Renewable Power Forecast to Scheduling of Thermal Units

Pedro M. Fonte[1,3], Bruno Santos[2], Cláudio Monteiro[2,3], João P.S. Catalão[4], and Fernando Maciel Barbosa[3,5]

[1] ISEL – Lisbon Superior Engineering Institute, Lisbon, Portugal
[2] Smartwatt,S.A, [3]University of Porto, [5]INESC TEC PORTO, Porto, Portugal
[4] University of Beira Interior, Covilhã, Portugal
pfonte@deea.isel.pt, brunosantos@smartwatt.com, catalao@ubi.pt
{cdm,fmb}@fe.up.pt

Abstract. In this work is discussed the importance of the renewable production forecast in an island environment. A probabilistic forecast based on kernel density estimators is proposed. The aggregation of these forecasts, allows the determination of thermal generation amount needed to schedule and operating a power grid of an island with high penetration of renewable generation. A case study based on electric system of S. Miguel Island is presented. The results show that the forecast techniques are an imperative tool help the grid management.

Keywords: Probabilistic renewable power forecast, kernel density estimator, Power scheduling.

1   Introduction

The increasing introduction of electric energy production with renewable sources, mainly those with high variability, has created several challenges to the energy networks operators, predominately in the scheduling chapter. This problem is potentiated in the low power networks, especially in island without any connection to continental networks. Due to its large implementation, generally, wind generation is considered the main source of variability on renewable generation but there are other sources that can introduce much faster variations, as solar generation [1]. On the other hand, hydro generation, although being easier to control its operation can depend on economic strategies, disconnected from available resource. If there is a small storage capacity, the production can be temporally disconnected form the rainfall. In an island context, a large variation on renewable production can introduce stability problems in the network, which can originate generation and/or load shed and, at limit, black-outs. By this sense and for security, the scheduling is generally done by a conservative way, with low risk but far away from an optimal operation.

   In this work is proposed an approach to a renewable power forecast method based on kernel density estimator (KDE).The main objective is to develop a short-term renewable power prediction to allow in the future the creation of a scheduling

approach for insular electricity networks. The developed methodology was tested in the S. Miguel Island in at Azores, Portugal.

2   Relationship to Collective Awareness Systems

Following the digital agenda for Europe "A Europe 2020" Initiative, the Collective Awareness Platforms (CAPS) are expected, among others, to support environmentally aware. By this sense, this work is based on the necessity of defining tools to allow a growth implementation of "clean" energy sources to produce electric energy. In spite of being renewable and "clean" techniques, due to their uncertainty creates great technical challenges to the power systems. In this work we present a set of tools in order to minimize the power production with pollutants fossil fuels and on other hand maximize the production with renewable and non-pollutants sources maintaining the technical requirements.

3   Problem Discussion

To decrease the fuel consumption by thermal generation, and consequently to decrease the costs of power production in S. Miguel Island, a large amount of renewable energy sources (RES), mainly wind generation, have been integrated. This is a reality in peak load periods but during off-peak periods the system is already saturated with RES. So, it's necessary, most of the times, to limit the wind power output in order to maintain the thermal generators operating above their technical minimum. Therefore, an efficient use of accurate short-term probabilistic forecast of RES can allow optimal committed/dispatched thermal generators, mainly during off-peak to minimize the necessity of renewable power derate. The production mix is composed by 1 thermal power plant with 8 units (divided in 2 groups with different rated power) 2 geothermal power plants with 5 units, 7 small hydro and 1 wind power plant. In table 1 is shown the rated power to each power source.

Table 1. Rated power of each power source

Source	(# units) total power
Fuel	(4) 28 MW
	(4) 64 MW
Wind	(9) 9 MW
Small hydro	(7) 5 MW
Geothermal	(5) 29,6 MW
other	*(2) < 1 MW*

During 2012 several measurements were done and it was concluded that load varied between a minimum of 17,3 MW and a maximum of 70,2 MW, while the maximum renewable production reached 62,2 MW. By operational security reasons it is mandatory that, at least, two thermal units must be on-line to avoid the loss of

thermal production. It means that several times, especially during off-peak periods, it was necessary to derate renewable production and, in some cases, work with thermal units below theirs minimum technical limits (12,26 MW). In these cases the thermal units are forced to work with poor efficiency and high fuel consumption whereas renewable sources are wasted. In 2012 these episodes occurred during 745 hours. In figure 1 is shown the thermal production as well as the sum of minimum technical limits of 2 units for 30 min intervals since 0h00, 1 Dec. up to 24h00, 31 Dec. During all over off-peak periods the thermal units had to work below their minimum limits.
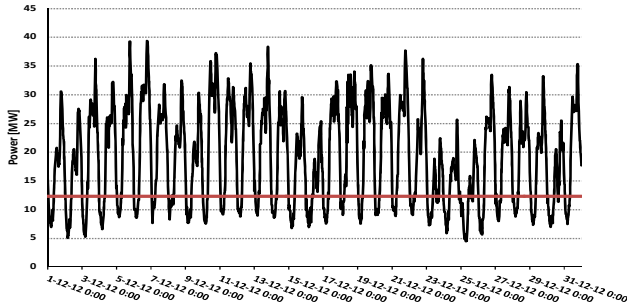
Fig. 1. Thermal production and technical limits

With an efficient forecast tool is possible to have an idea of expected renewable production in a forecast horizon and manage the mix of production, reducing with a certain degree of confidence the thermal production.

4   Renewable Power Probabilistic Forecasts

To ensure an optimized and secured scheduling operation of power grids, is mandatory to have an idea of load forecast, as well as the total renewable power forecast. With this information, is possible to know, under a certain degree of confidence, the thermal capacity needed to each time interval. Therefore, the wind and hydro and geothermal power forecast must be known as well as the load. Traditional forecasts techniques only forecast one value to each variable (*point forecast*) which doesn't afford any probability or confidence interval. Recent techniques present forecasts with uncertainties as densities estimations [2], probabilistic forecasts [3-7], quantiles regressions [8], among others. In this work it's proposed the use of KDE for the computing of predictive probability density function for each time-step of the prediction horizon.

4.1   Forecasting Formulation Modeling

In forecasting problems, the estimative of the future conditional density function has a very important role, since it describes the relation between explanatory and target variables. The conditional density estimation can be seen as a generalization of

regression, since conditional density estimation aims at obtaining the full probability density function $f_{y|x}(y|x)$, while the regression aims at estimating the conditional mean $E(y|x)$ [9]. In probabilistic forecasting problems the obtained *pdf* can be used to represent the uncertainty. In this work, we used a Nadaraya-Watson estimator represented by (1) which allow to estimating a random variable Y, when the explanatory random variable X is equal to x [2][10],

$$f_{Y|X}\left(y|x\right)=\frac{f_{Y,X}\left(y,x\right)}{f_X\left(x\right)}\,. \tag{1}$$

Where $f_{Y,X}(y,x)$ is the joint density function of (Y,X) and $f_X(x)$ is the marginal density function of X.

In the case of power and load forecast it consists in the estimation the future conditional *pdf* of a random variable 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 pairs consists on a set of explanatory variables X_n and corresponding value of variable to be predicted p_n. In this process is assumed the explanatory variables $x_{t+k|t}$ are known for each time-step ahead we want to forecast, resulting (2) where p_{t+k} is the power forecasted for look ahead time $t+k$,

$$\hat{f}_P\left(p_{t+k}\left|x_{t+k|t}\right.\right)=\frac{\hat{f}_{P,X}\left(p_{t+k},x_{t+k|t}\right)}{\hat{f}_X\left(x_{t+k|t}\right)}\,. \tag{2}$$

Since the joint and marginal densities aren't known, a nonparametric kernel estimation of the regression function can be used [10].

As renewable power production can depend on several variables (for instance wind speed and direction for wind power forecast or solar radiation and temperature for solar power forecast) a multivariate KDE was used [2][10]. For a given independent and identical distributed multivariate data $(X_{1d},\ldots,X_{nd})$ from d different variables from an unknown multivariate density function f, the multivariate KDE is given by (3)

$$\hat{f}\left(x_1,\ldots,x_d\right)=\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}\frac{1}{h_j}K_j\left(\frac{x_j-X_{ij}}{h_j}\right), \tag{3}$$

where n is the number of samples, d the number of variables and K_j is the kernel function to each variable j. In (3) h_j is the bandwidth (smoothing parameter) of each kernel around each sample X_{ij}.

Using Nadaraya-Watson estimator, the conditional density is given by (4)

$$\hat{f}_{P_{t+k}|X_{t+k|t}}\left(p,x\right)=\sum_{i=1}^{n}K\left(\frac{p-P_i}{h_p}\right)\cdot\frac{\prod_{j=1}^{d}K\left(\frac{x_j-X_{ij}}{h_j}\right)}{\sum_{i=1}^{n}\left[\prod_{j=1}^{d}K\left(\frac{x_j-X_{ij}}{h_j}\right)\right]}\,. \tag{4}$$

In the case of deterministic forecast, to estimate the conditional mean $P_{t+k}=E(p_{t+k}|x_{t+k|t})$, Nadaraya-Watson is also used to estimate P_{t+k} as a locally weighted average (5)

$$\hat{P}_{t+k} = E\left(p_{t+k}\left|x_{t+k|t}\right.\right) = \sum_{i=1}^{n} \frac{\prod_{j=1}^{d} K\left(\dfrac{x_j - X_{ij}}{h_j}\right)}{\sum_{i=1}^{n}\left[\prod_{j=1}^{d} K\left(\dfrac{x_j - X_{ij}}{h_j}\right)\right]}.P_i \qquad (5)$$

The first step is to choice the kernel, and in the literature there are several possible kernels, namely, Normal, Biweight, Epanechnikov, Logistic, among others, but the most important step is the chosen of bandwidth h. Small bandwidth values lead to an over fitted prediction function, while high values generalize too much [10]. In this work we used Normal Kernel and the bandwidths were calculated by the leave-one-out cross validation (LOO-CV) [10] with the hill climbing optimization algorithm.
After the definition of the expected value by (5) and standard deviation it is necessary to define the probability distribution. Following [2] *Beta pdf* (6) was found to be a good approximation for modelling variables with minimum and maximum limits, beyond to be very flexible.

$$f\left(x,\alpha,\beta\right) = \frac{1}{B[\alpha,\beta]}\left(x^{\alpha-1}\left(1-x\right)^{\beta-1}\right) \qquad (6)$$

Where $B[\alpha,\beta]$ is a normalization constant to ensure that the probability integrates to one, x is defined between [0,1] and shape parameters α and β are higher than 0. Allowing the evaluation of the distribution in a continuous way, this distribution fitting also makes possible to approximately simulate the obtained distribution using only two parameters. To define the distribution parameters α and β, calculated for every time-step of the forecast the moment's method was used.

4.2   Wind Power Forecast

Although depending on several variables, typically wind speed and direction are used as explanatory variables to wind power forecast. In this work the explanatory variables, wind speed and direction, are provided as spot forecast from numerical weather predictions (NWP) tools. As the NWP are one of the largest sources of uncertainty in wind power forecast, each kernel (Normal) is centered in spot forecast with a bandwidth h_x obtaining the prediction from (5).

4.3   Hydro and Geothermal Power Forecast

The geothermal power plant although being renewable, isn't characterized by its uncertainty, once the resource is relatively easy to control. Due to that, the geothermal power production remains relatively constant all over the year. It is assumed that the uncertainties are small oscillations around a specific value. Thus, the expected value

is modeled from a 24 hours moving average of past productions. To model the uncertainty, centered in the set point it's applied a kernel to obtain the uncertainty.

In the case of hydro power, the point forecast of the power is obtained by a tool named H4C proposed by [11]. The inputs of H4C are, data of hydro power production and precipitation forecast provided by NWP. Finally, to model the uncertainty the output variable of H4C will be used as input of KDE method.

4.4   Load Forecast

As in the generality of power grids, the power consumption in S. Miguel is affected by environmental factors, as temperature and social factors, as hour of the day, weekends, and holidays, among others. As previous forecasts, also in load forecast multivariable KDE was used. In this case, the explanatory variables are ambient temperature, hour of the day, day of the week and week of the year.

4.5   Aggregation of Forecasts

After the definition of all forecasts and distributions, is necessary to define the equivalent probabilistic model of the aggregations. Assuming that all renewable sources are statistically independent, their sum can be obtained by the convolution of all *pdf* [12] which resulted from RES forecasts. Finally, the convolution (subtraction) between aggregated RES forecast and load forecast (L-RES), results on the power which should be produced by thermal units. These values represent the remaining load to be feed by thermal units under a level of uncertainty defined by a *Beta* distribution.

5   Case Study

To perform the case study there is an off-line hourly data set of all variables necessary to the forecasts as well as the on-line thermal power units and total demand during 2012. In figure 2 is shown the forecast (24 hours ahead, refreshed at 0h00 of each day) of thermal power production resulting from (L-RES) and its confidence (uncertainty) interval from 5% up to 95%, since 0h00 November 15th up to 23h00 November 21st of 2012. In the same figure is shown the maximum and minimum limits of on-line thermal units committed by transmission system operator (TSO).
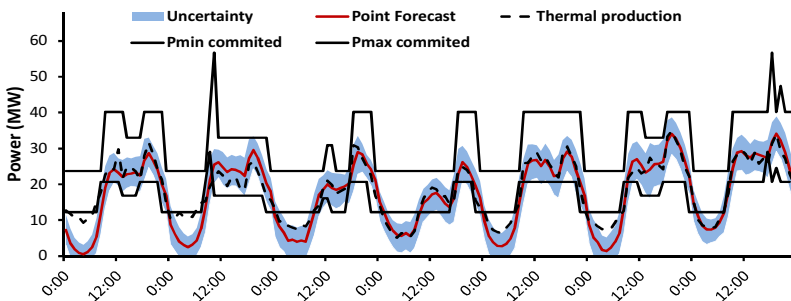

Fig. 2. Measured and forecasted thermal power production

It is shown that there is a good fitting between measured and forecasted thermal power. The large differences happen when the forecasts are below the lower technical limits of on-line thermal units. As a result of a rule of thumb, there are always two generators on-line whose sum of technical minima are 12,26 MW. So it means there was the necessity of wind power derating. In some of that cases, even with the derating there were the necessity of thermal units working below the minimum. In these cases the thermal units are working with low efficiency and high consumption of fuel and wasting "clear" energy.

In figure 3 is shown the result of a possible unit commitment based on thermal power forecast under uncertainty. Comparing with figure 2 is clear the decrease of thermal production mainly in off-peak periods. In the future works is intended to add more restrictions to the problem as contingencies and start/stop times.
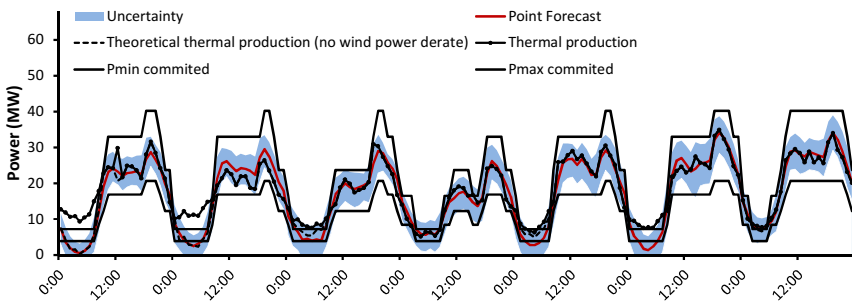

Fig. 3. Unit commitment and thermal forecast with uncertainty

Due to the real necessity of wind power derrating another source of uncertainty was introduced in the available off-line data set. There are pairs of explanatory variables which do not correspond to the real wind power available due to the derrating. To overcome this problem a real power curve of wind farm was modeled and during the periods where the power is limited or at least suspicious of limitation the obtained power curve was used as an estimation of wind power production. By this way another characteristic is shown in figure 3, a theoretical thermal production (without derrating).

6   Conclusions

There are vast consensuses on the need of forecasting techniques when there is a large penetration of renewable power plants in the electric grid. Nevertheless this issue becomes more important in the context of an island manly when there is low rated capacity. In this work was shown that accurate forecast techniques have an important role given important information to the scheduling and operation of power grid. Based on this work a short-term scheduling methodology based on risk analysis can be developed in order to create a tool which can optimize de unit commitment and due to that minimize the fuel consumption and maximize the weight of renewable production.

References

1. Sengupta, M., Keller, J.: PV ramping in a distributed generation environment: A study using solar measurements. In: 2012 38th IEEE Photovoltaic Specialists Conference (PVSC), June 3-8, pp. 586–589 (2012)
2. Bessa, R.J., Miranda, V., Botterud, A., Wang, J.: Time Adaptive Conditional Kernel Density Estimation for Wind Power Forecasting. IEEE Transactions on Sustainable Energy (October 2012)
3. Juban, J., Fugon, L., Kariniotakis, G.: Probabilistic short-term wind power forecasting based on kernel density estimators. In: European Wind Energy Conf. (May 2007)
4. Bludszuweit, H., Dominguez-Navarro, J.A., Llombart, A.: Statistical Analysis of Wind Power Forecast Error. IEEE Transactions on Power Systems, 983–991 (August 2008)
5. Jeon, J., Taylor, J.W.: Using Conditional Kernel Density Estimation forWind Power Density Forecasting. J. American Statistical Association 107, 66–79 (2012)
6. Al-Awami, A.T., El-Sharkawi, M.A.: Statistical characterization of wind power output for a given wind power forecast. In: 41st North American Power Symposium, vol. (1), pp. 1–4 (2009)
7. Matos, M.A., Bessa, R.J.: Setting the Operating Reserve Using Probabilistic Wind Power Forecasts. IEEE Transactions on Power Systems 26 (May 2011)
8. Bremnes, J.B.: Probabilistic wind power forecasts using local quantile regression. Wind Energy 1(1), 47–54 (2004)
9. Fu, G., Shih, F.Y., Wang, H.: A kernel-based parametric method for conditional density estimation. Pattern Recognition 44(2), 284–294 (2011)
10. Demir, S.: On the Adaptive Nadaraya-Watson. Journal of Mathematic and Statistics 39(3), 429–437 (2010)
11. Monteiro, C., Ramirez-Rosado, I.J., Fernandez-Jimenez, L.A.: Short-term forecasting model for electric power production of small-hydro power plants. Renewable Energy (6), 387–394 (2013)
12. Williamson, R.C.: Probabilistic Arithmetic, University of Queensland, Ph.D. Thesis (1989)