From Marginal to Simultaneous Prediction Intervals of Wind Power

Ricardo J. Bessa
Center for Power and Energy Systems
INESC TEC
Rua Dr. Roberto Frias
4200 - 465 Porto
Portugal
ricardo.j.bessa@inesctec.pt

Abstract—**The current literature in wind power forecast is focused in generating accurate uncertainty forecasts and communicating this information to the end-user. Multi-temporal decision-making problems require information about the temporal trajectory of wind power for the next hours. Presently, this information is provided through a set of temporal trajectories (or scenarios). This paper aims at contributing with an alternative approach for communicating this information through simultaneous prediction intervals. These intervals include the temporal dependency of forecast errors since they provide information about the probability of having the observed wind power trajectory fully inside the quantiles forming the interval. First, a learning sample of temporal trajectories are generated with the Gaussian copula method and using the marginal prediction intervals. Then, two methods proposed in the literature are used to construct the simultaneous intervals. The quality of these intervals is evaluated for three real wind farms.**

Keywords: Wind power; uncertainty; prediction intervals; decision-making; temporal trajectories.

I. INTRODUCTION

Driven by a new generation of decision-aid tools developed to support the renewable energy integration into the power system and electricity market, research on renewable energy forecasting evolved from point to uncertainty forecast [1]. Examples of tools within the stochastic optimization/decision-making paradigm are: setting the power system operating reserve [2], stochastic unit commitment [3] and optimal wind power bidding [4].

Recent research has been focused on producing uncertainty forecasts, taking into account the form of probabilistic forecasts, risk indices, ramps and trajectories (or short-term scenarios) of wind power generation.

In order to produce a set of prediction intervals, Pinson and Kariniotakis [5] proposed a fuzzy inference model combined with adapted resampling to determine the distributions of forecast errors associated with power output forecast. Bessa *et al.* [6] described a time-adaptive conditional kernel density estimation method with a non-parametric copula for modelling the dependency between numerical weather predictions (NWP) and wind power.

Pinson *et al.* [7] proposed a risk index, named Normalized Prediction Risk Index, which reflects the spread of an ensemble of wind power forecasts for a single look-ahead time or over a forecast period. It consists of a single numerical value (or qualitative value) that provides an *a priori* warning on the expected level of prediction error. Pinson *et al.* [8] proposed a method to generate temporal trajectories of wind power forecasts that include the temporal dependency of forecast errors; a similar method is also proposed by Ma *et al.* [9] to capture wind power variability and uncertainty. Numerical weather prediction (NWP) ensembles, which are a set of NWP produced by perturbing the initial conditions or result from a different parameterization of a NWP model, can be converted into power and also represent a set of temporal trajectories [10]. Bossavy *et al.* [11] described a method based on detecting ramp forecasts from the members of a wind power ensemble, which produces a forecast of wind power ramp event (i.e., prediction intervals with associated probabilities of ramp occurrence).

The EU Project Safewind explored alternative methods to communicate wind power uncertainty to end-users [12]: (a) communicating the variability in wind power fluctuations through the calculation of the cumulative probability of negative and positive changes and the comparison with a Gaussian distribution; (b) using the colour and risk levels currently employed by the Met Office for describing extreme weather events (e.g., ramps, cut-offs); (c) assessing the probability of specific extreme events leading to a binary output; (d) geographic map that enables end-user to identify areas with high forecast errors.

Communication of uncertainty forecasts to end-users is currently an active area of research in different areas. A general discussion about this topic can be found in [13] and [14].

This work was made in the framework of the BEST CASE project ("NORTE-07-0124-FEDER-000056") financed by the North Portugal Regional Operational Programme (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds, through Fundação para a Ciência e a Tecnologia (FCT).

978-1-5090-0191-0/15/$31.00 ©2015 IEEE

Traditionally, the uncertainty in multi-temporal decision-making problems is modelled with a set of temporal trajectories (or short-term scenarios) [15]. Two examples related to the power systems area are the stochastic unit commitment [3] and the storage-wind farm coordination [16]. With this type of representation, it may be computationally very demanding to find the optimal solution in a minimum running time and an end-user may have difficulties in interpreting the temporal trajectories of wind power.

The aim of this paper is to explore the simultaneous prediction intervals concept [17], which is an alternative approach to communicate uncertainty in multi-temporal stochastic decision-making problems. Simultaneous prediction intervals differ from marginal prediction intervals since they take into account the temporal dependency of forecast errors. Two methods proposed by Kolsrud in [18] are used in this paper to construct simultaneous prediction intervals for temporal trajectories generated with a Gaussian copula method [8] for the next 48 hours.

The paper is organized as follows: section II discusses the concept of simultaneous prediction intervals; section III describes the temporal trajectories generation method and the two methods to construct simultaneous prediction intervals; section IV presents results for three real wind farms; section V presents the conclusions and future work.

II. Simultaneous Prediction Intervals

Several methods were developed to generate wind power probabilistic forecasts, represented by quantiles, prediction intervals, probability density function (pdf), moments (e.g., skewness, kurtosis) [19]. A typical output from these methods is illustrated in Fig. 1 (frequently called “fan chart”) through a set of 48 hours-ahead prediction intervals for one wind farm, obtained from the conditional pdf predicted by the quantile-copula method described in [6].

These prediction intervals only represent information from marginal distributions, thus they can be named marginal or pointwise prediction intervals (MPI). Charts similar to the one in Fig. 1 may give misleading information to a decision-maker. For instance, the decision-maker may interpret each one of the quantiles as a temporal trajectory in time, which is not correct. In fact, the marginal intervals can only be interpreted individually for each hour, and are mathematically defined as:

$$\mathrm{Prob}\left\{\mathrm{P}_{\mathrm{t}}^{\tau^L} \le \mathrm{P}_{\mathrm{t}} \le \mathrm{P}_{\mathrm{t}}^{\tau^H}\right\} = \tau^H - \tau^L = 1-\alpha \qquad (1)$$

where τ^H and τ^L are the quantile nominal proportions of the interval limits, α the coverage rate and P_t the wind power in hour *t*.

For instance, in hour 12h00 of the first day there is a probability of *1-α=90%* (limited by quantiles 95% and 5%) that the observed value is within $\mathrm{P}_{\mathrm{t}}^{\tau^L} = 0.22$ and $\mathrm{P}_{\mathrm{t}}^{\tau^H} = 0.87$.

However, temporal trajectories, which include the temporal dependency of forecast errors, can be generated from the probabilistic forecast depicted in Fig.1. An example with 100 scenarios generated by a Gaussian copula (described in section III.A) is depicted in Fig. 2 for the same time horizon. In contrast to the quantiles illustrated in Fig.2, each trajectory (or scenario) represents a possible realization of wind power generation in the time horizon and is equally probable.

This information can be directly included in stochastic optimization problems, but it does not provide clear information to the end-user and can increase cognitive load (concept discussed in [13]). An alternative representation, which is a hybrid of the information contained in Fig. 1 and 2, is the simultaneous prediction intervals (SPI).

Fig. 3 depicts SPI constructed from the scenarios in Fig. 2 and using the Chebyshev distance-based method that will be described in section III.B. These intervals are mathematically defined as:

$$\mathrm{Prob}\left\{\mathrm{P}_{1\to t+T}^{\tau^L} \le \mathrm{P}_{1\to t+T} \le \mathrm{P}_{1\to t+T}^{\tau^H}\right\} = \tau^H - \tau^L = 1-\alpha \quad (2)$$

where *T* is the time horizon, $\mathrm{P}_{1\to t+T}^{\tau^L}$ and $\mathrm{P}_{1\to t+T}^{\tau^H}$ quantiles trajectories between 1 and *T*, and $\mathrm{P}_{1\to t+T}$ the observed wind power trajectory.

Eq. 2 means that the observed wind power is completely contained inside the SPI during all hours of the time horizon *T*. For example, in Fig. 3, the dark blue area means that there is a 10% probability of having the observed wind power trajectory completely inside the interval defined by quantiles 45% and 55%.

The comparison between Fig. 1 and 2 shows the following differences: (a) the SPI with lower coverage are wider than the MPI; (b) the shape of the quantile curves from the MPI is similar to the point (or deterministic) forecast.

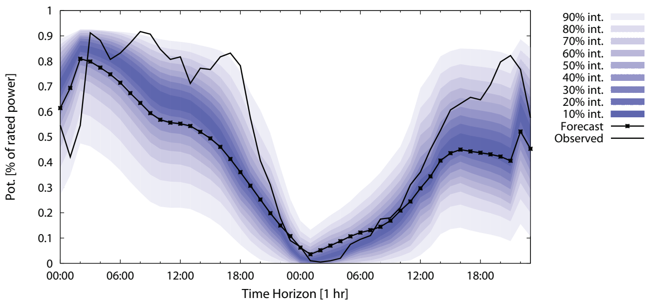

Figure 1. Marginal prediction intervals (MPI) for a 48 hour-ahead time horizon.

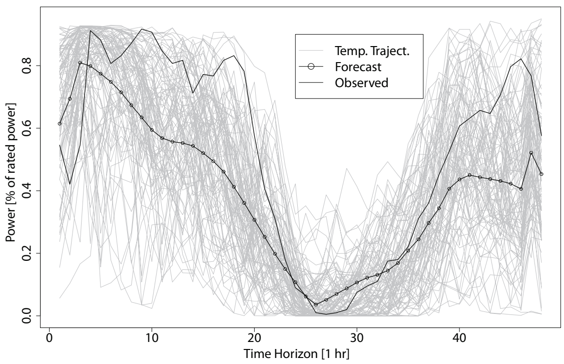

Figure 2. Temporal trajectories (or short-term scenarios) for a 48 hour-ahead time horizon.

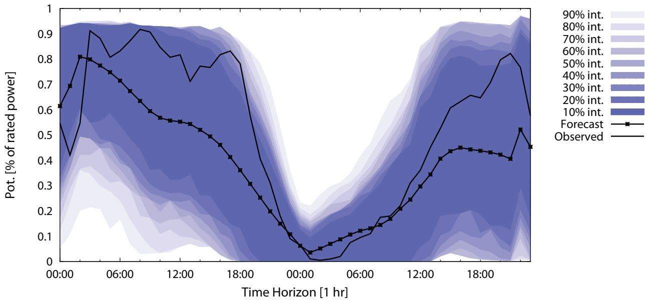

Figure 3. Simultaneous prediction intervals (SPI) for a 48 hour-ahead time horizon.

The SPI are wider in the centre intervals because of the variability in the scenarios depicted in Fig. 2.

The estimation of SPI also results in quantile trajectories that represent a certain probability of having an observed wind power lower or equal to that trajectory. This type of information can be included in multi-temporal decision-making problems, as an alternative to a set of temporal trajectories. However, it is not an objective of this paper to discuss how this information can be included in those problems.

III. Construction of Simultaneous Prediction Intervals

The construction of simultaneous prediction intervals is divided into three steps: (a) generation of probabilistic forecasts, i.e. marginal distributions (see references [5, 6]); (b) generation of a large set of temporal trajectories (described in section III.A); (c) construction of simultaneous intervals (described in section III.B).

A. Gaussian Copula to Generate Temporal Trajectories

The MPI are forecasted for each lead-time by employing the time-adaptive quantile-copula method with the point forecast and wind direction as inputs [6]. The temporal trajectories are generated with the NORTA (NORmal To Anything) method proposed by Cario and Nelson [20], which inspired the method described in [8].

This method generates N temporal trajectories of wind power ($\hat{\mathrm{P}}^{[n]}_{1\to t+T}$) by the following procedure:

1. generates N random vectors Z from a multivariate Gaussian distribution (i.e., Gaussian copula) with zero mean and empirical covariance matrix Σ_Z estimated from a learning sample;
2. transforms Z with Eq. 3 to obtain a random vector $\hat{\mathrm{P}}^{[n]}_{1\to t+T} = \left[\hat{P}_1^{[n]},\ldots,\hat{P}_t^{[n]}\right]$;

$$P_i^{[n]} = F_i^{-1}\left(\Phi\left(Z_i^{[n]}\right)\right) \quad (3)$$

where $F_i^{-1}(\cdot)$ is the inverse of the cumulative distribution function interpolated from the forecasted quantiles (i.e., inverse of the marginal distributions) and Φ the distribution function of a standard normal random variable.

The generated temporal trajectories ($\hat{\mathrm{P}}^{[n]}_{1\to t+T}$) will have the forecasted marginal distributions. The dependency structure between the lead-times is modelled with a Gaussian copula and the covariance matrix can be estimated with different correlation measures [21]. In this paper, the Pearson's correlation coefficient is used to compute the empirical covariance of variable Z_t defined as follows:

$$Z_t = \Phi^{-1}\left(\hat{F}\left(P_t\right)\right) \quad (4)$$

The covariance matrix is computed from a learning sample of past observations of Z.

B. Contruction of Simultaneous Intervals

Departing from a set of temporal trajectories, Kolsrud [18] proposed two methods to construct simultaneous prediction intervals, which are described in the next sub-sections.

1) Adjusted Intervals (AI)

The adjusted intervals (AI) method is very intuitive and basically consists in increasing the coverage of the MPI until the SPI coverage matches $1\text{-}\alpha$. It should be stressed that the learning sample is the set of N temporal trajectories generated by the method described in the previous section and the SPI coverage is given by the number of temporal trajectories completely inside the lower and upper quantiles of the SPI interval.

The method can be summarized as follows:

1. for each lead-time $t=1\ldots T$, increase the MPI uniformly by including the nearest sample point (from the set of N trajectories) above and below the interval limits;
2. compute the coverage of the SPI by counting the number of temporal trajectories completely inside the lower and upper quantiles;
3. if the coverage is less than the nominal value $1\text{-}\alpha$, go to step (1); else, end the algorithm.

Steps 1-3 can be repeated for any α. Note that the final SPI can have a coverage slightly greater than $1\text{-}\alpha$ since in each iteration at least two trajectories are included inside the SPI.

2) Chebyshev-based Intervals (CI)

The Chebyshev-based method (CI) sorts the trajectories according to the Chebyshev distance to the centre trajectory (mean value) of the learning sample. The SPI is the envelope of a subsample containing the $(1\text{-}\ \alpha)\ .N$ trajectories with the shortest distance to the mean.

The method can be summarized as follows:

1. compute the weighted Chebyshev distance (Eq. 5) of each trajectory to the mean trajectory $\bar{\hat{\mathrm{P}}}^{[n]}_{1\to t+T} = \left[\bar{\hat{P}}_1^{[n]},\ldots,\bar{\hat{P}}_t^{[n]}\right]$.

$$d_{WC}\left(\bar{\hat{\mathrm{P}}}^{[n]}_{1\to t+T}, \hat{\mathrm{P}}^{[n]}_{1\to t+T}, s\right) = \max_t\left(\left|\hat{\mathrm{P}}^{[n]}_t - \bar{\hat{\mathrm{P}}}^{[n]}_t\right| \middle/ \sigma_t\right) \quad (5)$$

where σ_t is the pointwise standard deviation calculated for each lead-time *t*;

2. sort the trajectories in ascending order according to its Chebyshev distance and construct a subsample with the first *(1-α).N*;
3. compute the envelope (Eq. 6 – the narrowest interval containing all trajectories in the subsample) of the *m* trajectories in the subsample. The SPI with coverage *1-α* is equal to the envelope.

$$\text{envelope} = \left(\left[\min_{[m]}\left(\hat{P}_t^{[m]}\right), \max_{[m]}\left(\hat{P}_t^{[m]}\right)\right]\right)_{t=1}^{T} \quad (6)$$

Steps 1-3 can be repeated for any *α*.

The distance of Eq. 5 is weighted by the pointwise standard deviation to account for possible heteroscedasticity in the learning sample.

IV. Case-Study Results

A. Description

The case-study consists of the first three real wind farms from the Global Energy Forecasting Competition (GEFCOM 2012) dataset, which is freely available in [22]. Three years of data are available and consists of historical power measurements and weather predictions extracted from the European Centre for Medium-range Weather Forecasts model (ECMWF) with hourly time resolution. The wind power values were normalized between 0 and 1 by the respective rated power of the wind farms.

The last year was used to generate a set of wind power trajectories, estimate the SPI prediction intervals and evaluation. The time horizon *T* is 48 hours-ahead.

In the literature, a vast number of metrics are available to evaluate the quality of MPI and quantiles. Some examples in [23] are calibration, sharpness, quantile score and continuous ranking probabilistic score (CRPS).

However, for SPI it is not straightforward to derive equivalent metrics. The only exception is calibration. For SPI, it is possible to measure calibration by counting the number of times the observed temporal trajectory is fully inside each interval. An SPI is perfectly calibrated if this empirical coverage matches the nominal coverage of each interval.

Sharpness is more difficult to quantify since it would be necessary to find a metric that summarizes the amplitude of the SPI. In this paper, the average size of the SPI is used to measure sharpness. Metrics, such as the quantile score or the CRPS, are strictly proper scoring rules that summarize into a single metric multiple aspects of the uncertainty forecast, such as calibration and sharpness. However, deriving such metric to SPI is still an open line of research and it is not possible to find a single publication covering this issue. Therefore, the evaluation of the SPI in the next section is conducted in terms of calibration and sharpness.

B. Results

This section evaluates the sharpness and calibration of MPI and SPI. The SPI were generated with the two methods described in section III.B [adjusted (AI) and Chebyshev (CI)] and using two different types of scenarios (or temporal trajectories): (a) generated with the Gaussian copula described in section III.A; (b) generated assuming independence between the lead-times.

Fig. 4 depicts, for wind farm 1, the deviation from perfect calibration (i.e., empirical minus nominal coverage) for intervals with nominal coverage ranging between 10% and 90% and obtained with the different models. The MPI show a poor performance in terms of calibration. Moreover, all these deviations are negative, meaning an overestimation of the coverage probability. For instance, for the nominal coverage 10%, the empirical value, which is the number of observed trajectories inside the 55% and 45% quantiles of the MPI, is zero. This means a -10% deviation. The same is valid for the 90% nominal coverage, with only 19.7% of the observed trajectories from the evaluation period inside the MPI.

The calibration of the SPI is much better compared to the one obtained by the MPI. The scenarios generated assuming independence resulted in SPI with the worst performance in calibration, compared to the SPI constructed from scenarios generated with the Gaussian copula, e.g. the highest deviation of the SPI-AI is 8%, while the one of SPI-AI (indep.) is 31%.

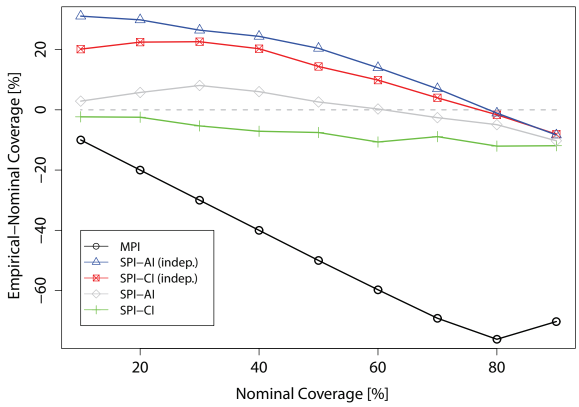

Figure 4. Calibration results for wind farm 1.

Finally, the two methods used to construct SPI achieved different results in terms of calibration. The SPI-AI shows, for coverages between 10% and 60%, an underestimation of the coverage probability (positive calibration) and the SPI-CI shows an overestimation (negative calibration). The maximum deviation is -10% and -12% for the AI and CI methods correspondingly and both for the 90% coverage.

Fig. 5 depicts the sharpness results for the methods discussed in Fig. 4. As discussed in section II, the SPI are wider than the MPI, which results in a higher sharpness.

Methods with a good performance in sharpness have a tendency to present a worse performance in calibration. However, there are also methods that show a poor performance in both metrics. In the case of Fig. 4 and 5, the SPI-AI (Indep.) and SPI-CI (Indep.) methods show a worse performance in both sharpness and calibration, when compared to the SPI constructed from the copula-based temporal trajectories. These results clearly show the importance of including information about the temporal dependency of forecast errors in the generation of temporal trajectories and construction of SPI.

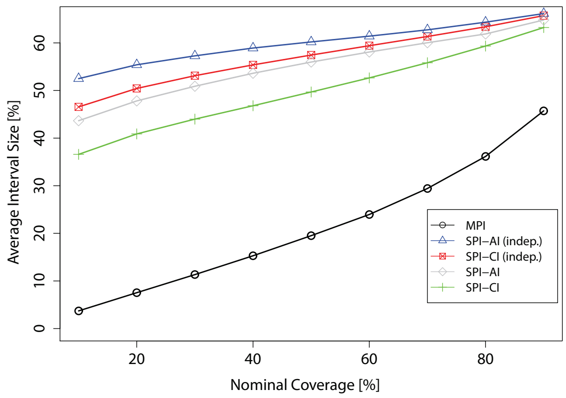

Figure 5. Sharpness results for wind farm 1.

Fig. 6 depicts the calibration results for wind farm 2. The conclusions are analogous to the ones obtained for wind farm 1. However, for this wind farm, the method SPI-CI achieves a lower deviation in calibration compared to SPI-AI. Furthermore, the sharpness of this method (depicted in Fig. 7) is also lower than SPI-AI.

Fig. 8 depicts the calibration results for wind farm 3. In this wind farm, the SPI-AI is the one that achieves the best performance in terms of calibration. However, this method shows a higher sharpness (depicted in Fig. 9) compared to SPI-CI.

These results were consistent with the ones obtained for the other two wind farms, which indicates that the conclusions derived in this paper are independent of the wind farm under analysis.

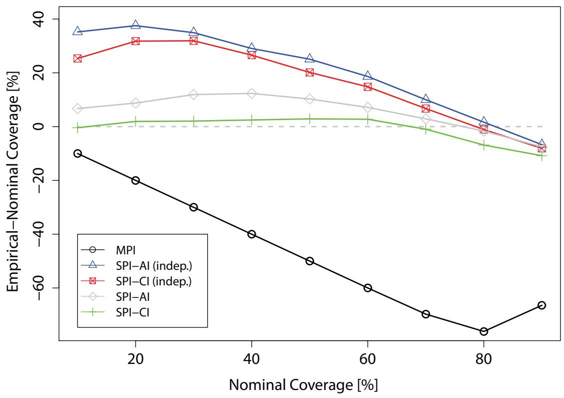

Figure 6. Calibration results for wind farm 2.

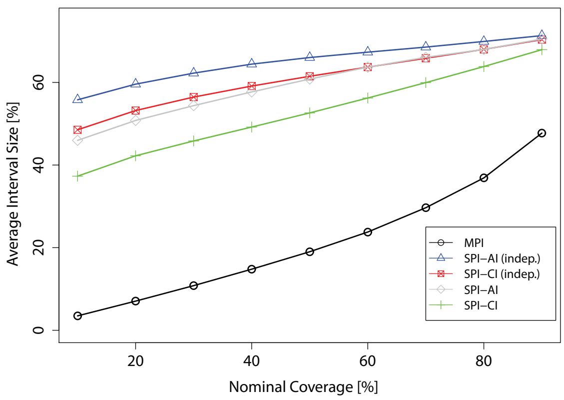

Figure 7. Sharpness results for wind farm 2.

In summary, these results show that SPI, compared to MPI, provide more accurate probabilistic information about the future trajectory of wind generation and the two methods proposed by Kolsrud [18] provide intervals with acceptable calibration. The Chebyshev-based Intervals (CI) showed a good performance in the three wind farms and it is computationally faster than the Adjusted Intervals (AI) method.

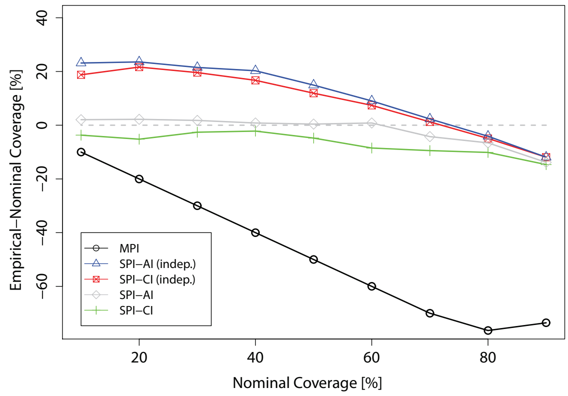

Figure 8. Calibration results for wind farm 3.

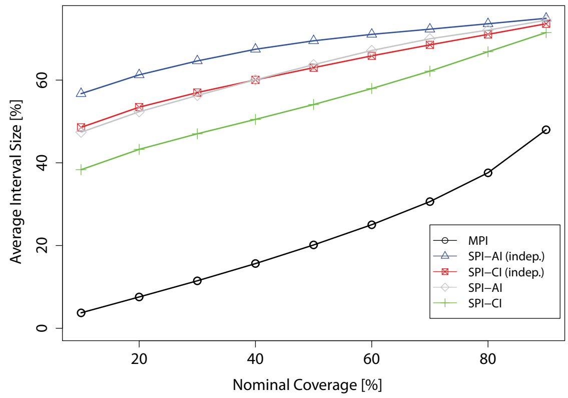

Figure 9. Sharpness results for wind farm 3.

V. CONCLUSIONS

This paper described a methodology for constructing simultaneous prediction intervals for wind power forecasts. The methodology is based on a learning sample of temporal trajectories (or short-term scenarios) of wind power and explores two methods proposed by Kolsrud [18]. The main goal of this work was to describe how the prediction interval should be constructed and represented, in order to provide information to the decision-maker about the temporal evolution of wind power and avoid misconceptions related to the classical representation through marginal prediction intervals.

The following conclusions were derived from the evaluation results for three real wind farms:

a) simultaneous prediction intervals are wider than marginal prediction intervals but provide more reliable probabilistic information about the temporal trajectory of wind power generation;

b) the temporal dependency of forecast errors cannot be neglected when generating the temporal trajectories since this also impacts the simultaneous prediction intervals quality (i.e., both calibration and sharpness);

c) the two methods proposed by Kolsrud [18] show an acceptable performance in terms of calibration (however, the suggested bootstrapping method for adjusting the intervals was not explored in this paper).

This work paves the way towards future avenues of research. The first topic is related to integrating the information from simultaneous prediction intervals in decision-making problems. This information is certainly valuable for multi-temporal problems, however it might also demand for significant changes in the formulation of such problems. An interesting development is provided in [24], where the concept of an uncertainty envelope is much related to these intervals.

The second topic for future work is to derive proper scoring rules to evaluate the overall quality of these intervals. Finally, the third topic is to develop new methods for generating simultaneous intervals, which might be estimated from a learning sample of random vectors or directly from the joint distribution function.

REFERENCES

[1] R.J. Bessa, C.L. Moreira, B. Silva, and M.A. Matos, "Handling renewable energy variability and uncertainty in power systems operation," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 3, no. 2, pp. 156–178, March/April 2014.

[2] M.A. Matos and R.J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp.594-603, May, 2011.

[3] J. Wang, A. Botterud, R. Bessa, H. Keko,V. Miranda, J.S. Akilimali, L. Carvalho, and D. Issicaba, "Wind power forecasting uncertainty and unit commitment," *Applied Energy*, vol. 88, no. 11, pp.4014-4023, Nov. 2011.

[4] A. Botterud, J. Wang, Z. Zhou, R.J. Bessa, H. Keko, J.S. Akilimali, and V. Miranda, "Wind power trading under uncertainty in LMP markets," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 894-903, May 2012.

[5] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1845-1856, 2010.

[6] R.J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renewable Energy*, vol. 40, no. 1, pp. 29-39, April 2012.

[7] P. Pinson, H.Aa. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," *Applied Energy*, vol. 86, nos. 7–8, pp. 1326-1334, July-August 2009.

[8] P. Pinson, H. Madsen, H.Aa. Nielsen, G. Papaefthymiou, and B. Klöckl, "From probabilistic forecasts to statistical scenarios of short-term wind power production," *Wind Energy*, vol. 12, no. 1, pp. 51-62, Jan. 2009.

[9] Xi-Yuan Mam, Y. Sun, and H. Fang "Scenario generation of wind power based on statistical uncertainty and variability," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 894-904, Oct. 2013.

[10] P. Pinson and H. Madsen, "Ensemble-based probabilistic forecasting at Horns Rev," *Wind Energy*, vol. 12, no. 2, pp. 137-155, 2009.

[11] A. Bossavy, R. Girard, and G. Kariniotakis, "Forecasting ramps of wind power production with numerical weather prediction ensembles," *Wind Energy*, vol. 16, no. 1, pp. 51-63, Jan. 2013.

[12] P. McSharry, L. von Bremen, E. Holmgren, N. Siebert, and G. Kariniotakis, Robin Girard, "Predictability measure and communication of information on uncertainty to forecast users," Deliverable Report DP-6.6, Safewind EU Project, Aug. 2012.

[13] A.E. Raftery, "Use and communication of probabilistic forecasts," Working Paper, University of Washington, 2014. arXiv:1408.4812

[14] D. Spiegelhalter, M. Pearson, and I. Short, "Visualizing uncertainty about the future," *Science*, vol. 333, no. 6048, pp. 1393-1400, Sept. 2011.

[15] R. T. Rockafellar and Roger J.-B. Wets, "Scenarios and policy aggregation in optimization under uncertainty," *Mathematics of Operations Research*, vol. 16, no. 1, pp. 119-147, 1991.

[16] E.D. Castronuovo, J. Usaola, R.J. Bessa, M.A. Matos, I.C. Costa, L. Bremermann, J. Lugaro, and G. Kariniotakis "An integrated approach for optimal coordination between wind power and hydro pumping storage," *Wind Energy*, vol. 17, no. 6, pp. 829-852, June 2014.

[17] V. Chew, "Simultaneous prediction intervals," *Technometrics*, vol. 10, no. 2, pp. 323-330, 1968.

[18] D. Kolsrud, "Time-simultaneous prediction band for a time series," *Journal of Forecasting*, vol. 26, no. 3, pp. 171-188, April 2007.

[19] C.Monteiro, R. Bessa, V.Miranda, A. Botterud, J.Wang, and G. Conzelmann, "Wind power forecasting: state-of-the-art 2009," Report ANL/DIS-10-1, Argonne National Laboratory, Nov. 2009.

[20] M.C. Cario and B.L. Nelson, "Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix," Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1997.

[21] S. Ghosh and S.G. Henderson, "Chessboard distributions and random vectors with specified marginals and covariance matrix," *Operations Research*, vol. 50, no. 5, pp. 820-834, Oct. 2002.

[22] T. Hong, P. Pinson, S. Fan, "Global energy forecasting competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 357-363, April-June 2014.

[23] R. Girard, P. Pinson, J. Juban, and G. Kariniotakis, "Towards the definition of a standardized evaluation protocol," European Project ANEMOS.plus, Deliverable D-1.3, Jan. 2009.

[24] H. Nosair and F. Bouffard, "Flexibility envelopes for power system operational planning," *IEEE Transactions on Sustainable Energy*, In Press, 2015.