Uncertainty and Prediction Risk Assessment of short-term wind power forecasts

Pierre Pinson, Georges Kariniotakis

Uncertainty and Prediction Risk Assessment of Short-term Wind Power Forecasts

Pierre Pinson
Ecole des Mines de Paris,
Center for Energy Studies,
pierre.pinson@ensmp.fr

George Kariniotakis
Ecole des Mines de Paris,
Center for Energy Studies,
georges.kariniotakis@ensmp.fr

Abstract

The paper introduces a new methodology for assessing on-line the prediction risk of short-term wind power forecasts. The first part of this methodology consists in computing confidence intervals with a confidence level defined by the end-user. Focus is given in this paper to the second part of this methodology, which consists in a quantification of the meteorological risk in order to give signals to the operator on the prediction risk - i.e. the probabilities for the occurrence of high prediction errors depending on the weather stability. For this purpose, two indices, named MRI and NPRI, are defined reflecting respectively the spread of the available Numerical Weather Predictions and of the wind power ensemble predictions generated from meteorological poor man's ensembles. A relation between these indices and the level of prediction error is shown. Evaluation results of this methodology over a three-year period on the case study of a Danish wind farm and over a one-year period on the case study of an Irish farm are given. The proposed methodology has an operational nature and can be applied to all kinds of wind power forecasting models.

Keywords: Wind power, short-term forecasting, confidence intervals, prediction risk, on-line software, numerical weather predictions, ensemble forecasting.

1 Introduction

NOWADAYS, wind farm installations in Europe exceed 28 GW. Motivated by the Kyoto Protocol, the European Commission has set the target of doubling the share of renewables in gross energy consumption from 6% in 1997 to 12% in 2010 [1]. This directive targets 22,1% indicative share of electricity produced from renewable energy sources in total Community electricity consumption by 2010. To achieve this share, installed wind power capacity in the Member States should increase to 45-60 GW. In 2003, the European Renewable Energy Council (EREC) revised upwards the 2010 target to 75 GW [2]. Such a large-scale integration of wind generation causes several difficulties in the management of a power system. Often, a high level of spinning reserve is allocated to account for the intermittent profile of wind production, thus reducing the benefits from the use of wind energy. Predictions of wind power production 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 one of the priorities in wind energy research needs for the coming years [3].

Apart from spot forecasts of the wind parks output in the next hours, of major importance is to provide tools for assessing on-line the accuracy of these forecasts. Tools for on-line evaluation of the prediction risk are expected to play a major role in trading wind power in a liberalized electricity market since they can prevent or reduce penalties in situations of poor prediction accuracy. In practice today, uncertainty is given in the form of confidence intervals or error bands around the spot wind power predictions.

Typical confidence interval methods, developed for models like neural networks [4–6], are based on the assumption that the prediction errors follow a Gaussian distribution. This however is often not the case for wind power predictions, where error distributions exhibit some skewness, while the confidence intervals are not symmetric around the spot prediction due to the shape of the wind turbines power curve. Moreover, the level of predicted wind speed introduces some non-linearity to the estimation of the intervals; i.e. at the cut-off speed, the lower power interval may switch to zero due to the cut-off effect. The limits introduced by the wind farm power curve (min, max power) are taken into account by the method proposed in [7], which is based on modeling errors using a β-distribution, the parameters of which have to be estimated by a post-processing algorithm. This approach however is applicable only to "physical"-type models since such models estimate power using an explicit wind turbine power curve - i.e. a function that gives the wind power output from the wind speed at the turbine level, which is not necessarily the case for statistical or artificial intelligence based models as the ones considered here [8].

In [9], [10] wind speed errors are classified as a function of look-ahead time and then they are transformed to power prediction errors using the wind turbine power curve vs wind speed. This method however is also limited for application to physical models rather than statistical ones since it requires local wind speed predictions (at the level of the wind farm), while it does not provide uncertainty as a function of a pre-specified confidence level. The wind speed errors are estimations provided by the Numerical Weather Prediction (NWP) model. As a consequence, this method does not take into account the modeling error itself that might be due to the spatial refinement of weather predictions or to the power curve used. Moreover, wind speed measurements are required, which might not be made available on-line. In a follow-up paper [11], the authors show a relation between specific meteorological patterns (defined from measurements) and various levels of forecasting error: this is a first step in the definition of risk indices in order to quantify the weather predictability.

In this paper we expose a part of a methodology for assessing on-line the uncertainty of wind power predictions by the joint use of appropriately defined confidence intervals and prediction risk indices.

In previous work [12], the authors proposed a generic approach for the estimation of confidence intervals, which can be applied to both ”physical” and ”statistical” wind power forecasting models. This is due to the fact that no hypothesis is made about the distribution of the prediction errors. The method accommodates both modeling errors and errors based on the NWPs. It uses past wind power data, which are often available on-line by a Supervisory Control And Data Acquisition (SCADA) system, as well as NWPs, which are nowadays the basic input to all models.

Generally, when confidence higher than 80% is required, the intervals are quite wide. This can lead to conservative or costly managing strategies of the predicted wind power (i.e. allocation of high spinning reserve). Given that confidence intervals are estimations of the uncertainty based on the past performance of the model, the objective of this work is to develop additional preventive tools able to assess on-line the prediction risk as a function of the forecast weather situation. The aim is to provide comprehensive information to the operators so that they are able to adjust the risk they are going to face when managing the predicted wind power, i.e. take low risk when forecast weather situation is unstable.

The paper presents detailed results from the application of the method on the case studies of Ireland and Denmark, where the aim is to predict the output of several wind farms for 48 hours ahead using on-line measurements and predictions from Hirlam NWP system. This evaluation is based on several years of data.

The prediction risk indices proposed in this paper together with the work presented in [12] on the estimation and the fine-tuning of prediction intervals constitute a complete methodology for assessing on-line the uncertainty of wind power predictions.

2 Errors in wind power predictions

Let us define the prediction error for the look-ahead time $t+k$ as following:

$$e_{t+k/t} \equiv p_{t+k} - \hat{p}_{t+k/t}, \tag{1}$$

where $\hat{p}_{t+k/t}$ is the forecast for look-ahead time $t+k$ produced by the model at time origin t, and p_{t+k} is the measured wind power. This error can vary between -100% and 100% of the nominal wind park power (P_n). For a non-bounded prediction model it can take values even outside that range. The possible error of the prediction model, defined as ”error margin”, depends on the level of measured wind power. Figure 1 represents graphically the error margin as a function of the wind turbine characteristic curve.

For wind speeds below cut-in speed, the error margin is maximal since the model can predict a production up to the nominal wind turbine power. On the contrary, for higher wind speeds the model will show a positive error margin, i.e. the generated power is likely to be greater than the one proposed by the prediction model. Close to the cut-off wind speed the uncertainty is again maximal since the model can switch from a positive error margin to a negative one, or the inverse.

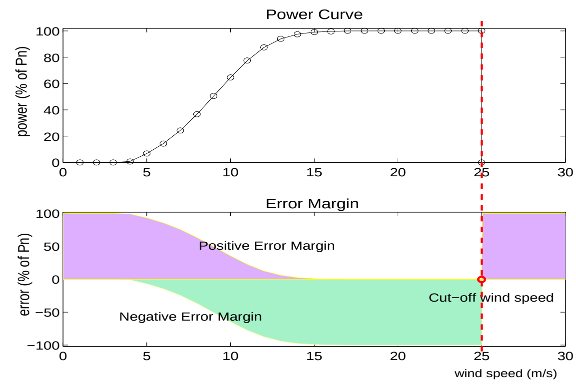


FIGURE 1: The error margin as a function of the wind turbine power curve.

The observed prediction error itself is in general the result of three factors: a modeling error e_{mod}, an error due to the accuracy of the input meteorological predictions e_{NWP} and finally, a stochastic component linked to the process itself e_s:

$$e_{t+k/t} = f(e_{mod}, e_{NWP}, e_s). \tag{2}$$

There are several ways to evaluate either offline or on-line the performance of wind power prediction systems. The common measures of forecast accuracy are the Normalized Mean Absolute Error (NMAE) and the Normalized Root Mean Square Error (NRMSE) - the normalization is done by using the wind park nominal power. The former reflects the model average performance while the latter provides complementary information by giving more weight to large errors. Some other statistical measures, like the skill score [13], the determination coefficient R^2, etc., are used in a more marginal way. Nevertheless, since several studies now focus on the value of wind power forecasting, an alternative manner to assess the performance of wind power predictors has appeared, which is the cost of the forecasting errors in an electricity market. In [14], the meaning of several criteria for power prognosis errors are analytically studied and compared. The authors show that depending on the criterion that is used, the value of wind power forecasts may not be the same. Therefore, the design/optimization of a wind generation prediction system should be seen as a multi-objective problem, and defined from the end-user specific objectives (cost minimization on a specific power market, control of the large forecasting errors, etc.). A model performance evaluation is done on a ”global” basis, i.e. over a long period of time. However, this performance is highly variable from one season to another, from one month to another, or even from one day to another. Therefore, in contrast to that ”global” performance evaluation, it is necessary to provide end-users with the possibility to assess the prediction accuracy in a more dynamic way. The two main features we aim to introduce here are the *uncertainty* and *prediction risk* estimation. These two are different concepts: the first corresponds to a visualization of the error distribution on an a-posteriori basis, while the second relates to preventive signals on the expected level of prediction error, depending on what one knows about the current and forecast situations.

The proposed methodology for assessing on-line both uncertainty and prediction risk includes the following features:

- Development of confidence intervals for the spot power prediction. The approach is based on the resampling method, which is applied to samples of errors. Errors are classified using fuzzy sets to account for the level of power and the risk for cut-off events. For that part, we refer to [12]. Figure 2 gives an example of wind power forecast for a wind farm in Denmark, with confidence intervals obtained with the adapted resampling method.

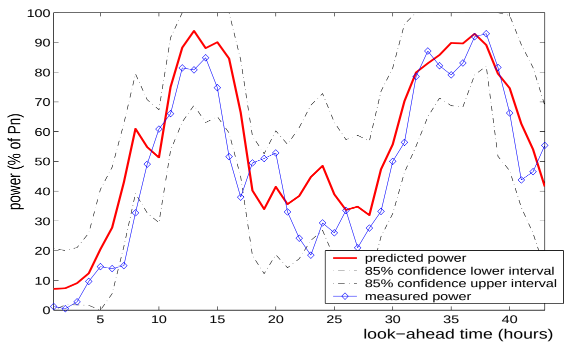


FIGURE 2: Wind power prediction with 85% confidence intervals for a Danish wind farm.

- Dynamic fine-tuning of the size of the intervals depending on the weather stability. This permits to avoid excessive risk or to take preventive actions in situations where high errors are expected ([12]).

- Development of on-line prediction risk indices based on ensembles of NWPs and wind power forecasts. These indices permit to derive rules for assessing the probability of high or extreme prediction errors due to unstable weather situations.

The last point is developed in the following sections.

3 Prediction risk assessment based on ensemble forecasts

Low quality forecasts are due partly to the power prediction model, and partly to the numerical weather prediction system (due in turn to low weather stability). Indeed, an unstable atmospheric situation can lead to very poor numerical weather predictions and thus to worthless wind energy ones. In contrast, when the atmospheric situation is stable, one can expect more accurate wind power predictions from the model.

In [11], methods from synoptic climatology are utilized to classify the local weather conditions based on measurements of wind speed and direction, as well as pressure. This classification, through principal component analysis and cluster analysis methods, permits to reveal characteristic meteorological patterns that can be associated to various levels of prediction error. Indeed, when low pressure systems are dominant the level of error is higher and inversely when high pressure systems govern, the wind prediction error is much lower. However, no link with wind power prediction errors is shown, and this study is based on meteorological measurements, which are often not available on-line. Therefore, the derived method does not appear directly applicable for detecting in a preventive way, and in an on-line environment, situations for which high level of forecasting error is expected. That is the reason why the ideas developed in the following paragraphs exploit the information included in the NWPs (and not in the measurements) in order to develop tools for on-line estimation of the meteorological risk in power predictions.

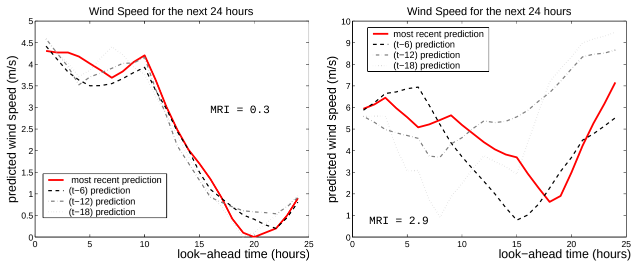
3.1 Wind speed ensemble forecasts for the assessment of weather stability

Meteorological Centres are able to produce different scenarios of Numerical Weather Predictions by perturbing the initial conditions of the forecasting model or by using different NWP models. These scenarios are called ensemble forecasts and permit to evaluate the stability of the weather regime as well as the meteorological predictability [15]. Both the U.S. National Center for Environmental Prediction (NCEP) and the European Center for Medium-Range Weather Forecasts (ECMWF) have produced operational ensembles for more than ten years. A review of their skills is presented in [16].

The NWP wind speed prediction error is composed by a part that is independent of the lead time and by an error that has a linear growth with the prediction horizon. The first includes the effects of weather disturbances that are smaller than the NWP resolution, while the second is due partly to the NWP model errors and partly to an error in the estimation of the initial state [17]. Ensemble forecasts permit to see the influence of this misestimation of the initial state in the weather forecasting evolution, and thus to preventively quantify the prediction uncertainty [18]. Moreover, recent studies have shown that this uncertainty assessment based on ensembles can have an economic value for end-users exposed to weather risk [19], [20].

For wind power applications only one forecast for the next 48 hours is often made available (or purchased) at a given time. For instance, Hirlam gives a unique 48-hour ahead forecast every 6 hours. Nevertheless, for a given hour, several predictions can be available from different time origins in the past (-6 hours, -12 hours, -18 hours, etc.). This kind of ensembles are known as poor man's ensemble forecasts. In a stable and well-predicted weather situation it is expected that these predictions will not differ significantly. Comparing all the available forecasts for the considered period, one can assess weather stability and predictability.

Because we want to have a general evaluation of that stability, 4 sets of predictions of various ages (0, 6, 12 and 18 hours) for the following 24 hours are compared. Figure 3 gives the examples of a stable atmospheric situation (left picture, the forecasts are quite close) and of an unstable one (right picture, spread forecasts).



(A) Stable weather situation (B) Unstable weather situation

FIGURE 3: Representative patterns of stable and unstable weather situations.

In previous work [12], a meteorological risk MRI-index was derived to reflect the spread of the available sets of wind speed forecasts, and it was shown that there is a relation between weather stability and the level of prediction error. In order to have a simple and clear illustration of that relation, we give the average performance of the Fuzzy-Neural Network (Fuzzy-NN) prediction model described in [8] for a wind farm in Ireland, which we compare to its performance for weather situations considered as stable or unstable, as described by the MRI-index. One can see that there is a great variation of the model performance depending on the weather conditions (up to $\pm$ 17% in this example).

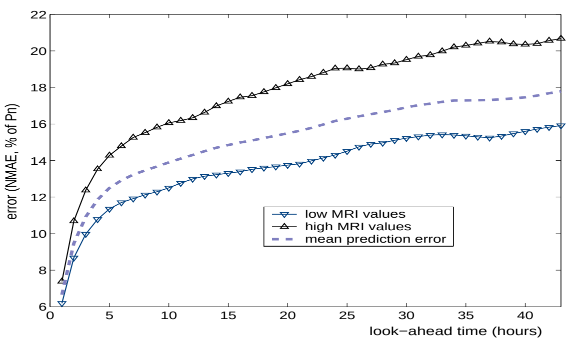


FIGURE 4: The level of prediction error depending on the meteo-risk as described by the MRI-index for a wind farm in Ireland.

3.2 Determining the prediction risk from wind power ensemble forecasts

Several projects focusing on ensemble forecasting for wind power are ongoing [17], [21]. Their aim is to exploit the information contained in multi-scenario forecasts to derive an uncertainty estimate for wind power spot predictions. The use of such a probabilistic approach has already given promising results for the load forecasting problem [22].

The reasoning developed by the authors in [12] for quantifying the meteorological risk was based on poor man's ensemble forecasts for wind speed. The idea was to quantify the meteorological risk and to determine how it relates to the error of power forecasts. However, since the relation between wind speed and wind power is not linear, certain situations, i.e. with high wind speeds, can be damped due to the flat part of the power curve. Inversely, situations with reasonable risk in NWPs can be deteriorated due to the high slope part of the power curve. Then, a way to consider those effects, and also to integrate the sensibility of the other NWP variables, would be to directly quantify the prediction risk for wind generation based on ensemble forecasts for wind power output.

Wind power ensemble forecasts are generated using the prediction model with input NWPs (wind speed, direction etc) provided at different lead times. In the same way real ensemble NWPs can be used if available. When the aim of a statistical prediction model is to produce output in the form of ensembles, then appropriate procedures are followed to optimize its architecture and to train its parameters but their presentation is out of the scope of this paper. In the case of a physical model the procedure is more straightforward.

There are several possibilities to measure the spread of the various sets of forecasts, which appear to be very close. In [23] the standard deviation of the forecasts for each time-step is mentioned as an example. Our aim here is to evaluate the situation for the whole range of prediction horizons in order to reflect the global weather situation. This is why a unique representative index is defined for the following N_h hours instead of indexes for every look-ahead time. In order to calculate the distance between two sets of forecasts, we propose a 2-norm between the N_h-valued vectors containing the predicted wind speed for the N_h following hours.

Define

$$WP_{t-\gamma_i} = \begin{pmatrix} \hat{p}_{t+1/t-\gamma_i} \\ \vdots \\ \hat{p}_{t+k/t-\gamma_i} \\ \vdots \\ \hat{p}_{t+N_h/t-\gamma_i} \end{pmatrix} \in \mathbb{R}^{N_h} \qquad (3)$$

to be the N_f sets of wind power forecasts, with γ_i being the age of each set. The values for γ_i can be 0, 6, 12, etc, for the case of using Hirlam as a NWP supplier: the ensemble member $WP_{t-\gamma_i}$ is the one obtained with the meteorological forecasts of age γ_i.

The distance between the predictions of ages γ_i and γ_j is given by:

$$d(WP_{t-\gamma_i}, WP_{t-\gamma_j}) = \Big(\frac{1}{N_h}\sum_{k=1}^{N_h}\big[\hat{p}_{t+k/t-\gamma_i} - \hat{p}_{t+k/t-\gamma_j}\big]^2\Big)^{\frac{1}{2}}. \quad (4)$$

Then, an index, called hereafter "production-risk" PRI-index, is defined to measure the spread of the power forecasts at a given time. It uses the most recent forecast as a reference and reflects the variability of the older forecasts:

$$\text{PRI} \equiv \sum_{i=1}^{N_f-1} w_i.d(WP_{t-\gamma_0}, WP_{t-\gamma_i}), \quad (5)$$

with w_i ($i \in \{1, 2, ..., N_f - 1\}$) being appropriate weights defined so that:

$$\begin{cases} w_i > w_{i+1}, & i \in \{1, 2, ..., N_f - 2\}, \\ \sum_{i=1}^{N_f-1} w_i = 1. \end{cases} \quad (6)$$

The use of the weights w_i permits to give more importance to the recent information we get from the weather predictions. In the case of real ensemble input rather than poor's man NWPs, the weights are equal for each ensemble member.

In the frame of the case studies of the paper, the horizon N_h for the calculation of the PRI-index is set to 24 hours. Since Hirlam forecasts are provided every 6 hours, four sets (N_f = 4) of wind power predictions covering the period can be produced. However, the same methodology could be applied to seven available sets of Hirlam forecasts on a 6-hour period for instance.

The information we aim to give on the prediction risk has to be normalized. Regarding the range of possible values for the "production-risk" index, the minimium value is clearly 0 for the case of the four sets of forecasts being identical. The maximum one is equal to the wind farm rated capacity, which corresponds to the case of the most recent forecast being 0 or the nominal power P_n while the older forecasts are respectively equal to P_n or 0:

$$\begin{cases} \text{PRI}_{min} = 0, \\ \text{PRI}_{max} = P_n. \end{cases} \quad (7)$$

Therefore, the normalization yields a modified formulation of the production risk index that we will call NPRI:

$$\text{NPRI} \equiv \frac{1}{P_n}\sum_{i=1}^{N_f-1} w_i.d(WP_{t-\gamma_0}, WP_{t-\gamma_i}), \quad (8)$$

As the meteorological risk MRI-index was seen to give an information on the expected level of prediction error [12], the relation between the NPRI-index and this level of prediction error will be shown in the next section. This relation can be used afterwards to define rules for the occurence of large errors depending on the NPRI-index value. In an on-line environment these rules will permit to derive signals or alarms for the end-user of the wind power prediction model. Then, the operator can consider such a signal for:

- taking preventive actions (i.e. increase spinning reserve),
- considering the lower interval (or other alternative strategies), rather than the spot prediction of power, when trading, in order to avoid penalties, etc.

The interest of these results is that they are obtained from single numerical weather predictions (that are used to generate poor man's ensemble forecasts) and not from 'real' meteorological ensembles. This is indeed an advantage for establishing operational tools since almost all the wind power prediction systems use single rather than ensemble NWPs as input. Morever,ensemble NWPs are often not available for purchase from meteorological services. Finally, the results presented here using the poor's man approach demonstrate the value of the methodology for the most pessimistic case. This value is expected to be higher in case that real ensembles are available.

4 Results

In this Section the results are presented from the validation of the developed methodology for a wind farm in Denmark (WF-A) and for one in Ireland (WF-B) with an installed capacity of several MWs each. The prediction model is the adaptive Fuzzy-NN model (F-NN) described in [8]. The available time-series cover a period of five years for WF-A, from which 12000 hours were used for training (learning set), 2000 hours for cross-validation and three years for testing the performance of the model. Regarding WF-B, the time-series cover a period of almost two years (learning: 6600 hours, cross-validation: 1000 hours, testing: one year). The results presented here are on the testing sets.

The prediction model provides forecasts for the next 43 hours with hourly time-steps. Forecasts are updated every hour using SCADA data as input. Hirlam NWPs that are used have a spatial resolution of around 15 km for WF-A and of around 30 km for WF-B. They are provided 4 times per day and at the level of wind farm as interpolated values.

In order to assess the relation between the NPRI-index and the level of prediction error, we collect wind power prediction errors as obtained by the Fuzzy-NN model and for the same period the NPRI-index values are estimated. The prediction error e_t^{24} for the next 24 hours, corresponding to the power forecast made at time t, is calculated as follows:

$$e_t^{24} = \frac{1}{24.P_n}\sum_{k=1}^{24} |e_{t+k/t}|. \quad (9)$$

Then, these errors are binned by NPRI-index values, and the average error $e_{t,j}^{24}$ ($j = 1, 2, ..., N_{bin}$) for the next 24 hours for each bin is computed. Finally, by comparing these averages to the global prediction error $\overline{e_t^{24}}$ of the model

$$\overline{e_t^{24}} = \frac{1}{N_p} \sum_t e_t^{24}, \tag{10}$$

where N_p is the total number of predictions made in the testing set, the representative points in Figures 5 and 6 are obtained. For each bin, 85% confidence intervals computed by a resampling method are also given in order to visualize the errors dispersion. One can notice from these Figures the prediction error increases with the NPRI-index, and the error dispersion too. This means that as the risk index gets higher the prediction error is likely to be greater, as well as the uncertainty on this level of prediction error.

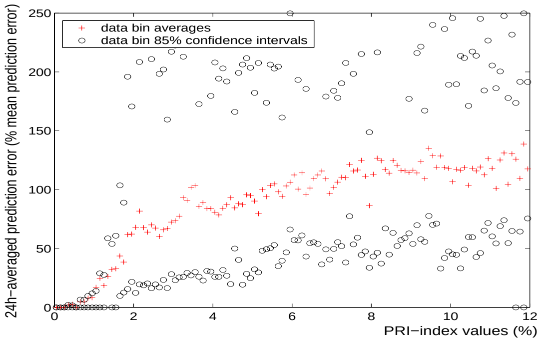


FIGURE 5: Prediction errors vs NPRI-index values over a one year dataset for WF-B: data bin averages and 85% confidence intervals.

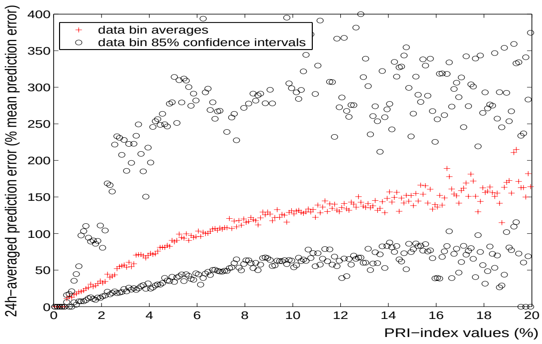


FIGURE 6: Prediction errors vs NPRI-index values over a three year dataset for a WF-A: data bin averages and 85% confidence intervals.

Another way to illustrate that relation is to calculate the cumulative distribution function of the prediction errors for various bins of NPRI-index values (Figure 7).

These curves give the probability with which an error larger than a defined threshold occurs, depending on the value of the NPRI-index. For instance, if at a certain time, the index takes a value between 0 and 2.5, there will be a probability of 1% that an error e_t^{24} larger than the global prediction error $\overline{e_t^{24}}$ occurs. However, if at that same time the value of the index is within the $[15, 20[$ bin, the probability for the same kind of error is much larger (78%):

$$\begin{aligned} P(e_t^{24} > \overline{e_t^{24}} \,|\, \text{NPRI} \in [0, 2.5[\,) &= 1\%, \\ P(e_t^{24} > \overline{e_t^{24}} \,|\, \text{NPRI} \in [15, 20[\,) &= 78\%. \end{aligned} \tag{11}$$

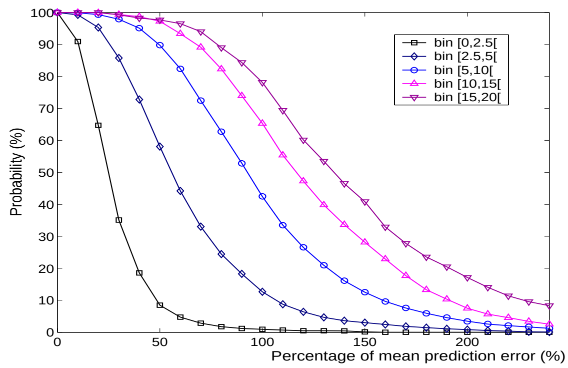


FIGURE 7: Cumulative distribution functions of the prediction error depending on the NPRI-index bin. These curves are based on three years of simulation for WF-A.

Table 1 gives the probabilities for errors to be larger than $\frac{1}{2}$, 1, $\frac{3}{2}$ and 2 times the average error depending on the range of the NPRI-index. The Table is estimated for the case of a wind farm Denmark (WF-A). Based on such a Table, several rules similar to the one given above can be derived.

Table 1 also provides information on the probability of extreme prediction errors to happen (extremes are defined as errors larger than twice the global prediction error of the model). Actually, for WF-A, when the NPRI-index takes low values (between 0 and 5 %) an extreme prediction error is unlikely to happen, and that is not the case if this one is within the bin $[15, 20[$ (18% probability of occurrence). On the other hand, if NPRI $> 10\%$, an error of at least 50% of the global prediction error is expected.

	Ranges of NPRI-index (%)				
Probability (%) of occurrence of errors larger than n times the global prediction error	Bin $[0, 2.5[$	Bin $[2.5, 5[$	Bin $[5, 10[$	Bin $[10, 15[$	Bin $[15, 20[$
$n = \frac{1}{2}$	8	59	90	98	98
$n = 1$	1	13	42	65	78
$n = \frac{3}{2}$	0	4	13	29	40
$n = 2$	0	1	4	8	18

TABLE 1: Rules for the occurrence of larger errors depending on the value of the NPRI-index for WF-A.

Finally, Figures 8 and 9 illustrate ensemble forecasts based on poor man's ensemble of NWPs. For these examples, which are related to WF-A, there are 4 ensemble members of 24 hours ahead each: the spot prediction based on the most recent NWPs and three sets of power forecasts based on Hirlam NWPs supplied 6, 12 and 18 hours ago. The Figure displays the predictions for the window of the next 24 hours. The NPRI-index can be estimated from these ensembles and probabilities of errors larger than defined bounds can be derived from the curves of Figure 7.

These examples illustrate two contrasting cases: the first one shows a situation where wind power predictability is quite high (low NPRI-index value) while the second shows a les predictable situation (high NPRI). The phase shifts between members in the second case warn on deteriorated forecasting accuracy during this period. In fact, the errors e_t^{24} for the next 24 hours for the two cases are respectively equal to 5.83% and 15.42% (of the wind farm installed power) respectively. The average error $\overline{e_t^{24}}$ of the model for this case-study is indeed 9.23%.

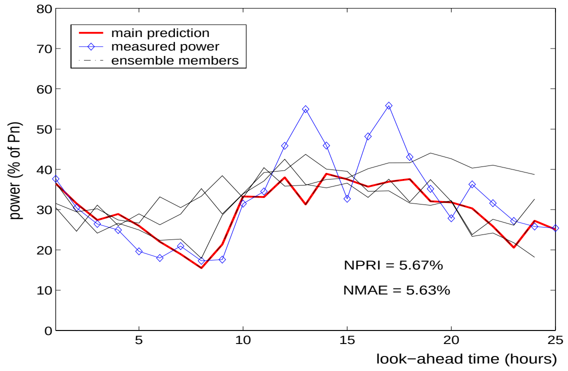


FIGURE 8: Wind power spot prediction and the ensemble members for WF-A (NPRI = 5.67% and NMAE = 5.83% of P_n).

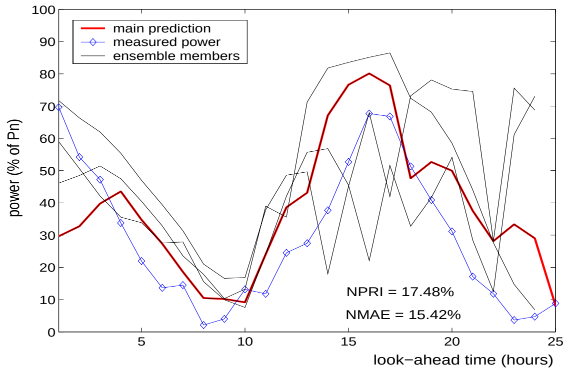


FIGURE 9: Wind power spot prediction and the ensemble members for WF-A (NPRI = 17.48% and NMAE = 15.42% of P_n).

5 Conclusions

A generic methodology for assessing on-line the prediction risk of short-term wind power forecasts has been presented. Focus was given to the second part of this methodology, which consists in providing end-users with information on the expected level of prediction error. For this, a meteorological risk MRI-index has been introduced to evaluate the weather stability. Then, this meteorological risk has been integrated in the wind power prediction process itself by producing multi-scenario wind power predictions from NWP poor man's ensemble forecasts. Another index, named as NPRI, has been derived in order to reflect the production risk. Such an index can forewarn end-users about the probability of small, large or even extreme prediction errors to occur.

The methodology that was described in this paper was tested over a one-year evaluation dataset for a wind farm located in Ireland and over a three-year dataset for a Danish wind farm. The results are encouraging and comprise a first step in the development of on-line tools that can be used in a complementary way to the prediction model itself. The developed methods were implemented in the form of on-line modules and integrated in the Armines Wind Power Prediction System (AWPPS). The prediction modules of AWPPS are integrated in the More-Care Energy Management System and installed for on-line operation in Ireland and other sites such as Crete, Madeira, etc.

Those methods are based on NWP poor man's ensemble forecasts and give promising results. The advantage of such techniques is their operational nature: since most of the state-of-the-art wind power prediction systems use single NWPs, they can be adapted for generating ensemble forecasts for wind power. However, it will be of particular interest in the future to evaluate the gain of using 'real' ensembles produced by meteorological research centers in this methodology.

Finally, as the value of wind power forecasting models has to be assessed, we will evaluate and quantify the benefits (and short-comings) of the presented methodology for the uncertainty and prediction risk management, for both the reserve management and the wind power trading problems.

Acknowledgements

This work was performed in the frame of the ANEMOS Project (ENK-CT2002-00665) funded in part by the European Commission. The authors gratefully acknowledge ESB National Grid and ELSAM for providing data for the realization of the study.

References

[1] Official Journal of the European Commission, "Directive 2001/77/EC of the European Parliament and of the Council, on the promotion of electricity produced from renewable energy sources in the internal electricity market." Sept. 2001.

[2] A. Zervos, "Developing wind energy to meet the Kyoto targets in the European Union," *Wind Energy*, vol. 6, pp. 309–319, July-September 2003.

[3] 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.

[4] J. G. Carney, P. Cunningham, and U. Bhagwan, "Confidence and prediction intervals for neural network ensembles," in *Proc. of the International Joint Conference on Neural Networks 1999*, 1999. paper 2090.

[5] T. Heskes, "Practical confidence and prediction intervals," in *Advances in Neural Information Processing Systems 9* (M. J. M. Mozer and T. Petsche, eds.), MIT Press, Cambridge, 1997.

[6] R. De Veaux, J. Schumi, J. Schweinberg, D. Shellington, and L. H. Ungar, “Prediction intervals for neural networks via nonlinear regression,” *American Statistical Association*, vol. 40, Nov. 1998.

[7] A. Luig, Bofinger, and H. G. Beyer, “Analysis of confidence intervals for the prediction of regional wind power output,” in *Proc. of the 2001 European Wind Energy Association Conference, EWEC’01, Copenhagen, Denmark*, pp. 725–728, June 2001.

[8] G. Kariniotakis and D. Mayer, “An advanced on-line wind resource prediction system for the optimal management of wind parks,” in *Proc. of the 2002 Global Windpower Conference, Paris, France*, April 2002.

[9] M. Lange and H.-P. Waldl, “Assessing the uncertainty of wind power predictions with regard to specific weather conditions,” in *Proc. of the 2001 European Wind Energy Association Conference, EWEC’01, Copenhagen, Denmark*, pp. 695–698, June 2001.

[10] M. Lange and H.-P. Waldl, “Accuracy of short-term wind power prediction depending on meteorological conditions,” in *Proc. of the 2002 Global Windpower Conference, Paris, France*, April 2002.

[11] M. Lange and D. Heinemann, “Relating the uncertainty of short-term wind speed prediction to meterological situations with methods from synoptic climatology,” in *Proc. of the 2003 European Wind Energy Association Conference, EWEC’03, Madrid, Spain*, June 2003.

[12] P. Pinson and G. Kariniotakis, “On-line assessment of prediction risk for wind power production forecasts,” *Wind Energy*, to appear in 2004.

[13] L. Landberg, G. Giebel, H. A. Nielsen, H. Madsen, and T. S. Nielsen, “Short-term prediction - An overview,” *Wind Energy*, vol. 6, pp. 273–280, July-September 2003.

[14] C. S. Nielsen and H. F. Ravn, “Criteria in short-term wind power prognosis,” in *Proc. of the 2003 European Wind Energy Association Conference, EWEC’03, Madrid, Spain*, June 2003.

[15] J. M. Murphy, “The impact of ensemble forecasts on predictability,” *Q. J. R. Meteorol. Soc.*, vol. 114, pp. 463–493, 1988.

[16] F. Atger, “The skill of ensemble prediction systems,” *Monthly Weather Review*, vol. 127, Sept. 1999.

[17] J. U. Joergensen and C. Moehrlen, “Verification of ensemble prediction systems for a new market: wind energy.” Special project interim report, available online, June 2003.

[18] Z. Toth, Y. Zhu, and T. Marchock, “The use of ensembles to identify forecasts with small and large uncertainty,” *Weather and Forecasting*, vol. 16, no. 4, pp. 463–477, 2001.

[19] T. N. Palmer, “The economic value of ensemble forecasts as a tool for risk assessment: from days to decades,” *Q. J. R. Meteorol. Soc.*, vol. 128, no. 581, pp. 747–774, 2002.

[20] S. Jewson and C. Ziehmann, “Using ensemble forecasts to predict the size of forecast changes, with application to weather swap value at risk,” *Atmospheric Science Letters*, vol. 4, pp. 15–27, February 2003.

[21] G. Giebel, L. Landberg, J. Badger, K. Sattler, H. Feddersen, T. S. Nielsen, H. A. Nielsen, and H. Madsen, “Using ensemble forecasting for wind power,” in *Proc. of the 2003 European Wind Energy Association Conference, EWEC’03, Madrid, Spain*, June 2003.

[22] J. W. Taylor and R. Buizza, “Neural network load forecasting with weather ensemble predictions,” *IEEE Trans. on Power Systems*, vol. 17, pp. 626–632, August 2002.

[23] L. Landberg, G. Giebel, L. Myllerup, J. Badger, H. Madsen, and T. S. Nielsen, “Poor-man’s ensemble forecasting for error estimation,” in *Proc. of the 2002 American Wind Energy Association conference, Portland, Oregon*, June 2002.