Nongaussian time series model via Quantile Regression

Marcelo Ruas, *Student Member, IEEE,* and Alexandre Street, *Member, IEEE*

***Abstract*—Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.**

***Keywords*—Quantile Regression, Model Identification, Nongaussian time series model**

TODO LIST

I. Introduction

Renewable energy power is a hot topic in the recent years. It is a much cleaner way of producing energy than by using other sources such as coal and gas, and with less hazard potential than nuclear power plants. The installed capacity of renewable plants has been increasing in a fast pace and projections point that wind power alone will account to 18% of global power by 2050 [1]. In spite of being a welcoming technology, it brings a new problem for energy planners: the complex behavior of wind is very difficult to understand and predict. Due to its unpredictability, the cost for the system as a whole may be high. Having better prediction models can help the planner to make better and less risky decisions, increasing the attractiveness of renewable energy to the energy system. In this work we will investigate how to model dynamics of renewable energy time series in both short and long terms.

Conventional statistics are often focused on estimating the conditional mean of a given random variable. This is not very useful when dealing with renewable energy, as the variability and the notion of risk is extremely important for planning. Our attention is, then, turned to probabilistic forecasting models. Among these, there are many possibilities. [2] reviews the commonly used methodologies for this task, separating them in parametric and nonparametric classes. Main characteristics of **parametric models** are (i) assuming a distribution shape and (ii) low computational costs. GAS models [3] and ARIMA-GARCH, for example, model the renewable series by assuming the distribution *a priori*. On the other hand, **nonparametric models** (i) don't require a distribution to be specified, (ii) needs mode data to produce a good approximation and (iii) have a higher computational cost. Popular methods are Quantile Regression, Kernel Density Estimation, Artificial Intelligence or a mix of them.

A common finding is that wind and solar time series don't have a Gaussian behavior. Furthermore, we can't point any distribution that fits on the data undoubtedly. For this reason, we choose to use a nonparametric approach, more specifically the Quantile Regression (QR), as defined in [4]. As we are working with time series, we rely on the results of [5], which defines the quantile autoregression, extending the application of QR on cases where the covariates are lagged values of y_t. QR is a powerful tool for measuring quantile others than the median. By estimating many quantiles on a thin grid of probabilities, one can have as many points as desired of the estimated conditional distribution function.

However, when estimating a distribution function, as each quantile is estimated independently, the monotonicity of the distribution function may be violated. This issue is a known as crossing-quantiles. To get around it we propose to either add a constraint on the optimization model or make a transformation as in [6].

Using QR as in [4] is not a novelty when predicting the conditional distribution of wind power time series, as is already done in [10]–[13]. In this article, we combine QR with regularization techniques. On the Mixed Integer Linear Programming (MILP) approach, we use the best subset selection, which [7] does for minimizing the quadratic error and the ℓ_1 penalization as in [8] or [9] (the AdaLasso variant, where each coefficient may have a different weight on the objective function to ensure oracle properties). On the advantages of QR we highlight the fact that it is a direct way of finding a given quantile directly and it is data driven, so it doesn't depend on distribution assumptions.

Review of quantile regression for wind and solar - procurar mais referências?

We propose a new combination of methods to predict the k-step ahead conditional distribution. By using MILP, we achieve a solution which is optimal for the given objective. In order to improve the quality of predictions and interpretability, we incorporate a joint regularization by specifying the existence of groups among the probabilities α. We could not find any other work in the literature that interpreted different quantiles as models depending on one another. The objective of this paper is to propose and test different techniques of predicting the conditional distribution based on QR.

II. Quantile Regression based time series model

Let the conditional quantile function $Q_{Y|X=x} : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ (in short, from now on, $Q_{Y|X}(\cdot,\cdot)$) be defined as

$$Q_{Y|X}(\alpha, x) = F^{-1}_{Y|X}(\alpha, x) = \inf\{y : F_{Y|X}(y, x) \geq \alpha\}, \quad (1)$$

where Y is a random variable and X a random vector. In this work, we use the sample quantile, where we calculate the optimization based on a finite number of observations, instead of integrating over all domain of random variable Y. The sample quantile is given by

Definir Y_t no lugar de Y?

$$\hat{Q}_{Y|X}(\alpha, \cdot) \quad \in \quad \arg\min_{q_\alpha(\cdot)} \sum_{t \in T} \rho_\alpha(y_t - q_\alpha(\cdot)), \quad (2)$$

$$q_\alpha \in \mathcal{Q}, \quad (3)$$

and ρ is the check function defined as

$$\rho_\alpha(x) = \begin{cases} \alpha x & \text{if } x \geq 0 \\ (1-\alpha)x & \text{if } x < 0 \end{cases}. \quad (4)$$

Let Y be a real valued random variable. The quantile function $Q_Y : [0,1] \to \mathbb{R}$ is defined pointwise by its α-quantile, which is given by where F_Y is the distribution function of random variable Y and $\alpha \in [0,1]$. Equation 1 defines what we call from now on as the quantile function $Q_Y(\cdot)$, in relation to random variable Y. In this article, we are interested in a conditional quantile function $Q_{Y|X=x} : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ (in short, from now on, $Q_{Y|X}(\cdot,\cdot)$), where X can be a vector.

In this work, we use the sample quantile, where we calculate the optimization based on a finite number of observations, instead of integrating over all domain of random variable Y. For the specific case where the random variable is a time series y_t, quantiles are estimated from a n size sample of observations of y_t and a explanatory variable x_t for each t, such that our random sample is formed by the sequence $\{y_t, x_t\}_{t=1}^n$. To estimate the α-quantile from a sample, we change **??** for the following optimization problem:

where $T = \{1, \ldots, n\}$, $|x|^+ = \max\{0, x\}$ and $|x|^- = -\min\{0, x\}$. The solution from the above problem is an estimator $\hat{Q}_{Y|X}$ for the quantile function $Q_{Y|X}$.

To model this problem as a Linear Programming problem, thus being able to use a modern solver to fit our model, we create variables ε_t^+ e ε_t^- to represent $|y - q(\cdot)|^+$ and $|y - q(\cdot)|^-$, respectively. The optimal argument $q^*_\alpha(\cdot)$ on the Linear Programming problem 5 is the estimated α-quantile for the given random sample.

$$\begin{aligned} q^*_\alpha(.) \in \arg\min_{q_\alpha(\cdot), \varepsilon_t^+, \varepsilon_t^-} & \sum_{t \in T} \left(\alpha\varepsilon_t^+ + (1-\alpha)\varepsilon_t^-\right) \\ \text{s.t. } & \varepsilon_t^+ - \varepsilon_t^- = y_t - q_\alpha(x_t), \qquad \forall t \in T, \\ & \varepsilon_t^+, \varepsilon_t^- \geq 0, \qquad \forall t \in T. \end{aligned} \quad (5)$$

Let A be a set containing a sequence of probabilities α_i such that $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_Q < 1$. This set represents a finite discretization of the interval $[0,1]$. One of our goals with quantile regression is to estimate a quantile function $\hat{Q}_{Y|X}$ of a given real valued random variable X from a sequence of quantiles $q_{\alpha_1}(x_t) \leq q_{\alpha_2}(x_t) \leq \cdots \leq q_{\alpha_{|A|}}(x_t)$, with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{|A|} < 1$, for any given t. The process of fitting $\hat{Q}_{Y|X}$ is by mapping every α_i with its estimated quantile $\hat{q}_{\alpha_i}(x_t)$. The denser the grid of values in A, better is the approximation of $Q_{Y|X}$. Thus, the distribution found for Y is nonparametric, as no previous assumptions are made about its shape, and its form is fully recovered by the data we have.

A typical problem, however, arises when working with quantile regression. When quantiles are estimated independently, it is possible to find $q_\alpha(x_t) > q_{\alpha'}(x_t)$, for a given t, when $\alpha_1 < \alpha_2$. An example can be seen on Figure **??**, where quantiles $\alpha = 0.95$ and $\alpha = 0.9$ cross. This problem, called *crossing quantiles*, can be prevented by estimating all quantiles with a single minimization problem.

As opposed to traditional parametric models where one has to assume the error distribution on the model

$$y_t = h(F_t) + \varepsilon_t, \quad (6)$$

as we choose the nonparametric methodology we are released from this step. A distribution for ε_t is defined within the estimation procedure. The predictive distribution is derived from the individual quantiles

$$\hat{q}^\alpha_{t+k|k} = \hat{y}_{t+k|t} + \hat{F}^\varepsilon_{t,k}(\alpha). \quad (7)$$

[13] defines a model based on extreme machine learning to estimate every α-quantile in a thin grid $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_Q < 1$.

A. Linear parametric explanatory model

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

B. General non-linear nonparametric model

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

REFERENCES

[1] International energy agency. [Online]. Available: https://www.iea.org/newsroom/news/2013/october/wind-power-seen-generating-up-to-18-of-global-power-by-2050.html

[2] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, Apr. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032114000446

[3] D. Creal, S. J. Koopman, and A. Lucas, "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, vol. 28, no. 5, pp. 777–795, 2013.

[4] R. Koenker, *Quantile regression*. Cambridge university press, 2005, no. 38.

[5] R. Koenker, Z. Xiao, J. Fan, Y. Fan, M. Knight, M. Hallin, B. J. M. Werker, C. M. Hafner, O. B. Linton, and P. M. Robinson, "Quantile Autoregression [with Comments, Rejoinder]," *Journal of the American Statistical Association*, vol. 101, no. 475, pp. 980–1006, 2006. [Online]. Available: http://www.jstor.org/stable/27590777

[6] V. Chernozhukov, I. Fernández-Val, and A. Galichon, "Quantile and Probability Curves Without Crossing," *Econometrica*, vol. 78, no. 3, pp. 1093–1125, May 2010. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.3982/ECTA7880/abstract

[7] D. Bertsimas, A. King, and R. Mazumder, "Best Subset Selection via a Modern Optimization Lens," *arXiv:1507.03133 [math, stat]*, Jul. 2015, arXiv: 1507.03133. [Online]. Available: http://arxiv.org/abs/1507.03133

[8] A. Belloni and V. Chernozhukov, "L1-Penalized Quantile Regression in High-Dimensional Sparse Models," *arXiv:0904.2931 [math, stat]*, Apr. 2009, arXiv: 0904.2931. [Online]. Available: http://arxiv.org/abs/0904.2931

[9] G. Ciuperca, "Adaptive LASSO model selection in a multiphase quantile regression," *Statistics*, vol. 50, no. 5, pp. 1100–1131, Sep. 2016. [Online]. Available: http://dx.doi.org/10.1080/02331888.2016.1151427

[10] J. K. Møller, H. A. Nielsen, and H. Madsen, "Time-adaptive quantile regression," *Computational Statistics & Data Analysis*, vol. 52, no. 3, pp. 1292–1303, Jan. 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167947307002502

[11] 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, no. 1-2, pp. 95–108, 2006.

[12] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, Jan. 2004. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/we.107/abstract

[13] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct Quantile Regression for Nonparametric Probabilistic Forecasting of Wind Power Generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767–2778, Jul. 2017.