Nongaussian time series model via Quantile Regression

Marcelo Ruas 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 an emergent topic which is demanding attention from the academic community. The installed capacity of renewable energy plants has been increasing in a fast pace and projections point out that wind power alone will account to 18% of global power by 2050 [1]. In spite of its virtues, several new challenges are inherent when dealing with such power source. New statistical models capable to handle with such difficulties are an emerging field in power systems literature . The main objective in such literature is to propose new models capable of generating scenarios of renewable energy source which are demanded in (i) energy trading, (ii) unit commitment, (iii) grid expansion planning, and (iv) investment decisions (see ([2]–[5]) and references therein). To provide good scenarios from an array of potential influential factors, one has to properly select which features are relevant and create a good model for the conditional distribution. Notwithstanding, a little attention is devoted to addressing both at the same time.

Conventional statistical models are often focused on estimating the conditional mean of a given random variable. By reducing the outcome to a single statistic, we loose important informations about the series random behavior. In order to account for the process inherent variability we focus our work on probability forecasting. [6] reviews the commonly used methodologies regarding probabilistic forecasting models, splitting them in parametric and nonparametric classes. Main characteristics of **parametric models** are (i) assuming a distribution shape and (ii) low computational costs. 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 (QR), Kernel Density Estimation, Artificial Intelligence or a mix of them.

Most time series methods rely on the assumption of Gaussian errors. However, renewable series such as wind and solar are reported as non-Gaussian [7]–[10]. To circumvent this problem, the usage of nonparametric methods, which doesn't rely on assuming any previously assumed distribution of errors, is adequate. We choose to use Quantile Regression (QR) as a tool for constructing a model for time series where errors are non-Gaussian, because of its facility to implement on commercial solvers and to extend the original model. However, when estimating a distribution function, as each quantile is estimated independently, the monotonicity of the distribution function may be violated. To get around this issue (known as crossing-quantiles) we propose to either add a constraint on the optimization model (which is more computationally intensive) or making a transformation as in [23], which can then be estimated independently.

The seminal work [11] defines QR as we use today. By this formulation, the conditional quantile is the solution of an optimization problem where we minimize the sum of the check function (defined formally in the next session). Instead of using the classical regression to estimate the conditional mean, the QR determines any quantile from the conditional distribution. Applications are enormous, ranging from risk measuring at financial funds (the Value-at-Risk) to a central measure robust to outliers. By estimating many quantiles on a thin grid of probabilities, one can have as many points as desired of the estimated conditional distribution function. In [12], the application of QR is extended to time series, when the covariates are lagged values of y_t. In our work, beyond autoregressive terms, it is also considered other exogenous variables as covariates.

In [13]–[17], QR is employed to model the conditional distribution of Wind Power Time Series. An updating quantile regression model is presented by [14]. The authors present a modified version of the simplex algorithm to incorporate new observations without restarting the optimization procedure. In [15], the authors build a quantile model from already existent independent Wind Power forecasts. The approach by [13] is to use QR with a nonparametric methodology. The authors add a penalty term based on the Reproducing Kernel Hilbert Space, which allows a nonlinear relationship between the explanatory variables and the output. This paper also develops an online learning technique, where the model is easily updated after each new observation. In [17], wind power probabilistic forecasts are made by using QR with a special type of Neural Network (NN) with one hidden layer, called extreme learning machine. In this setup, each quantile is a different linear combination of the features of the hidden layer. The authors of [18] use the weighted Nadaraya-Watson to estimate the conditional function in the time series.

Even though regularization is a topic already explored in previous QR articles, none of the aforementioned works that use QR to estimate the conditional distribution make use of regularization to do variable selection and arrive in a parsimonious model. The work by [19] defines the proprieties and convergence rates for QR when adding a penalty proportional to the ℓ_1-norm to perform variable selection, using the same idea as the LASSO [20]. The ADALASSO equivalent to QR is proposed by [21]. In this variant, the penalty for each variable has a different weight, and this modification ensures that the oracle propriety is being respected.

For the best of the authors knowledge, no other work has developed a methodology where regularization and estimation of the conditional distribution using QR is carried on at the same time. We propose to attack both problems simultaneously by using either Mixed Integer Linear Programming (MILP) or a LASSO penalization. On the LASSO formulation, regularization is performed for an individual quantile as described in [19], with the difference that all quantiles are estimated at the same time. In [22], the best subset with size K is selected by solving a MILP problem to minimize the sum of squared errors. The idea is straightforward: integer variables are used to count whether a variable is included or not in the model; a total number of K variables is allowed. Model selection for QR is performed using this same approach. The advantage we highlight on using the latter methodology is that the solution provided is optimal in the sense of minimizing the check function for a given number K of variables.

The objective of this paper is, then, to propose a new methodology to address nonparametric time-series focused on renewable energy. In our analysis, we develop both nonlinear and linear models for QR. The main contributions are:

- A nonparametric methodology to model the conditional distribution of time series.
- On the linear case, we propose a parsimonious methodology that selects the global optimal solution.
- Regularization techniques applied to an ensemble of quantile functions to estimate the conditional distribution.

The remaining of the paper is organized as follows. In section II, we present both the linear parametric and the nonlinear QR based time series models. In section III, we discuss the estimation procedures for them. The regularization strategies are also presented on this section. Finally, in section IV, a case study using real data from both solar and wind power is presented in order to test our methodology. Section V will conclude this article.

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] A. Moreira, D. Pozo, A. Street, and E. Sauma, “Reliable renewable generation and transmission expansion planning: Co-optimizing system’s resources for meeting renewable targets,” *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–1, 2016.

[3] R. Jabr, “Robust transmission network expansion planning with uncertain renewable generation and loads,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4558–4567, 2013.

[4] C. Zhao and Y. Guan, “Data-driven stochastic unit commitment for integrating wind generation,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2587–2596, July 2016.

[5] A. C. Passos, A. Street, and L. A. Barroso, “A dynamic real option-based investment model for renewable energy portfolios,” *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 883–895, March 2017.

[6] 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

[7] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and M. Constantinescu, “Time adaptive conditional kernel density estimation for wind power forecasting,” *Sustainable Energy, IEEE Transactions on*, vol. 3, no. 4, pp. 660–669, 2012.

[8] J. Jeon and J. W. Taylor, “Using conditional kernel density estimation for wind power density forecasting,” *Journal of the American Statistical Association*, vol. 107, no. 497, pp. 66–79, 2012.

[9] J. W. Taylor and J. Jeon, “Forecasting wind power quantiles using conditional kernel estimation,” *Renewable Energy*, vol. 80, pp. 370–379, 2015.

[10] 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, July 2017.

[11] R. Koenker and G. Bassett Jr, “Regression quantiles,” *Econometrica: journal of the Econometric Society*, pp. 33–50, 1978.

[12] 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

[13] C. Gallego-Castillo, R. Bessa, L. Cavalcante, and O. Lopez-Garcia, “On-line quantile regression in the rkhs (reproducing kernel hilbert space) for operational probabilistic forecasting of wind power,” *Energy*, vol. 113, pp. 355–365, 2016.

[14] 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

[15] 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.

[16] 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

[17] 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.

[18] Z. Cai, “Regression Quantiles for Time Series,” *Econometric Theory*, vol. 18, no. 1, pp. 169–192, 2002. [Online]. Available: http://www.jstor.org/stable/3533031

[19] 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

[20] R. Tibshirani, “Regression shrinkage and selection via the lasso,”

Journal of the Royal Statistical Society. Series B (Methodological), pp. 267–288, 1996.

[21] 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

[22] 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

[23] 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

[24] M. Efroymson, "Multiple regression analysis," *Mathematical methods for digital computers*, vol. 1, pp. 191–203, 1960.

[25] Y. Li and J. Zhu, "L1-norm quantile regression," *Journal of Computational and Graphical Statistics*, 2012.

[26] A. Belloni and V. Chernozhukov, "Least squares after model selection in high-dimensional sparse models," 2009.

[27] J. A. Machado, "Robust model selection and m-estimation," *Econometric Theory*, vol. 9, pp. 478–493, 1993.