

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Probabilistic Forecasting of Day-ahead Electricity Prices for the Iberian Electricity Market

Rui Moreira

FEP - Faculty of Economics, University of Porto
Porto, Portugal
199803899@fep.up.pt

Ricardo Bessa

INESC TEC - INESC Technology and Science
Porto, Portugal
ricardo.j.bessa@inesctec.pt

João Gama

INESC TEC and FEP - Faculty of Economics, University of Porto
Porto, Portugal
jgama@fep.up.pt

Abstract— With the liberalization of the electricity markets, price forecasting has become crucial for the decision-making process of market agents. The unique features of electricity price, such as non-stationary, non-linearity and high volatility make this a very difficult task. For this reason, rather than a simple point forecast, market participants are more interested in a probabilistic forecast that is essential to estimate the uncertainty involved in the price. By focusing on this issue, the aim of this paper is to analyze the impact of external factors in the electricity price and present a methodology for probabilistic forecasting of day-ahead electricity prices from the Iberian electricity market. The models are built using regression techniques and aim to obtain, for each hour, the quantiles of 5% to 95% by steps of 5%.

Index Terms—Day-ahead forecast; Iberian electricity market; machine learning; probabilistic forecasting; quantile regression.

I. INTRODUCTION

The exchange electricity markets are characterized mainly by the extreme price volatility. Electricity price time series exhibit specific characteristics such as daily, weekly and annual seasonality, non-stationarity, positive skewness, high kurtosis, intra- and inter-day correlation, and short-lived mean reverting price spikes [1]. Many of these characteristics result from the unique features of electricity as commodity, with such is economically non-storable and therefore production have to match constantly and instantaneously the demand.

The market risk is therefore quite high, so for the various market participants involved, it is essential to have reliable and accurate electricity price forecasts that enable risk assessment of its bidding strategies and investment decisions. By far most of the researchers in this field focus on point forecasting, determining the expected value of the price at each time point for a given horizon [2]. But now, this information is not enough to make a proper risk analysis. Rather than forecasting the value, market participants are more interested in estimating the uncertainty involved in the price, knowing the degree of dispersion of the predicted value and their probabilities. Thus, it has been proved that it is increasingly important to conduct probabilistic forecasts. However relatively little attention has been paid to probabilistic forecasting of electricity prices. The

different approaches used in the few studies are prediction intervals or density forecasts, with the first being the most common approach [2]. Serinaldi [3] introduced the Generalized Additive Models for Location, Scale and Shape (GAMLSS) to model the distribution of electricity price, whose parameters varies dynamically according to explanatory variables via a number of linear and/or nonlinear relationships. These models forecast the entire distribution of price and allow the incorporation of the seasonality, trends and abrupt changes both in terms of location parameter (expected value) as the scale and shape (linked to price volatility, skewness and kurtosis). Wan *et al.* [4] proposes a hybrid method to construct prediction intervals. First use an extreme learning machine for estimate point forecasts and model uncertainties via a complex bootstrapping approach, and then the noise variance of forecasting results is estimate using a neural network trained with the maximum likelihood method. The prediction intervals are constructed based on the variance of model uncertainty and the noise, and are evaluated with respect to calibration and sharpness. Recently, Jónsson *et al.* [5] presented a semi-parametric methodology to estimate the density of the day-ahead electricity prices. It includes a time-adaptive quantile regression model for quantiles ranging from 5% to 95%, and a description of the distribution tails by exponential distributions.

It is a known fact that electricity prices are influenced by the level of demand, which in turn depends on the intensity of business activities and weather. However, that is no longer the only variable to be considered. Generation from renewable energies has been growing significantly, especially the wind power generation, and has taken particular importance, in periods of high generation, the price of electricity tends to drastically fall. By focusing on this issue, the aim of this paper is to analyze the impact of external factors, such as wind power generation and weather variables, in the electricity prices and present a methodology for probabilistic forecasting of day-ahead electricity prices from the Iberian Electricity Market (MIBEL).

The rest of paper is organized as follow: Section II presents a brief description of the probabilistic forecast algorithms; the test case results are presented in Section III and

This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme within project «POCI-01-0145-FEDER-006961», and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) as part of project UID/EEA/50014/2013.

Section IV presents the conclusions and suggestions for future research.

II. QUANTILE REGRESSION FRAMEWORK

A. Definition of Quantile Forecast

A prediction interval $\hat{I}_{t+k|t}^{(\alpha)}$, estimated at time step t for look-ahead time $t+k$, is a range of values within the true outcome y_{t+k} is expected to lie with a pre-assigned probability $(1-\alpha)$:

$$P\left(y_{t+k} \in \hat{I}_{t+k|t}^{(\alpha)}\right) = P\left(y_{t+k} \in \left[\hat{L}_{t+k|t}^{(\alpha)}, \hat{U}_{t+k|t}^{(\alpha)}\right]\right) = 1-\alpha \quad (1)$$

where $\hat{L}_{t+k|t}^{(\alpha)}$ and $\hat{U}_{t+k|t}^{(\alpha)}$ denotes, respectively, the lower and upper bounds of the prediction interval. In central prediction intervals these bounds correspond to the quantiles with proportion $(\alpha/2)$ and $(1-\alpha/2)$ of the predictive distribution [6].

B. Linear Quantile Regression (LQR)

In the quantile regression, introduced by Koenker et Basset [7], the quantile $Q^{(\tau)}$ is modelled as

$$Q^{(\tau)} = \beta_0(\tau) + \beta_1(\tau)x_1 + \dots + \beta_p(\tau)x_p \quad (2)$$

where x_i are explanatory variables and $\beta_i(\tau)$ are unknown coefficients depending of τ , to be determined from the data.

Given the loss function,

$$\rho_\tau(\varepsilon) = \begin{cases} \tau\varepsilon & \text{se } \varepsilon \geq 0 \\ (\tau-1)\varepsilon & \text{se } \varepsilon < 0 \end{cases} \quad (3)$$

and considering $\beta(\tau)$ as the vector containing the unknown coefficients, the best estimator of $\beta(\tau)$ is the solution of the loss optimization problem

$$\hat{\beta}(\tau) = \arg \min_{\beta} \sum_{i=1}^N \rho_\tau \left[y_i - (\beta_0 + \beta_1 x_{i,1} + \dots + \beta_p x_{i,p}) \right]. \quad (4)$$

Since in practical applications strictly linear processes are rare, should be allowed some flexibility to the model. For that, one possible solution is to consider $Q^{(\tau)}$ as an additive model:

$$Q^{(\tau)} = \alpha(\tau) + \sum_{j=1}^p f_j(x_j; \tau) \quad (5)$$

where $\alpha(\tau)$ is a constant and each of the functions $f_j(x_j; \tau)$ can be approximated by linear combinations of known basis functions of the corresponding explanatory variable [8].

The LQR is available in the R package *quantreg* [9]. For the basis functions were considered cubic splines, with four degrees of freedom using the function *bs* from the R package *splines*.

C. Quantile Regression Boosting

Another approach for the estimation of quantile regression additive models is the implementation of boosting algorithms, such as the component-wise gradient boosting (GB) [10], an optimization method that combines base learners via gradient descent techniques for fitting linear (LM-GB) and linear additive models (LAM-GB). This algorithm can be used for a large number of different loss functions, and for quantile regression the choice is the function defined in (3). A great advantage of the algorithm is that performs automatic variable selection during the fitting process. The GB is implemented in the R package *mboost* [11]. The step length is fixed at 0.1 and the optimal number of iterations was determined by evaluating the empirical risk through 5-fold cross-validation. For the LAM-GB we used cubic-penalized spline base learners with four degrees of freedom, a second-order difference penalty and 20 inner knots.

D. Quantile Regression Forests (QRF)

Quantile regression forests, introduced by Meinshausen [12], is a generalization of machine learning algorithm random forests. Unlike other techniques, quantile regression forests does not directly employ minimization of the loss function defined in (3), and estimate the full conditional distribution function in one step. QRF are implemented in the R package *quantregForest* [13]. The size of random subset of the explanatory variables was determined by 5-fold cross-validation using the function *train* from the R package *caret* [14].

E. Quantile Regression Neural Networks (QRNN)

Neural networks for quantile regression were introduced by Taylor [15], and consist in adopting the function defined in (3), as a special case of cost function of standard neural network algorithm. However, the optimization process theoretically requires a differentiable function at all points, which is not the case for the function $\rho_\tau(\varepsilon)$ that is not differentiable at the origin. To address this problem Cannon [16] replace the loss function $\rho_\tau(\varepsilon)$ by a differentiable approximation defined as

$$\rho_\tau^*(\varepsilon) = \begin{cases} \tau h(\varepsilon) & \text{se } \varepsilon \geq 0 \\ (\tau-1)h(\varepsilon) & \text{se } \varepsilon < 0 \end{cases} \quad (6)$$

where $h(u)$ is the Huber norm. Software for QRNN is available in the R package *qrnn* [17]. The number of hidden nodes and weight penalty for weight decay regularization are determined by 3-fold cross-validation using the function *train* from the R package *caret* [14].

F. Quantile Regression on Reproducing Kernel Hilbert Space (KQR)

The quantile regression on Reproducing Kernel Hilbert Space is intrinsically connected to the framework of support vector machine and is defined as [18]

$$\min_{f \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - f(x_i)) + \frac{\lambda}{2} \|g\|_{\mathcal{H}}^2 \quad (7)$$

where H is a Reproducing Kernel Hilbert Space (RKHS) generated by a kernel, $f = g + b$ with $b \in \mathbf{R}$ and $g \in H$, $\|\cdot\|_H$ is RKHS norm and λ a regularization parameter to prevent overfitting. KQR is implemented in the R package *kernlab* [19]. The kernel selected was the Gaussian Radial Basis with automatic estimation of the hyper-parameter sigma.

III. CASE STUDY

A. Description

The Iberian Electricity Market covers the Iberian Peninsula (Portugal and Spain control areas), and started its operation with a common platform for both countries in, July 1, 2007.

The data used include historical prices of electricity in Portugal, and forecasts of electricity demand, wind power generation, wind speed, precipitation, temperature and solar irradiance. They refer to each 24-hour period between 1 January 2013 and 15 September 2014. The market prices were obtained from the Iberian Energy Market Operator website [20]. The forecasts of demand and wind power generation are for the total in the Iberian Peninsula and were obtained from *Red Eléctrica de España* website [21]. Weather forecasts were generated by the Weather Research & Forecasting Model (WRF) [22] and reports to eight different regions of Iberian Peninsula.

The data from 2013 were used to estimation of parameter of the models. The remaining data were used as an independent test set. The explanatory variables are lagged values of price (with lags of 1, 24, 48,72 and 168h), time of the day, month, day of the year, forecasts of electricity demand, wind power production, wind speed, precipitation and average of forecasted temperatures.

Solar irradiance data were not used because they presented a weak correlation with the price, and installed capacity (in relative terms) of photovoltaic power generation in Iberian Peninsula, especially in Portugal, is minimal.

The quantile forecast, at time step t , for each look-ahead time $t+k$, $k=1, \dots, 24$, was forecasted with a iterative strategy. For the weather variables at look-ahead time $t+k$ were considered available forecasts at time step t and for the variable "price lag" at look-ahead time $t+k$, $k \geq 2$ was adopted the value of forecasted median (quantile 50%).

Probabilistic prediction of the electricity price for each of the 24 hours of the following day is generated in the form of a set of quantiles between 5% and 95 % by increments of 5%, using the techniques described in Section II.

Point forecast results, associated to the quantile 50%, are evaluated with the Root Mean Square Error (RMSE) normalized by the maximum price. For evaluation purpose of probabilistic forecasts were considered three metrics: calibration, sharpness and continuous ranked probability score (CRPS). A completed description of these metrics is given in [5],[23].

B. External Factors Impacting the Energy Prices

Wind power production differs substantially from conventional energy sources. Relying exclusively on wind, it is very

volatile and the absence of fuel costs allows a very low marginal cost, which has a strong influence on the market-close price.

In order to perform the analysis of the external factors in the price distribution, the dataset was divided into bins and the properties are estimated within each bin. In Fig.1 is presented the histogram of the electricity prices for different levels of forecasted wind power production.

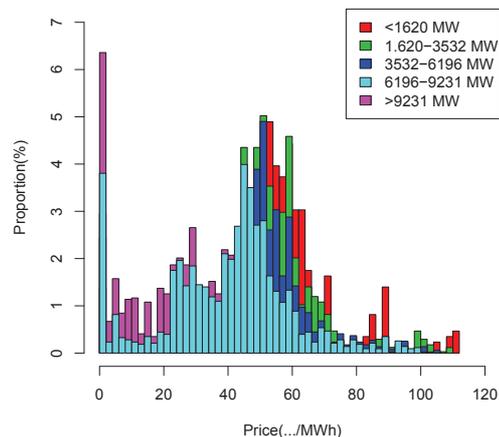


Figure 1. Distribution of prices for different levels of forecasted wind power production.

The Fig.1 shows that an increase in the forecasted wind power production leads to a decrease of the mean price and the weights of the tails of the price distribution. Therefore, the probability of occurrence of extremely high prices is much lower when is predicted to be high the wind power production. The differences in distribution properties are summarized in Table I.

TABLE I. PROPERTIES OF PRICE DISTRIBUTION FOR DIFFERENT LEVELS OF FORECASTED WIND POWER PRODUCTION (MW)

	<1620	1620-3532	3532-6196	6196-9231	>9231
Mean	55.78	53.23	46.36	40.04	30.28
Std. Dev.	18.68	16.29	18.69	20.03	19.32
Skewness	-0.44	-0.14	-0.47	-0.06	0.08
Kurtosis	3.22	3.37	1.25	0.37	-0.53

Linked to wind power production, wind speed has a strong impact in the electricity price. From Table II is obvious that the average spot price tend to decrease as the forecasted wind speed increase, which becomes more significantly for values above 3.96 m/s. For that values it is also observable an increase in the skewness and a decrease in the kurtosis.

TABLE II. PROPERTIES OF PRICE DISTRIBUTION FOR DIFFERENT LEVELS OF FORECASTED WIND SPEED (m/s)

	<1.42	1.42-3.96	3.96-6.44	6.44-9	>9
Mean	52.49	50.87	45.36	36.31	26.59
Std. Dev.	19.38	18.23	17.67	21.23	20.57
Skewness	0.03	-0.02	-0.35	-0.15	0.30
Kurtosis	1.18	1.79	1.03	-0.31	-0.47

The temperature is another weather factor impacting the price of electricity, especially when it is high. For predicted values above 20°C there is a significant increase in the average price and a significant reduction of the standard deviation, which indicates a high concentration of values around the mean. The last two lines in Table III also show that the distributions are rather skewed and with high kurtosis as the predicted temperature increases.

TABLE III. PROPERTIES OF PRICE DISTRIBUTION FOR DIFFERENT LEVELS OF FORECASTED TEMPERATURE (°C)

	<5	5-20	>20
Mean	42.44	42.08	52.73
Std. Dev.	24.93	20.69	7.73
Skewness	0.12	-0.16	-0.89
Kurtosis	-0.36	1.40	2.11

For short-term forecasting, the precipitation does not have the same importance as the other weather factors since the impact on the water level of hydropower plants reservoirs is “slow”.

C. Forecasting Evaluation Results

Fig. 2 depicts the RMSE for each look-ahead time step.

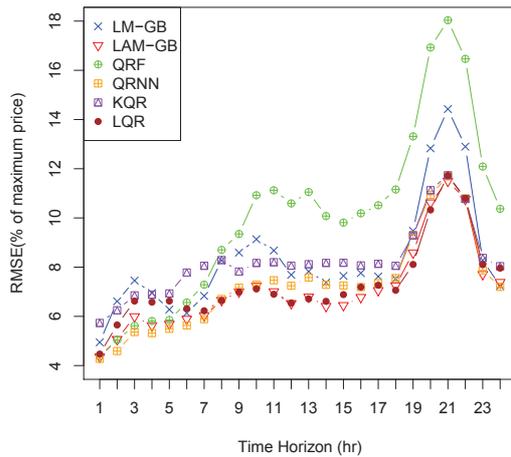


Figure 2. Root Mean Square normalized by maximum price.

There is no model that surpasses the others in all look-ahead time steps, and the LM-GB and QRF models clearly stand out negatively. In the first seven hours the best performance, in terms of RMSE, is obtained by QRNN. But from there, it is supplanted in most time horizons by the LQR and the LAM-GB, which generally is the one with better performance. The growing trend of error over the time horizon can be partially explained by the forecast iterative strategy. The sharp increase between 18th and 21st hours could be associated with the existence of many outliers in those hours.

Fig.3 depicts the difference from the “perfect calibration” for the whole time horizon.

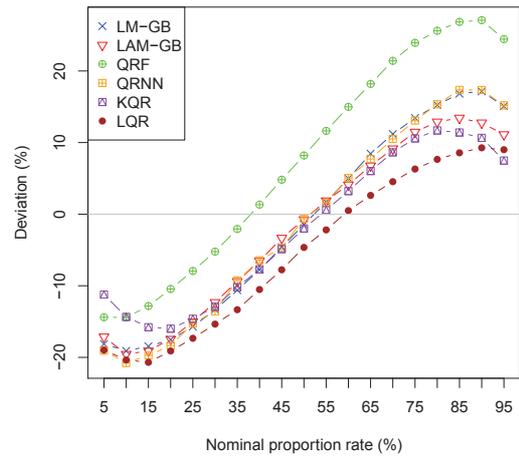


Figure 3. Calibration diagram.

The QRF presents the lowest deviations for the quantiles between 15-40%, however for higher quantiles presents the worst results of all models considered. For the extreme quantiles – 5, 10 and 95% – the KQR present the lowest deviations and for the quantiles between 65-90% the lowest deviations are obtained by LQR. On average, the models underestimate the quantiles for values below the 50% quantile and overestimate the quantiles for greater values.

In terms of sharpness, depicted in Fig. 4, is desired to have intervals with smaller size for all coverages rates.

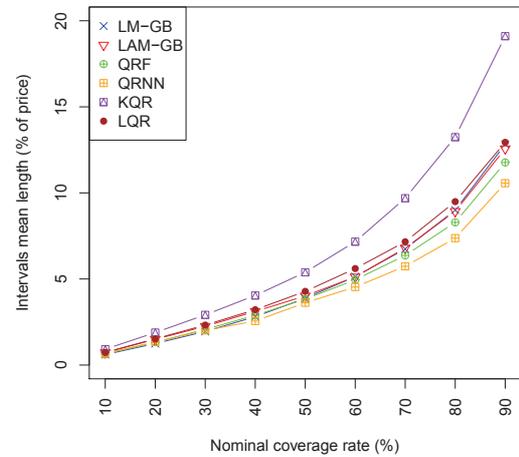


Figure 4. Sharpness diagram.

The KQR present the worst performance, while the remaining have a similar performance, although QRNN presented a slightly lower sharpness.

Fig. 5 depicts the CRPS for each look-ahead time step.

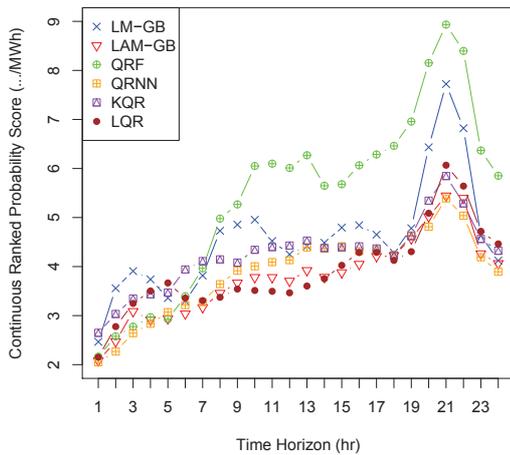


Figure 5. CRPS diagram.

In the first four and the last five hours the QRNN present the best performance. In the remaining hours prevail the LAM-GB or the LQR. Note that as in RMSE, between the 18th and 21st hour there is a significant increase in the values of CRPS.

IV. CONCLUSIONS

The analysis of the relation between the external factors and the price leads to the conclusion that they significantly influence not only the average price, but also the shape and form of the price distribution, highlighting the wind power levels. Scenarios that wind production is high leading to a decrease in prices, resulting even, on many occasions, in zero or close to zero.

In terms of calibration, important in this context for the bidding problem, the KQR model, with values between 11.7% and 17%, has the best performance. However, still slightly higher than what would be ideal.

The quality of the estimated models shows that the use of quantile regression methods is a plausible strategy for obtaining probabilistic forecasts. The LAM-GB, slightly stand out above the rest. In addition to presenting on average the best CRPS performance, performs automatic variable selection, compared with the QRNN and KQR require less computational effort and produce models much easier of understanding.

This paper does not address the potential use of this information by different market participants. Since the value of information depends entirely on the decision-making process and decision-maker attitude towards risk.

For future work, since the occurrence of extreme prices appears to have a strong impact on the quality of the prediction, the employment of a spike detection/forecast method may lead to improved results. Another possibility involves the application of multiple models for combination of different quantiles forecasts. Finally, the inclusion of forecasted solar irradiance should be considered in a near future due to the massive deployment of self-consumption strategies in the Iberian Peninsula.

REFERENCES

- [1] T. Jónsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Forecasting electricity spot prices accounting for wind power predictions," *IEEE Transactions on Sustainable Energy*, vol. 4, pp. 210-218, 2013.
- [2] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, pp. 1030-1081, 2014.
- [3] F. Serinaldi, "Distributional modeling and short-term forecasting of electricity prices by Generalized Additive Models for Location, Scale and Shape," *Energy Economics*, vol. 33, pp. 1216-1226, 2011.
- [4] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A hybrid approach for probabilistic forecasting of electricity price," *IEEE Transactions on Smart Grids*, vol. 5, no. 1, pp. 463 - 470, 2014.
- [5] T. Jónsson, P. Pinson, H. Madsen, and H. Nielsen, "Predictive densities for day-ahead electricity prices using time-adaptive quantile regression," *Energies*, vol. 7, pp. 5523-5547, 2014.
- [6] P. Pinson, "Estimation of the uncertainty in wind power forecasting," Phd Thesis, École Nationale Supérieure des Mines de Paris, 2006.
- [7] R. Koenker and G. Bassett Jr, "Regression quantiles," *Econometrica: journal of the Econometric Society*, vol. 46, no. 1, pp. 33-50, 1978.
- [8] H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, pp. 95-108, 2006.
- [9] R. Koenker, *quantreg: Quantile Regression*. R package version 5.11. <http://CRAN.R-project.org/package=quantreg>
- [10] P. Buehlmann, "Boosting for high-dimensional linear models," *The Annals of Statistics*, pp. 559-583, 2006.
- [11] B. Hofner, A. Mayr, N. Robinzonov, and M. Schmid, "Model-based boosting in R: a hands-on tutorial using the R package mboost," *Computational Statistics*, vol. 29, pp. 3-35, 2014.
- [12] N. Meinshausen, "Quantile regression forests," *The Journal of Machine Learning Research*, vol. 7, pp. 983-999, 2006.
- [13] N. Meinshausen and M. N. Meinshausen, *quantregForest: Quantile Regression Forests*. R package version 1.0. <http://CRAN.R-project.org/package=quantregForest>
- [14] M. Kuhn, J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, *et al.*, *caret: Classification and Regression Training*. R package version 6.0-47. <http://CRAN.R-project.org/package=caret>
- [15] J. W. Taylor, "A quantile regression neural network approach to estimating the conditional density of multiperiod returns," *Journal of Forecasting*, vol. 19, no. 4, pp. 299-311, 2000.
- [16] A. J. Cannon, "Quantile regression neural networks: Implementation in R and application to precipitation downscaling," *Computers & Geosciences*, vol. 37, pp. 1277-1284, 2011.
- [17] A. J. Cannon, "Quantile regression neural networks: implementation in R and application to precipitation downscaling," *Computers & Geosciences*, vol. 37, pp. 1277-1284, 2011.
- [18] I. Takeuchi, Q. V. Le, T. D. Sears, and A. J. Smola, "Nonparametric quantile estimation," *The Journal of Machine Learning Research*, vol. 7, pp. 1231-1264, 2006.
- [19] A. Karatzoglou, A. Smola, K. Hornik, A. Zeileis (2004). *kernlab - An S4 Package for Kernel Methods in R*. *Journal of Statistical Software*, vol. 11, no. 9, pp. 1-20, 2004.
- [20] *Iberian Energy Market Operator* Available: <http://www.omie.es> (accessed on Jan. 2015).
- [21] *Red Eléctrica de España*. Available: <http://www.esios.ree.es/web-publica/> (accessed on Jan. 2015).
- [22] *The Weather Research & Forecasting Model*. Available: <http://www.wrf-model.org/index.php> (accessed on Jan. 2015).
- [23] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: required properties and evaluation," *Wind Energy*, vol. 10, pp. 497-516, 2007.