the midday forecast origin and predictions made for midday on each of the 10 days in the forecast origin. The archived weather variables include both upper level variables (typically temperature, wind speed, humidity and vertical velocity at different heights) and surface variables (e.g. temperature, wind speed, precipitation, cloud cover). Benefit has been found in the use of weather ensemble predictions for electricity demand forecasting and temperature derivative pricing [4, 18].

Following the success of medium-range ensemble prediction systems, ECMWF have introduced a new higher-resolution ensemble prediction system that extends the forecast horizon from 10 to 15 days [22]. Another development in this area is that national meteorological services are developing short-range ensemble prediction systems. Three of these systems are the COSMO-LEPS system, developed by a consortium of countries that include Germany, Greece, Italy and Switzerland, the MOGREPS system, developed by the UK Met Office, and PEACE, the French system. These institutes are investigating the possibility of generating ensemble forecasts with a more accurate spatial resolution, more frequently (possibly up to eight times a day), and for lead times at three hourly intervals up to three days ahead. This change should provide their customers, such as wind farm operators, with more frequent forecast updates, and with more detailed predictions of future weather changes. With regard to our focus on midday data and predictions, the development of these new ensemble prediction systems implies that our analysis can be extended for other periods of the day and for more frequent lead times.

Ensemble predictions have been considered for wind energy forecasting in [11]. We extend this work in the following respects: we use data for five wind farm locations, rather than just Heathrow; we use six years of ensemble predictions, while only two years were used in [11]; we calibrate and smooth the ensemble-based density forecast in a novel way; and we compare ensemble-based prediction with density forecasts from sophisticated times series models.

B. Calibration and Smoothing of Ensemble-Based Wind Speed Density Forecasts

As discussed in the previous section, a density forecast can be constructed as the histogram of the 51 ensemble members. However, previous work has revealed that the resultant density forecast tends to underestimate the uncertainty in the weather variable [18, 23]. In view of this, we calibrated our wind speed ensemble predictions before conversion into wind energy density forecasts. To be consistent with the time series analysis of Section III, we worked with the square root of the wind speed ensemble forecasts. The calibration approach employed in this paper differs from the approach used in previous studies in three respects. First, both the level and the spread are rescaled, as our analysis revealed a bias in both the mean and the variance of the 51 ensemble members. Second, the approach incorporates kernel density estimation in order to smooth the histogram of calibrated ensemble members [24]. Third, both the kernel smoothing and the calibration are jointly estimated using a maximum likelihood approach.

The ith member of the calibrated square root of wind speed ensemble forecast, $\hat{y}^i_{t+k|t}$, is given by

$$\hat{y}^i_{t+k|t} = \mu^{ENS}_{t+k|t} - b_k + \lambda_k \left(\tilde{y}^i_{t+k|t} - \mu^{ENS}_{t+k|t} \right),$$

where $\tilde{y}^i_{t+k|t}$ is the square root of the original wind speed ensemble member; $\mu^{ENS}_{t+k|t}$ is the mean of the 51 ensemble members for the square root of wind speed; t is the forecast origin; k is the forecast horizon; and b_k and λ_k are calibration parameters. The density forecast for the square root of wind speed is estimated using a kernel smoothing technique, which averages over the contributions from each of the 51 ensemble members. The density forecast of the square root of wind speed at value y is given by the following expression:

$$\hat{p}_{t+k|t}(y) = \frac{1}{51} \sum_{i=1}^{51} K(y, \hat{y}^i_{t+k|t}, h_k),$$

where K is the kernel function and h_k is its bandwidth. The size of the bandwidth reflects the uncertainty associated with each ensemble member. Note that each kernel is normalized so that

$$\int_0^\infty K(y, \hat{y}^i_{t+k|t}, h_k)dy = 1,$$

and, therefore, the forecast density has this same property, which is required of a probability density function. In order to account for the fact that wind speed is bounded below by zero, we employed truncated normal distributions with truncation at zero. Using historical ensemble forecasts and verifications, we calculated the sum of the log-likelihood function,

$$L(b_k, \lambda_k, h_k) = \sum_t \ln \hat{p}_{t+k|t}(y_{t+k}),$$

for each forecast horizon k. By maximizing the likelihood using a nonlinear optimization algorithm (Nelder-Mead simplex), we obtained estimates for b_k, λ_k and h_k. Interestingly, we found that the effective spreading of the ensemble caused by the kernel smoothing meant that the estimated value of the calibration factor, λ_k, was typically below one.

V. Comparison of Wind Power Point Forecasting

In this section, we compare 10 methods for wind power point forecasting. We produced forecasts from one to 10 days ahead for each of the five wind locations using as forecast origin each day in the 18-month post-sample period.

Several of the methods use the wind power curve of Fig. 3 to convert wind speed point forecasts to wind power point forecasts. However, this is rather simplistic because a nonlinear function of the mean of a random variable is not necessarily equal to the mean of a nonlinear function of the random variable. In view of this, we also included methods that first generate the wind power density forecast and then use its mean as the wind power point forecast.

A. Methods

The first three methods listed below are simple benchmark methods; the next four involve simulation based on the sophisticated time series models presented in Section III; and the final three use predictions from an atmospheric model.

1. A random walk forecast was created by using the actual wind power for the forecast origin. Actual wind power was calculated from actual wind speed using the power curve.

2. The wind power corresponding to the mean of the actual wind speed for the five most recent periods prior to the forecast origin.

3. The wind power corresponding to the mean of the actual wind speed for the same day of the year in each of the previous five years.

4. The AR-GARCH models in Table II were used to produce mean and variance forecasts for one to 10 days ahead. Using these forecasts and a Gaussian distribution, 10,000 values were simulated for the square root of wind speed for each lead time. Using the power curve in Fig. 3, these values were converted into 10,000 simulated wind power realizations. The mean of these was used as the point forecast.

5. This method was identical to Method 4 except it used the empirical distribution of standardized AR-GARCH residuals, instead of a Gaussian distribution.

6. This method was identical to Method 4 except it used the ARFI-GARCH models of Table III.

7. This method was identical to Method 5 except it used the ARFI-GARCH models of Table III.

8. Using the power curve in Fig. 3, we calculated wind power predictions based on traditional meteorological wind speed point forecasts, which are generated by running the atmospheric model once at high resolution with best estimates for the initial conditions.

9. Using the uncalibrated 51 wind speed ensemble members, we produced 51 wind power scenarios for each lead time. The mean of these was used as the point forecast. Because this method derives a point forecast from the wind power density forecast, it is more statistically appealing than Method 8. Furthermore, the mean of the 51 weather ensemble members is often found to be a more accurate weather point forecast than the traditional high resolution forecast [25].

10. Using the calibrated and kernel smoothed ensemble-based densities for the square root of wind speed, 10,000 values were sampled for each lead time. Using the power curve in Fig. 3, these values were converted into simulated wind power. The mean of these was used as the point forecast.

B. Results

For seven of the 10 methods, Fig. 6 shows the mean absolute error (MAE) for the post-sample forecast errors averaged over the five series. The relative performances of the 10 methods were very similar for each of the five locations, and so the average MAE results in Fig. 6 provides a useful summary. The MAE results were similar for the AR-GARCH and ARFI-GARCH models with Gaussian or empirical distribution, so we show only the results for one of these methods, Method 6. It is disappointing to see that beyond two days ahead this method is matched by the simplistic approach that uses the mean of the actual wind power for the same day in each of the previous five years.

For prediction up to four days ahead, the best results were produced by the three methods that use predictions from the atmospheric model. Comparing these three approaches, Fig. 6 shows that there is benefit in using the ensemble predictions, as in Methods P9 and P10, rather than the naïve use of the high resolution atmospheric model wind speed point forecasts as in Method 8. The results for Method 10 are an improvement on those for Method 9, showing that there is benefit in using the calibrated and smoothed ensemble-based density forecast. Overall, of the 10 methods, Method 10 has the best results, although it is slightly outperformed by the 5-year mean and the sophisticated statistical models for the three longest lead times.

The relative performances of the methods were similar when evaluated using root mean square error (RMSE). The one exception to this was that the sophisticated time series methods outperformed the 5-year mean for all lead times. For simplicity we do not report in any further detail the RMSE results.
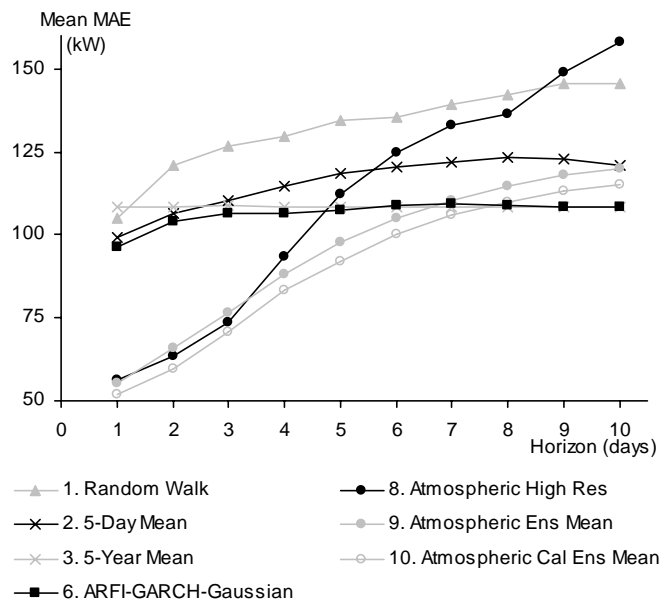

Fig. 6. Post-sample MAE results averaged over the five wind farm locations.

VI. Comparison of Wind Power Density Forecasting

We generated density forecasts using Methods 4, 5, 6, 7, 9 and 10, described in Section V. We evaluated forecast

quality using the log likelihood and the probability integral transform [26]. For the latter, the relative ranking of the methods was unclear, and so for simplicity, we present here only the log-likelihood results. The log-likelihood for verification time $t+k$ and forecast horizon k is $L_{t+k|t} = \ln \hat{p}_{t+k|t}(z_{t+k})$, where $\hat{p}_{t+k|t}(z_{t+k})$ is the probability estimate, provided by the density forecast, evaluated at the observed wind power, z_{t+k}. This can be empirically calculated by estimating the derivative of the cumulative distribution function (CDF) of wind power. For Methods 4, 5, 6, 7 and 10, we constructed the CDF using M=10,000 sampled values from the density forecast, and for Method 9, we used the M=51 wind power scenarios. Let Δz be the difference between the values on either side of the observed wind power. The empirically calculated log-likelihood is given by $\ln 1/((M+1)\Delta z)$. The sum of these log-likelihoods provides a score for each forecast horizon.

The log-likelihood results were similar for the time series models with Gaussian or empirical distribution, so we report only the results for the models with a Gaussian assumption, Methods 4 and 6. Fig. 7 shows the sum of the log-likelihoods averaged over the five series. The averaging of the results is reasonable because the relative performances of the methods were very similar for each of the five locations. A higher value of the log-likelihood measure is preferable. The method that overall performs the best in Fig. 7 is Method 10, which is the calibrated and smoothed distribution of the 51 ensemble members. The quality of the density forecast will depend in part on the quality of the estimate of the mean. However, it is interesting to note that the accuracy for the early lead times of Method 9 for forecasting the mean, as shown in Fig. 6, is not repeated in Fig. 7. This indicates that the relative rankings of the methods in Fig. 7 is certainly not entirely due to the ability of the methods to estimate the mean of the density and must, therefore, also be due to the quality of the methods for estimating the other features of the density.
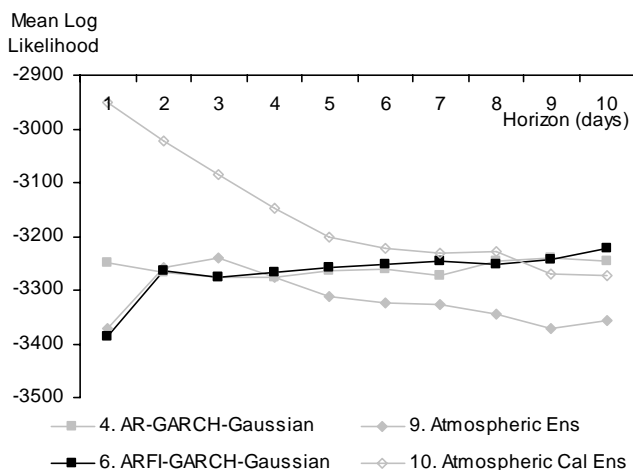

Fig. 7. Post-sample log-likelihood results averaged over the five locations.

VII. SUMMARY AND CONCLUDING COMMENTS

In this paper, we have shown how wind power density forecasts, for lead times from one to 10 days ahead, can be generated from wind speed ensemble predictions produced by an atmospheric model. The systematic bias in the location and scale of the distribution of the 51 ensemble members was corrected using a calibration approach that also incorporated kernel smoothing, with parameters optimized using maximum likelihood. We compared the resultant density forecasts with those from sophisticated time series models built using wind speed observations. The calibrated and smoothed ensemble-based density forecasts were found to be more accurate up to a lead time of about a week. The resultant point forecasts were comfortably superior to those generated by the time series models and those based on traditional high resolution wind speed point forecasts from an atmospheric model. It is, therefore, our conclusion that weather ensemble predictions have strong potential for wind power forecasting.

VIII. REFERENCES

[1] R. Doherty, and M. J. O'Malley, "A new approach to quantify reserve demand in systems with significant installed wind capacity," *IEEE Transactions on Power Systems*, vol. 20, pp. 587-595, 2005.

[2] E. D. Castronuovo, and J. A. P.Lopes, "On the optimization of the daily operation of a wind-hydro power plant," *IEEE Transactions on Power Systems*, vol. 19, pp. 1599-1606, 2004.

[3] A. Fabbri, T. Gomez San Roman, J Rivier Abbad, and V. H. Mendez, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Transactions on Power Systems*, vol. 20, pp. 1440-1446, 2005.

[4] J. W. Taylor, and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions on Power Systems*, vol. 17, pp. 626-632, 2002.

[5] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Transactions Power Systems*, vol. 20, pp. 1166-1172, 2005.

[6] G. N. Karniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural network models", *IEEE Transactions on Energy Conversion*, vol. 11, pp. 762-767, 1996.

[7] M.C. Alexiadis, P. S. Dokopoulos, and H. S. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models," *IEEE Transactions on Energy Conversion*, vol. 14, pp. 836-842, 1999.

[8] P. Pinson, and G. N. Kariniotakis, "Wind power forecasting using fuzzy neural networks enhanced with on-line prediction risk assessment," *Proceedings of IEEE Power Tech Conference*, Bologna, vol. 2, pp. 8, 23-26 June 2003.

[9] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models," *IEEE Transactions on Energy Conversion*, vol. 21, pp. 273-284, 2006.

[10] P. M. Lange, "On the uncertainty of wind power predictions – analysis of the forecast accuracy and statistical distributions of errors," *Journal of Solar Energy Engineering*, vol. 127, pp. 177184, 2005.

[11] M. S. Roulston, D. T. Kaplan, J. Hardenberg, and L. A. Smith, "Using medium-range weather forecasts to improve the value of wind energy production, " *Renewable Energy*, vol. 28, pp. 585-602, 2003.

[12] I. Sánchez, "Short-term prediction of wind energy production," *International Journal of Forecasting*, vol. 22, pp. 43-56, 2006.

[13] R. S. J. Tol, "Autoregressive conditional heteroscedasticity in daily wind speed measurements," *Theoretical Applied Climatology*, vol. 56, pp. 113-122, 1997.

[14] E. Cripps, and W. T. M. Dunsmuir, "Modeling the variability of Sydney Harbor wind measurements," *Journal of Applied Meteorology*, vol. 42, pp. 11313-1138, 2003.

[15] B. T. Ewing, J. Brown Kruse, and J. L. Schroeder, "Time series analysis of wind speed with time-varying turbulence", *Environmetrics*, vol. 17, pp. 119-127, 2006.

[16] P. H. Franses, J. Neele, and D. van Dijk, "Modeling asymmetric volatility in weekly Dutch temperature data," *Environmental Modelling and Software*, vol. 16, pp. 131-137, 2001.

[17] S. Campbell, and F. X. Diebold, "Weather forecasting for weather derivatives," *Journal of the American Statistical Association*, vol. 100, pp. 6-16, 2005.

[18] J. W. Taylor, "Forecasting weather variable densities for weather derivatives and energy prices," in *Modelling Prices in Competitive Electricity Markets*, D. W. Bunn, Ed. Chichester, UK: Wiley, 2004, pp. 307-330.

[19] J. Haslett, and A. E. Raftery, "Space-time modelling with long-memory dependence: assessing Ireland's wind power resource," *Applied Statistics*, vol. 38, pp. 1-50, 1989.

[20] J. Geweke, and S. Porter-Hudak, "The estimation and application of long memory time series models," *Journal of Time Series Analysis*, vol. 4, pp. 221-238, 1983.

[21] R. Buizza, T. Petroliagis, T.N. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons, and N. Wedi, "Impact of model resolution and ensemble size on the performance of an ensemble prediction system," *Quarterly Journal of the Royal Meteorological Society*, vol. 124, pp. 1935-1960, 1998.

[22] R. Buizza, J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, "The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System)," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, pp. 681-695, 2007.

[23] J. W. Taylor, and R. Buizza, "Density forecasting for weather derivative pricing," *International Journal of Forecasting*, vol. 22, pp. 29-42, 2006.

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

[25] F. Molteni, R. Buizza, T.N. Palmer, and T. Petroliagis, "The new ECMWF ensemble prediction system: methodology and validation," *Quarterly Journal of the Royal Meteorological Society*, vol. 122, pp. 73-119, 1996.

[26] M. P. Clements, *Evaluating Econometric Forecasts of Economic and Financial Variables*, Hampshire, UK: Palgrave Macmillan, 2005.

IX. BIOGRAPHIES

James W. Taylor is a Professor of Decision Science at the Saïd Business School, University of Oxford. His research interests include exponential smoothing, prediction intervals, quantile regression, volatility forecasting, energy forecasting and weather ensemble predictions.

Patrick E. McSharry is a Royal Academy of Engineering/EPSRC Research Fellow at the Department of Engineering Science, University of Oxford, and a Senior Member of IEEE. He is currently supported by a Marie Curie Research Fellowship, funded by the European Union's Sixth Framework. His research interests include signal processing, data analysis, complex dynamical systems, mathematical modelling and forecasting.

Roberto Buizza is a principal scientist and the Leader of the EPS Group at the European Centre for Medium-Range Weather Forecasts. His research interests include instability theory, chaos and predictability of the atmospheric flow, diagnostics of deterministic and probabilistic numerical weather forecasting systems, weather-risk management and economic value of weather forecasts.