



## **MODELOS ADITIVOS GENERALIZADOS NA PREVISÃO DE QUANTIS DE GERAÇÃO EÓLICA**

**José Francisco Moreira Pessanha**

Centro de Pesquisas de Energia Elétrica - Eletrobras Cepel  
Av. Horácio Macedo 354, Cidade Universitária, Ilha do Fundão, Rio de Janeiro - RJ  
francisc@cepel.br

**Valk Luiz de Oliveira Castellani**

Centro de Pesquisas de Energia Elétrica - Eletrobras Cepel  
Av. Horácio Macedo 354, Cidade Universitária, Ilha do Fundão, Rio de Janeiro - RJ  
valk@cepel.br

**Victor Andrade de Almeida**

Centro de Pesquisas de Energia Elétrica - Eletrobras Cepel  
Av. Horácio Macedo 354, Cidade Universitária, Ilha do Fundão, Rio de Janeiro - RJ  
andrade@cepel.br

### **RESUMO**

Previsões de curto prazo da geração eólica constituem informações fundamentais para a integração segura e econômica de parques eólicos à um sistema elétrico de potência. No presente trabalho apresenta-se um modelo aditivo generalizado para previsão dos quantis da geração eólica (regressão quantílica), a partir dos quais se obtém uma previsão da densidade de probabilidade da produção de um parque eólico. A metodologia foi implementada no programa computacional VENTOS. Para ilustrar a aplicação da metodologia e as funcionalidades do programa apresentam-se os resultados de um experimento computacional com dados reais de um parque eólico localizado na Galícia, Espanha.

**PALAVRAS CHAVE. Energia eólica, Regressão quantílica, Previsão probabilística.**

**Tópicos EN, EST, AD&GP**

### **ABSTRACT**

Short-term wind power forecasts are fundamental information for the safe and economic integration of wind farms into an electric power system. In this work we present a generalized additive model to predict the wind power quantiles (quantile regression) from which we obtain a prediction of the wind power probability density function in a wind farm. The methodology was implemented in the VENTOS computational program. In order to illustrate the application of the methodology as well as the functionalities of the program this work presents the results achieved by a computational experiment based on real data from a wind farm located in Galicia, Spain.

**KEYWORDS. Wind power, Quantile regression, Probabilistic forecasting.**

**Paper topics EN, EST, AD&GP**



## 1. Introdução

O comportamento intermitente da velocidade do vento impõe desafios para a integração segura e econômica dos aproveitamentos eólicos aos sistemas de potência. Para viabilizar a integração da geração eólica, os centros de operação demandam novas ferramentas de gerenciamento de energia elétrica, entre as quais se destaca a previsão da geração eólica [Estanqueiro et al. 2008]. Em geral, os métodos de previsão privilegiam a predição do valor esperado da geração eólica sem, contudo, quantificar e informar apropriadamente a incerteza da previsão. Em virtude do comportamento aleatório e intermitente do vento, a previsão do valor esperado da geração eólica sem uma quantificação adequada da incerteza não atende completamente às necessidades da operação. Por exemplo, em sistemas com grande participação de fontes eólicas, a reserva operativa deve ser dimensionada de maneira a cobrir as incertezas nas previsões de carga e de geração eólica [Matos e Bessa 2011].

A densidade de probabilidade contém toda a informação de uma variável aleatória, portanto, para lidar com a incerteza, deve-se considerar as previsões de densidade de probabilidade da geração eólica fornecidas por métodos de previsão probabilística [Zhang et al. 2014]. Entre os principais métodos de previsão probabilística da geração eólica destacam-se a regressão quantílica [Nielsen et al. 2006], a estimação não paramétrica de densidade [Bessa et al. 2011] e o processo gaussiano [Chen et al. 2014].

Neste artigo apresentam-se as linhas gerais de uma metodologia de previsão probabilística da geração eólica, útil na programação diária da operação e na operação em tempo real de sistemas elétricos. A metodologia baseia-se na equação de regressão quantílica especificada por Nielsen et al. (2006), um modelo aditivo generalizado [Wood 2006] para previsão de quantis de geração eólica. A opção por esta abordagem reside na sua relativa facilidade de implementação computacional e na agilidade em ajustar os modelos de previsão. A metodologia de previsão foi implementada em C++ e encontra-se disponível no Programa VENTOS, o módulo de previsão de geração eólica que integra o Sistema de Encadeamento de Modelos Energéticos - ENCAD desenvolvido pelo CEPEL (Figura 1).

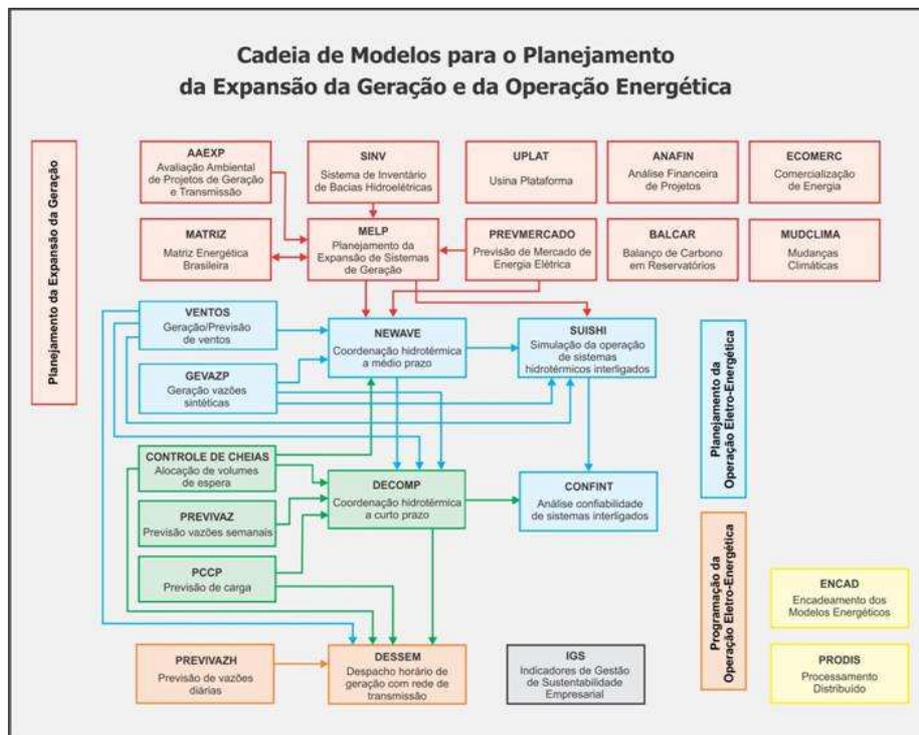


Figura 1 - Sistema de Encadeamento de Modelos Energéticos (ENCAD)  
<http://www.cepel.br/areas-de-atuacao/otimizacao-energetica-e-meio-ambiente.htm>

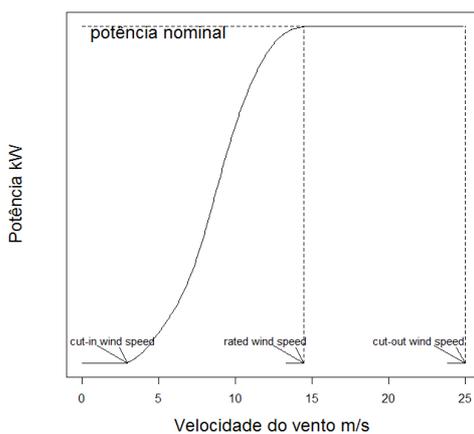


Em sua primeira versão, o programa VENTOS fornece as previsões pontuais (medianas) da geração eólica consideradas pelo modelo DESSEM na otimização da programação diária da operação.

A seguir, na seção 2, tem-se uma breve introdução à previsão da geração eólica. Na sequência, na seção 3, apresenta-se a especificação do modelo de previsão probabilística da geração eólica, bem como uma descrição do método de interpolação utilizado na obtenção de previsões com resolução temporal semi-horária a partir de dados horários. Na seção 4 tem-se a especificação dos dados de entrada e resultados gerados pelo programa VENTOS. A aplicação da metodologia é ilustrada na seção 5 por meio dos resultados de um experimento computacional com dados reais provenientes de um parque eólico localizado na Galícia, Espanha. Por fim, na seção 6, tem-se as principais conclusões do trabalho.

## 2. Energia eólica

Sob condições ideais, a relação entre a velocidade do vento  $v$  e a produção de energia  $P$  é definida pela curva de potência  $P(v)$  indicada na Figura 2. Conforme ilustrado nessa figura, a produção de energia inicia-se em velocidades de vento entre 2 e 3 m/s (*cut-in wind speed*). Em seguida, a produção de energia cresce rapidamente com o aumento da velocidade do vento até encontrar a potência nominal (*rated power*), quando a velocidade alcança um valor entre 12 e 17 m/s (*rated wind speed*). Para velocidades acima de 25 m/s (*cut-out wind speed*), a turbina eólica deve ser desligada para evitar risco de danos ao rotor [Custódio 2009].



$$P(v) = \begin{cases} 0, & v \leq V_{in} \\ 0.5C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v^3, & V_{in} < v \leq V_r \\ 0.5C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot V_r^3, & V_r < v \leq V_{out} \\ 0, & v > V_{out} \end{cases}$$

em que

$V_{in}$  = *cut-in wind speed* (m/s)

$V_r$  = *rated wind speed* (m/s)

$V_{out}$  = *cut-out wind speed* (m/s)

$C_p$  = coeficiente de performance

$\eta_m$  = eficiência do gerador

$\eta_g$  = eficiência da transmissão

$\rho$  = densidade do ar (kg/m<sup>3</sup>)

$A$  = área de varredura do rotor (m<sup>2</sup>)

Figura 2 - Curva de potência

Assim, uma forma simples de gerar a previsão da produção eólica consiste em prever a velocidade do vento e encontrar na curva de potência a geração correspondente. Porém, conforme ilustrado pelo diagrama de dispersão na Figura 3, construído a partir de medições de velocidade do vento e geração, para uma dada velocidade do vento, a produção eólica pode apresentar diferentes valores.

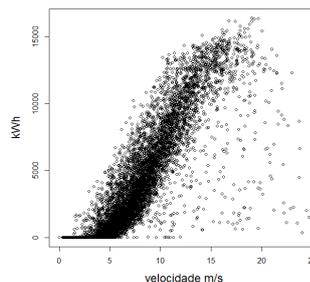


Figura 3 - Medições de velocidade e geração



Portanto, a relação entre a velocidade do vento e geração eólica é estocástica, logo obter apenas uma previsão pontual da produção eólica sem conhecer a incerteza da previsão pode não ser suficiente. Por exemplo, em sistemas com grande participação de fontes eólicas, a reserva operativa deve ser dimensionada de maneira a cobrir as incertezas na previsão de carga e na previsão da produção eólica [Matos e Bessa 2011]. A incerteza da produção eólica pode ser quantificada por meio da previsão da distribuição de probabilidade (previsão probabilística) da geração eólica em cada instante do horizonte de previsão.

A função densidade de probabilidade contém toda a informação de uma variável aleatória e a partir dela é possível obter uma previsão pontual (por exemplo, a média ou a mediana) e medidas de incerteza e risco (por exemplo, o desvio-padrão e a distância interquartílica). Os quantis de uma distribuição de probabilidade constituem uma forma simples de descrevê-la. Por exemplo, a distribuição padrão  $N(0,1)$  pode ser descrita pelos quantis indicados na Figura 4.

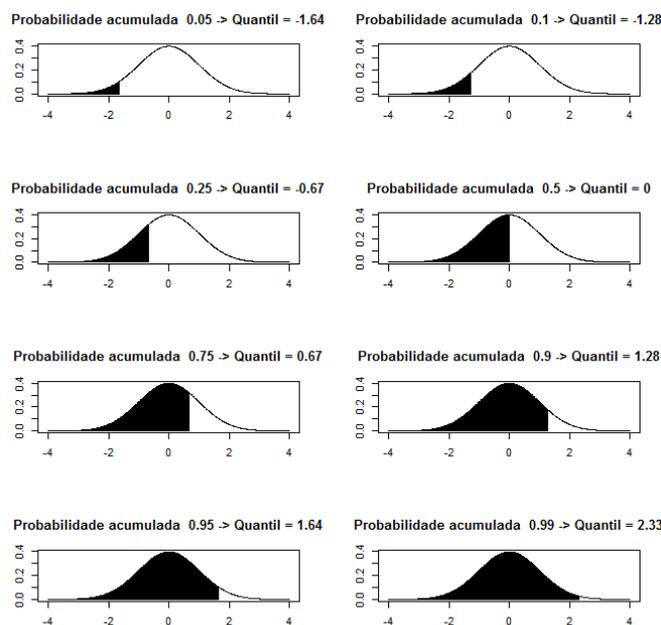


Figura 4 - Probabilidades acumuladas e quantis da distribuição normal padrão

As incertezas nas previsões de vento indicam que a abordagem de previsão mais adequada para a operação de sistemas elétricos com expressiva participação eólica consiste em prever a distribuição de probabilidade da geração eólica, um resultado que pode ser alcançado por meio de modelos de previsão probabilística, por exemplo, modelos de regressão quantílica [Nielsen et al. 2006],[Anastasiades e McSharry 2013].

### 3. Modelo de previsão probabilística da geração eólica

Em um modelo de regressão linear, a equação de regressão estimada fornece uma estimativa do valor esperado da variável resposta  $y$  em função da variável explicativa  $X$ . Já em um modelo de regressão quantílica, a equação de regressão fornece uma estimativa do quantil  $\tau$  (por exemplo,  $\tau=0,5$  para a mediana) da variável resposta em função da variável explicativa. Assim, o conjunto de equações de regressão estimadas para diferentes valores de  $\tau$ , permite caracterizar a distribuição de probabilidade da variável resposta condicionada aos valores das variáveis explicativas.

De forma distinta do estimador mínimos quadrados, na regressão quantílica para um quantil  $\tau$  o estimador dos coeficientes de regressão busca minimizar a seguinte função objetivo em uma amostra com  $n$  pares  $(y_i, X_i) \forall i=1, n$ :



$$\text{Min}_{\beta} \sum_{i=1}^n \rho_{\tau}(y_i - X_i^T \beta) \quad (1)$$

em que  $\rho_{\tau}(z) = z \cdot [\tau - I(z < 0)]$  e  $I(z < 0)$  é uma função indicadora que assume valor unitário para  $z < 0$ , caso contrário, é igual a zero.

A função objetivo (1) pode ser inserida no seguinte problema de programação linear (PPL), cuja solução fornece as estimativas dos  $K+1$  coeficientes de regressão:

$$\begin{aligned} \text{Min}_{u,v,\beta} \quad & \tau \sum_{i=1}^n u_i + (1-\tau) \sum_{i=1}^n v_i \\ \text{s.a} \quad & y_i - X_i^T \beta = u_i - v_i \quad \forall i = 1, n \\ & u_i \geq 0 \quad \forall i = 1, n \\ & v_i \geq 0 \quad \forall i = 1, n \\ & \beta \in R^{K+1} \end{aligned} \quad (2)$$

Para  $\tau = 0,5$  (mediana) o PPL em (2) corresponde ao estimador *least absolute deviations* (LAD) [Alcântara et al. 2003]. Alternativamente, a regressão quantílica pode ser ajustada por meio do algoritmo *Iteratively Reweighted Least Squares* (IRLS), um procedimento no qual o estimador mínimos quadrados ponderados é aplicado iterativamente [Kutner et al. 2004], conforme descrito a seguir:

- 1) Faça  $h=0$  (contador de iterações),  $\beta_h = 1_{K+1}$  (vetor com todos elementos iguais a 1) e  $\Omega = I_{K+1}$  (matriz identidade).
- 2) Faça  $h=h+1$  e calcule  $\beta_h = (X^T \Omega X)^{-1} X^T \Omega Y$  (mínimos quadrados ponderados).
- 3) Pare se o critério de convergência for alcançado, caso contrário continue.
- 4) Calcule o vetor de resíduos  $\hat{\epsilon} = Y - X\beta_h$  e construa o vetor de pesos  $w$  a partir dos resíduos:  
 Se  $\hat{\epsilon}_i < 0$  então  $w_i = \tau \cdot \hat{\epsilon}_i \quad \forall i = 1, n$   
 Se  $\hat{\epsilon}_i \geq 0$  então  $w_i = (1-\tau) \cdot \hat{\epsilon}_i \quad \forall i = 1, n$
- 5) Faça  $\Omega_{ii} = 1/w_i$  (diagonal da matriz  $\Omega$  igual a  $1/w$ ) e volte para o passo 2.

Em geral, as previsões anemométricas são informadas na forma de coordenadas polares: módulo da velocidade ( $V$ ) e direção ( $\theta$ ). A direção é um ângulo, uma variável aleatória circular [Pewsey et al. 2013], e por esta razão é mais conveniente trabalhar com as coordenadas cartesianas da velocidade do vento:  $u = V \cos(\theta)$  e  $w = V \sin(\theta)$ .

No programa VENTOS adota-se a especificação do modelo de regressão quantílica proposta em [Nielsen et al. 2006], em que o quantil da geração eólica  $P$  depende das coordenadas cartesianas ( $u, w$ ) da velocidade do vento. Os quantis estimados pelo conjunto de equações de regressão descrevem a densidade condicionada  $f(P|u, w)$ . Para capturar as não linearidades nas relações entre a geração eólica e as componentes  $u$  e  $w$ , Nielsen et al. (2006) incluem como variáveis explicativas os coeficientes das componentes cartesianas  $u$  e  $w$  em dez bases *B-Splines* [Wood 2006] totalizando 20 variáveis explicativas. A especificação do modelo resultante corresponde ao modelo aditivo generalizado indicado a seguir:

$$\hat{Q}_t(\tau) = \beta_0(\tau) + \sum_{j=1}^{10} b_j(u_t^*) \beta_j^u(\tau) + \sum_{j=1}^{10} b_j(w_t^*) \beta_j^w(\tau) + \epsilon_t \quad (3)$$



em que  $\varepsilon_t$  denota um erro aleatório,  $\beta^u(\tau)$  e  $\beta^w(\tau)$  denotam os coeficientes de regressão para o quantil  $\tau$  e  $b(u^*)$  e  $b(w^*)$  representam as variáveis explicativas, neste caso, os coeficientes das componentes cartesianas  $u$  e  $w$  nas 10 bases *B-Splines*.

O modelo em (3) deve ser ajustado separadamente para cada quantil, por exemplo, para cada um dos nove quantis do conjunto {5%, 10%, 25%, 50%, 75%, 90%, 95%}. As equações resultantes formam um conjunto de equações de previsão. Vale ressaltar que o mesmo histórico de valores verificados da velocidade/direção do vento e da geração eólica é utilizado na estimação de todas as equações de previsão.

A produção de um parque eólico está confinada ao intervalo entre zero e a capacidade do parque. Porém, o estimador de mínimos quadrados ponderados usado na estimação dos coeficientes de regressão admite que a variável resposta é irrestrita. A aplicação dos algoritmos de estimação diretamente aos valores brutos da geração eólica pode resultar em quantis negativos ou com valores acima da capacidade do parque. Para contornar o problema, os dados de geração eólica são previamente transformados por meio da função arco tangente hiperbólica, aplicada aos valores normalizados da geração eólica:

$$y_t^{**} = 0,5 \cdot \ln \left[ \frac{1 + y_t^*}{1 - y_t^*} \right] \quad (4)$$

em que  $y_t^* = (y_t - \min(y)) / (\max(y) - \min(y))$ .

Os quantis expressos em MW ( $Q$ ) são obtidos pela aplicação da função tangente hiperbólica aos quantis estimados ( $q$ ) pelas equações de regressão e na sequência os valores resultantes desta transformação ( $q^*$ ) são escalonados para o intervalo definido pelos valores mínimo e máximo da geração eólica:

$$q^* = [\exp(q) - \exp(-q)] / [\exp(q) + \exp(-q)] \quad (5)$$

$$Q = [\max(y) - \min(y)] \cdot q^* + \min(y) \quad (6)$$

A aplicação do conjunto de equações estimadas (3) à uma dada previsão da velocidade/direção do vento fornece as previsões dos quantis da geração eólica. Por exemplo, a partir das previsões anemométricas horárias disponibilizadas pelo Operador Nacional do Sistema ([http://www.ons.org.br/operacao/previsao\\_horaria\\_vento.aspx](http://www.ons.org.br/operacao/previsao_horaria_vento.aspx)), pode-se prever os quantis da geração eólica em bases horárias até cinco dias à frente. Contudo, para as decisões a serem tomadas pelo operador do sistema elétrico é fundamental que também sejam disponibilizadas previsões com resolução semi-horária (30 minutos). A estratégia adotada pelo programa VENTOS para obter as previsões semi-horárias da geração eólica consiste em aplicar o modelo de regressão quantílica em (3) às previsões anemométricas horárias e na sequência fazer uma interpolação por partes com *splines cúbicas* [Dubinsky e Elperin 1997].

Seja o conjunto  $y_1, y_2, \dots, y_H$  das previsões horárias para o horizonte de previsão  $H$  passos à frente nos instantes  $t_1, t_2, \dots, t_H$  respectivamente. Seja  $f(t)$  uma função suave tal que o seu valor médio em um intervalo de tempo é igual ao valor horário previsto, ou seja, a função  $f(t)$  corresponde à geração instantânea.

$$\frac{1}{H_j} \int_{t_{j-1}}^{t_j} f(t) dt = y_j \quad (7)$$

onde  $H_j = t_j - t_{j-1}$ ,  $j=1, \dots, H$ .

Admitindo que todo período horário possa ser dividido em  $m$  intervalos menores, por exemplo, intervalos de 30 minutos, a demanda média em cada um dos  $m$  intervalos é dada por:

$$y_{ji} = \frac{1}{h_j} \int_{t_{ji-1}}^{t_{ji}} f(t) dt \quad (8)$$

onde  $h_j = H_j/m$ ,  $t_{jk} = t_j + kh_j$ ,  $j=1, \dots, n$  e  $k=1, \dots, m$ .

No lugar da função  $f(t)$ , pode-se considerar a sua função  $F(t)$  definida a seguir:



$$F(t) = \int_{t_0}^t f(z) dz \quad (9)$$

onde  $F(t_0)=0$ .

Os valores da função  $F(t)$  nos pontos  $t=1,2,\dots,H$  são obtidos a partir das previsões horárias:

$$F(t_j) = \sum_{k=1}^j \int_{t_{k-1}}^{t_k} f(t) dt = \sum_{k=1}^j H_j y_j \quad (10)$$

Dada a função  $F(t)$ , a função  $f(t)$  que descreve a evolução instantânea da geração eólica é obtida por diferenciação, ou seja,  $f(t) = dF/dt$ . Assim, o valor médio, em um intervalo com duração menor que uma hora, é calculado da seguinte forma:

$$y_{j_i} = \frac{1}{h_j} \left[ \int_{t_{j_i-1}}^{t_{j_i}} f(t) dt - \int_{t_0}^{t_{j_i-1}} f(t) dt \right] = \frac{1}{h_j} [F(t_{j_i}) - F(t_{j_i-1})] \quad (11)$$

Portanto, somente a função  $F(t)$  é necessária ao invés da curva de carga instantânea  $f(t)$ . Uma aproximação da função  $F(t_j)$ , cujos valores são conhecidos em  $n = H+1$  pontos ( $t_0, t_1, \dots, t_{H+1}$ ), pode ser obtida por meio de uma interpolação. Dado que o número de pontos a serem interpolados pode ser elevado, a interpolação por um único polinômio de ordem elevada poderia produzir erros significativos. Nestas situações, recomenda-se que a interpolação seja realizada por meio de muitos polinômios de baixa ordem, cada um válido no intervalo entre dois pontos, cujos coeficientes devem ser determinados de maneira a garantir a suavidade da curva ajustada, bem como das respectivas derivadas de primeira e segunda ordens, por exemplo, a interpolação por partes com *splines* cúbicas [Gilat e Subramaniam 2008].

Para um conjunto de dados com  $n$  pontos há  $n-1$  intervalos, a interpolação por partes com *splines* cúbicas implica no ajuste de um polinômio de terceiro grau ( $p_i(t)=a_i t^3 + b_i t^2 + c_i t + d_i$ ) em cada intervalo entre as horas  $t_i$  e  $t_{i+1}$  e requer a determinação de  $4(n-1)$  coeficientes de maneira que a curva resultante seja contínua. Uma especificação mais sofisticada consiste em utilizar polinômios na forma de Lagrange, uma opção que requer a solução de um sistema de equações lineares com apenas  $n-2$  equações. A segunda derivada de um polinômio de terceira ordem é linear, logo, dentro de cada intervalo entre  $t_i$  e  $t_{i+1}$ , a segunda derivada em um ponto  $t \in [t_i, t_{i+1}]$  pode ser expressa na forma de Lagrange:

$$d^2 p_i(t)/dt^2 = (t - t_{i+1})/(t_{i+1} - t_i) d^2 p_i(t_i)/dt^2 + (t - t_i)/(t_{i+1} - t_i) d^2 p_i(t_{i+1})/dt^2 \quad (12)$$

Integrando-se duas vezes a equação (12) e determinando-se as constantes de integração a partir dos valores conhecidos nos extremos do intervalo, ou seja,  $p_i(t_i) = y_i$  e  $p_{i+1}(t_{i+1})=y_{i+1}$ , obtém-se a equação do polinômio de terceiro grau  $p_i(t)$  para  $t \in [t_i, t_{i+1}]$ :

$$p_i(t) = \frac{p_i''(t_i)}{6(t_{i+1}-t_i)} (t_{i+1}-t)^3 + \frac{p_i''(t_{i+1})}{6(t_{i+1}-t_i)} (t-t_i)^3 + \left[ \frac{y_i}{t_{i+1}-t_i} - \frac{p_i''(t_i)(t_{i+1}-t_i)}{6} \right] (t_{i+1}-t) + \left[ \frac{y_{i+1}}{t_{i+1}-t_i} - \frac{p_i''(t_{i+1})(t_{i+1}-t_i)}{6} \right] (t-t_i) \quad \forall i=1, n-1 \quad (13)$$

Como há  $n-1$  intervalos, o conjunto de equações (13) forma um sistema com  $n-1$  equações lineares, cada uma com 2 incógnitas definidas pelas segundas derivadas  $d^2 p_i(t)/dt^2$  e  $d^2 p_i(t_{i+1})/dt^2$ . Assumindo a continuidade na curvatura da função ajustada, as segundas derivadas de polinômios adjacentes são iguais em cada ponto  $t_i, i=2, \dots, n-1$ . Conseqüentemente, há  $n$  valores que precisam ser determinados correspondentes às derivadas segundas  $d^2 p_i(t_i)/dt^2$  para  $i=1, \dots, n$ . Para garantir a continuidade da função ajustada aos  $n$  pontos conhecidos, as primeiras derivadas no ponto de intercessão de dois polinômios adjacentes devem ser iguais, ou seja,  $dp_i(t_{i+1})/dt = dp_{i+1}(t_{i+1})/dt$  para  $i=2, \dots, n-2$ . Derivando as equações (13) e aplicando a condição de igualdade das



derivadas primeiras obtém-se o sistema tridiagonal com  $n-2$  equações lineares e  $n$  incógnitas (as derivadas segundas) a seguir:

$$(t_{i+1} - t_i)p_i''(t_i) + 2(t_{i+2} - t_i)p_i''(t_{i+1}) + p_i(t_{i+2} - t_{i+1})p_i''(t_{i+2}) = 6 \begin{bmatrix} y_{i+2} - y_{i+1} & -y_{i+1} - y_i \\ t_{i+2} - t_{i+1} & t_{i+1} - t_i \end{bmatrix} \quad \forall i=1, n-2 \quad (14)$$

Para completar um sistema com  $n$  equações e  $n$  incógnitas admite-se que as segundas derivadas são nulas nos pontos inicial e final (*splines* cúbicas naturais):  $d^2p_1(t_1)/dt^2 = 0$  e  $d^2p_n(t_n)/dt^2 = 0$ . A solução do sistema de equações lineares resultante pode ser obtida de forma eficiente pelo algoritmo tridiagonal [Press et al. 1992] e na sequência ser introduzida nos polinômios de terceiro grau na equação (13) para obter o polinômio interpolador.

#### 4. Programa Ventos

O programa VENTOS é o módulo do sistema ENCAD responsável pelo fornecimento de previsões probabilísticas da geração eólica. Na versão atual o programa fornece previsões com resolução temporal horária e semi-horária. Basicamente, o programa possui dois modos de execução: ajuste de modelos de previsão (treinamento) e previsão, sendo que em ambos os casos, os dados de geração eólica e velocidade/direção do vento devem ser fornecidos em bases horárias.

No modo de treinamento, os dados históricos da geração eólica verificada, acompanhados dos valores verificados de velocidade/direção do vento, são utilizados com a finalidade de estimar os coeficientes dos modelos de regressão quantílica. Visando a previsão probabilística, os mesmos dados são utilizados na estimação de equações de regressão, uma para cada quantil  $\tau$ , por exemplo,  $\tau \in \{5\%, 10\%, 25\%, 50\%, 75\%, 90\%, 95\%\}$ . As equações são estimadas separadamente e ao final tem-se um conjunto de equações que informam as estimativas dos quantis da geração eólica em função da velocidade e da direção do vento.

No ajuste (treinamento) dos modelos de regressão quantílica, pode-se utilizar todo o conjunto de dados, porém para avaliar a capacidade preditiva do modelo ajustado recomenda-se segmentá-lo em dois conjuntos: *insample* e *outsample*. O conjunto *insample* é utilizado na estimação dos coeficientes de regressão, enquanto o conjunto *outsample* tem por finalidade avaliar a capacidade preditiva do modelo em novos dados não utilizados na estimação.

Na modo de previsão, as previsões de velocidade e direção do vento fornecidas por modelos NWP (*Numerical Weather Prediction*) são importadas pelo programa e introduzidas nas equações previamente estimadas para que sejam calculadas as previsões dos quantis que caracterizam a distribuição de probabilidade da geração eólica.

Os dados necessários para o ajuste dos modelos de regressão quantílica (treinamento) incluem o histórico da geração eólica verificada em bases horárias e o histórico da velocidade/direção do vento verificada em bases horárias (Figura 5a). Adicionalmente, no ajuste dos modelos, pode-se considerar o histórico das previsões anemométricos no lugar do histórico das medições anemométricas.

A especificação dos quantis a serem modelados e a escolha do modo de execução (treinamento ou previsão) é efetuada na janela ilustrada na Figura 5b, na qual foram selecionados apenas três quantis: 25%, 50% e 75%.

Os modelos de regressão quantílica são ajustados separadamente (Figura 6a) e ao final apresentam-se relatórios que permitem avaliar a qualidade do ajuste dos modelos e visualizar as estimativas dos coeficientes das equações de regressão (Figura 6b).

Os dados requisitados pelo programa VENTOS para a previsão probabilística da geração eólica incluem as previsões da velocidade do vento em bases horárias, acompanhadas das respectivas previsões da direção do vento (Figura 7a). As previsões dos quantis de geração eólica com resolução semi-horária são apresentadas na forma gráfica e em tabela, conforme ilustrado na Figura 7b.



**(a) dados para o ajuste dos modelos de regressão**

Ordem	Direção Verificada	Velocidade Verificada	Geração	Direção Prevista	Velocidade Prevista
1	200.0000	9.5300	10249.0200	207.6853	11.1491
2	205.0000	10.4800	10965.3600	207.8301	12.0896
3	195.0000	10.6600	12298.2700	209.5942	13.2891
4	192.0000	10.9700	12554.9800	210.9193	14.5889
5	195.0000	11.5300	13190.8800	219.3481	15.0840
6	199.0000	11.5500	12185.3000	219.9942	14.9930
7	200.0000	11.9300	13382.8700	220.7802	15.6115
8	204.0000	14.7400	11842.1500	223.1104	15.7368
9	205.0000	16.0500	12101.2300	217.5838	17.0528
10	213.0000	16.8600	14824.7900	213.3092	18.1149
11	212.0000	18.3500	13230.7000	209.7891	18.4572

**(b) configuração do tipo de estudo**

Parque Eólico: **Dados Gerais**

Número da Barra: 45  
 Data/Hora do Início da Previsão: Qua 11/01/2017  
 Tipo de Estudo: Ajuste/Treinamento do Modelo  
 Método: Regressão Quantílica  
 Dados para Treinamento (n-1): 50  
 Tipo de Execução: Treinamento/Teste com velocidades previstas  
 Horizonte: 24  
 Número de Iterações: 5.000  
 Não Soltes: 7  
 Número de Quantis: 3

Seleção dos quantis a serem previstos: 75.0 %, 50.0 %, 25.0 %

(a) dados para o ajuste dos modelos de previsão (b) configuração do tipo de estudo  
Figura 5 - Especificação dos modelos de regressão quantílica

**(a) execução do treinamento**

```

V V DRRRRR 300 8 TTTTTT XXXX SSSSS
V V SSS 8 8 8 T 0 0 0 SSSSS
V V SSSSS 8 8 8 T 0 0 0 S
V SSSSSSS 8 8 8 T 00000 SSSSSSS
V SSSSSSS 8 8 8 T 00000 SSSSSSS

Loading general data
Loading wind power data from dataset.txt
Forecasting method: Quantile Regression Model
Building B-splines basis
Fitting quantile regression model, quantile 0.500000
Fitting quantile regression model, quantile 0.750000
Fitting quantile regression model, quantile 0.250000
Writing file with the regression coefficients estimator
Computing the wind power quantiles - example
Writing file with the estimated quantiles
Computing the wind power quantiles - example
Writing file with the forecasted quantiles (out-sample)
Cubic Spline Interpolation
Program successfully finished.
    
```

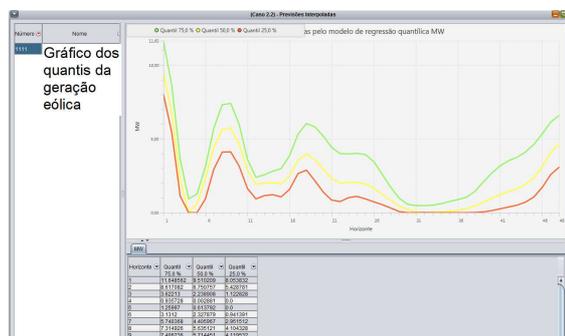
**(b) estimativas dos coeficientes de regressão**

Nome do Arquivo	Quantil 75.0 %	Quantil 50.0 %	Quantil 25.0 %
coeficientes.M	0.799880	2.387725	2.296528
estruturas.M	-0.125448	4.152425	2.394487
performance_exemplo.M	1.968652	-1.968652	0.202628
performance_subamostra.M	1.968652	-1.968652	0.202628
previsoes_exemplo_horarios.M	0.202628	-0.202628	0.202628
previsoes_subamostra_horarios.M	0.202628	-0.202628	0.202628
previsoes_subamostra_interpoladas.M	0.202628	-0.202628	0.202628

(a) execução do treinamento (b) estimativas dos coeficientes de regressão  
Figura 6 - Ajuste dos modelos de regressão quantílica

**(a) dados anemométricos para a previsão**

Data/Hora	Velocidade	Ângulo
20/01/2017 01:00	6.570	47
20/01/2017 02:00	7.400	41
20/01/2017 03:00	6.780	31
20/01/2017 04:00	5.970	26
20/01/2017 05:00	5.270	22
20/01/2017 06:00	4.550	17
20/01/2017 07:00	3.850	10
20/01/2017 08:00	3.720	347
20/01/2017 09:00	3.880	323
20/01/2017 10:00	3.430	309
20/01/2017 11:00	5.050	284



(a) dados anemométricos para a previsão (b) previsão dos quantis da geração eólica  
Figura 7 - Modo previsão

## 5. Experimento computacional

Para ilustrar a aplicação da metodologia de previsão e as funcionalidades do programa VENTOS concebeu-se um experimento computacional com dados provenientes de um parque eólico localizado na Galícia, Espanha. Os dados disponibilizados abrangem o período de janeiro de 2014 até janeiro de 2015 e incluem as médias horárias previstas e verificadas da velocidade/direção do vento e a série histórica da geração eólica. A capacidade instalada do parque é da ordem de 17,56 MW distribuída em 24 aerogeradores de 5 tecnologias e 9 fabricantes diferentes. A diversidade de tecnologias e o terreno complexo do parque [Martí et al. 2006],[Pinson 2006] explicam a grande dispersão da nuvem de observações ilustrada na Figura 3.

Na Tabela 1 apresentam-se as principais estatísticas das médias horárias verificadas da velocidade do vento e da geração. As estatísticas revelam um baixo fator de capacidade do parque eólico (21%) e uma grande variabilidade da geração (coeficiente de variação igual a 1,14).



Tabela 1 - Estatísticas da velocidade do vento e da geração eólica

Estatística	Velocidade m/s	Geração MWh/h
Máximo	24,6	16,370
3º quartil	9,2	5,556
Mediana	6,3	1,800
1º quartil	4,4	0,158
Mínimo	0	0
Média	7,1	3,427
Desvio padrão	3,8	3,920

Para caracterizar a densidade de probabilidade da geração eólica em cada hora do horizonte de previsão foram selecionados os seguintes quantis: 5%, 10%, 25%, 50%, 75%, 90%, 95%, previstos por meio de sete equações de regressão quantílica ajustadas separadamente. No ajuste das equações de regressão foram consideradas as médias horárias de vento/direção e geração ao longo do período entre 1/1/2014 e 9/11/2014 (período *insample*). Já os registros do período de 10/11/2014 até 14/11/2014 compõem o período *outsample*.

Os conjuntos *insample* e *outsample* podem ser constituídos pelos valores verificados de velocidade/direção e geração eólica. Porém, a previsão da geração eólica é obtida pela aplicação das equações de regressão aos valores das previsões anemométricas, oriundas de modelos de previsão climática. Visando uma avaliação mais realista da capacidade preditiva do modelo, o programa VENTOS disponibiliza três modalidades para o ajuste das equações de regressão:

- M1 - Valores verificados da velocidade/direção do vento nos conjuntos *insample* e *outsample*.
- M2 - Valores previstos da velocidade/direção do vento nos conjuntos *insample* e *outsample*.
- M3 - Valores verificados da velocidade/direção no conjunto *insample* e valores previstos no conjunto *outsample*.

A performance do previsor foi avaliada com base nos desvios entre as médias horárias verificadas da geração eólica e as respectivas medianas previstas. Vale destacar que nos experimentos realizados foram consideradas previsões anemométricas atualizadas apenas uma vez ao dia. A raiz do erro quadrático médio (REQM) e o erro absoluto médio (MAD) nas diferentes modalidades de ajuste encontram-se na Tabela 2. Conforme esperado, a melhor performance *outsample* foi alcançada na modalidade M1, na qual consideram-se previsões anemométricas perfeitas, i.e., iguais aos valores verificados. A modalidade M3 fornece uma avaliação mais justa da performance *outsample* dos modelos ajustados com dados anemométricos verificados, pois no conjunto *outsample* consideram-se as previsões anemométricas oriundas de modelos NWP. Os resultados da modalidade M2 sugerem que ajustar os modelos de previsão com base nos valores previstos das variáveis anemométricas pode resultar em uma performance preditiva superior à alcançada pelo modelo ajustado com dados anemométricos verificados.

Tabela 2 - Estatísticas de velocidade e geração eólica

Modalidade	Estatística	<i>Insample</i>	<i>Outsample</i>
M1	REQM	1,71 MW	2,09 MW
	MAD	0,97 MW	1,51 MW
M2	REQM	2,25 MW	2,60 MW
	MAD	1,49 MW	1,98 MW
M3	REQM	1,71 MW	4,18 MW
	MAD	0,97 MW	3,44 MW

O principal resultado gerado pelo programa VENTOS é uma planilha contendo as previsões horárias e semi-horárias dos quantis da geração eólica. Na Figura 8 tem-se a representação gráfica das previsões semi-horárias dos quantis da geração eólica resultantes da modalidade M2. As planilhas podem ser facilmente exportada para o MS Excel com a finalidade de construir outros gráficos, conforme ilustrado na Figura 9.

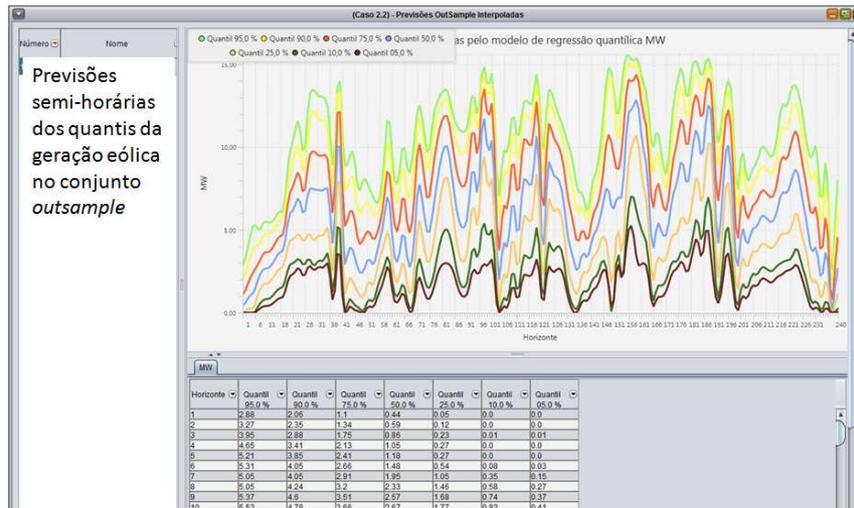


Figura 8 – Previsões semi-horárias dos quantis da geração eólica no período *outsample*

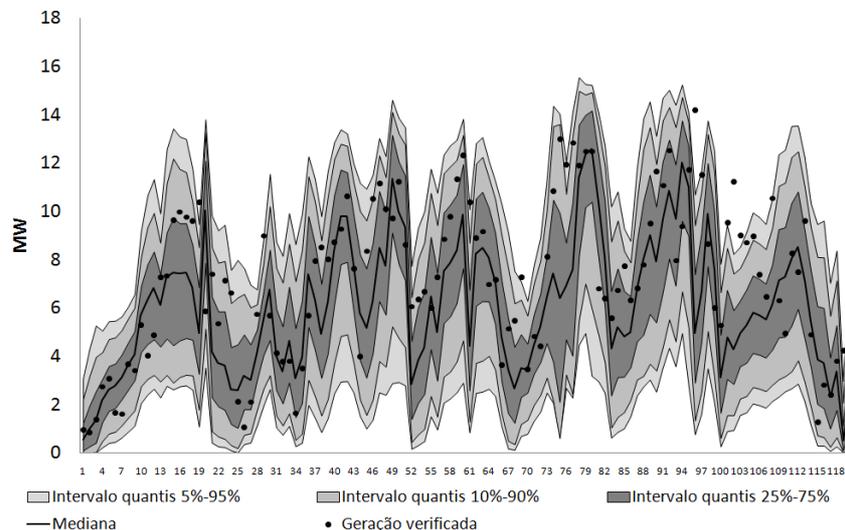


Figura 9 – Previsões horárias dos quantis da geração eólica no período *outsample*

## 6. Conclusões

A perspectiva de maior relevância da energia eólica nos sistemas elétricos aponta para a necessidade da operação contar com modelos capazes de fornecer previsões de curto-prazo da geração eólica. Para lidar com a variabilidade do vento e a incerteza nas previsões anemométricas, a previsão da geração eólica deve seguir uma abordagem probabilística, i.e., o predictor deve fornecer previsões da distribuição de probabilidade da geração eólica e não apenas informar previsões pontuais do seu valor esperado. Neste trabalho apresenta-se uma metodologia de previsão probabilística baseada em modelos aditivos generalizados para previsão dos quantis da geração eólica. A metodologia foi implementada no programa VENTOS e os resultados oriundos de experimentos computacionais mostram o potencial da abordagem adotada. Apesar dos resultados satisfatórios, mais investigações devem ser efetuadas visando o aprimoramento do modelo de previsão e da apresentação dos resultados aos usuários.

## Referências

Alcântara, A.A.M., Sant'Anna, A.P., Lins, M.P.E. (2003). Restringendo flexibilidade de pesos em DEA utilizando análise de regressão MSEA. *Pesquisa Operacional*. Agosto, vol. 23, 2, 347-357.



Anastasiades, G., McSharry, P. (2013). Quantile Forecasting of Wind Power Using Variability Indices. *Energies*, vol. 6, 662-695.

Bessa, R.J., Mendes, J., Miranda, V., Botterud, A., Wang, J., Zhou, Z. (2011) Quantile-copula density forecast for wind power uncertainty modeling, *IEEE Trondheim PowerTech*, Norway, Jun.

Chen, N., Qian, Z., Nabney, I.T., Meng, X. (2014) Wind Power Forecasts Using Gaussian Processes and Numerical Weather Prediction, *IEEE Transactions on Power Systems*, v. 29, n. 2, March, pp. 656-665.

Custódio, R.S. (2009). Energia eólica para produção de energia elétrica. Energia eólica para produção de energia elétrica. Rio de Janeiro: Eletrobras.

Dubinsky, A., Elperin, T. (1997). A method for calculating a load curve using average values of load over time intervals. *Electrical Power & Energy Systems*. vol. 19, 6, 393-395.

Estanqueiro, A., Castro, R., Flores, P., Ricardo, J., Pinto, M., Rodrigues, R., Peças Lopes, J. (2008). How to prepare a power system for 15% wind energy penetration: the Portuguese Case Study. *Wind Energy*. v. 11, 75-84.

Gilat, A., Subramaniam, V. (2008) Métodos numéricos para engenheiros e cientistas: uma introdução com aplicações usando Matlab, Bookman.

Kutner, M.H., Nachtsheim, C.J., Neter, J. (2004) Applied Linear Regression Models. McGraw-Hill Irwin.

Marti, I., Kariniotakis, G., Pinson, P., Sanchez, I., Nielsen, T.S., Madsen, H., Giebel, G., Usaola, J., Palomares, A. M., Brownsword, R., Tambke, J., Focken, U., Lange, M., Sideratos, G., Descombes, G. (2006) Evaluation of Advanced Wind Power Forecasting Models – Results of the Anemos Project, European Wind Energy Conference, Athènes, Greece, February.

Matos, M.A., Bessa, R.J. (2011). Setting the Operating Reserve Using Probabilistic Wind Power Forecasts. *IEEE Transactions on Power Systems*. 2, May, vol. 26, 594-602.

Nielsen, H.A., Madsen H., Nielsen, T.S. (2006). Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts. *Wind Energy*. April, vol. 9, 95-108.

Pewsey, A., Neuhauser, M., Ruxton, G.D. (2013) Circular Statistics in R, Oxford.

Pinson, P. (2006) Estimation of the uncertainty in wind power forecasting, These, Ecole des Mines de Paris.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B. P. (1992). Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press.

Wood, S.N. (2005). Generalized Additive Models: an introduction with R. Chapman & Hall.

Zhang, Y., Wang, J., Wang, X. (2014). Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*. vol. 32, 255-370.