Simulating scenarios and prediction intervals in wind power forecasting with the Beta distribution

Alvaro Jaramillo[1], Ismael Sánchez[2], Edgardo Castronuovo[1] and Julio Usaola[1]

[1] Department of Electrical Engineering
[2]Deparment of Statics
Carlos III of Madrid University
Campus of Leganés – 28911 Madrid (Spain)
Phone:+34 91 624 9999, fax:+34 91 624 9430
E-mail: ajaramil@ing.uc3m.es, ismael@est-econ.uc3m.es, castronuovo@ieee.org and jusaola@ing.uc3m.es

1. Introduction

When dealing with wind power generation in a market environment, one of its most challenging issues is the uncertainties associated to the wind power production forecasts. These uncertainties increase the economic risk of the wind power trade and could cause penalties to be assumed by the producer. Wind power forecast tools give an expected production for a short-term period, two or three days ahead. This information is very useful when the agents go to the energy market. However, errors in the forecasting values must be always assumed.
The current work presents a statistical tool based in historical data to generate scenarios of wind power production, which can represent a realistic assumption about the future generation. The power forecast is obtained from SIPREOLICO [1], a prediction tool very used in Spain. Instead of using Gaussian distributions, Beta distributions are utilised in the representation. These distributions allow represent more accurately the real behaviour of the wind power production, [2]. The scenarios can be used to determine the best operation of renewable production in one area of the power system, sizing storage devices, programming reserves and develop market strategies.

Keywords: wind power, error analysis, forecast wind production, market policies.

2. Proposed methodology

In [3], Non-Parametric statistics for the power wind forecast are used, aiming to obtain prediction intervals in the forecasting production. The present analysis considers statistical distributions to determine the expected variation in the wind power production in the near future.
From the historical data, the real productions and the wind power forecasts for each interval of the past, calculated previously to the operation, are obtained. With these values, the forecasting errors in the wind power production, as the difference between real and estimated productions, are calculated. The forecasting errors in the past could be assumed as normally distributed [4] with $\mu_\varepsilon = 0$, and can be calculated through equation (1).

$$\begin{aligned} \hat{e}_{t|t-k} &= p_t - \hat{p}_{t|t-k} \\ \varepsilon_t &= \left[\hat{e}_{t|t-1}, \hat{e}_{t|t-2}, \ldots, \hat{e}_{t|t-k}\right]^T \end{aligned} \tag{1}$$

The prediction errors are assumed to be stationary of second order, with covariance matrix Σ_t, as showed in (2).

$$\Sigma_t = E(\varepsilon_t \varepsilon_t^T) \tag{2}$$

The covariance matrix Σ_t concentrates all the information available in the historical data and the interdependence characteristics of the forecasting errors. Using this matrix and the wind power forecast, $\hat{p}_{t+k|t}$, a number of scenarios representing most of the possibilities of wind power production can be calculated. The first step is generating a multivariate Gaussian variable that simulate the behaviour of the forecasted errors, X_k^s. This multivariate variable is transformed in Uniform distribution, as in (3).

$$Y_k^s = normcdf(X_k^s, \mu_k, \sigma_k) \tag{3}$$

Where μ_k is a zero vector and σ_k is a vector of standard deviations errors, $std(\hat{e}_{t|t-k})$.
The next, the information obtained from the historical data allows calculating a Beta distribution, more adequate to represent the wind power production forecast [2]. The possible scenarios in the future for the wind power production, $p_{t+k|t}^s$, are calculated by using a statistical combination of the information present in both Beta parameters at each time, and the interdependence coming from the forecasting errors, present in the variable Y_k^s. Therefore, the possible wind power production in a period of the future $p_{t+k|t}^s$ is calculated as in (4).

$$p_{t+k|t}^s = beta^{-1}(Y_k^s, \hat{\alpha}_{t,k}, \hat{\beta}_{t,k}) \tag{4}$$

In (4), Beta parameters $\hat{\alpha}_{t,k}$ and $\hat{\beta}_{t,k}$ can be calculated by using (5) and (6).

$$\hat{\alpha}_{t,k} = \hat{\mu}_{t,k}\left(\frac{\hat{\mu}_{t,k}(1-\hat{\mu}_{t,k})}{\sigma_{t,k}^2} - 1\right) \tag{5}$$

$$\hat{\beta}_{t,k} = (1-\hat{\mu}_{t,k})\left(\frac{\hat{\mu}_{t,k}(1-\hat{\mu}_{t,k})}{\sigma_{t,k}^2} - 1\right) \tag{6}$$

Where the mean value is the expected production in the next future, $\hat{\mu}_{t,k} = \hat{p}_{t+k|t}$; and the variance is a approximation obtained from the sample of the errors in the past $\sigma_{t,k}^2 = \mathrm{var}(\hat{e}_{t|t-k})$. To exactly calculate the variance, the expected value could be also considered.

3. Results

The Beta distribution is used to generate possible scenarios of wind power generation in the time ahead, $t+k$. In Fig. 1, a hundred of these scenarios and the prediction, $\hat{p}_{t+k|t}$, are shown. In Fig. 2, the confidence interval created by the all scenarios available, the power forecast, $\hat{p}_{t+k|t}$, and the real measure, p_{t+k}, are represented.

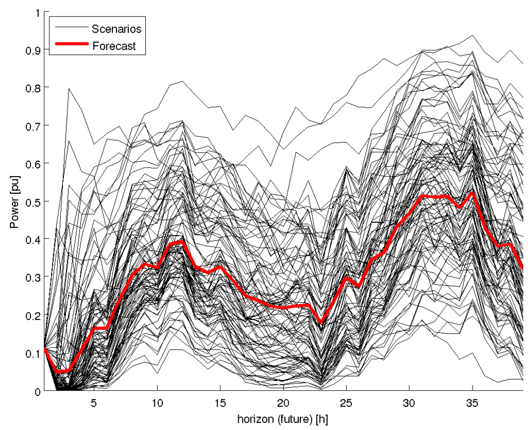

Fig. 1. Possible power scenarios and the power forecast.

In Fig. 2, the error intervals in the wind power production are represented as bands, showing the probable excursion of the forecasting values in the near future. The principal advantage of this methodology is the possibility of quantify economical risks, making easier to develop operational strategies.

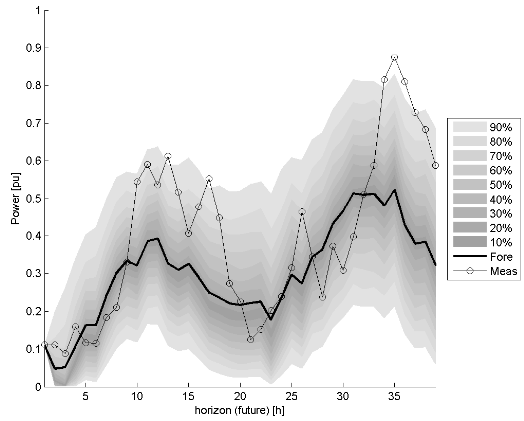

Fig. 2. Confidence intervals based on the simulated power scenarios.

4. Conclusion

The methodology presented here allows calculating the prediction intervals for the wind power production, through the scenarios based on the historical data and the power forecasts obtained by a statistical tool. The calculation of prediction intervals can be very promising, making easier to understand the behaviour of the real variables in the near future. These prediction intervals can be used to calculate the optimal operation of distributed Control Centres for renewable production, economical investment decisions, optimal operation of systems with storage capabilities and other issues.

References

[1] G. Gonzalez, B. Diaz-Guerra, F. Soto, S. Lopez, I. Sanchez, J. Usaola, M. Alonso, y M.G. Lobo, “SIPREOLICO-Wind power prediction tool for the Spanish peninsular power system,” *Proceedings of the CIGRE 40th General Session & Exhibition. Paris (France)*, 2004.

[2] H. Bludszuweit, J. Dominguez-Navarro, y A. Llombart, “Statistical Analysis of Wind Power Forecast Error,” *Power Systems, IEEE Transactions on*, vol. 23, 2008, págs. 983-991.

[3] P. Pinson, H. Madsen, H.A. Nielsen, G. Papaefthymiou, y B. Klöckl, “From probabilistic forecasts to statistical scenarios of short-term wind power production,” *Wind Energy*, vol. 12, 2009, págs. 51-62.

[4] R.D. Cook y S. Weisberg, *Residuals and Influence in Regression*, 1982.