

Hybrid methodology for modeling short-term wind power generation using conditional Kernel density estimation and singular spectrum analysis

Soraida Aguilar-Vargas ^a, Reinaldo Castro-Souza ^a, José Francisco Pessanha ^b
& Fernando Luiz Cyrino-Oliveira ^c

^a Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil. saguilar@ele.puc-rio.br, reinaldo@ele.puc-rio.br

^b Institute of Mathematics and Statistics, Rio de Janeiro State University, Rio de Janeiro, Brazil. professorjfm@hotmail.com

^c Department of Industrial Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil. cyrino@puc-rio.br

Received: August 11th, 2016. Received in revised form: December 10th, 2016. Accepted: January 19th, 2017

Abstract

A fundamental part of the probabilistic forecasting of wind energy process is to take into account wind speed forecasts. To achieve accurate probabilistic forecast of wind output, it is developed a hybrid methodology using a nonparametric techniques known as SSA (Singular Spectrum Analysis) and (CKDE) Conditional Kernel Density Estimation. SSA is employed to forecast wind speed and CKDE to obtain probabilistic forecasts of wind energy, based on the fact that wind power generation has a nonlinear relation with the wind speed and both are random variables distributed according to a joint density function. A Brazilian hourly wind dataset including wind speed and wind power is used to illustrate the approach. Once the wind speed forecasts are obtained the corresponding probabilistic forecast of the wind power generation is estimated for a lead time of 24 hours ahead. The results obtained are compared with other existing methodologies.

Keywords: Wind power generation; SSA; CKDE; time series; forecasting.

Metodología híbrida para el modelado de la generación de energía eólica de corto plazo mediante la estimación de la densidad condicional por Kernel y análisis espectral singular

Resumen

Una parte fundamental del proceso de previsión probabilística de energía eólica es tener en cuenta las previsiones de la velocidad del viento. Para obtener pronósticos probabilísticos precisos de la producción eólica ha sido desarrollada una metodología híbrida utilizando técnicas no paramétricas conocidas como SSA (Análisis Singular Espectral) y Estimación Condicional de la Densidad por Kernel (CKDE). SSA es empleada para predecir la velocidad del viento y CKDE para obtener previsiones probabilísticas de la energía eólica, dado que la generación de la energía eólica tiene una relación no lineal con la velocidad del viento y ambas son variables aleatorias distribuidas que siguen una función de densidad conjunta. Haciendo uso de una base de datos brasilera horaria que incluye la velocidad del viento y la energía eólica es ilustrada dicha metodología. Una vez que las previsiones de la velocidad del viento son obtenidas, los correspondientes pronósticos probabilísticos de la generación de energía eólica son estimados para un horizonte de 24 horas. Los resultados obtenidos son comparados con otras metodologías existentes.

Palabras clave: Generación de energía eólica; SSA; estimación condicional de la densidad por kernel; series temporales; previsión.

1. Introduction

Wind energy is the fastest growing source of power generation in Brazil as a result of the implementation of The

Brazilian government's Decennial Energy Plan – PDE 2024 [1], which indicates that the participation of wind energy in electricity generation matrix is expected to reach nearly 12% by 2024 with the expansion of at least 24 GW of installed

How to cite: Aguilar-Vargas, S., Castro-Souza, R., Pessanha, J. F., and Cyrino-Oliveira, F. L., Hybrid methodology for modeling short-term wind power generation using conditional Kernel density estimation and singular spectrum analysis. *DYNA* 84(201), pp. 145-154, 2017.

wind farms. For this, the Brazilian wind industry should try to keep a steady 2 GW annual growth, as well as to guarantee a competitive wind energy price at the auctions. This way the industry, as a whole, should attract new investments.

Forecasts of wind energy depend mainly on the wind speed, and given its stochastic nature, in order to include the wind speed into the optimal dispatch, reliable and accurate predictions are required. Therefore, the integration of the wind power into the Brazilian system is a rather difficult task. For operational purposes this kind of forecasts are based on short-term horizons, for example, up to 24 hours ahead.

Different methods for hourly time series forecasting [2] could be employed, i.e., statistical methods (ARIMA, Kalman filter, probabilistic forecasting, etc.), computational intelligence methods (artificial neural networks, support vector machines, fuzzy logic and neuro-fuzzy systems) and hybrid methods that combine approaches based on statistical techniques and computational intelligence and other techniques.

Nonetheless, almost all of these approaches provide a single expected value for each forecast horizon, i.e., a point forecast. The technical literature reports that point forecasts, which are also deterministic forecasts, are the major approaches employed [2]. Despite of that, point forecast has a relevant shortcoming, any information about the deviation from the predicted values is not provided, and therefore, the distribution of the forecasting errors is ignored, what limit its use in decision-making processes.

Using point forecast could lead to not accurate predictions of wind power generation, making the power system unreliable. Underestimated forecasts of wind power might cause wind power curtailment or that conventional plants operate in part-load due to the system operator commit fewer wind power units than necessary, and consequently, increase generation costs. Otherwise, if the forecast of the wind power is overestimated, the system will have a power supply shortage unless sufficient spinning reserve has been committed to the power system.

The existing approaches for wind speed and wind power do not respond satisfactorily the accuracy required to be part of the optimal Brazilian hydro-thermal wind dispatch. Therefore, as an alternative to overcome this difficulty one could use probabilistic forecasting through the probability density function, instead of the point forecasts provided by the existing approaches to model the wind speed / wind energy.

Indeed, the majority of the available methods are carried out in two stages: in the first a model (be it statistical, computational intelligence or even hybrid) is fitted to predict the wind speed in order to produce hourly wind speed point forecasts. On the second stage, these forecasts are taken to the wind turbine power curve that outputs the corresponding wind energy [3,4]. See also [5] for a state-of-the-art report on the subject.

The main criticism to these approaches lies on the fact that there are other meteorological variables, besides the wind speed, that are not taken into account when using solely the power curve on the second stage to produce the wind energy forecast. Such one-to-one relationship is just a guideline. For the same wind speed there are different wind

energy possibilities as can be checked on wind farm in operation.

Also, instead of using conventional time series methods to model the wind speed, such forecast will be produced by SSA (Singular Spectrum Analysis) to extract the noise-free signal of the wind speed series [6,7] to produce more accurate wind speed forecasts.

This article proposes a hybrid methodology for modeling the wind power generation, through three stages, to obtain the full probability density function of the wind energy. For this, in the first stage using SSA (Singular Spectrum Analysis) it is obtained the wind speed forecasts. In the second stage is calculated the estimation of the density function of the wind power employing Conditional Kernel Density Estimation; and in the third stage, for a given particular wind speed forecast is calculated the density forecast function of the wind power. The remainder of this paper is organized as follows: in section 2 presents a brief literature review, and gives a description of SSA and kernel density estimation. Then, in section 3 the methodology to modelling wind speed and wind power is defined. Next, in section 4 it is shown the results based on kernel density estimators and wind speed forecast. The paper ends with some conclusions and remarks in section 5.

2. Literature review

The wind speed forecasts for wind power generation and operation planning in power systems focuses mainly on short-term forecast, due to the power system operations, such as electricity market clearing, regulation actions, power system planning for unit commitment and dispatch are held within specified periods ranging from 1 to 24 hours ahead.

Nonetheless, due to the fact that wind power generation depends on wind speed, several methods have been developed. Usually, these methods are classified into three categories: physical, statistical and hybrid approaches. The first are physical methods, which establishes a lot of considerations to predict the wind speed according to the physical description of the atmosphere. In this case, information provided by the weather service in most of the procedures transform the coarse grid of the wind speed and other climate variables to the characteristics of the terrain where is located the wind farm. The second are the statistical methods, which use approaches like ARIMA models, Artificial Neural Network (ANN) models or combination of both, to obtain the relationship of the measured power or wind speed data. The third are hybrid models, which can be considered as a combination of different approaches of physical and statistical models or combination of these models or combination of models for the short-term and for the medium-term, in order to improve overall performance forecasts.

Traditionally, the most widely used model for prediction has been ARIMA, although there had been previous attempts to predict the wind speed. The first work considering wind power forecasts was proposed by [8], in which it is fitted an autoregressive process (AR) to wind speed data transformed to make their distribution approximately Gaussian and standardized to remove diurnal nonstationarity. Despite the wind speed time series and wind energy present a highly

nonlinear dynamic, different works have opted for linear approaches, such as in [9-12].

Similarly, for wind speed prediction it was used the Kalman filter to estimate the parameters of AR and ARMA models, as indicated in [13,14]. In [15] it is also used a state-space model to be optimized with a Kalman filter to predict the wind speed over the North Atlantic Ocean. Other approaches like ARFIMA in [16] and ARFIMA-FIGARCH in [17] are employed for the same purpose. In the case of non-linear statistical models, STAR and SETAR are applied in [18].

Different Artificial Neural Network models architecture such as feed-forward neural networks (FNNs), multi-layer perceptrons (MLP), recurrent neural networks (RNNs), radial basis function (RBF) NNs, Adaline networks, have been used not only for wind speed forecast, but also for wind power generation [18-20], where its performance is improved. In [21] is used Adaptive Neuro-Fuzzy Interface System (ANFIS), to forecast wind speed and wind power through a training set that includes wind speed and direction data.

In the case of physical models, most of the physical approaches used to forecast wind speed is known as a Numerical Weather Prediction (NWP), which solves complex mathematical models, i.e., conservation equations numerically using current weather conditions like temperature, direction, pressure, surface roughness and obstacles at the given site. These methods increase the real resolution of NWP model in order to achieve accurate prediction of the weather [22]; however, they are not effective for short-term forecasting due to their computational costs, see [23].

All of these described approaches offer a point forecast, while methods like quantile regression, copulas or conditional density kernel estimation give a probabilistic forecast; specifically, different kernel estimators have been used to compute the conditional density of the wind power output. In [27] it was employed an adaptation of the classic Nadaraya-Watson kernel density, while in [28] it was applied the two-step CKD estimator of [30] and the Parzen-Rosenblatt estimator [30,31], in which the two-step CKD estimator led them to results slightly more accurate than those produced by Parzen-Rosenblatt estimator. In [32] the estimator used was Nadaraya-Watson with time adaptive. Following the same approach of those works this paper will be focused on Nadaraya-Watson estimator.

2.1. Singular Spectrum Analysis (SSA)

SSA is a non-parametric technique for analyzing and forecasting time series, which relies exclusively on data [24]. Furthermore, the SSA does not require the stationary assumption of time series, and allows the decomposition of a time series into various additive components [7].

Basically, the implementation of SSA involves three stages: decomposition, reconstruction and forecasting. These three stages are described in the following subsections.

2.1.1. Decomposition

The decomposition stage has two steps: *embedding and singular value decomposition (SVD)*.

2.1.1.1. Embedding

Embedding consists of a moving window of length L that runs through the time series $Y_T = (y_1, y_2, \dots, y_T)$, obtaining in each step a column of the trajectory matrix as eq. (1), whose total number of columns is equal to $K = T - L + 1$.

$$\mathbf{X} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ y_3 & y_4 & y_5 & \cdots & y_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{bmatrix} \quad (1)$$

The number of components L extracted from the time series is determined by the window length. The parameter L is a sufficiently large integer value, but no greater than $T/2$, i.e., $2 \leq L \leq T/2$ [6,7].

2.1.1.2. Singular value decomposition (SVD)

Using the SVD the trajectory matrix \mathbf{X} in eq. (1) can be expressed by a sum of the elementary matrices and represents a sum of rank-one bi-orthogonal elementary matrices as follow:

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_d \quad (2)$$

where each component in eq. (2) is expressed by $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$, d is the number of nonzero eigenvalues, of the matrix $\mathbf{X}\mathbf{X}^T$, denoted by $\lambda_1, \dots, \lambda_L$ in decreasing order of magnitude ($\lambda_1 \geq \dots \lambda_L \geq 0$). U_i and V_i denote the left and right eigenvectors of the trajectory matrix, and U_1, \dots, U_L is an orthonormal system, i.e., $(U_i, U_j) = 0$ for $i \neq j$ and $\|U_i\| = 1$, and corresponds to the eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$. The set $\{U_i, V_i, \lambda_i\}$ corresponds to an eigentriple $\forall i = 1, \dots, d$.

2.1.2. Reconstruction

The reconstruction stage is composed of two steps: *grouping and diagonal averaging*. In the grouping step the d elementary matrices obtained from the SVD step are grouped into mutually exclusive groups (*clusters*). Furthermore, the matrices are converted into time series through the so-called average diagonal procedure and, at the end it is obtained the additive components that make up the time series.

2.1.2.1. Grouping of the eigentriples

In this step, the elementary matrices \mathbf{X}_i are split into several groups and summed within each group. For this, the indexes $\{1, \dots, d\}$ are segmented into m ($m < d$) disjoint subsets $\{I_1, \dots, I_m\}$, such that the corresponding elementary matrices \mathbf{X}_i indices, in the same group, are classified into the same *cluster* and added in this sequence. Thus, the trajectory matrix can be expressed by the sum of m matrices:

$$\mathbf{X} = \mathbf{X}_{I_1} + \mathbf{X}_{I_2} + \cdots + \mathbf{X}_{I_m} \quad (3)$$

For example, if $d = 6$ and $m = 3$, the six elementary matrices are grouped into three clusters, as illustrated in Fig.

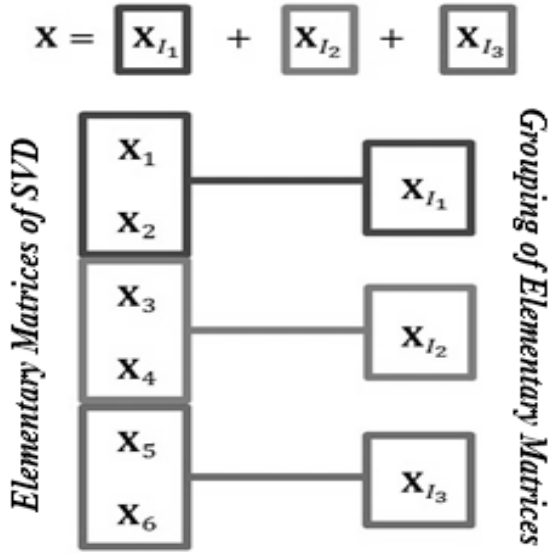


Figure 1. Grouping of elementary matrices.
Source: The authors.

1. The procedure of choosing the sets I_1, \dots, I_m is called the eigentriple grouping.

2.1.2.2. Diagonal averaging

In this stage, each matrix obtained by clustering process is transformed into a time series of length T [25]. Consider a matrix $\mathbf{X}_{I_s} \forall s=1, \dots, m$ with elements $x_{ij}, 1 \leq i \leq L, 1 \leq j \leq K$. Let be $L^* = \min(L, K)$, $K^* = \max(L, K)$ and $T = L + K - 1$. Additionally, let be $x_{ij}^* = x_{ij}$, if $L < K$ and $x_{ij}^* = x_{ji}$, otherwise. Thus, the matrix \mathbf{X}_{I_s} is transform into the series $\tilde{x}_1, \dots, \tilde{x}_T$ as follows:

$$\tilde{x}_k^{(s)} = \begin{cases} \frac{1}{k} \sum_{m=1}^k x_{m, k-m+1}^* & \text{para } 1 \leq k < L^*, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} x_{m, k-m+1}^* & \text{para } L^* \leq k \leq K^*, \\ \frac{1}{T-k+1} \sum_{m=k-K^*+1}^{T-K^*+1} x_{m, k-m+1}^* & \text{para } K^* < k \leq T. \end{cases} \quad (4)$$

The diagonal average of the matrix $\mathbf{X}_{I_s} \forall s=1, \dots, m$ builds the series $\tilde{X}^{(s)} = (\tilde{x}_1^{(s)}, \dots, \tilde{x}_T^{(s)})$. Therefore, the initial series y_1, \dots, y_T is decomposed into a sum of m series:

$$y_t = \sum_{j=1}^m \tilde{x}_t^{(j)}; \quad \forall t = 1, \dots, T \quad (5)$$

2.1.3. Forecasting

To obtain forecasts using SSA the basic requirement is that the time series satisfies the relations *Linear Recurrent Formula (LRF)*, as indicated in [25]. The serie $Y_T = (y_1, y_2, \dots, y_T)$ satisfies the *LRF* order $L - 1$ if:

$$y_i = \sum_{j=1}^{L-1} a_j y_{i-j}; \quad i = T + 1, \dots, T + M \quad (6)$$

The main assumption is that the window length L is chosen in such way to separate the signal from noise. The predictions are performed using the r chosen eigentriples as follows:

$$\hat{y}_i = \begin{cases} \tilde{x}_i, & i = 1, \dots, T \\ \sum_{j=1}^{L-1} a_j \hat{y}_{i-j}, & i = T + 1, \dots, T + M \end{cases} \quad (7)$$

where $\tilde{x}_1, \dots, \tilde{x}_T$ are the values obtained from the reconstructed series in eq. (4), while $\hat{y}_{T+1}, \dots, \hat{y}_{T+M}$ indicate the forecasts M steps ahead.

Also in eq. (7), the coefficients form the vector $R = (a_1, \dots, a_{L-1})^T$ are determined as follows [26]:

$$R = \frac{1}{1 - v^2} \sum_{j=1}^r \pi_j U_j^{\nabla}, \quad (8)$$

where the vector U_j^{∇} denotes the first $L - 1$ components of the eigenvector U_j , π_j is the last element of U_j ($j = 1, \dots, r$) and $v^2 = \sum_{j=1}^r \pi_j^2$.

2.2. Conditional Kernel density forecast methodology

The basic idea of the conditional kernel estimation is to provide a density (pdf) of a random variable Y given a random variable $X = x$. This technique is classified as non-parametric, and presents an advantage over other approaches because it estimates the underlying distribution from the data, without supposing a family distribution. The kernel density estimator computes a smooth density estimation from the data sample by placing to each sample point a function representing its contribution to the density. The distribution is obtained by summing all these contributions.

2.2.1. The Nadaraya-Watson estimator

Conditional density estimation provides the assessment of the pdf of a random variable Y , given an explanatory variable X with the value of x known as follows:

$$f_Y(Y|X = x) = \frac{f_{Y,X}(y, x)}{f_X(x)} \quad (9)$$

The expressions in (9) is unknown, for this reason, this density function is estimated from the data, i.e., with the sample available and using the non-parametric estimator known as Nadaraya-Watson. According to [29], the conditional density estimation is given by

$$\hat{f}(y|x) = \sum_{j=1}^n \mathcal{W}_j(x) K_{h_y}(y - Y_j), \quad (10)$$

where

$$\mathcal{W}_j(x) = \frac{K_{h_x}(x - X_j)}{\sum_{i=1}^n K_{h_x}(x - X_i)}. \quad (11)$$

and where N is the length of the sample, $K_h(\cdot) = K(\cdot/h)/h$ is a kernel function and h the bandwidth parameter. This estimator has two bandwidth, h_x and h_y , that controls the amount of smoothing; i.e., h_y controls the smoothing of each conditional density and h_x controls the smoothing of the explanatory variable.

It is possible to observe that the Nadaraya-Watson estimator requires double kernel estimation, i.e., kernel density estimation in the y direction as indicated in eq. (10) and the other in the x direction as in eq. (11). For a given x , the density function of the random variable Y at the value y is built up by applying kernel density estimation to the sample of values of Y , weighting each Y value according with the contribution of X given by the value x .

3. Methodology

This session brings an overview of the non-parametric proposal to obtain probabilistic short-term forecasts of wind power. This methodology is based on the non-parametric techniques SSA and CDKE described previously. The overall methodological strategy of probabilistic forecasts of wind generation used in this paper is an adaptation of the two stages of point forecast approach of wind power generation, whose structure is shown in Fig. 2 and is detailed as follows:

- First stage: use SSA sequentially to generate hourly forecasts of wind speed h steps ahead.
- Second stage: based on historical data of wind speed and wind power estimate the stochastic power curve, i.e., the conditional density of $P(\text{wind power} | \text{wind speed})$ through the kernel estimator Nadaraya-Watson, which is, basically, the conditional probability density of power at a certain wind speed value.
- Third stage: use the wind speed predicted values at the first stage in combination with the stochastic power curve estimated at the second stage to generate the wind energy density forecasting.

4. Results

4.1. The wind power and wind speed data

To illustrate the application of the described method it was considered a Brazilian wind farm, but for confidentiality reasons it is not possible to publish its name and the location of the park. The dataset is composed of observations ranging from January 1, 2007 up to January 1, 2008, a total of 8784 hourly data of wind speed in m/s (Fig. 3) and wind power in kWh for one turbine, whose capacity is around 2,200kW. In

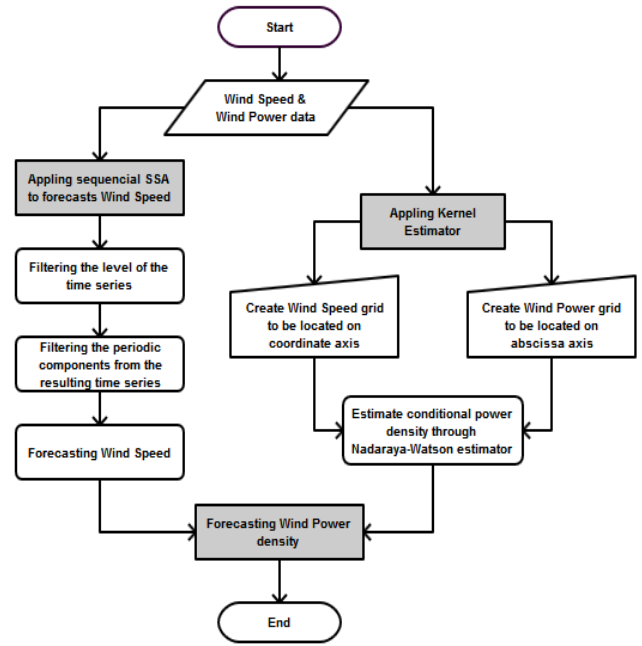


Figure 2. Flowchart of the propose methodology. Source: The authors.

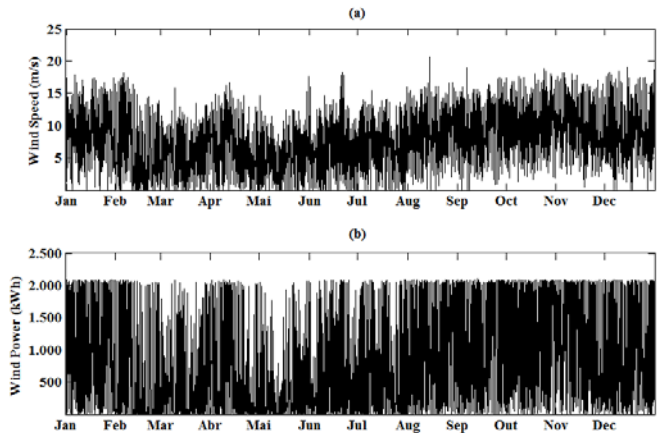


Figure 3. (a) Wind speed (m/s), (b) Power (kWh). Source: The authors.

Fig. 4 it is shown the wind power output versus wind speed for the entire period, thereby generating the power curve covering the whole dataset. In this work it is used both: Rssa [33] and hrdcde [34] packages available in R software.

In what follows, the three stages approach proposed in this paper is applied to the Brazilian wind farm dataset.

4.2. First stage: Wind speed forecast

The first stage of the modelling process to forecast wind speed starts by choosing the L and r hyperparameters. To get better separability of the periodic components it is recommended that the window length L be proportional to the seasonal period (24 hours), as indicated in the first column in Table 1.

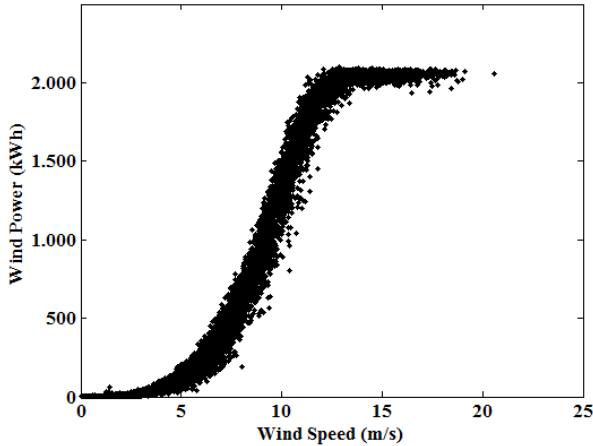


Figure 4. Power Curve (kWh).
Source: The authors.

Table 1.

Lag parameter L .

Parâmetro L	Parâmetro r	RMSE
24	1 até 10	1,9557
36	1 até 16	2,0143
168	1 até 25	1,9217
720	1 até 18	<u>1,4822</u>
2920	1 até 28	2,6318
4380	1 até 29	2,4642

Source: The authors.

For each value of L it was applied SSA and the hyperparameter r was chosen by enumeration, i.e., it was incorporated one by one component until one finds a value that minimizes the root mean square error (RMSE) between the observed series and the forecasts given by eq. (12). Thus the hyperparameters L and r were set at 720 and 18 respectively, as indicated in Table 1. Notice that the r value in this process is only used with the purpose of determining L .

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_t - y_t)^2}{N}} \quad (12)$$

With the identification of $L = 720$, starts the second step of SSA decomposition stage, i.e., the *grouping* of eigentriples via cluster analysis, in particular, using agglomerative hierarchical method. This information is used in SSA procedure to extract the level (first chart of Fig. 5), that corresponds to the first component of the time series as can be seen in Fig. 6 identified as Level Component.

Having obtained extracted the level, the second stage of SSA is applied to reconstruct the time series. The residuals are considered the new time series on which it is applied again SSA to extract the periodic components that maybe present, as can be identified in Fig. 6 for almost all the eigenvectors. Reconstructing the time series was obtained from 259 out of 720 eigentriples. The reconstructed time series is shown on Fig. 7 and the resulting wind speed forecasting for 24 hours ahead on Fig. 8.

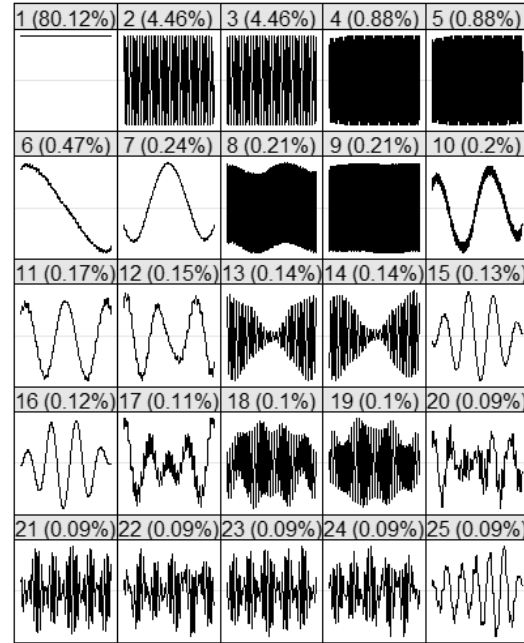


Figure 5. Eigenvectors.
Source: The authors.

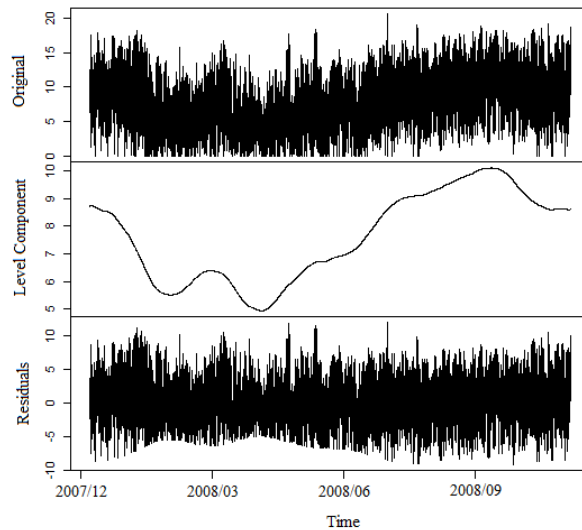


Figure 6. Original time series, Reconstructed series for level component and Residuals.

Source: The authors.

In order to compare the performance of the SSA forecast, other models like double seasonal Holt-Winters and SARIMA were fitted to the data. Naïve estimator is also used for this purpose and all these results are presented in Fig. 9. A numerical comparison is made via some error measures as mean absolute percentage error (MAPE), root mean square error (RMSE), mean absolute error (MAE) and U-THEIL depicted in Table 2 and given by eq. (13), (14) and (15), respectively. The results of these statistics, in bold, indicate that SSA have a better performance in comparison with the other models.

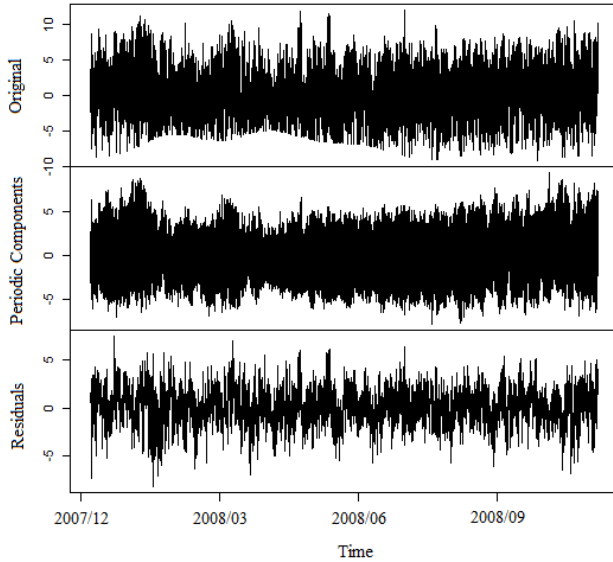


Figure 7. Original time series, Reconstructed series for periodic components and Residuals.

Source: The authors.

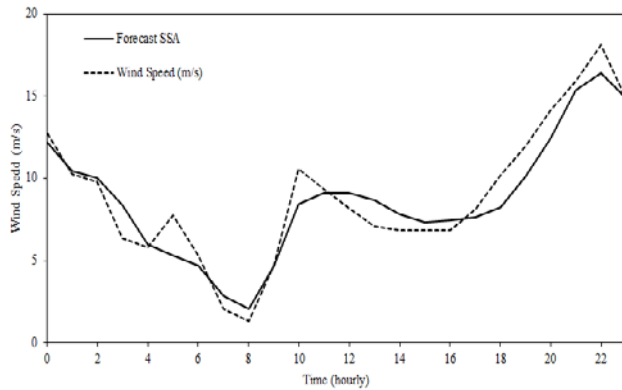


Figure 8. Wind speed (m/s) forecasting 24 hours ahead.

Source: The authors.

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (13)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (14)$$

$$U - THEIL = \sqrt{\frac{\sum_{t=1}^{N-1} \left(\frac{\hat{y}_{t+1} - y_{t+1}}{y_t} \right)^2}{\sum_{t=1}^{N-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}} \quad (15)$$

4.3. Second stage: Stochastic power curve estimation

The second stage employs kernel estimators to obtain the stochastic power curve, where the full density function is

estimated by repeating the conditional kernel density given by eq. (10) and eq. (11). Thus, it is necessary to create a grid with the wind power values (p) ranging from zero up to the wind farm capacity, or like in this study up to the turbine capacity. Variations in p are obtained by small increments until the established capacity.

Similarly, it is required to create a grid for axis x with the wind speed values ranging from zero up to few meters above the maximum wind speed recorded in the database used. This work chooses the speed of 20 m/s as the maximum possible wind speed, because the maximum wind speed recorded in database is very close to this value. Notice that the discretization for both wind power and wind speed must be carried out at regular increments.

Using these grids and eq. (10) and (11), it is estimated the conditional density of wind power. This result is depicted in Fig. 10.

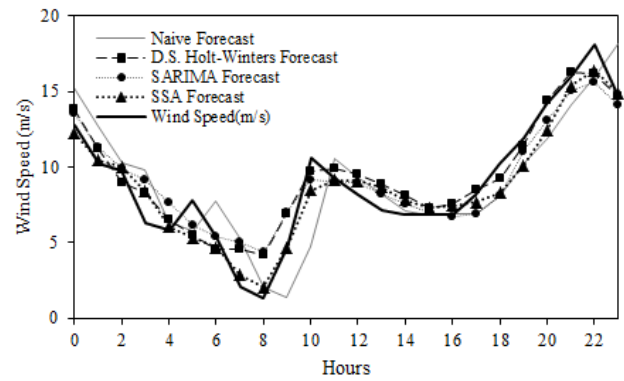


Figure 9. Wind speed (m/s) forecasting 24 hours ahead for different methodologies.

Source: The authors.

Table 2.

Error measures of wind speed forecast.

H = 24	SSA	SARIMA	D.S. HW	NAIVE
MAPE	0,1382	0,2751	0,3208	0,2832
RMSE	1,2050	1,5630	1,8467	2,3121
MAD	0,9585	1,2692	1,4500	1,8911
U-THEIL	0,5345	0,6932	0,8191	1,0255

Source: The authors.

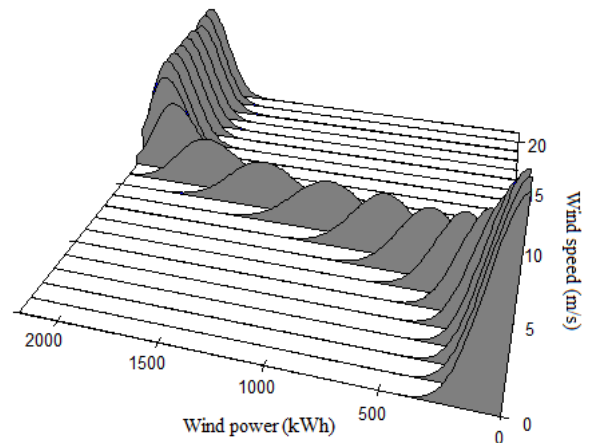


Figure 10. Conditional density estimation of wind power on wind speed.

Source: The authors.

To produce an estimate of the conditional density of wind power production in each discrete value, the wind power should be conditional on a specific wind speed. These estimates are stored to be used in the next stage of the methodological framework.

4.4. Third stage: Wind power density forecast

Considering the estimation of the conditional probability density of wind output, calculated in the previous stage, and the wind speed forecasts generated by applying *sequential SSA* in the first stage, it is obtained the corresponding conditional wind power density forecast. In particular, assuming wind speed forecasts values of 6 m/s and 11 m/s and using the estimated stochastic power curve, one can obtain the respective probability density forecast of wind power conditioned for these two wind speeds, which is shown in Fig. 11.

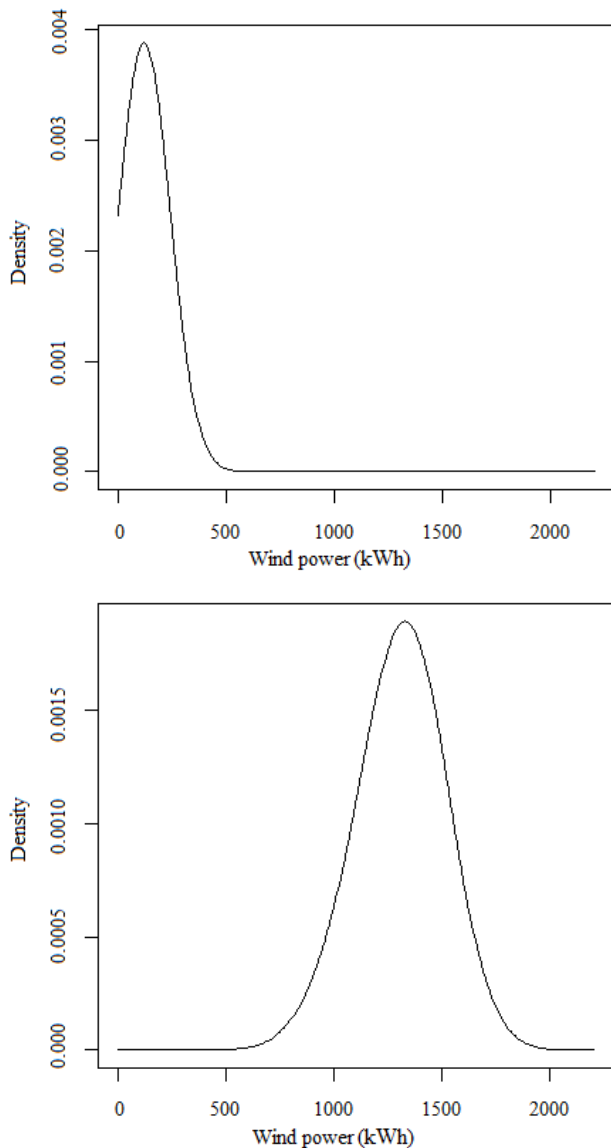


Figure 11. Wind power density for a wind speed of 6 and 11 m/s. Source: The authors.

For each forecast of the wind speed CKDE provides a full probability density function for wind generation. Even though, the purpose of this work is density forecasting, it was evaluated point forecast accuracy using the mean of the distribution as the central location, i.e., a possible point forecast. Analogous to the wind speed forecast, it will be used RMSE, MAE and U-THEIL to evaluate wind power point forecast. Fig. 12 shows wind power forecast take into a count the wind speed forecasted by *sequential SSA*.

Using wind speed forecasts provide by the Naïve method, Double seasonal Holt-Winters and SARIMA model and CKDE it is possible to generate the point forecast for the wind energy production as can be seen in Fig. 13 for each model, while in Table 3 are displayed the results of the error measures for the different approaches.

According to [35,36], the mean of a density forecast is considered the optimal point forecast for a quadratic loss function. This suggests the use of the MAE and the RMSE as the goodness of fit criteria. The results of the error measures in bold in Table 3, show that the sequential SSA for wind speed provide a better performance.

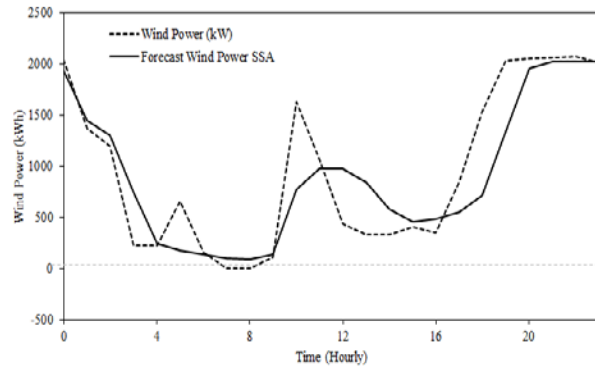


Figure 12. Wind power (kW) forecasting 24 hours ahead. Source: The authors.

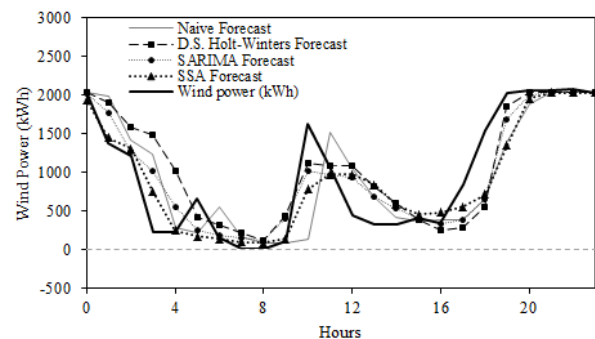


Figure 13. Wind power (kWh) forecasting 24 hours ahead for different methodologies. Source: The authors.

Table 3.

Error measures for wind power forecast.

H = 24	SSA	SARIMA	D.S. HW	NAIVE
RMSE	375,73	380,49	458,96	498,24
MAD	252,66	264,47	324,84	329,52
U-THEIL	0,7541	0,7637	0,9212	1,0000

Source: The authors.

5. Conclusions

This paper proposes a new methodological framework for modeling short-term wind power generation and captures the uncertainty concerning to wind speed. The developed work provides insight on how to obtain forecasts of wind power probability density function by applying sequential SSA and CKDE. For short-term wind speed forecast it is used SSA approach that proved to be a robust alternative to produce wind speed forecast. The comparison with other methodologies to produce wind speed forecast, show that an improvement of the forecasts for the wind energy is improved considerably using the proposed method in this paper. As a future work, it is suggested to incorporate other explanatory variables such as direction, temperature and air pressure to modeling wind speed and evaluate other kernel density estimator that allows time-varying for the parameters.

Acknowledgment

The author S. Aguilar would like to thank FAPERJ/CAPES (research projects E-26/202.168/2015) for their financial support. F. L. Cyrino Oliveira also thanks the support of CNPq (research project 443595/2014-3) and FAPERJ (research project E-26/202.806/2015).

References

- [1] Brasil, Ministério de Minas e Energia, Empresa de Pesquisa Energética. Plano Decenal de Expansão de Energia 2024, Brasília: MME/EPE 2015. [online]. [date of reference March 24th of 2016]. Available at: <http://www.epe.gov.br/PDEE/Relat%C3%B3rio%20Final%20do%20PDE%202024.pdf>
- [2] Zhu, X. and Genton M.G., Short-Term wind speed forecasting for power system operations. *International Statistical Review*, 80(1), pp. 2-23, 2012. DOI: 10.1111/j.1751-5823.2011.00168.x
- [3] Madsen, H., Wind power prediction tool in control dispatch centres. ELSAM, Skaerbaek, Denmark, 1995, ISBN 87-87090-25-2.
- [4] Nielsen, T.S. and Madsen, H., Using meteorological forecasts in on-line predictions of wind power, ELSAM, Skaerbaek, Denmark, 1996.
- [5] Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J. and Conzelmann, G., Wind power forecasting: State-of-the-art 2009, Argonne National Laboratory ANL/DIS-10-1, 2009.
- [6] Golyandina, N., Nekrutkin, V. and Zhigljavsky, A., Analysis of time series structure: SSA and related techniques. Chapman & Hall/CRC, New York, 2001.
- [7] Hassani, H., Singular spectrum analysis: Methodology and comparison. *Journal of Data Science*, 5, pp. 239-257, 2007.
- [8] Brown, B.G., Katz, R.W. and Murphy, A.H., Time series models to simulate and forecast wind speed and wind power. *Journal of Climate and Applied Meteorology*, 23(8), pp. 1184-1195, 1984. DOI: 10.1175/1520-0450(1984)023<1184:TSMTSA>2.0.CO;2
- [9] Kamal, L. and Jafri, Y.Z., Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Solar Energy*, 61(1), pp. 23-32, 1997. DOI: 10.1016/S0038-092X(97)00037-6
- [10] Poggi, P., Muselli, M., Notton, G., Cristofari, C. and Louche, A., Forecasting and simulating wind speed in Corsica by using an autoregressive model. *Energy Conversion and Management*, 44(20), pp. 3177-3196, 2003. DOI: 10.1016/S0196-8904(03)00108-0
- [11] Philippopoulos, K. and Deligiorgi, D., Statistical simulation of wind speed in Athens, Greece based on Weibull and ARMA models. *International Journal of Energy and Environment*, 3(4), pp. 151-158, 2009.
- [12] Erdem, E. and Shi, J., ARMA based approaches for forecasting the tuple of wind speed and direction. *Applied Energy*, 88(4), pp. 1405-1414, 2011. DOI: 10.1016/j.apenergy.2010.10.031
- [13] Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I. and Kallos, G., Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(12), pp. 2348-2362, 2008. DOI: 10.1016/j.jweia.2008.03.013
- [14] Liu, H., Tian, H. and Li, Y., Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Applied Energy*, 98(0), pp. 415-424, 2012. DOI: 10.1016/j.apenergy.2012.04.001
- [15] Malmberg, A., Holst, U. and Holst, J., Forecasting near-surface ocean winds with Kalman filter techniques. *Ocean Engineering*, 32, pp. 273-291, 2005. DOI: 10.1016/j.oceaneng.2004.08.005
- [16] Kavasseri, R.G. and Seetharaman, K., Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy*, 34(5), pp. 1388-1393, 2009. DOI: 10.1016/j.renene.2008.09.006
- [17] Caporin, M. and Preš, J., Modelling and forecasting wind speed intensity for weather risk management. *Computational Statistics and Data Analysis*, 56(11), pp. 3459-3476, 2012. DOI: 10.1016/j.csda.2010.06.019
- [18] Cadenas, E. and Rivera, W., Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. *Renewable Energy*, 34(1), pp. 274-278, 2009. DOI: 10.1016/j.renene.2008.03.014
- [19] Barbounis, T.G., Theocharis, J.B., Alexiadis, M.C. and Dokopoulos, P.S., Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE, Transactions on Energy Conversion*, 21(1), pp. 273-284, 2006. DOI: 10.1109/TEC.2005.847954
- [20] Sideratos, G. and Hatzigryriou, N.D., Probabilistic wind power forecasting using radial basis function neural networks. *IEEE, Transactions on Power Systems*, 27(4), pp. 1788-1796, 2012. DOI: 10.1109/TPWRS.2012.2187803
- [21] Damousis, I.G., Alexiadis, M.C., Theocharis, J.B. and Dokopoulos, P., A fuzzy model for wind speed prediction and power generation in wind farms using spatial correlation. *IEEE, Transactions on Energy Conversion*, 19, pp. 352-361, 2004. DOI: 10.1109/TEC.2003.821865
- [22] Zhao, X., Wang, S.X. and Li, T., Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia*, 12, pp. 761-769, 2011. DOI: 10.1016/j.egypro.2011.10.102
- [23] Genton, M.G. and Hering, A.S., Blowing in the wind. *Significance*, 4, pp. 11-14, 2007. DOI: 10.1111/j.1740-9713.2007.00212.x
- [24] Beneki, C., Eeckels, B. and Leon, C., Signal extraction and forecasting of the UK tourism income time series: A singular spectrum analysis approach. *Journal of Forecasting*, 31(5), pp. 391-400, 2012. DOI: 10.1002/for.1220
- [25] Golyandina, N. and Zhigljavsky, A., Singular spectrum analysis for time series. New York: Springer, 2013.
- [26] Mahmoudvand, R., Najari, N. and Zokaei, M., On the optimal parameters for reconstruction and forecasting in the singular spectrum analysis. *Communication in Statistics – Simulations and Computations*, 42, pp. 860-870, 2013. DOI: 10.1080/03610918.2012.659822
- [27] Juban, J., Siebert, N. and Kariniotakis, G.N., Probabilistic short-term wind power forecasting for the optimal management of wind generation. *IEEE Lausanne Power Tech*, pp. 683-688, 2007. DOI: 10.1109/PCT.2007.4538398
- [28] Jeon, J. and Taylor, J.W., Using conditional Kernel density estimation for wind power density forecasting. *J. Am. Stat. Assoc.*, 107(497), pp. 66-79, 2012. DOI: 10.1080/01621459.2011.643745
- [29] Hyndman, R.J., Bashtannyk, D.M. and Grunwald, G.K., Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics*, 5(4), pp. 315-336, 1996. DOI: 10.2307/1390887
- [30] Parzen, E., On estimation of a probability density function and model. *The Annals of Mathematical Statistics*, 33, pp. 1065-1076, 1962. DOI: 10.1214/aoms/1177704472
- [31] Rosenblatt, M., Conditional probability density and regression estimates. *Multivariate Analysis II*, Ed. P.R. Krishnaiah, New York: Academic Press, pp. 25-31, 1969.

- [32] Bessa, R.J., Miranda, V., Botterud, A., Wang, J. and Constantinescu, E.M., Time adaptive conditional Kernel density estimation for wind power forecasting. *IEEE Trans. Sustain. Energy*, 3(4), pp. 660-669, 2012. DOI: 10.1109/TSTE.2012.2200302
- [33] Golyandina, N. and Korobeynikov, A., Basic singular spectrum analysis and forecasting with R. *Computational Statistics & Data Analysis*, 71, pp. 934-954, 2014. DOI: 10.1016/j.csda.2013.04.009
- [34] Hyndman, R.J., Highest density regions and conditional density estimation. R package version 3.1, 2015. [online]. [date of reference June 26th of 2016]. Available at: <https://cran.r-project.org/web/packages/hdr/de/hdrde.pdf>
- [35] Gneiting, T., Quantiles as optimal point forecasts. *International Journal of Forecasting*, 27, pp. 197-207, 2011. DOI: 10.1016/j.ijforecast.2009.12.015
- [36] Gneiting, T., Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), pp. 746-762, 2012. DOI: 10.1198/jasa.2011.r10138

S. Aguilar-Vargas, received the BSc. Eng in Management Engineering in 2005, the MSc. degree in Systems Engineering in 2009, both from the Universidad Nacional de Colombia, Medellin, Colombia. She obtained her PhD. degree in Electrical Engineering from Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil, in 2015. She is currently a post-doctoral research at PUC-Rio and her current research interests cover a range of topics relevant to modelling and forecasting time series.
ORCID: 0000-0002-5177-8336

R. Castro-Souza, received a PhD in Bayesian forecasting from Warwick University, UK. Presently, he is a full professor in statistics and time series analysis at Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil and his research interests are in the area of forecasting with applications to the energy market.
ORCID: 0000-0001-9614-0507

J.F. Pessanha, has a BSc in Statistics by National School of Statistical Science - Ence, 1992 and he is Electrical Engineer by Rio de Janeiro State University - UERJ, Brazil in 1994. In addition, he is MSc. by Federal University of Rio de Janeiro - UFRJ, in 1999, and PhD. by Pontifical Catholic University of Rio de Janeiro, Brazil, in 2006, both in Electrical Engineering. He is professor at Institute of Mathematics and Statistics at the UERJ and Researcher at the Brazilian Electric Research Center - CEPEL.
ORCID: 0000-0002-7134-2388

F.L. Cyrino-Oliveira, is a Senior Lecturer in the Industrial Engineering Department of PUC-Rio (Pontifical Catholic University of Rio de Janeiro) and Deputy Head for Graduate Studies. He obtained his undergraduate degree in Industrial Engineering at Federal University of Juiz de Fora and his Master and Doctoral degrees in Electrical Engineering at PUC-Rio. His research interests include Time Series, Forecasting, Data Analytics, Operations Research, Simulation and Statistical Learning.
ORCID: 0000-0003-1870-9440



UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería
de Sistemas e Informática

Oferta de Posgrados

Especialización en Sistemas
Especialización en Mercados de Energía
Maestría en Ingeniería - Ingeniería de Sistemas
Doctorado en Ingeniería- Sistema e Informática

Mayor información:

E-mail: acsei_med@unal.edu.co
Teléfono: (57-4) 425 5365