A non-parametric quantile autoregressive model for wind power Firm Energy Certificate

Henrique Hoeltgebaum, Marcelo Ruas, Cristiano Fernandes and Alexandre Street, *Member, IEEE*

***Abstract*—In this article we use the framework of a non-parametric quantile regression model to generate forecasts of wind capacity factors of several quantiles. Such scenarios are then used as input to raise the distribution of the quantiles associated with each wind plant. In our proposed model we introduce a ℓ_1 penalty term in the objective function, in order to properly adress an adaptive dependency structure for the α-quantile time series. Hence, in contrast to the well-known quantile regression model, our proposed framework is not limited to linear functions. Computation experiments show a slight improvement when comparing our non-parametric framework to the benchmark.**

***Index Terms*—Quantile regression, quantile autoregressive models, firm energy certificate, wind power, ℓ_1-penalty term, time varying quantiles.**

I. Introduction

Wind Firm Energy Certificate (FEC) estimation impose several challenges. First and foremost, it is a quantile function of an aleatory quantity, named here on wind capacity factor (WP). Due to its non-dispachable profile, accurate scenario generation model could reproduce a fairly dependence structure in order to the estimation of FEC. Secondly, as a quantile function, the more close to the extremes, more sensitive to sampling error.

In this work, we introduce a new non-parametric quantile autoregressive model with ℓ_1-penalty term, in order to properly simulate FEC densities for several α-quantiles.

A. Review of the Brazilian Electricty Sector

In response to the growing demand of energy, Brazil began a period of major structural reforms in the 1990s [?]. Such reforms were positively accepted by private investors, which leads to numerous concession auctions for new projects [?]. Such enviroment had a mood turn over during 2001-2002, since the former security criteria were fully based on market mechanism and lead to a serious supply crisis. The outcomes of such crisis were the reduction of total load by 20% and an economic loss of tens of billions US dollars [?].

Therefore, as a response to aforementioned crisis, brazilian government developed a new power sector model in 2004 [?]. The implementation of the reformed framework for Brazil's electricity sector generally aims to provide the long-term system's expansion incentive, reduce uncertainty in generation companies future revenues, and mitigate unfavourable effects of the short-term market [?].

In addition, such system was applied to the Brazilian case based on mandatory reability contracts as an incentive to investors. These contracts are considered to be financial instruments, in the same spirit as forward contracts. Moreover, in order to provide a confidence sensibility over generation, all contracts ought to be covered by so-called "firm energy certificates" (FEC). These FECs are defined in GWh/year and represent the plant's physical energy production capacity. In a nutshell, the FEC of a certain plant is the maximum amount of energy that can be sold through contracts and establishes the reliability assured by the generator backing the contract. It is thus a critical parameter for the power plant's economic feasibility.

B. Motivation and objetives

Accurate estimation of FEC values provide the investors safety regarding his cash-flows, since the main issue about introducing renewable energy into energy matrixes is its intermitent[1] profile. Regarding wind power, its production is fully conditioned to the wind speed, which is caused by diference in atmospheric pressure. Hence, prior knowledge over its future behaviour is essential when considering the generation of this renewable source.

Regarding wind plants, its FEC where fully detailed on a document produced by ANEEL in July of 2008 [XXX], which defines the firm energy certificate of a wind plant as the mean of the monthly production. As described in the mentioned document, the estimation of wind plants FEC is given by

$$FEC^{(\alpha)} = \sum_{m=1}^{12} \frac{E_m^{(\alpha)} \times h_m}{8760}, \tag{1}$$

where $FEC^{(\alpha)}$ denotes the FEC of the wind plant certified with α% of confidence, h_m is the number of hours in the month m and $E_m^{(\alpha)}$ denotes the mean of the monthly quantiles certified at α%. In practice, the generator provides 12 values of energy certified at α%. refering to the month m.

To give the reader a further understanding about Equation (1), we present in Figure 1 the calculated FEC's for a wind-plant located at the brazilian northeast, named Icaraizinho. The estimatives $E_m^{(\alpha)}$ for several $\alpha \in \{50, 55, ..., 95\}$ are ploted.

This work was partially supported by the R&D project ANEEL PD-7625-0001/2013.

Henrique Hoeltgebaum, Marcelo Ruas, Cristiano Fernandes and Alexandre Street are with the Electrical Engineering Department, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brazil (e-mail: hhhelfer@hotmail.com; mcruas@gmail.com; cris@ele.puc-rio.br; street@ele.puc-rio.br).

[1]When there is no human intervention over its production.

II. Parametric model

Here we donoted as parametric model the well-known quantile regression model [?]. In contrast to the linear regression model through ordinary least squares (OLS), which provides only an estimation of the dependent variable conditional mean, quantile regression model yields a much more detailed information concerning the complex relationship about the dependent variable and its covariates, as defined by Equation (2),

$$\mathcal{Q}_y(\alpha|x_1,x_2,...,x_n) = \beta_0(\alpha) + \beta_1(\alpha)x_1 + \beta_2(\alpha)x_2 + ... \\ +\beta_n(\alpha)x_n + F_\varepsilon^{-1}(\alpha), \tag{2}$$

where F_ε denotes the error density function.

A. Parameter estimation

The parameters of the quantile regression are estimated by a linear programming optimization problem that can be formulated as

$$\min_{\beta \,\in\, \mathbb{R}^p \times \mathbb{R}^{2n}_+} \{\tau \mathbf{1}_n^T u + (1-\tau)\mathbf{1}_n^T v | X\beta + u - v = y\}. \tag{3}$$

Here, X stands for the usual covariates matrix, the residual vector $y - X\beta$ splits itself into negative and positive parts (u and v).

III. Quantile Autoregression

Our main objective in this paper is estimating the p-step ahead α-quantile function $\mathcal{Q}^\alpha_{y_t|y_{t-p}}(t)$ for a given set of time series data y_t, as the one on figure 2. So, given a sequence $\{y_t\}$, we can pair an observation y_t with its p-lagged correspondent y_{t-p}. Figure 3 shows this relationship. We will assume that all information regarding the estimated quantile value are past observations, being in accordance with other pure autoregressive models.

We will investigate two ways of estimating the quantiles for the aforementioned relationship: we will use a linear and a nonparametric model.

IV. Linear-QAR

V. NP-QAR

Fitting a linear estimator for the Quantile Auto Regression isn't appropriate when nonlinearity is present in the data. This nonlinearity may produce a linear estimator that underestimates the quantile for a chunk of data while overestimating for the other chunk (we illustrate this in figure 4). To prevent this issue from occurring we propose a modification which we let the prediction $\mathcal{Q}^\alpha_{y_t|y_{t-1}}(t)$ adjust freely to the data and its nonlinearities. To prevent overfitting and smoothen our predictor, we include a penalty on its roughness by including the ℓ_1 norm of its second derivative. For more information on the ℓ_1 norm acting as a filter, one can refer to [?].

Let $\{\tilde{y}_t\}_{t=1}^n$ be the sequence of observations in time t. Now, let $\tilde{x}_t$ be the $p-$lagged time series of $\tilde{y}_t$, such that $\tilde{x}_t = L^p(\tilde{y}_t)$, where L is the lag operator. Matching each observation $\tilde{y}_t$ with its $p-$lagged correspondent $\tilde{x}_t$ will produce $n-p$ pairs $\{(\tilde{y}_t,\tilde{x}_t)\}_{t=p+1}^n$ (note that the first p observations of y_t must be discarded). Consider J to be the set of indexes such that

$$\tilde{x}_{J_1} \le \tilde{x}_{J_2} \le \cdots \le \tilde{x}_{J_{n-P}}.$$

Now, we define $\{x_i\}_{i=1}^{n-p} = \{\tilde{x}_{J_t}\}_{t=p+1}^n$ and $\{y_i\}_{i=1}^{n-p} = \{\tilde{y}_{J_t}\}_{t=p+1}^n$ and $I = \{2,\ldots,n-p-1\}$. As we need the second difference of q_i, I have to be shortened by two elements.

Our optimization model to estimate the nonparametric quantile is as follows:

$$\mathcal{Q}^\alpha_{y_t|y_{t-1}}(i) = \arg\min_{q_i} \sum_{i\in I} \left(|y_i - q_i|^+\alpha + |y_i - q_i|^-(1-\alpha)\right) \\ +\lambda\sum_{i\in I}|D^2 q_i|, \tag{4}$$

where $D^2 q_t$ is the second derivative of the q_t function, calculated as follows:

$$D^2 q_i = \left(\frac{q_{i+1}-q_i}{x_{i+1}-x_i}\right) - \left(\frac{q_i - q_{i-1}}{x_i - x_{i-1}}\right).$$

The first part on the objective function is the usual quantile regression condition for $\{q_i\}$. The second part is the ℓ_1-filter. The purpose of a filter is to control the amount of variation for our estimator q_i. When no penalty is employed we would always get $q_i = y_i$. On the other hand, when $\lambda \to \infty$, our estimator approaches the linear quantile regression. The penalty for variation on the values of

The output of our optimization problem is a sequence of ordered points $\{(x_i,q_i)\}_{i\in I}$. The next step is to interpolate these points in order to provide an estimation for any other value of x. To address this issue, we propose a B-splines interpolation, that will be discussed in another subsection.

When estimating quantiles for a few different values of α, however, sometimes we find them overlapping each other, which we call crossing quantiles. To prevent this, we include a non-crossing constraint:

$$q_i^\alpha \le q_i^{\alpha'}, \quad \forall i \in I, \alpha < \alpha'. \tag{5}$$

Acknowledgment

The authors would like to thank FICO (Xpress-MP developer) for the academic partnership program with the Electrical Engineering Department of Pontifical Catholic University of Rio de Janeiro, Brazil (PUC-Rio). The authors would also like thank the LAMPS researchers for the daily exchanges and their insightful considerations.

References

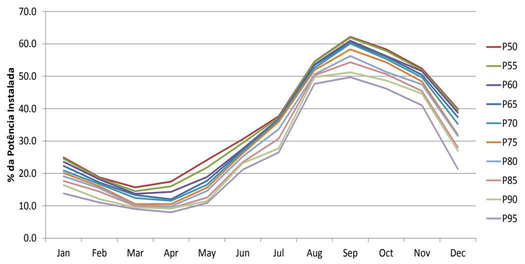

Fig. 1. Percentiles $E_m^{(\alpha)}$ for each month m and confidence criteria α.

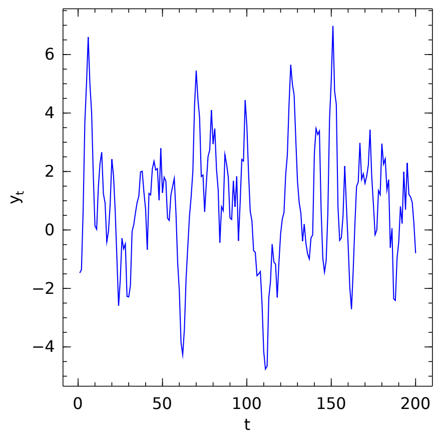

Fig. 2. Time series y_t

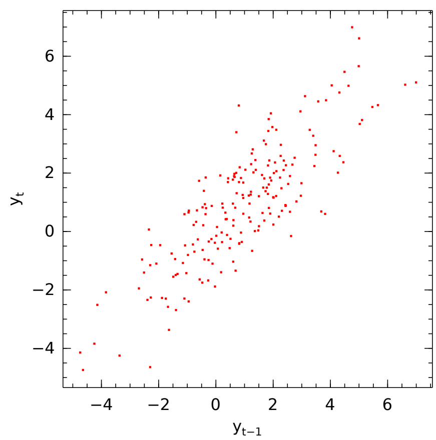

Fig. 3. Relationship between y_t and its first lag y_{t-1}

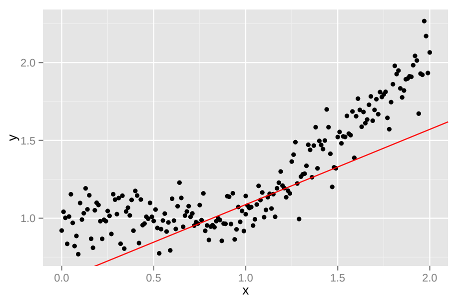

Fig. 4. Example of data where nonlinearity is present and a linear quantile estimator is employed