

Net load forecasting in presence of renewable power curtailment

P. M. Fonte

ISEL-Instituto Superior de Engenharia de Lisboa,
Instituto Politécnico de Lisboa
Lisboa, Portugal
pfonte@deea.isel.ipl.pt

¹Cláudio Monteiro

^{1,2}Fernando Maciel Barbosa
¹University of Porto, Faculty of Engineering
²INESC TEC
Porto, Portugal
cdm@fe.up.pt, fmb@fe.up.pt

Abstract— This paper analyzes a real case study based on an islanding power grid, where there is wind power curtailment during the grid operation. This curtailment skews the wind power production database creating a huge challenge to the overall power production forecast. Thus, it is presented a solution which has allowed more accurate forecasts in order to improve the renewable production and reduce the fuel consumption in thermal power plants. The proposed filtering approach demonstrated to be a good solution allowing wind power forecasts with less error and *net load* forecasts with more accuracy.

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

I. INTRODUCTION

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

The developed models for load forecast generally get good results with lower deviations from the measured values since load profiles follow a characteristic pattern. When compared with load forecast, the prediction of RES, due to its variability and intermittence, presents much bigger challenges. Thus, beyond the developed point forecasts techniques there is also the need to incorporate uncertainty in the forecasting. Despite addressing other types of RES (namely solar and hydro), wind generation is presented as the main source of generation uncertainty in power systems scheduling, grid operation and market environment. There are three major factors influencing the uncertainty of wind power

forecast, namely, the Numerical Weather Prediction (NWP), the conversion of wind-to-power (W2P) due to the nonlinearity of the power curve and terrain complexity. Being the NWP and the clouds' dynamic the main source of uncertainty in the case of solar photovoltaic since conversion is well defined. Regarding hydro power forecasting systems, the uncertainty generally propagates from the NWP model through the rainfall-run-off model. These models are limited by their representation of flow dynamics, which main problem is not the representation of the dynamic but knowing the local parameters [3]. Uncertainty created by these errors has a great impact on power systems scheduling since the forecasted values at the beginning of the scheduling process can be quite different from those on operation periods. In a traditional and conservative point of view, generally, the uncertainties are compensated using conservative decisions, like over-designing the equipment or overestimating the operational parameters based on worst-case. This approach, though being secure, may lead to significant results' deterioration of the optimization problem. To overcome this situation, several uncertainty models are provided in literature, as moments of distributions, set of quantiles or interval forecasts, probability mass functions and probability density functions (parametric and non-parametric)[4],[5]. Due to its high installed power capacity all over the World, wind power forecast gathers the majority of researchers' attention, and the higher number of published works.

This paper is organized as follows: section II characterizes the problem; section III presents the proposed forecasting methodology; in section IV the case study is presented and the performances of spot and probabilistic forecasts are assessed; section VI presents the conclusions.

II. PROBLEM CHARACTERIZATION

Scheduling challenges are enhanced in island's systems with low rated power, without connection to the large continental grids and without storage capacity. Great variations on renewable production may introduce stability problems in the network, which can originate generation or load shed and, at limit, blackouts.

The island power system under study is composed by one thermal power plant with 8 units based on heavy fuel, plus 2 geothermal with 5 units, 7 mini-hydro and 1 wind power plants with 9 units. The yearly average geothermal power production is 19.2 MW (meaning approximately 42% of the yearly average load). This generation acts as base of the load diagram and does not contribute to the load follow. On the other hand, hydro generation with 3 MW of rated power has a small weight in the overall production. Therefore, the effective load follow has to be done by an efficient management between thermal and wind production. Analysis of production's datasets confirms that all this renewable production helps to decrease the thermal production during peak load periods. However, during off-peak periods the system is already saturated with renewable power production. Additionally, to the system operator, for operational security reasons it is mandatory that, at least, two thermal units must be on-line. It is an attempt to avoid the complete loss of thermal production due to outages and representing a minimum production of 12.85 MW. In fig. 1, the hourly average thermal production, as well as the sum of minimum technical limits of 2 units, from 0h00 of December 1st until 23h00 of December 31st of 2013 is shown.

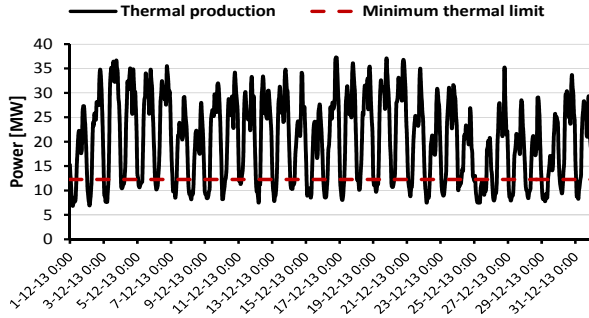


Figure 1. Thermal production and minimum technical limits

As observable in fig.1, during the off-peak periods, renewable production is generally so high (or the load so low) that it obligates the thermal units to work below the minimum technical limits with poor efficiency and high fuel consumption. During 2012 the thermal machines worked below their minimum limits for 17% of the year while in 2013 this percentage was 20.6%. To minimize this situation, especially during off-peak periods, the wind power production is preventively limited with the consequent waste of renewable resources (which occurred during 30.5% of the set of 2012 and 2013). On the other hand, an extreme reduction of the thermal committed capacity can lead to a situation wherein the spinning reserves are not sufficient to handle great variations of load, renewable production or generation outages. Therefore, due to the uncertainty in load and renewable production, sometimes it's hard to find a complete robust/economic scheduling solution. To face this problem, there is a clear necessity of an efficient method to forecast load and renewable production in order to know the real thermal production's needs. This will allow costs' and emissions' reduction and the optimization of the number and the allocated power of thermal units.

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

III. FORECASTING METHODOLOGY

Thermal production forecasting is characterized as *net load* [7],[8] and it is calculated by the differences between forecasted load and the sum of forecasted renewable production. These forecasts, further than spot values, incorporate uncertainty by probabilistic forecasting with the *net load* resulting from the convolution of the various forecasting probabilistic distributions. The forecasted *pdf* were based on the *Nadaraya-Watson* estimator (1) which allows the estimation of a random variable Y , when the explanatory random variable X is equal to x [9]. This conditional density estimation can be seen as a generalization of regression, since conditional density estimation aims at obtaining the full probability density function $\hat{f}_{y|x}(y|x)$ [10].

$$\hat{f}_Y(y_{t+k} | X = x_{t+k|t}) = \frac{\hat{f}_{Y,X}(y_{t+k}, x_{t+k|t})}{\hat{f}_X(x_{t+k|t})} \quad (1)$$

In (1) $\hat{f}_{Y,X}(y_{t+k}, x_{t+k|t})$ is the estimated joint density function and $\hat{f}_X(x_{t+k|t})$ is the marginal density of X . However, since the joint and marginal densities are not known, they can be determined with a nonparametric kernel estimator [9]. In case of power forecast, it consists on the estimation of the future conditional *pdf* of power ($p_{t+k|t}$) for each look-ahead time step $t+k$, given a set with N pairs of samples (p_n, x_n) summarizing all information available up to instant t . Each pair consists on a set of explanatory variables X_n (variables which can explain the power production) and the corresponding value of variable to be predicted P_n . In this process it is assumed that explanatory variables $x_{t+k|t}$ are known at time t for each time-step ahead $t+k$ and p_{t+k} is the power forecasted for look ahead time $t+k$. As the random variable can depend on several explanatory variables, a multivariate KDE can be used and applied to (1). The conditional density estimator results from (2)[9],[11].

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

In (2) N is the number of samples, D is the number of variables and K_j is the kernel function for each variable X_j . The kernel function chosen for all variables was the normal

distribution. Although, in the case of wind direction, wrapped normal distribution was the choice. The parameter h_j is the bandwidth of each kernel around each sample and it controls the smoothness of the estimation. The optimization of the bandwidth h_j was done with Leave-One-Out Cross Validation (LOOCV) technique [12].

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

The load, hydro and geothermal production forecasts do not exhibit considerable challenges since the explanatory variables are based in NWP and historical datasets. The considered variables for load forecasts were hour of the day, day of the week, week of the year and hourly temperature. Regarding hydro, rain forecasts and hydrological power potential (*HPP*) [14] were considered. For geothermal power forecasts a set point resulting from a moving average was considered. In case of load, greater errors may arise if the real conditions are not sufficiently envisaged in the dataset for a certain forecast moment; for instance, national or regional holidays, abnormal temperatures for a certain period of the year, general strike, among others. Concerning geothermal units, the power production is dependent on a renewable, but easy to control, resource. Here, the production is defined by set points and it remains relatively constant around the set point. In this circumstance, the main source of the deviation between what was forecasted and the real production was the unexpected outages of some units. The hydro power plants present a negligible storage capacity meaning that power production cannot be delayed from the moment when it rains until the moment where there is the necessity of power production. On the other hand, as the watersheds are not large enough to introduce a significant delay between the rain period and the production, the hydro power forecast depends merely from the rain forecasts. These are the main source of errors for which an accurate forecast is needed.

Wind power forecasts introduce a different kind of challenge as the measured values of production cannot be fully linked with the explanatory variables due to wind power curtailment. Even with accurate forecasts of the explanatory variables (wind speed and direction), remarkable errors can occur. In fig. 2 the measured hourly average wind speed and measured power is shown. It is clear that there is a large amount of wind power values which do not correspond to measured wind values. These differences can result from malfunctions of the measuring equipment, unexpected units outages and wind power curtailment. Very high values of wind speed are another source of uncertainty, since installed turbines are equipped with “*software for storm regulation*”. In spite of cutting production for velocities above the maximum, it regulates the pitch angle of the blades in order to reduce the power production. Without this information, power forecasts for velocities above the maximum value

become hard to forecast. That said, it is clear that the curtailment process will introduce very significant errors between the wind speed prediction and the measured power, skewing the dataset. In this case it was necessary to do some data pre-processing, gathering the information disclosed by the system operator about wind generation limits in each hour. Thus, there was the necessity to filter all these occurrences in the dataset, replacing the wind power curtailed values by “real” values which should be measured in absence of limitation.

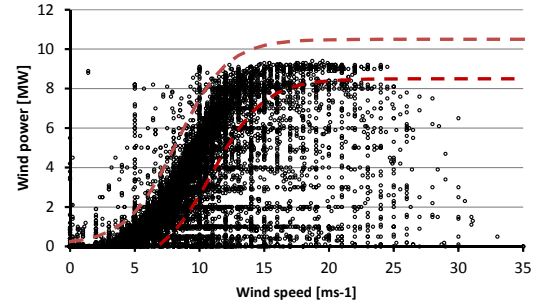


Figure 2. Measured wind speed and power with outlier filtering

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

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

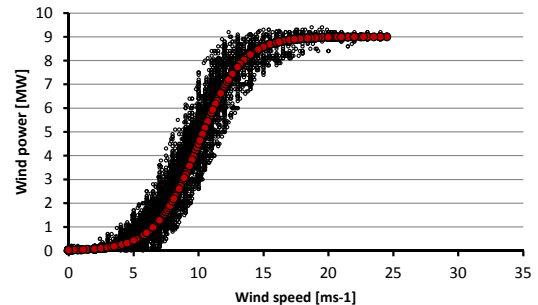


Figure 3. Resulting “theoretical” power curve

One of the informations given by the system operator, beyond the power produced and wind measured, is the wind power limitation. After the removal of the values resulting from bad measurements and outages, it is possible, knowing the periods where there was curtailment and applying (3) to those periods, to have an idea of the wind power forecast in absence of curtailment. With this, the skewness of the dataset can be minimized reducing the non-controlable errors and only focusing on the errors that were delivered from forecastings. In fig. 4 the wind power production, the wind power limitation, as well as the theoretical wind power

production which results from equation (3) from 0h00 of 28th January until 23h00 of 3rd February 2014 is depicted. It is observable that, mainly during off-peak periods, there are notorious differences. The theoretical wind power production should be understood as the wind power values that would be measured in the absence of limitation. The difference between the theoretical and the measured wind energy during this period was approximately 278 MWh.

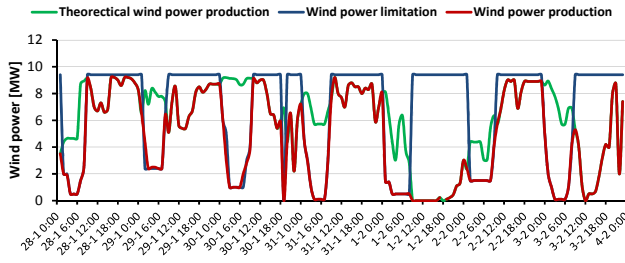


Figure 4. Measured and theoretical wind power

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

IV. CASE STUDY

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

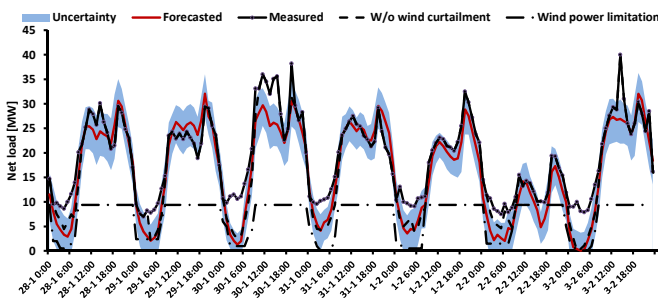


Figure 5. Forecasted and measured *net load*

As a first analysis it is clear that the system operator picked wind power limitation during all off-peak periods. Considering that, when there is an effective power curtailment, the *net load* tends to grow. The fact that in all off-peak periods the *net load* with wind power curtailment tends to be higher than the forecasted is explained. Notice that this analysis is done under the assumption that there were no notable errors in the remaining load, hydro and geothermal forecasts. Excluding some cases, as in 30th January and in 3rd February, the forecasts outside the off-peak periods, present a reasonable fitting with the measures. The measured values are also reasonably covered by the uncertainty interval. The

performances of the spot forecasts can be assessed by some indicators as the BIAS, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Standard Deviation of Errors (SDE) [15]. In table I the assessment of spot forecasts against the real measured and the theoretical values in absence of curtailment are shown.

TABLE I. SPOT FORECASTS ASSESSMENT

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

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

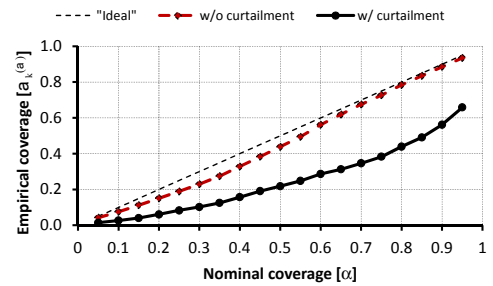


Figure 6. Reliability diagram

It is clear that the forecasting method used in this approach tends to systematically underestimate the uncertainty since forecasted quantiles proportions (nominal coverage) are lower than the empirical ones. Thus, the values of *net load* outperform all estimated quantiles, meaning that the probabilistic forecasts have an associated bias. However, these results are much lower than those reached with wind curtailment. A more intuitive way to analyze the bias of the probabilistic forecasting methods is subtracting the empirical coverage $\hat{a}_k^{(\alpha)}$ to the nominal coverage α [11],[15] resulting in the reliability diagram shown in fig. 7.

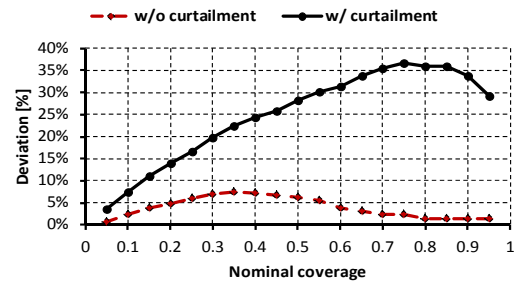


Figure 7. Reliability diagram (deviation from “ideal” reliability)

It is visible that the uncertainty was underestimated for all predicted quantiles. It should be noticed that the *net load* results from four different variables forecasts, with different values and profiles of uncertainty. For the same dataset, the sharpness was calculated, as shown in fig. 8. It is an indicator of the usefulness level of the predictions, which are correspondent to the ability of probabilistic forecasts to concentrate the probabilistic information, in other words the “level” of uncertainty [4]. The values of the sharpness are relatively low, with a nominal coverage of 0.9, corresponding to 19.6% of the maximum *net load* of validation set (48.3 MW). Analyzing the results it was concluded that it must be a trade-off between the reliability and the sharpness, because improving reliability usually will worsen the sharpness [11],[15]. Low values of sharpness can lead to “narrow” uncertainty intervals which can result in underestimation or overestimation of the uncertainty, with consequent degradation of the reliability.

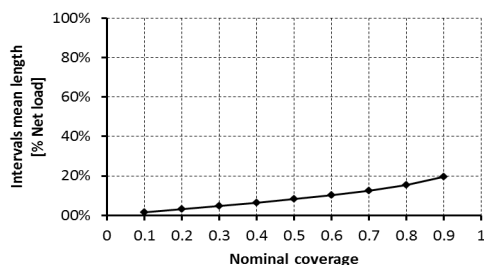


Figure 8. Sharpness diagram of the probabilistic *net load* forecast

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

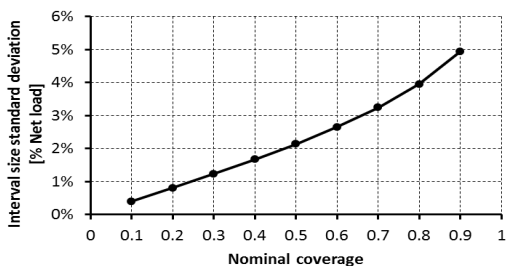


Figure 9. Resolution of the probabilistic *net load* forecast

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

and, consequently, the uncertainty profile does not significantly change along the dataset.

V. CONCLUSIONS

With the proposed technique it is possible to minimize net load forecasting errors introduced by wind power curtailment. Results showed a forecasting errors reduction, especially the bias. Reliability reveals a clear improvement in comparison with the situation of wind power curtailment. A sharpness value of 20% of the maximum net load presents an acceptable trade-off with the reliability. It is conceivable to reduce the waste of “clean” energy and to improve the scheduling of thermal units because the accuracy of the thermal necessities is improved. A deeper analysis of each forecasted variables may provide even better results.

REFERENCES

- [1] S. Pelland, J. Remund, J. Kleissl, T. Oozeki, and K. De Brabandere, “Photovoltaic and Solar Forecasting: State of the Art,” Int. Energy Agency. Rep. *IEA PVPS T14*, Oct. 2013.
- [2] M. Lydia and S. Suresh Kumar, “A comprehensive overview on wind power forecasting,” in *Proc. 2010 IPEC Conf.*, pp. 268–273.
- [3] R. Hostache, P. Matgen, M. Montanari, C. Fosty, and L. Pfister, “Assessment and reduction of hydro-meteorological model forecasting uncertainty using an autoregressive error model and bivariate meta-gaussian density,” in *Proc. 2009 4th IASME / WSEAS International Conference on Water Resource, Hydraulics & Hydrology*, pp. 110–118.
- [4] J. Juban, “Uncertainty estimation of wind power forecasts: Comparison of Probabilistic Modelling Approaches,” in *Proc. 2008 EWEC - European Wind Energy Conference & Exhibition*, pp. 1–10.
- [5] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, “Wind Power Forecasting: State-of-the-art 2009” Argonne National Laboratory., Chicago, IL, Tech. Rep. ANL/DIS-10-1, Nov. 2009.
- [6] S. Fink, C. Mudd, K. Porter, and B. Morgenstern, “Wind Energy Curtailment Case Studies May 2008 — May 2009,” Exeter Associates, Inc., Columbia, MD, Tech. Rep. NREL/SR-550-46716, Oct. 2009.
- [7] Y. V. Makarov, P. V. Etingov, and J. Ma, “Incorporating Uncertainty of Wind Power Generation Forecast Into Power System Operation, Dispatch, and Unit Commitment Procedures,” *IEEE Trans. Sustain. Energy*, vol. 2, no. 4, pp. 433–442, Oct. 2011.
- [8] S. Jin, A. Botterud, and S. M. Ryan, “Temporal Versus Stochastic Granularity in Thermal Generation Capacity Planning With Wind Power,” *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2033–2041, Sept. 2014.
- [9] S. Demir and O. Toktamis, “On The Adaptive Nadaraya-Watson Kernel Regression Estimators,” *Hacetep J. Math. Stat.*, vol. 39, no. 3, pp. 429–437, Mar. 2010.
- [10] G. Fu, F. Y. Shih, and H. Wang, “A kernel-based parametric method for conditional density estimation,” *Pattern Recognit.*, vol. 44, no. 2, pp. 284–294, Feb. 2011.
- [11] R. J. Bessa, V. Miranda, A. Botterud, and J. Wang, “Time Adaptive Conditional Kernel Density Estimation for Wind Power Forecasting,” *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, 2012.
- [12] J. M. Liu, R. Chen, L.-M. Liu, and J. L. Harris, “A semi-parametric time series approach in modeling hourly electricity loads,” *J. Forecast.*, vol. 25, no. 8, pp. 537–559, Dec. 2006.
- [13] P.M. Fonte, “Advanced Forecast and Scheduling of power Systems with Highly Variable Sources,” Ph.D. dissertation, Dept. Electrical Eng., Faculty of Engineering, Univ. of Porto, Porto, 2015.
- [14] C. Monteiro, I. J. Ramirez-Rosado, and L. A. Fernandez-Jimenez, “Short-term forecasting model for electric power production of small-hydro power plants,” *Renew. Energy*, vol. 50, pp. 387–394, Feb. 2013.
- [15] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. Skov, “A Protocol for Standardizing the performance evaluation of short term wind power prediction models,” in *Wind Engineering, Multi-Science Publishing*, 2005, pp. 475–489.