



PREVISÃO PROBABILÍSTICA DE CURTO PRAZO DA GERAÇÃO DE ENERGIA EÓLICA

Soraida Aguilar Vargas

Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio
saguilar@ele.puc-rio.br

José Francisco Pessanha

Universidade do Estado do Rio de Janeiro - UERJ
professorjfm@hotmail.com

RESUMO

A entrada em operação de fontes eólicas em sistemas elétricos de potência cria uma série de desafios para a operação destes sistemas. Neste contexto, a previsão da velocidade do vento nos aproveitamentos eólicos passa a ser uma informação fundamental para os modelos de apoio à decisão que suportam a operação econômica e segura dos sistemas elétricos. Dada a incerteza associada ao vento, os métodos para previsão probabilística têm se mostrado uma boa opção para a previsão dos recursos eólicos, pois fornecem previsões pontuais e são capazes de fornecer previsões da função densidade de probabilidade da produção eólica. Neste trabalho apresenta-se um método não paramétrico para calcular a função densidade de probabilidade da produção eólica em qualquer horizonte de tempo, condicionada à previsão da velocidade do vento obtida por um modelo paramétrico (ARIMA).

PALAVRAS-CHAVE. Energia eólica, previsão probabilística, estimação da densidade de núcleo.

ABSTRACT

The entry into operation of wind sources in electric power systems creates a number of challenges for the operation of these systems. In this context, the forecast wind speed in wind farms becomes a crucial information for decision support models that support the economic and safe operation of electrical systems. Given the uncertainty associated with wind, methods for probabilistic forecasts have proved a good choice for the prediction of wind resources because they provide point predictions and are able to provide estimates of the probability density function of the wind power production. This paper presents a nonparametric method to calculate the probability density function of wind power production at any time horizon, conditional on the forecasting of wind speed obtained by a parametric model (ARIMA).

KEYWORDS. Wind Energy, probabilistic forecasting, density kernel estimation.

1. Introdução

A produção de energia eólica está crescendo rapidamente, e devido a isto, as previsões desse recurso renovável tornam-se essenciais para a operação econômica e segura de um sistema elétrico de potência.

A incerteza associada ao vento tem sido tratada através de várias técnicas estatísticas, entre elas encontram-se algumas ferramentas de previsão para diferentes horizontes de tempo. A literatura técnica reporta que a previsão pontual é a principal abordagem empregada na previsão de curto prazo da velocidade do vento para geração de energia eólica (Zhu e Genton, 2012). No entanto, este tipo de previsão tem uma deficiência importante, pois as informações sobre qualquer desvio em relação aos valores previstos não são fornecidas, o que limita o seu uso em processos de tomada de decisão. Por esta razão, a previsão probabilística tornou-se uma opção muito mais atraente, dado que a função densidade de probabilidade (*probability density function* - pdf) proporciona a flexibilidade necessária para o tratamento de diversos problemas de tomada de decisão e permite a incorporação da incerteza. Por exemplo, a previsão da densidade de probabilidade dos recursos eólicos permite que os operadores dimensionem a reserva de potência necessária para operação segura do sistema.

A grande maioria das publicações recentes dos modelos para previsão da geração eólica é feita em dois estágios: no primeiro estágio ajusta-se um modelo (seja ele estatístico, de inteligência computacional ou mesmo híbrido) para prever a velocidade do vento incidente nos aerogeradores, e em seguida, utilizando a curva de potência fornecida pelo fabricante da turbina, obtém-se a previsão da produção de energia (Madsen 1995; Nielsen e Madsen, 1996). Monteiro et al., 2009 apresentam uma descrição dos algoritmos físicos e estatísticos mais importantes para a previsão da velocidade do vento. Os modelos baseados na análise das séries históricas da velocidade do vento normalmente usam um enfoque estatístico para prever a velocidade, e posteriormente prever a produção de energia eólica.

Por outro lado, a regressão quantílica, a estimação da densidade de núcleo, a previsão probabilística baseada em conjuntos e a inteligência artificial são modelos não-paramétricos disponíveis na atualidade para a previsão probabilística da geração de energia eólica, os quais permitem especificar quantis ou intervalos (Zhang et al., 2014). A estimação da densidade condicional de núcleo fornece uma ferramenta para construir modelos como o que são mostrados em Juban et al. (2007) e Bessa et al. (2011), os quais têm como saída a função de densidade de probabilidade prevista para a produção eólica. Este resultado pode ser transformado em várias representações de incerteza, tais como quantis, desvio padrão, ou assimetria. Esta abordagem, portanto, fornece uma representação completa e robusta da previsão probabilística de energia eólica.

A proposta deste trabalho está focada na previsão da função densidade de probabilidade da produção eólica, dada uma previsão da velocidade do vento para diferentes horizontes de previsão. O trabalho está organizado em cinco seções. A seguir, na seção 2 apresenta-se uma breve descrição da estimação da densidade de núcleo. Na sequência, na seção 3 tem-se uma breve introdução aos modelos ARIMA e na seção 4 são descritos os dados do experimento computacional, cujos resultados são descritos na seção 5. Por fim, na seção 6 são apresentadas as principais conclusões.

2. Metodologia de Previsão Densidade de Núcleo

2.1. Estimação da Densidade de Núcleo

A ideia básica da estimação de núcleo é proporcionar uma estimativa da função densidade de probabilidade (pdf) de uma variável aleatória Y . Os métodos de estimação de densidade de probabilidade podem ser classificados em paramétricos e não-paramétricos. A abordagem paramétrica tem como objetivo escolher uma família de distribuição conhecida, por

exemplo, uma distribuição de Weibull ou uma distribuição Gaussiana; enquanto que a abordagem não-paramétrica consiste em ajustar uma função densidade de probabilidade aos dados, porém sem um modelo previamente definido.

O estimador de densidade de núcleo estima a densidade suavizada a partir de uma amostra de dados, colocando em cada ponto amostral uma função que representa sua contribuição à densidade. A distribuição é obtida somando todas estas contribuições.

Diferentes estimadores de núcleo têm sido usados para calcular a densidade condicional da produção eólica. Em Juban et al. (2007) empregou-se uma adaptação do estimador de densidade clássico de Nadaraya-Watson, enquanto Jeon e Taylor (2012) aplicaram o estimador de dois-passos CKD de Hyndman et al. (1996) e o estimador Parzen-Rosenblatt (Parzen, 1962; Rosenblatt, 1969), onde estimador CKD de dois-passos obteve resultados ligeiramente mais precisos do que os produzidos pelo estimador Parzen-Rosenblatt, porém não se tem argumentos sólidos teóricos que façam com que a utilização do estimador mais complexo seja melhor. Em Bessa et al. (2012), o estimador utilizado foi o Nadaraya-Watson com tempo adaptativo. Seguindo a mesma abordagem, o presente estudo centra-se no uso do estimador Nadaraya-Watson.

2.2. O Estimador Nadaraya-Watson

A densidade condicional fornece a estimativa da pdf de uma variável aleatória Y , dada uma variável explicativa X com o valor de x conhecido a ser previsto no tempo $t + k$, dada a informação disponível no momento t :

$$f_Y(Y_{t+k} | X_t) = \frac{f_{Y,X}(Y_{t+k}, X_t)}{f_X(X_t)} \quad (1)$$

A aplicação do estimador Nadaraya-Watson aos dados disponíveis oferece a possibilidade de estimar a densidade condicionada em (1), no caso das densidades $f_{Y,X}$ e f_X serem desconhecidas. Segundo Hyndman (1996), a estimativa da densidade condicional é dada por

$$\hat{f}(y | X = x) = \sum_{i=1}^N K_{h_y}(y - Y_i) \cdot w_i(x) \quad (2)$$

onde

$$w_i(x) = \frac{K_{h_x}(x - X_i)}{\sum_{i=1}^N K_{h_x}(x - X_i)} \quad (3)$$

em que N é o tamanho da amostra, $K_{h(\cdot)} = K(\cdot/h)/h$ é uma função de núcleo (*kernel*) e h o parâmetro do tamanho da janela.

Este estimador tem duas janelas, h_x e h_y , que controlam a quantidade de suavização; isto é, h_y controla a suavização de cada densidade condicional e h_x controla a suavização da variável explicativa.

Usando as equações (2) e (3), tem-se que a estimativa da densidade condicional resultante está representada na Fig. 1. Adicionalmente, estabelecendo um valor de 7 m/s e outro

de 12 m/s para a velocidade do vento, as respectivas densidade de probabilidade da produção de energia eólica condicionadas às velocidades do vento são ilustradas nas Figuras 2 e 3.

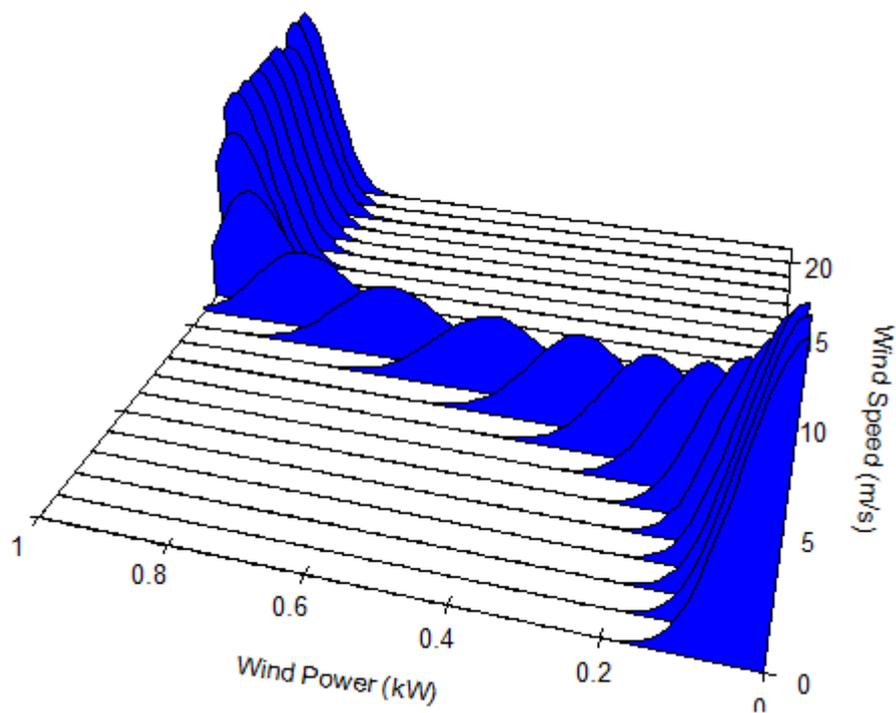


Fig. 1. Estimação da densidade condicional da produção eólica dada a velocidade do vento.

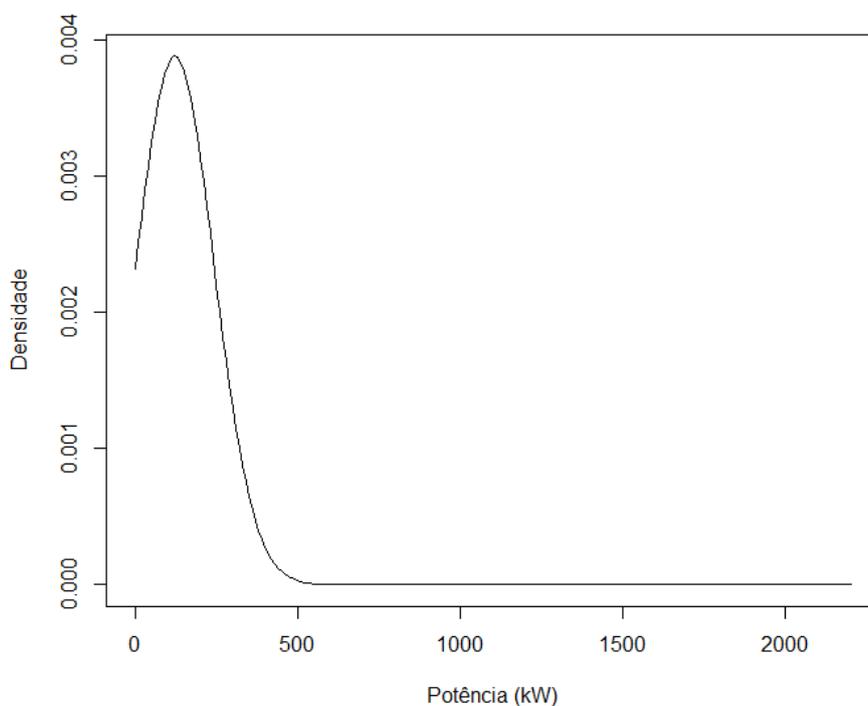


Fig. 2. Exemplo da pdf para a geração eólica dada uma velocidade do vento de 7 m/s.

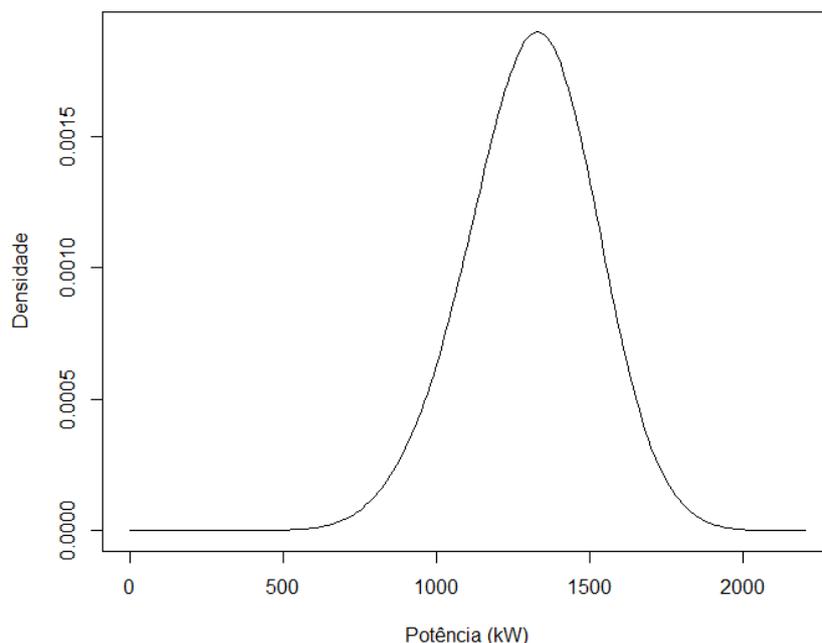


Fig. 3. Exemplo da pdf para a geração eólica dada uma velocidade do vento de 12 m/s.

3. Modelagem da Velocidade do Vento

Tradicionalmente o modelo mais usado para previsão tem sido o modelo ARIMA de Box e Jenkins (1970). Embora as séries temporais da velocidade do vento e da geração de energia eólica apresentem uma dinâmica altamente não-linear, os trabalhos de Blanchard e Desrochers (1984), Balouktsis et al. (1986), Daniel e Chen (1991), Tantareanu (1992), Huang e Chabali (1995), Kamal e Jafri (1997), Schwartz e Milligan (2002), Kennedy e Rogers (2003), Poggi et al. (2003), Torres et al. (2005), Philippoulos e Deligiorgi (2009) e Erdem e Shi (2011) têm optado por esta abordagem linear.

A velocidade do vento apresenta padrões sazonais e uma alta variabilidade, por isso neste trabalho é usado um modelo ARIMA multiplicativo para dados sazonais, conhecido como modelo SARIMA, apesar da série temporal apresentar uma alta não-linearidade. Uma vez ajustado o modelo, este é utilizado para prever um-passo-à-frente a velocidade do vento. Na sequência, tal previsão será utilizada na estimativa da densidade condicional para prever a densidade de probabilidade da produção eólica.

3.1 Modelo Multiplicativo ARIMA

O modelo ARIMA (*AutoRegressive Integrate Moving Average*) pertence à família de modelos proposto por Box e Jenkins (1970), o qual é o resultado da procura de um modelo que represente o processo estocástico gerador da série temporal. Nestes modelos a série temporal y_t é explicada pelos seus valores passados defasados e o termo do erro estocástico. O modelo ARMA é um submodelo do modelo ARIMA, uma vez que só é aplicável a séries estacionárias, isto é, a série não possui raízes unitárias; enquanto que o modelo ARIMA permite a modelagem de séries que apresentam uma média não-estacionária. A maioria das séries de temporais, especialmente as séries físicas tais como as séries hidrológicas são não-estacionárias, o que quer dizer que a média (e as vezes a variância) da série muda ao longo do tempo. A diferenciação muitas das vezes é usada para remover o componente de tendência de uma série temporal antes da construção do modelo para descrever a série. No entanto, neste trabalho a série da velocidade do vento já é estacionária, e por esta razão o modelo ajustado foi um AR na parte não sazonal.

Geralmente, um processo multiplicativo SARIMA (Box e Jenkins, 1976) pode ser representado pela equação:

$$\phi_p(B)\Phi_p(B^L)(1-B)^d y_t = c + \theta_q(B)\Theta(B^L)\varepsilon_t \quad (4)$$

onde $\phi_p(B)$ e $\theta_q(B)$ são os parâmetros não-sazonais do processo Autorregressivo e de Médias Móveis de ordem p e q , respectivamente. $\Phi_p(B)$ e $\Theta_q(B)$ são os parâmetros sazonais do processo Autorregressivo e de Médias Móveis de ordem P e Q , respectivamente. B é um operador de defasagem e ε_t corresponde ao termo de erro de eventos aleatórios que não podem ser explicados pelo modelo; c é uma constante, e y_t é a séries de temporal.

4. Dados de potência e de velocidade do vento

A base de dados utilizada neste trabalho pertence a um parque eólico. O conjunto de dados contém registros horários durante o período de 1 de janeiro de 2007 até 1 de Janeiro de 2008; i.e. um total de 8,784 médias horárias da velocidade do vento (m/s) e da potência (kW) para uma turbina eólica com potência nominal de 2.200 kW. Desta amostra, 8760 observações são utilizadas na modelagem (in-sample) e as 24 restantes na validação do modelo (out-of-sample). Na Figura 4 mostra-se o diagrama de dispersão da velocidade média horária contra a potência média horária para o período descrito acima, gerando assim a curva de potência da base de dados históricos.

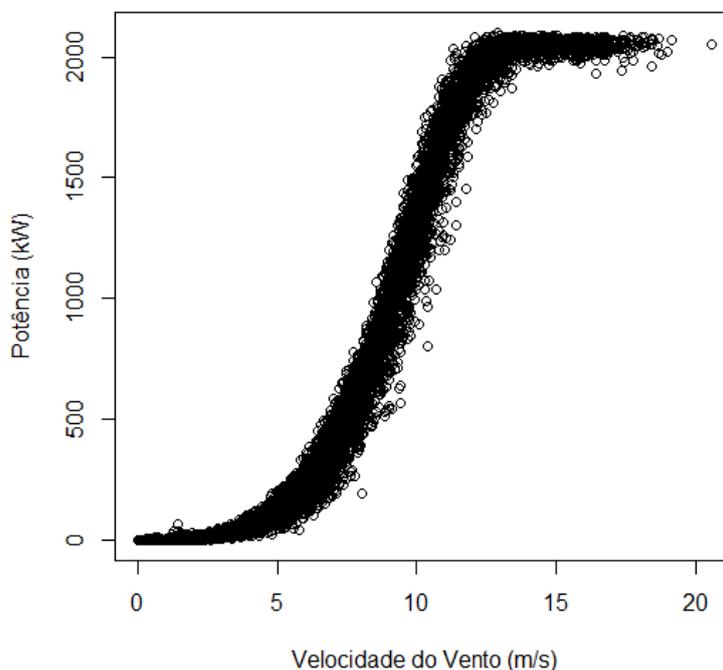


Fig. 4. Curva de Potência.

5. Resultados

Embora o modelo SARIMA seja utilizado para prever a velocidade do vento, este trabalho se concentra na previsão da densidade de probabilidade da produção de energia eólica. A previsão da velocidade é um passo intermediário, isto é, ela é necessária a fim de obter o valor futuro da velocidade do vento para ser usado no cálculo da estimação da densidade condicional da geração eólica, e assim fornecer a previsão probabilística da produção eólica. Na Figura 5 é

ilustrada a função de autorrelação (FAC) e autocorrelação parcial (FACP), a qual indica que esta série temporal possui um padrão sazonal.

Por outra parte, o ajuste do modelo é apresentado na Tabela 1. No processo de estimação destaca-se que várias configurações de modelos foram testadas previamente, e escolheu-se o modelo que minimiza os critérios de informação de Akaike, Hannan-Quinn e de Schwarz. O melhor modelo foi o SARIMA(2,0,0)x(1,1,1)₂₄. Uma vez estimado o modelo, este foi utilizado para prever a velocidade do vento 24 horas à frente, i. e. uma previsão horária para o dia 01 de Janeiro de 2008, os resultados desta previsão apresentam-se na Figura 6.

Tabela 1. Modelo ajustado SARIMA(2, 0, 0)x(1,1,1)₂₄

	coeficiente	erro padrão	z	p-valor	
ϕ_1	0,799451	0,0108263	73,84	0,0000	***
ϕ_2	-0,0256117	0,0106182	-2,412	0,0159	**
Φ_1	0,0679778	0,0116126	5,854	4,80e-09	***
Θ_1	-0,941079	0,00428697	-219,5	0,0000	***
Loq da verossimilhança	-16706,96	Critério de Akaike	33423,92		
Critério de Schwarz	33459,30	Critério Hannan-Quinn	33435,97		

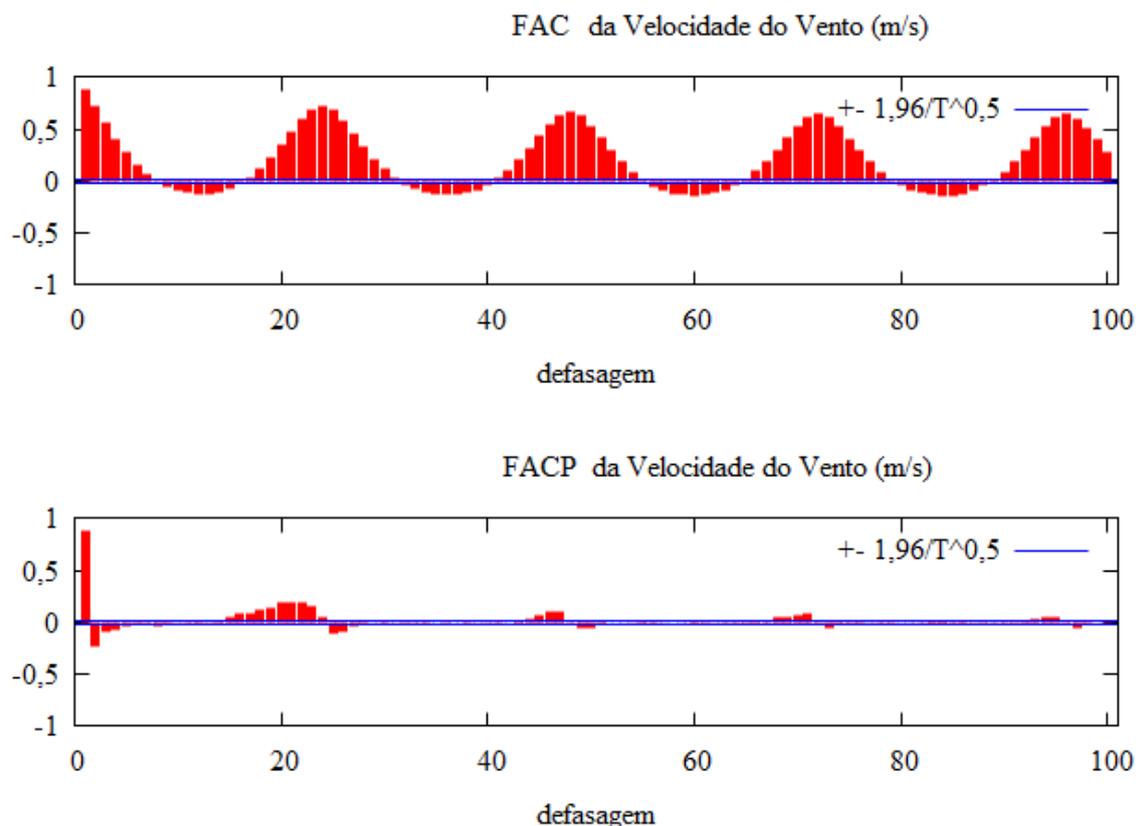


Fig. 5. Função de Autocorrelação (FAC) e função de autocorrelação parcial (FACP) da série temporal da velocidade do vento.

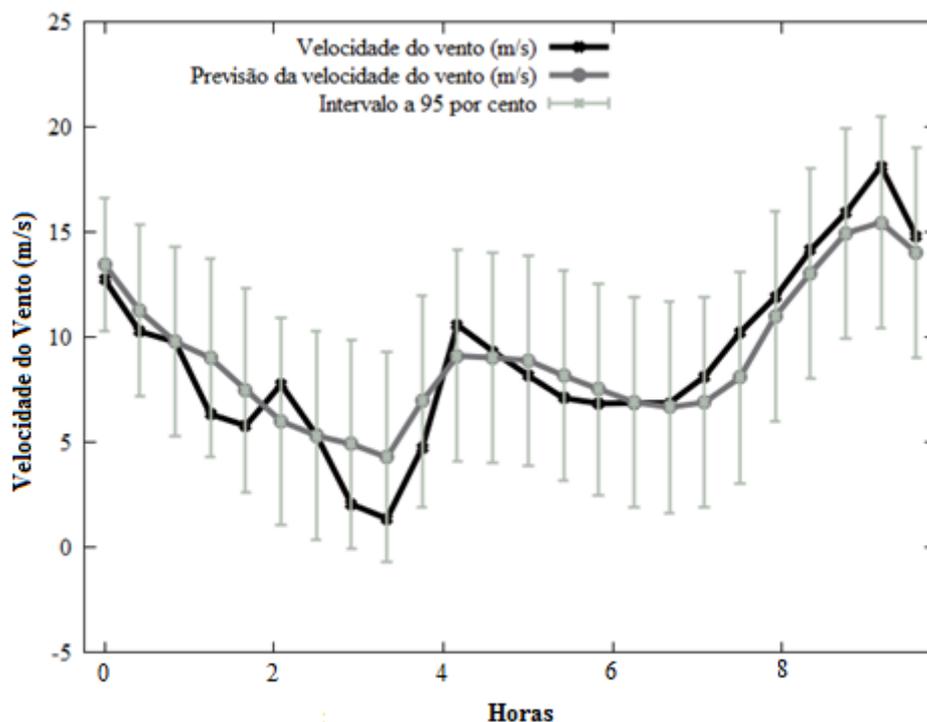


Fig. 6. Previsão da velocidade do vento (m/s) 24 horas à frente.

Para avaliar o modelo foram calculadas medidas de erro dentro da amostra (in-sample), ou seja, no ajuste do modelo e fora da amostra (out-of-sample) quando é feito o processo de previsão. As estatísticas MSE, MAD e U-Theil apresentadas na Tabela 2 indicam a boa qualidade do modelo ajustado à série temporal de velocidade do vento.

Tabela 2. Medidas de erro para a velocidade do vento

	In-Sample	Out-of-Sample
RMSE (m/s)	1,6397	1,5629
MAD (m/s)	1,2358	1,2693
U-THEIL	0,36939	0,6933

As velocidades do vento previstas são então utilizadas para gerar a previsão probabilística da produção eólica. A previsão da velocidade do vento 24 horas à frente proporciona a previsão da densidade de probabilidade condicional da produção eólica ao longo de cada hora do dia.

Na estimação da densidade condicional de núcleo empregou-se o núcleo gaussiano e criou-se uma grade com valores da potência média horária que variam desde 0 até a capacidade da turbina (2.200 kW) com intervalos de 10 kW e uma grade com valores da velocidade do vento que vão de 0 até 21 m/s com incrementos de 0,1 m/s. O resultado obtido é formado por 221 valores diferentes para a potência média gerada e 212 valores da velocidade do vento, constituindo assim uma matriz de 212x221 a qual possui a forma da Figura 1. Simulações adicionais com a discretização da potência média em intervalos menores foram realizadas, porém a maior resolução resultou apenas no aumento na carga computacional sem uma apreciável melhoria.

Uma vez obtidas as densidades de probabilidade da produção eólica, dadas as velocidades previstas com o modelo SARIMA, calculou-se a previsão pontual como sendo a média e a mediana de cada densidade condicional. Na Tabela 3 são apresentadas as medidas de erro segundo cada previsão pontual da geração de energia eólica. A Figura 7 apresenta as previsões pontuais 24 horas à frente da produção eólica definidas pelas médias das distribuições condicionais.

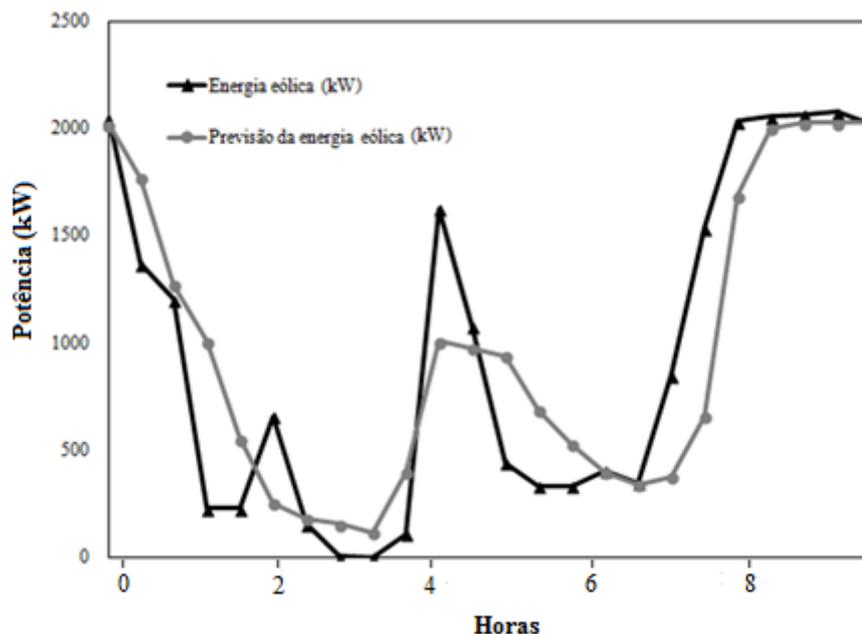


Fig. 7. Previsão da potência média horária (kW) gerada pela turbina eólica 24 horas à frente.

Tabela 3. Medidas de erro para a produção eólica (potência média horária)

	Previsão Pontual - Média	Previsão Pontual - Mediana
RMSE (kW)	359,0157	357,7933
MAD (kW)	259,0089	256,2467
U-THEIL	0,7226	0,7202

Os valores resultantes da medida U-Theil indicam que as previsões pontuais da produção eólica determinadas pela média ou mediana das densidades condicionais são melhores que a previsão ingênua. O RMSE e o MAD apresentam valores altos como consequência de duas situações: a primeira é o erro na previsão da velocidade do vento e a segunda reside na incapacidade do estimador de núcleo conseguir reproduzir fielmente a densidade condicional, conforme evidenciado na Figura 8, onde se relaciona a curva de potência real com a média condicional calculada acima para obter as previsões pontuais, na qual se observa claramente que a média condicional da potência média para pequenos valores de velocidade do vento assume valores superiores aos valores observados. Aqui é de grande importância sublinhar que os erros na previsão da velocidade são amplificados na previsão da geração de energia eólica, dado que a potência gerada é proporcional ao cubo da velocidade.

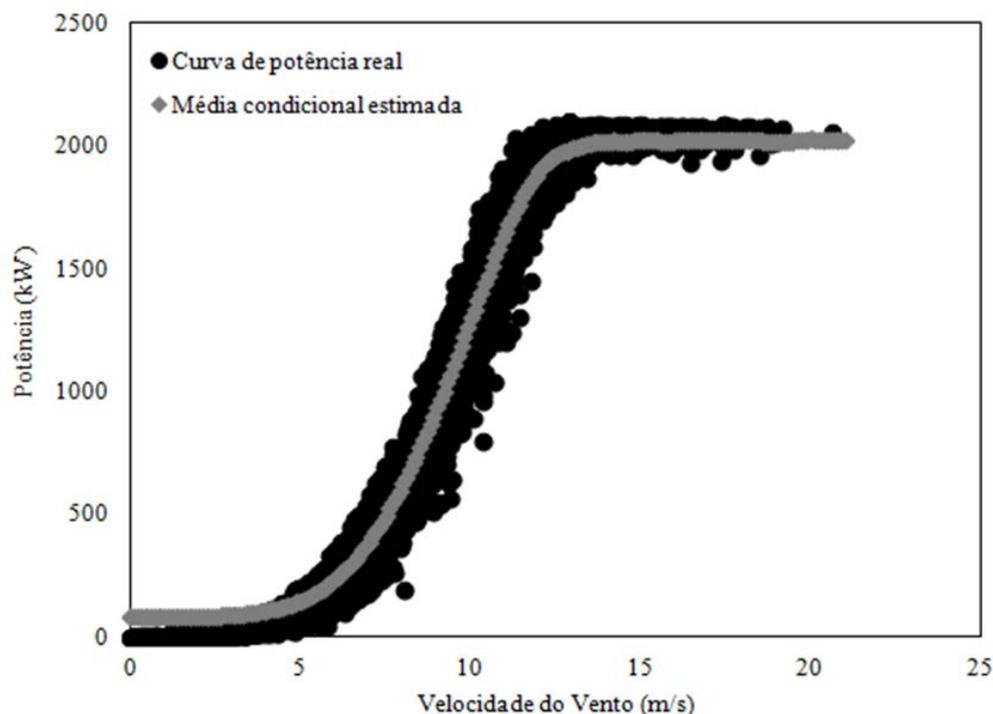


Fig. 8. Curva de potência obtida pelo estimador Nadaraya-Watson.

De forma similar, na Figura 9 são apresentados os diagramas de dispersão dos valores reais contra os previstos, tanto da velocidade do vento quanto da potência média horária gerada, neles pode-se observar a boa qualidade do ajuste. No entanto, em termos de previsão pontual, as dispersões observadas no diagrama da geração eólica é maior que aquela obtida com a previsão pontual da velocidade do vento. De certa forma, isto corrobora as estatísticas de ajuste mostradas acima, em particular, o U-Theil que foi de 0.39 para a velocidade e 0.72 para a geração.

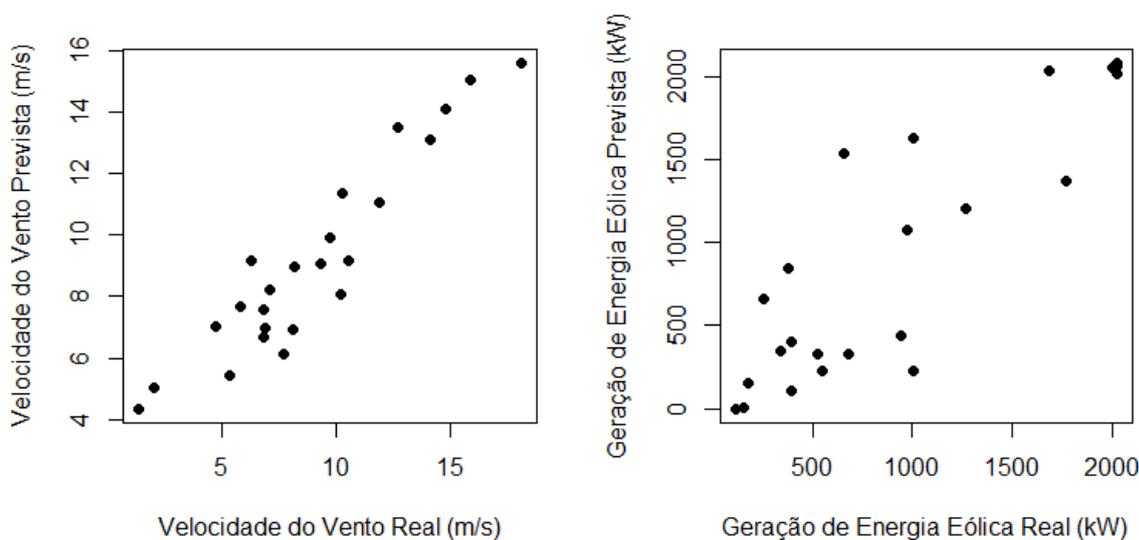


Fig. 9. Diagrama de dispersão da velocidade do vento e da potência média horária

6. Conclusões

Neste trabalho foi introduzida uma abordagem para prever a densidade de probabilidade da potência média horária gerada por uma turbina eólica, a qual captura a incerteza associada à natureza estocástica da curva de potência e que é modelada através do uso da estimação da densidade condicional de núcleo.

Os resultados obtidos mostram que o processo de previsão probabilística da geração eólica é consistente com o cálculo da densidade condicional da potência dada a velocidade do vento, pois esta densidade tem um bom ajuste se comparado com o gráfico da curva de potência da base de dados, e com as medidas de erro da previsão que foram razoavelmente boas; embora o modelo de previsão da velocidade do vento tenha sido um modelo SARIMA multiplicativo. A natureza intrínseca altamente não-linear da série da velocidade do vento é uma indicação de que outros métodos devem ser utilizados, além de testar outros tipos de núcleo diferentes ao gaussiano, por exemplo, o núcleo de Epanechnikov ou Biweight.

Em trabalhos futuros recomenda-se investigar outros modelos que levem em consideração as características da série da velocidade do vento, assim como o uso de outras variáveis físicas que possam permitir uma melhor representação do fenômeno, por exemplo, a temperatura e a direção do vento.

Agradecimentos

Os autores gostariam de agradecer à CAPES PEC-PG pelo apoio financeiro.

References

- Balouktsis, A., Tsanakas, D., e Vachtsevanos, G. (1986). Stochastic simulation of hourly and daily average wind speed sequences. *Wind Engineering*, 10(1), 1-11.
- Bessa, R. J., Mendes, J., Miranda, V., Botterud, A., Wang, J., e Zhou, Z. (2011). Quantile-copula density forecast for wind power uncertainty modeling. In *2011 IEEE Trondheim PowerTech*, Trondheim, Norway, (pp. 1-8).
- Bessa, R. J., Miranda, V., Botterud, A., Wang, J., e Constantinescu, E. M. (2012). Time adaptive conditional kernel density estimation for wind power forecasting. *IEEE Transactions on Sustainable Energy*, 3(4), 660-669.
- Blanchard, M. e Desrochers G. (1984). Generation of autocorrelated wind speeds for wind energy conversion system studies. *Solar Energy*, 33, 571-579.
- Box, G. E. E. e Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, Cal.
- Daniel, A. e Chen, A. (1991). Stochastic simulation and forecasting of hourly wind speed sequences in Jamaica. *Solar Energy*, 46, 1-11.
- Erdem, E., e Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind speed and direction. *Applied Energy*, 88(4), 1405-1414.
- Hyndman, R. J., Bashtannyk, D. M., e Grunwald, G. K. (1996). Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics*, 5(4), 315-336.
- Huang, Z., e Chalabi, Z. S. (1995). Use of time-series analysis to model and forecast wind speed. *Journal of Wind Engineering and Industrial Aerodynamics*, 56, 311-322.
- Jeon, J., e Taylor, J. W. (2012). Using Conditional Kernel Density Estimation for Wind Power Density Forecasting. *Journal of the American Statistical Association*, 107(497), 66-79.
- Juban, J., Siebert, N., e Kariniotakis, G. N. (2007). Probabilistic Short-term Wind Power Forecasting for the Optimal Management of Wind Generation. In *2007 IEEE Lausanne Power Tech* (pp. 683-688).
- Kamal, L., e Jafri, Y. Z. (1997). Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Solar Energy*, 61(1), 23-32.

- Kennedy, S., e Rogers, P. (2003). A Probabilistic Model for Simulating Long-Term Wind-Power Output. *Wind Engineering*, 27(3), 167–181.
- Madsen, H. (1995). Wind Power Prediction Tool in Control Dispatch Centres. ELSAM, Skaerbaek, Denmark, ISBN 87-87090-25-2.
- Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., e Conzelmann G. (2009). Wind Power Forecasting: State-of-the-Art 2009. Argonne National Laboratory ANL/DIS-10-1.
- Nielsen, T.S., e Madsen, H. (1996). Using Meteorological Forecasts in On-line Predictions of Wind Power. ELSAM, Skaerbaek, Denmark.
- Parzen, E. (1962). On estimation of probability density function and mode. *Annals of Mathematical Statistics*, 33(3), 1065-1076.
- Philippopoulos, K., e Deligiorgi, D. (2009). Statistical simulation of wind speed in Athens, Greece based on Weibull and ARMA models. *International Journal of Energy and Environment*, 3(4), 151-158.
- Poggi, P., Muselli, M., Notton, G., Cristofari, C., e Louche, A. (2003). Forecasting and simulating wind speed in Corsica by using an autoregressive model. *Energy Conversion and Management*, 44(20), 3177–3196.
- Rosenblatt, M. (1969). Conditional probability density and regression estimates. *Multivariate Analysis II*, Ed. P.R. Krishnaiah, pp. 25-31. New York: Academic Press.
- Schwartz, M., e Milligan M. (2002). Statistical Wind Forecasting at the U.S. National Renewable Energy Laboratory. In *Proceedings of the First IEA Joint Action Symposium on Wind Forecasting Techniques*, Norrköping, Sweden, (pp. 115-124B). Published by FOI - Swedish Defence Research Agency.
- Tantareanu, C. (1992). Wind Prediction in Short Term: A first step for a better wind turbine control. Nordvestjysk Folkecenter for Vedvarende Energi, ISBN 87-7778-005-1.
- Torres, J. L., García, A., De Blas, M., e De Francisco, A. (2005). Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy*, 79(1), 65–77.
- Zhu, X., e Genton, M. G. (2012). Short-Term Wind Speed Forecasting for Power System Operations. *International Statistical Review*, 80(1), 2–23.
- Zhang, Y., Wang, J., e Wang, X. (2014). Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32, 255–270.