

## **Comparação da capacidade preditiva de modelos ARIMA para o consumo de energia elétrica no Brasil**

### ***Comparison of ARIMA models predictive capacity for the consumption of electricity in Brazil***

#### **Raphael Silveira Amaro**

Universidade Federal de Santa Maria - UFSM

[r.s.a1234@hotmail.com](mailto:r.s.a1234@hotmail.com)

Mestre em Administração pela Universidade Federal de Santa Maria (UFSM). Especialista em Gestão Financeira, Controladoria e Auditoria pela Fundação Getúlio Vargas (FGV). Graduado em Economia pela Universidade Federal de Santa Maria (UFSM).

#### **Paulo Sergio Ceretta**

Universidade Federal de Santa Maria - UFSM

[ceretta10@gmail.com](mailto:ceretta10@gmail.com)

Doutor em Engenharia de Produção pela Universidade Federal de Santa Catarina (UFSC). Professor associado do departamento de Ciências Administrativas da Universidade Federal de Santa Maria (UFSM).

#### **Daniel Arruda Coronel**

Universidade Federal de Santa Maria - UFSM

[daniel.coronel@uol.com.br](mailto:daniel.coronel@uol.com.br)

Doutor em Economia Aplicada pela Universidade Federal de Viçosa (UFV). Professor associado do departamento de Ciências Administrativas da Universidade Federal de Santa Maria (UFSM).

#### **Reisoli Bender Filho**

Universidade Federal de Santa Maria - UFSM

[reisolibender@yahoo.com.br](mailto:reisolibender@yahoo.com.br)

Doutor em Economia Aplicada pela Universidade Federal de Viçosa (UFV). Professor associado do departamento de Ciências Administrativas da Universidade Federal de Santa Maria (UFSM).

#### **Michele Gendelsky de Oliveira**

Centro Universitário Franciscano - UNIFRA

[michelygendelsky@gmail.com](mailto:michelygendelsky@gmail.com)

Graduação em Ciências Contábeis.

## RESUMO

Devido ao fato de que a energia elétrica, em grandes quantidades, não pode ser armazenada de forma viável e eficiente, torna-se extremamente importante mensurar de forma precisa e confiável a sua demanda futura. Deste modo, o objetivo desta pesquisa é comparar a capacidade preditiva de modelos *Autoregressive Integrated Moving Average* para a série de demanda de energia elétrica de médio prazo do Brasil. Para que tal objetivo seja alcançado, realizam-se previsões da demanda de energia elétrica mensal para 50 passos à frente, em cada modelo concorrente, com reajuste dos parâmetros a cada passo. Como resultado, as evidências empíricas encontradas demonstraram que o melhor modelo para realizar previsões eficientes e confiáveis para a série é um processo com dois vetores autorregressivos, um vetor de média móvel e ordem de integração unitária.

**Palavras-chave:** Energia elétrica. Modelos ARIMA. Capacidade preditiva.

## ABSTRACT

Due to the fact that large quantities of electrical energy can not be stored in a viable and efficient way, it is extremely important to accurately and reliably measure their future demand. Thus, the objective of this research is to compare the predictive capacity of *Autoregressive Integrated Moving Average* models for the series of medium-term electric energy demand in Brazil. For this objective to be achieved, monthly electricity demand forecasts are made for 50 steps forward, in each concurrent model, with readjustment of the parameters at each step. As a result, the empirical evidence found has shown that the best model to perform efficient and reliable predictions for the series is a process with two autoregressive vectors, one moving average vector and unit integration order.

**Keywords:** Electricity. ARIMA models. Predictive ability.

## 1 INTRODUÇÃO

A energia elétrica desempenha um papel cada vez mais importante na economia mundial. O desenvolvimento da sociedade e o crescimento das atividades econômicas impactam de forma expressiva a demanda por eletricidade, como recurso energético. As empresas utilizam a energia elétrica como um fator de produção, enquanto o bem-estar das famílias dependem, direta ou indiretamente, do consumo de eletricidade (KHEIRKHAH et al., 2013). Em vista de sua importância, o mercado de energia elétrica necessita de um planejamento de abastecimento de eletricidade, para que ocorra uma gestão eficiente dos sistemas de geração de energia e de tomadas de decisões pontuais relativas à otimização de aumentos da capacidade de geração atual (TAYLOR; MENEZES; MCSHARRY, 2006).

Compreender e mensurar de maneira correta a demanda de energia elétrica é vital para a sobrevivência das empresas que a produzem, a entregam e a revendem, pois a eletricidade, em grandes quantidades, não pode ser armazenada de forma viável e eficiente e,

consequentemente, a sua demanda deve ser quase que instantaneamente atendida (VILAR; CAO; ANEIROS, 2012). Tanto a superestimação quanto a subestimação da geração de energia elétrica geram custos para os agentes desse mercado e devem ser evitadas. Um planejamento eficiente sobre o comportamento da demanda de energia elétrica é indispensável para que os agentes possam desenvolver estratégias voltadas para a maximização de lucro e minimização de riscos (UNSIHUAY *et al.*, 2010).

A forma de previsão da demanda por energia elétrica depende da precisão exigida e do tipo de planejamento que se deseja realizar. A previsão de energia de curto prazo é uma ferramenta fundamental para controlar e programar os sistemas de energia, uma vez que fornece um algoritmo de escalonamento para a geração e transmissão de energia elétrica para as próximas horas ou para os próximos dias. Este mecanismo ajuda a determinar quais dispositivos devem funcionar em um determinado período, de modo a minimizar os custos do sistema e assegurar a demanda mesmo se ocorrer alguma falha no sistema local (TAYLOR; MENEZES; MCSHARRY, 2006). Outros horizontes de previsão, como os de médio e o de longo prazo, podem ser bem interessantes, pois possibilitam o planejamento da manutenção de redes elétricas e fornecem informações valiosas sobre as necessidades futuras de energia do mercado, ajudando os agentes na tomada de decisões, na redução de riscos financeiros e no cálculo sobre a viabilidade de novos investimentos no setor elétrico (ROMERA; MORÁN; FERNÁNDEZ, 2008).

A previsão da demanda de energia de curto prazo é complexa de ser estimada, pois envolve fatores sociais e meteorológicos, tais como: a temperatura, a umidade, os eventos esporádicos, a hora do dia, o dia da semana e também o seu consumo passado. Percebe-se que uma grande quantidade de dados e de informações deve ser utilizada para que se possa obter previsões precisas e confiáveis sobre a demanda de curto prazo. No entanto, quando se utilizam períodos de previsões mais extensos, como a previsão de médio e de longo prazo, a influência dos fatores citados anteriormente é diluída, pois os picos e os vales específicos encontrados na série de curto prazo são amenizados ao longo do tempo (ROMERA; MORÁN; FERNÁNDEZ, 2008). Com isso, modelos univariados, que utilizam apenas o consumo passado de energia, surgem como uma ferramenta que pode ser utilizada para prever de forma eficiente o futuro nível de energia elétrica que será demandado.

Conhecer a previsão da demanda de energia elétrica de um período mais extenso do que o de curto prazo é vital para o progresso do desenvolvimento econômico do país, uma vez que os agentes de mercado precisam dessa informação para mapear o perfil de consumo de

eletricidade e, com isso, dimensionar recursos para investimentos nos setores críticos do processo, diminuindo o risco de desequilíbrio entre a oferta e a demanda de energia elétrica.

Na tentativa de contribuir com novos estudos sobre este tema, o objetivo deste trabalho é comparar a capacidade preditiva de modelos *Autoregressive Integrated Moving Average* (ARIMA) para a série de demanda de energia elétrica de médio prazo do Brasil. Para que tal objetivo seja alcançado, realizam-se previsões da demanda de energia elétrica mensal para 50 passos à frente, em cada modelo concorrente, com reajuste dos parâmetros a cada passo. Uma vez que todas as previsões foram computadas para todos os passos à frente, é possível realizar a comparação da capacidade preditiva dos modelos concorrentes estudados por meio de uma determinada função de perda específica.

Este estudo contribui para a literatura existente pois fornece uma importante ferramenta de auxílio para que se possa calcular uma medida mais precisa da demanda de energia elétrica para o Brasil. As evidências encontradas podem servir para que as empresas minimizem custos de produção e perdas de energia nos sistemas de geração, de transmissão e de distribuição. Uma maior eficiência no planejamento sobre o comportamento da demanda de energia elétrica pode proporcionar para as empresas um incremento nas suas margens de lucro, que pode viabilizar novos investimentos no setor e elevar a otimização dos níveis de qualidade do serviço prestado. Ademais, conhecer o comportamento futuro da demanda de eletricidade pode ser útil para prevenir crises energéticas, assim como pode servir para auxiliar as empresas ou os governantes do país no cálculo sobre a viabilidade de investimentos em fontes alternativas de geração de energia.

Além desta introdução, este artigo está organizado da seguinte forma: a seção 2, expõe alguns estudos empíricos realizados anteriormente; a seção 3, apresenta, em detalhes, o método a ser utilizado; a seção 4, fornece as evidências empíricas encontradas; e por fim, a seção 5, apresenta as considerações finais sobre os resultados.

## **2 ESTUDOS EMPÍRICOS ANTERIORES**

O conhecimento sobre a previsão da demanda de energia elétrica é fundamental para o desenvolvimento do setor elétrico de um determinado país. As empresas de geração, de transmissão, de distribuição e de comercialização precisam conhecer o comportamento da demanda de eletricidade de modo a minimizar seus custos de produção e aumentar a sua lucratividade. O intuito é possibilitar a manutenção da qualidade no fornecimento de energia e proporcionar planejamentos no setor sobre a viabilidade de novos investimentos, como, por

exemplo, de fontes energéticas alternativas. Deste modo, estudos empíricos sobre este tema vêm crescendo dentro do campo profissional e acadêmico. Nesta seção, apresenta-se uma síntese de trabalhos sobre previsão da demanda de energia elétrica nos últimos anos.

Em 2013, Godinho, Milani e Pereira realizaram uma análise da série de consumo de energia elétrica na região Sudeste do Brasil, no período compreendido entre janeiro de 2002 e janeiro de 2003, com o intuito de encontrar os graus dos polinômios do modelo ARIMA que melhor se ajusta aos dados dentro da amostra. Como resultado, os autores verificaram que o modelo que melhor se ajustou aos dados foi o modelo ARIMA(1,1,1).

Em 2015, Passos realizou uma análise da série anual de consumo residencial de energia elétrica no Brasil, no período compreendido entre 1963 e 2012. Tendo como objetivo ajustar um modelo ARIMA e empregar o método de suavização exponencial para posteriormente comparar os métodos na obtenção do melhor modelo e na obtenção de previsões futuras. Como resultado, os autores verificaram que o modelo que melhor se ajustou aos dados foi o modelo ARIMA(0,1,1).

Em 2007, Ediger e Akar apresentaram um estudo sobre a previsão da demanda futura de energia elétrica da Turquia para o período entre os anos de 2005 e 2020. Os autores utilizaram modelos ARIMA e *Seasonal* ARIMA (SARIMA) para prever, individualmente, as fontes de energia de carvão, de gás natural e solar. Ademais, também utilizaram os modelos para prever a demanda total de energia elétrica do país. As evidências empíricas encontradas sugerem que o modelo ARIMA é superior ao modelo SARIMA para realizar previsões da demanda total de energia e é mais confiável do que o somatório de cada previsão individual das fontes de energia elétrica utilizando modelos ARIMA e SARIMA.

Em 2011, Kandanand comparou a capacidade preditiva de três modelos distintos: ARIMA, *Artificial Neural Network* (ANN) e *Multiple Linear Regression* (MLR), para prever a demanda de energia elétrica da Tailândia com base nos dados históricos entre os anos de 1986 e 2010. Os resultados encontrados mostraram que não existe diferença estatisticamente significativa entre os erros de previsão obtidos pelos três modelos concorrentes. Deste modo, de acordo com o princípio da parcimônia, o autor sugere que os modelos ARIMA e MLR sejam preferíveis ao modelo ANN devido às suas estruturas simplificadas de cálculo.

Inúmeros estudos sobre diferentes métodos de previsão da demanda de energia elétrica foram realizados ao longo da última década. Vilar, Cao e Aneiros (2012) previram o preço e a demanda de eletricidade, utilizando uma regressão não-paramétrica com dados explicativos funcionais e um modelo linear parcial semi-funcional, para um dia à frente dentro do mercado de eletricidade da Espanha, no período compreendido entre os anos de 2008 e 2009. Kavasseri

e Seetharaman (2009) realizaram um estudo de previsão, utilizando modelos *fractional-ARIMA*, para um e para dois dias à frente, da velocidade do vento para geração de energia eólica dentro do estado da Dakota do Norte nos Estados Unidos. Em 2013, Kheirkhah et al. propuseram um novo algoritmo para realizar previsões que utiliza as seguintes abordagens econométricas: ANN, *Principal Component Analysis* (PCA), *Data Envelopment Analysis* (DEA) e *ANOVA method*. Este algoritmo foi empregado para estimar e prever a demanda de energia elétrica do Irã, utilizando os dados históricos de consumo elétrico iraniano do período compreendendo entre abril de 1992 e fevereiro de 2004.

Assim sendo, percebe-se que existe um farto número de estudos, para prever a demanda futura de eletricidade, que utilizam o modelo univariado ARIMA. A principal justificativa para o seu uso é que este modelo apresenta uma estrutura menos complexa e mais acessível para ser calculada. Além disso, como exposto, o desempenho da previsão do modelo ARIMA mostrou-se eficiente para prever o futuro nível de energia elétrica que será demandado em períodos de médio e de longo prazo.

### 3 MÉTODO

Os processos *Autoregressive Moving Average* (ARMA) são definidos por equações diferenciais lineares com coeficientes constantes e desempenham um papel fundamental na modelagem de dados de séries temporais (EDIGER; AKAR, 2007). Uma das razões da popularidade dos processos ARMA é que para um grande número de classe de funções de autocovariância  $\gamma(\cdot)$  é possível encontrar um processo ARMA, dito  $\{X_t\}$ , com a função de autocovariância  $\gamma_x(\cdot)$ , de tal modo que  $\gamma(\cdot)$  seja bem próxima de  $\gamma_x(\cdot)$ . Esta definição indica que para qualquer número inteiro positivo  $K$  existe um processo ARMA  $\{X_t\}$ , de modo que  $\gamma_x(h) = \gamma(h)$  para  $h = 0, 1, \dots, K$ . (BROCKWELL; DAVIS, 2002). A estrutura de um processo ARMA( $p, q$ ) leva a uma simplificação dos métodos gerais para predição linear e pode ser definida pela seguinte expressão matemática:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (1)$$

$$\phi(B)X_t = \theta(B)\varepsilon_t.$$

Em que  $(\varepsilon_t) \sim RB(0, \sigma^2)$ ,  $B$  é o operador de defasagem ( $B^j X_t = X_{t-j}$ ,  $B^j \varepsilon_t = \varepsilon_{t-j}$ ,  $\forall j \geq 0$ ),  $\phi(\cdot)$  e  $\theta(\cdot)$  são os polinômios de grau  $p$  e  $q$ , representados, respectivamente, por  $\phi(\varepsilon) = 1 - \phi_1 \varepsilon - \dots - \phi_p \varepsilon^p$  e  $\theta(\varepsilon) = 1 + \theta_1 \varepsilon + \dots + \theta_q \varepsilon^q$ . Observe que caso  $\theta(\varepsilon) \equiv 1$ , a série temporal  $\{X_t\}$  segue um processo autorregressivo de ordem  $p$  e pode ser representado por  $AR(p)$ . Já na hipótese de  $\phi(\varepsilon) \equiv 1$ , a série segue um processo de média móvel de ordem  $q$  e pode ser representado por  $MA(q)$  (CHAVEZ; BERNAT; COALLA, 1999).

A definição do processo  $ARMA(p, q)$  impõe que  $\{X_t\}$  seja estacionário para que a sua variância seja não negativa e finita. Esta condição é perfeitamente atendida quando  $\phi(\varepsilon) = 1 - \phi_1 \varepsilon - \dots - \phi_p \varepsilon^p \neq 0$ , para todo  $\varepsilon$  complexo em que  $|\varepsilon| = 1$ . A utilização de  $\varepsilon$  como número complexo vem do fato de que os zeros de um polinômio de grau  $p > 1$  pode ser tanto um número real quanto complexo (BROCKWELL; DAVIS, 2002). Outra restrição imposta pelo processo  $ARMA(p, q)$  é que o modelo  $MA(q)$  seja capaz de ser expresso como um modelo  $AR(\infty)$ . Esta restrição é chamada de condição de invertibilidade e requer que  $\theta(\varepsilon) = 1 + \theta_1 \varepsilon + \dots + \theta_q \varepsilon^q \neq 0$ , para todo  $|\varepsilon| \leq 1$ . Observe que, caso esta condição seja atendida, todas as raízes de  $\theta(\varepsilon) = 0$  devem estar fora do círculo unitário (CHAVEZ; BERNAT; COALLA, 1999).

Um problema enfrentado pelo cálculo dos parâmetros do processo  $ARMA(p, q)$  ocorre quando o conjunto de dados utilizados apresenta desvios aparentes da condição de estacionariedade. Se uma série temporal, representada por um conjunto de observações  $\{X_1, \dots, X_n\}$ , seguir um processo estocástico linear não-estacionário, precisa-se transformar os dados para gerar uma nova série que possua as propriedades de estacionariedade. Este processo de transformação pode ser alcançado pelo método de diferenciação. O número de diferenças necessárias para tornar uma série estacionária é denominado ordem de integração e é representado por  $d$ . Caso  $d$  seja um número inteiro não negativo, então  $\{X_t\}$  segue um processo *Autoregressive Integrated Moving Average* (ARIMA) de ordem  $(p, d, q)$  se  $Y_t := (1 - B)^d X_t$  for um processo ARMA de ordem  $(p, q)$  (BROCKWELL; DAVIS, 2002). Esta definição indica que  $\{X_t\}$  satisfaz a seguinte equação diferencial:

$$\phi^*(B)X_t \equiv \phi(B)(1 - B)^d X_t = \theta(B)\varepsilon_t. \quad (2)$$

Em que  $B$  é o operador de defasagem,  $(\varepsilon_t) \sim RB(0, \sigma^2)$ ,  $\phi(\varepsilon)$  e  $\theta(\varepsilon)$  são polinômios de grau  $p$  e  $q$ , respectivamente, e  $\phi(\varepsilon) \neq 0$  para  $|\varepsilon| \leq 1$ . Neste caso,  $\{X_t\}$  é dito um processo estacionário se  $d = 0$  e pode ser reduzido a um processo  $ARMA(p, q)$ . Supondo que  $\{X_t\}$  seja uma série temporal Gaussiana, com média zero e função de autocovariança  $k(i, j) = E(X_i X_j)$ ,

sendo  $X_n = (X_1, \dots, X_n)'$  e  $\hat{X}_n = (\hat{X}_1, \dots, \hat{X}_n)'$ , em que  $\hat{X}_1 = 0$  e  $\hat{X}_j = E(X_j | X_1, \dots, X_{j-1})$ ,  $j \geq 2$ , as estimativas da média, da variância  $\sigma^2$  e dos coeficientes  $\{\phi_i, i = 1, \dots, p\}$  e  $\{\theta_i, i = 1, \dots, q\}$  de um modelo ARMA( $p, q$ ) podem ser obtidas pelo procedimento conhecido como estimação por máxima verossimilhança, que é representado pela seguinte expressão:

$$L(\phi, \theta, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^n r_0 \dots r_{n-1}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^n \frac{(X_j - \hat{X}_j)^2}{r_{j-1}} \right\}. \quad (3)$$

Sendo  $L(\cdot)$  a função de verossimilhança,  $n$  o número de funções de densidade individuais e  $r_n = E(X_{n+1} - \hat{X}_{n+1})^2$ . Nota-se que os parâmetros estimados são os valores que maximizam  $L(\cdot)$  para um dado conjunto de dados. A solução da maximização da função de probabilidade condicional  $L(\cdot)$  é obtida numericamente, derivando-a e igualando-a a zero. Para maiores detalhes, consultar Brockwell e Davis (2002). Uma vez estimados os parâmetros dos modelos por meio de  $L(\cdot)$ , pode-se realizar as previsões para  $h$ -passos à frente.

### 3.1 Previsão de $h$ -passos à frente utilizando modelos ARIMA( $p, d, q$ )

Dada uma série temporal estacionária com média  $\mu$  e com função de autocovariância  $\gamma$  em termos dos valores  $\{X_n, \dots, X_1\}$ , até o instante  $n$ , o intuito deste trabalho é encontrar a melhor combinação linear de  $1, X_n, X_{n-1}, \dots, X_1$  que preveja  $X_{n+h}$  com o menor erro quadrado médio. O melhor predictor linear encontrado será denotado por  $P_n X_{n+h}$  e representado pela seguinte fórmula matemática:

$$P_n X_{n+h} = \mu + \sum_{i=1}^n a_i (X_{n+1-i} - \mu). \quad (4)$$

Em que  $a$  expõe os coeficientes da equação. Deste modo, as previsões  $h$  passos à frente, para  $n > m = \max(p, q) \forall h \geq 1$ , podem ser descritas pela função  $g(h) := P_n X_{n+h}$ , que deve satisfazer a homogeneidade da equação diferencial linear:

$$g(h) - \phi_1^* g(h-1) - \dots - \phi_{p+d}^* g(h-p-d) = 0, \quad h > q. \quad (5)$$



Sendo que  $\phi_1^*, \dots, \phi_{p+d}^*$  são coeficientes de  $\varepsilon, \dots, \varepsilon^{p+d}$  em  $\varepsilon^*(z) = (1 - \varepsilon)^d \phi(\varepsilon)$ . Caso os zeros de  $\phi(\varepsilon)$  (representados por  $\omega_1, \dots, \omega_p$ ) sejam todos distintos, a solução de (5) pode ser representada por:

$$g(h) = a_0 + a_1 h + \dots + a_d h^{d-1} + b_1 \omega_1^{-h} + \dots + b_p \omega_p^{-h}, \quad h > q - p - d. \quad (6)$$

Em que os coeficientes  $a_1, \dots, a_d$  e  $b_1, \dots, b_p$  podem ser determinados por meio da equação  $p + d$ , obtida igualando o lado direito de (6) por  $q - p - d < h \leq q$  com o valor numérico calculado de  $g(h)$ . Nota-se que, para  $h \leq 0$ ,  $P_n X_{n+h} = X_{n+h}$ , e para  $1 \leq h \leq q$ ,  $P_n X_{n+h}$  pode ser calculado pela seguinte expressão:

$$P_n X_{n+h} = \sum_{j=1}^{p+d} \phi_j^* P_n X_{n+h-j} + \sum_{j=h}^q \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j}). \quad (7)$$

Conhecida a fórmula de cálculo para realizar previsões utilizando processos ARIMA, pode-se definir o método para comparar a capacidade preditiva dos modelos concorrentes, pois, não necessariamente, o modelo que oferece o melhor ajuste dentro da amostra produzirá a melhor previsão para fora da amostra (WENNSTRÖM, 2014).

### 3.2 Método para comparar a capacidade preditiva de modelos concorrentes

A série completa do consumo de energia elétrica é composta de um número  $T$  de observações e é representada por  $X_1, X_2, \dots, X_T$ . O número total de observações se divide em dois subconjuntos:  $\{X_1, X_2, \dots, X_n\}$  e  $\{X_{n+1}, X_{n+2}, \dots, X_T\}$ , em que  $n$  representa o início da previsão. O primeiro subconjunto refere-se ao período denominado dentro da amostra, em que os parâmetros dos modelos concorrentes são estimados. O segundo subconjunto refere-se ao período denominado fora da amostra, em que as previsões dos modelos concorrentes são realizadas e avaliadas. Após a divisão da amostra, o processo de previsão e de avaliação da previsão para fora da amostra se inicia com os seguintes passos:

1. Define-se  $z = n$  para ser o ponto de origem da previsão. Logo após, estima-se os parâmetros dos modelos concorrentes usando o subconjunto  $\{X_1, X_2, \dots, X_z\}$ .

2. Realiza-se a previsão para um passo à frente, para cada modelo concorrente, usando os seus parâmetros estimados no passo (1).
3. Aumenta-se a origem da previsão em uma observação ( $z = z + 1$ ) e recomeça-se o processo no passo (1). Este procedimento é repetido até que se chegue na última observação disponível da série de energia elétrica, em que a origem da previsão  $z$  seja igual ao ponto  $T$ .

Conhecidos os passos do processo de previsão, realizam-se previsões para 50 passos à frente com reajuste dos parâmetros a cada passo para todos os modelos concorrentes. A escolha de 50 observações para o período para fora da amostra estabeleceu-se de forma arbitrária. Uma vez que os modelos utilizados, para realizar previsões, referem-se aos processos ARIMA( $p, d, q$ ), os modelos concorrentes são representados pelos diferentes graus dos polinômios do processo, respeitando as limitações:  $0 \geq p \geq 3$  e  $0 \geq q \geq 3$ . Estas limitações impostas fazem com que o número total de modelos univariados concorrentes seja limitado a 15, que resulta em uma matriz com um total de 750 previsões estimadas.

Após todas as previsões serem computadas para todos os passos à frente, realiza-se a comparação da capacidade preditiva dos modelos concorrentes estudados a partir de uma medida estatística estabelecida. Neste trabalho, utilizaram-se as seguintes funções de perda: *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Mean Absolute Percentage Error* (MAPE) e *Root Mean Square Error* (RMSE). As suas expressões matemáticas são descritas a seguir:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{X}_i - X_i|, \quad MSE = \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2, \quad (8)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{X}_i - X_i}{X_i} \right|, \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2}.$$

Em que  $\hat{X}_i$  é o vetor de  $n$  previsões e  $X_i$  é o vetor dos valores observados que correspondem às entradas para a função que gerou as previsões. Percebe-se que as funções de perda utilizadas possuem como intuito avaliar a qualidade das previsões obtidas pelos modelos

concorrentes por meio do cálculo do erro das previsões. Sendo assim, o modelo que fornecer os menores valores para as funções estabelecidas possuirá a melhor capacidade preditiva.

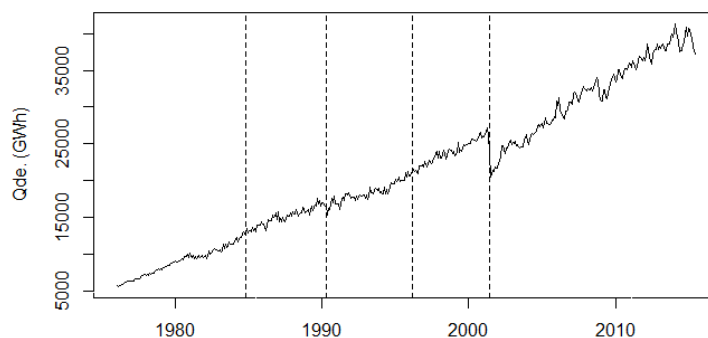
### 3.3 Fonte de dados

A base de dados utilizada neste trabalho refere-se ao consumo de energia elétrica total do Brasil, medido em Gigawatt-hora (GWh). A série de dados foi coletada no sitio do Instituto de Pesquisa Econômica Aplicada<sup>1</sup> (Ipea) e corresponde a informações mensais sobre o consumo de eletricidade no período compreendido entre janeiro de 1976 e junho de 2015, totalizando 474 observações.

## 4 EVIDÊNCIAS EMPÍRICAS

Nesta seção, as evidências empíricas da performance de previsão dos modelos concorrentes para a série de consumo de energia elétrica do Brasil são examinadas em detalhes. Inicia-se verificando a presença de quebra estrutural na série coletada por meio da realização dos seguintes testes estatísticos: *cumulative sum control chart* (CUSUM), desenvolvido por Page (1954); *moving sums of residuals* (MOSUM), proposto por Bauer e Hackl (1978); *supF test*, desenvolvido por Andrews (1993); e, por fim, os testes *aveF* e *expF*, propostos por Andrews e Ploberger (1994). Os resultados encontrados, para todos os testes aplicados, indicaram a rejeição da hipótese nula da não existência de quebras estruturais na série de consumo de energia elétrica a 1% de significância. Os pontos em que ocorrem mudanças no comportamento da evolução da série estudada são destacados na Figura 1.

Figura 1 – Evolução da série temporal de consumo de energia elétrica no Brasil e seus pontos de quebras estruturais (destacados pelas linhas verticais), no período compreendido entre janeiro de 1976 e junho de 2015.



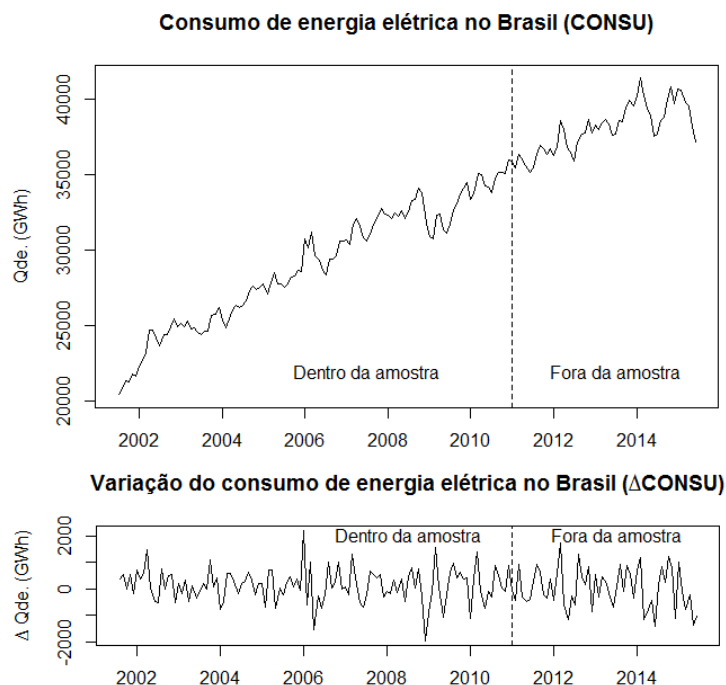
Fonte: elaboração dos autores.

<sup>1</sup> Disponível em <<http://www.ipea.gov.br>>.

A partir dos resultados encontrados pelos testes de quebras estruturais, percebe-se que a última mudança do comportamento da série analisada ocorreu em junho de 2001. A ocorrência desta quebra estrutural pode ser explicada pela crise energética, providenciada e maximizada por fatores climáticos, vivenciada pelo Brasil no ano de 2001, em que modificou a postura de consumo de eletricidade dos brasileiros (SOUZA; RODRIGUES; REIS, 2004). Deste modo, para alcançar o objetivo proposto, utiliza-se a série de consumo de energia elétrica total do Brasil (Consu) no período compreendido entre julho de 2001 e junho de 2015, totalizando 168 observações mensais.

A Figura 2 expõe a evolução da série ao longo do tempo, juntamente com a evolução da sua variação, representada pela primeira diferença da série ( $\Delta$ Consu). Nota-se que a série Consu apresenta uma trajetória ascendente ao longo do período analisado, mostrando que o consumo de eletricidade no Brasil vem aumentando de forma gradual e consistente. Entre julho de 2001 e junho de 2015, a demanda de energia teve um crescimento expressivo, totalizando 56,7% no acumulado do período. A evolução deste crescimento, de certa forma, mostrou-se estável ao longo do tempo, exibindo um comportamento sem grandes oscilações no período analisado. A variação mais significativa da série encontra-se em dezembro de 2008, em que o vale deixado pelo gráfico reflete, possivelmente, algum efeito relativo à crise financeira mundial. Embora tenha ocorrido uma queda do consumo de eletricidade no final do ano de 2008, o impacto da oscilação não foi suficiente para provocar uma mudança expressiva no comportamento da série analisada, pois a tendência do consumo de eletricidade continuou ascendente dentro de um canal de alta sem expressar nenhuma quebra estrutural na série.

Figura 2 – Evolução do consumo de energia elétrica no Brasil, em Gigawatt-hora (GWh), juntamente com a evolução da sua primeira diferença, no período compreendido entre julho de 2001 e junho de 2015. A linha vertical pontilhada diferencia o período dentro da amostra (lado esquerdo) do período fora da amostra (lado direito) para 50 passos à frente.



Fonte: elaboração dos autores.

As estatísticas descritivas da série são apresentadas na Tabela 1. Como esperado, percebem-se evidências de normalidade, pois a série não possui excesso de curtose e não apresenta assimetria elevada. Este resultado demonstra que a função de distribuição de probabilidade do *Consu* possui o mesmo achatamento que uma distribuição Gaussiana. Além disso, o teste *Jarque-Bera* aceita a hipótese nula de normalidade para a série, tanto no período dentro da amostra quanto no período fora da amostra, confirmando que a sua função de distribuição de probabilidade é mesocúrtica.

Tabela 1 – Estatísticas descritivas para a série de consumo de energia elétrica no Brasil, em Gigawatt-hora (GWh), para o período compreendido entre julho de 2001 e junho de 2015.

| Estatísticas       | 50 passos à frente |                 |
|--------------------|--------------------|-----------------|
|                    | Dentro da amostra  | Fora da amostra |
| <b>Consu</b>       |                    |                 |
| Observações        | 118                | 50              |
| Média              | 29.228,36          | 38.196,86       |
| Mediana            | 29.528,00          | 38.253,50       |
| Máximo             | 36.352,00          | 41.383,00       |
| Mínimo             | 20.447,00          | 35.161,00       |
| Desvio Padrão      | 4.081,00           | 1.511,88        |
| Assimetria         | -0,19              | 0,02            |
| Excesso de Curtose | 2,01               | 2,30            |
| Jarque-Bera        | 5,55               | 1,02            |

Fonte: elaboração dos autores.

Quando se trabalha com modelagem de série temporal, precisa-se analisar se a série é estacionária, pois, caso não o seja, torna-se impossível realizar inferências estatísticas sobre ela. Portanto, com o intuito de verificar se as séries *Consu* e  $\Delta$ *Consu* possuem estacionariedade, aplicou-se o teste *Augmented Dickey–Fuller* (ADF), desenvolvido por Dickey e Fuller (1979), e, também, o teste *Kwiatkowski–Phillips–Schmidt–Shin* (KPSS), desenvolvido por Kwiatkowski, Phillips, Schmidt e Shin (1992). Os resultados destes dois testes mostraram que a série é integrada de ordem um,  $I(1)$ , pois a série *Consu* mostrou-se não estacionária a 1% de significância, apresentando valores para a estatística ADF de -3,7886 e para a estatística KPSS de 5,5411. Enquanto a sua primeira diferença,  $\Delta$ *Consu*, mostrou-se estacionária a 1% de significância, apresentando valores para a estatística ADF de -6,2501 e para a estatística KPSS de 0,1478.

Uma vez que o comportamento da série foi analisado, inicia-se o processo descrito na seção 3 para obter as previsões dos processos ARIMA e comparar a capacidade preditiva dos modelos concorrentes. A Tabela 2 apresenta os resultados encontrados para as quatro funções de perda calculadas por meio das 750 previsões realizadas pelos 15 modelos concorrentes estabelecidos.

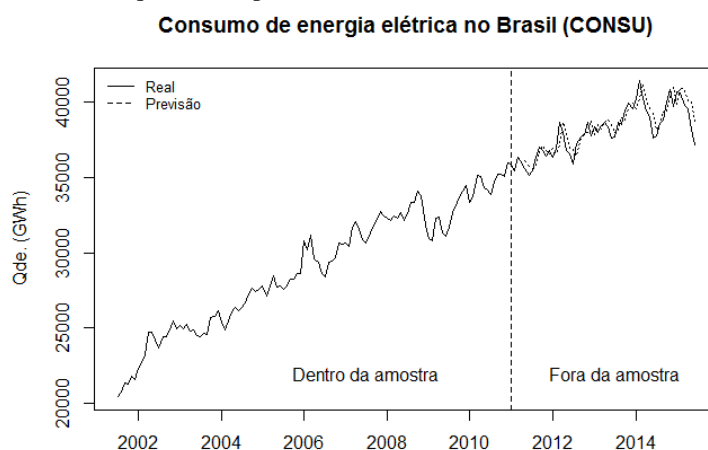
Tabela 2 – Resultados obtidos pelas quatro funções de perda utilizadas para comparar a capacidade preditiva das 750 previsões realizadas pelos 15 modelos concorrentes.

| Modelos Concorrentes | RMSE        |         | MSE         |         | MAE         |         | MAPE        |         |
|----------------------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|
|                      | Estatística | Posição | Estatística | Posição | Estatística | Posição | Estatística | Posição |
| ARIMA(2,1,1)         | 758,27      | 1       | 574.977,32  | 1       | 630,44      | 1       | 0,01648     | 1       |
| ARIMA(3,1,2)         | 758,64      | 2       | 575.533,31  | 2       | 632,67      | 3       | 0,01654     | 3       |
| ARIMA(1,1,2)         | 759,70      | 3       | 577.142,15  | 3       | 631,46      | 2       | 0,01651     | 2       |
| ARIMA(2,1,2)         | 762,37      | 4       | 581.212,08  | 4       | 634,03      | 4       | 0,01658     | 4       |
| ARIMA(3,1,1)         | 766,31      | 5       | 587.231,77  | 5       | 638,40      | 6       | 0,01669     | 6       |
| ARIMA(2,1,3)         | 770,78      | 6       | 594.107,93  | 6       | 639,93      | 8       | 0,01672     | 8       |
| ARIMA(1,1,3)         | 773,76      | 7       | 598.703,20  | 7       | 636,80      | 5       | 0,01662     | 5       |
| ARI(2)               | 774,80      | 8       | 600.310,58  | 8       | 661,36      | 13      | 0,01727     | 13      |
| ARI(1)               | 774,93      | 9       | 600.519,39  | 9       | 658,07      | 11      | 0,01719     | 11      |
| IMA(1)               | 775,90      | 10      | 602.026,50  | 10      | 658,85      | 12      | 0,01721     | 12      |
| ARIMA(3,1,3)         | 776,83      | 11      | 603.472,38  | 11      | 638,87      | 7       | 0,01670     | 7       |
| IMA(2)               | 779,41      | 12      | 607.486,94  | 12      | 657,57      | 10      | 0,01718     | 10      |
| IMA(3)               | 799,11      | 13      | 638.570,88  | 13      | 654,15      | 9       | 0,01708     | 9       |
| ARI(3)               | 804,63      | 14      | 647.430,63  | 14      | 663,17      | 14      | 0,01728     | 14      |
| ARIMA(1,1,1)         | 805,71      | 15      | 649.164,33  | 15      | 668,92      | 15      | 0,01748     | 15      |

Fonte: elaboração dos autores.

Embora os valores encontrados pelas funções de perda diverjam em algumas posições, o modelo referente ao processo ARIMA(2,1,1) demonstrou ter a melhor capacidade preditiva para prever a série Consu, pois apresentou menores valores para todas as quatro diferentes funções de perda utilizadas. A Figura 3 expõe o comportamento da evolução do Consu, juntamente com as 50 previsões obtidas com o processo ARIMA(2,1,1).

Figura 3 - Evolução do consumo de energia elétrica no Brasil, em Gigawatt-hora (GWh), no período compreendido entre julho de 2001 e junho de 2015, juntamente com as previsões obtidas com o processo ARIMA(2,1,1) para 50 passos à frente, com reajuste a cada passo, para o período fora da amostra compreendido entre maio de 2011 e junho de 2015. A linha vertical pontilhada diferencia o período dentro da amostra (lado esquerdo) do período fora da amostra (lado direito).

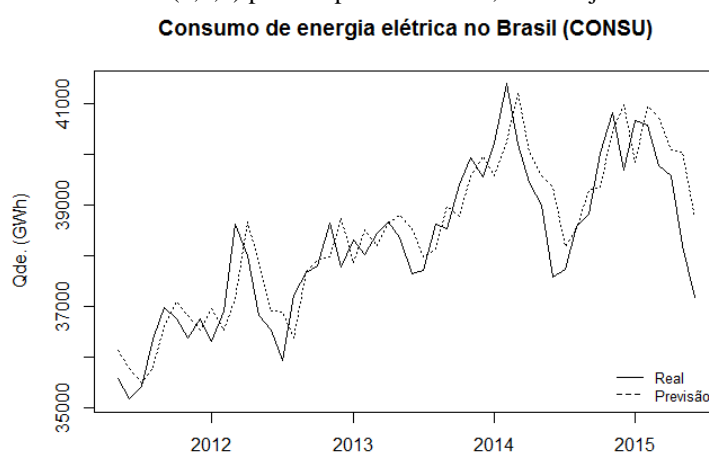


Fonte: elaboração dos autores.

Os graus dos polinômios do processo ARIMA encontrados para o modelo que representa a melhor capacidade preditiva diferem-se dos encontrados por Passos (2015) e Godinho, Milani e Pereira (2013), que foram, respectivamente, (0,1,1) e (1,1,1). Esta divergência ocorre devido as diferentes metodologias empregadas para a seleção do melhor modelo. Os autores citados não realizaram previsões para períodos para fora da amostra e sim, escolheram os graus dos polinômios do processo ARIMA através da estimação dos parâmetros que melhor se ajustaram aos dados no período dentro da amostra. Entretanto, como já dito anteriormente, o modelo que oferece o melhor ajuste dentro da amostra, não necessariamente, produzirá a melhor previsão para fora da amostra (WENNSTRÖM, 2014).

O procedimento de previsões realizado neste trabalho evidencia de fato o modelo que possui a melhor capacidade preditiva, uma vez que realiza 50 previsões para cada modelo, com reajuste a cada passo, e aponta o modelo em que as previsões mais se aproximam dos valores reais da série temporal. Percebe-se que as previsões obtidas pelo processo ARIMA(2,1,1) são muito próximas dos valores reais da série do consumo de energia elétrica, evidenciando o bom desempenho de previsão do modelo proposto. A Figura 4 evidencia, em maiores detalhes, a evolução do Consu dentro do período fora da amostra, juntamente com as 50 previsões obtidas com o processo ARIMA(2,1,1).

Figura 4 - Evolução do consumo de energia elétrica no Brasil, em Gigawatt-hora (GWh), para o período fora da amostra compreendido entre maio de 2011 e junho de 2015, juntamente com as previsões obtidas com o processo ARIMA(2,1,1) para 50 passos à frente, com reajuste a cada passo.



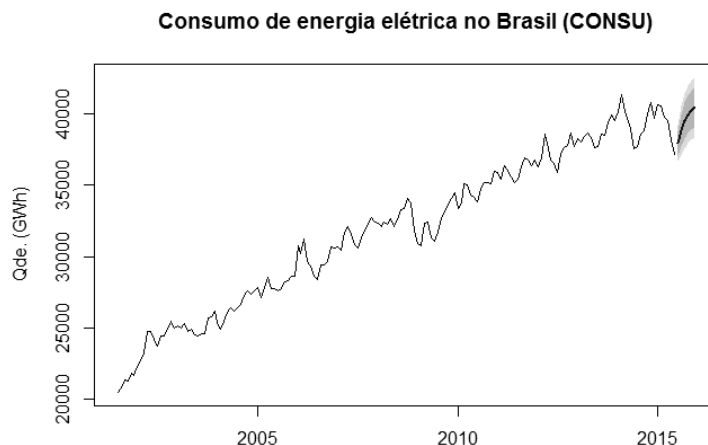
Fonte: elaboração dos autores.

Uma vez que se descobre os graus dos polinômios do processo ARIMA que oferecem o melhor modelo, pode-se utilizá-los para calcular uma medida mais precisa da demanda futura de eletricidade para o Brasil. Um exemplo gráfico pode ser visto na Figura 5, em que se



mensura, por meio do processo ARIMA(2,1,1), as previsões da série Consu para os seis próximos meses do ano de 2015.

Figura 5 - Evolução do consumo de energia elétrica no Brasil, em Gigawatt-hora (GWh), no período compreendido entre julho de 2001 e junho de 2015, juntamente com as previsões obtidas com o processo ARIMA(2,1,1) para 6 passos à frente, sem reajuste a cada passo.



Fonte: elaboração dos autores.

A Figura 5 demonstra que a tendência de alta do consumo de energia elétrica tende a continuar válida para os próximos meses de 2015, embora os últimos meses observados, antes da realização da previsão, tenham apresentado variações negativas. Um ponto a acrescentar sobre o exemplo gráfico, exposto na Figura 5, é que as previsões calculadas referem-se a 6 passos à frente sem reajuste dos parâmetros a cada passo, o que diminui a eficiência do processo de previsão a cada passo à frente calculado. Para que o modelo indicado ofereça previsões confiáveis e eficientes, o ideal é que sejam realizadas previsões somente para um passo à frente e depois reajustar o processo, como praticado ao longo desse trabalho.

## 5 CONSIDERAÇÕES FINAIS

Este trabalho buscou realizar uma comparação da capacidade preditiva de modelos ARIMA( $p,d,q$ ) para prever a série de consumo de energia elétrica de médio prazo do Brasil, utilizando, para isto, 50 previsões com reajuste dos parâmetros a cada passo para cada modelo concorrente. Como resultado, as evidências empíricas encontradas demonstraram que o melhor modelo para realizar previsões eficientes e confiáveis para a série é o processo ARIMA(2,1,1).

Os graus dos polinômios encontrados oferecem um caminho para calcular uma medida mais precisa da demanda futura de eletricidade para o Brasil, que é vital para o progresso do desenvolvimento econômico do país. O modelo proposto pode servir para que as empresas do

mercado energético minimizem custos de produção e perdas de energia nos sistemas de geração, de transmissão e de distribuição. Além disso, aumentar a eficiência no planejamento deste mercado, pode proporcionar para as empresas um incremento nas suas margens de lucro, que pode viabilizar novos investimentos no setor e elevar a otimização dos níveis de qualidade do serviço prestado.

Por outro lado, o modelo proposto pode ser utilizado para mapear o perfil de consumo de eletricidade do país e, com isso, dimensionar recursos para investimentos nos setores críticos do processo, diminuindo o risco de desequilíbrio entre a oferta e a demanda de energia elétrica. Desta forma, o modelo pode ser útil para prevenir crises energéticas ocasionadas pelo excesso não esperado de demanda, assim como, pode servir para auxiliar as empresas ou os governantes do país no cálculo sobre a viabilidade de investimentos em fontes alternativas de geração de energia.

Este trabalho, ao analisar o horizonte de tempo mensal, de médio prazo, teve como intuito encontrar uma ferramenta para auxiliar os agentes de mercado no planejamento energético futuro e, com isso, tornar a gestão dos sistemas energéticos mais eficiente. Como alternativa, sugere-se que sejam realizados estudos de previsão para outros horizontes temporais, como, por exemplo, os de curto prazo. A previsão de curto prazo oferece um algoritmo de escalonamento para a geração e transmissão de energia elétrica para as próximas horas ou para os próximos dias, tornando o seu conhecimento essencial para controlar e programar os sistemas de energia de modo a minimizar custos e assegurar a demanda imediata de energia.

A previsão de demanda de energia de curto prazo é mais complexa de ser estimada do que a previsão para prazos mais extensos, pois envolve fatores sociais e meteorológicos que em horizontes temporais longos são diluídos ao longo do tempo. Portanto, propõem-se a utilização e a comparação de outros métodos estatísticos alternativos que possam capturar comportamentos distintos aos observados somente pelos valores passados e pelos termos de erros estocásticos fornecidos pelos processos ARIMA.

## REFERÊNCIAS

ANDREWS, D. W. K., Tests for parameter instability and structural change with unknown change point. **Econometrica**, v. 61, n. 4, p. 821-856, July 1993.

ANDREWS, D. W. K.; PLOBERGER, W., Optimal tests when a nuisance parameter is present only under the alternative. **Econometrica**, v. 62, n. 6, p. 1383-1414, Nov. 1994.

- BAUER, P.; HACKL, P., The use of Mosums for quality control. **Technometrics**, v. 20, n. 4, p. 431-436, Nov. 1978.
- BROCKWELL, P. J.; DAVIS, R. A., **Introduction to Time Series and Forecasting**. 2. ed., Springer-Verlang, 2002.
- CHAVEZ, S. G.; BERNAT, J. X.; COALLA, H. L., Forecasting of energy production and consumption in Asturias (northern Spain). **Energy**, v. 24, n. 3, p. 183-198, Mar. 1999.
- DICKEY, D. A.; W.A. FULLER., Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, v. 74, p. 427-431, 1979.
- EDIGER, V. Ş.; AKAR, S., ARIMA forecasting of primary energy demand by fuel in Turkey. **Energy Policy**, v. 35, n. 3, p. 1701-1708, Mar. 2007.
- GODINHO, T. S.; MILANI, L. L.; PEREIRA, G. A., Análise e previsão do consumo de energia elétrica da região sudeste usando a metodologia de Box e Jenkins. **Anais do XII Encontro Mineiro de Estatística - MGEST 2013**. Uberlândia - 05 e 06 de Setembro de 2013.
- Ipea - **Instituto de Pesquisa Econômica Aplicada**, 2015. Disponível em: <<http://www.ipea.gov.br/>>. Acesso em: Mai. 2015.
- KANDANANOND, K., Forecasting electricity demand in thailand with an artificial neural network approach. **Energies**, v. 8, n. 4, p. 1246-1257, Aug. 2011.
- KAVASSERI, R. G.; SEETHARAMAN, K., Day-ahead wind speed forecasting using f-ARIMA models. **Renewable Energy**, v. 34, n. 5, p. 1388-1393, May. 2009.
- KHEIRKHAH, A.; AZADEH, A.; SABERI, M.; AZARON, A.; SHAKOURI, H., Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis. **Computers & Industrial Engineering**, v. 64, n. 1, p. 425-441, Jan. 2013.
- KWIATKOWSKI, D.; P. C. B. PHILLIPS; P. SCHMIDT; Y. SHIN. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. **Journal of Econometrics**, v. 54, n. 1-3, p. 159-178, Dec. 1992.
- PAGE, E. S., Continuous Inspection Schemes. **Biometrika**, v. 41, n. 1-2, p. 100-115, 1954.
- PASSOS, F. F., **Análise temporal da série de consumo residencial de energia elétrica no Brasil no período de 1963 a 2012**. Monografia (Bacharel em Ciências Econômicas) - Instituto de Ciências Sociais Aplicadas. Universidade Federal de Alfenas, Minas Gerais, Brasil, 2015.
- ROMERA, E. G.; MORÁN, M. A. J.; FERNÁNDEZ, D. C., Monthly electric energy demand forecasting with neural networks and Fourier series. **Energy Conversion and Management**, v. 49, n. 11, p. 3135-3142, Nov. 2008.
- SOUZA, D. L. O.; RODRIGUES, M.; REIS, D. R., Crise energética 2001: providencial e reflexiva. **Revista Educação e Tecnologia**, v. 5, n. 8, p. 27-40, Set. 2004.
- TAYLOR, J. W.; MENEZES, L. M.; MCSHARRY, P. E., A comparison of univariate methods for forecasting electricity demand up to a day ahead. **International Journal of Forecasting**, v. 22, n. 1, p. 1-16, Jan.–Mar. 2006.
- UNSIHUAY, V. C.; ZAMBRONI, A. C.; MARANGON, L. J. W.; BALESTRASSI, P. P., Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. **International Journal of Electrical Power & Energy Systems**, v. 32, n. 2, p. 108-116, Feb. 2010.

VILAR, J. M.; CAO, R.; ANEIRO, G., Forecasting next-day electricity demand and price using nonparametric functional methods. **International Journal of Electrical Power & Energy Systems**, v. 39, n. 1, p. 48-55, July 2012.

WENNSTRÖM, A. **Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices**. 2014. 61 p. Dissertação (Master of Science) - Department of Mathematics. Royal Institute of Technology (KTH), Stockholm, Sweden, 2014.