Net load forecasting in presence of renewable power curtailment

P. M. Fonte
Instituto Superior de Engenharia de Lisboa
Lisboa, Portugal
pfonte@deea.isel.ipl.pt

[1]Cláudio Monteiro
[1,2]Fernando Maciel Barbosa
[1]University of Porto, Faculty of Engineering
[2]INESC TEC
Porto, Portugal
cdm@fe.up.pt, fmb@fe.up.pt

Abstract— **In this paper is analyzed a real case study based on an islanding power grid where there is the necessity of wind power curtailment during the operation of the power grid. This curtailment skews the wind power production database and creates a huge challenge to the overall power production forecast. Thus is presented a solution which allowed making more accurate forecasts in order to improve the renewable production and the reduction of fuel consumption in thermal power plants.**

Index Terms-- **Kernel Density Estimator, Power curtailment, Power forecasting.**

I. INTRODUCTION

One of the issues of an optimal power systems scheduling is the correct forecasting. A few decades ago the main concerning was with the load demand and some hydro power production. With the increase of other renewable energy sources (RES) as wind, solar and mini hydro new challenges arise leading to several research works[1],[2]. When the explanatory variables results from meteorological forecasts there are always associated errors. These errors can result from several factors, such as incorrect or incomplete models, incorrect starting conditions, wrong parameters, extreme events, variations of sources dynamics over the forecast period, amongst others.

The models developed for load forecast generally get good results with lower deviations from the measured values because load profiles follow a characteristic pattern. Comparing with load forecast, the prediction of RES, due to its variability and intermittence, presents much bigger challenges. Thus, beyond the developed point forecasts techniques there is the necessity to incorporate uncertainty in the forecasting. Though addressing other types of RES (namely solar and hydro), wind generation is presented as the main source of generation uncertainty in power systems scheduling, grid operation and market environment. There are three major factors which have influence on the uncertainty of wind power forecast, namely, the Numerical Weather Prediction (NWP), the conversion of wind to power (due to the nonlinearity of the power curve) and terrain complexity. On the other hand, the NWP and the clouds dynamic are the main source of uncertainty in the case of solar photovoltaic since conversion is well defined. In the case of hydro power forecasting systems, the uncertainty generally propagates from the NWP model through the rainfall-run-off model. The rainfall-run-off models are limited by their representation of flow dynamics, whose main problem is not the representation of the dynamic but knowing the local parameters [3]. However, due to its high installed power capacity all over the World, wind power forecast gathers the majority of the attention of researchers, and the major number of published works.

The uncertainty created by these errors has a great impact on power systems scheduling since the forecasted values at the beginning of the scheduling process can be quite different from those in the operation periods. In a traditional and conservative point of view, generally, the uncertainties are compensated using conservative decisions, like over-designing the equipment or overestimating the operational parameters basing them on worst-case. This approach, though being secure, may lead to significant results' deterioration from the optimization problem. To overcome this situation, several uncertainty models are provided in literature, as moments of distributions, set of quantiles or interval forecasts, probability mass functions and probability density functions (parametric and non-parametric)[4],[5].

II. PROBLEM CHARACTERIZATION

The scheduling challenges are enhanced in islanding systems with low rated power, without connection to the large continental grids and without storage capacity. Large variations on renewable production can introduce stability problems in the network, which can originate generation or load shed and, at limit, black-outs. The island power system under study is composed by one thermal power plant with 8 units based on heavy fuel plus 2 geothermal, 7 mini-hydro and 1 wind power plants.

The yearly average geothermal power production is 19,2 MW (meaning approximately 42% of the yearly average load). This generation acts as base of the load diagram and do not contributing to the load follow. On the other hand hydro generation (3 MW of rated power) is inexpressive. Therefore the effective load following has to be done by an efficient management between thermal and wind production. Analyzing the production datasets it is verified that all this renewable production helps to decrease the thermal production during peak load periods, but during off-peak periods the system is already saturated with renewable power production. Additionally, to the system operator, for operational security reasons it is mandatory that, at least, two thermal units must be on-line. It is to avoid the complete loss of thermal production due to outages and represents a minimum production of 12,85 MW. In fig. 1, the hourly average thermal production, as well as the sum of minimum technical limits of 2 units, since 0h00 of December 1st up to 23h of December 31st of 2013 is shown.

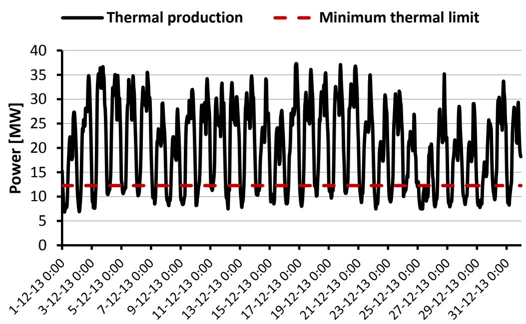

Figure 1. Thermal production and minimum technical limits

As can be seen in fig.1, during the off-peak periods, generally the renewable production is so high (or the load so low) that obligates the thermal units work below the minimum technical limits with poor efficiency and high fuel consumption. During 2012 the thermal machines worked below their minimum limits 17% of the year while in 2013, 20,6%. To minimize this situation, especially during off-peak periods, the wind power production is preventively limited with consequent waste of renewable resources (which happened during 30,5% of the set of 2012 and 2013). On the other hand, an extreme reduction of the thermal committed capacity can lead to a situation wherein the spinning reserves are not sufficient to handle with great variations of load, renewable production or generation outages. Therefore, due to the uncertainty in load and renewable production, generally, sometimes is hard to find a completely robust/economic scheduling solution. To face this problem, it is clear the necessity of an efficient method to forecast load and renewable production in order to know the real thermal production necessities. It allows costs and emissions reduction and the optimization of the number and the allocated power of the on-line thermal units.

The wind power limitation does not occur only in this case study, in [6] are described several case concerning wind power curtailment, where it is stated that wind curtailment occurs for two primary reasons: 1) lack of available transmission during a particular time to incorporate some or all of the wind generation; or 2) high wind generation at times of minimum or low load, and excess generation cannot be exported to other balancing areas due to transmission constraints. In these instances, wind generation may be curtailed after other generation is running at minimum and imports reduced or curtailed as well.

III. FORECASTING METHODOLOGY

The thermal production forecasting is characterized as *net load* [7],[8] and it is calculated by the differences between forecasted load and the sum of forecasted renewable production. These forecasts, further than spot values, incorporate uncertainty by probabilistic forecasting with the *net load* resulting from the convolution of the various forecasting probabilistic distributions. The forecasted *pdf* were based on the *Nadaraya-Watson* estimator (1) which allows estimating a random variable Y, when the explanatory random variable X is equal to x [9]. This conditional density estimation can be seen as a generalization of regression, since conditional density estimation aims at obtaining the full probability density function $f_{y|x}(y|x)$ [10]. In the case of power forecast, it consists on the estimation of the future conditional *pdf* of power for each look-ahead time step $t+k$ ($p_{t+k|t}$), given a set with N pairs of samples (p_n,x_n) summarizing all information available up to instant t. Each pair consists on a set of explanatory variables X_n and the corresponding value of variable to be predicted P_n. In this process it is assumed that explanatory variables $x_{t+k|t}$ are known for each time-step ahead and p_{t+k} is the power forecasted for look ahead time $t+k$.

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

In (1) $\hat{f}_{p,X}\left(p_{t+k},X_{t+k|t}\right)$ is the estimated joint density function and $\hat{f}_X\left(X_{t+k|t}\right)$ is the marginal density of X. However, since the joint and marginal densities are not known, they can be determined with a nonparametric kernel estimator [9]. As the random variable can depend on several explanatory variables a multivariate KDE can be used, and applied to the *Nadaraya-Watson* estimator of (1). The conditional density estimator results from (2)[9],[11].

$$\hat{f}_P\left(p_{t+k},x_{t+k|t}\right)=\sum_{i=1}^{N}K\left(\frac{p-P_i}{h_p}\right).\frac{\prod_{j=1}^{D}K_j\left(\frac{x_j-X_{ij}}{h_j}\right)}{\sum_{i=1}^{N}\left[\prod_{j=1}^{D}K_j\left(\frac{x_j-X_{ij}}{h_j}\right)\right]} \tag{2}$$

In (2) N is the number of samples, D is the number of variables and K_j is the kernel function to each variable j. The parameter h_j is the bandwidth of each kernel around each sample X_{ij} and controls the smoothness of the estimation. For all explanatory variables the kernel function chosen was the normal distribution though in the case of wind direction a wrapped normal distribution was chosen. The optimization of the bandwidth h_j was done with Leave-One-Out Cross Validation (LOOCV) technique [12].

To develop this process, the datasets used in this work contain hourly average values since 0:00 of 1st January 2012 up to 23h00 of 30th June 2014. The training/parameterization dataset contain 17520 hourly average values since 0:00 of 1st January up to 23h00 of 31st December 2013 and the test/validation set is composed by 4344 hourly average values since 0:00 of 1st January 2014 up to 23h00 of 30th June 2014. The NWP forecasts have 1 hour of temporal resolution and the forecast are available at 00h00 for 00h00 up to t+24[13].

The load, hydro and geothermal production forecasts do not exhibit considerable challenges since they all are based NWP and historical datasets strongly connected with the explanatory variables. In the case of load greater errors may arise if the real conditions are not sufficiently envisaged in the dataset for a certain forecast moment. For instance, national or regional holidays, abnormal temperatures for a certain period of the year, general strike, among others. In the case of geothermal units, the power production is dependent on a renewable, but easy to control, resource. The production is defined by set points and it remains relatively constant around the set point. In this case, the main source of the deviation between what was forecasted and the real production was the unexpected outages of some units. The hydro power plants present a negligible storage capacity meaning the power production cannot be delayed from the moment when it rains until the moment when there is the necessity of power production. On the other hand, as the watersheds are not big enough to introduce a significant delay between the rain period and the production, the hydro power forecast depends only from the rain forecasts, which are the main source of errors being necessary an accurate forecast.

The wind power forecasts introduce a different kind of challenges, because the measured values of production could not be fully linked with the explanatory variables due to wind power curtailment. Even with accurate forecasts of the explanatory variables, remarkable errors can occur. In fig. 2 the measured hourly average wind speed and measured power is shown. It is clear that there is a large amount of wind power values which do not correspond to measured wind values. These differences can result from malfunctions of the measuring equipment, unexpected units outages and wind power curtailment. Very high values of wind speed are another source of uncertainty, since the installed turbines are equiped with "*software for storm regulation*". Instead of cutting the production to velocities above the maximum, it regulates the pitch angle of the blades in order to reduce the rotation velocity and, consequently, the power production. Without this information, the power forecasts for velocities above the maximum values become hard to forescast. From the above, it is clear that the curtailment process will introduce very significant errors between the wind speed prediction and the measured power, skewing the dataset. In this case it was necessary to do some data pre-processing, gathering the information disclosed by the system operator about wind generation limits in each hour. Thus, there was the necessity to filter all these situations in the dataset, replacing the wind power curtailed values by "real" values which should be measured in absence of limitation.

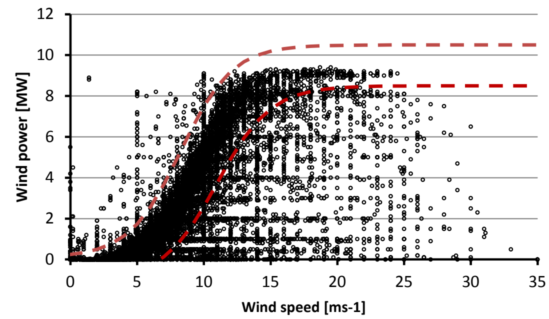

Figure 2. Measured wind speed and power with outlier filtering

This can be achieved computing a theorectic wind-to-power (W2P) function. As shown in fig. 2 , to overcome some of these problems, two sigmoid functions were modelled in order to act as filters. With this fuctions, it was intended to filter "abnormal" wind power production values. This process was applied with measured wind speed in order to avoid forecasting and W2P errors. After filtering the values outside the limits, with the least squared method, it was possible to achieve a "theorectical" relation between wind speed and power production, given by (3) as shown in fig. 3.

$$P_W = \frac{9,2}{1+e^{(-0,6v+6)}} \tag{3}$$

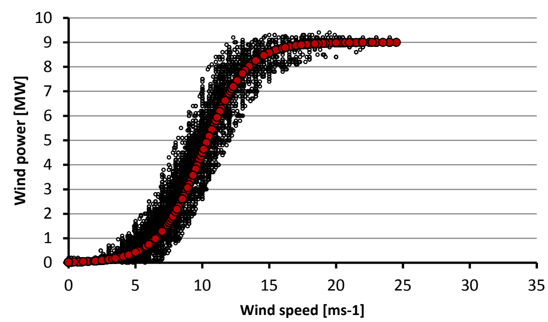

Figure 3. Resulting "theorectical" power curve

One of the informations given by the system operator, behind the power produced and wind measured is the wind power limitation. After the remotion of the values resulting from bad measurements and outages, is possible, knowing the periods where there was curtailment, and applying (3) to those periods, to have an ideia of the wind power forecast in absence of curtailment. By this way the skewness of the dataset can be minimized reducing the non-controlable errors and focusing only in the errors that deliveried from the forecatings. In fig. 4 the wind power production, the wind power limitation, as well as the theoretical wind power production which results from equation (3) from 0h00 of 28th January up to 23h00 of 3rd February is depicted. It is observable that there are notorious differences mainly during off-peak periods. The theoretical wind power production should be understood as the wind power values that would be measured in the absence of limitation. The difference between the theoretical and the measured wind energy during this period was approximately 278 MWh.

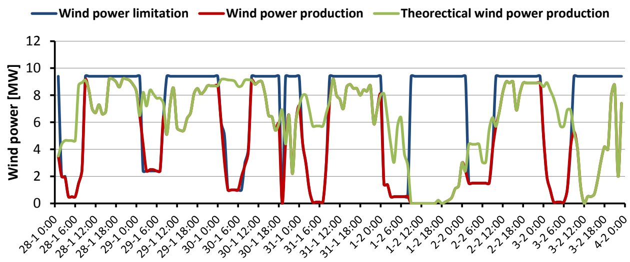

Figure 4. Measured and theorectical wind power

This is a clear sign that curtailment introduces remarkable errors in dataset and the potential wind power production that is wasted.

IV. CASE STUDY

To demonstrate the results of the proposed technique was done a *net load* forecast for 24 hours ahead during a week since 0h00 of 28th January up to 23h00 of 3rd February. Fig. 5 shows the spot forecasted *net load* with the respective uncertainty interval, as well as the measured values. The nominal coverage rate of interval is 0.8. To feature the wind curtailment accomplished during the period under study, the wind power limitations decided by the system operator are also depicted. Additionally, the possible measured values that *net load* could present, in absence of wind curtailment are shown too.

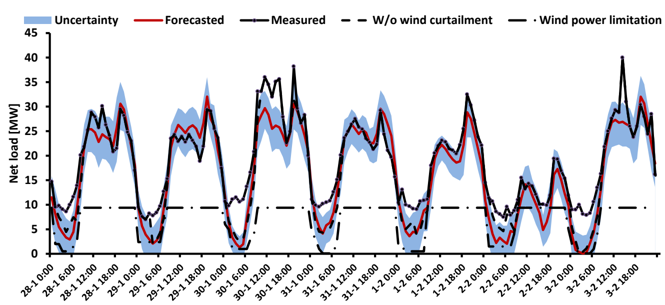

Figure 5. Forecasted and measured *net lod*

In a first analysis it is clear that the system operator opted for wind power limitation during all off-peak periods. Considering that when there is an effective power curtailment, the *net load* tends to grow, it is explained why, in all off-peak periods the *net load* with wind power curtailment tends to be higher than those forecasted. Notice that this analysis is done under the assumption that there were no notable errors in the remaining load, hydro and geothermal forecasts. Excluding some cases, as 30th January and 3rd February, the forecasts outside the off-peak periods present a reasonably fitting with the measures. The measured values also are reasonably covered by the uncertainty interval. The performances of the spot forecasts can be assessed by some indicators as the BIAS, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Standard Deviation of Errors (SDE). In table I the results of spot forecasts are shown [14].

TABLE I. SPOT FORECASTS ASSESSMENT

	BIAS	***MAE***	***RMSE***	***SDE***
measured	2,18	3,19	3,91	4,02
theoretical	0,79	2,14	2,75	3,22

On the other hand, the performances of probabilistic forecast can be assessed by other indicators such as reliability, sharpness and resolution [4],[11],[15],[16]. The dataset under study was composed by hourly forecasted and measure values of *net load* with and without wind curtailment, between January 1st and June 30th 2014. In fig. 6 the reliability of the *net load* probabilistic forecast is shown as well as the "ideal" reliability. The reliability is calculated with the real measured *net load* and the theoretical *net load* that should be measured in absence of curtailment.

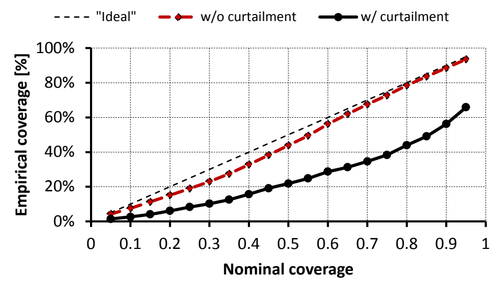

Figure 6. Reliability diagram

It is clear that the forecasting method used in this approach tends to systematically underestimate the uncertainty [11],[15],[17] since forecasted quantiles proportions are lower than the empirical ones. Thus, the values of *net load* outperform all estimated quantiles, meaning that the probabilistic forecasts have an associated bias. A more intuitive way to analyze the bias of the probabilistic forecasting methods is representing it as a deviation from the "ideal" reliability. It is done subtracting the nominal proportion α to the empirical coverage $\hat{a}_k^{(\alpha)}$ resulting the reliability diagram shown in fig. 7.

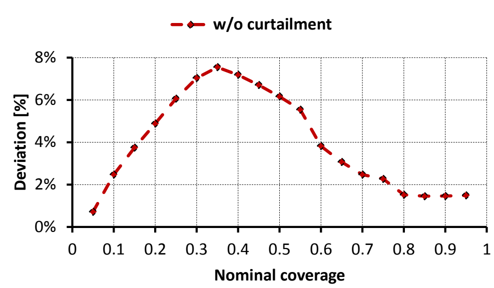

Figure 7. Reliability diagram (deviation from "ideal" relaiability)

It is visible that the uncertainty was underestimated for all predicted quantiles. It should be noticed that the *net load* results from four different variables forecasts, with different values and profiles of uncertainty. For the same dataset, it was calculated the sharpness, as shown in fig. 8. The values of the sharpness are relatively low, with a nominal coverage of 0.9, corresponding to 25,8% of the rated *net load*. For rated *net load*, it was considered the highest value registered in the dataset (36,8 MW). Analyzing the results it is clear that it must be a trade-off between the reliability and the sharpness, because improving reliability will usually worsen the sharpness [11],[15],[17]. Low values of sharpness can lead to "narrow" uncertainty intervals which can result in underestimation or overestimation of the uncertainty, with consequent degradation of the reliability.

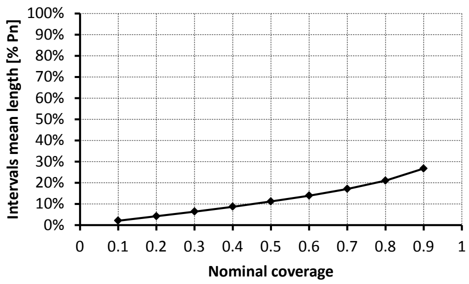

Figure 8. Sharpness diagram of the probabilistic *net load* forecast

Another criterion which can be used for the evaluation of probabilistic forecasts is the resolution [16]. It represents the capacity of the forecasting model to provide situation dependent forecasts. It can be measured by the standard deviation of the predictive intervals size since it is not possible to directly verify this property. Fig. 9 shows the resolution of the probabilistic *net load* forecast. In general and contrarily to sharpness, increasing the resolution gives more value to the probabilistic forecasting method [19]. Large standard deviations reveal that the probabilistic forecasting method has the capacity to represent a wide set of real situations.

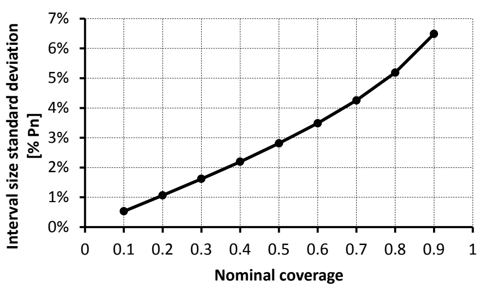

Figure 9. Resolution of the probabilistic *net load* forecast

As seen in fig. 9 the standard deviation interval size is relatively low with a resolution of 6,5% to a nominal coverage of 0.9. This value results from the smoothing effect of the aggregation of renewable production and load forecasts. Throughout the dataset it is verified that the *net load* does not reveal notable changes when submitted to identical inputs and, consequently, the uncertainty profile does not significantly change along the dataset.

V. Conclusions

This work presented a real case study which by its own characteristics presents several challenges to the system operator. The wind power curtailment skewed the dataset which makes the forecasting quite challenging. With the proposed filtering technique it is already possible to do wind power forecasts with less error and forecast the *net load* with more accuracy. With these forecasts is possible to improve the scheduling of the power system in order to optimize the number and the allocated power for each thermal unit.

References

[1] S. Pelland, J. Remund, J. Kleissl, T. Oozeki, and K. De Brabandere, "Photovoltaic and Solar Forecasting: State of the Art," *Int. Energy Agency Photovolt. Power Syst. Program. Rep. IEA PVPS T14*, pp. 1–40, 2013.

[2] M. Lydia and S. Suresh Kumar, "A comprehensive overview on wind power forecasting," *2010 Conf. Proc. IPEC*, pp. 268–273, Oct. 2010.

[3] R. Hostache, P. Matgen, M. Montanari, C. Fosty, and L. Pfister, "Assessment and reduction of hydro-meteorological model forecasting uncertainty using an autoregressive error model and bivariate meta-gaussian density," in *4th IASME / WSEAS International Conference on Water Resource, Hydraulics & Hydrology*, 2009, pp. 110–118.

[4] J. Juban, "Uncertainty estimation of wind power forecasts: Comparison of Probabilistic Modelling Approaches," in *EWEC - European Wind Energy Conference & Exhibition*, 2008.

[5] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, "Wind Power Forecasting : State-of-the-art 2009 Technical Report," 2009.

[6] S. Fink, C. Mudd, K. Porter, and B. Morgenstern, "Wind Energy Curtailment Case Studies May 2008 — May 2009," 2009.

[7] Y. V Makarov, P. V Etingov, and J. Ma, "Incorporating Uncertainty of Wind Power Generation Forecast Into Power System Operation , Dispatch , and Unit Commitment Procedures," *IEEE Trans. Sustain. Energy*, vol. 2, no. 4, pp. 433–442, 2011.

[8] S. Jin, A. Botterud, and S. M. Ryan, "Temporal Versus Stochastic Granularity in Thermal Generation Capacity Planning With Wind Power," *IEEE Trans. Power Syst.*, pp. 1–9, 2014.

[9] S. Demir and O. Toktamis, "On The Adaptive Nadaraya-Watson Kernel Regression Estimators," *Hacettep J. Math. Stat.*, vol. 39, no. 3, pp. 429–437, 2010.

[10] G. Fu, F. Y. Shih, and H. Wang, "A kernel-based parametric method for conditional density estimation," *Pattern Recognit.*, vol. 44, no. 2, pp. 284–294, Feb. 2011.

[11] R. J. Bessa, V. Miranda, A. Botterud, and J. Wang, "Time Adaptive Conditional Kernel Density Estimation for Wind Power Forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, 2012.

[12] J. M. Liu, R. Chen, L.-M. Liu, and J. L. Harris, "A semi-parametric time series approach in modeling hourly electricity loads," *J. Forecast.*, vol. 25, no. 8, pp. 537–559, Dec. 2006.

[13] "P.M. Fonte, 'Advanced Forecast and Scheduling of power Systems with Highly Variable Sources,' Ph.D. dissertation, Dept. Electrical Engineering. Portugal., Faculty of Engineering, Univ. of Porto, Porto, 2015."

[14] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. Skov, "A Protocol for Standardazing the performance evaluation of short term wind power prediction models," in *Wind Engineering, Multi-Science Publishing*, 2005, pp. 475–489.

[15] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: required properties and evaluation," *Wind Energy*, vol. 10, no. 6, pp. 497–516, Nov. 2007.

[16] P. Pinson, J. Juban, and G. N. Kariniotakis, "On the Quality and Value of Probabilistic Forecasts of Wind Generation," in *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on*, 2006, pp. 1–7.

[17] R. J. Bessa, M. Matos, I. C. Costa, L. Bremermann, I. G. Franchin, R. Pestana, N. Machado, H.-P. Waldl, and C. Wichmann, "Reserve Setting and Steady-State Security Assessment Using Wind Power Uncertainty Forecast: A Case Study," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 827–836, Oct. 2012.

[19] P. Pinson, G. Kariniotakis, T. S. Nielsen, and H. Madsen, "Properties of Quantile and Interval Forecasts of Wind generation and their Evaluation," in *Europe's Premier Wind Energy Event*, 2006, no. 1.