

USO DO GRÁFICO DE CONTROLE PARA MINIMIZAR ERROS DE PREVISÃO EM SÉRIES DE ENERGIA ELÉTRICA

*Eliane da Silva Christo¹
Mirley Bitencourt Ferreira²*

Resumo: Em 2001, o setor elétrico sofreu uma grave crise de abastecimento que culminou em um plano de racionamento de energia elétrica. Esse acontecimento gerou uma série de questionamentos sobre os rumos que o setor elétrico estava trilhando. O reforço dos instrumentos de planejamento energético pelo governo brasileiro permitiu a recuperação de uma responsabilidade do Estado indispensável para que a expansão da oferta de energia se realize de forma satisfatória, em conformidade com o interesse público e a estratégia de desenvolvimento econômico e social. Em termos econômicos, a previsão e o controle estatístico da demanda de energia se tornaram a base da composição de preço no mercado e os erros de previsão e falhas passaram a resultar em perdas financeiras. O problema maior no estudo de modelos é otimizar os dados previstos, ou seja, minimizar os erros de previsão. Estes erros são causados pela má elaboração dos modelos ou pela grande variabilidade do processo. O objetivo deste trabalho é minimizar os erros de previsão de uma série de energia elétrica através do método de amortecimento exponencial aplicada aos pontos fora dos limites do gráfico de controle dos resíduos gerados por um modelo de previsão ARIMA. O método foi comparado a um trabalho anterior onde os dados que ficaram fora de controle foram retirados da série. Verificou-se que o resultado obtido neste trabalho apresentou um erro de previsão menor, além de manter todos os dados da série.

Palavras-Chave: Previsão de Séries Temporais; Controle Estatístico de Qualidade; Energia.

Abstract: In 2001, the electricity sector suffered a serious supply crisis that culminated in a plan of electricity rationing. This event generated a series of questions about the direction the electricity sector was treading. The strengthening of the tools for energy planning by the Brazilian government allowed the recovery of a state responsibility essential so that the expansion of energy to take place satisfactorily in accordance with the public interest and the strategy of economic and social development. In economic terms, forecasting and statistical control of energy demand became the basis of the composition of the market price and the forecast errors and failures began to result in financial losses. The biggest problem in the study of models is to optimize the predicted data, i.e., minimize errors of prediction. These errors are caused by poor preparation of the models or the great variability of the process. The objective is to minimize the forecast errors from a series of electricity through the exponential smoothing method applied to points outside the limits of the control chart of the waste generated for an ARIMA forecasting model. The method was compared to a last work where the data that were out of control were removed the series. It was found that the results obtained in this study showed a lesser prediction error, besides maintaining all data's series.

Keywords: Time Series Forecasting; Statistical Quality Control; Energy.

¹ Departamento de Engenharia de Produção – UFF, Volta Redonda, RJ, e-mail: elianechristo@id.uff.br
² Departamento de Engenharia de Produção – UFF, Volta Redonda, RJ, e-mail: mirleybit@hotmail.com

1. INTRODUÇÃO

Nos últimos anos, a indústria de energia elétrica de diversos países vem sofrendo diversas transformações, principalmente no que diz respeito à estrutura do mercado e regulamentação. O principal objetivo deste processo de reestruturação é promover a eficiência econômica através da competição. Um dos componentes mais importantes de infra-estrutura no processo de reestruturação reside na oferta regular e confiável de energia elétrica, um equilíbrio entre oferta e demanda. Compete, portanto, a cada operação estabelecer critérios operativos que possam assegurar a integridade do sistema, garantindo deste modo o fornecimento de energia a todos os consumidores, dentro de padrões de qualidade aceitáveis com relação às oscilações de tensão e frequência e tempo de interrupção do serviço (Cascaes, 2003).

Em termos de modelo matemático, uma margem de segurança é obtida impondo-se restrições de segurança. Cada uma destas restrições tem a função de prevenção contra a ocorrência de algum evento crítico. Em termos econômicos, a previsão e o controle estatístico da demanda de energia se tornaram a base da composição de preço no mercado e os erros de previsão e falhas passaram a resultar em perdas financeiras. Diante dessa situação, eliminar desperdícios, adotar tecnologias avançadas, desenvolver novos produtos, envolver os colaboradores e buscar a melhoria contínua dos processos de produção tornaram-se a base de sustentação dos negócios (Bolle, 2001). O Controle Estatístico do Processo concentra-se na diminuição ou eliminação da incidência de erros, defeitos e falhas em um processo.

O problema maior no estudo de modelos de previsão é aperfeiçoar os dados previstos, ou seja, minimizar as diferenças entre os valores reais e os previstos, isto é, minimizar os erros de previsão. Estes erros são causados pela má elaboração dos modelos ou pela grande variabilidade do processo. Neste contexto, tornam-se necessários estudos aprofundados para elaboração do modelo mais preciso de previsão.

Neste trabalho, será utilizada uma série de cargas ativas mensais da demanda de energia fora dos horários de ponta (pico) dos anos 1982 a 2000 de uma empresa de energia elétrica. O horário de ponta é o período de 3 (três) horas durante o dia no qual o consumo de energia elétrica está em seu ápice, aumentando fortemente a carga demandada do sistema de geração, transmissão e distribuição de energia elétrica. Por isso, neste período a energia elétrica tem um preço significativamente maior, implicando em um custo de energia elevado

para aqueles consumidores que a usam no período. Corresponde ao intervalo de 3 horas consecutivas, ajustado de comum acordo entre a concessionária e o cliente, situado no período compreendido entre as 18h e 21h e durante o horário de verão e das 19h à 22h, com exceção de sábados, domingos e feriados nacionais. A análise será feita em duas etapas: na primeira etapa, se realizará a identificação do modelo de previsão que melhor se ajusta aos dados e a partir daí calcular os erros percentuais absolutos médios (MAPE); na segunda etapa, será feito o gráfico de controle dos resíduos do modelo inicial para verificar se todos os dados estão sob controle estatístico. Caso isso não seja confirmado, se aplicará o método de amortecimento exponencial nos pontos fora dos limites do gráfico a fim de tornarem os resíduos sob controle e reduzir os erros percentuais absolutos médios.

2. METODOLOGIA

2.1 HISTÓRICO DO SETOR ELÉTRICO BRASILEIRO

A reforma do Setor Elétrico Brasileiro começou em 1993 com a Lei nº 8.631, que extinguiu a equalização tarifária vigente e criou os contratos de suprimento entre geradores e distribuidores, e foi marcada pela promulgação da Lei nº 9.074 de 1995, que criou o Produtor Independente de Energia e o conceito de Consumidor Livre (Câmara e Soares, 2012).

Em 1996 foi implantado o Projeto de Reestruturação do Setor Elétrico Brasileiro (Projeto RE-SEB), coordenado pelo Ministério de Minas e Energia (Silva, 2012).

As principais conclusões do projeto foram a necessidade de implementar a desverticalização das empresas de energia elétrica, ou seja, dividi-las nos segmentos de geração, transmissão e distribuição, incentivar a competição nos segmentos de geração e comercialização, e manter sob regulação os setores de distribuição e transmissão de energia elétrica, considerados como monopólios naturais, sob regulação do Estado (Moreira, 2011).

Foi também identificada a necessidade de criação de um órgão regulador (a Agência Nacional de Energia Elétrica - ANEEL), de um operador para o sistema elétrico nacional (Operador Nacional do Sistema Elétrico - ONS) e de um ambiente para a realização das transações de compra e venda de energia elétrica (o Mercado Atacadista de Energia Elétrica - MAE).

Concluído em agosto de 1998, o Projeto RE-SEB definiu o arcabouço conceitual e institucional do modelo a ser implantado no Setor Elétrico Brasileiro (Silva, 2011, Silva e Gómez, 2011).

Em 2001, o setor elétrico sofreu uma grave crise de abastecimento que culminou em um plano de racionamento de energia elétrica. Esse acontecimento gerou uma série de questionamentos sobre os rumos que o setor elétrico estava trilhando. Visando adequar o modelo em implantação, foi instituído em 2002 o Comitê de Revitalização do Modelo do Setor Elétrico, cujo trabalho resultou em um conjunto de propostas de alterações no setor elétrico brasileiro (Rosim, 2011).

Durante os anos de 2003 e 2004 o Governo Federal lançou as bases de um novo modelo para o Setor Elétrico Brasileiro, sustentado pelas Leis nº 10.847 e 10.848, de 15 de março de 2004 e pelo Decreto nº 5.163, de 30 de julho de 2004.

Em termos institucionais, o novo modelo definiu a criação de uma instituição responsável pelo planejamento do setor elétrico a longo prazo (a Empresa de Pesquisa Energética - EPE), uma instituição com a função de avaliar permanentemente a segurança do suprimento de energia elétrica (o Comitê de Monitoramento do Setor Elétrico - CMSE) e uma instituição para dar continuidade às atividades do MAE, relativas à comercialização de energia elétrica no sistema interligado (a Câmara de Comercialização de Energia Elétrica - CCEE).

Em relação à comercialização de energia, foram instituídos dois ambientes para celebração de contratos de compra e venda de energia, o Ambiente de Contratação Regulada (ACR), do qual participam Agentes de Geração e de Distribuição de energia elétrica, e o Ambiente de Contratação Livre (ACL), do qual participam Agentes de Geração, Comercialização, Importadores e Exportadores de energia, e Consumidores Livres (Silva, 2011, de Souza et al., 2011).

2.2 CENÁRIO BRASILEIRO DE ENERGIA ELÉTRICA

O reforço dos instrumentos de planejamento energético pelo governo brasileiro permitiu a recuperação de uma responsabilidade do Estado indispensável para que a expansão da oferta de energia se realize de forma satisfatória, em conformidade com o interesse público e a estratégia de desenvolvimento econômico e social.

O marco regulatório estabelecido nos últimos anos visa assegurar de forma sustentável a

expansão do sistema energético nacional, por um processo de planejamento participativo que promova a diversificação da matriz energética, o uso de fontes renováveis, a eficiência e conservação de energia, incentivando a integração energética sul-americana, a pesquisa e o desenvolvimento tecnológico do setor (dos Santos e de Souza, 2011, Silva e Gómez, 2011).

Além disso, as atividades de planejamento procuram garantir no médio e longo prazos o equilíbrio entre oferta e demanda de energia, com qualidade e confiabilidade, alicerçado em regras econômico-financeiras que permitam a remuneração atrativa dos investimentos, em harmonia com os princípios da modicidade tarifária, do desenvolvimento sustentável e da universalização e inclusão social (Albuquerque, 2012).

Os estudos de inventário hidrelétrico foram retomados e leilões de energia para atendimento do mercado regulado passaram a ser promovidos de forma periódica e sistemática. Também foi intensificado o acompanhamento da implementação de novos empreendimentos, por meio do Programa de Aceleração do Crescimento, viabilizando projetos fundamentais para o atendimento da demanda energética brasileira (Silva, 2011).

2.3 PREVISÃO DE ENERGIA ELÉTRICA

A previsão de energia elétrica tem merecido relevante atenção dos pesquisadores da área, muitas técnicas e algoritmos foram desenvolvidos.

(Hippert et al., 2001) apresentam diversos métodos utilizados na literatura para a previsão de demanda de energia a curto prazo. Neste trabalho o autor relata as principais metodologias utilizadas entre 1991 e 1999 nos quais estão incluídos métodos estatísticos, métodos baseados em redes neurais, lógica fuzzy e sistemas de especialistas.

A maioria dos modelos usuais de previsão já foi experimentada na previsão de carga: modelos autorregressivos multiplicativos (El-Keib et al., 1995), modelos dinâmicos lineares (Douglas et al., 1998), modelos não lineares (Sadownik e Barbosa, 1999) e métodos baseados no filtro de Kalman (da Silva, 2001). Entre os modelos causais, nos quais a carga é modelada como função de alguma variável exógena, têm sido tentadas as funções de transferências de Box & Jenkins (Box et al., 2011), modelos ARMAX (Yang et al., 1995), técnicas de otimização (Yu, 1996) e modelos estruturais (Harvey, 1991, Harvey e Koopman, 1993).

Apesar do grande número de alternativas, contudo, os modelos causais mais freqüentes encontrados na literatura são os modelos lineares aditivos que decompõem a carga, geralmente, somente em um componente básico (Christo e Souza, 2006) ou também com um componente relacionado com o clima (Bunn, 1985). O componente básico deve refletir o comportamento normal da série de cargas, é equivalente ao valor esperado da carga, dados os valores passados da série, se não houver nenhuma circunstância anômala. O componente relacionado com o clima deve modelar a influência das variáveis meteorológicas na carga. É evidente que, parte desta influência já foi explicada pelo componente básico, já que o efeito do clima tende a se inserir nos valores passados da série, no entanto, este componente em geral só é significativo quando ocorrem mudanças meteorológicas inesperadas. Estes métodos de decomposição são bastante úteis, porque é possível atribuir aos componentes uma interpretação física, o que permite aos operadores do sistema entender melhor o comportamento da série, e permite aos agentes do mercado entender melhor os mecanismos causadores da volatilidade.

Na década de 90, a grande presença dos sistemas especialistas é marcante. Os modelos são baseados em rede neural (Hippert et al., 2001, Shrivastava et al., 2012), lógica nebulosa (Hsu e Ho, 1992), sistemas neuro-nebulosos (Nguyen e Liao, 2011) e modelagem estatística combinada com uma ou com algumas dessas técnicas (Lauret et al., 2012, Yang et al., 1995, Harvey, 1991, El-Keib et al., 1995).

Parte dos artigos investigados se interessava por previsões pontuais: da carga uma hora à frente, da carga de pico ou da energia total um dia à frente. A outra parte dos artigos tratava do problema mais complexo de previsão: a curva de carga horária para o dia seguinte (da Silva, 2001).

Quatro abordagens diferentes foram mencionadas nos artigos, para possibilitar a previsão da curva de carga horária para o dia seguinte (Hagan e Behr, 1987):

- Previsão Iterativa: consiste em prever um valor horário de cada vez e agregar este valor à série de entradas do modelo, de modo que, as previsões para os valores seguintes são baseadas nas previsões anteriores (incerteza crescerá);
- Previsão por meio de curva padrão: Consiste em criar um modelo para o perfil diário e explicar os desvios de carga em relação a este modelo como consequência do efeito de variáveis exógenas (Douglas et al., 1998, Sousa et al., 2011, Yu, 1996).

- Previsão por Modelos em Paralelo: Consiste em modelar perfil por meio de conjuntos de modelos similares em paralelo (um modelo para cada hora do dia). Uma desvantagem desta abordagem é a de que, por considerar cada hora em separado, os modelos não exploram suficientemente a autocorrelação entre as cargas em diferentes horas do perfil (Paparoditis e Sapatinas, 2012).
- Previsão por modelo único (saída multivariada): Consiste em usar um modelo único com 24 saídas, cada uma responsável pela previsão do perfil diário e as entradas do modelo em geral incluem todas as cargas do perfil do dia antes da previsão. Esta última forma pode apresentar melhor desempenho, que um modelo para cada hora, já que leva em conta a dependência serial entre os dados (Sadownik e Barbosa, 1999).

2.4 AVALIAÇÃO DA AUTOCORRELAÇÃO DA SÉRIE

Para iniciar a análise da série é preciso detectar se há autocorrelação entre os dados (Souza e Camargo, 1996, Montgomery e Johnson, 1976). A saber, autocorrelação caracteriza uma série temporal não-estacionária, isto é, não possui média ou variância constante ao longo do tempo e a covariância entre dois períodos de tempo não depende apenas da distância entre os dois períodos. Pode-se detectar a presença de autocorrelação através dos seguintes meios: (i) o gráfico de correlograma, com base no cálculo da função de autocorrelação (FAC ou ACF – *autocorrelation function*) e da função de autocorrelação parcial (FACP ou PACF – *partial autocorrelation function*), (ii) as estatísticas Q de Box-Pierce, (iii) e de Ljung-Box, (iv) o teste da raiz unitária, (v) o teste da correlação cruzada, ou (vi) o teste do periodograma acumulado.

Os gráficos de correlograma da FAC e da FACP são métodos visuais fáceis e eficientes de detectar a presença de autocorrelação. Eles mostram a autocorrelação da série em várias defasagens (*lag*), indicando em quais pontos esse autocorrelação é significativa, com base nos teste Q de Box-Pierce e no teste de Ljung-Box (LB).

A autocorrelação está presente em séries temporais, e existem dois tipos destas séries, estacionárias e não-estacionárias.

A série temporal estacionária refere-se a uma série em que sua média fica em torno de um determinado valor fixo como na equação [1]:

$X_t = \mu + \epsilon_t$ [1]
 X_t : uma característica da qualidade
 t : tempo no instante 1,2,3...
 μ : média do processo
 ϵ_t : variação aleatória em torno da média considerada como sendo $N(0, \sigma_t^2)$ com σ_t^2 constante.

Já a série temporal não-estacionária tem sua média oscilando, assim sendo, não é possível determinar um valor para esta. Tal série não é previsível, é impossível determinar seu comportamento num tempo adiante, logo, não é coerente tentar determinar sua capacidade de processo. Neste caso, usa-se a transformação de Box & Cox (Box e Cox, 1964) que, na maioria dos casos, tornam a série estacionária.

O processo será autocorrelacionado caso os valores da variação aleatória forem interdependentes, sendo os valores independentes o processo será não autocorrelacionado. A correlação de uma série temporal pode ser medida através do cálculo de função de autocorrelação.:

$\rho_k = \text{cov}(X_t, X_{t-k}) / V(X)$ [2]
 para $k = 0, 1, 2, \dots$
 $\text{cov}(X_t, X_{t-k})$ – covariância das observações em k períodos de tempo (ou *lags*)
 $V(X)$ – variância das observações, assumida como constante.

A equação [2] equivale a equação [3]:

$\rho_k = \sum (X_t - \bar{X})(X_{t-k} - \bar{X}) / \sum (X_t - \bar{X})^2$ [3]
 para $k = 0, 1, 2, \dots$

É aconselhável efetuar a expressão utilizando valores baixos para k , em geral, $k \leq n/4$, sendo n o número total de observações.

2.5 MODELOS DE PREVISÃO

Serão apresentados a seguir os modelos utilizados neste trabalho para a previsão mensal da carga ativa de energia elétrica (Geurts e Ibrahim, 1975).

2.5.1 AMORTECIMENTO EXPONENCIAL

Nesses modelos assume-se que o nível médio das observações pode ser descrito, a cada instante de tempo, por uma função conhecida (de Miranda, 2007, Souza, 1983).

Para uma série representada por Z_t para $t=1, 2, \dots$, sendo $\mu(t)=E[Z_t]$ o seu nível médio no instante t , tem-se $\mu(t)=f(t)$

O objetivo dos métodos é estimar os parâmetros que caracterizam a função $f(t)$. Para a série Z_t será considerado um modelo estocástico como a equação [4].

$$Z_t = \mu(t) + \epsilon_t \quad [4]$$

onde

ϵ_t é o ruído do sistema no instante t de média nula e variância constante para todo t .

O modelo da equação [4] passa a ser representado pela equação [5].

$$Z_t = f(t) + \epsilon_t \quad [5]$$

resultando na equação [6] de previsão τ passos-à-frente no instante T

$$Z_{T+\tau} = \phi(T+\tau) + \epsilon_{T+\tau} \quad [6]$$

A previsão de $Z_{T+\tau}$ (equação [7]) é obtida em função da expressão anterior

$$\begin{aligned} \hat{Z}_T(\tau) &= E[Z_{T+\tau} | Z_T] = \\ &= E[f(T+\tau) + \epsilon_{T+\tau} | Z_T] = \\ &= E[f(T+\tau)] \end{aligned} \quad [7]$$

Para séries não-sazonais, a função normalmente adotada para $f(t)$ é do tipo polinomial como a equação [8].

$$f(t) = \sum_i a_{i+1} t^i \quad [8]$$

O modelo fica como a equação [9]:

$$Z_t = \sum_{i=0}^n a_{i+1} t^i + \epsilon_t \quad [9]$$

e a equação [10] representa a previsão

$$Z_{T+\tau} = \sum_{i=0}^n a_{i+1} (T+\tau)^i + \epsilon_{T+\tau} \quad [10]$$

Fazendo uma translação da origem de modo que $t=0$ coincida com o instante T , a previsão pontual τ passos-à-frente de T fica sendo a equação [11].

$$\hat{Z}_T(\tau) = E \left[\sum_{i=0}^n a_{i+1} \tau^i \right] \quad [11]$$

Desenvolvendo chega-se a equação [12].

$$\hat{Z}_T(\tau) = \sum_{i=0}^n \tau^i E[a_{i+1}] \quad [12]$$

Designando $\hat{a}_{i+1}(T)$ o estimador de a_{i+1} feito no instante T tem-se a equação [13].

$$\hat{Z}_T(\tau) = \sum_{i=0}^n \hat{a}_{i+1}(T) \tau^i \quad [13]$$

Pelo método de amortecimento exponencial, esses estimadores são obtidos em função de médias móveis de tamanho N. De acordo com o número de parâmetros, são efetuadas médias da média anterior para compor as equações desses parâmetros. Essas médias são calculadas pelas equações [14].

$$M_T = \alpha Z_T + (1 - \alpha) M_{T-1}$$

$$M_T^{[i]} = \alpha M_T^{[i-1]} + (1 - \alpha) M_{T-1}^{[i]}, \quad i > 1$$

$$M_T = \frac{1}{N} \sum_{j=0}^{N-1} Z_{T-j} \quad [14]$$

$$M_T^{[i]} = \frac{1}{N} \sum_{j=0}^{N-1} M_{T-j}^{[i-1]}, \quad i > 1$$

onde

α - constante de amortecimento

As expressões dos estimadores para polinômios de ordem 0, 1 e 2 podem ser conseguidas na referência. O procedimento de determinação desses parâmetros constitui o Método de Brown.

Caso as observações sejam provenientes de uma série temporal que possui uma repetição periódica definida, deve-se incluir esta informação determinística ao nível médio da série. Séries com tal característica são conhecidas como séries sazonais, sendo o comprimento do ciclo sazonal denotado por S. Assim para o ciclo sazonal de 1 ano tem-se S=12 para séries mensais, S=4 para séries trimestrais e S=52 para séries semanais, etc.

2.5.2 BOX & JENKINS

A filosofia da modelagem Box & Jenkins baseia-se em duas idéias (Box et al., 2011): o princípio da parcimônia (escolher um modelo com o menor número de parâmetros possíveis para uma representação matemática adequada) e a construção de modelos através de um ciclo iterativo (estratégia de seleção de modelos até a obtenção de um modelo satisfatório).

A modelagem Box & Jenkins segue a equação [15].

$$w_t = \sum_{k=0}^{\infty} \Psi_k a_{t-k} \quad [15]$$

onde:

a_t - ruído branco

B – operador retardo que representa um atraso de um período de tempo

($B_k Z_t = Z_{t-k}$)

Ψ_k - filtro linear definido como

$$\frac{\theta_q(B)}{\phi_p(B)}$$

$$\phi_p(B)$$

Desta forma os modelos Box & Jenkins são dados pela equação [16].

$$\phi_p(B) w_t = \theta_q(B) a_t \quad [16]$$

onde

$\phi(\cdot)$ e $\theta(\cdot)$ são polinômios de graus p e q, respectivamente

Como na vida prática nem todos os processos são estacionários, deve-se procurar algum tipo de operador que produza a partir de w_t um processo não estacionário. Nesta modelagem consideram-se somente os processos chamados não estacionários homogêneos, isto é, aquela classe de processos não estacionários para os quais diferenças sucessivas produzem um processo estacionário.

Utilizando-se o operador soma ($\nabla_n Z_t = Z_t - Z_{t-n}$) e aplicando-o na série disponível obtém-se o modelo ARIMA(p,d,q) representado na equação [17] (Souza e Camargo, 1996).

$$\Phi(B) \nabla_d Z_t = \Theta(B) a_t \quad [17]$$

Nesse modelo supõe-se que a d-ésima diferença de Z_t possa ser representada por um modelo ARMA(p,q).

O passo seguinte na metodologia é a identificação do modelo, ou seja, a sua ordem. Para isso são utilizados os conceitos de função de autocorrelação e autocorrelação parcial. Após a identificação da ordem do modelo, é necessário obter as estimativas dos parâmetros desse modelo. A técnica utilizada para as estimativas é a da máxima verossimilhança.

Para comprovar a validade do modelo selecionado, são realizados alguns testes estatísticos. Finalmente, após a obtenção da estimativa do modelo,

procede-se à previsão de valores futuros da série e de seus limites de confiança. Destaca-se que uma análise mais profunda desse método pode ser encontrada nas referências.

2.6 GRÁFICOS DE CONTROLE PARA DADOS CORRELACIONADOS

A verificação de que o processo está ou não sob controle é feita pelo exame dos itens das amostras extraídas periodicamente. Se o processo estiver sob controle, as amostras apresentarão aquela variabilidade correspondente a amostras extraídas de uma população normal, isto é, a variabilidade devida apenas ao acaso na amostragem. O processo sob controle supõe, portanto, que o característico de qualidade do conjunto de itens produzidos possua distribuição normal.

Além disso, supões também que essa distribuição permaneça estável, isto é, que seus dois parâmetros, a média e o desvio padrão, permaneçam constantes, o que é verificado periodicamente pela extração de uma seqüência de amostras. Por isso, diz-se que, em um processo sob controle, a variabilidade é devida tão somente a causas aleatórias. Estas causas de variação não provocam alterações apreciáveis na qualidade do produto, por isso, as causas aleatórias são consideradas como parte natural do processo de fabricação.

Quando a variabilidade se torna anormal, as alterações nos característicos de qualidade do produto são sensíveis. As amostras indicarão que o processo de fabricação (isto é, “a população”) se modificou, e ficou fora de controle. As causas da modificação podem ser descobertas e, por isso, são denominadas causas identificáveis. Sua influência, rápida e crescente, exige pronta ação corretiva, no sentido de eliminar a presença dessas causas identificáveis. Contudo, quando acarretam melhoria de qualidade, sua identificação e adaptação contribuem para o aperfeiçoamento do processo de fabricação.

A presença de causas identificáveis é indicada pela ocorrência de diferenças insignificantes entre o valor observado e a média do processo, isto é, de valores amostrais fora da taxa de controle. Essa presença pode ser revelada por meio de um dispositivo prático, o chamado gráfico ou carta de controle ou ainda gráfico de Shewhart.

Se os dados que representam o processo forem autocorrelacionados pode ser que muitas causas especiais, detectadas durante a avaliação da estabilidade,

de, sejam apenas falsos alarmes. Quando isto ocorre, as cartas de controle convencionais de Shewhart não devem ser aplicadas diretamente tanto para a análise da estabilidade quanto para o controle de processos (Montgomery, 2009, Ishikawa e Loftus, 1990).

Uma alternativa para o controle estatístico de processos autocorrelacionados, consiste em espaçar as medidas por um intervalo de tempo suficientemente longo e os tradicionais gráficos de controle de “X-bar”, “R” e “S” são substituídos pelos gráficos de observações individuais “I” e amplitude móvel “MA”. Utiliza-se o gráfico da amplitude móvel com o objetivo de monitorar a variabilidade do processo e o gráfico de controle para observação individual com o objetivo de monitorar o nível do processo.

O gráfico da média móvel exponencialmente ponderada (EWMA) é também indicado para detecção de pequenas mudanças na média do processo, além de ser recomendado para a análise de variáveis autocorrelacionadas (Lu e Reynolds, 1999).

(Montgomery, 2009) propõe três alternativas para tratar os dados autocorrelacionados: (i) ajuste de um modelo ARIMA (auto-regressivo integrado e de média móvel) e aplicação de um gráfico de controle padrão para resíduos (por exemplo, de observações, média móvel, CUSUM – soma acumulada, EWMA – média móvel ponderada exponencialmente), os gráficos de controle EWMA com linha central móvel, e uso do Controle da Engenharia do Processo (EPC – *Engineering Process Control*).

(Wardell et al., 1994) estudaram o desempenho do gráfico de resíduos na presença de autocorrelação, que é o gráfico tradicional de Shewhart aplicado aos resíduos. O desempenho do CUSUM em processos autocorrelacionados foi estudado por (Yashchin, 1993, Lu e Reynolds, 2001). O gráfico CUSUM é utilizado no lugar do gráfico de Shewhart quando se quer detectar mudança na média do processo menor que $1,5\sigma$ mais rapidamente. O gráfico EWMA possui a mesma finalidade do CUSUM.

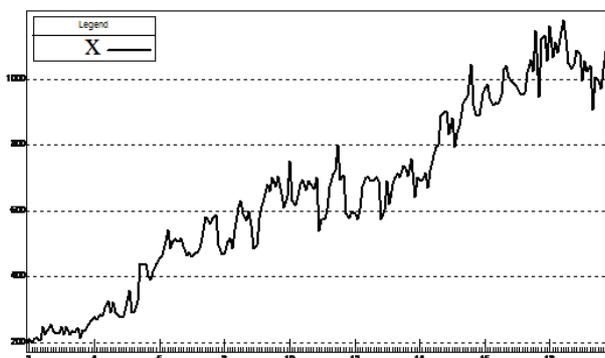
A autocorrelação tem sido reconhecida como um fenômeno natural nas indústrias onde parâmetros como temperatura e pressão variam lentamente para a taxa ao qual eles são medidos. Quando os gráficos de controle de Shewhart são construídos para dados autocorrelacionados, os limites de controle tradicionais aumentam a probabilidade de ocorrer alarmes falsos.

3 RESULTADOS E DISCUSSÕES

A série utilizada são cargas mensais da demanda de fora da ponta (pico) dos anos 1982 a 2000 de uma empresa de energia “X”. Corresponde ao intervalo de 3 horas consecutivas, ajustado de comum acordo entre a concessionária e o cliente, situado no período compreendido entre as 18h e 21h e durante o horário de verão e das 19h à 22h, com exceção de sábados, domingos e feriados nacionais.. A análise de previsão foi feita para o último ano (Janeiro à Julho de 2000), o qual não entrou no modelo para identificação e estimação dos parâmetros.

A Figura 1 apresenta um gráfico seqüencial da série original. Nota-se, que a série é não estacionária, esse fato será confirmado com a Figura 2 que apresenta a função de autocorrelação (FAC) da série.

Figura 1- Série de Empresa de energia “X”



Antes de aplicar as metodologias, deve-se verificar se a série é homocedástica, ou seja, variância constante. Esta verificação é feita através do histograma da série ou através do teste de homogeneidade de variância. Quando uma série é heterocedástica deve-se aplicar a função não linear de Box & Cox para normalizá-la (Box e Cox, 1964).

Pelo teste de homogeneidade de variância, tabela 1, verifica-se que a série é homocedástica, ou seja, não há necessidade da transformação de Box & Cox.

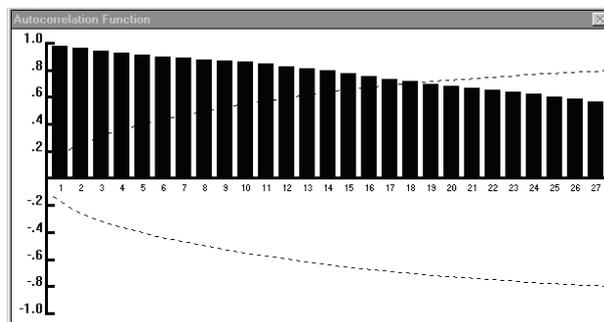
Tabela 1 – Testes de variância

LEVENE TEST FOR HOMOGENEITY OF VARIANCES ¹			
STATISTIC	DF1	DF2	2-TAIL SIG.
1.4529	1	221	0.229

¹H₀: Variância Homocedástica;
H₁: Variância Heterocedástica.

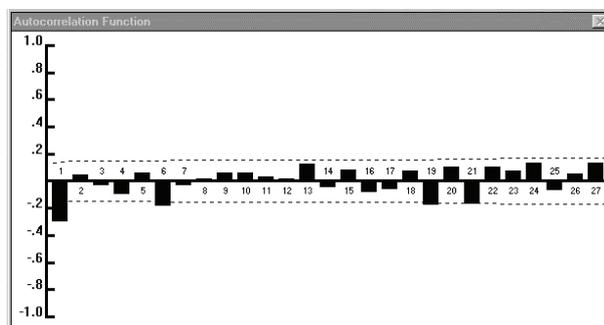
Para identificação do modelo, inicialmente, analisa-se o correlograma da série, figura 2:

Figura 2 – Função de autocorrelação



Na figura 2 observa-se que as autocorrelações estimadas apresentam valores absolutos altos (até o “lag” 18), pois estes intervalos estão fora do intervalo de confiança, além de decaírem lentamente. Isto indica, que se deve aplicar o operador diferença a série e fazer o cálculo novamente. O resultado se encontra na figura 3.

Figura 3 – Função de autocorrelação

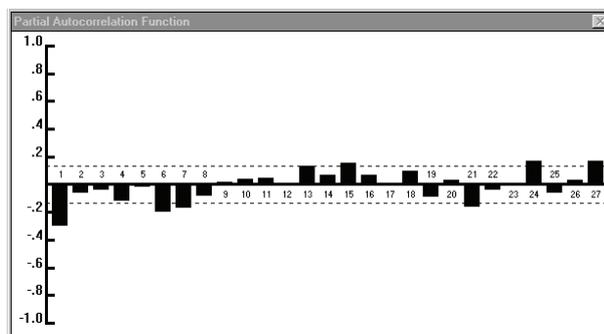


Aplicando a 1ª diferença na série e estimando sua função de autocorrelação observa-se que ela atingiu as características de um processo estacionário, sem necessidade de mais diferenciação. Com isto, o grau de diferenciação “d” é 1.

Além do gráfico da autocorrelação, determinou-se o gráfico da autocorrelação parcial (figura 4).

Através dos gráficos representados pelas figuras 3 e 4 respectivamente, verificam-se que o “lag 1” nos 2 gráficos é significativo. Com isto, tem-se, a princípio, o modelo ARIMA (1,1,1)

Figura 4 – Função de autocorrelação parcial



Na tabela 2 estão apresentados os testes estatísticos aplicados ao modelo selecionado.

Ao estimar os parâmetros de 3 modelos candidatos ARIMA(1,1,1), ARIMA(0,1,1) e ARIMA(1,1,0) verificou-se que todos os parâmetros foram significativos ao nível de 5%. Porém, o primeiro modelo apresentou estatísticas melhores que as demais. O coeficiente de determinação (R^2) e o Critério de Informação Bayesiano (BIC) foram

maiores enquanto que o Erro Percentual Absoluto Médio (MAPE) foi menor. Apesar de apresentar número maior de parâmetros (Schwarz, 1978).

Pelos resultados apresentados, com R^2 e BIC maiores, o modelo escolhido como representativo do processo gerador da série em estudo é o ARIMA(1,1,1) mesmo não sendo o mais parcimonioso. A figura 5 apresenta função de autocorrelação dos resíduos deste modelo.

Tabela 2- Modelos ARIMA

MODEL ARIMA	PARÂMETROS ESTIMADOS	SIGNIFICÂNCIA DOS PARÂMETROS	R2-AJUSTADO	LJUNG-BOX	BIC	MAPE
(1,1,1)	$\phi_1=0.4172$ $\theta_1=0.7226$	0.9933 1.0000	0.9697	37.47	48.44	0.0567
(1,1,0)	$\phi_1=-0.2956$	1.0000	0.9688	31.81	48.18	0.05749
(0,1,1)	$\theta_1=0.3230$	1.0000	0.969	38.48	48.04	0.05743

Figura 5 – Função de autocorrelação dos erros

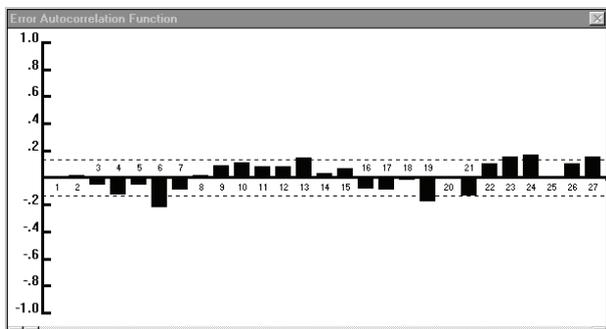
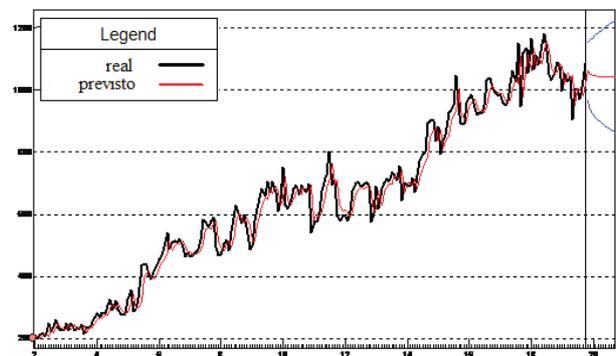


Figura 6 – Previsão Box & Jenkins



Após o término do processo iterativo de identificação, estimação e diagnóstico, emprega-se o modelo gerador da série que melhor se ajusta aos dados reais para prever valores futuros.

A previsão obtida através do modelo escolhido encontra-se na tabela 3.

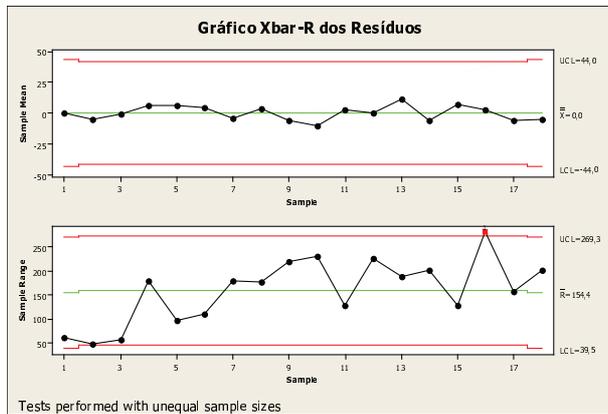
A figura 6 apresenta o gráfico de previsão do modelo escolhido.

Através da tabela 3 e da figura 6, observa-se que as previsões estão bem próximas dos valores reais, porém vale ressaltar que à medida que o horizonte de previsão vai aumentando, os valores previstos tendem a não serem tão precisos. Eles tendem para uma reta.

Tabela 3 – Previsão Box&Jenkins

PREVISÃO DA SÉRIE PARA O PERÍODO DE JANEIRO A JULHO DE 2000				
PERÍODO	LOWER 2.5	PREVISÃO	UPPER 97.5	ATUAL
2000-01	968.066	1060.649	1153.231	890.00
2000-02	937.761	1050.489	1163.217	980.00
2000-03	921.889	1046.251	1170.612	970.00
2000-04	911.324	1044.482	1177.641	1104.00
2000-05	903.013	1043.745	1184.476	1244.00
2000-06	895.780	1043.437	1191.093	1238.00
2000-07	889.141	1043.308	1203.628	1300.00

Figura 7 – Gráfico de Controle dos Resíduos

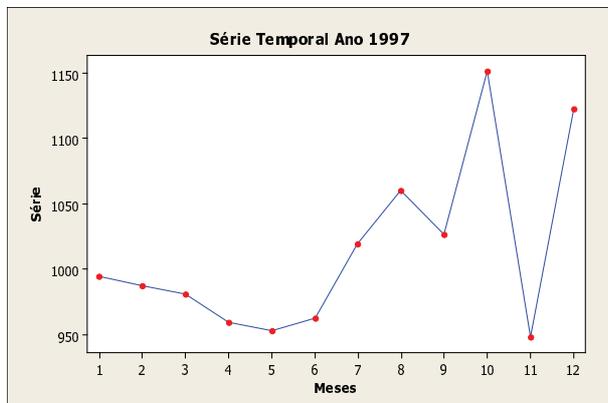


Como os resíduos do modelo são independentes e normalmente distribuídos, o gráfico de controle de Shewhart pôde ser aplicado para identificar se todos os pontos destes resíduos estão sob controle. Através da análise dos gráficos de controle da média

e amplitude ($\bar{X} - R$), figura 7, observa-se que um ponto (amostra 16) se encontra fora dos limites de controle.

Como o ponto fora de controle corresponde a amostra 16 (ano 1997) da série de energia, foi feita uma análise mais detalhada através do gráfico da série somente deste ano específico. A figura 8 apresenta esta análise.

Figura 8 – Gráfico do Ano 1997



Observa-se que, pela figura 8, o ano 1997 apresenta variações significativas nos níveis, principalmente, nos últimos meses (Outubro, Novembro e Dezembro). Por isso, foram feitas várias análises específicas para este ano.

Em uma primeira análise feita por (Christo e Ferreira, 2011), foram simplesmente retirados os pontos referentes aos meses que apresentaram maiores variações no ano de 1997. Como Novembro foi o mês que apresentou mudança brusca no crescimento da série, este foi retirado isoladamente, e depois retirado conjuntamente com seus meses

anterior e posterior. Foram feitas novas previsões, mantendo o modelo inicial ARIMA (1,1,1). O resultado apresentou uma redução de 2,74% do erro percentual absoluto médio (MAPE) do modelo inicial quando se retiraram os meses Outubro e Novembro conjuntamente.

Apesar da redução do erro de previsão e dos resíduos dos dados atuais terem ficado sob controle no gráfico de Shewhart. Quando se tiram dados da série, há perdas de informações que podem causar distorções na análise final.

A fim de minimizar as mudanças nos níveis, sem precisar eliminar nenhuma informação da série, foi aplicado um amortecimento exponencial simples no ano 1997, que se apresentou fora de controle no gráfico de Shewhart. Foram feitas duas análises considerando pesos diferentes ($\alpha = 0,2$ e $\alpha = 0,5$).

Após a suavização dos dados do ano 1997, os valores amortecidos foram incorporados à série original e feita uma nova análise das autocorrelações para descobrir o modelo de previsão ideal. Com isso, verificou-se que, o modelo mais adequado seria o mesmo utilizado no início, ARIMA (1,1,1). Uma comparação entre os erros percentuais médios é apresentada na tabela 4. O Modelo I é aquele onde os dados do ano 1997 foram amortecidos exponencialmente com $\alpha = 0,2$, e no Modelo II os dados foram amortecidos exponencialmente com $\alpha = 0,5$.

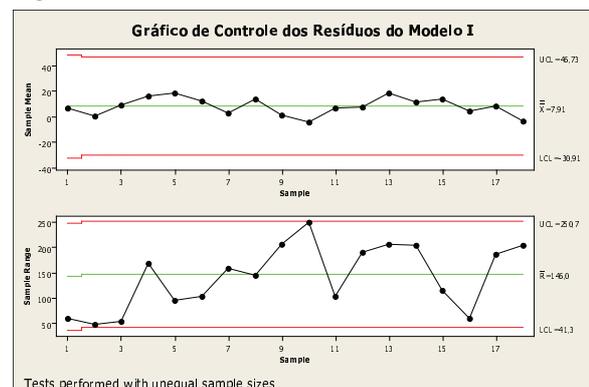
Tabela 4 – MAPES

SÉRIES	MODELO ORIGINAL	MODELO I	MODELO II
MAPE	0,0567	0,054191	0,05479

Observa-se que, o menor erro percentual absoluto médio é alcançado no Modelo I, quando os dados do ano 1997 foram amortecidos exponencialmente com $\alpha = 0,2$. A redução do erro em relação ao modelo original foi de 4,41%.

Os gráficos de controle ($\bar{X} - R$) dos resíduos independentes e normalmente distribuídos do Modelo I são apresentados na figura 9.

Figura 9 – Gráficos de Controle dos Resíduos do Modelo I



Observa-se, na figura 9, que, todos os pontos ficaram sob controle estatístico. Alguns pontos dos resíduos gerados pela previsão se encontraram bem próximos aos limites do gráfico de controle da amplitude, o que pode estar influenciando negativamente na previsão. Provavelmente, uma análise mais detalhada desses pontos possa melhor ainda mais os resultados.

4 CONCLUSÕES

Quando os dados não são independentes, ou seja, dados correlacionados, as cartas de controle convencionais de Shewhart não devem ser aplicadas diretamente tanto para a análise da estabilidade quanto para o controle de processos Shewhart. Com isso, usa-se um modelo de previsão que mais se ajusta aos dados correlacionados, e depois se pode fazer um gráfico de controle dos resíduos deste modelo que é independente e normalmente distribuído.

No artigo foi feito, primeiramente, um modelo ARIMA(1,1,1) na série de energia, logo após observou-se no gráfico de controle tradicional se todos os dados estavam sob controle. E verificou-se que, um ponto, correspondente ao ano de 1997, estava fora de controle. Com isso, foram aplicados dois modelos de amortecimento exponencial nos níveis com pesos diferentes ($\alpha = 0,2$ e $\alpha = 0,5$) somente nos dados de 1997. Os valores amortecidos foram incorporados à série original e feita novamente uma análise de previsão para os dois modelos criados (Modelo I e Modelo II). Constatou-se, através dos erros médios percentuais (MAPE) das previsões que o melhor modelo foi I onde considerou $\alpha = 0,2$.

A nova série, que também foi prevista pelo ARIMA (1,1,1), apresentou além de uma redução do MAPE de 4,41%, todos os resíduos ficaram sob controle estatístico. O que comprova o bom desempenho da análise.

Quando comparado o resultado encontrado neste trabalho com o resultado de (Christo e Ferreira, 2011), verificou-se uma melhora de quase 2% na redução do MAPE e ainda preservando todas as informações da série.

Alguns outros pontos dos resíduos gerados pela previsão se encontraram bem próximos aos limites do gráfico de controle da amplitude, o que pode estar influenciando negativamente na previsão. Uma sugestão para trabalhos futuros seria uma análise mais detalhada desses pontos da mesma maneira que foi feita para o ponto que se encontrava fora dos limites de controle.

5 REFERÊNCIAS

- ALBUQUERQUE, C. L. 2012. O impacto da introdução da concorrência no mercado de geração de energia elétrica no Brasil.
- BOLLE, F. 2001. Competition with supply and demand functions. *Energy Economics*, 23, 253-277.
- BOX, G. E. P. & COX, D. R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society (B)*, 26 (2), 41.
- BOX, G. E. P., JENKINS, G. M. & REINSEL, G. C. 2011. *Time series analysis: forecasting and control*, Wiley.
- BUNN, D. W. 1985. Statistical efficiency in the linear combination of forecasts. *International Journal of Forecasting*, 1, 151-163.
- CÂMARA, A. A. F. & SOARES, P. B. D. O Direito da Eletricidade e sua Regulação no Brasil. 2012.
- CASCAES, J. C. 2003. Critérios de Planejamento. *Canal Energia*. Rio de Janeiro.
- CHRISTO, E. S. & FERREIRA, M. B. 2011. Previsão e Controle de Energia.
- CHRISTO, E. S. & SOUZA, R. C. 2006. Uma abordagem estatística para a previsão de potência reativa em sistemas elétricos. *Pesquisa Operacional*, 26, 361-381.
- DA SILVA, H. F. 2001. *Um Sistema Integrado de Monitoração e Previsão de Carga Elétrica de Curto Prazo*. University of London.
- DE MIRANDA, C. V. C. 2007. *Previsão de dados de alta frequência para carga elétrica usando holt-winters com dois ciclos*. Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro.
- DE SOUZA, L. L. C., MALDONADO, M. U. & RADOS, G. J. V. 2011. Gestão da terceirização no setor brasileiro de distribuição de energia elétrica. *RAE - Revista de Administração de Empresas*, 188-201.
- DOS SANTOS, S. M. & DE SOUZA, M. P. 2011. Análise das contribuições potenciais da Avaliação Ambiental Estratégica ao Plano Energético Brasileiro. *Eng Sanit Ambient*, 16, 369-378.
- DOUGLAS, A. P., BREIPOHL, A. M., LEE, F. N. & ADAPA, R. 1998. The impacts of temperature forecast uncertainty on Bayesian load forecasting. *Power Systems, IEEE Transactions on*, 13, 1507-1513.
- EL-KEIB, A., MA, X. & MA, H. 1995. Advancement of statistical based modeling techniques for short-term load forecasting. *Electric Power Systems Research*, 35, 51-58.

- GEURTS, M. D. & IBRAHIM, I. 1975. Comparing the Box-Jenkins approach with the exponentially smoothed forecasting model application to Hawaii tourists. *Journal of Marketing Research*, 182-188.
- HAGAN, M. T. & BEHR, S. M. 1987. The time series approach to short term load forecasting. *Power Systems, IEEE Transactions on*, 2, 785-791.
- HARVEY, A. & KOOPMAN, S. J. 1993. Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, 88, 1228-1236.
- HARVEY, A. C. 1991. *Forecasting, structural time series models and the Kalman filter*, Cambridge university press.
- HIPPERT, H. S., PEDREIRA, C. E. & SOUZA, R. C. 2001. Neural networks for short-term load forecasting: A review and evaluation. *Power Systems, IEEE Transactions on*, 16, 44-55.
- HSU, Y. Y. & HO, K. L. Fuzzy expert systems: an application to short-term load forecasting. 1992. IET, 471-477.
- ISHIKAWA, K. & LOFTUS, J. H. 1990. *Introduction to quality control*, Tokyo, 3A Corporation.
- LAURET, P., DAVID, M. & CALOGINE, D. 2012. Nonlinear Models for Short-time Load Forecasting. *Energy Procedia*, 14, 1404-1409.
- LU, C. W. & REYNOLDS, M. R. 1999. EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, 31, 166-188.
- LU, C. W. & REYNOLDS, M. R. 2001. Cusum charts for monitoring an autocorrelated process. *Journal of Quality Technology*, 33, 316-334.
- MONTGOMERY, D. C. 2009. *Introduction to statistical quality control*, Hoboken, N.J., Wiley.
- MONTGOMERY, D. C. & JOHNSON, L. A. 1976. *Forecasting and time series analysis*, New York, McGraw-Hill.
- MOREIRA, M. M. S. C. 2011. Reestruturação e privatização no setor elétrico brasileiro: impactos sobre as relações de trabalho.
- NGUYEN, T. & LIAO, Y. 2011. Short-Term Load Forecasting Based on Adaptive Neuro-Fuzzy Inference System. *Journal of Computers*, 6, 2267-2271.
- PAPARODITIS, E. & SAPATINAS, T. 2012. Short-Term Load Forecasting: The Similar Shape Functional Time Series Predictor. *arXiv preprint arXiv:1201.2617*.
- ROSIM, S. O. 2011. Geração de energia elétrica-Um enfoque histórico e institucional das questões comerciais no Brasil.
- SADOWNIK, R. & BARBOSA, E. P. 1999. Short-term forecasting of industrial electricity consumption in Brazil. *Journal of Forecasting*, 18, 215-224.
- SCHWARZ, G. 1978. Estimating the dimension of a model. *The annals of statistics*, 6, 461-464.
- SHRIVASTAVA, V., MISRA, R. & BANSAL, R. 2012. An assessment of electrical load forecasting using artificial neural network. *International Journal of Computer Aided Engineering and Technology*, 4, 80-89.
- SILVA, B. G. 2012. Evolução do setor elétrico brasileiro no contexto econômico nacional: uma análise histórica e econométrica de longo prazo.
- SILVA, E. F. 2011. Principais condicionantes das alterações no modelo de comercialização de energia elétrica: retrospectiva e análise crítica.
- SILVA, M. E. & GÓMEZ, C. P. 2011. O papel do governo ea prática do consumo sustentável: como esse stakeholder atua no setor elétrico?[doi: 10.5329/RECADM. 20111002008]. *Revista Eletrônica de Ciência Administrativa-RECADM*, 10, 107-123.
- SOUSA, J., JORGE, H. & NEVES, L. Forecasting the next day load profile using load profiling information and meteorological variables. 2011. IEEE, 1-6.
- SOUZA, R. 1983. Métodos automáticos de Amortecimento Exponencial para previsão de séries temporais. *Monografia GSM-10/83, maio*.
- SOUZA, R. C. & CAMARGO, M. E. 1996. Análise e previsão de séries temporais: os modelos ARIMA. *Ijuí, RS: Sedigraf*, 242.
- WARDELL, D. G., MOSKOWITZ, H. & PLANTE, R. D. 1994. Run length distributions of residual control charts for autocorrelated processes. *J QUAL TECHNOL*, 26, 310-317.
- YANG, H. T., HUANG, C. M. & HUANG, C. L. Identification of ARMAX model for short term load forecasting: an evolutionary programming approach. 1995. IEEE, 325-330.
- YASHCHIN, E. 1993. Performance of CUSUM control schemes for serially correlated observations. *Technometrics*, 35, 37-52.
- YU, Z. 1996. A temperature match based optimization method for daily load prediction considering DLC effect. *Power Systems, IEEE Transactions on*, 11, 728-733.