Downloaded from orbit.dtu.dk on: Apr 20, 2024

DTU Library

On the quality and value of probabilistic forecasts of wind generation

Pinson, Pierre; Juban, Jeremie; Kariniotakis, Georges

Published in:
International Conference on Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006.

Link to article, DOI:
10.1109/PMAPS.2006.360290

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Pinson, P., Juban, J., & Kariniotakis, G. (2006). On the quality and value of probabilistic forecasts of wind generation. In *International Conference on Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006.* IEEE. https://doi.org/10.1109/PMAPS.2006.360290

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

9th International Conference on Probabilistic Methods Applied to Power Systems
KTH, Stockholm, Sweden – June 11-15, 2006

On the Quality and Value of Probabilistic Forecasts of Wind Generation

P. Pinson, J. Juban and G. N. Kariniotakis, *Member, IEEE*

(Invited Paper)

***Abstract*— While most of the current forecasting methods provide single estimates of future wind generation, some methods now allow one to have probabilistic predictions of wind power. They are often given in the form of prediction intervals or quantile forecasts. Such forecasts, since they include the uncertainty information, can be seen as optimal for the management or trading of wind generation. This paper explores the differences and relations between the quality (i.e. statistical performance) and the operational value of these forecasts. An application is presented on the use of probabilistic predictions for bidding in a European electricity market. The benefits of a probabilistic view of wind power forecasting are clearly demonstrated.**

***Index Terms*— Wind power, probabilistic forecasting, evaluation methods, decision-making, operational value.**

I. Introduction

WIND POWER is the fastest-growing renewable electricity-generating technology. The targets for the next decades aim at high share of wind power in electricity generation in Europe [1] (up to 75GW for 2010). However, such a large-scale integration of wind generation capacities induces difficulties in the management of a power system. Also, the deregulation of European electricity markets makes that wind generation is expected to be traded through electricity pools, in which energy bids are settled several hours before actual delivery. Imbalances lead to financial penalties that may substantially tighten the revenue of wind power producers. Owing to these two reasons, predictions of wind generation up to 48 hours ahead contribute to a secure and economic power system operation. Increasing the value of wind generation through the improvement of prediction systems' performance is considered as one of the priorities in wind energy research needs for the coming years [2].

So far, most wind power prediction tools give an estimate of the future power generation without addressing the issue of the forecast uncertainty [3]: we usually say that such prediction systems provide the 'most likely outcome' for each look-ahead time. Given the significant variability of the level of forecasting errors, end-users have expressed the need for an on-line estimation of forecast uncertainty. This additional information would then be introduced in decision-making processes. Such an information can indeed be given by prediction risk indices that tell what may be the expected level of forecasting error [4], or alternatively by quantile or interval forecasts, which translates to quoting particular points of predictive distributions of wind power (see [5]–[7] among others). In the present paper, we leave aside the case of risk indices and concentrate on probabilistic predictions.

Here, our aim is to describe what may be the *quality* and *value* of probabilistic forecasts of wind generation. These two concepts have been introduced in the meteorological forecasting literature by Murphy [8]: the quality aspect stands for the statistical performance of the probabilistic forecasts, evaluated with a set of measures and diagrams, while the value aspect corresponds to the increased benefits (economical or not) from the use of these probabilistic forecasts in an operational context. We base the present study on uncertainty estimates provided by a method previously developed by the authors [4] (i.e. adapted resampling), which can be considered for post-processing state-of-the-art wind power point prediction methods. The test case consists in a multi-MW wind farm located in Denmark, for which probabilistic forecasts are obtained by applying the adapted resampling approach to point predictions given by three different methods, over a period of one year. Regarding the decision-making process, we concentrate on simulating the participation of the operator of the considered wind farm in the Dutch electricity market, composed by the day-ahead pool APX (standing for Amsterdam Power eXchange) and by the real-time regulation market run by TenneT, the Transmission System Operator (TSO) for the Netherlands. We compare several bidding strategies, based on point predictions only, or alternatively on more advanced strategies derived from probabilistic forecasts.

II. Probabilistic forecasts of wind generation and their required properties

A. Definitions

In this paper, we consider that wind generation for a given time t can be seen as a random variable p_t. Hence, the measured power output p_t^* is a realization of p_t. Denote by f_t^p the probability density function of that random variable. If a point forecasting method is developed for providing point predictions that would minimize a mean square error criterion, the resulting point forecast $\hat{p}_{t+k/t}$ made at time t for lead time $t+k$ corresponds to an estimate of the expectation of p_{t+k}. Such an information could be seen as sufficient for the management or trading of wind generation. But, associated uncertainty estimates would permit to optimize decisions resulting from the use of these predictions. Ideally, an estimate $\hat{f}^p_{t+k/t}$ of f^p_{t+k} would give the whole information on what *could* be the wind generation at time $t+k$. Today,

P. Pinson is with the Informatics and Mathematical Modeling department of the Technical University of Denmark, Richard Petersens Plads, DK-2800 Kgs. Lyngby, Denmark (e-mail: pp@imm.dtu.dk). J. Juban and G. Kariniotakis are with the Center for Energy and Processes of École des Mines de Paris, B.P. 207, F-06904 Sophia-Antipolis, France. (e-mail: jeremie.juban@ensmp.fr; georges.kariniotakis@ensmp.fr).

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

probabilistic forecasts of wind generation are given in the form of quantile or interval forecasts, produced with statistical post-processing methods [4], [6], or from wind power ensembles [5]. Fig. 1 depicts an episode with hourly point predictions for the following 48 hours, given by a state-of-the-art forecasting approach, compared to the measured power values. The prediction intervals, which are estimated consequently by the adapted resampling method, are shown here in the form of a fan chart. Hereafter, all variables are normalized by the nominal power P_n of the considered wind farm.

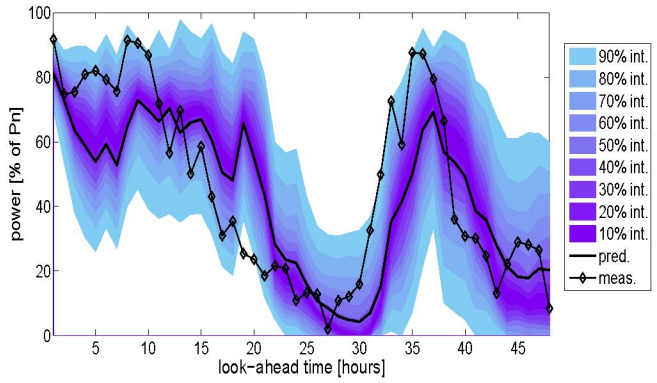

Fig. 1. Example of a 48-hour point prediction of wind generation, associated with a set of interval forecasts. The point predictions are given by a state-of-the-art method and interval forecasts are estimated consequently with the adapted resampling approach.

Prediction intervals give the range of possible values within which the true effect is expected to lie with a pre-assigned probability, known as their nominal coverage rate. Here we concentrate on central prediction intervals $\hat{I}^{(\alpha)}_{t+k/t}$, estimated at time t for lead time $t+k$ with a nominal coverage rate $(1-\alpha)$, the bounds of which correspond to the $(\alpha/2)$ and $(1-\alpha/2)$ quantiles of the predictive distribution $\hat{f}^{p}_{t+k/t}$ of expected power generation at that lead time:

$$\hat{I}^{(\alpha)}_{t+k/t} \equiv [\hat{r}^{(\alpha/2)}_{t+k/t}, \hat{r}^{(1-\alpha/2)}_{t+k/t}]. \quad (1)$$

Estimating at once a sequence of n interval forecasts with various nominal coverage rates varying over all values in the unit interval permits to construct predictive distributions of wind generation. Denote by $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ the set of nominal coverage rates (with the α_i in the ascending order). Probabilistic wind power forecasts made at time t for lead time $t+k$ are then given by the set of corresponding $2n$ predictive quantiles:

$$\hat{f}^{p}_{t+k/t} \equiv \{\hat{r}^{(\alpha_n/2)}_{t+k/t}, \hat{r}^{(\alpha_{n-1}/2)}_{t+k/t}, \ldots, \hat{r}^{(\alpha_1/2)}_{t+k/t}, \hat{r}^{(1-\alpha_1/2)}_{t+k/t}, \ldots, \hat{r}^{(1-\alpha_{n-1}/2)}_{t+k/t}, \hat{r}^{(1-\alpha_n/2)}_{t+k/t}\}, \quad (2)$$

which can also be written as

$$\hat{f}^{p}_{t+k/t} \equiv \{\hat{r}^{(\tilde{\alpha}_1)}_{t+k/t}, \hat{r}^{(\tilde{\alpha}_2)}_{t+k/t}, \ldots, \hat{r}^{(\tilde{\alpha}_{2n})}_{t+k/t}\}. \quad (3)$$

B. Required properties

The first requirement for interval forecasts is that their measured coverage should be close to the nominal one. Actually, if considering infinite series of interval forecasts, that empirical coverage should exactly equal the pre-assigned probability. That first property is referred to as 'reliability' or 'calibration' in the forecasting literature.

Besides this first requirement, it is necessary that prediction intervals provide a situation-dependent assessment of the forecast uncertainty. Their size should then vary depending on various external conditions. For the example of wind prediction, it is intuitively expected that prediction intervals (for a given nominal coverage rate) should not have the same size when predicted wind speed equals zero and when it is near cut-off speed. The most simple type of intervals is constant-size intervals (e.g. produced from climatology). Advanced methods for their estimation are expected to produce variable-size intervals. This property is commonly named 'sharpness' or 'resolution'. Note that here, we will introduce a nuance between sharpness and resolution: the former will relate to the average size of intervals while the latter is associated to their size variability.

Actually, the traditional view of interval forecast evaluation, which mainly comes from the econometric forecasting community, is based on the testing of correct conditional coverage. This means intervals have to be unconditionally reliable, and independent (see [9] for instance). In the case of wind power forecasting, we know there exists a correlation among forecasting errors (at least for short time-lags). Thus, we do not expect prediction intervals to be independent. Then, it appears preferable to develop an evaluation framework that is based on an alternative paradigm. We propose to consider reliability as a primary requirement and then sharpness and resolution as an added value. It should be noted here that reliability can be increased by using some re-calibration methods (e.g. conditional parametric models [5] or smoothed bootstrap [10]), while sharpness/resolution cannot be enhanced with post-processing procedures. This second aspect is the inherent (and invariant) ability of a probabilistic forecasting method to distinctly resolve future events [11].

III. Evaluating the Quality of Wind Power Probabilistic Forecasts

A. Focus on reliability

Firstly, let us introduce the indicator variable $\mathcal{I}^{(\tilde{\alpha})}_{t,k}$, which is defined for a quantile forecast $\hat{r}^{(\tilde{\alpha})}_{t+k/t}$ made at time t and for horizon k as follows:

$$\mathcal{I}^{(\tilde{\alpha})}_{t,k} = \begin{cases} 1, & \text{if } p_{t+k} \leq \hat{r}^{(\tilde{\alpha})}_{t+k/t} \\ 0, & \text{otherwise} \end{cases}. \quad (4)$$

Consequently, denote by $n^{(\tilde{\alpha})}_{k,1}$ the sum of hits and $n^{(\tilde{\alpha})}_{k,0}$ the sum of misses (for a given horizon k) over the N realizations:

$$n^{(\tilde{\alpha})}_{k,1} = \#\{\mathcal{I}^{(\tilde{\alpha})}_{t,k} = 1\} = \sum_{t=1}^{N} \mathcal{I}^{(\tilde{\alpha})}_{t,k}, \quad (5)$$

$$n^{(\tilde{\alpha})}_{k,0} = \#\{\mathcal{I}^{(\tilde{\alpha})}_{t,k} = 0\} = N - n^{(\tilde{\alpha})}_{k,1}. \quad (6)$$

The easiest way to check the calibration of probabilistic forecasts is to compare the empirical coverages of the various quantiles to the nominal ones (i.e. the required probability

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

$\tilde{\alpha}_i$). An estimation $\hat{a}_k^{(\tilde{\alpha})}$ of the actual coverage $a_k^{(\tilde{\alpha})}$, for a given horizon k, is obtained by calculating the mean of the $\{\mathcal{I}_{t,k}^{(\tilde{\alpha})}\}_{t=1,\ldots,N}$ time-series over the test set:

$$\hat{a}_k^{(\tilde{\alpha})} = \frac{1}{N}\sum_{t=1}^{t=N}\mathcal{I}_{t,k}^{(\tilde{\alpha})} = \frac{n_{k,1}^{(\tilde{\alpha})}}{n_{k,0}^{(\tilde{\alpha})}+n_{k,1}^{(\tilde{\alpha})}}. \quad (7)$$

This standard approach to the evaluation of prediction intervals was proposed by Ballie et al. [12] and by McNees [13]. This is the idea used in reliability diagrams which give the empirical coverage versus the nominal coverage for various nominal coverage rates. The closer to the diagonal the better. They can alternatively be depicted as the deviation from the 'perfect reliability' case for which empirical coverage would equal the nominal one (calculated as the difference between these two quantities). This idea is similar to the use of Probability Integral Transform (PIT) histograms as proposed in [14], except that reliability diagrams directly provide that additional information about the magnitude of the deviation from the 'perfect reliability' case.

B. Focus on sharpness/resolution

When dealing with sharpness or resolution, focus is given to the size of the prediction intervals. Let us define

$$\delta_{t,k}^{(\alpha)} = \hat{r}_{t+k/t}^{(1-\alpha/2)} - \hat{r}_{t+k/t}^{(\alpha/2)} \quad (8)$$

the size of the central interval forecast $\hat{I}_{t+k/t}^{(\alpha)}$.

If two uncertainty estimation methods provide intervals at an acceptable level of reliability, it is the method that yields the narrowest intervals that is to be preferred. Hence, we relate the sharpness aspect to the average size $\bar{\delta}_k^{(\alpha)}$ of the prediction intervals for a given horizon k:

$$\bar{\delta}_k^{(\alpha)} = \frac{1}{N}\sum_{t=1}^{N}\delta_{t,k}^{(\alpha)} = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{r}_{t+k/t}^{(1-\alpha/2)} - \hat{r}_{t+k/t}^{(\alpha/2)}\right). \quad (9)$$

In parallel, the resolution concept is standing for the ability of providing a situation-dependent assessment of the uncertainty. It is not possible to directly verify that property, though we may study the variation in size of the intervals. The standard deviation $\sigma_k^{(\alpha)}$ of the interval size (for a given horizon k and nominal coverage rate $(1-\alpha)$) calculated as

$$\sigma_k^{(\alpha)} = \left(\frac{1}{N-1}\sum_{t=1}^{N}\left(\delta_{t,k}^{(\alpha)} - \bar{\delta}_k^{(\alpha)}\right)^2\right)^{\frac{1}{2}} \quad (10)$$

provides that information. Because of the non-linear and conditionally heteroskedastic nature of the wind generation process, the forecast uncertainty is highly variable and it is thus expected that the interval size also greatly varies. The way both $\bar{\delta}_k^{(\alpha)}$ and $\sigma_k^{(\alpha)}$ can be considered for evaluating the quality of predictive distributions is described in [15].

Alternatively, one may envisage to use skill scores for evaluating the sharpness and resolution of probabilistic forecasts. Skill scores are obtained from scoring rules that associate a single numerical value $S(\hat{f}_{t+k/t}^{p}, p_{t+k}^{*})$ to a predictive distribution $\hat{f}_{t+k/t}^{p}$ if the event p_{t+k}^{*} materializes. A scoring rule should reward a forecaster that expresses his true beliefs. It is said to be proper if it does so. Gneiting et al. [16] recently showed that any scoring rule of the form

$$S(\hat{f}_{t+k/t}^{p}, p_{t+k}^{*}) = \sum_{i=1}^{2n}\big(\alpha_i s_i(\hat{r}_{t+k/t}^{(\tilde{\alpha}_i)}) + (s_i(p_{t+k}^{*}) - s_i(\hat{r}_{t+k/t}^{(\tilde{\alpha}_i)}))\mathcal{I}_{t,k}^{(\tilde{\alpha}_i)} + h(p_{t+k}^{*})\big) \quad (11)$$

with $\mathcal{I}_{t,k}^{(\tilde{\alpha}_i)}$ the indicator variable (for the quantile with proportion $\tilde{\alpha}_i$) introduced above, s_i non-decreasing functions and h arbitrary, was proper for evaluating predictive distributions expressed as a set of quantiles. The final score value is obtained by averaging the score values for all predictive distributions over the evaluation set, as a function of the look-ahead time k. Here S is a positively-oriented score: a higher score value stands for a higher skill.

A unique score does not tell what are the contributions of reliability or sharpness/resolution to the skill of a given probabilistic forecasting method, because it encompasses all the aspects of probabilistic forecast evaluation (see discussion in [15]). Though, if reliability is assessed in a first stage, then relying on skill scores such as the one given by Eq. (11) allows one to compare the overall skill of rival methods.

C. Evaluation of the quality of the probabilistic forecasts obtained by adapted resampling

Adapted resampling is a statistical distribution-free approach to the estimation of quantiles and interval forecasts suitable for non-linear and bounded processes [4]. It serves to post-process point forecasting methods and enhances them with probabilistic forecasts of wind generation. This approach has been applied and evaluated on a variety of case-studies consisting of periods ranging from several months to several years for various European wind farms. Here, adapted resampling is applied for post-processing point predictions provided by 3 methods from different research centers in Europe (denoted by M1, M2, and M3). The evaluation period is composed by 8760 series of 43-hour ahead hourly predictions. Predictive distributions are built by estimating 9 central predictions intervals with nominal coverage rates 10, 20,..., and 90%. Fig. 1 gives an example of such series of probabilistic forecasts of wind generation. Regarding the settings of the method, we follow the sensitivity analysis carried out in [15]: we set the sample size to 300 items, the number of bootstrap replications to 50, and we use 5 triangular fuzzy sets for mapping the range of power values.

For assessing in a first stage the reliability of the three series of probabilistic predictions, we use reliability diagrams which give the deviations from 'perfect reliability', as a function of the nominal coverage rates of the quantiles (Fig. 2). Deviation values are the averages over the 43 look-ahead times. The figures in the legend give the average absolute deviations over the range of nominal coverage rates. One sees from Fig. 2 that the deviations from nominal coverage are comprised between -1.8 and 0.3% whatever the point forecasts used as input. Owing to these low deviations from 'perfect reliability', we assume that the three sets of probabilistic forecasts are

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

well-calibrated. However, there is a difference in the overall reliability of predictive distributions: the ones produced from the M1 forecasts are more reliable than the two others. Predictive quantiles resulting from M2 and M3 forecasts are slightly underestimated.

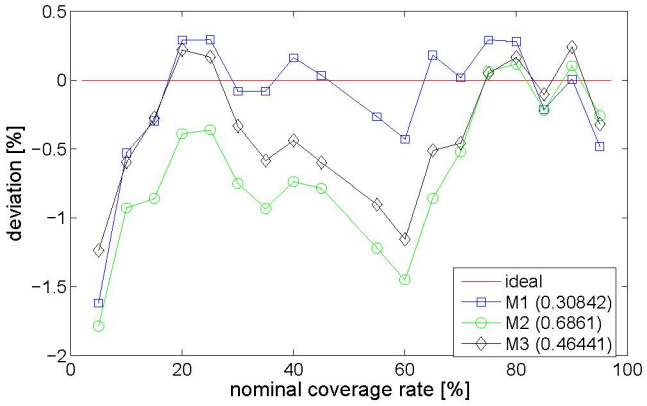

Fig. 2. Reliability diagrams for evaluating the reliability of the three series of probabilistic forecasts of wind generation. Figures in the legend give the average absolute deviation from 'perfect reliability' over the range of nominal coverage rates.

In a second stage, because we have assumed the three sets of probabilistic forecasts to be well-calibrated, we turn our attention to the use of the skill score introduced above (Eq. (11)), for assessing their overall skill. We put $s_i(p) = 4p$, $(i = 1, \ldots, 2n)$, and $h(p) = -2p$, following [16]. Fig. 3 depicts the evolution of the resulting score, as a function of the look ahead time. Figures in the legend give the average score values over the range of horizons. There is a trend that the score values decrease for longer horizons. This meets the general statement that it is harder to forecast events that are further in the future. One notices that the skill of the three sets of probabilistic forecasts is rather close. But, even if the average score values tell that M1 predictive distributions have higher skill than the two others on average, one sees that the best and worst predictive distributions are not obviously the same over the whole range of horizons. Also, although M3 probabilistic forecasts are more reliable than the M2 ones, the overall skill of the latter is slightly higher. This is because predictive distributions obtained from M2 point predictions prove to be sharper.

IV. Evaluating the value of wind power probabilistic forecasts

For studying the operational value of predictive distributions of wind generation, we turn our attention to the simulation of the participation of the operator of the considered wind farm in the Dutch electricity market over the year 2002.

A. Assumptions and problem formulation

Each electricity pool has its own rules and regulations that determine the way energy is to be sold or purchased, how the market prices are settled, and the obligations that the participants (producers or consumers) are committed to.

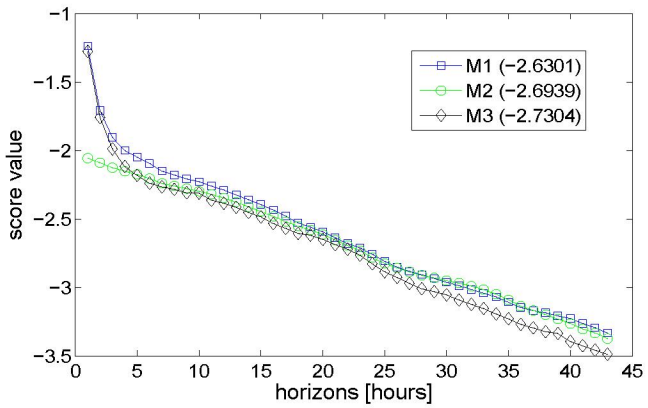

Fig. 3. Skill score values as a function of the prediction horizon for the three sets of probabilistic forecasts. Figures in the legend give the average score values over the range of look-ahead times.

Here, we consider that all energy producers participate in the electricity market under the same rules, i.e. they have to propose their bids on the day-ahead market, and they are then financially responsible of their deviations from schedule. The costs of keeping the balance are charged to the participants, proportionally to their imbalance. A description of APX and TenneT markets is given in [17]. In the Netherlands, bids are to be given before 10:30 for the following day from midnight to midnight. Relevant predictions horizons are then between 14 and 38 hours ahead (if the last available forecast is provided at 10:00). The length of the Program Time Unit (PTU) on APX is of 1 hour. But then, the PTU length on the regulation market run by Tennet is of 15 minutes only. Therefore, we consider that bids proposed on APX are composed by four equal amounts of energy for each of the TenneT PTUs composing an APX trading hour. Wind energy predictions are obtained by integrating the power forecasts over the related PTUs.

A crucial assumption for this study is that the wind power producer is a price-taker in the Dutch electricity market. This means that the amounts of energy he proposes on the market cannot impact the market clearing price π_k^c. We formulate the same assumption regarding the regulation market: the wind farm operator is a small entity on that market and have no influence on imbalance prices.

Whatever the considered electricity market, the revenue R_k for a given PTU k of a participant proposing an amount of energy E_k^c but actually generating E_{t+k}^* can be formulated as

$$R_k(E_k^c, E_{t+k}^*) = \pi_k^c E_{t+k}^* - T_k^*(E_k^c, E_{t+k}^*). \qquad (12)$$

This revenue is composed by the income resulting from the selling of actual wind generation at the spot price, minus the cost $T_k^*(E_k^c, E_{t+k}^*)$ associated to the deviation from contract d_k defined as

$$d_k = E_{t+k}^* - E_k^c. \qquad (13)$$

In turn, the regulation costs are a function of the spot price π_k^c, the regulation unit costs $\pi_k^{*,+}$ and $\pi_k^{*,-}$ for downward and upward dispatch respectively, and the amount of energy produced in imbalance d_k:

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

$$T_k^*(E_k^c, E_{t+k}^*) = \begin{cases} \pi_k^{*,+} d_k & , \ d_k \geq 0 \\ \pi_k^{*,-} d_k & , \ d_k < 0 \end{cases} . \qquad (14)$$

B. Definition of bidding strategies

Different types of bidding strategies may be defined depending on the available information about future wind generation. If only point predictions are provided by a forecasting method, then these point predictions appear as the best bids one may propose on the electricity pool [18]. However, if the market participant has predictive distributions, he can develop more advanced strategies, reflecting either his aim to maximize his revenue over a certain period of time or a risk aversion for large regulation costs [19], [20]. Here we concentrate on the first type of strategies, referred to as Probabilistic Choice (PC) strategies.

Applying a PC strategy directly translates to maximizing the expectation of the revenue for each PTU k, formulated by Eq. (12). Because in that formulation all wind generation is sold at the spot price, and the level of contracted energy only influences the regulation costs, an optimal energy bid $\tilde{E}_k^c$ is determined by minimizing the expectation of these costs:

$$\tilde{E}_k^c = \underset{E_k^c}{\arg\min} \, \mathbb{E}\left[T_k^*(E_k^c, E_{t+k})\right], \qquad (15)$$

where this expectation can be written as

$$\begin{aligned} \mathbb{E}\left[T_k^*(E_k^c, E_{t+k})\right] &= \int_0^{E_k^c} \pi_k^{*,-}(E_k^c - x) f_{t+k}^E(x) dx \\ &+ \int_{E_k^c}^1 \pi_k^{*,+}(x - E_k^c) f_{t+k}^E(x) dx. \qquad (16) \end{aligned}$$

Because neither the regulation unit costs nor the probability distributions of future wind generation can be known when proposing bids on the electricity pool, it is necessary to consider probabilistic forecasts $\hat{f}_{t+k/t}^E$ as well as a model of the participant's sensitivity to regulation costs. This latter may be derived from estimates of regulation unit costs only, or more generally by modeling the way the wind power producer will face deviations from contract in the most cost-efficient manner, if he has the possibility to couple his production with conventional generation, or to use storage devices. Here, we consider regulation penalties only in this model. The way the optimization problem given by Eq. (15) integrates these two information, and is consequently solved, is described in [20].

C. Results and discussion

Table I gathers the quarterly averages of both the market clearing prices and regulation unit costs for upward and downward dispatch over 2002. The average spot price on APX proves to be significantly variable throughout the year, being 4 times higher during the third quarter (corresponding to the summer period) than during the first one (corresponding in turn to the winter period). The average market clearing price over 2002 is equal to 29.99€/MWh. In parallel, the regulation market behavior exhibits also substantial variations. Although the average unit costs for upward and downward dispatch are rather similar for the last two quarters, the latter is much higher than the former over the first six months. One notices that the unit cost for positive deviations from contract is on average higher than the unit cost for negative deviations. The average values for $\pi_k^{*,-}$ and $\pi_k^{*,+}$ over 2002 are of 4.03 and 10.93€/MWh respectively. Hence, it would appear preferable to overestimate expected wind generation.

TABLE I
DUTCH ELECTRICITY MARKET CHARACTERISTICS FOR 2002: QUARTERLY AVERAGES OF SPOT PRICES AND REGULATION PENALTIES FOR UPWARD AND DOWNWARD DISPATCH.

Quarter	$\bar{\pi}_k^c$ [€/MWh]	$\bar{\pi}_k^{*,-}$ [€/MWh]	$\bar{\pi}_k^{*,+}$ [€/MWh]
1	11.65	0.33	16.22
2	38.38	1.34	11.13
3	41.17	8.22	8.51
4	29.38	6.97	7.61

For setting up the model of regulation penalties, we assume it is possible to perfectly predict these quarterly averages. This consequently yields four different models of the regulation unit costs, depending on the quarter of the year. Besides that, we consider the three sets of probabilistic forecasts, the quality of which have been assessed in the previous Section. Out of the 8760 predictions series used for the quality evaluation, only 365 are actually considered for defining optimal bids on the electricity pool. They are the forecast series produced everyday at 10:00. The PC strategies resulting from the use of predictive distributions obtained by post-processing M1, M2, and M3 are denoted by PC1, PC2, and PC3 respectively. For comparison, participation strategies directly derived from the point predictions are also evaluated over 2002, and denoted by N1, N2 and N3.

Results from the application of the various bidding strategies are gathered in Table II. In this Table, η corresponds to the overall quantity of energy that is traded on the regulation market, in percentage of the total amount of energy produced over that year. It is then decomposed in the parts that are subject to downward η^+ and upward η^- dispatch (such that $\eta = \eta^+ + \eta^-$). γ is a performance ratio, which gives the income of the wind farm operator as a percentage of the revenue he would have obtained by using perfect forecasts (and thus without any cost for regulation).

TABLE II
RESULTS FOR THE VARIOUS BIDDING STRATEGIES.

	N1	N2	N3	PC1	PC2	PC3
η^+ [%]	23.96	23.30	23.33	15.73	15.97	17.17
η^- [%]	23.45	23.50	26.36	45.75	45.99	43.96
η [%]	47.41	46.80	49.69	61.48	61.96	61.17
γ [%]	84.88	85.31	85.74	90.44	90.06	90.08

One sees from Table II that whatever the considered point prediction method, applying the more advanced strategy instead of using the point forecasts only permits to augment the market participant's income. The revenue of the operator is increased by 5-6 points, which corresponds to reducing the

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

regulation costs by one third over that year. It is not achieved by decreasing the amount of energy subject to the regulation mechanism — this quantity is substantially higher — but by orientating the imbalances. While η^+ and η^- are at similar levels for the bidding strategies N1, N2 and N3, the latter is approximately three times higher than the former when applying the more advanced bidding strategies.

Predictive distributions produced from M1 were found to be the most reliable and to have the highest overall skill. Here, this translates to a higher value from their use when bidding on the Dutch electricity market. For the two other series of probabilistic forecasts, the ones (M3) that were more reliable had a lower skill. The performance ratios corresponding to their related bidding strategies are almost equal. In general, the differences between γ values for the three advanced bidding strategies are lower than the ones between strategies N1, N2, and N3.

V. Conclusions

Several methods exist today for the probabilistic forecasting of wind generation. Such a kind of forecasts provide the whole information on what could be the wind power production in the following hours (typically up to 48-hour ahead). We have investigated on two aspects of the evaluation of predictive distributions: their quality, which corresponds to their statistical performance, and their value, which relates to the resulting benefits from their use in an operational context. This evaluation was based on three series of probabilistic forecasts obtained by the adapted resampling approach.

The quality of predictive distributions of wind generation has been assessed in term of their reliability and sharpness/resolution. Here, we have shown that the three forecast series could be considered as reliable, and that their overall skill (then encompassing both reliability and sharpness/resolution) was pretty similar.

Then, the evaluation of their value showed firstly that considering bidding strategies derived from probabilistic forecasts allows one to increase the revenue resulting from his participation in an electricity pool. Moreover, the predictive distributions that proved to be the most reliable and skillful had a higher value for trading through the electricity market.

Note that the value of probabilistic forecasts does not lie only in this revenue-maximization aspect but also in the fact that the resulting trading or management strategies of wind generation can be specially tailored to the forecast user needs. This is hardly possible if relying on point predictions only. In the future, we will further develop on methodologies based on predictive distributions, in a stochastic programming framework, for an optimal management of wind power.

Acknowledgment

This work was performed in the frame of the ANEMOS Project (ENK5-CT2002-00665) funded in part by the European Commission and by ADEME, the French Environment and Energy Management Agency. The authors gratefully acknowledge ELSAM for providing data for the realization of the study. Acknowledgements are finally due to the modelers who provided their wind power point predictions.

References

[1] A. Zervos, "Developing wind energy to meet the Kyoto targets in the European Union," *Wind Energy*, vol. 6, no. 3, pp. 309–319, July-September 2003.

[2] S.-E. Thor and P. Weis-Taylor, "Long-term research and development needs for wind energy for the time frame 2000-2020," *Wind Energy*, vol. 5, pp. 73–75, April-June 2003.

[3] G. Giebel, R. Brownsword, and G. Kariniotakis, "State of the art on short-term wind power prediction," June 2003, Anemos Project, Deliverable Report D1.1, available online: http://anemos.cma.fr.

[4] G. Kariniotakis and P. Pinson, "Uncertainty of short-term wind power forecasts - A methodology for on-line assessment," in *Proc. of the IEEE/PMAPS Conference 2004, 8th Conference on 'Probabilistic Methods Applied to Power Systems', Ames, Iowa (USA)*, September 2004, pp. 729–736.

[5] H. A. Nielsen, T. S. Nielsen, H. Madsen, and K. Sattler, "Wind power ensemble forecasting," in *Proc. of the 2004 Global Windpower Conference, Chicago, Illinois (USA)*, March 2004.

[6] H. A. Nielsen, T. S. Nielsen, and H. Madsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, no. 1-2, pp. 95–108, January-April 2006.

[7] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119–132, May-June 2004.

[8] A. H. Murphy, "What is a good forecast? An essay on the nature of goodness in weather forecasting," *Wea. Forecasting*, vol. 8, no. 2, pp. 281–293, 1993.

[9] P. F. Christoffersen, "Evaluating interval forecasts," *International Economic Review*, vol. 39, no. 4, pp. 841–862, 1998.

[10] P. Hall and A. Rieck, "Improving coverage accuracy of nonparametric prediction intervals," *J. Royal Stat. Soc.*, vol. 63, no. 4, pp. 717–725, 2001.

[11] Z. Toth, O. Tallagrand, G. Candille, and Y. Zhu, "Probability and ensemble forecasts," in *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. Jolliffe and D. Stephenson, Eds. John Wiley & Sons Ltd, 2003.

[12] R. T. Baillie and T. Bollerslev, "Prediction in dynamic models with time-dependant conditional variances," *J. Econometrics*, vol. 51, pp. 91–113, 1992.

[13] S. McNees, "Forecast uncertainty: can it be measured?" Federal Reserve Bank of Boston, Boston" mimeo, 1995.

[14] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," University of Washington, Department of Statistics, Tech. Rep., May 2005, technical report no. 483.

[15] P. Pinson, H. A. Nielsen, T. S. Nielsen, G. Kariniotakis, and H. Madsen, "Properties of quantile and interval forecasts of wind generation and their evaluation," in *Proc. of the 2006 European Wind Energy Association Conference, EWEC'06, Athens, Greece*, February 2006, Scientific Track.

[16] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," University of Washington, Department of Statistics, Tech. Rep., 2004, technical report no. 463.

[17] A. Boogert and D. Y. Dupont, "On the effectiveness of the anti-gaming policy between the day-ahead and real-time electricity markets in the Netherlands," *Energy Economics*, vol. 27, no. 5, pp. 752–770, September 2005.

[18] J. Usaola, O. Ravelo, G. Gonzalez, F. Soto, C. Davila, and B. Diaz-Guerra, "Benefits for wind energy in electricity markets from using short-term wind power prediction tools - A simulation study," *Wind Engineering*, vol. 28, no. 1, pp. 119–128, 2004.

[19] G. N. Bathurst, J. Weatherhill, and G. Strbac, "Trading wind generation in short-term energy markets," *IEEE Trans. on Power Syst.*, vol. 17, no. 3, pp. 782–789, August 2002.

[20] P. Pinson, C. Chevallier, and G. Kariniotakis, "Optimizing benefits from wind power participation in electricity markets using advanced tools for wind power forecasting and uncertainty assessment," in *Proc. of the 2004 European Wind Energy Association Conference, EWEC'04, London, United Kingdom*, November 2004.

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.

Pierre Pinson was born in Poitiers, France. He received the M.Sc. degree in applied mathematics from the National Institute for Applied Sciences (INSA Toulouse) in 2002 and the Ph.D. degree in energetics from Ecole des Mines de Paris in 2006. Part of his education was made at University of Leicester (UK), TU Delft (The Netherlands), and the Technical University of Denmark. He is currently with the Informatics and Mathematical Modeling department of the Technical University of Denmark as a junior scientist. His research interests include among others forecasting methods, optimization under uncertainty, decision sciences, and renewable energies.

Jérémie Juban was born in Saint-Étienne, France. He received the M.Sc. degree in applied mathematics from the National Institute for Applied Sciences (INSA Rouen) in 2004. He spent his last academic year at McGill University (Canada). He is currently a Ph.D. student at the Center for Energy and Processes of Ecole des Mines de Paris. His research interests include artificial intelligence, forecasting methods, risk assessment, and renewable energies.

Georges N. Kariniotakis was born in Athens, Greece. He received his Engineering and M.Sc. degrees from the Technical University of Crete, Greece and his Ph.D. degree from Ecole des Mines de Paris in 1996. He is currently with the Center for Energy and Processes of Ecole des Mines de Paris as a senior scientist. He is the scientific and technical coordinator of the EU project Anemos. His research interests include among others renewable energies, wind power forecasting, distributed generation and artificial intelligence. He is a member of IEEE.

© Copyright KTH 2006

Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on November 20, 2009 at 10:32 from IEEE Xplore. Restrictions apply.