



## Previsão do consumo de eletricidade no nordeste brasileiro Forecast of electricity consumption in Brazilian Northeast

Luiz Moreira Coelho Junior<sup>1\*</sup>

Thiago Freire Melquíades<sup>2</sup>

Kalyne de Lourdes da Costa Martins<sup>3</sup>

Edvaldo Pereira Santos Júnior<sup>4</sup>

Graziela Pinto de Freitas<sup>5</sup>

**Resumo:** O crescimento do consumo de energia elétrica no Brasil gera impactos no modo de vida da sociedade moderna, quando mostra sua relevância e apresenta diversas possibilidades de estudos. Este artigo analisou a série temporal do consumo de energia elétrica no Nordeste brasileiro, no período de janeiro de 2004 a dezembro de 2012 e estimou a previsão para o ano de 2013. O método utilizado foi preconizado por Box & Jenkins, por meio dos modelos da família ARIMA. Para a análise da série e a escolha do modelo para previsões do ano de 2013, foi usado como base o Critério de Informação de Akaike, no teste Box-Pierce e no Erro Quadrático Médio de Previsão. Os modelos identificados pelas funções de autocorrelação e de autocorrelação parcial não apresentaram ajustes adequados, sendo necessário buscar novos modelos, de acordo com princípio da parcimônia. O modelo SARIMA(1,1,0)(0,0,1)<sub>12</sub> foi que forneceu a melhor previsão para o consumo de energia elétrica no nordeste brasileiro.

**Palavras-chave:** Economia da energia, séries temporais, modelo Box-Jenkins

<sup>1</sup> Economista, Dr., Professor do Departamento de Engenharia de Energias Renováveis, Universidade Federal da Paraíba, Campus I, João Pessoa, PB - Brasil. E-mail: luiz@cear.ufpb.br

<sup>2</sup> Administrador, Mestre pelo Programa de Pós-Graduação em Energias Renováveis, Universidade Federal da Paraíba, Campus I, João Pessoa, PB - Brasil. E-mail: thiago-melquiades@hotmail.com

<sup>3</sup> Economista, Mestre pelo Programa de Pós-Graduação em Energias Renováveis, Universidade Federal da Paraíba, Campus I, João Pessoa, PB - Brasil. E-mail: kalyne\_lourdes@hotmail.com

<sup>4</sup> Engenheiro de Energias Renováveis, Universidade Federal da Paraíba, Campus I, João Pessoa, PB - Brasil. E-mail: edvaldo.junior@cear.ufpb.br

<sup>5</sup> Engenheira Ambiental, Mestranda pelo Programa de Pós-Graduação em Energias Renováveis, Universidade Federal da Paraíba, Campus I, João Pessoa, PB - Brasil. E-mail: freitas.graziela@cear.ufpb.br

\*Autor para correspondência.

**Abstract:**

The growth in electricity consumption in Brazil generates impacts on the way of modern society's life, whereas it shows its relevance and approaches to studies. This research analyzed the time series of electricity consumption in Northeast Brazil, from January 2004 to December 2012 and estimate for the year 2013. The method used was recommended by Box and Jenkins, through the ARIMA family models. The analysis of the series and the choice of model for 2013 year forecasts were based on the Akaike Information Criterion, the Box-Pierce test and the mean squared error of forecast. The models, which were identified by the autocorrelation and partial autocorrelation functions, did not show proper adjustments, thus it was necessity to seek new models according to the parsimony principle. The SARIMA model  $(1,1,0) (0,0,1)_{12}$  provided the best prediction for the consumption of electricity in Brazilian northeast. The ARIMA family models offered good predictions about the electricity consumption in the Brazilian northeast.

**Keywords:** Energy economy, times series, Box-Jenkins model.

---

## 1. Introdução

No Brasil, a demanda por energia elétrica vem crescendo exponencialmente nos últimos anos. Para a Empresa de Pesquisa Energética – EPE (2012), este crescimento é proporcionado pela mudança do perfil da população, que apresentou um maior incremento na renda per capita, e pela necessidade de expansão industrial.

As indústrias brasileiras de aço bruto, celulose e papel e cimento são os maiores consumidores da energia elétrica. Esses setores encabeçaram o fornecimento de grande parte dos insumos básicos requeridos por outros segmentos da economia, como a construção civil e a produção de máquinas e equipamentos. Os produtos fabricados a partir desses insumos compõem o estilo de desenvolvimento da sociedade moderna, e, conseqüentemente, geram consideráveis impactos ao consumo de energia elétrica (EPE, 2012).

Em 2012, o consumo de energia elétrica no Nordeste foi de aproximadamente 76 GWh, com 19 milhões de consumidores, o que representa 26,2% do total nacional. Desta participação, a indústria obteve maior consumo de eletricidade, com 38,2% do total, seguido do consumo residencial, com 28,3% e do setor comercial, com 15,4%. Os principais segmentos da indústria nordestina foram metalurgia, fabricação de produtos químicos, fabricação de produtos minerais não metálicos e fabricação de produtos têxteis. Já para o setor comercial, os principais segmentos foram o varejo, o de alojamento, o de telecomunicações e as atividades de atenção à saúde humana (EPE, 2013).

A previsão da consiste em um fator importante para tomada de decisão e orientação de políticas públicas, por prever tendências, sazonalidades e ciclos futuros, promovendo estimativas para um determinado momento (REZENDE *et al.*, 2005; COELHO JUNIOR *et al.*, 2006; SLACK *et al.*, 2009; COELHO JUNIOR *et al.*, 2009; COELHO JUNIOR *et al.*, 2013). Diante deste contexto, prever a demanda de energia elétrica constitui uma importante análise em relação ao comportamento futuro da indústria e da população brasileiras.

Uma técnica bastante difundida para construção de uma previsão é a metodologia de Box & Jenkins (1976), desenvolvida na década de 70, com a publicação do livro *Time Series Analysis, Forecasting and Control*. Essa abordagem consiste na análise de séries temporais, onde o comportamento de uma variável, essencialmente, pode ser definido pelo seu próprio comportamento no passado.

A aplicação da metodologia Box-Jenkins para o consumo de energia elétrica foi utilizada por Gomes (1989); Lobo (1991); Zandonade (1993); Neto *et al.* (2006) e Campos (2008). Os estudos dimensionaram as vantagens e desvantagens do método como instrumento de previsão da demanda fundamentado em séries temporais. Já na pesquisa desenvolvida por Ediger & Akar (2007), os resultados despontaram para a qualidade dos modelos ARIMA na previsão da demanda de energia, apresentando conclusões satisfatórias. Contudo, este trabalho analisou a série temporal do Consumo de Energia Elétrica no Nordeste (CEEN), no período de janeiro de 2004 a dezembro de 2012, por meio dos modelos da família ARIMA e estimou a previsão para 2013.

## **2. Materiais e Métodos**

Os procedimentos metodológicos utilizados foram subdivididos em dois tópicos: objeto de estudo e séries temporais.

### **2.1. Objeto de estudo**

Os dados empregados neste trabalho foram obtidos do relatório mensal de consumo de energia por classe, fornecido pela Empresa de Pesquisa Energética – EPE. Contudo, a atualização deste material é de responsabilidade da Superintendência de Estudos Econômicos e Energéticos – SEE. As informações podem ser obtidas segmentadas por classes, como residencial, industrial, comercial, rural, serviço público e iluminação pública. Também podem ser classificadas por região geográfica e subsistemas elétricos.

A série temporal estudada corresponde a 108 observações do registro mensal de consumo de energia elétrica da Região Nordeste, em Megawatt/hora (MWh), no período de janeiro de 2004 a dezembro de 2013. Para o ajuste do modelo, consideraram-se os meses de janeiro de 2004 a dezembro de 2012. Reservaram-se 12 observações, que correspondem aos meses de Janeiro a Dezembro de 2013 para validação da previsão, como fonte de comparação da escolha do melhor modelo.

## 2.2. Séries temporais

Uma série temporal pode ser definida como um composto de informações de uma variável dispostas logicamente ao longo do tempo, formando uma determinada trajetória de observações. As séries temporais podem ser classificadas como discretas ou contínuas, e, por muitas vezes, uma série discreta pode ser obtida através de uma amostragem de uma série contínua, para intervalos de tempo fixos (FISCHER, 1982).

A série pode ser obtida por  $Z(t_1), \dots, Z(t_n)$ , observada no instante  $t_1, \dots, t_n$ . A análise de uma série temporal pode resultar na extração de componentes com a presença de sazonalidade, ciclos e tendências. Esses modelos podem ser probabilísticos ou estocásticos, no sentido de serem simples e parcimoniosos, ou seja, a melhor forma de representação para o fenômeno estudado (MORETTIN & TOLOI, 2006).

Outra característica de uma série temporal, diz respeito à sua estacionariedade ou não estacionariedade. No entendimento de Bueno (2011), as séries estacionárias apresentam média e variância constantes ao longo do tempo, ao contrário das séries não estacionárias. As combinações dos tipos de séries temporais são apresentadas na Tabela 1, onde  $C$  é uma constante e  $\varepsilon_t$ , um componente aleatório extraído de uma distribuição normal  $\varepsilon_t \sim N(0,1)$ .

**Tabela 1:** Tipos de séries temporais

Tipo de Série		Estocástica	Determinística
Estacionária		$y_t = c + \varepsilon_t$	$y = c + 0,5 y_{t-1}$
Não Estacionária	Tendência Estacionária	$y_t = c + \delta t + \varepsilon_t$	$y = c + \delta t$
	Tendência Estocástica	$y_t = y_{t-1} + \varepsilon_t$	

Fonte: Bueno, 2011, p.02.

Outro conceito bastante importante de séries temporais é o de ruído branco (RB), que consiste em uma série estocástica discreta com sequência  $\{\varepsilon_t\}$  apresentar média

zero, variância constante e elas não estarem correlacionadas com a própria série, ou seja, a representação de RB  $(0, \sigma^2)$ .

### 2.2.1. Box e Jenkins

O método Box-Jenkins foi desenvolvido na década de 1970, pelos Professores George E. P. Box e Gwilym M. Jenkins. Consiste na formação de uma estrutura de ciclos interativos, determinados pela seguinte forma:

- I. Determina a classe do modelo;
- II. Faz-se a identificação do modelo proposto por meio da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) e outros critérios;
- III. Consiste na determinação dos parâmetros do modelo na qual foi identificado;
- IV. Obtêm-se o diagnóstico do modelo ajustado, averiguando se os resíduos são ruído branco e adequados para previsão.

Caso o modelo se apresente adequado, o processo é finalizado e ocorre a previsão, caso contrário, o processo é reiniciado a partir da etapa de identificação (MORETTIN & TOLOI, 2006). A estrutura que representa esse processo é demonstrada na Figura 1.

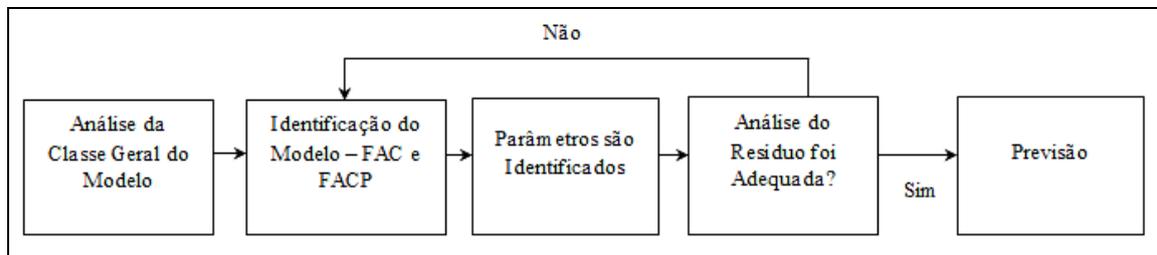


Figura 1. Metodologia Box & Jenkins  
Fonte: Morettin & Tolo, 2006.

### 2.2.2. Família ARIMA

Dado que o modelo Autoregressivo Integrado a Média Móvel (ARMA) parte da combinação dos processos AR e MA, ele também pode apresentar a combinação de ambos, denominado ARMA, descrito por Lee & Ko (2011), o ARMA  $(p, q)$  na Equação 1. No entanto, os processos AR, MA e ARMA, são estacionários. No caso de processos não estacionários, necessita-se do processo de diferenciação, com isso, obtém-se a estabilização do modelo, formando o ARIMA  $(p, d, q)$ , como disposto na Equação 2.

Observa-se que a equação  $\Delta^d z_t = (1 - B)^d z_t$  refere-se à ordem de diferenciação.

$$\phi_p(B)z_t = \theta_q(B)\alpha_t \quad (1)$$

$$\varphi_p(B)\Delta^d Z_t = \theta_q(B)\alpha_t \quad (2)$$

O modelo também pode avaliar a presença de sazonalidade de uma série temporal. Para tal, acrescenta-se um “S” ao modelo. O SARIMA (p,d,q)(P,D,Q)<sub>s</sub>, que representa o modelo ARIMA sazonal multiplicativo (Equação 3). Onde  $\Delta_S^D Z_t = (1-B^S)^D Z_t$  representa a ordem de diferenciação sazonal de D. Também,  $\varphi_p(B^S)$  representa AR(p) sazonal e  $\theta_q(B^S)$ , o MA(q) sazonal, esses são definidos nas Equações 4 e 5, respectivamente. Os parâmetros sazonais de AR(p) são descritos por  $\Phi_1 \dots \Phi_p$ , já MA(q) por  $\Theta_1 \dots \Theta_q$  (Lee & Lee, 2005; Coelho Junior *et al.*, 2006).

$$\varphi_p(B)\Phi_P(B^S)\Delta^d \Delta_S^D Z_t = \theta_q(B)\Theta_Q(B^S)\alpha_t \quad (3)$$

$$\Phi_P(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS} \quad (4)$$

$$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \quad (5)$$

### 2.2.3. Teste da raiz unitária

Para Gujarati (2006), considerou a equação do modelo de passeio aleatório (Equação 6), onde  $u_t$  é um termo de ruído branco e para  $\rho = 1$ , chama-se raiz unitária, que consiste em uma situação de não estacionariedade. No entanto, se  $|\rho| \leq 1$ , tem-se uma situação de estacionariedade, ou seja, a série temporal  $Y_t$  é estacionária. Para observar a transformação de série não estacionária em uma série estacionária, deve-se subtrair  $Y_{t-1}$  de ambos os lados da equação 6, desta maneira, obtém a equação 7. A nova equação consiste em  $\delta = (\rho - 1)$ , com  $\Delta$  representando a primeira diferença, que é uma série estacionária.

$$Y_t = \rho Y_{t-1} + u_t \quad -1 \leq \rho \leq 1 \quad (6)$$

$$\Delta Y_t = \delta Y_{t-1} + \mu_t \quad (7)$$

Para a comprovação formal da hipótese nula, foram utilizados alguns testes, como o teste *Dickey-Fuller* Aumentado (ADF), desenvolvido por Dickey & Fuller (1981), o teste *Phillips-Perron* (PP); desenvolvido por Phillips & Perron (1988) e o teste *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS); desenvolvido por Kwiatkowski *et al.* (1992). Esses testes confirmam se a série  $Y_t$  possui ou não raiz unitária, com isso, comprova-se ou não a estacionariedade da série.

### 2.2.4. Autocorrelação de uma série

Segundo Gujarati (2006), a Função de Autocorrelação (FAC) é dada por  $k$  defasagens, com denotação de  $\rho_k$  (Equação 8). O termo  $\gamma_k$  representa a covariância com  $k$  defasagens e  $\gamma_0$  representa a variância. Isso caracteriza que quando a defasagem  $k = 0$ , em que  $\rho_0 = 1$ , considerando que  $\rho_k$  é um número puro que se situa entre  $\pm 1$  para qualquer coeficiente de correlação. Assim, deve-se calcular a função de autocorrelação amostral, representada pela covariância com defasagem  $k$  (Equação 9), a variância amostral (Equação 10), na qual  $n$  representa o tamanho da amostra e  $\bar{Y}$  é considerada a média amostral. Chega-se à equação de autorrelação amostral (Equação 11), cujo gráfico representante é chamado de correlograma amostral. Conclui-se que uma série temporal pode ser considerada estacionária quando se assemelha ao correlograma de uma série temporal de ruído branco.

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{8}$$

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n} \tag{9}$$

$$\hat{\gamma}_0 = \frac{\sum (Y_t - \bar{Y})^2}{n} \tag{10}$$

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \tag{11}$$

A Função de Autocorrelação (FAC) denota a autocorrelação *versus* as defasagens, ela ajudará a identificar o processo de Médias Móveis de Ordem  $q$  – MA( $q$ ), tendo uma participação bastante importante no processo de estimação de séries temporais. Considerando a Equação 12, onde  $\varepsilon_t$  representa ruído branco, e  $y_t$  depende do erro  $\varepsilon_t$ , seguidamente do erro do passado, denominado Médias Móveis de Ordem 1 – MA(1), no entanto, se o processo depende de  $\varepsilon_{t-2}$ , passa-se a chamá-lo de Médias Móveis de Ordem 2 – MA(2). Já a Função de Autocorrelação Parcial (FACP), é caracterizada por um processo de filtragem da FAC, demonstrado por um Processo Autorregressivo de Ordem  $p$  – AR( $p$ ), ou seja, um processo de filtragem, no qual se expurgam as correlações implícitas, conservando as relações puras entre as observações. A FACP denota  $\hat{\phi}_{j,j}$ , *versus*  $j$ , onde se extrai a regressão de  $y_t$  em detrimento de  $y_{t-1}$ ,

obtendo  $\phi_{1,1}$ , e, por conseguinte,  $y_t$  versus  $y_{t-1}$  e  $y_{t-2}$ , assim consegue  $\phi_{2,1}$  e  $\phi_{2,2}$ , como mostra a Equação 13 (BUENO, 2011).

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \quad (12)$$

$$y_t = \phi_{j,1}y_{t-1} + \phi_{j,2}y_{t-2} + \dots + \phi_{j,j}y_{t-j} + e_{t,j}, j = 1,2,\dots \quad (13)$$

### 2.2.5 Escolha do modelo

Inicialmente, o processo consiste na identificação dos parâmetros da série, que pode ser representada por um modelo ARIMA (p,d,q), já se apresentar indícios de sazonalidade, será representada por um modelo de classe SARIMA (p,d,q)(P,D,Q)<sub>s</sub>. O processo parte da análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP), assim como os testes ADF, PP e KPSS, para verificação da raiz unitária da série. Logo após, escolhem-se os modelos pretendentes utilizando o *Akaike Information Criterion* (AIC), seguindo o princípio da parcimônia. Com isso, faz-se a análise dos resíduos da série ajustada verificando a existência de ruído branco.

Na interpretação de Bueno (2011), os critérios de informação são importantes para balancear a redução de erro, fornecendo os parâmetros aceitáveis para um modelo, como exposto na equação 14.

$$AIC(p,q) = \ln \hat{\sigma}^2 + n \frac{2}{T} \quad (14)$$

Outra proposta é *Portmanteau Test*, que, para Tsay (2005), consiste na equação 15. Onde rejeita-se a hipótese de  $H_0$  se  $Q(m) > x_{\alpha}^2$ . O teste permite reforçar a existência de ruído branco no modelo escolhido.

$$Q(m) = T(T+2) \sum_{\lambda=1}^m \frac{\hat{\rho}_{\lambda}^2}{\lambda} \quad (15)$$

Assim sendo, verifica-se o Erro Quadrático Médio (EQM) das previsões obtidas, que permite a comparação dos valores previstos e observados da série ajustada (16), sendo escolhido o modelo que apresentar o menor EQM.

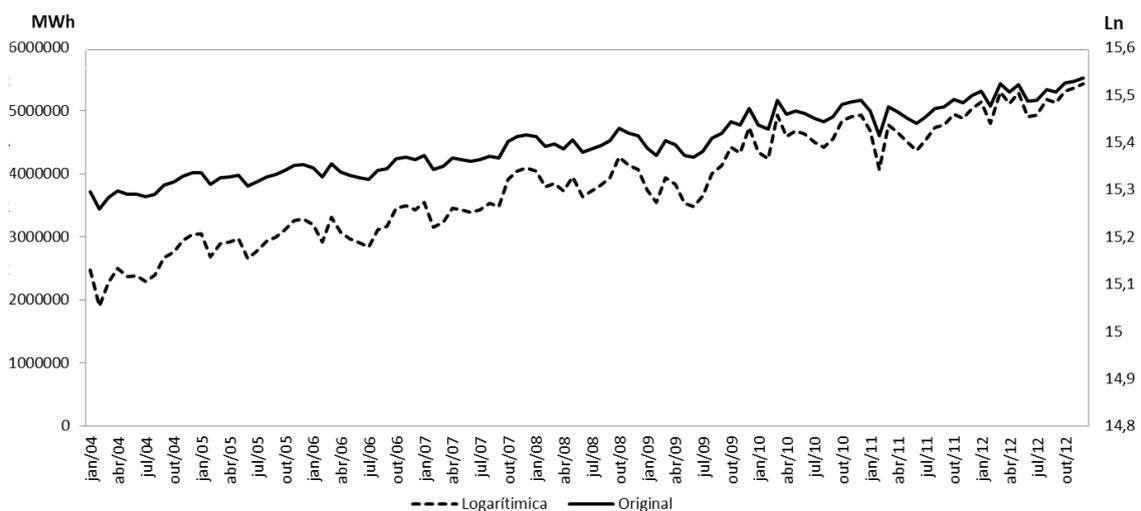
$$EQM = \frac{\sum (y_t - y_t^e)^2}{n} \quad (16)$$

### 2.2.6. Tratamento dos dados

Para operacionalização e tratamento dos dados, utilizou-se o *software* R, como o auxílio do *Package tseries* (R Development Core, 2014). Complementando, Torgo (2009), ressalta que o R consiste em uma ferramenta bastante útil no que diz respeito ao nível de programação, fator atenuado pela presença do *package*, que são funções de determinadas funcionalidades e tópicos, disponíveis gratuitamente.

## 3. Resultados e Discussão

A Figura 2 apresenta o comportamento da série original do consumo de energia elétrica da Região Nordeste (CEEN), em MWh, e da série logarítmizada [ $\ln(\text{CEEN})$ ] no período de janeiro de 2004 a dezembro de 2012. A logaritimização neperiana foi necessária para estabilização da variância, preservando as propriedades dos dados da série. A CEEN possui média estimada em 4.490.502, mediana de 4.446.000, valor mínimo de 3.445.000 e valor máximo de 5.519.000.



**Figura 2:** Evolução do Consumo de energia elétrica da Região Nordeste (CEEN) original e logarítmica, em MWh, no período de 2004 a 2012.

Pode-se observar uma tendência de crescimento do CEEN, isso fornece indícios que ela não pode ser estacionária. A série também apresenta sazonalidade com efeitos aleatórios ao longo do tempo. Apenas com a análise visual da Figura 2, não se pode afirmar se a série  $\ln(\text{CEEN})$  é estacionária ou não. Assim, para uma constatação mais formal, utilizaram-se os testes *Dickey-Fuller* Aumentado (ADF), o teste *Phillips-Perron* (PP) e o teste *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS), para verificação da presença de raiz unitária, apresentados na Tabela 2.

**Tabela 2:** Testes de raiz unitária para a série do consumo de energia elétrica do Nordeste logarítmica [Ln(CEEN)] e a primeira diferença {1ª Dif [Ln(CEEN)]}

Testes	[Ln(CEEN)]				1ª Dif [Ln(CEEN)]			
	1%	5%	10%	t	1%	5%	10%	T
<b>ADF</b>	-3,99	-3,43	-3,13	-3,8599	-3,99	-3,43	-3,13	-5,5058
<b>KPSS Test</b>	0,739	0,463	0,347	2,1822	0,739	0,463	0,347	0,029
<b>Phillips-Perron</b>	-4,046	-3,451	-3,151	-5,855	-4,046	-3,451	-3,151	-14,949

O teste ADF mostra a presença da estacionariedade quando há rejeição hipótese nula ( $H_0$ ), ou seja, quando a série tem presença de raiz unitária. E para a hipótese alternativa ( $H_1$ ) é a constatação da estacionariedade da série, não incorrendo na região de raiz unitária. O teste ADF mostra que se  $|\alpha| < |t|$ , aceita-se  $H_0$ . Desta forma, o Ln(CEEN) ao nível de significância de 1% aceita  $H_0$ , pois há presença de raiz unitária, ou seja, é não estacionária, sendo necessário transformar pela 1ª diferença para torná-la estacionária.

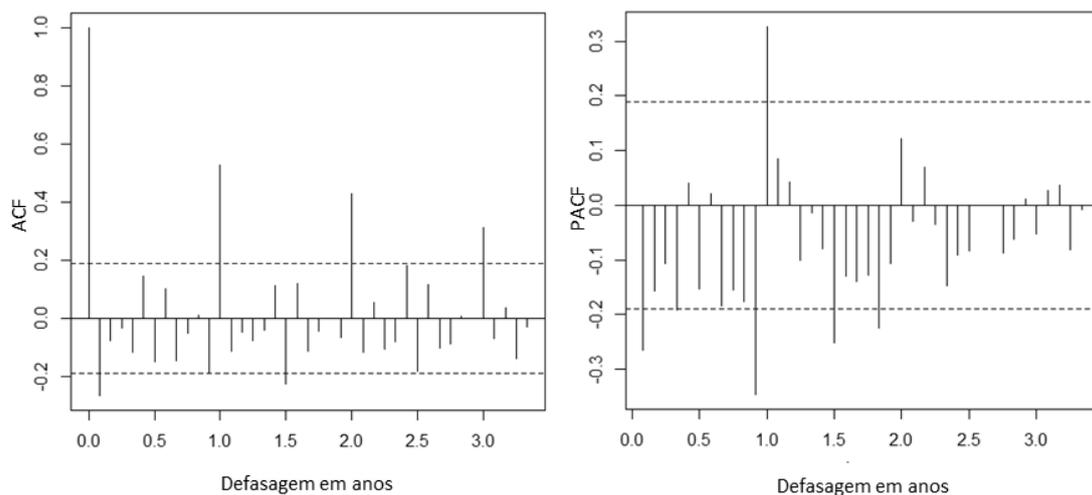
Aplicando o teste ADF à 1ª diferença da série {1ª Dif [Ln (CEEN)]}, observou-se que para os níveis de significâncias estudados, o valor de t foi maior do que qualquer um dos valores críticos, logo, pode-se afirmar que não tem raiz unitária, desta forma, afirmando que a 1ª Dif [Ln(CEEN)] é estacionária. Já o teste KPSS afirma que a hipótese nula é a não existência de raiz unitária, desta forma, a [LN(CEEN)] apresentou  $t = 2,1822$ , maior do que qualquer um dos valores críticos do teste.

Isso evidenciou a rejeição da hipótese nula, mostrando que a [LN(CEEN)] é não estacionária. Para a 1ª Dif [Ln(CEEN)] o  $t = 0,029$ , menor do que qualquer um dos valores críticos, caracterizando uma possível estacionariedade da série em primeira diferença. O teste PP da 1ª Dif [Ln(CEEN)] comprova a estacionariedade, demonstrando um  $t = -14,9495$ .

Para identificação da ordem modelo foram realizadas as Funções de Autocorrelação (ACF) e Autocorrelação Parcial (PACF) da 1ª Dif [Ln(CEEN)], de acordo com a Figura 3. Observa-se que o modelo apresenta efeitos sazonais 12 passos para frente (lag =12), identificando um SARIMA (p,d,q)(P,D,Q)<sub>12</sub>. Verificou-se pela PACF, a identificação dos termos autoregressivos (AR) de ordem 1 para a parte não sazonal e ordem 1 para a parte sazonal, ou seja, tem-se  $p = 1$  e  $P = 1$ . Já os termos de médias móveis (MA) foram identificados pela ACF. Para o termo não sazonal tem-se o

lag 2 significativo, indicando a ordem de  $q = 2$  e a ordem de MA sazonal verificou os lags 12, 24, 36 significativo, sendo  $Q = 3$ .

Identificados os modelos SARIMA  $(1,1,2)(1,1,3)_{12}$  e SARIMA  $(1,1,2)(1,0,3)_{12}$ , foram estimados os seus parâmetros. Estes modelos apresentaram coeficientes não significativos, ou seja, evidenciaram propriedades estatísticas insatisfatórias para fornecer boas previsões. Logo, foi necessário selecionar novos modelos. Seguindo o princípio da parcimônia, delimitou-se a ordem do modelo de no máximo de 3 defasagens para os processos autoregressivos [AR ( $p = 3$  e  $P = 3$ )] e de 3 defasagens para os processos de médias móveis [MA ( $q = 3$  e  $Q = 3$ )]. Foi constituído um espaço amostral de 322 modelos SARIMA  $(p,d,q) (P,D,Q)_{12}$ , dentre os quais utilizou-se o critério de informação *Akaike Information Criterion* (AIC) como fonte de seleção dos modelos.



**Figura 3:** Função de Autocorrelação (ACF) e função de autocorrelação parcial (PACF) da primeira diferenciação da série logarítmica para o consumo de energia elétrica do Nordeste, no período de 2004 a 2012..

O teste de AIC considera que quanto menor, melhor, foram pré-selecionados 250 modelos que apresentaram os maiores valores negativos. Destes selecionados, apenas 38 modelos SARIMA  $(p,d,q) (P,D,Q)_{12}$  apresentaram coeficientes significativos em que foram realizadas as verificações dos resíduos, averiguando a presença de ruído branco por meio do teste de Box-Pierce. Dos 38 modelos, apenas 14 apresentaram ruído branco, dentre os quais foram retirados cinco modelos, como mostra a Tabela 3.

O teste Box-Pierce mostrou a qualidade nos ajustes nos modelos sazonais autoregressivos integrados e de médias móveis (SARIMA), caracterizando que seus

resíduos são ruídos brancos para todos, desta forma, temo que  $Q(m) < \chi^2_\alpha$ . Com essa análise, realizou-se a previsão dos modelos selecionados na Tabela 3 e observou-se, por meio do Erro Quadrado Médio de Previsão (EQMP), qual foi o mais adequado.

**Tabela 3:** Modelos pré-selecionados para a previsão do consumo de energia elétrica do Nordeste, no período de 2004 a 2012.

SARIMA (p,d,q)(P,D,Q) <sub>s</sub>	AIC	Test Box-Pierce	
		Q (m)	$\chi^2_\alpha$
(1,1,0)(2,0,0) <sub>12</sub>	-490,296	0,224	43,1842
(1,1,0)(0,0,1) <sub>12</sub>	-461,167	0,1745	46,0089
(1,1,0)(2,1,0) <sub>12</sub>	-436,255	0,0656	50,7356
(0,1,1)(0,0,1) <sub>12</sub>	-462,804	0,1468	47,1479
(0,1,1)(0,0,2) <sub>12</sub>	-475,943	0,3367	40,0414

A Tabela 4 mostra as previsões para o consumo de energia elétrica na Região Nordeste no período de janeiro a dezembro de 2013.

**Tabela 4:** Valores observados e previsões pelos modelos SARIMA (p,d,q)(P,D,Q)<sub>12</sub> para ano de 2013 do consumo de energia elétrica da Região Nordeste.

Período	Observad	(1,1,0)(2,0,0)	(1,1,0)(0,0,1)	(1,1,0)(2,1,0)	(0,1,1)(0,0,1)	(0,1,1)(0,0,2)
		12	12	12	12	12
jan/13	5.734.052	5.475.123	5.566.361	5.418.839	5.557.895	5.540.670
fev/13	5.585.774	5.219.673	5.533.771	5.139.867	5.525.427	5.355.647
mar/13	5.774.357	5.546.559	5.625.217	5.593.355	5.615.959	5.493.994
abr/13	5.945.444	5.465.393	5.570.905	5.451.057	5.562.616	5.470.535
mai/13	5.742.036	5.467.037	5.646.226	5.464.036	5.636.394	5.492.736
jun/13	5.482.722	5.329.648	5.548.965	5.320.937	5.540.371	5.351.407
jul/13	5.597.801	5.380.688	5.528.052	5.349.799	5.520.169	5.415.111
ago/13	5.624.175	5.501.501	5.567.903	5.454.648	5.560.002	5.534.186
set/13	5.645.512	5.498.927	5.555.095	5.477.173	5.547.306	5.513.984
out/13	5.950.440	5.598.372	5.599.053	5.630.945	5.590.728	5.552.640
nov/13	5.913.715	5.591.548	5.621.415	5.632.918	5.612.658	5.533.372
dez/13	5.897.244	5.656.923	5.618.234	5.706.550	5.609.954	5.571.612
<b>Total</b>	68.893.27 2	65.731.392	66.981.195	65.640.124	66.879.477	65.825.893
<b>EQMP</b>		7,93	4,25	8,76	4,53	7,85
<b>Erro (%)</b>		-4,59%	-2,78%	-4,72%	-2,92%	-4,45%

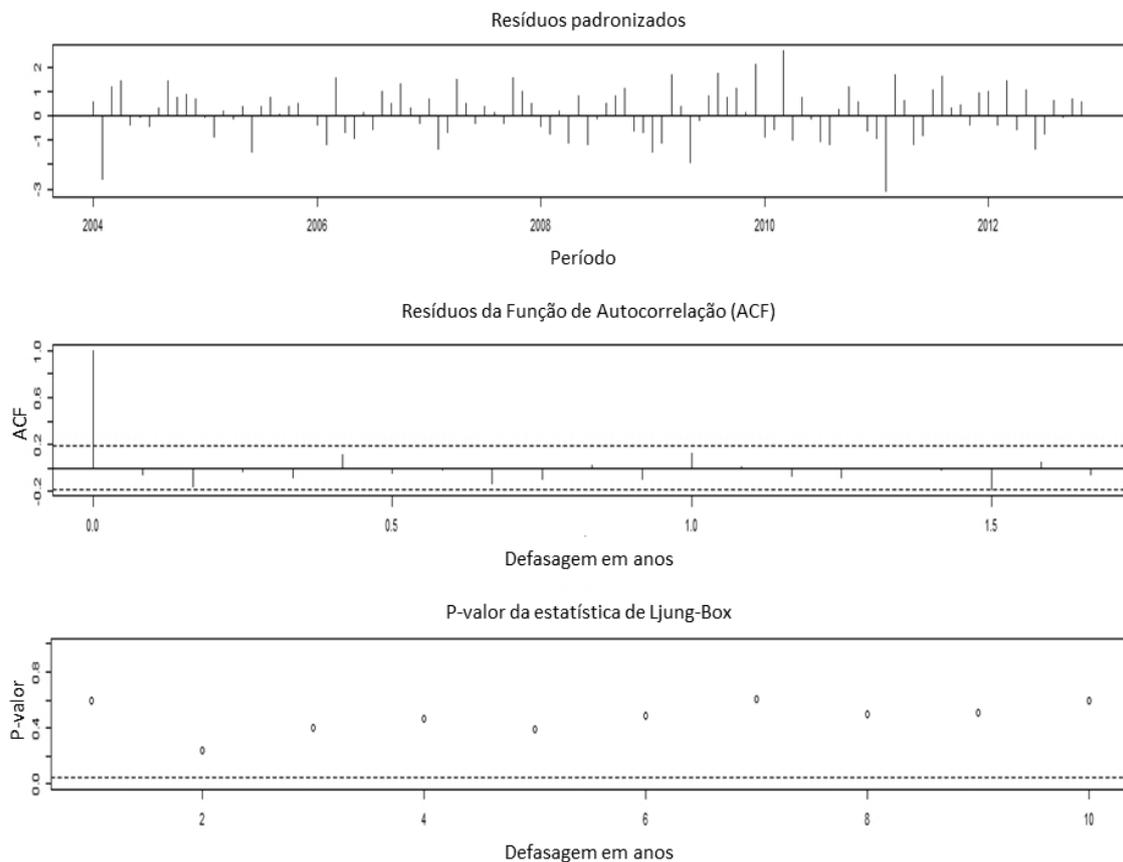
O modelo SARIMA (1,1,0) (0,0,1)<sub>12</sub> foi que apresentou menor EQMP para ano de 2013, sendo o modelo escolhido, cuja equação é apresentado da seguinte forma:

$$Y_t = \frac{(1 - \Theta_1 B^{12}) \alpha_t}{(1 - \phi_1 B)(1 - B)} \quad (17)$$

Logo, a equação do modelo SARIMA (1,1,0) (0,0,1)<sub>12</sub> com seus coeficientes assume a expressão:

$$Y_t = \frac{(1 - 0.3884^{12})\alpha_t}{(1 + 0.2390B)(1 - B)} \quad (18)$$

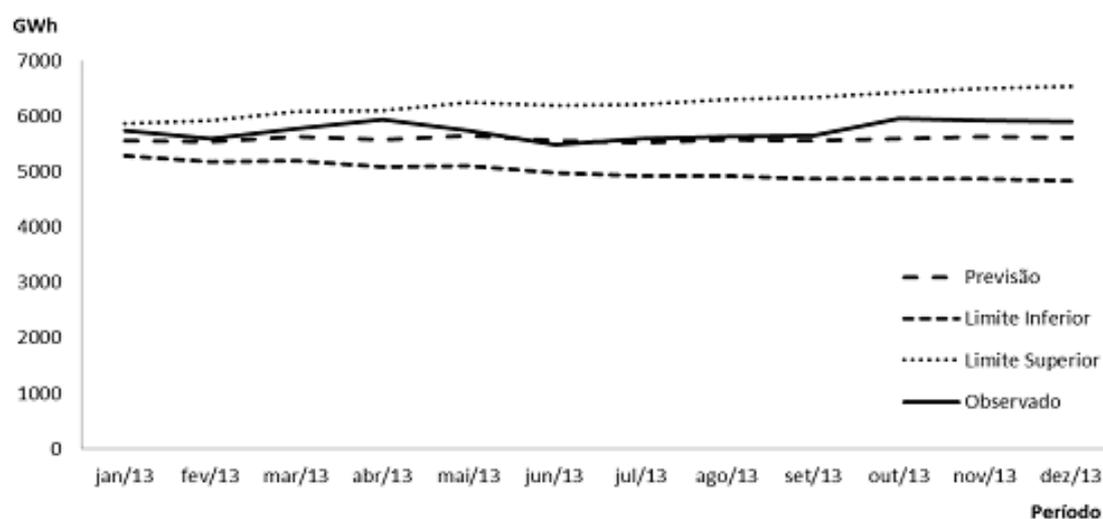
Após a estimação do modelo, faz-se necessário averiguar os resíduos da série ajustada, por meio da análise do correlograma dos resíduos. Para confirmar a presença de ruído branco no modelo SARIMA (1,1,0) (0,0,1)<sub>12</sub>, a Figura 4 apresenta os valores residuais do modelo SARIMA (1,1,0)(0,0,1)<sub>12</sub>. Além do teste Box-Pierce, o correlograma sugere independência dos resíduos para diversas defasagens, onde os limites de controle do gráfico da FAC (Função de autocorrelação) corroboram que o modelo escolhido mostrou-se adequado.



**Figura 4:** Valores residuais do modelo SARIMA (1,1,0)(0,0,1)<sub>12</sub>

Projetou os valores para 12 meses pra frente, ou seja, para o ano de 2013. A Figura 5 demonstra o comportamento da previsão, considerando um intervalo de confiança de 95%, além dos valores observados para o ano de 2013 em meses. O

modelo apresentou previsões satisfatórias, com um EQMP (Erro do Quadrado Médio de Previsão) de 4,25.



**Figura 5:** Previsão de janeiro a dezembro de 2013 para o consumo de energia elétrica no nordeste, em gigawatts hora (GWh).

#### 4. Conclusões

Com base nos resultados obtidos por este artigo, conclui-se que:

- Foi fornecida uma melhor compreensão a respeito do comportamento do consumo de energia elétrica residencial e industrial no nordeste, visando auxiliar no planejamento de políticas públicas.
- Constatou que os modelos da família ARIMA apresentaram resultados satisfatórios para a previsão do consumo de energia elétrica.
- Os modelos SARIMA  $(1,1,2)(1,1,3)_{12}$  e SARIMA  $(1,1,2)(1,0,3)_{12}$ , identificados através da ACF e PACF não apresentaram ajustes adequados, sendo necessário buscar novos modelos.
- De acordo com princípio da parcimônia, o modelo SARIMA  $(1,1,0)(0,0,1)_{12}$  foi o que forneceu a melhor previsão para o consumo de energia elétrica no nordeste brasileiro.

#### 5. Referências

BOX, G. & JENKINS, G. 1976. *Time series analysis: forecasting and control*. rev. ed., Oakland, California: Holden-Day.

BUENO, R. L. S. 2011. *Econometria de séries temporais*. São Paulo: Cengage Learning.

CAMPOS, R. J. 2008. *Previsão de séries temporais com aplicações a séries de consumo de energia elétrica*. Dissertação de Mestrado em Engenharia Elétrica. Universidade de Minas Gerais, Belo Horizonte.

COELHO JUNIOR, L. M.; REZENDE, J. L. P.; BATISTA, A. L. F.; MENDONÇA, A. R.; LACERDA, W. S. 2013. Use of artificial neural networks for prognosis of charcoal prices in Minas Gerais. *Cerne*, 19, 281-288.

COELHO JUNIOR, L. M.; REZENDE, J. L. P.; SÁFADI, T.; CALEGARIO, N. 2006. Análise temporal do preço do carvão vegetal oriundo de floresta nativa e de floresta plantada. *Scientia Forestalis*, 70, 43-53.

COELHO JUNIOR, L. M.; REZENDE, J. L. P.; SÁFADI, T.; CALEGARIO, N. 2009. Análise do comportamento temporal dos preços da borracha natural no mercado internacional. *Ciência Florestal*, 19, 293-303.

DICKEY, D. A. & FULLER, W. A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1071.

EDIGER, V. S. & AKAR, S. 2007. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 35, 1701-1708.

EPE - Empresa de Pesquisa Energética. *Anuário estatístico de energia elétrica de 2012*. Disponível em: [http://www.epe.gov.br/AnuarioEstatisticodeEnergiaEletrica/20130909\\_1.pdf](http://www.epe.gov.br/AnuarioEstatisticodeEnergiaEletrica/20130909_1.pdf). Acesso em 25 de outubro de 2016.

EPE - Empresa de Pesquisa Energética. *Banco de dados*. Disponível em [http://www.epe.gov.br/mercado/Paginas/Consumomensaldeenergiael%C3%A9tricaporclasse\(regi%C3%B5esesubistemas\)%E2%80%932011-2012.aspx](http://www.epe.gov.br/mercado/Paginas/Consumomensaldeenergiael%C3%A9tricaporclasse(regi%C3%B5esesubistemas)%E2%80%932011-2012.aspx). Acesso em 16 de junho de 2016.

FISCHER, S. *Séries univariantes de tempo metodologia de Box e Jenkins*. 1982. Porto Alegre: Fundação de Economia e Estatística.

GOMES, F. C. 1989. Os modelos ARIMA e a abordagem de Box-Jenkins uma aplicação na previsão do IBOVESPA a curtíssimo prazo. *Revista de Administração de Empresas*, 29, 63-70.

GUJARATI, D. N. 2006. *Econometria básica*. Rio de Janeiro: Elsevier.

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P. & SHIN, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159-178.

LEE, H. J. & LEE, J. K. 2005. An effective customization procedure with configurable standard models. *Decision Support Systems*, v.41, pp. 262-278.

LOBO, R. M. P. O. 1991. *Modelos para previsão de cargas a curto prazos aplicados ao sistema elétrico do Paraná*. Dissertação de Mestrado em Engenharias de Sistemas. Universidade Estadual de Campinas, São Paulo.

MORETTIN, P. A. & TOLOI, C. M. C. 2006. *Análise de séries temporais*. São Paulo: Edgard Blucher.

LAGO NETO, J. C., CORREIA, J. C., PARENTE, R. C. P. & PEREIRA, J. R. G. 2006. Modelo de projeção para o consumo de energia elétrica e óleo diesel no interior do estado do Amazonas. *Revista brasileira de energia*, 12, 1-9.

PHILLIPS, P. C. B., & PERRON, P. 1988. Testing for a unit root in time series regression. *Biometrika*, 75, 335–346.

R DEVELOPMENT CORE TEAM. 2014. *Software for computing and graphics*. Disponível em <<http://www.R-project.org>>. Acesso em 16 de junho de 2016.

REZENDE, J. L. P.; OLIVEIRA, A. D.; COELHO JUNIOR, L. M. 2005. Análise do preço do carvão em quatro regiões de Minas Gerais. *Cerne*, 11, 318-335,

SLACK, N., CHAMBERS, S. & JOHNSTON, R. 2009. *Administração da produção*. 3. ed. São Paulo: Atlas.

TORGO. L. 2009. *A linguagem em R: Programação para análise de dados*. Lisboa: Escolar editor.

TSAY, R. S. 2005. *Analysis of financial time series*. 2. ed. New Jersey: John Wiley.

ZANDONADE, E. 1993. *Aplicação da metodologia de redes neurais em previsão de séries temporais*. Dissertação de Mestrado em Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.