

Wind Power Forecast Uncertainty Using Dynamic Combination of Predictions

Nuno Pinho da Silva^{1*}, Luis Rosa^{1,2}, Wang Zheng^{1,3}, Rui Pestana^{1,2}

RESEARCH ARTICLE

Received 15 May 2015; accepted 21 August 2015

Abstract

The system operators rely on forecasting tools to promote security of supply in the case of contingent renewable generation upheaval, thus decreasing the chance of counter trading in the intraday markets. This work introduces a self-adaptive ensemble based method providing optimal point predictions under the square loss function constrained over the probability simplex. The output is used to centre a new nonparametric probabilistic power forecast that leverages linear interpolation of the order statistics, thus providing forecast uncertainty estimations. The proposed methodology shows competitive reliability, with coverage and sharpness characteristics that compare favourably with reference methods, thus enabling the perusal of forecast uncertainty in operations planning.

Keywords

Renewable Energy Integration, System Operation, Dynamic Wind Power Forecast, Forecast Uncertainty

1 Introduction

The supply of electricity generated from renewable energy sources plays a crucial role in the implementation of the strategic European energy targets for decreasing carbon dependency. To the political incentive adds the fact that, in general, power production from renewable energy sources has lower marginal cost than carbon based technologies, which is an economic reason for the penetration of renewables in the European power system.

The technology of converting the wind kinetic energy into electricity is one of the most mature forms of power production from renewable sources. It thus has the potential to provide an important contribution for achieving the target of 20% share of renewable energy sources in the overall European Community energy consumption by 2020.

However, the dispatch of wind energy is challenging due to the volatile nature of the wind speed at short-term timescales. Therefore, the system operators use forecasting tools to improve the security of electricity supply, thus promoting the integration of wind in the mix of technologies satisfying the demand.

To promote wind energy integration in the Portuguese electric system, REN – the Portuguese Transmission System Operator (TSO) – developed a wind power forecast tool to support decision-making in its control centre. The wind power prediction tool is continuously improving since 2004. In 2014, there was 4536 MW of wind power connected to the Portuguese grid and 24% of the electricity consumption was supplied by wind generation.

This work presents and assesses a new methodology leveraging a self-adaptive ensemble based method to dynamically combine the wind power forecasts, thus providing the optimal point prediction under the square loss function constrained over the probability simplex. The output is used to estimate the forecast uncertainty in a nonparametric fashion by using order statistics. Relying on standard evaluation metrics, the results show improvements in the performance of both the deterministic forecast and forecast uncertainty estimation over reference methods.

2 Methodology

The proposed methodology is depicted in Fig. 1. REN – the Portuguese TSO – uses a dedicated wind power forecast tool

¹R&D Nester, Centro de Investigação em Energia REN – State Grid, S.A.
Rua Cidade de Goa, nº 4 – B, 2685-038 Sacavém, Portugal

²REN – Rede Eléctrica Nacional, S.A.
Rua Cidade de Goa, nº 4 – B, 2685-038 Sacavém, Portugal

³CEPRI – China Electric Power Research Institute,
No.15 Xiaoying East Road, Qinghe, Beijing 100192, P.R.China

*Corresponding author, e-mail: nuno.pinhosilva@rdnester.com

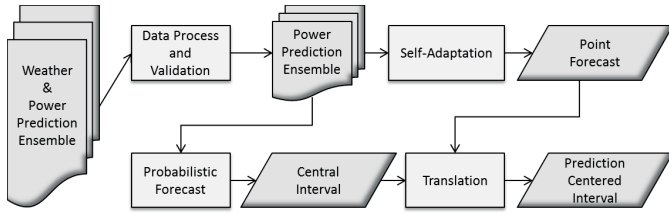


Fig. 1 Wind Power Forecast Methodology

at its control centre. The tool converts weather forecasts into power forecasts and combines them with independent power forecasts, using fixed weights for providing the final wind power forecasts.

In order to provide higher weights to the most accurate elements of the ensemble, this work proposes leveraging the error measurements to dynamically update the weights of the ensemble predictions by using the self-adaptation procedure described in Section 2.2. The result is a univariate time-series of wind power forecasts.

Univariate time-series of point forecast are also called deterministic forecasts, i.e. there is one value for each time index. In complement to this single value information, it is useful for decision-making to provide means for assessing the uncertainty of a given prediction.

Forecast uncertainty estimation relies on the probabilistic forecast framework. It allows producing a bivariate time-series with the upper and lower bound of an interval containing the true power production with a given probability. The intervals' extreme points are the probability distribution's quantiles. A comprehensive review of deterministic and probabilistic wind power forecast systems can be found in [1, 2]. Market driven applications are considered in [3, 4].

2.1 Deterministic Forecast

Deterministic forecasts, or point forecasts, are provided in the form of univariate time-series. The power forecast is given by the expected value of the power production, p , conditioned by the model at the time of issuing the prediction, t , and the look-ahead time, h . Denoting the forecast model at computation time by θ_t , the power forecast may be written as

$$f(t+h) = E\{p|\theta_t, h\}. \quad (1)$$

The REN wind power prediction tool uses raw forecast data provided by four different forecasters. Two forecasters issue weather forecast data (wind speed and air density) while the other two issue wind power forecast data.

The weather forecast providers use only numerical weather prediction (NWP) models while the wind power forecast providers use statistical models together with the NWP models. Each one of the four forecasters issues four deterministic forecasts per day, with 15 minutes time resolution and 7 days ahead time horizon.

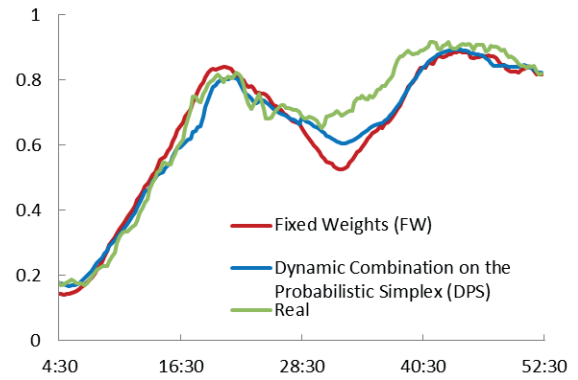


Fig. 2 Deterministic Forecast over horizon. Nominal capacity is 3094.3 MW.

Every time raw data is issued, the forecast tool processes and validates the provider's raw data in several ways, such as wind farm's permit limitation, grid and wind farm maintenance database, and includes online power data (SCADA). For example, considering the weather raw data, one of those processes is the conversion of wind to power by means of the corresponding manufacturer's power curve.

The final deterministic wind power forecast is obtained through a linear combination of the last processed power prediction of each provider.

2.2 Dynamic Combination Forecast via Self-adaptation

The deterministic forecast resulting from the method explained in the previous section is a fixed weighted linear combination of the elements of the ensemble. Thus, the weight of each element is insensitive to the observed forecast error. However, there are forecasters in the ensemble better than others and these are not the same forecasters at all times.

To fully exploit the predictive power of the ensemble, the contribution of each element in the ensemble to the output should be updated over time according to the measured forecast error. This is called dynamic combination forecast. Furthermore, the system should be aware of what forecasters are better at a given time in order to provide the best possible forecast in an automatic fashion. This is self-adaptation.

The self-adaptive ensemble based method provides the deterministic wind power forecasts according to

$$f(t+h) = \sum_{i=1}^N a_i(t,h) f_i(t+h), \quad (2)$$

where the weights $\{a_i(t,h)\}_{i=1}^N$ are functions of the forecast horizon h and the computation time t , such that the resulting deterministic forecast $f(t+h)$ is the optimal point prediction under the square loss function constrained over the probability simplex.

Let $\mathcal{C} = \{x \in \mathbb{R}_+^n : \mathbf{1}^T x = 1\}$ be the probability simplex (where $\mathbf{1}$ is a vector of ones) and $\mathbf{a}(t,h)$ be a vector collecting the combination weights $\{a_i(t,h)\}_{i=1}^N$. The optimal coefficients are given by

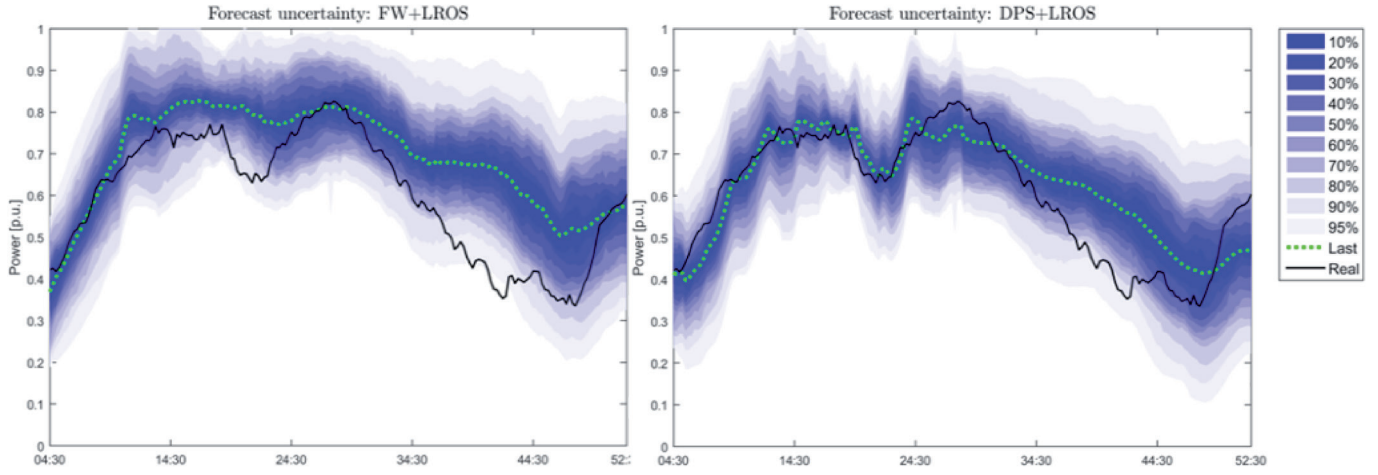


Fig. 3 Forecast uncertainty estimation over horizon. The intervals were computed with the probabilistic forecast approach exploiting the quantile linear regression over the order statistics presented in this work (LROS). The left panel shows intervals centred on the fixed weighted point prediction (FW+LROS), while the right panel shows intervals centred on the optimal dynamic combination over the probabilistic simplex point forecast (DPS+LROS). The nominal capacity is 3094.3 MW.

$$\mathbf{a}^*(t, h) = \arg \min_{\mathbf{a} \in \mathcal{C}} RMSE(\mathbf{F}, \mathbf{r}, \mathbf{a}, t, h), \quad (3)$$

where RMSE is the root mean squared error, \mathbf{F} is a matrix collecting historic records of the ensemble forecasters and \mathbf{r} is a vector with the corresponding historic records of the real wind power, measured at the control centre.

The model variables (\mathbf{F}, \mathbf{r}) change according to a moving window over time (e.g. keeping records up to 15 days previous to the target $t + h$). This moving window defines the internal state of the forecast system. Every time a new element arrives at the ensemble, a deterministic wind power forecast is issued by updating (\mathbf{F}, \mathbf{r}) and computing the weights in (3). In other words, the system outputs the dynamic combination of the ensemble by self-tuning the weights according to the minimization of the RMSE error over the probability simplex, i.e. the forecast system internal state changes over time in an automatic fashion.

Figure 2 shows an example where the improved performance of the self-adaptive dynamic combination over the probability simplex presented here can be qualitatively assessed. Quantitative evaluation is shown in Section 3.1. A side benefit of constraining over the probability simplex is to provide a straightforward interpretation for the combination coefficients: each a_i is the rate of each ensemble predictor in the final deterministic wind power forecast.

2.3 Probabilistic Forecast and Forecast Uncertainty

Probabilistic forecast delves with the prediction of the probability distribution of a given random variable (r.v.). As a probability distribution may be expressed by its quantiles, one may estimate the wind power distribution's quantiles to issue the probabilistic wind power forecast [5, 6].

Let p be the r.v. representing the wind power production and $q_{1-\alpha}$ be the quantile of the wind power production

probability distribution with nominal coverage rate $1-\alpha$, i.e. $P\{p \leq q_{1-\alpha}\} = 1-\alpha$. The quantile estimation retrieves the interval $[0, q_{1-\alpha}]$ of wind power with probability $1-\alpha$; however, this is only one of many intervals with probability $1-\alpha$; another is the following central interval $I_{1-\alpha} = [q_{\alpha/2}, q_{1-\alpha/2}]$.

Central intervals are centred on the median of the wind power distribution and the extreme points of the intervals correspond to quantiles of the probability distribution of the power production.

Another kind of prediction intervals is the intervals centred on specific predictions and largely used when applying the Box-Jenkins method for uncertainty estimation [7]. A prediction centred interval is centred on a given power point forecast and bounds two regions of equal probability, above and below the given central power prediction. Hence, the wind power forecast uncertainty is expressed as central, or prediction-centred, intervals where the power production is expected to be found with a given probability. This probability is called nominal coverage rate. In other words, the output of the forecast uncertainty estimation system is a bivariate time-series of interval forecasts for a given nominal coverage rate rather than univariate time-series of point forecasts.

This work uses a nonparametric approach to wind power probabilistic forecast. Central intervals are established by processing the deterministic forecast ensemble using order statistics. In particular, the incomplete beta function allows to determine the minimum required sample size for a given nominal coverage rate [8]. For example, 39 is the sample size for predicting a central interval with a given nominal coverage rate of 95%. Quantiles are modelled as linear functions of the order statistics. It is shown by Hall and Rieck [9] that the coverage error of quantile estimation by linearly interpolating the sample order statistics is $O(N^{-2})$, $O(N^{-3})$ and $O(N^{-4})$, using, respectively, two, three and five order statistics.

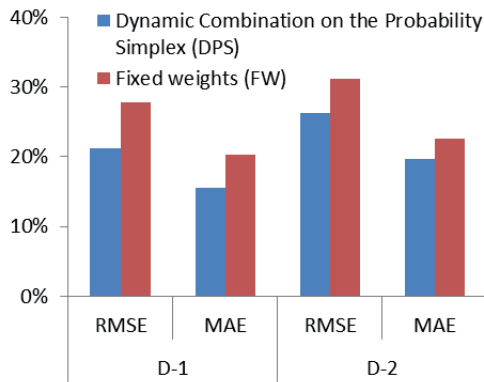


Fig. 4 Deterministic Forecast Evaluation

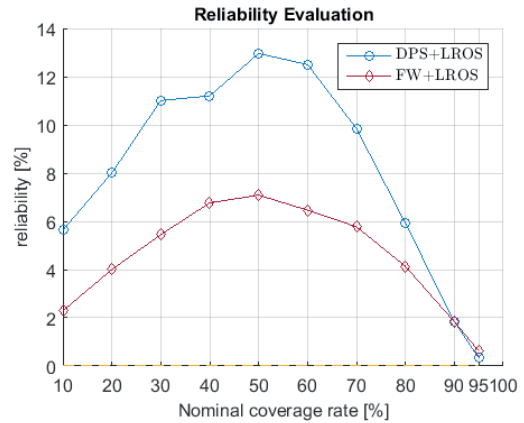


Fig. 5 Reliability evaluation of forecast uncertainty estimation.

Therefore, the presented approach based on linear interpolation of the order statistics enables an efficient distribution-free probabilistic power forecast system, with sample size and coverage error controlled by the required coverage rate.

The forecast uncertainty estimation is provided by prediction-centred intervals. As shown in [4], prediction-centred intervals improve the method’s forecast ability when compared to centred intervals, by improving its empirical coverage rate. Here, the uncertainty power intervals are centred on the optimal point forecast provided by the self-adaptive ensemble based method detailed in the previous section.

Figure 3 shows an example of the forecast uncertainty estimation obtained with the nonparametric approach presented here. The improved performance of the intervals centred on the deterministic forecast provided by self-adaptive dynamic combination over the probability simplex can be qualitatively assessed. In particular, true power is always within the uncertainty range and it can be found more frequently within lower confidence intervals when DPS+LROS is applied. Quantitative evaluation is shown in Section 3.2.

3 Evaluation

The evaluation relies on standard metrics and experiments including three months of daily data from Portugal, namely March, June and September of 2014.

The forecast systems consider 3094.3 MW out of the 4536 MW connected to the grid in 2014. The lower value is used as the nominal power in the results.

3.1 Evaluating the deterministic forecast

The scoring functions are the RMSE (and the MAE) relative to the mean wind power production in each month and averaged over the three months considered on the experiments.

Figure 4 compares the results. It can be seen that the self-adaptive dynamic combination over the probability simplex (DPS) improves the forecast performance.

3.2 Evaluating the forecast uncertainty estimation

The reliability is a nonparametric metric for assessing the quality of probabilistic wind power forecast [10]. It measures how much the method’s empirical coverage rate departs from the nominal coverage rate. Therefore, this metric is also known as calibration.

Positive values represent the method’s ability to guarantee the requested nominal coverage rate, by including more samples within the interval than the percentage defined by the nominal coverage rate, while negative values represent the method’s inability to meet the requested nominal coverage rate.

Figure 5 compares the reliability of the forecast uncertainty estimation when the uncertainty intervals are centred on the deterministic forecast with fixed weights (FW+LROS) and when the uncertainty intervals are centred on the optimal point forecast detailed in Section 2.2 (DPS+LROS).

It can be seen that the forecast uncertainty estimation relying on the self-adaptive method improves the empirical coverage rate. As the reliability characteristic is positive, the nominal coverage rate is guaranteed.

From this result, the proposed forecast uncertainty estimation methodology is to use the optimal point prediction under the square loss function constrained over the probability simplex provided by the self-adaptive ensemble based method in Section 2.2 to centre the uncertainty intervals provided by the linear regression of the order statistics in Section 2.3.

4 Related work

The work in [10] presents results from the ensemble-based approach (EB) [11], the quantile regression approach using splines (QRS) [5] and the ZEPHYR [12] using adaptive resampling, i.e. bootstrap [13] (ZAR). Regarding bootstrap based methods, a recent approach (BELM) [6] exploits bootstrap together with neural networks, trained with extreme learning machine algorithm, to perform quantile regression and provide central uncertainty intervals.

Table 1 Comparison with reference methods G, EB, QR and BELM. The better empirical coverages rates and sharpness are in bold

	Reference Methods						DPS+LROS			
	EB		QR		BELM					
NCR [%]	50	75	50	75	90	95	50	75	90	95
Reliability [%]	13.68	4.57	5.58	9.34	2.46	0.30	12.59	7.88	1.83	0.33
Sharpness [p.u.]	0.28	N/A	0.22	N/A	N/A		0.11	0.18	0.24	0.28

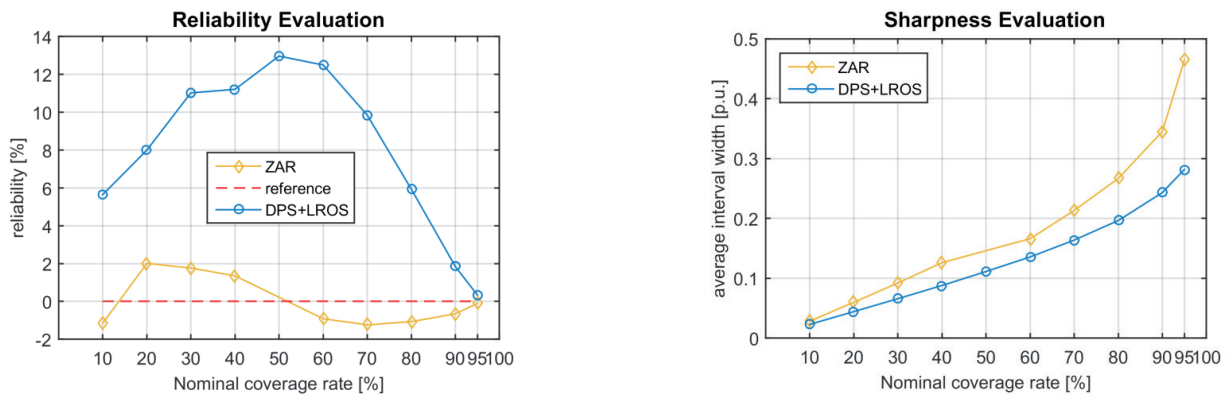


Fig. 6 Comparison with reference method ZAR.

The ensemble-based method share some features with the proposed methodology. It is, therefore, instructive to salient the differences. The EB approach [11] exploits an ensemble of predictions to estimate the power distribution quantiles, as the proposed methodology. However, the EB method provides central uncertainty intervals by using cubic splines and logit regression, which is a different kind of uncertainty interval and a different ensemble processing than the prediction-centred intervals and the linear regression of ordered statistics exploited in this work (c.f. Section 2.3).

The proposed methodology (DPS+LROS) is compared with reference methods using the reliability and sharpness of the estimated uncertainty intervals.

As it was seen in the previous section, reliability provides a way of assessing the method’s ability to guarantee the requested nominal coverage rate, where positive reliability shows a safe margin of operation.

The sharpness metric provides a way to measure the triviality of the forecasted intervals. It is computed by the average length of the intervals: the lower it is, the less uncertain the method is regarding the true power production value. Therefore, a sharp method is a method presenting interval forecasts with low average length.

Table 1 shows the comparison with reference methods EB, QRS and BELM, where NCR stands for Nominal Coverage Rate and N/A means Not Available in the data source. It can be seen that the DPS+LROS reliability is consistently competitive across the range of nominal coverage rates while keeping the forecasted intervals sharp.

Figure 6 shows comparative results with the ZAR method. The proposed methodology presents better empirical coverage characteristics while forecasting sharper intervals.

5 Conclusion

This work introduced a self-adaptive ensemble based method providing optimal point predictions under the quadratic error constrained over the probability simplex.

The forecast uncertainty estimation is obtained by centring a new nonparametric probabilistic wind power forecast approach that leverages the linear interpolation of order statistics to compute the quantiles of the predictive distribution. It is thus efficient and provides sample size and coverage error control by the required coverage rate.

Using standard metrics and 3 months daily data from Portugal, the evaluation of the proposed approach shows the competitive reliability of its prediction intervals, with coverage and sharpness characteristics that compare favourably with the reference methods, thus enabling the perusal of forecast uncertainty in decision-making. Reliable and sharp prediction intervals reduce the uncertainty in market and system operation planning.

Acknowledgement

This work was supported by CEPRI (Chinese Electric Power Research Institute), in the scope of the R&D Nester (Centro de Investigação em Energia REN-SGCC, S.A.) project “Renewable Energy Dispatch Tool”.

References

- [1] Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., Draxl, C. "The State of the Art in Short-Term Prediction of Wind Power." In: *ANEMOS.plus and SafeWind*, 2011.
- [2] Foley, A. M., Leahy, P. G., Marvuglia, A., McKeogh, E. J. "Current methods and advances in forecasting of wind power generation." *Renewable Energy*. 37 (1). pp. 1-8. 2012. DOI: [10.1016/j.renene.2011.05.033](https://doi.org/10.1016/j.renene.2011.05.033)
- [3] Pinson, P. "Wind Energy: Forecasting Challenges for Its Operational Management." *Statistical Sciences*. 28 (4). pp. 564-585. 2013. DOI: [10.1214/13-STS445](https://doi.org/10.1214/13-STS445)
- [4] Pinho da Silva, N., Rosa, L., Pestana, R. "Ensemble-based estimation of wind power forecast uncertainty." In: *European Energy Market (EEM), 2015 12th International Conference on the*. pp. 1-5, 19-22 May 2015. DOI: [10.1109/EEM.2015.7216622](https://doi.org/10.1109/EEM.2015.7216622)
- [5] Nielsen, T. S., Nielsen, H. A., Madsen, H. "Using quantile regression to extend an existing wind power forecast system with probabilistic forecasts." *Wind Energy*. 9 (1-2). pp. 95-108. DOI: [10.1002/we.180](https://doi.org/10.1002/we.180)
- [6] Wan, C., Xu, Z., Pinson, P., Dong, Z. Y., Wong, K. P. "Probabilistic Forecasting of Wind Power Generation Using Extreme Learning Machine." *Power Systems, IEEE Transactions on*. 29 (3). pp. 1033-1044. 2014. DOI: [10.1109/TPWRS.2013.2287871](https://doi.org/10.1109/TPWRS.2013.2287871)
- [7] Box, G. E. P., Jenkins, G. M., Reinsel, G. C. "*Time Series Analysis: forecasting and Control*." 4th ed., Wiley Series in Probability and Statistics, 2008.
- [8] Wilks, S. S. "Statistical prediction with special reference to the problem of tolerance limits." *The Annals of Mathematical Statistics*. 13 (4). pp. 400-409. 1942. DOI: [10.1214/aoms/1177731537](https://doi.org/10.1214/aoms/1177731537)
- [9] Hall, P., Rieck, A. "Improving coverage accuracy of nonparametric prediction intervals." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 63 (4). pp. 717-725. 2001. DOI: [10.1111/1467-9868.00308](https://doi.org/10.1111/1467-9868.00308)
- [10] Pinson, P., Kariniotakis, G., Nielsen, H. A., Nielsen, T. S., Madsen, H. "Properties of quantile and interval forecast of wind generation and their evaluation." In: *Proceedings of the European Wind Energy Conference & Exhibition*. Athens, 2006.
- [11] Nielsen, H. A., Nielsen, T. S., Madsen, H., Badger, J., Giebel, G., Landberg, L., Sattler, K., Feddersen, H. "Wind power ensemble forecasting." In: *Global Wind Power Conference*. Chicago, 2004.
- [12] Nielsen, T. S., Madsen, H., Nielsen, H. A., Landberg, L., Giebel, G., "ZEPHYR - The prediction models." In: *European Wind Energy Association Conference*. Denmark, 2001.
- [13] Pinson, P. "*Estimation of the uncertainty in wind power forecasting*." Ph.D. Thesis - Ecole des Mines de Paris, Paris, 2006.