Untitled

Marcelo

November 16, 2016

Introduction

Introduction

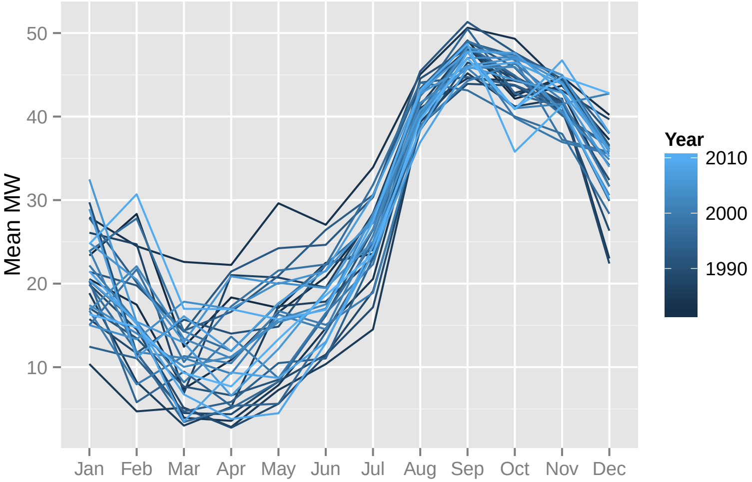
- Wind Firm Energy Certificate (FEC) (Porrua, 2010) estimation imposes several challenges.
- First, it is a quantile function of an aleatory quantity, named here on wind capacity factor (WP). Due to its non-dispachable profile, accurate scenario generation models could reproduce a fairly dependence structure in order to the estimation of FEC.
- Second, as it is a quantile functions, the more close to the extremes of the interval, the more sensitive to sampling error.

Introduction

- The main frameworks we investigate are parametric linear models and a non-parametric regression.
- In all approaches we use the time series lags as the regression covariates.
- To study our methods performance, we use the mean power monthly data of Icaraizinho, a wind farm located in the Brazilian northeast.

The Icaraizinho dataset

- The Icaraizinho dataset consists of monthly observations from 1981 to 2011 of mean power measured in Megawatts.



Quantile Regression

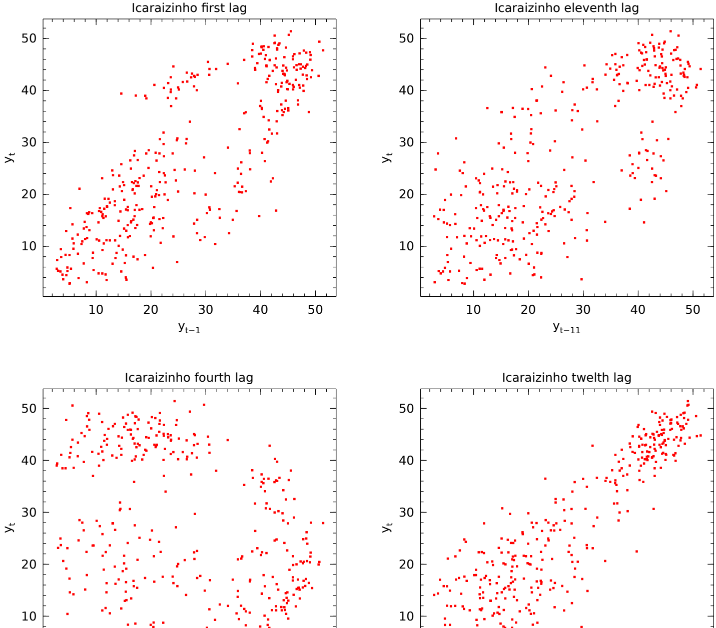
A Quantile Regression for the α-quantile is the solution of the following optimization problem:

$$\min_{q} \sum_{t=1}^{n} \alpha|y_t - q(x_t)|^+ + (1-\alpha)|y_t - q(x_t)|^-,$$

where $q(x_t)$ is the estimated quantile value at a given time t and $|x|^+ = \max\{0, x\}$ and $|x|^- = -\min\{0, x\}$. To model this problem as a Linear Programming problem, thus being able to use a modern solver to fit our model, we can create variables ε_t^+ e ε_t^- to represent $|y - q(x_t)|^+$ and $|y - q(x_t)|^-$, respectively. So we have:

$$\begin{aligned}
\min_{q,\varepsilon_t^+,\varepsilon_t^-} \; & \sum_{t=1}^{n} \left(\lambda\varepsilon_t^+ + (1-\lambda)\varepsilon_t^-\right) & \\
\text{s.t. } \; & \varepsilon_t^+ - \varepsilon_t^- = y_t - q(x_t), & \forall t \in \{1,\dots,n\}, \\
& \varepsilon_t^+, \varepsilon_t^- \geq 0, & \forall t \in \{1,\dots,n\}.
\end{aligned}$$

Relationship between y_t and some lags



Linear Models for the Quantile Autoregression

Best subset selection with Mixed Integer Programming

- We investigate the usage of Mixed Integer Programming to select which variables are included in the model, up to a limit of inclusions imposed *a priori*. The optimization problem is described below:

$$
\begin{aligned}
\min_{\beta_0,\beta,z,\varepsilon_t^+,\varepsilon_t^-} \quad & \textstyle\sum_{t=1}^{n} \left(\alpha\varepsilon_t^+ + (1-\alpha)\varepsilon_t^-\right) & \\
\text{s.t} \quad & \varepsilon_t^+ - \varepsilon_t^- = y_t - \beta_0 - \textstyle\sum_{p=1}^{P} \beta_p x_{t,p}, & \forall t \in \{1,\dots,n\}, \\
& \varepsilon_t^+, \varepsilon_t^- \geq 0, & \forall t \in \{1,\dots,n\}, \\
& -M_U z_p \leq \beta_p \leq M_U z_p, & \forall p \in \{1,\dots,P\}, \\
& \textstyle\sum_{p=1}^{P} z_p \leq K, & \\
& z_p \in \{0,1\}, & \forall p \in \{1,\dots,P\}.
\end{aligned}
$$

Best subset selection with Mixed Integer Programming

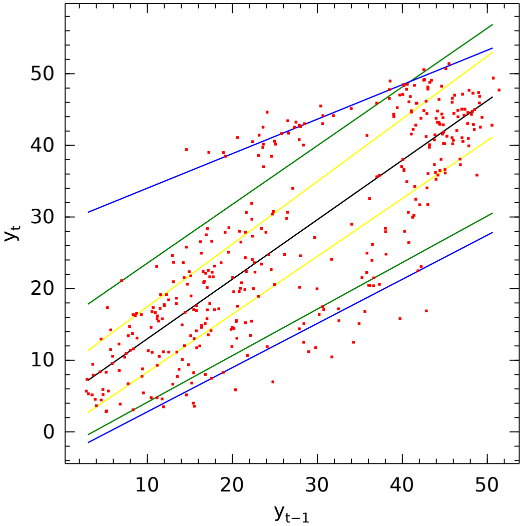


Figure 3: Linear Quantile Regression when only y_{t-1} is used

Best subset selection with Mixed Integer Programming

	K=1	K=2	K=3	K=4	K=5	K=6	K=7	K=8
β_0	-15.33	9.38	1.48	1.34	8.72	-1.68	4.94	0.65
β_1	-0.00	0.79	0.66	0.58	0.46	0.40	0.48	0.46
β_2	-0.00	-0.00	-0.00	-0.00	-0.00	0.33	-0.00	-0.00
β_3	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.20
β_4	-0.00	-0.47	-0.28	-0.27	-0.29	-0.35	-0.31	-0.40
β_5	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
β_6	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.11	0.08
β_7	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
β_8	-0.00	-0.00	-0.00	-0.00	-0.15	-0.00	-0.31	-0.26
β_9	-0.00	-0.00	-0.00	-0.00	-0.00	0.14	0.16	0.20
β_{10}	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
β_{11}	-0.00	-0.00	0.26	0.17	0.21	0.08	0.16	0.19
β_{12}	1.17	-0.00	-0.00	0.18	0.15	0.19	0.22	0.20

Table 1: Coefficients for quantile $\alpha = 0.05$

Best subset selection with a ℓ_1 penalty

- Another way of doing regularization is including the ℓ_1-norm of the coefficients on the objective function.
- By lowering the penalty we impose on the ℓ_1-norm, more variables are being added to the model.
- This is the same strategy of the LASSO, and its usage for the quantile regression is discussed in Li and Zhu (2012).
- The proposed optimization problem to be solved is:

$$\min_{\beta_0,\beta} \sum_{t=1}^{n} \alpha |y_t - q(x_t)|^+ + (1-\alpha)|y_t - q(x_t)|^- + \lambda \|\beta\|_1$$

$$q(x_t) = \beta_0 - \sum_{p=1}^{P} \beta_p x_{t,p},$$

Best subset selection with a ℓ_1 penalty

- In order to represent the above problem to be solved with linear programming solver, we restructure the problem as below:

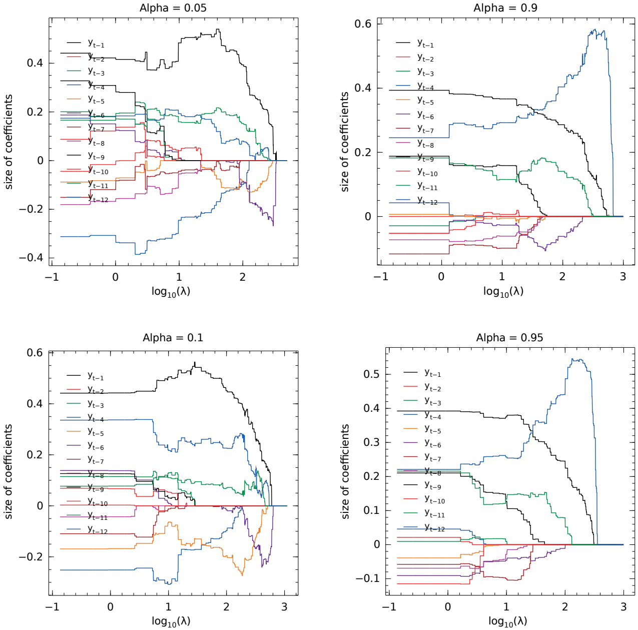
$$\beta_\lambda^{*LASSO} = argmin_{\beta_0,\beta,\varepsilon_t^+,\varepsilon_t^-} \sum_{i=1}^{n} \left(\alpha\varepsilon_t^+ + (1-\alpha)\varepsilon_t^-\right) + \lambda \sum_{p=1}^{P} \xi_p$$

$$\begin{aligned}
&\varepsilon_t^+, \varepsilon_t^- \geq 0, \qquad \forall t \in \{1,\ldots,n\}, \\
&\varepsilon_t^+, \varepsilon_t^- \geq 0, \qquad \forall t \in \{1,\ldots,n\}, \\
&\xi_p \geq \beta_p, \qquad \forall p \in \{1,\ldots,P\}, \\
&\xi_p \geq -\beta_p, \qquad \forall p \in \{1,\ldots,P\},
\end{aligned}$$

Best subset selection with a ℓ_1 penalty

- For low values of λ, the penalty is small and thus we have a model where all coefficients have a nonzero value. - When λ is increased, the coefficients shrink towards zero (as an extreme case we have a constant model) - The linear coefficient β_0 is not penalized. - We make this experiment For the same quantiles values of α we experimented on the previous section ($\alpha \in \{0.05, 0.1, 0.5, 0.9, 0.95\}$).

Best subset selection with a ℓ_1 penalty



Simulation Study

- We propose simulating an AR(1) model

$$y_t = \phi_0 + \phi y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \tag{1}$$

and test two approaches to predict the one-step ahead quantile. - On the first one, we consider known this process true model, thus estimating values for $\hat{\phi}_0$, $\hat{\phi}$ and $\hat{\sigma}_\epsilon^2$. - On the second approach, we use the quantile regression to make a direct estimation of the quantiles.

Simulation Study

Simulation Study

Next steps

Testing methods with high-frequency data

- All methods already discussed

Local Quantile Regression

- Based on Bremmes (2004). Being the estimation of the θ quantile $q_\theta()$:
$$q_\theta(x; \alpha_0, \alpha) = \alpha_0 + \alpha^T x$$
we define the local quantile regression by solving the following optimization problem:
$$argmin_{(\alpha_0,\alpha)} \sum_{i=1}^{n} \rho_\theta(e_{\rho,i} - q_\theta(x_i - x; \alpha_0, \alpha)) w\left(\frac{\|x_i - x\|_2}{h_\lambda(x)}\right)$$
where the loss function is defined by
$$\rho_\theta(u) = \begin{cases} u\theta & \text{if } u \geq 0 \\ u(\theta - 1) & \text{otherwise} \end{cases}$$
and the weight function
$$w(u) = \begin{cases} (1 - u^3)^3 & \text{if } u \in [0, 1) \\ 0 & \text{otherwise} \end{cases}$$

Generating Scenarios of Wind Power Prediction

- Following recommendation of Pinson and Madsen (2009). This method was employed for short term monte carlon simulation.

1. One must have a different model for each horizon $k \in \{1, \ldots, K\}$. Residuals $\boldsymbol{X}$ comes from a multivariate Gaussian distribution $\boldsymbol{X} \sim N(\mu_0, \Sigma_{t-k})$.
2. By applying the inverse probit function Φ to each component of $\boldsymbol{X}^{(i)}$, we obtain the random variable $Y_k^{(t)} = \hat{F}_{t+k|t}(p_{t+k}), \forall t$. must be estimated for each desired value of k. As we want to generate d different scenarios, we have to sample
3.

Simulation Study