Quantile-Copula Density Forecast for Wind Power Uncertainty Modeling

R. J. Bessa, J. Mendes, V. Miranda, *Fellow*, *IEEE*, A. Botterud, *Member*, *IEEE*, J. Wang, *Member*, *IEEE*, and Z. Zhou, *Member*, *IEEE*

***Abstract*—A probabilistic forecast, in contrast to a point forecast, provides to the end-user more and valuable information for decision-making problems such as wind power bidding into the electricity market or setting adequate operating reserve levels in the power system. One important requirement is to have flexible representations of wind power forecast (WPF) uncertainty, in order to facilitate their inclusion in several problems. This paper reports results of using the quantile-copula conditional Kernel density estimator in the WPF problem, and how to select the adequate kernels for modeling the different variables of the problem. The method was compared with splines quantile regression for a real wind farm located in the U.S. Midwest.**

***Index Terms*-- Wind power, forecasting, probabilistic, uncertainty, kernel density, copula.**

I. Introduction

A branch of current research in wind power forecasting is devoted to developing models that forecast the wind power uncertainty. Wind power uncertainty can take the form of probabilistic forecasts, risk indices, or scenarios for short-term wind power generation [1]. Furthermore, the representation of the uncertainty usually relies on the algorithm used. For instance, uncertainty is represented by a set of quantiles, when using quantile regression [2].

The representation of uncertainty is an important feature to consider when building an algorithm for wind power uncertainty forecasting. To accommodate use in different decision problems it is an advantage with a flexible representation of wind power uncertainty. The work reported in this paper represents wind power uncertainty as a probability density function (*pdf*).

The proposed uncertainty forecast model can take as input deterministic wind power point forecasts or Numerical Weather Prediction (NWP) data, and produce a *pdf* from which deterministic and probabilistic forecasts could be computed.

From an information theory perspective, the *pdf* contains all the information associated with a random variable, as it enables to compute the moments of the forecasted distribution [3]. Therefore, since the *pdf* representation is generic and can be transformed into several forms such as quantiles, standard deviation, or skewness, it is the uncertainty representation adopted in this work.

Uncertainty represented by a *pdf* can be easily integrated into several decision-making problems. For instance, it can be used to compute the expected utility or the value-at-risk when bidding in the electricity market [4], or to characterize the tails of the system generation margin for setting the operating reserve [5], or to represent wind power generation in network nodes for power flow analysis [6].

Moreover, the *pdf* representation copes with multimodal distributions, and allows the computation of the modes instead of just computing the expected value (which in this case is not a good summary of the distribution). It is unlikely to find wind power multimodal density distributions, but the mode or the median is still a better deterministic forecast, because normally the wind power distributions are highly skewed.

Kernel Density Forecast (KDF) is not new in the wind power forecast state-of-the-art; the first KDF model was introduced by Juban *et al.* [7]. The main difference to the model presented in this paper is in the mathematical formulation. The method from Juban *et al.* is an adaptation of the classic Nadaraya-Watson conditional kernel density estimator (cKDE) [8]. In contrast, we use the quantile-copula [9], which is a different mathematical formulation. Instead of a multivariate kernel density estimator [10] for the joint density function, the quantile-copula uses a copula for modeling the dependency structure between the variables.

According to Faugeras [9] the main advantages of the quantile-copula approach over the Nadaraya-Watson estimator are: the methods based on the Nadaraya-Watson estimator are numerically unstable when the denominator is close to zero; for a problem with several explanatory variables, the method has only one kernel product, instead of two; at a conceptual level, density estimation should only be based on density estimation methods and not on regression approaches.

The main contributions from this paper are: i) application of the quantile-copula approach for the first time to the wind power forecasting problem; ii) selecting adequate kernels for modeling the different variables types, e.g. bounded variables and circular variables.

The full paper is organized as follows: section II describes the Quantile-Copula estimator, which is applied for the first time to the wind power problem; section III presents results for a real wind farm located in the U.S. Midwest, obtained with

The authors R.J. Bessa, J. Mendes, V. Miranda are with Instituto de Engenharia de Sistemas e Computadores do Porto (INESC Porto)and Faculty of Engineering, University of Porto, Portugal (emails: rbessa@inescporto.pt, joana.p.mendes@inescporto.pt and vmiranda@inescporto.pt). The authors A. Botterud, J. Wang and Z. Zhou are with Argonne National Laboratory, USA (e-mails: abotterud@anl.gov, jianhui.wang@anl.gov, zzhou@anl.gov).

The work of R. J. Bessa was supported by Fundação para a Ciência e Tecnologia (FCT) Ph.D. Scholarship SFRH/BD/33738/2009.

different kernels, different bandwidth sizes, different input variables, and compared to a method from the state-of-the-art (i.e. splines quantile regression [2]); section IV summarizes the conclusions.

II. Kernel Density Forecast (KDF) Methodology

A. Formulation and Output

The wind power probability density forecast problem consists in determining the wind power *pdf* at time step *t* for each look-ahead time step *t+k* of a given time-horizon (*e.g.* up to 72 hours ahead), knowing *a priori* a set of explanatory variables (NWP forecasts, wind power measured values).
This is formulated as:

$$\hat{f}_P\left(p_{t+k} \mid X = x_{t+k|t}\right) = \frac{f_{P,X}\left(p_{t+k}, x_{t+k|t}\right)}{f_X\left(x_{t+k|t}\right)} \tag{1}$$

where p_{t+k} is the wind power forecasted for look-ahead time $t+k$, $x_{t+k|t}$ are the explanatory variables forecasted for look-ahead time step $t+k$ and available/launched at time step t, $f_{P,X}$ is the joint density function of the forecasted wind power and explanatory variables, f_X is the density function of the set of explanatory variables, and $\hat{f}_P\left(p_{t+k} \mid X = x_{t+k|t}\right)$ is the forecasted wind power density function for look-ahead time step $t+k$.

A possible representation of a forecasting output is given by the stacked conditional plot depicted in Fig. 1.

This figure depicts the changes in the wind power density function for different values of wind speed (ranging from 0 to 20 m/s). In fact, if for instance the forecast wind speed is 15 m/s, then the *pdf* of the forecast wind power corresponds to the distribution in the 15 m/s line.

Note that the conditional densities for intermediate values of wind speed are very broad. Moreover, we may detect a higher concentration of density in the tails for lower and higher values of wind speed.

B. Quantile-Copula (QC) Estimator

The basic theoretical framework for KDF is the kernel density estimation (KDE), which is a non-parametric estimator of the *pdf* given by [11]:

$$\hat{f}_x(x) = \frac{1}{N \cdot h}\sum_{i=1}^{N} K\left(\frac{x - X_i}{h}\right) \tag{2}$$

where N is the number of samples, K is a Kernel function and h the bandwidth parameter. This equation, places a Kernel around each sample X_i.

Moreover, the multivariate kernel density estimator [10] is given for the bivariate case by:

$$\hat{f}(x_1,\ldots,x_d) = \frac{1}{N}\sum_{i=1}^{N}\prod_{j=1}^{d}\frac{1}{h}K_j\left(\frac{x_j - X_{ij}}{h_j}\right) \tag{3}$$

where K is a multivariate Kernel function and $h_1,\ldots,h_d$ a bandwidth vector.

Conditional density estimation consists of estimating the density of a random variable Y conditioned on $x=X$, which can be formulated as follows:

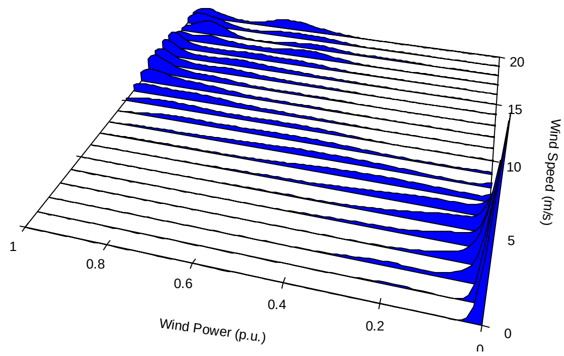

Fig. 1. Example of tacked conditional plot for wind power (axis to the left) and forecasted wind speed (axis to the right).

$$\hat{f}(y \mid X = x) = \frac{\hat{f}_{XY}(x,y)}{\hat{f}_X(x)} \tag{4}$$

where $\hat{f}_{XY}(x,y)$ is the multivariate density function of X and Y (joint distribution function), and $\hat{f}_X(x)$ is the marginal density of X.

Almost at the same time, Faugeras [9] and Bouezmarni and Rombouts [12] proposed the idea of using a copula for modeling the dependency structure between Y and X. Regarding copulas, the Sklar theorem [13] states the following for the bivariate case:

"Let H be a two-dimensional distribution function with marginal distribution functions F and G. Then there is a copula C such that

$$H(x,y) = C(F(x), G(y)) \tag{5}$$

Conversely, for any univariate distribution functions F and G and any copula C, the function H is a two-dimensional distribution function with marginals F and G. Furthermore, if F and G are continuous, then C is unique."

This theorem means that the multivariate distribution function can be separated into two parts: i) marginal functions that can be estimated separately; ii) dependency structure between the marginal which is modeled by the copula. For more details about copulas see [14]. Since we know that

$$F_{XY}(x,y) = C(F_X(x), G_Y(y)) \tag{6}$$

then Eq. 3 can be replaced by:

$$f(x,y) = \frac{\partial^2}{\partial u \cdot \partial v}C(u,v) = f_X(x) \cdot f_Y(y) \cdot c(u,v) \tag{7}$$

where u and v are quantile transforms of the data, $u=F_X(x)$ and $v=F_Y(y)$, and c is the copula density function.

Replacing Eq. 7 in Eq. 4, we have the following conditional density estimator:

$$f(y \mid X = x) = f_Y(y) \cdot c(u,v) \tag{8}$$

Now it is necessary to build an estimator for the copula c in Eq. 8.

The idea proposed in [12] was to consider a parametric model for the copula, with the marginal distributions being represented by a nonparametric model (empirical distribution function). However, we followed the idea described in [9], where the copula density is estimated with KDE. Hence, the estimator for $f_Y(y)$ is the KDE in Eq. 2, and the copula density estimator is the estimator in Eq. 3 as follows (for the bivariate case):

$$\hat{c}(u,v)=\frac{1}{N}\sum_{i=1}^{N}K\left(\frac{u-U_i}{h}\right)\cdot K\left(\frac{v-V_i}{h}\right) \tag{9}$$

where U_i and V_i are the data transformed by the empirical cumulative distribution function, $U_i=F_X^e(X_i)$ and $V_i=F_Y^e(Y_i)$.

Fig. 2 depicts the copula density computed with Eq. 9, for the quantile transform of the wind speed and wind power (*i.e.* using *F*), for a real wind farm. From Fig. 2 we see that there is a strong dependence in the two extreme corners, e.g. when there are lower quantiles in wind speed (lower wind speed values) the wind power quantiles also present lower values with a higher probability.

Hence, substituting Eq. 9 and 2 in Eq. 8, the estimator becomes:

$$\hat{f}(y\mid X=x)=\frac{1}{N\cdot h}\cdot\sum_{i=1}^{N}K\left(\frac{y-Y_i}{h}\right)\cdot \frac{1}{N}\cdot\sum_{i=1}^{N}K\left(\frac{F_X^e(u)-F_X^e(U_i)}{h_x}\right)\cdot K\left(\frac{F_Y^e(v)-F_Y^e(V_i)}{h_y}\right) \tag{10}$$

In the wind power problem the variable *Y* is the wind power, and the explanatory variables *X* could be NWP variables (*e.g.* wind speed, wind direction, pressure, temperature), wind power point forecast, and measured wind power.

C. *Choice of the Kernel Function*

The choice of the kernel function for the wind power forecasting problem depends on the type of variable and data. The wind power is bounded between zero generation and rated power. Here, two Chen beta kernels [15] were used to model the variables with support [0,1] (wind power and quantile transformed variables):

$$K_1(x)=\frac{1}{N}\sum_{i=1}^{N}K_{x/b+1,(1-x)/b+1}(X_i) \tag{11}$$

where $K_{p,q}$ is the density function of a *Beta(p,q)* random variable with *p* and *q* as the two positive shape parameters, and *b* being the bandwidth parameter of $K_{p,q}$; and

$$K_2(x)=\frac{1}{N}\sum_{i=1}^{N}K_{x,b}^{*}(X_i) \tag{12}$$

$K_{x,b}^{*}$ are boundary kernels defined as

$$K_{x,b}^{*}(t)=\begin{cases}K_{x/b,(1-x)/b}(t)\ if\ \ x\in[2b,1-2b]\\ K_{\rho(x),(1-x)/b}(t)\ if\ \ x\in[0,2b)\\ K_{x/b,\rho(1-x)}(t)\ if\ \ x\in(1-2b,1]\end{cases} \tag{13}$$

where $\rho(x,b)=2b^2+2.5-\sqrt{4b^4+6b^2+2.25-x^2-x/b}$ and *Kp,q* is a *Beta(p,q)* density function.

The swapped version between x and X and the bivariate Gaussian copula [16] were also considered as Kernels for this type of variables.

The wind direction and hour of the dayare circular variables and a von Mises distribution was therefore used [17]:

$$g(\theta;\mu,\kappa)=\frac{1}{2\pi\cdot I_0(\kappa)}e^{\kappa\cdot\cos(\theta-\mu)} \tag{14}$$

where I_0 is the modified Bessel function of the first kind and order 0 and defined by

$$I_0(\kappa)=\frac{1}{2\pi}\int_0^{2\pi}e^{\kappa\cdot\cos(\theta)}d\theta \tag{14}$$

The parameter μ is the directional center of the distribution, κ is the concentration parameter and θ belongs to any interval of length 2π. Note that the circular kernels are used for variables *u* and *v* in the copula, hence it is necessary to perform a change of scale from [0,1] to [0, 6.266] (in radians).

III. Case-Study

A. *General Description*

The wind power data is taken from a large wind farm located in flat terrain in the U.S. Midwest. The NWP data was generated with the Weather Research and Forecasting (WRF) model [18] by Argonne National Laboratory and consists of several weather variables (e.g. wind speed, direction, temperature) for one reference point inside the wind farm, launched at 6:00 AM.

The complete dataset (SCADA and NWP) correspond to the period between January 1st 2009 and February 20th 2010.

For wind power predictions, the temporal horizons were day-ahead forecasts (section III.C) and between look-ahead time step t+6 and t+48 (section III.D). The temporal resolution of the forecasts is one hour.

As for both SCADA and NWP data, they have a temporal resolution of 10 minutes, so that a simple averaging of the 10 minutes data was performed to produce hourly data.

B. *Evaluation Metrics*

The outputs obtained from the Quantile-Copula (QC) estimator are compared with a method from the state-of-the-art, i.e. splines quantile regression [2]. The aim is to analyze the difference between the two methods, and also to show that the performance results obtained with QC are consistent with what is found in the state-of-the-art methods.

The output of both methods is represented by a set of quantiles, e.g. ranging between 5% and 95% with a 5% increment.

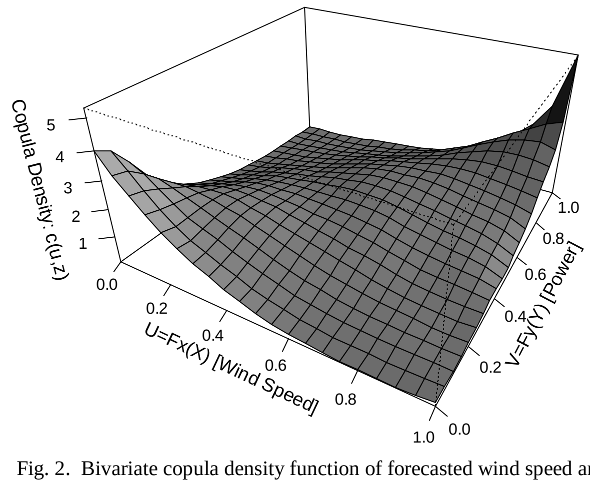

Fig. 2. Bivariate copula density function of forecasted wind speed and measured wind power.

Results are then assessed using the evaluation framework detailed in [19], which consists of a series of quantile forecasts for unique or varying nominal proportions and observations (measured values).

In this study, three evaluation metrics were considered. The first one is calibration, and it is related to the level of agreement between nominal proportions (forecasted probabilities) and the ones computed from the evaluation sample. This means that, for instance, an 85% quantile should contain 85% of the observed values lower or equal to its value.

The second metric is sharpness, which measures the tendency of probability forecasts towards discrete forecasts. It is measured by the mean size of the forecast intervals (distance between quantiles). Quantiles are gathered by pairs in order to obtain intervals with different nominal coverage rate. This gives an indication on the level of usefulness, where narrow intervals are desired.

The third metric is a skill score, which gives combined information about a model's probabilistic forecast evaluation in a single measure. According to [19], the skill score of an event p_{t+k} with predictive distribution $\hat{f}_{t+k}$ can be computed, for each look-ahead time-step k, as the summation over m quantiles, $\hat{q}_{t+k}^{\alpha_i}$ with nominal proportion α_i and actual outcome p_{t+k}:

$$Sc(\hat{f}_{t+k}, p_{t+k}) = \sum_{i=1}^{m} (\xi^{(\alpha_i)} - \alpha_i)(p_{t+k} - \hat{q}_{t+k}^{(\alpha_i)}) \qquad (15)$$

where ξ is the indicator variable, indicating whether the quantile covers the actual outcome ("hit") or not ("miss"). The higher the scoring rule the better, with the maximum value 0 obtained for perfect probabilistic forecasts. In [19] it is suggested that calibration should be assessed in a first analysis (as the primary requirements), and then the information provided by skill score allows to derive conclusions about the remaining metrics (e.g. sharpness).

C. *Results: Different Kernels and Bandwidth Sizes*

The time horizon of this section is day-ahead forecasts (i.e. between t+18 and t+42) and the input variables are the forecasted wind speed from the NWP only.

1) *Evaluation with different kernel types:* Comparisons between different kernel functions were performed for various combinations of the kernel bandwidths (only the results for two combinations are presented).

The Kernel functions used for both wind speed and wind power variables were: Chen's beta kernels (*Chen 1* and *Chen 2* from Eqs. 11 and 12, respectively); the bivariate Gaussian copula (see [16]); Chen's beta kernels with X swapped with x (*Beta 1x* and *Beta 2x* from Eqs. 11 and 12). The kernel size values were determined experimentally (by trial-error), testing it offline.

Results comparing different kernels under distinct kernel sizes proved that *Chen 1* is the estimator with the best overall calibration performance. Figs. 3 and 4 show the calibration results for (h_{Power};$h_{WindSpeed}$) = (0.01; 1.2) and (h_{Power};$h_{WindSpeed}$) = (0.004; 1), respectively. The figures show the deviation from perfect calibration (i.e. where empirical proportions match nominal or forecasted proportions). It is clear that quantiles are, in general, under-estimated for quantiles below 60%, as the deviation between nominal and empirical proportion rate is negative.

Figs. 5 and 6 depict the sharpness diagram, where lower interval lengths correspond to a better performance in terms of sharpness. The x-axis is the nominal coverage of the forecast interval (1-α) and the y-axis is the average size of the intervals.

All kernels have similar sharpness results, *Beta 1x* having the best performance and the *Gaussian copula* the worst.

Fig. 7 and 8 depicts the skill score computed for each look-ahead time step for several Kernels. Chen 1 proved to have the best overall performance, as its skill score value is the largest (closest to zero) for any kernel size chosen.

To summarize, the beta kernel *Chen 1* from Eq. 11 has a better performance not only in calibration, but also in skill score. All the kernels presented similar sharpness results, although *Chen 1* and *2* tend to have a better performance.

2) *Evaluation with different bandwidth sizes:* The impact of different Kernel bandwidth sizes were assessed for both wind speed and power variables. Results from this evaluation, using the *Chen 1* kernel, have shown a similar calibration performance for both wind speed kernel sizes lower and higher (or equal) than 1. Fig. 9 depicts this behavior, showing that quantiles are over-estimated in the former case and under-estimated in the latter.

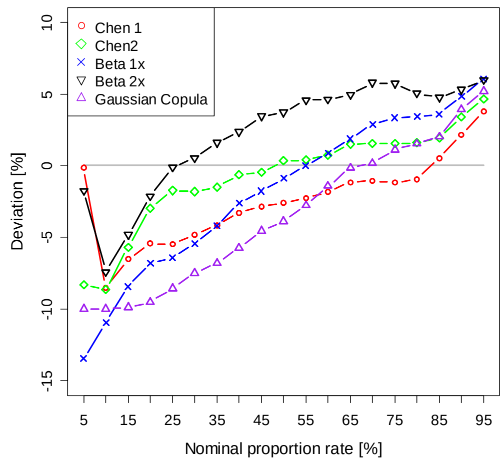

Fig. 3. Calibration diagram for (h_{Power};$h_{WindSpeed}$) = (0.01; 1.2).

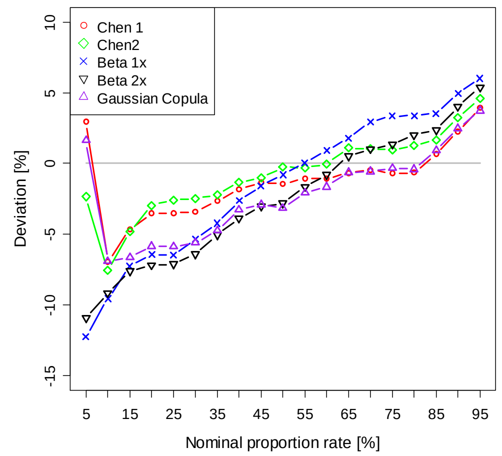

Fig. 4. Calibration diagram for (h_{Power};$h_{WindSpeed}$) = (0.004; 1).

Fig. 10 illustrates sharpness results. This performance was worse for wind speed kernel sizes equal to or larger than 1. The same behavior is obtained in terms of the skill score performance, as shown in Fig. 11.

To summarize, different kernel sizes lead to distinct results. By changing the kernel size it is possible to move from quantile underestimation to overestimation, and vice-versa. The impact is also significant in terms of sharpness and skill score.

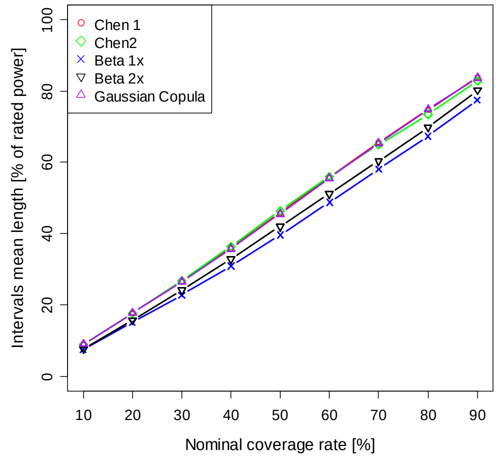

Fig. 5. Sharpness diagram for (h_{Power};$h_{WindSpeed}$) = (0.01; 1.2).

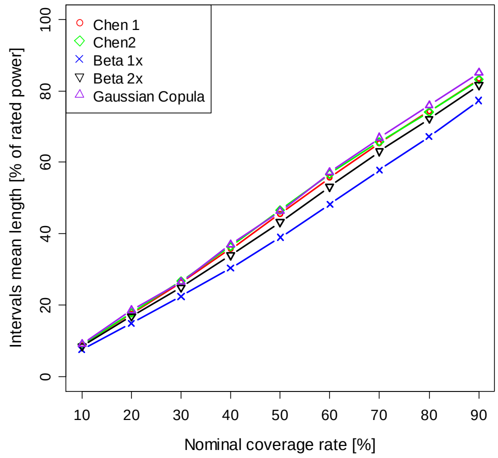

Fig. 6. Sharpness diagram for (h_{Power};$h_{WindSpeed}$) = (0.004; 1).

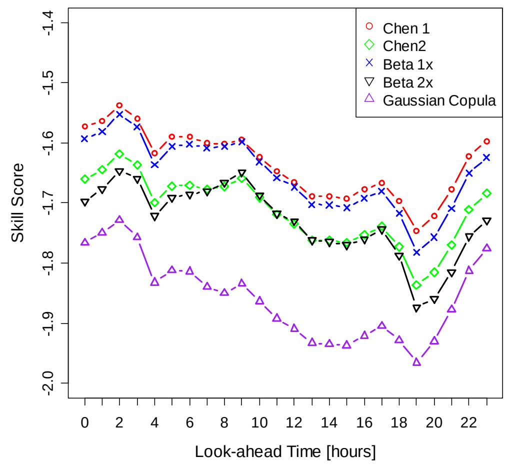

Fig. 7. Skill score diagram for (h_{Power};$h_{WindSpeed}$) = (0.01; 1.2).

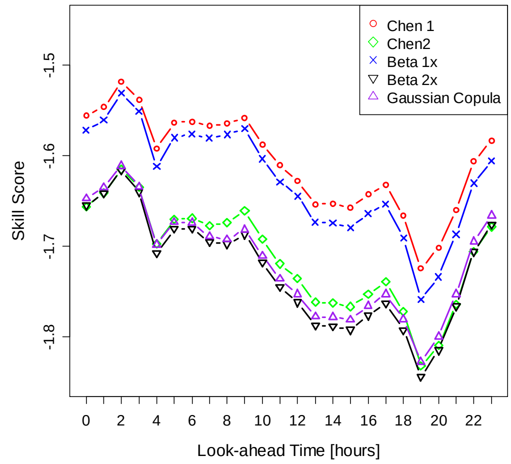

Fig. 8. Skill score diagram for (h_{Power};$h_{WindSpeed}$) = (0.004; 1).

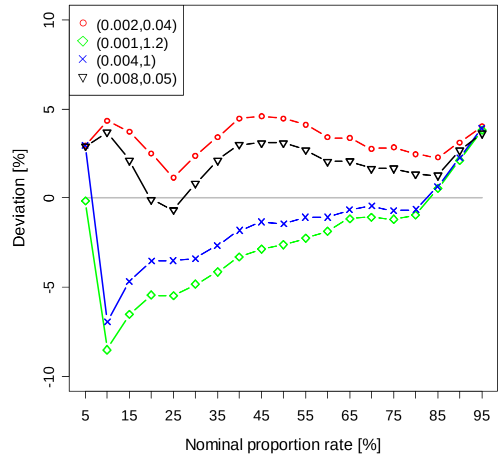

Fig. 9. Calibration diagram for different kernel bandwidths (h_{Power};$h_{WindSpeed}$).

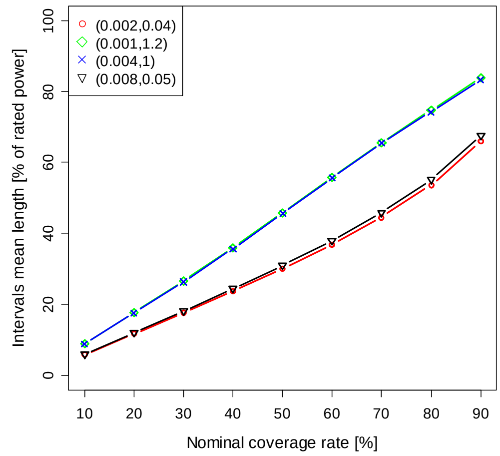

Fig. 10. Sharpness diagram for different kernel bandwidths (h_{Power};$h_{WindSpeed}$).

D. Results: Different Inputs and 42hrs Ahead Forecasts

In this section the impact of different input variables forecasted will be tested. The following combination of input variables were considered and compared: M0) wind speed; M1) wind speed + direction; M2) wind speed + hour of the day; M3) wind speed + look-ahead time step; M4) wind speed + direction + hour of the day; M5) wind speed + direction + look-ahead time step.

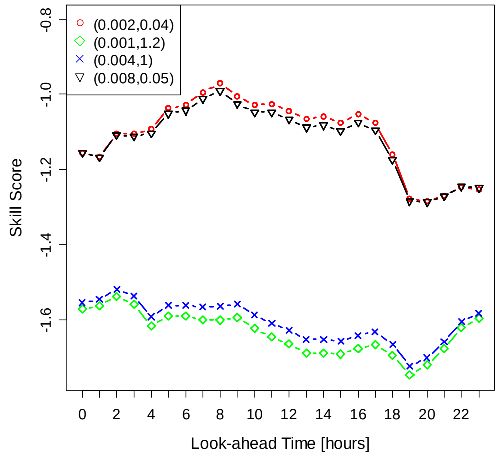

Fig. 11. Skill score diagram for different kernel bandwidths (h_{Power};$h_{WindSpeed}$).

The kernel bandwidth values were determined experimentally (trial-error) and using as starting point the values suggested by the function “cde.bandwidths” from the R package “hdrcde” [20].

The following kernel functions were used in the Quantile-Copula (QC) estimator:

- Wind power and wind speed: Chen’s beta kernel from Eq. 11 with a bandwidth equal to 0.008;
- Wind direction and hour of the day: von Mises distribution from Eq. 14 with a bandwidth equal to 1.0;
- Look-ahead time step: Chen’s beta kernel from Eq. 11 with a bandwidth equal to 0.2.

Fig. 12 and 13 depict the calibration and skill score correspondingly, obtained with the five different models. The best performance is from models M3 (wind speed + look-ahead time step) and M2 (wind speed + hour of the day) for both studies. The performance of all methods is rather similar, and there are no significant differences.

Note that the inclusion of wind direction in the model (M1) only improves slightly the calibration performance for some quantiles. This could be justified by the fact that the wind farm is located in flat terrain, and this could mitigate the impact of wind direction on the uncertainty forecast.

The results in terms of sharpness are also rather similar and consequently are not depicted in this paper.

Fig. 14-16 present the comparison between model M3 for QC and the splines QR estimator. Note that the calibration is presented for quantiles between 1% and 5% in 1% steps, then from 5% to 95% in 5% steps, and finally from 95% to 99% in 1% steps.

The results with splines quantile regression were obtained with a spline basis with 8 degrees of freedom (determined by trial-error). The wind direction was modeled with a periodic cubic spline basis with equidistant knots. This is done by the S-PLUS/R functions “pb.bse”, “pb.h” and “bint0” available from [21].

In terms of calibration, QC presents the best overall calibration performance. QR presents the best calibration for the left tail, while QC is better in the right tail. QR presents better performance in sharpness.

QC has almost the same performance as QR in terms of skill score. QR is better than QC for some look-ahead steps, but it is also worse in others.

These results confirm, as mentioned in [7], that methods with better calibration present a worse performance in terms of sharpness.

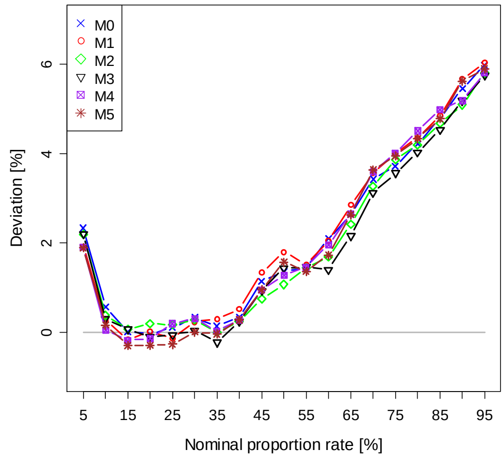

Fig. 12. Calibration diagram for models M0-M5.

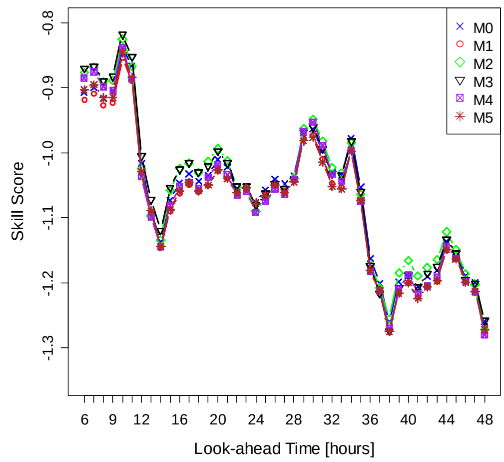

Fig. 13. Skill score diagram for models M0-M5.

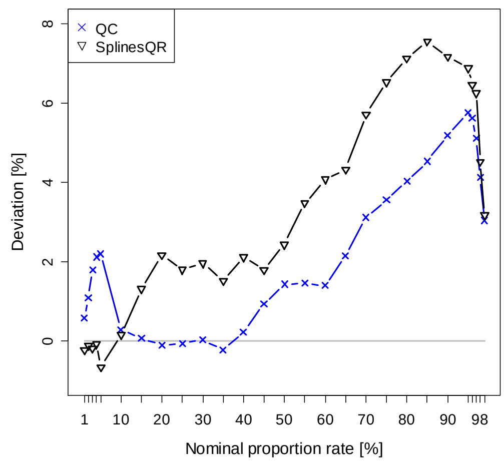

Fig. 14. Calibration diagram for QC and splines QR models.

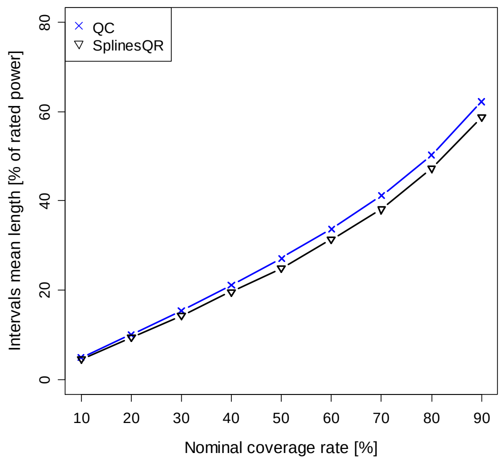

Fig. 15. Sharpness diagram for QC and splines QR models.

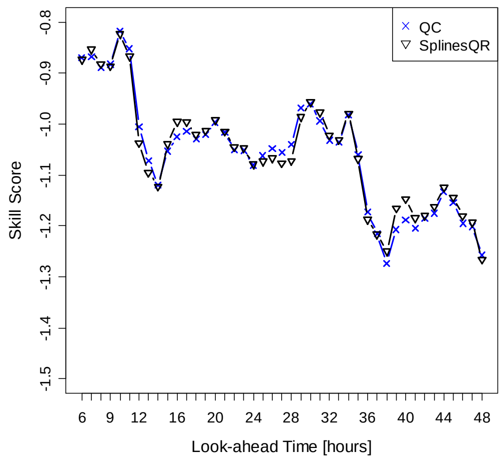

Fig. 16. Skill score diagram for QC and splines QR models.

IV. CONCLUSIONS

This paper presents results of using the quantile-copula conditional density estimator in the wind power uncertainty forecasting problem, and how to select the adequate kernel for modeling the different variables of the problem.

From the results obtained for a real wind farm located in the United States, it is possible to conclude that the quantile-copula method leads to a better calibration performance, while the splines quantile regression methods present the best sharpness performance. The skill score performance is rather similar for both approaches.

We only assess the *quality* of the probabilistic forecast in this paper, and only evaluate the correspondence between the forecasts and the reality. No considerations are made about the impact on decision-making problems, i.e. the forecast *value*.

The evaluation of the probabilistic forecast *value* can solely be done with the use of probabilistic forecasts in a specific decision-making problem and subsequent evaluation of the decision's quality. Nevertheless, the *quality* metrics used in this paper will impact the forecast *value*, but the link between quality and value will differ from problem to problem.

For instance, the calibration metric is particularly important for the wind power bidding problem. As proven in [22] for convex loss functions, the optimal solution that minimizes the expected loss is a quantile related with an asymmetry parameter that reflects the possibly distinct and different costs of underforecast and overforecast. Hence, in electricity markets with different penalty factors for forecast error signs, the optimal quantile differs from the median and needs to be computed to derive the optimal wind power bid [23].

Under this situation, a probabilistic bias (translated by the calibration performance) may lead to an under or overestimation of the optimal quantile, and consequently to a lower income than expected. Hence, only a good calibration guarantees the accuracy in the optimal quantile computation. The sharpness is not very relevant for this problem.

Having said this, the results presented in this paper show evidence that the QC estimator improves the calibration performance compared to the QR approach. This can increase the forecasting *value*, e.g. in the wind power bidding problem. Moreover, a KDF is characterized by a *pdf*, which is a very flexible uncertainty representation. Hence, such forecasts can be tailored for use in a range of decision problems within the wind power and power system domains.

V. ACKNOWLEDGMENT

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (''Argonne''). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up non-exclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

The authors acknowledge Horizon Wind Energy for providing the wind farm data used in the analysis.

VI. REFERENCES

[1] C. Monteiro, R.J. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, "Wind power forecasting: state-of-the-art 2009," Report ANL/DIS-10-1, Argonne National Laboratory, 2009. http://www.dis.anl.gov/projects/windpowerforecasting.html

[2] H.A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9(1-2), pp. 95-108, 2006.

[3] R.J. Bessa, V. Miranda, and J. Gama, "Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting," *IEEE Transactions on Power Systems*, vol. 24(4). Nov. 2009.

[4] A. Botterud, J. Wang, R.J. Bessa, H. Keko, and V. Miranda, "Risk management and optimal bidding for a wind power producer," in *Proceedings of the IEEE PES General Meeting*, Minneapolis, USA, 2010.

[5] M.A. Matos and R.J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Transactions on Power Systems*, 2011, *In Press*. DOI: 10.1109/TPWRS.2010.2065818

[6] J. Usaola, "Probabilistic Load Flow with Correlated Wind Power Injections", *Electric Power Systems Research*, vol. 80(5), pp. 528-536, May 2010.

[7] J. Juban, N. Siebert, and G. Kariniotakis, "Probabilistic short-term wind power forecasting for the optimal management of wind generation," in

Proceedings of the IEEE PowerTech Conference, Lausanne, Switzerland, July 2007.
[8] M. Rosenblatt, *Conditional probability density and regression estimators*, in Multivariate Analysis II, New York: Academic Press, pp. 25-31, 1969.
[9] O.P. Faugeras, “A quantile-copula approach to conditional density estimation,” *Journal of Multivariate Analysis*, vol. 100, pp. 2083-2099, Oct. 2009.
[10] M.P. Wand and M.C. Jones, “Multivariate plug-in bandwidth selection,” *Computational Statistics*, vol. 9, pp. 97-116, 1994.
[11] M. Rosenblatt, “Remarks on some nonparametric estimates of a density function,” *The Annals of Mathematical Statistics*, vol. 27(3), pp. 832-837, 1956.
[12] T. Bouezmarni and J. V. K. Rombouts, “Semiparametric multivariate density estimation for positive data using copulas,” *Computational Statistics & Data Analysis*, vol. 53(6), pp. 2040-2054, Apr. 2009.
[13] M. Sklar, “Fonctions de répartition à n dimensions et leurs marges,” *Publications de l’Institut de Statistique de L’ Université de Paris*, vol. 8, pp. 229–231, 1959.
[14] R. Nelson, *An Introduction to Copulas. Lecture Notes in Statistics*, New York: Springer, vol. 139, 1999.
[15] S.X. Chen, “Beta kernel estimators for density functions,” *Computational Statistics & Data Analysis*, vol. 31(2), pp. 131-145, 1999.
[16] M.C. Jones and D.A. Henderson, “Miscellanea kernel-type density estimation on the unit interval,” *Biometrika*, *Oxford University Press for Biometrika Trust*, vol. 94(4), pp. 977-984, 2007.
[17] K.V. Mardia and P.E. Jupp, *Directional Statistics*, Wiley: New York, 2000.
[18] W. Skamarock, J, B. Klemp, J. Dudhia, et al., “A Description of the Advanced Research WRF Version 2”, NCAR/TN–468+STR Technical note, Jun. 2005.
[19] P. Pinson, H.A. Nielsen, J.K. Moller, H. Madsen, and G. Kariniotakis, “Nonparametric probabilistic forecasts of wind power: required properties and evaluation,” Wind Energy, vol. 10(6), pp. 497–516, Nov. 2007.
[20] R. J. Hyndman, J. Einbeck and M. Wand, “Package ‘hdrcde’,” R Manual, April 2010. (www.robjhyndman.com/software/hdrcde)
[21] http://www.imm.dtu.dk/~han/pub (accessed on Feb. 2011).
[22] T. Gneiting, “Quantiles as optimal point forecasts,” *International Journal of Forecasting*, vol. 27(2), pp. 197-207, Apr.-June 2011.
[23] J.B. Bremnes, “Probabilistic wind power forecasts using local quantile regression,” *Wind Energy*, vol. 7, pp. 47-54, 2004.

VII. Biographies

Ricardo J. Bessa received his Licenciado (five years) degree from the Faculty of Engineering of the University of Porto, Portugal (FEUP) in 2006 in Electrical and Computer Engineering. In 2008 he received his Master degree in Data Analysis and Decision Support Systems on the Faculty of Economics of the University of Porto (FEP). Currently, he is a researcher at INESC Porto in its Power Systems Unit and a PhD student of the Doctoral program in Sustainable Energy Systems (MIT Portugal) at FEUP. His research interests include wind power forecasting, electric vehicles, data mining and decision-aid methods.

Joana Mendes received the Licenciado degree in Physics from the Faculty of Sciences of the University of Porto, Porto, Portugal (FCUP), in 2008, and the M.Sc. degree in Applied Meteorology from the University of Reading, England, UK, in 2009. She is now a Researcher at INESC Porto in its Power Systems Unit.

Vladimiro Miranda (M’90–SM’04–F’05) received the Licenciado, Ph.D., and Agregado degrees from the Faculty of Engineering of the University of Porto, Porto, Portugal (FEUP), in 1977, 1982, and 1991, respectively, all in electrical engineering. In 1981, he joined FEUP and currently holds the position of Professor Catedrático. He is also currently a Director of INESC Porto, an advanced research institute in Portugal. He has authored many papers and been responsible for many projects in areas related with the application of computational intelligence to power systems.

Audun Botterud is an Energy Systems Engineer in the Center for Energy, Environmental, and Economic Systems (CEEESA) at Argonne National Laboratory. He has a M.Sc. in Industrial Engineering (1997) and a Ph.D. in Electrical Power Engineering (2003), both from the Norwegian University of Science and Technology (NTNU). He was previously with SINTEF Energy Research in Trondheim, Norway, and has also spent one year and a half as a visitor at the MIT Laboratory for Energy and the Environment of Cambridge, Massachusetts, USA. His research interests include power systems planning and economics, renewable energy systems, wind power systems, wind power integration, stochastic optimization, and agent-based modeling.

Jianhui Wang received his B.Sc. degree in Management Science and Engineering and a M.Sc. degree in Technical Economics and management from North China Power University, China, in 2001 and 2004, respectively, and his Ph.D. in Electrical Engineering from Illinois Institute of Technology, USA, in 2007. Presently, he is an Assistant Computational Engineer – Energy Systems in CEEESA. His research interests include energy economics and policy, agent-based modeling and simulation, and electrical power systems optimization and economics. He is chair of IEEE PES power system operation methods subcommittee and co-chair of the task force on integration of wind and solar power into power system operations.

Zhi Zhou received his Ph.D. in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute, USA, in 2010. Presently, he is a postdoctoral appointee at the Decision and Information Sciences (DIS) Division at Argonne National Laboratory (Argonne). His research interests include agent-based modeling and simulation, electricity markets, and renewable energy.