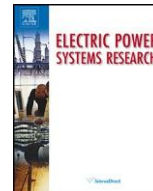


Listas de conteúdos disponíveis no ScienceDirectInvestigação de Sistemas de
Energia Eléctricao u r n a l h o m e p á g i n a : www.elsevier.com/locate/epsr

Um novo modelo à escala multitemporal para previsão da procura de energia eléctrica: Do horizonte de curto a médio prazo

Kianoosh G. Boroojenia, M. Hadi Aminib^{c,d,e,*}, Shahab Bahramif, S.S. Iyengara,
Arif I. Sarwatg, Orkun Karabasoglu^{b,c,d,e,*}^a School of Computing and Information Sciences, Florida International University, Miami, FL 33199, EUA^b Sun Yat-Sen University - Carnegie Mellon University Joint Institute of Engineering, Guangzhou, Guangdong 510006, China^c Departamento de Engenharia Electrotécnica e Informática, Carnegie Mellon University, Pittsburgh, PA 15213, EUA^d SYSU-CMU Shunde International Joint Research Institute, Shunde, China^e School of Electronics and Information Technology, SYSU, Guangzhou, China^f Departamento de Engenharia Electrotécnica e Informática, Universidade de British Columbia, Vancouver, BC, Canadá^g Departamento de Engenharia Electrotécnica e Informática, Florida International University, Miami, FL 33174, EUA

a r t i c l e i n f o

História do artigo:

Recebido a 22 de Abril de 2016

Recebido na forma revista 23 de Agosto

de 2016 Aceite 26 de Agosto de 2016

Palavras-chave:

Modelo autoregressivo

Modelo de média

móvel Previsão da série

temporal

Previsão da procura de energia

eléctrica Critério de

informação Akaike Critério de

informação Bayesian

a b s t r a c t

A previsão de carga a curto prazo é essencial para o funcionamento fiável e económico dos sistemas de energia. A previsão a curto prazo cobre uma gama de previsões desde uma fracção de uma hora de antecedência até uma previsão de um dia de antecedência. Uma previsão de carga precisa resulta no estabelecimento de práticas operacionais e estratégias de licitação adequadas, bem como na programação de transacções de energia adequadas. Este documento apresenta uma técnica generalizada de modelação de dados históricos de carga sob a forma de séries temporais com diferentes ciclos de sazonalidade (por exemplo, diário, semanal, trimestral, anual) numa determinada rede eléctrica. O método proposto modela separadamente os ciclos não sazonais e sazonais dos dados de carga utilizando componentes auto-regressivos (AR) e de média móvel (MA), que apenas se baseiam em dados históricos de carga sem exigir quaisquer entradas adicionais, tais como dados meteorológicos históricos (que podem não estar disponíveis na maioria dos casos). A precisão da modelação de dados é examinada utilizando os critérios de informação Akaike/Bayesian (AIC/BIC), que são dois métodos eficazes quantificação para avaliação da previsão de dados. A fim de validar a eficácia e exactidão da previsão proposta, utilizamos os dados de carga horária da rede PJM como um conjunto de dados de entrada no mundo real.

© 2016 Elsevier B.V. Todos os direitos reservados.

1. Introdução

1.1. Motivação

A previsão da procura de electricidade desempenha um papel fundamental na gestão dos sistemas de energia, especialmente para fins de operação e manutenção [1]. É particularmente importante para sistemas de energia desregulamentados, onde as imprecisões da previsão têm implicações em significativo para operadores de mercado, proprietários de transmissões e participantes no mercado. A previsão de carga é categorizada com base na escala temporal em curto, médio, e longo prazo.

* Autores correspondentes em: Sun Yat-sen University-Carnegie Mellon University Joint Institute of Engineering (JIE), School of Electronics and Information Technology, SYSU, Guangzhou, Guangdong, China; Carnegie Mellon University, Pittsburgh, PA, EUA; SYSU-CMU Shunde International Joint Research Institute, Guangdong, China. Tel.: +86 13226996631.

Endereços electrónicos: kghol002@fiu.edu (K.G. Boroojenia), amini@cmu.edu (M.H. Amini), bahramis@ece.ubc.ca (S. Bahrami), iyengar@cis.fiu.edu (S.S. Iyengar),

asarwat@fiu.edu (A.I. Sarwat), karabasoglu@cmu.edu (O. Karabasoglu).

previsão. Estes três tipos são utilizados para programação e controlo do sistema de potência, operação e planeamento, e planeamento da geração/ expansão da transmissão, respectivamente [1,2]. A previsão de carga a curto prazo (STLF) é necessária para a programação da geração, avaliação de segurança da operação do sistema, e informação de expedição económica horária [3]. Do ponto de vista da procura, existe uma exigência de estimar com precisão a procura para se conseguir uma operação mais fiável do sistema de potência [4,5]. Futuros sistemas de energia, nomeadamente redes inteligentes, estão a emergir com o conceito de infra-estruturas

avanzadas de medição para melhorar a fiabilidade dos sistemas convencionais de energia no lado da procura [6,7]. Além disso, a gestão do lado da procura é amplamente utilizada para aplicações residenciais [8,9] e de controlo de carga industrial [10] e aplicações de centros de energia inteligentes [11]. Embora a utilização da resposta da procura e de outros recursos do lado da procura melhore o funcionamento fiável dos sistemas de energia, a previsão precisa da procura é uma obrigação inevitável para manter o equilíbrio entre a carga e a geração. A previsão precisa da procura irá melhorar o desempenho em tempo real e a longo prazo dos sistemas de energia, com base nos dados históricos disponíveis.

<http://dx.doi.org/10.1016/j.epsr.2016.08.031>

0378-7796/© 2016 Elsevier B.V. Todos os direitos reservados.

Nomenclatura

L^i	o operador de desfasamento na série temporal x_t devolvendo o seu i -ésimo valor anterior x_{t-i}
O	operador de diferenciação na série temporal x_t retornando $x_t - x_{t-1}$
$\text{corr}(\cdot, \cdot)$	a função de correlação que recebe duas variáveis aleatórias como input e devolve as suas correlações
AC_x	a função de auto-correlação de um processo de stationário de amplo sentido x_t
PAC_x	a função de auto-correlação parcial de um processo estacionário de amplo sentido x_t
$P_{t,l}(x)$	aprojecção da série temporal x para o espaço abrangido por $x_{t+1}, x_{t+2}, \dots, x_{t+l-1}$
X_t	série temporal representando a carga diária original
\tilde{X}_t	séries temporais representando a carga diária prevista
r_d	transformação logarítmica Box-Cox para o homogéneo - dimensionando a variância do conjunto de treino diário
$D_{\text{semanal}}^{(d)}$	transformação diferenciada semanal num tempo
$D_{\text{anual}}^{(d)}$	transformação anual diferenciadora num tempo
$D_{\text{diária}}^{(h)}$	transformação diária diferenciada numa série temporal com valores horários
$D_{\text{semanal}}^{(h)}$	transformação diferenciada semanal num tempo
I_t	valores de carga horária transformados como um tempo estacionário
ϵ	transformou os valores de carga diária como uma série temporal estacionária
L	probabilidade máxima de um modelo de previsão
K	número de parâmetros estimados de um modelo de previsão
t	série cronológica de previsão correspondente a x_t
$\phi_t^{(h)}$	auto-regressivo de hora em hora
O_t	média móvel diária coefficient de I_t
$O_t^{(w)}$	média móvel semanal coefficient de I_t
$\epsilon_t^{(w)}$	valores de erro de I_t no modelo ARMA
$\epsilon_t^{(d)}$	valores de erro de I_t no modelo ARMA
ϕ_t	auto-regressivo diário coefficient de
$y_t^{(w)}$	média móvel semanal coefficient de I_t
$y_t^{(d)}$	média móvel anual coefficient de I_t
Y_t	série cronológica <i>ptresidual</i> da previsão do dia
\tilde{Y}_t	série cronológica representando os valores originais de carga horária
\tilde{Y}_t	séries temporais que representam a carga horária prevista
r_h	transformação logarítmica Box-Cox para o homogéneo - dimensionando a variância do conjunto de treino horário

previsão (LTLF) [19,20] que um inquérito bibliográfico representou 6%, 58%, 20% e 16% dos esforços de investigação anteriores centrados nestes horizontes de previsão de carga, respectivamente. Os autores em [13] aplicaram artificial neural networks to model load dynamics for VSTLF application. A abordagem VSTLF proposta foi testada para a previsão de carga em linha numa empresa de electricidade nos Estados Unidos. A rede neural ensacada é implantada para STLTF em [21]. Em [14], são utilizadas janelas de 5 minutos de tempo em movimento para determinar a carga horária à frente. Este documento adoptou redes neuronais wavelet com dados pré-filtering para prever a carga em espaços de tempo muito curtos, minimizando o efeito de dados ruidosos. O conjunto de dados ISO New England foi utilizado para a validação da abordagem proposta em [14]. Piras et al. em [15] propuseram uma arquitectura de rede neural heterogénea - composta de uma parte não supervisionada para detectar algumas características dos dados e sugerir variáveis de regressão. Para obter uma transição suave entre submodelos, foi utilizada uma média fuzzy ponderada para integrar as saídas de cada submodelo.

Amjady em [16] compôs um estimador estatal de previsão (FASE) e a rede neural multi-layer perceptron (MLP) para construir uma previsão de carga a curto prazo. O método treina o MLP para determinar a função de mapeamento que é necessária para a FASE (input características) e a saída (carga real). Dove et al. implementaram dif...

procedimentos de regularização ferente para fins de formação em redes neurais para previsão de carga a médio prazo. Esta abordagem referia-se a como modelo de rede neural de alimentação para a frente (FNN) [18]. Em [19], foi desenvolvido um sistema especializado baseado no conhecimento para a escolha do modelo de previsão anual mais apropriado. O modelo seleccionado foi então utilizado para o planeamento do sistema de potência a médio/longo prazo. Em [22], são utilizados métodos de aprendizagem multi-tarefa baseados em kernel para prever a procura de energia eléctrica na rede de distribuição. Os autores em [20], propuseram um método de previsão para prever o pico de procura a longo prazo. Diferentes incertezas que podem afectar o pico de procura foram considerados durante o LTLF, incluindo o crescimento populacional, condições económicas, e condições climáticas. Além disso, a utilização em larga escala de veículos eléctricos pode aumentar a incerteza devido a diferentes comportamentos de condução e padrões de carregamento [23–26].

Também pode ser utilizado para os problemas de expansão do sistema eléctrico [27].

Para além da faixa horária baseada em classification de previsão métodos, podemos classificar as abordagens existentes de acordo com os

técnicas aplicadas. Um método é utilizar o padrão histórico de carga como uma série cronológica para prever a procura utilizando a série cronológica anal...

métodos ysis. O segundo método é baseado na correlação entre o padrão de carga e as variáveis meteorológicas. Esta abordagem constrói a relação entre a carga histórica e as condições climáticas para prever a procura.

Abordagens tradicionais de previsão de carga, tais como

regressão e interpolação, pode não conduzir a resultados precisos. Por outro lado, métodos de previsão complexos, que são computacionalmente...

carregados convergem mais lentamente. Vários estudos têm sido centrados em técnicas de previsão, incluindo abordagem lógica difusa e artificial rede neural [16,28], regressão linear [29], e mineração de dados [30], funções de transferência [31], estatísticas Bayesianas [32], previsão crítica [33], e modelos dinâmicos cinzentos [34]. Em [20], uma metodologia para prever a procura de electricidade até 10 anos à frente foi proposto. Artificial redes neurais (ANN) tem sido amplamente

$p_t^{(h)}$

série temporal residual da previsão de hora de antecedência

utilizado para previsão da procura de electricidade [35–37]. De acordo com [35], a regra geral Delta (GDR) foi utilizada para o treino de neurónios em ANN e o vector de saída é utilizado como um padrão de entrada para a rede. Em [36], a ANN é utilizada para a previsão da procura a curto prazo. de Lee et al. usaram ANN para os dias de semana e fins-de-semana separadamente para alcançar resultados mais precisos. Em [38], a rede neural artificial é utilizada para prever as interrupções nas redes inteligentes com base nas condições meteorológicas e interrupções históricas na rede analisada. Em [37], é fornecido um levantamento bibliográfico abrangente da aplicação da rede neural na previsão da procura a curto prazo. Amjady em

1.2. Revisão literária

Um inquérito exaustivo sobre abordagens de previsão da procura foi fornecido em [12]. Este documento classifi ed os métodos em quatro grupos: previsão de carga a muito curto prazo (VSTLF) [13,14], STLF [15–17], previsão de carga a médio prazo (MTLF) [18], e carga a longo prazo

[39] propôs uma previsão de curto prazo que diferencia entre dias de semana, fins de semana e feriados públicos para melhorar a precisão. Existem quatro grandes preocupações relacionadas com os estudos anteriores sobre a previsão da procura: (1) longo processo de previsão de carga (flow do processo de previsão de carga é demasiado longo) como [40]; (2) pequeno conjunto de dados é utilizado para validação do modelo [41]; (3) grande quantidade de dados é necessária na fase de formação [42]; e (4) grande erro na previsão [43].

A procura de carga eléctrica durante um período de tempo é uma série temporal sazonal não estacionária. Muitas tentativas foram feitas por estatísticos para criar modelos de previsão para este tipo de séries temporais. Dudek em [44] propôs modelos de regressão linear para previsão de carga de curto prazo baseada em padrões, nos quais múltiplos ciclos sazonais da série temporal de previsão é filtered fora e não-estacionários em média e a variância é eliminada. Foi desenvolvido um procedimento passo a passo [45] para a modelação sazonal não estacionária da série temporal aplicada seguindo a metodologia Box-Jenkins, que é uma metodologia de modelação conhecida first created by Box e Jenkins em 1976 [46]. Além disso, Pappas et al. em [47] apresentou um novo método de previsão da procura de electricidade - utilizando a teoria da partição multi-modelo, first filters filters para a sazonalidade e não-estacionária dos dados reais utilizando a caixa de ferramentas modeling e previsão de cargas e preços de electricidade (MFE) para Matlab e depois aplica um modelo multi-variável de média móvel auto-regressiva (ARMA) para prever a procura de electricidade do sistema eléctrico helénico. Além disso, a Desouky e a Elkateb, em [48] utilizados um híbrido dos métodos ARMA e ANN, a fim de obter uma previsão mais promissora em comparação com trabalhos anteriores usando o método da série. Além disso, Huang et al. propuseram uma previsão de carga a curto prazo baseada no modelo ARMA, incluindo considerações não-Gaussianas do processo [49]. Além disso, Contreras et al. em [50] analisaram diferentes modelos ARIMA para prever a carga de electricidade no dia seguinte.

1.3. A nossa contribuição

Este documento apresenta uma técnica generalizada para modelar os seus dados de carga tórica sob a forma de séries temporais com diferentes ciclos de sazonalidade (por exemplo, diário, semanal, trimestral, anual) numa determinada rede eléctrica. O método proposto modela separadamente os ciclos não sazonais e sazonais dos dados de carga utilizando componentes auto-regressivos (AR) e de média móvel (MA), que se baseiam apenas em dados históricos de carga sem requerer quaisquer entradas adicionais, tais como como dados meteorológicos históricos (que podem não estar disponíveis na maioria dos casos). A precisão da modelação de dados é examinada não só por calculando os erros de previsão convencionais mas utilizando os critérios de informação Akaike/Bayesian (AIC/BIC) que são dois métodos eficazes quantificação que penalizam a complexidade de um modelo e recompensam a sua fitness e exactidão. Se a previsão flexibility é a nossa preocupação mais importante, ou seja, o nosso principal objectivo é desenvolver uma previsão que possa ser amplamente utilizada para diferentes conjuntos de formação de dados sem deteriorar o desempenho da previsão, o BIC é uma alternativa apostada (em comparação com o AIC) para quantificar a utilidade da previsão, uma vez que o BIC penaliza mais a complexidade dos modelos de previsão do que o AIC. Contudo, o AIC é um critério melhor se optarmos por comprometer o modelo flexibility em certa medida para ganhar mais fitness e melhorar a precisão da previsão através da redução do erro de previsão. As principais contribuições deste documento são as seguintes:

- Melhorar a metodologia Box-Jenkins para modelar os seus dados de carga de electricidade tórica durante um período de tempo, a

- Usando o teste baseado no periodograma de Bartlett para provar que a série temporal residual (erro de previsão) é um ruído branco e significa - menos. Ou seja, o modelo proposto extraiu qualquer informação significativa do conjunto de formação de entrada, a fim de obter a previsão mais precisa.
- Usando critérios de informação Akaike/Bayesian (AIC/BIC) em vez de as medições de erro convencionais para avaliação de modelos e fine-tuning: A AIC/BIC trata do compromisso entre a apropriação do modelo fitness e a complexidade do modelo. Considera algum valor de penalização para modelos complexos, que só podem funcionar precisamente para o conjunto de formação sub-estudado; contudo, as medições de erros convencionais dependem consideravelmente do conjunto de formação dado e medem o fitness do modelo apenas no que diz respeito ao conjunto de formação.
- Finalmente, mostramos que, alterando o tamanho ou o tempo de formação o modelo permanece robusto e o erro de previsão quase se mantém constante.

1.4. Organização do jornal

O resto do presente documento está organizado da seguinte forma. Secção 2 introduz preliminares da modelação da série temporal. A secção 3 apresenta uma explicação detalhada do modelo proposto à escala temporal múltipla. Na Secção 4, é apresentado um estudo de caso prático e é ilustrado o desempenho superior da nossa metodologia inovadora. A Secção 5 apresenta o papel e os resultados da previsão. A representação detalhada do conjunto de dados utilizados é fornecida no apêndice.

2. Preliminares

Uma série temporal é uma sequência de pontos de dados, normalmente constituída por medições sucessivas feitas ao longo de um intervalo de tempo. As séries temporais são muito frequentemente traçadas através de gráficos de linhas e são utilizadas em qualquer field que envolva medições temporais. A análise baseada em séries temporais compreende métodos de análise de dados, a fim de extrair estatísticas significativas e outras características dos dados. É também utilizada para a previsão que se baseia na utilização de um modelo para prever valores futuros com base em valores previamente observados. A série temporal x_t é estacionária de largo sentido (WSS) se a sua média não variar ao longo do tempo ($E[x_t] = \mu_x$) e a sua função de autocorrelação, defined as:

$$R_x(t_1, t_2) = \frac{1}{E[(x_{t_1} - \mu_x)(x_{t_2} - \mu_x)]}$$

a fim de criar uma previsão precisa da procura de energia eléctrica.

- Acrescentar múltiplos ciclos de sazonalidade para criar uma escala temporal múltipla modelação para previsão da procura de energia eléctrica, que é mais

depende apenas do valor de $(t_2 - t_1)$ para cada ponto de observação. Uma série temporal sazonal do ciclo sazonal T é chamada estacionária ao longo do seu ciclo sazonal se

$$\mathbb{E}[x_t] = \mathbb{E}[x_{t+T}],$$

e $R_x(t, t + nT)$ é independente de t para cada $t > 0$; por exemplo, se uma série temporal com valor diário e sazonal semanal for *semanal-estacionária* se tiver constantes autocorrelações em múltiplos de 7.

flexível para a natureza diária, semanal e anual sazonal dos dados de electricidade. Consequentemente, o nosso modelo pode ser considerado como um modelo de média móvel integrada auto-regressiva sazonal alargada (SARIMA) [51].

In statistics, an auto-regressive (AR) model is a representation of a type of random process showing (wide-sense) stationary behavior. In contrast with the regression analysis which models the output variable as a function of multiple independent variables, the autoregressive model considers the output variable as a linear function of its own previous values and on a stochastic term (a non-deterministic, imperfectly predictable term). Thus, the model is in the form of a stochastic difference equation:

$$x_t = \sum_{i=1}^p cpiL^i x_t + \varepsilon_t,$$

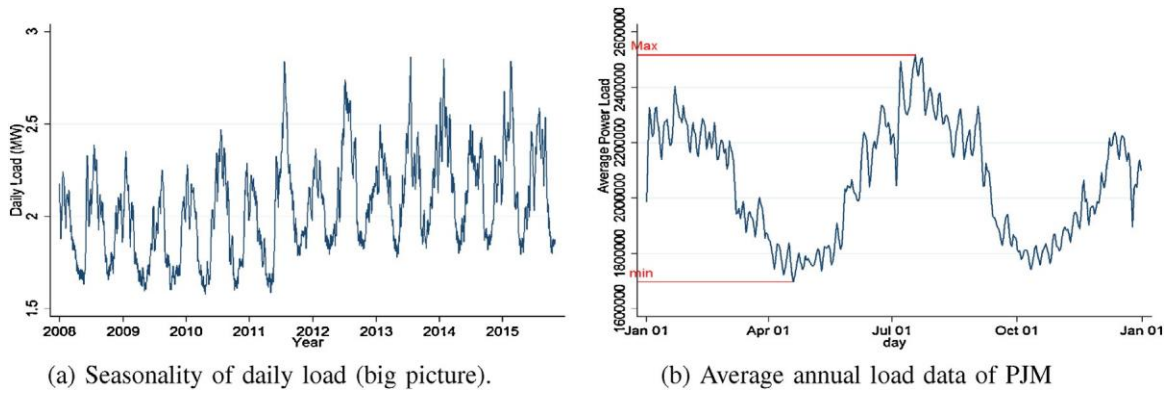


Fig. 1. Análise estatística dos dados de carga da rede PJM durante um período de 8 anos.

onde x_t especifica o processo aleatório de uma dada série temporal, cp_i é o coeficiente do termo i th AR, L é o operador de atraso agindo em x_t como $L : x_t \rightarrow x_{t-1}$, e ε_t denota o termo estocástico chamado *valor de erro* do modelo.

AR model is a special case of the more general ARMA model of time-series, which has a more complicated stochastic structure. Given a time-series of data x_t where t is an integer index and the x_t are real numbers, an ARMA(p, q) model is given by:

$$x_t = \underbrace{cp_1 L^1}_{i=1} x_t + 1 + \underbrace{0iL^i}_{i=1} \varepsilon_t, \quad (1)$$

AR-PartMA-Part

onde 0_i 's são os parâmetros da parte média móvel.

Estes modelos são fitted para dados da série temporal para prever pontos futuros na série (previsão). Um modelo ARIMA é um extensão de um ARMA. São aplicados em alguns casos em que os dados são não estacionariedade e uma etapa inicial de diferenciação (correspondente à parte "integrada" do modelo) pode ser aplicada para reduzir a não estacionariedade. De facto, o modelo ARIMA(p, d, q) pode ser visto como uma "cascata" de dois modelos. O first aplica-se a uma série temporal não-estacionária x_t :

$$x_t I = (1 - L)^d x_t$$

enquanto a segunda parte modela a série temporal WSS $x_t I$ usando ARMA(p, q) descrito em Eq. (1).

Quando dois dos três termos são zeros, o modelo pode ser referido com base no parâmetro não-zero, deixando cair "AR", "I" ou "MA" do acrónimo correspondente. Por exemplo, ARIMA(1,0,0) é AR(1), ARIMA(0,1,0) é I(1), e ARIMA(0,0,1) é MA(1). Antes de explicando a metodologia proposta, introduzimos duas funções que são atribuídas a uma série temporal WSS x_t : função de autocorrelação1 (ACF) e função de autocorrelação parcial2 (PACF). O comportamento destas duas funções revela como uma série temporal de A&S pode ser formulada como um modelo ARMA. Mais detalhes sobre como o comportamento da ACF e PACF é interpretado são apresentados na secção seguinte.

¹ A ACF de uma série temporal WSS x_t é defined como função $AC_X(I) = \text{corr}(x_{t+I}, x_t)$.

$\text{corr}(\cdot, \cdot)$ é a função de correlação defined como $\text{corr}(x_{t+I}, x_t) = \frac{E[(x_t - \mu_X)(x_{t+I} - \mu_X)]}{\sigma_X^2}$,

onde μ_X e σ_X^2 são a média e a variação do processo estacionário x_t , respectivamente.

3. A metodologia proposta

Os valores de carga de potência ordenados cronologicamente de uma parte specific dos sistemas de potência é uma série temporal que ocorre ao longo de um período de tempo de uma forma sazonal. Para modelar uma série temporal como um funcionamento dos seus valores passados, analisamos o padrão com o pressuposto de que o padrão geral persistirá no futuro. Esta secção é dedicada a identificar o melhor modelo que corresponda ao comportamento estatístico que os dados da procura de electricidade observada mostram ao longo de alguns anos. A qualidade da modelação da série temporal da carga de energia determina a precisão da estimativa e previsão dos dados da carga de energia no futuro. A fim de avaliar o modelo identification, a secção utiliza os critérios de informação Akaike/Bayesian que nos ajudam a figure para saber até que ponto a série temporal estimada fits encerra os dados da carga de energia observada. A análise da série temporal residual (a diferença real entre a potência de carga observada e estimada) pode também determinar a qualidade de um modelo em que, se a série temporal residual não for determinista e não tiver parte significativa, o modelo representa com precisão os dados observados. Durante esta secção, utilizamos os dados de carga diária real da rede PJM durante o período de 2008–2014 como o nosso conjunto de formação, a fim de ilustrar os passos da criação do modelo. Em seguida, utilizamos os dados de carga da mesma rede em 2015 para avaliar o modelo e comparar os valores observados e previstos. A figura 1 mostra uma visão estatística sobre os dados de carga da rede PJM.

3.1. Esboço da metodologia proposta

Em primeiro lugar, verificamos se a série temporal que representa os dados de carga tem uma variação homogênea. Caso contrário, aplicamos uma transformação logarítmica adequada à série temporal para a tornar homogênea a variância. Para obter a transformação eficaz, aplicamos o método Box-Cox [52], que é explicado mais tarde.

Em segundo lugar, examinamos a parcela ACF da série temporal transformada a nível não sazonal e ciclos sazonais (semanal, anual, etc.) para find qualquer indicação de ser não estacionário. Se a ACF da série temporal cair ou declinar no first poucos desfasamentos, é considerada como estacionária. Se os valores ACF ou caiem após um número considerável de desfasamentos ou declinam bastante lentamente, deve ser marcado como não estacionário. Note que a ACF de uma série temporal pode apresentar comportamentos diferentes em diferentes ciclos sazonais; por exemplo, a ACF de uma série cronológica com desfasamentos semanais pode mostrar um sinal estacionário; enquanto,

considerando os desfasamentos diários, o mesmo ACF pode morrer rapidamente (não estacionário). A fim de filter, os índios não-estacionários...

cations, aplicamos repetidamente transformações diferentes sobre os

onde μx é a média de x_t .

z O PACF de uma série temporal estacionária de largo sentido x_t é representado pelo $PACx(l)$ e é definido como $ACx(l)$ se $l = 1$, e $\text{corr}(x_{t+l} - P_{t,l}(x_{t+1}), x_t - P_{t,l}(x_t))$ caso contrário; onde $P_{t,l}(x)$ denota a projecção de x para o espaço abrangido por $x_{t+1}, x_{t+2}, \dots, x_{t+l-1}$.

Finalmente, a série temporal estacionária transformada é modelada para um modelo de média móvel e/ou autoregressivo baseado na forma como ACF e PACF da série temporal se comportam. Finalmente, os parâmetros detalhados

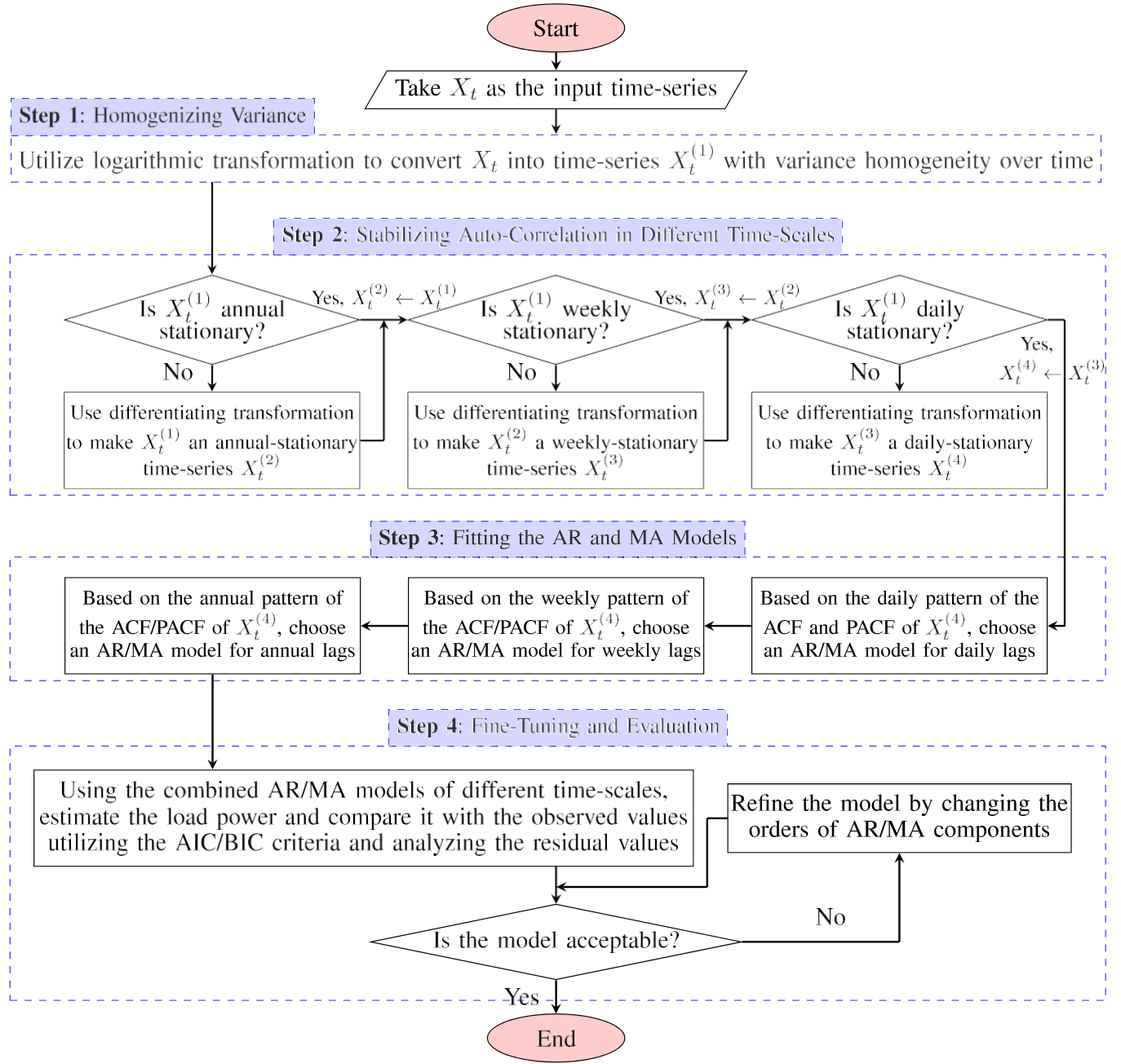


Fig. 2. A representação flowchart da metodologia proposta para a criação de um modelo de previsão de valores de carga diária.

do modelo é estimado através da tentativa de múltiplos valores e da validação das previsões utilizando o AIC. A figura 2 mostra a representação da metodologia proposta em flowchart.

3.2. Variação homogeneizante

Quando uma série temporal não tem homogeneidade de variância,³ algumas transformações são aplicadas na série temporal a fim de homogeneizar (estabilizar) a sua variância ao longo do tempo. Aqui, utilizamos um caso especial de transformação Box-Cox [52] para obter uma série temporal homogênea. É comum que o desvio padrão de uma série temporal (especialmente dados de carga eléctrica) mude proporcionalmente com a sua média ao longo do tempo,

i.e. $Var[x_t] = (a\mu_t + b)^2$ para alguns a e b . Let r_d denotam a função que transforma x_t a uma série temporal com variação homogênea: $Var[r_d(x_t)] = \text{constante}$. Assumindo que rdI representa a derivada função de r_d , aproximamos $r_d(x_t)$ pela sua série first-order Taylor em torno da sua média μ_t :

$$rd(x_t) \approx rd(\mu_t) + rdI(\mu_t)(x_t - \mu_t).$$

Consequentemente, obtemos a seguinte aproximação para a variância da série temporal transformada $r_d(x_t)$:

$$\begin{aligned} Var[rd(x_t)] &\approx Var[rd(\mu_t) + rdI(\mu_t)(x_t - \mu_t)] \\ &= (r_d'(\mu_t))^2 Var[x_t] = (r_d'(\mu_t)(a\mu_t + b))^2. \end{aligned} \quad (2)$$

Para obter a série temporal $r_d(x_t)$ com variância-homogeneidade, a condição de dobradiças deve manter-se: $rdI(\mu_t)(a\mu_t + b) = \text{constante}$. Ou seja,

³ Uma série temporal tem homogeneidade de variância se a sua variância não se alterar ao longo do tempo.

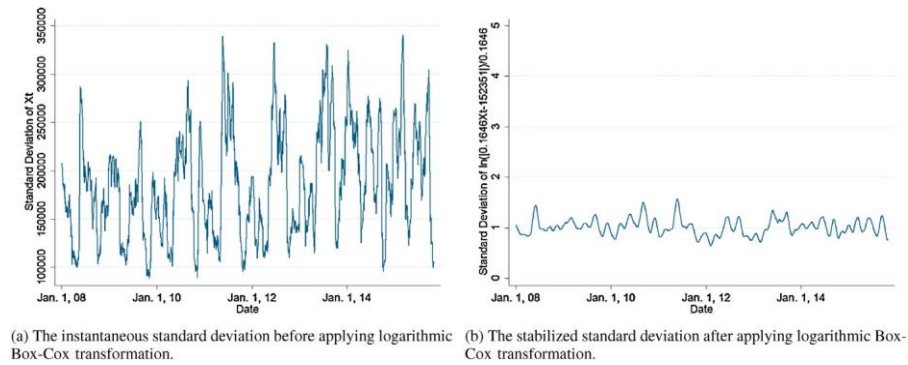


Fig. 3. Homogeneização da variação dos valores diários de carga eléctrica na rede eléctrica PJM durante o período de 2008–2014.

r_d é uma transformação logarítmica: $r_d(X_t) = \ln(|ax_t + b|)/a$. Os valores de a e b são obtidos por análise de regressão linear de a_t (desvio padrão de X_t); onde μ_t é a variável independente.

In order to illustrate the aforementioned variance-homogenizer method, consider the daily values of electric load in PJM power network over the period of 2008–2014. Let X_t denote the time-series representation of the daily load values. Fig. 3 shows the parcela de desvio-padrão a_{X_t} de X_t ($\text{Var}(X_t)$) ao longo do tempo. Pode

ver que a série temporal não é estacionária em variação. Utilizando uma regressão linear, obtemos a relação de $\ln(X_t) = -1646\mu_{X_t} + 152351$ entre média instantânea e desvio-padrão de a série temporal onde a^2 de determinação para a regressão linear é R coefficient = 0.86. Como resultado, a transformação Box-Cox para tornar a série temporal estacionária está na seguinte forma: $r_d(X_t) = \ln(0,1646X_t - 152351)/0,1646$.

3.3. Estabilização da auto-correlação em diferentes escalas de tempo

Qualquer série temporal que mostre semelhanças periódicas ao longo do tempo é referida como sazonal. Os valores da procura de energia eléctrica seguem múltiplos padrões periódicos. Por exemplo, os valores da procura de carga aos domingos seguem o mesmo padrão; enquanto que são diferentes das segundas-feiras". Além disso, os feriados especiais têm especific anualmente-repetitivos

padrões de carga. Os comportamentos sazonais aparecem na época exacta

desfasamentos L , $2L$, $3L$, $4L$, e $5L$. Para os dados de carga diária, a série temporal mostra um padrão repetitivo nos desfasamentos 7, 14, 21, etc. (sazonalidade semanal). Também, para as séries horárias de carga, os dados mostram a sazonalidade diária com $L = 24$, ou seja, os desfasamentos sazonais são 24, 48, 72, e assim por diante. Em geral, os valores da série temporal transformada são considerados estacionários se a ACF

satisfies as seguintes condições:

- (1) Caindo repentinamente ou diminuindo rapidamente ao nível não sazonal.

- (2) Caindo repentinamente ou diminuindo rapidamente nos ciclos sazonais (desfasamentos sazonais exactos ou próximos dos sazonais).

Caso contrário, estes valores não são considerados como (de largo sentido) stationários. Por exemplo, considerar novamente os dados de carga diária da rede PJM. Como mencionado anteriormente, $X^{(1)}$ é a série logarítmica de tempo de carga transformada com variância estabilizada. A figura 4(a) mostra a ACF parcela de $X^{(1)}$ para o first 1200 lags. Pode-se ver que o auto a correlação de dados não cai ou diminui após o first poucos anuais; em vez disso, diminui com um ritmo lento. Consequentemente, $X^{(1)}$ mostra um comportamento não estacionário no ciclo sazonal anual. Additionally, considerando a Fig. 4(b) que representa a parcela ACF de $X^{(1)}$ no (1) poucos desfasamentos semanais, a auto-correlação X first em semanal e os atrasos diários declinam muito lentamente. Como resultado, $X^{(1)}$ não é uma série temporal estacionária tanto em desfasamentos semanais (ciclo sazonal semanal) como em desfasamentos diários (nível não sazonal).

Doravante, a fim de tornar $X^{(1)}$ estacionário, precisamos de o transformar usando múltiplas transformações diferenciadoras a nível anual, semanal e diário. Transformando a série temporal $X_t^{(1)}$ utilizando a diferenciação anual $D_{\text{anual}}^{(d)} = 1 - L^{364}$, os valores anuais de correlação automática caem após o desfasamento anual first (ver o gráfico ACF de $X^{(2)}$ representado na Fig. 5(a)). Note-se que consideramos

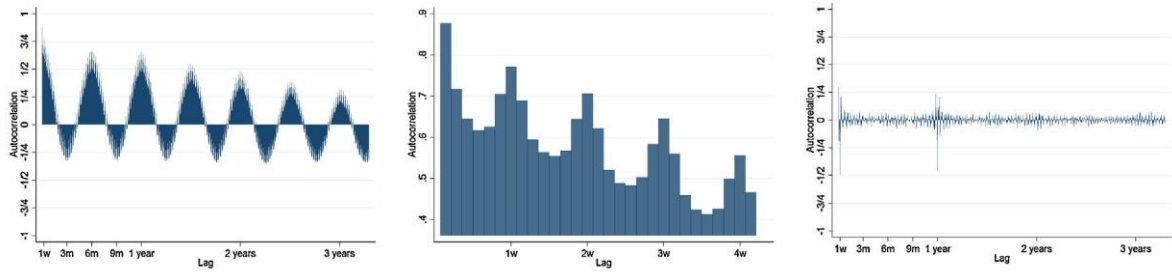
364 como o número de dias num ano, uma vez que é o múltiplo mais próximo de 7 ao tamanho de um ano e simplificará os cálculos posteriores. Além disso, ao transformar a série temporal $X^{(2)}$ utilizando o dif...

ferentificação $D_{\text{semanal}}^{(d)} = 1 - L^7$, os valores semanais de auto-correlação $D^{(d)}$ $(X^{(2)})$ (3)

cair substancialmente após o atraso de first-semana (ver o ACF de X_t =

$D_{\text{semanal}}^{(d)}(X_t^{(2)})$ na Fig. 5(b)). A diferenciação diária $D = 1 - L$ trans-

formulários $X^{(3)}$ numa série temporal com comportamento estacionário em todos os desfasamentos anuais, semanais, e diários. Como mostra a Fig. 5(c), a auto-correlação de



(a) ACF of $X_t^{(1)}$. The quarterly and annual seasonal patterns are illustrated. (b) ACF of $X_t^{(1)}$ showing its non-stationary behavior in the first few weekly and daily lags. (c) ACF of $X_t^{(4)}$ which is obtained by equation $X_t^{(4)} = D_{\text{annual}}^{(d)}(D_{\text{weekly}}^{(d)}(D_{\text{daily}}^{(d)}(X_t^{(1)})))$.

Fig. 4. Lotes ACF de dados de carga diária na rede PJM. Sub-figure (a) mostram a sazonalidade trimestral, e anual, dos dados de carga estabilizada de variância. Sub-figure (b) mostra o comportamento não estacionário dos dados, tanto no nível diário como no ciclo semanal. A sub-figure (c) ilustra como múltiplas transformações diferenciadoras nos levaram a obter uma série temporal estacionária em todos os níveis diários, semanais e anuais.

$$\text{dos dias seguintes (a parte não sazonal da série temporal) de } \frac{1}{2} + yI_{t-364} + \epsilon_t. \quad (4)$$

(ver os valores da segunda coluna do **Quadro 1**). Além disso, as autocorrelações parciais de ι_t nos dias seguintes são reduzidas exponencialmente (os valores da terceira coluna da mesma tabela suportam a reivindicação). O comportamento combinado das funções ACF e PACF a nível não sazonal, que é representado nas **Figs. 6c** e **f**, implica que os valores da carga diária transformada (ι_t) seguem uma

$$y_{t-1} = \underbrace{\alpha_0 + \alpha_1 y_{t-1}}_{\text{non-seasonal}} + \underbrace{\alpha_2 y_{t-12}}_{\text{weekly-seasonal}} + \underbrace{\alpha_3 y_{t-364}}_{\text{annual-seasonal}} + \varepsilon_t. \quad (4)$$

3.5. Afinação e avaliação

Na etapa de construção do modelo final, os parâmetros detalhados dos componentes AR/MA são determinados tentando múltiplos valores e finding o modelo de previsão com valores mínimos AIC/BIC.

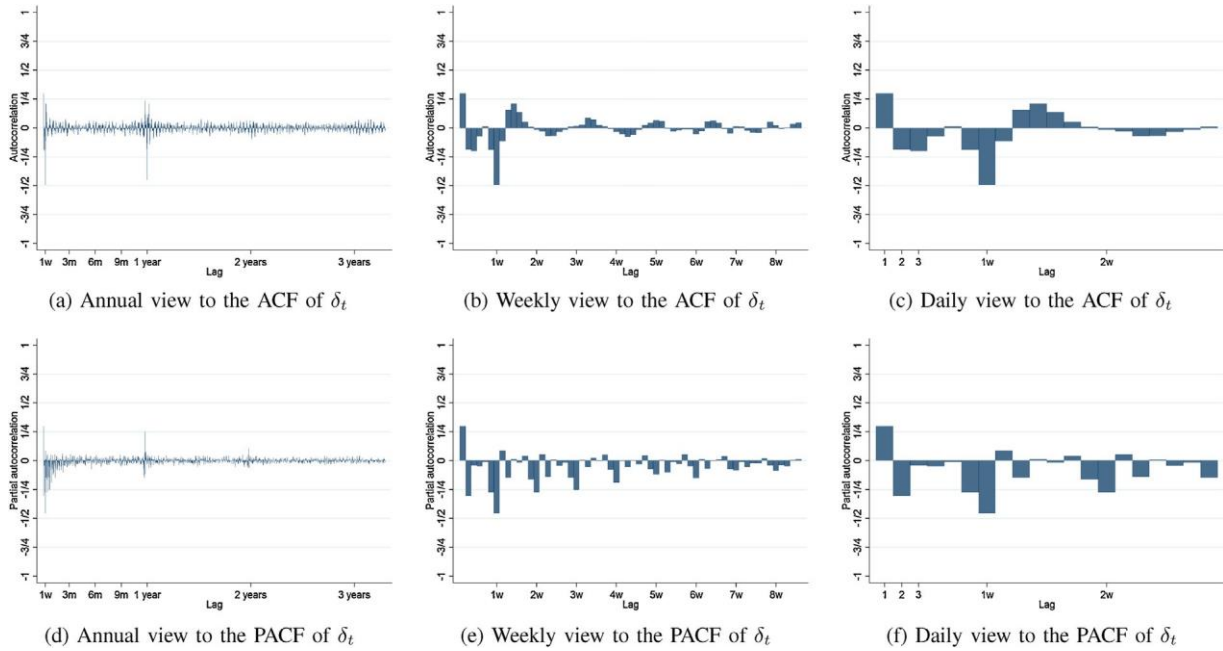


Fig. 6. Lotes ACF e PACF da série temporal estacionária δ_t obtidos através da transformação dos valores de carga diária da rede PJM.

3.5.1. AIC/BIC

O AIC é uma medida da qualidade relativa dos modelos estatísticos para um dado conjunto de dados. Dada uma colecção de modelos para os dados, a AIC estima a qualidade de cada modelo, em relação a cada um dos outros modelos. Assim, a AIC fornece um meio para a selecção de modelos. A AIC baseia-se na teoria da informação: oferece uma estimativa relativa da informação perdida quando um determinado modelo é utilizado para representar o processo que gera os dados. Para o efeito, trata do compromisso entre o fitness apropriado do modelo e o modelo com complexidade. A AIC não fornece um teste de um modelo no sentido de testar uma hipótese nula, ou seja, a AIC não pode dizer nada sobre a qualidade do modelo num sentido absoluto. Se todos os modelos candidatos \hat{f}_L são bons, a AIC não dará qualquer aviso a esse respeito. Suponhamos que temos um modelo estatístico de alguns dados. Que seja o valor máximo da função de probabilidade para o modelo; que seja o número de parâmetros estimados no modelo. Depois, o valor da AIC do modelo é o seguinte:

$$AIC = 2K - 2 \ln(L).$$

Dado um conjunto de modelos candidatos para os dados, o modelo preferido é o que tem o valor mínimo de AIC. Assim, a AIC recompensa a bondade de fit (avaliada pela função de probabilidade), mas também inclui uma penalização que é uma função crescente do número de esti-

parâmetros cruzados. A penalização desencoraja overfitting (o aumento do número de parâmetros no modelo melhora quase sempre a bondade do fit).

BIC é um critério de selecção de modelos entre um conjunto finito de modelos; o modelo com o BIC mais baixo é o preferido. Baseia-se, em parte, na função de probabilidade e está intimamente relacionado com o AIC. Quando modelos fitting, é possível aumentar a probabilidade adicionando parâmetros, mas fazê-lo pode resultar em sobre-fitting. Tanto o BIC como o AIC resolvem este problema introduzindo um termo de penalização para o número de parâmetros no modelo; o termo de penalização é maior no BIC do que no AIC. O BIC é formalmente definido como

$$BIC = -2 - \ln L + K - \ln(n),$$

onde n é o tamanho dos dados históricos em que o modelo se baseia.

3.5.2. Aperfeiçoamento do modelo

Como mencionado anteriormente na Eq. (4), o comportamento da ACF e PACF da série temporal δ_t sugere que deveria haver dois termos daily-AR, um termo semanal-MA, e um termo anual-MA no modelo. Nesta etapa, construímos diferentes modelos de previsão com múltiplos número de termos AR e MA e comparamos os seus correspondentes valores AIC/BIC com find para determinar qual o modelo que tem menos AIC/BIC e, portanto, previsões da forma mais precisa. Por outras palavras, uma sensibilidade

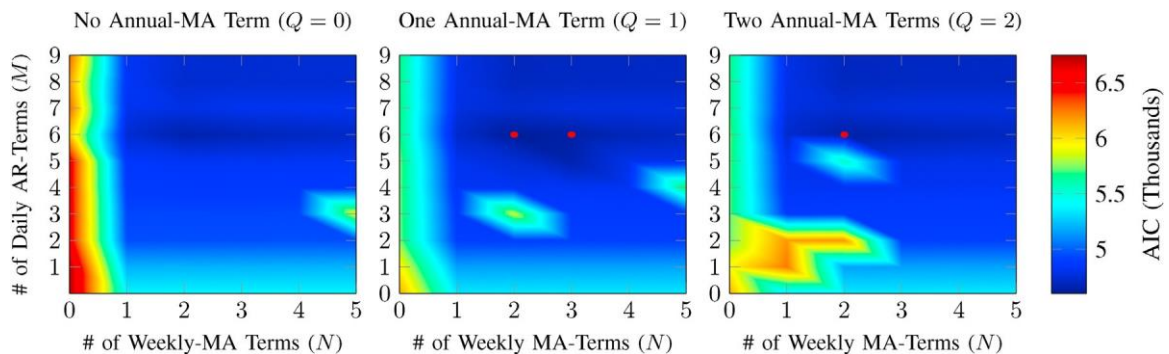


Fig. 7. Valores AIC para diferentes configurações do modelo AR/MA proposto. Os marcadores vermelhos mostram três valores AIC mais pequenos. (Para interpretação das

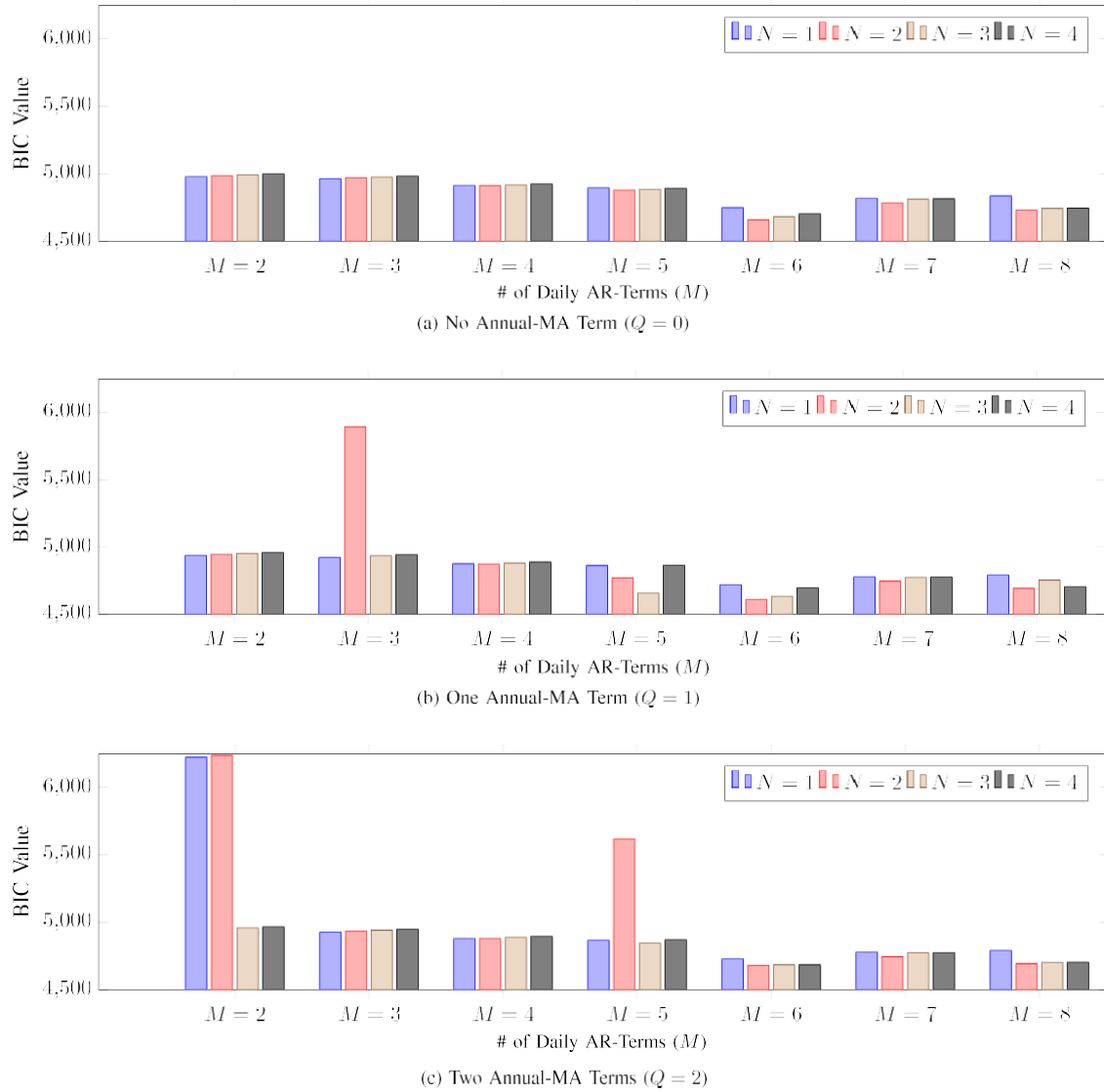


Fig. 10. Gráfico de barras dos valores BIC correspondentes a diferentes definições do modelo AR/MA proposto.

regiões nestas figuras, marcadas por pontos vermelhos, especificam as definições dos parâmetros do forecaster que conduzem aos valores mais baixos AIC/BIC. A figura 9 mostra as alterações do AIC em três valores diferentes de Q . Cada uma das subfiguras representa também uma análise de sensibilidade sobre M e N para uma specific Q . Da mesma forma, na figura 10, o padrão de alteração dos valores BIC é mostrado em três valores diferentes de Q . Cada uma das subfiguras representa também uma análise de sensibilidade sobre M e N para uma specific Q .

A figura 11 compara os valores de três tipos de erro diferentes para múltiplas configurações do modelo proposto. Fig. 11a-c especifica os valores de Erro de Raiz-Quadrado (RMSE), Erro Médio-Absoluto-Percentual-Erro (MAPE), e Erro Médio-Absoluto-Erro (MAE)⁴ respectivamente para diferentes valores de M , N e Q . Os valores mínimos de erro do nosso modelo proposto acontecem nos pontos mínimos das parcelas representadas nas Figs. 7, 8 e 11. Quadro 2

resume as melhores escolhas para o triplo (M, N, Q) que minimizam os valores de AIC/BIC e três tipos de erro diferentes e AIC/BIC.

O quadro 2 sugere que o modelo com definição $(M, N, Q) = (2, 6, 1)$ tem o menor valor de AIC e BIC e o penúltimo valor de erro

entre todos os cenários considerados. Além disso, como mencionado anteriormente, AIC e BIC não só recompensam a bondade do modelo de fit, mas também incluem a função de penalização que está a aumentar as funções do número de parâmetros estimados e desencorajam overfitting. Consequentemente, são melhores critérios para avaliar os modelos de previsão do que os valores de erro que se limitam a quantificar como fit o modelo é para o conjunto de formação. Como resultado, consideramos $(M, N, Q) = (6, 2, 1)$ como a melhor configuração do nosso modelo proposto para a previsão da carga eléctrica diária. Após criar o modelo baseado em AR/MA para t e calcular os valores de previsão de x_{t+1} , os valores de carga previstos x_{t+1} são obtidos aplicando as transformações inversas dos passos um e dois em x_{t+1} , ou seja, em relação à Eq. (3), a série temporal x_{t+1} é calculada pela seguinte equação:

$$x_{t+1} = r_d^{-1}(x_t + [(L^1 + L^7 - L X^* t + L^{364} - L^{365} - L^{371} + L^{372}) o_{rd}] X_t), \quad (6)$$

onde r_d é a transformação Logarithmic Box-Cox introduzida anteriormente e

$$x_t = (d)_t + \sum_{i=1}^6 (d)_t + \sum_{j=1}^2 y(w)_j^{(d)} + y_t^{(y)} + y_t^{(d)} + \varepsilon_t^{(d)}. \quad (7)$$

⁴ Assuming o_i and f_i as the observed and forecast values ($i = 1, 2, \dots, n$ and $o_i, f_i > 0$), o RMSE, MAPE, e MAE são respectivamente defined com o $\sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2}$, $\frac{1}{n} \sum_{i=1}^n \frac{|f_i - o_i|}{o_i}$, e $\frac{1}{n} \sum_{i=1}^n |f_i - o_i|$.

Quadro 2

Escolha da melhor configuração do modelo proposto para a previsão dos valores de carga diária da rede PJM em 2015.

Daily-AR (M), # de	Crítérios de selecção#dos termos termos semanais-MA (N) e # de termos anuais-MA (Q)	AIC (milhares)	BIC (milhares)	Erro RMS (%)	Erro MAP (%)	Erro de MA (%)
MinimizingM AICM	$= 2, N = 6, Q = 1$	4.61	4.76	6.60	4.49	4.68
	$= 3, N = 6, Q = 1$	4.63	4.79	6.80	4.59	4.73
	$M = 2, N = 6, Q = 2$	4.64	4.80	7.02	4.75	5.42
MinimizingM BICM	$= 2, N = 6, Q = 1$	4.61	4.76	6.60	4.49	4.68
	$= 3, N = 6, Q = 1$	4.63	4.79	6.80	4.59	4.73
	$M = 3, N = 5, Q = 1$	4.65	4.81	7.09	4.82	4.98
MinimizingM	$= 5, N = 9, e Q = 1$	4.71	4.85	6.60	4.46	4.67
RMS/MAP/MAM	$= 2, N = 6, e Q = 1$	4.61	4.76	6.60	4.49	4.68
errosM	$= 4, N = 8, e Q = 2$	4.71	4.87	6.62	4.49	4.70

3.5.3. Comparar a série temporal residual com o ruído branco

Como já foi mencionado, uma forma de avaliar a bondade de um forecaster é comparar a sua série temporal residual $p^{(d)} X^t x_t$ (valores de erro ao longo do tempo) com o ruído branco. O modelo de previsão ideal extrai qualquer informação significativa do conjunto de formação de entrada para

obter a previsão mais precisa. Daí em diante, nenhum sinal significativo permanecerá na série temporal residual, ou seja, a série temporal residual será um *ruído branco*. Aqui, utilizamos o periodograma cumulativo de $p^{(d)}$ para mostrar quão perto a série temporal residual está de um *ruído branco*.

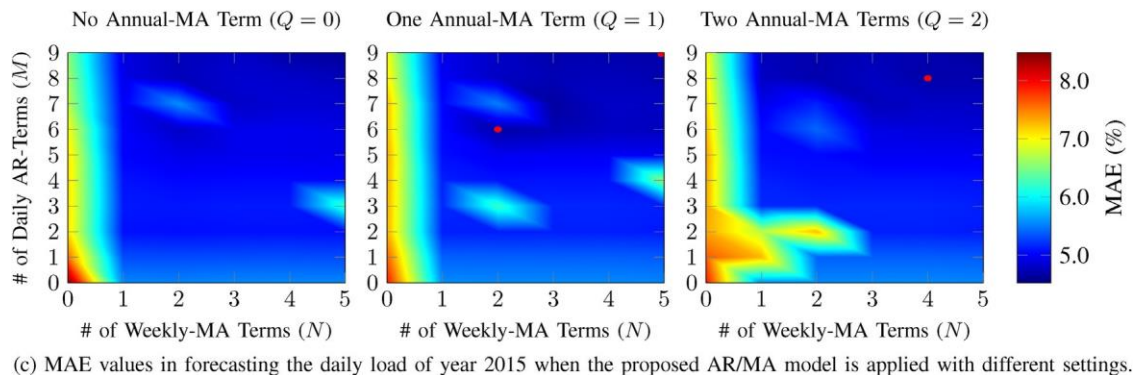
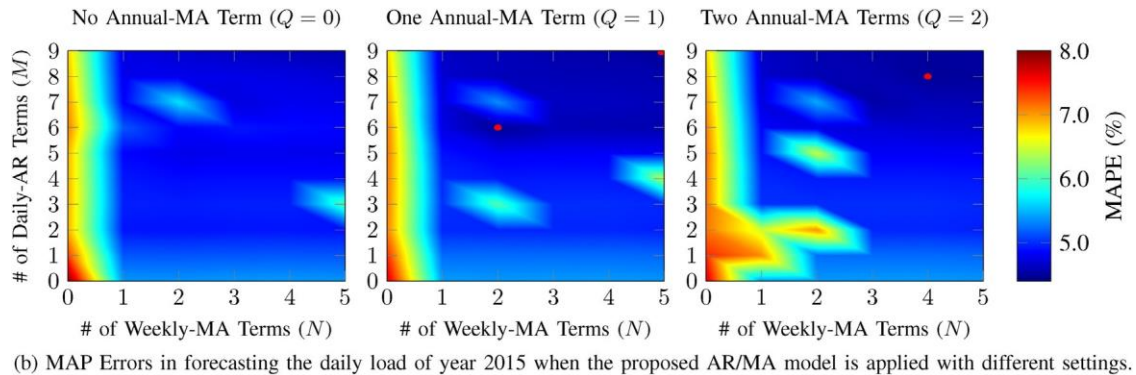
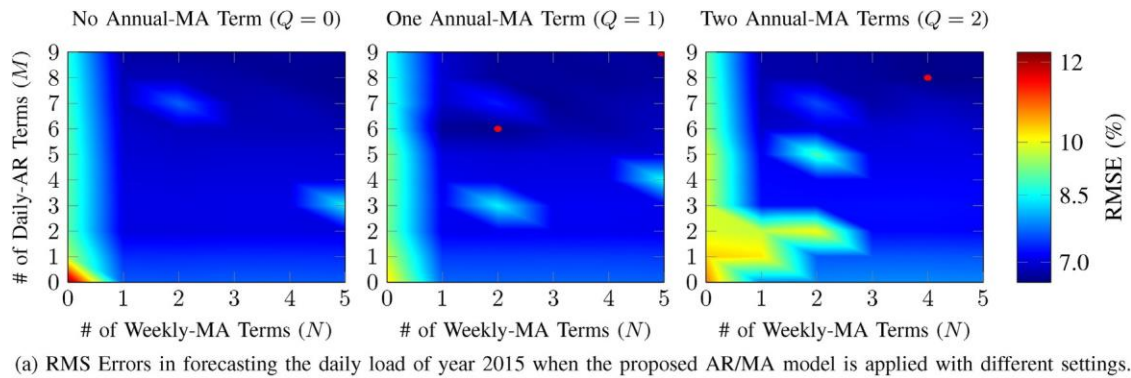


Fig. 11. Comparando os valores de três tipos diferentes de erros para diferentes configurações do modelo AR/MA proposto. Os pontos com três menores valores de erro foram specified com marcadores vermelhos. (Para interpretação das referências a cores nesta legenda figure, o leitor é remetido para a versão web deste artigo).

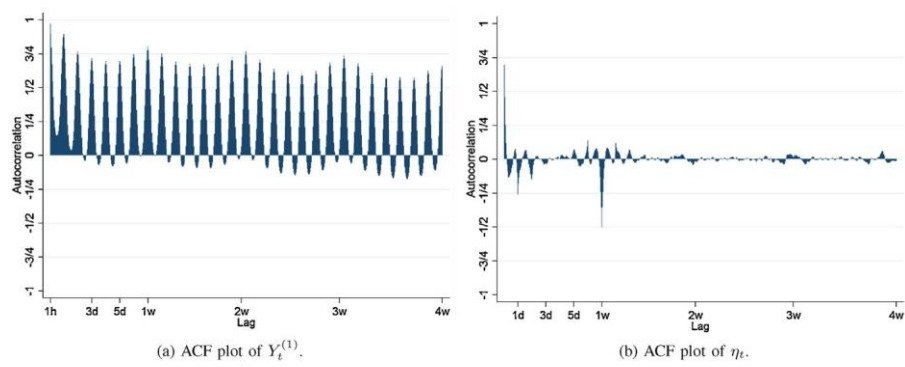


Fig. 13. Lotes ACF de $Y_t^{(1)}$ e a sua série temporal transformada $17t$.

Quadro 3

Comportamento das funções de autocorrelação e de autocorrelação parcial dos dados de carga horária em níveis não sazonais e sazonais múltiplos.

Não sazonal		(horária)		Sazonalidade diária		Sazonalidade semanal
Lag	ACFPACFLagACFPFLagACFLagACFPACF					
Onda sinérgica em declínio	Caindo em <i>Lag</i> = 1	Caindo em <i>Lag</i> = 48	Exponencialmente em declínio	Caindo em <i>Lag</i> = 168	Exponencialmente em declínio	
2+0.		1+0.697+0.69724-0.		281-0.	229168-0.	490-0.156
3+0.		428-0.11148-0.		135-0.	130336-0.	009-0.089
4+0.		218-0.07172-0.			033-0.096504+0.	013-0.061
5-0.		084-0.02096+0.				008-0.064672--
6-0.		001-0.030120+0.				050-0.019840--
					060-0.045144+0.149+0.0901008--	

Os valores das duas últimas colunas do **Quadro 3** determinam que o ACF e PACF de I_t em desfasamentos semanais ($L = 168, 336, \dots$) mostram um comportamento semelhante ao dos desfasamentos diários. De facto, os valores de ACF cairão subitamente após o desfasamento semanal first; enquanto o PACF declina exponencialmente em relação aos desfasamentos semanais. Esta combinação de comportamentos leva-nos à conclusão de que a série temporal I_t deve incluir a média móvel da ordem first no ciclo sazonal semanal, ou seja, I_t tem uma parte de média móvel não-determinista sob a forma de $O_{1t}^{(w)} \varepsilon_t^{(h)}$ onde $O^{(w)}$ representa o coefficient da mudança de first... prazo médio em desfasamentos semanais.

O **quadro 3** sugere o seguinte modelo baseado em AR/MA para a série temporal I_t de A&S_t

$$I_t = \underbrace{c p^{(h)} I_{t-1}}_{\text{não sazonal-sazonal-sazonal}} + \underbrace{O_1^{(d)} \varepsilon_{t-24}}_{\text{sazonal-sazonal}} + \underbrace{O_2^{(d)} \varepsilon_{t-48}}_{\text{sazonal-sazonal}} + \underbrace{O_{1t-168}^{(w)} \varepsilon_t^{(h)}}_{\text{semanal-sazonal}} + \varepsilon_t^{(h)}. \quad (8)$$

4.4. Afinação e avaliação

Como mencionado anteriormente, na etapa final da construção do modelo, os parâmetros detalhados dos componentes AR/MA são determinados experimentando múltiplos valores e finding o modelo de previsão com valores mínimos AIC/BIC. No caso da previsão horária, nós figured out que as seguintes configurações do modelo darão o valor mínimo AIC/BIC ao prever a carga eléctrica horária da rede PJM em 2015:

$$I_t = \sum_{i=1}^{12} c p^{(h)} I_{t-i} + \sum_{j=1}^4 O_j^{(d)} \varepsilon_{t-24j} + O_{1t-168}^{(w)} \varepsilon_t^{(h)} + \varepsilon_t^{(h)} \quad (9)$$

O **Quadro 4** resume o desempenho do nosso modelo comparando os outros analistas da literatura actual.

Quadro 4

Comparação da exactidão das previsões apresentadas em Pappas et al. [47], Shaker et al. [54], e Dudek [44].

Trabalho	Ano	Localização	Método	WN Testi	MAPE (%)	RMSE (%)	MAE (%)	Tipo	AIC (1000's)	BIC (1000's)
Pappas et al.	2010	Grécia	MMPFa	Testado	1.87	Não testado	Não testado	Curto prazo	Não testado	Não testado
Pappas et al.	2010	Grécia	AICCb	Não testado	1.98	Não testado	Não testado	Curto prazo	Não testado	Não testado
Shaker et al.	2014	Alberta Canadá	WNNc	Não testado	Não testado	1.328	0.983	Curto prazo	Não testado	Não testado
Shaker et al.	2014	Alberta Canadá	MLPNNd	Não testado	Não testado	3.724	2.556	Curto prazo	Não testado	Não testado
Shaker et al.	2014	Alberta Canadá	RBFNNe	Não testado	Não testado	2.287	1.712	Curto prazo	Não testado	Não testado
Dudek	2016	Rede polaca	PCRF	Não testado	1.15	Não testado	Não testado	Curto prazo	Não testado	Não testado
Dudek	2016	Rede polaca	PLSRg	Não testado	1.09	Não testado	Não testado	Curto prazo	Não testado	Não testado
Este trabalho	2016	Rede PJM	MTS ARMAh	Testado	0.86	1.24	0.92	Curto prazo	Testado	Testado

5. Resumo e perspectivas

Este trabalho propõe uma nova técnica passo-a-passo para a criação de um modelo de previsão baseado em AR/MA. Este modelo é utilizado para a previsão da procura de energia eléctrica desde o horizonte de curto a médio prazo. No passo first, a série temporal correspondente aos dados de carga histórica é avaliada para garantir que tem variância homo-genética ao longo do tempo. No caso de a série temporal ser

não variância homogénea, aplicamos um logaritmo apropriado

Transformação Box-Cox, a fim de a converter numa série temporal com homogeneidade de variância. No passo seguinte, os valores ACF da série temporal transformada são examinados para ver se existe alguma indicação de não ser estacionário a nível não sazonal e ciclos sazonais (diários, semanais, e anuais). Se a ACF valores ou caem ou diminuem no first poucos desfasamentos, deve ser considerado estacionário. Por outro lado, se os valores ACF ou caem após um número considerável de desfasamentos ou diminuem muito lentamente, deve ser considerado como não estacionário. Para eliminar

as indicações não estacionárias, aplicamos de forma iterativa transformações diferenciadoras na série temporal. No passo seguinte, a série temporal estacionária, resultante do segundo passo, é modelada para um modelo de média móvel e/ou autoregressivo baseado no comportamento das suas correspondentes ACF e PACF em diferentes resoluções de tempo (horário, diário, semanal, e anual). Para uma determinada resolução de tempo, se os valores ACF da série temporal forem cortados no first poucos desfasamentos, enquanto os valores PACF são reduzidos exponencialmente ou de forma sinusoidal, a série temporal nesse specific resolução de tempo é melhor fit para um modelo de média móvel. Por outro lado, se os valores do PACF de a série temporal cai substancialmente no first poucos desfasamentos, enquanto a

Os valores ACF são reduzidos exponencialmente ou de forma sinérgica, o melhor modelo para essa resolução temporal specific é autoregressivo.

A fim de afinar o modelo, utilizamos o AIC/BIC que negocia com o trade-off entre fitness apropriado do modelo e a complexidade do modelo. Para este efeito, consideramos um valor de penalização para

78^{ma} partição Multi-modelo filter.

^b Critério de informação Aikake corrigido.

^c Rede neural Wavelet.

^d Rede neural perceptron multicamadas.

^e Rede neural com função de base radial.

^f Regressão da componente principal.

^g Regressão parcial mínima quadrática.

^h ARMA multitemporal.

Teste ⁱ WN: teste de ruído branco.

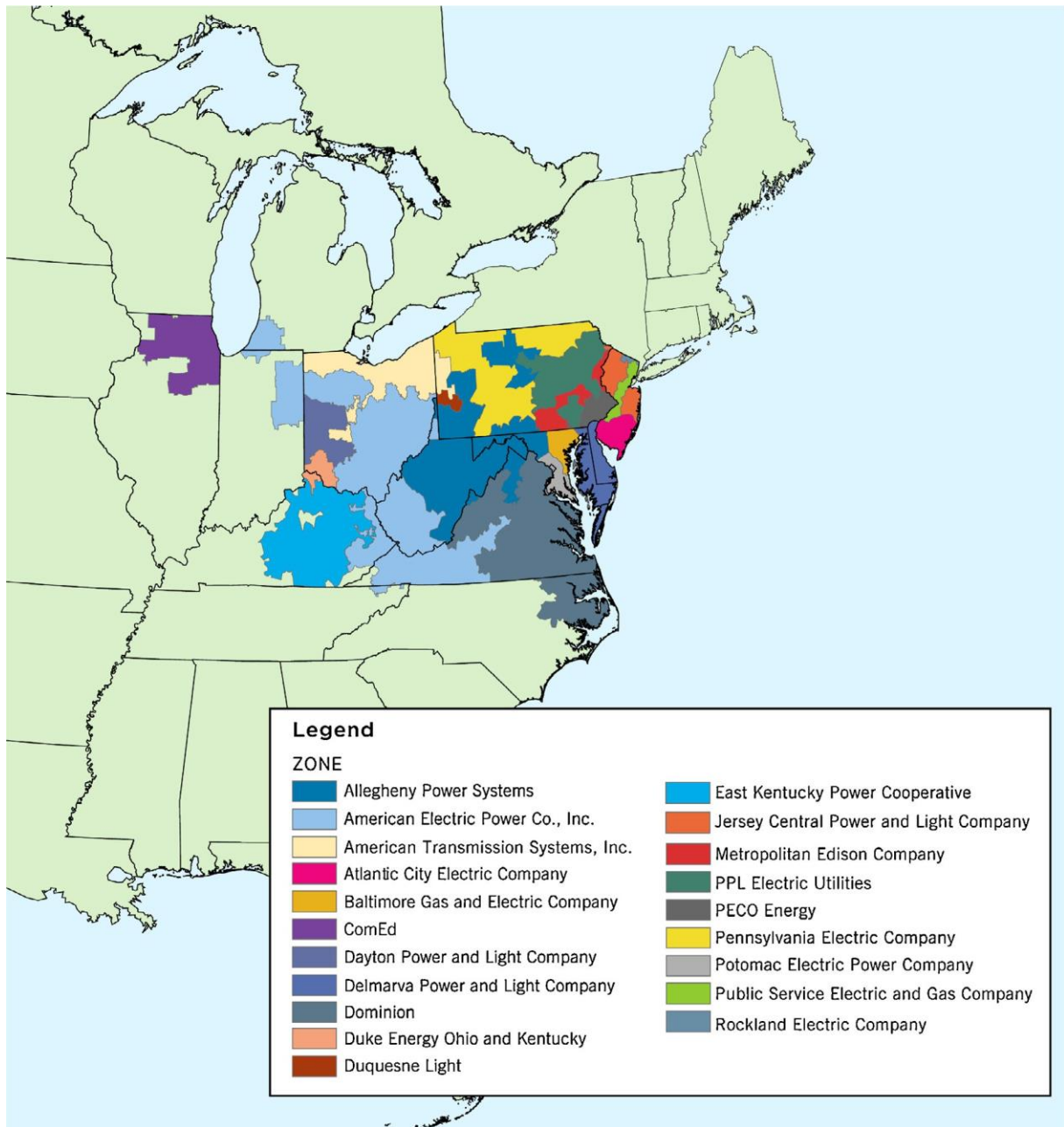


Fig. 14. Regiões geográficas abrangidas pela interligação PJM [56].

modelos complexos que só podem funcionar precisamente para o conjunto de formação analisado; no entanto, as medições de erro convencionais dependem consideravelmente do conjunto de formação dado e medem o fitness do modelo apenas no que diz respeito ao conjunto de formação. Se a previsão flexibility é a nossa preocupação mais importante, ou seja, o nosso principal objectivo é desenvolver uma previsão que possa ser amplamente utilizada para diferentes conjuntos de formação de dados sem deteriorar o desempenho da previsão, o BIC é uma alternativa melhor (em comparação com o AIC) para quantificar a utilidade da previsão, uma vez que o BIC penaliza mais a complexidade dos modelos de previsão do que o AIC. Contudo, o AIC é um critério melhor se optarmos por comprometer o modelo flexibility em algum grau, a fim de ganhar mais fitness e melhorar a precisão da previsão através da redução do erro de previsão.

A fim de validar a eficácia do forecaster proposto, implementamo-lo para prever a procura de energia eléctrica de um exemplo do mundo real a partir da interligação PJM. A precisão da

O forecaster proposto é avaliado pelo teste baseado no periodograma de Bartlett e quantificado por múltiplos tipos de erros convencionais. O nosso desempenho do forecaster proposto, comparado com os estudos prévios aos nossos, é validado utilizando três métricas de erro (MAPE, RMSE, e MAE). Por exemplo, o valor MAPE do nosso forecaster para PJM dados de carga horária medida é de 0,86%, o que mostra uma redução de 21% em comparação com o forecaster de Dudek que é desenvolvido em 2016 [44].

Apêndice A. Informação detalhada sobre os dados de carga horária do PJM

A interligação PJM é uma organização de transmissão regional (RTO) que gere a electricidade nas seguintes regiões dos Estados Unidos: Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, Carolina do Norte, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia e o Distrito de Columbia [55]. Em

Quadro 5

A carga diária no first seis meses de 2015 em PJM. Os números são mostrados em kWh.

Dia/mês	Janeiro	Fevereiro	Março	Abril	Maiο	Junho
1	2,224,247.0	2,287,549.2	2,442,177.8	2,024,486.3	1,847,056.1	2,101,982.5
2	2,214,133.4	2,467,346.6	2,437,075.9	1,974,545.2	1,681,862.1	1,943,814.3
3	2,146,834.0	2,619,797.2	2,529,970.7	1,807,215.0	1,657,509.5	1,922,430.1
4	1,974,231.4	2,445,748.3	2,351,650.9	1,730,857.6	1,944,146.7	2,008,393.8
5	2,445,435.6	2,571,644.5	2,564,235.5	1,692,526.7	2,016,051.7	2,033,350.8
6	2,652,063.8	2,681,963.0	2,705,979.7	1,898,950.8	2,030,419.7	1,871,852.8
7	2,843,836.9	2,243,782.4	2,369,589.7	1,922,330.2	2,067,433.1	1,868,774.4
8	3,023,713.9	2,025,259.8	1,954,528.4	1,970,665.4	2,131,769.8	2,238,112.9
9	2,736,306.6	2,347,400.7	2,139,678.0	2,028,948.3	1,927,307.7	2,270,840.2
10	2,679,408.0	2,504,609.6	2,165,532.0	1,943,293.5	1,931,900.8	2,382,590.1
11	2,463,715.8	2,459,178.1	2,055,011.6	1,718,914.8	2,239,332.3	2,542,460.0
12	2,461,286.0	2,547,267.4	2,090,851.4	1,688,730.1	2,169,028.7	2,578,623.9
13	2,561,561.2	2,755,521.6	2,127,633.3	1,886,606.5	1,922,147.6	2,352,410.3
14	2,690,702.6	2,543,273.5	1,932,114.2	1,879,084.5	1,890,764.5	2,353,527.6
15	2,547,062.4	2,809,745.1	1,884,176.9	1,882,273.2	1,928,452.4	2,621,394.3
16	2,453,213.0	3,010,474.6	2,060,271.8	1,884,624.7	1,904,483.3	2,590,952.4
17	2,304,689.7	2,802,988.6	2,017,453.2	1,876,043.3	2,001,731.2	2,332,716.6
18	2,159,352.6	2,803,212.0	2,184,042.3	1,705,664.2	2,281,909.0	2,393,511.2
19	2,297,892.3	3,052,242.2	2,190,832.1	1,629,643.8	2,160,130.8	2,338,444.6
20	2,345,274.7	3,104,390.6	2,226,302.0	1,901,434.6	1,952,688.8	2,161,401.8
21	2,412,450.4	2,720,441.3	1,958,744.1	1,879,815.0	1,881,833.7	2,312,095.6
22	2,416,012.7	2,306,819.3	1,896,000.7	1,926,903.3	1,847,969.6	2,638,997.9
23	2,399,030.1	2,670,971.2	2,203,770.5	1,974,713.6	1,665,668.7	2,694,198.8
24	2,247,946.0	2,868,672.0	2,220,515.3	1,947,838.4	1,701,091.1	2,399,544.4
25	2,152,295.5	2,609,601.0	2,179,801.4	1,804,096.3	1,910,615.1	2,307,415.4
26	2,492,879.1	2,609,484.5	2,066,330.6	1,707,445.2	2,285,864.7	2,234,547.0
27	2,571,906.4	2,640,380.4	2,137,934.4	1,919,708.2	2,350,697.4	1,929,125.4
28	2,607,071.1	2,539,066.3	2,161,144.5	1,902,106.7	2,359,577.4	1,851,272.4
29	2,562,531.5	–	2,101,593.8	1,882,956.0	2,336,808.4	2,087,940.5
30	2,481,083.7	–	2,093,609.3	1,867,538.4	2,178,553.2	2,237,095.2
31	2,484,570.9	–	2,048,286.0	–	2,056,848.5	–

algumas das regiões mencionadas, a PJM está a coordenar a electricidade por grosso para toda a região. No entanto, para outras áreas, serve apenas uma parte da região. A sua principal tarefa é fornecer electricidade de confiança a mais de 61 milhões de clientes. Fig. 14

representa as regiões abrangidas pela interligação PJM. As tabelas 5 e 6 representam a carga histórica diária para o first e o segundo semestre de 2015 nas regiões abrangidas pela interligação PJM, respectivamente.

Quadro 6

A carga diária nos últimos seis meses de 2015 em PJM. Os números são mostrados em kWh.

Dia/mês	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
1	2,235,981.1	2,326,506.7	2,675,402.2	1,871,454.0	1,625,687.1	2,089,866.7
2	2,098,594.8	2,326,394.3	2,686,777.1	1,848,997.8	1,900,526.7	2,076,625.0
3	1,947,908.6	2,560,275.1	2,719,161.1	1,756,196.2	1,932,687.6	2,161,429.0
4	1,897,211.1	2,574,999.1	2,599,482.7	1,674,459.5	1,925,894.8	2,175,708.5
5	2,000,731.8	2,498,275.3	2,280,105.6	1,869,228.5	1,949,722.0	2,055,484.6
6	2,428,637.2	2,325,697.5	2,215,356.6	1,897,800.1	1,931,240.4	2,013,747.7
7	2,523,706.6	2,284,944.6	2,375,704.9	1,914,453.2	1,739,918.6	2,173,588.0
8	2,399,394.5	2,172,007.5	2,679,940.3	1,930,268.7	1,746,607.9	2,164,163.3
9	2,428,976.7	2,169,595.7	2,630,928.8	1,918,961.5	2,026,658.3	2,168,256.1
10	2,366,755.8	2,391,373.9	2,366,200.6	1,667,591.6	1,969,193.8	2,065,124.2
11	2,169,871.9	2,437,264.5	2,193,611.3	1,641,818.7	1,945,546.9	1,995,847.2
12	2,157,110.2	2,334,476.4	1,888,958.6	1,874,666.6	1,970,728.4	1,797,150.0
13	2,442,000.9	2,317,973.4	1,750,362.8	1,892,341.3	1,951,856.9	1,746,418.1
14	2,495,718.8	2,400,619.0	1,954,813.6	1,866,502.2	1,858,388.4	1,963,965.5
15	2,338,597.1	2,336,837.2	2,081,851.1	1,873,318.1	1,800,851.9	2,004,113.1
16	2,221,536.9	2,409,186.7	2,178,468.4	1,844,691.9	1,964,457.0	2,048,928.6
17	2,404,239.5	2,698,831.6	2,240,706.4	1,750,889.6	1,961,147.6	2,087,605.0
18	2,478,797.6	2,636,636.9	2,230,933.6	1,791,688.1	1,941,140.4	2,187,556.3
19	2,589,602.2	2,634,215.7	2,039,646.9	2,015,386.1	1,928,128.6	2,176,180.8
20	2,824,010.1	2,497,295.3	1,847,147.5	1,952,358.9	1,975,289.0	2,107,844.3
21	2,675,592.0	2,347,128.5	1,964,309.2	1,917,727.2	1,927,928.7	2,155,731.4
22	2,412,955.5	2,094,643.9	2,006,255.5	1,903,754.3	1,952,532.7	2,011,890.0
23	2,379,320.3	2,079,893.8	2,036,401.2	1,849,517.5	2,237,398.5	1,958,759.2
24	2,408,843.6	2,363,988.9	2,043,572.7	1,718,982.9	2,181,561.0	1,767,741.6
25	2,352,382.1	2,274,723.0	1,980,467.2	1,657,375.6	2,097,915.4	1,669,885.1
26	2,361,044.4	2,128,817.6	1,767,919.0	1,895,559.4	1,745,687.3	1,726,500.0
27	2,638,585.8	2,121,712.1	1,752,623.6	1,963,221.4	1,745,832.1	1,743,846.6
28	2,770,611.5	2,141,847.0	2,119,168.7	1,953,809.3	1,758,048.5	2,052,214.7
29	2,813,593.6	2,086,999.4	2,172,743.8	1,913,205.3	1,820,186.8	2,028,147.0
30	2,725,597.2	2,190,522.9	2,073,200.0	1,902,573.2	2,115,151.5	2,011,894.9
31	2,562,081.4	2,561,310.6	–	1,781,476.8	–	1,968,810.2

Referências

- [1] E. Gonzalez-Romera, M.A. Jaramillo-Moran, D. Carmona-Fernandez, Monthly electric energy demand forecasting based on trend extraction, *IEEE Trans. Power Syst.* 21 (4) (2006) 1946–1953.
- [2] J.W. Taylor, P.E. McSharry, Short-term load forecasting methods: an evaluation based on European data, *IEEE Trans. Power Syst.* 22 (4) (2007) 2213–2219.
- [3] M. Shahidehpour, H. Yamin, Z. Li, Operações de Mercado em Sistemas de Energia Elétrica: Forecasting, Scheduling, and Risk Management, Wiley Online Library, 2002.
- [4] K. Moslehi, R. Kumar, A reliability perspective of the smart grid, *IEEE Trans. Smart Grid* 1 (1) (2010) 57–64.
- [5] M.H. Amini, O. Karabasoglu, M.D. Ilic, K.G. Boroojeni, S.S. Iyengar, ARIMA-based método de previsão da procura considerando o modelo probabilístico de vehi eléctrico - estacionamento dos cles, in: *IEEE PES Assembleia Geral 2015*, Denver, CO, EUA, 26–30 de Julho de 2015.
- [6] F. Rahimi, A. Ipakchi, Demandar resposta como um recurso de mercado sob o paradigma smart grid, *IEEE Trans. Smart Grid* 1 (1) (2010) 82–88.
- [7] S. Bahrami, V.W.S. Wong, An autonomous demand response program in smart grid with foresighted users, in: *Proc. of IEEE Smart Grid Communications (SmartGridComm)*, Miami, FL, 2015, pp. 205–210.
- [8] F. Kamyab, M.H. Amini, S. Sheykha, M. Hasanpour, M.M. Jalali, Programa de resposta à procura em smart grid usando mecanismo de licitação da função de oferta, *IEEE Trans. Smart Grid* 7 (2) (2016) 1277–1284.
- [9] N. Mahmoudi, New Demand Response Framework and Its Applications for Electricity Markets (PhD Dissertation), The University of Queensland, 2015.
- [10] X. Zhang, G. Hug, J. Zico Kolter, I. Harjunkoski, Model predictive control of industrial loads and energy storage for demand response, in: *Reunião Geral do IEEE PES*, 2016.
- [11] S. Bahrami, A. Sheikhi, From demand response in smart grid towards integrated demand response in smart energy hub, *IEEE Trans. Smart Grid* 7 (2) (2016) 650–658.
- [12] L. Hernandez, et al., A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings, *IEEE Commun. Surv. Tutor.* (2014) 1–36.
- [13] W. Charytoniuk, M.-S. Chen, Previsão de carga a muito curto prazo utilizando redes neurais artificial, *IEEE Trans. Power Syst.* 15 (1) (2000) 263–268.
- [14] C. Guan, P.B. Luh, L.D. Michel, Y. Wang, P.B. Friedland, Very short-term load previsão: redes neurais wavelet com dados pré-filtering, *IEEE Trans. Power Syst.* 28 (1) (2013) 30–41.
- [15] A. Piras, A. Germond, B. Buchenel, K. Imhof, Y. Jaccard, Heterogeneous artificial rede neural para previsão de carga eléctrica a curto prazo, *IEEE Trans. Power Syst.* 11 (1) (1996) 397–402.
- [16] N. Amjadi, Previsão de curto prazo da carga de autocarros de sistemas de energia por um novo método híbrido, *IEEE Trans. Power Syst.* 22 (1) (2007) 333–341.
- [17] L. Zjavka, V. Snásel, Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks, *Electr. Power Syst. Res.* 137 (2016) 113–123.
- [18] E. Doveh, P. Feigin, D. Greig, L. Hyams, Experience with FNN models for medium term power demand predictions, *IEEE Trans. Power Syst.* 14 (2) (1999) 538–546.
- [19] M.S. Kandil, S.M. El-Debeiky, N.E. Hasanien, Long-term load forecasting for fast desenvolvendo utilitário usando um sistema especializado baseado no conhecimento, *IEEE Trans. Power Syst.* 17 (2) (2010) 491–496.
- [20] R.J. Hyndman, S. Fan, Density forecasting for long-term peak electricity demand, *IEEE Trans. Power Syst.* 25 (2) (2010) 1142–1153.
- [21] A.S. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos, B. Venkatesh, Melhoria da previsão de carga a curto prazo utilizando redes neuronais ensacadas, *Electr. Power Syst. Res.* 125 (2015) 109–115.
- [22] J. Fiot, F. Dinuzzo, Electricity demand forecasting by multi-task learning, *IEEE Trans. Smart Grid* 99 (2016).
- [23] M.H. Amini, A. Islam, Alocação de estacionamentos de veículos eléctricos na rede de distribuição, in: *IEEE Innovative Smart Grid Technologies Conference (ISGT)*, 2014.
- [24] I. Rahman, et al., Review of recent trends in optimization techniques for plug-in hybrid, and electric vehicle charging infrastructures, *Renew. Sustentar. Energy Rev.* 58 (2016) 1039–1047.
- [25] S. Bahrami, M. Parniani, Game theory based charging strategy for plug-in hybrid electric vehicles, *IEEE Trans. Smart Grid* 5 (5) (2014) 2368–2375.
- [26] S. Bahrami, V.W.S. Wong, A potential game framework for charging PHEVs in smart grid, in: *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, Victoria, Canada, 2015, pp. 28–33.
- [27] H. Arasteh, et al., SoS baseado no plano de expansão do sistema de distribuição multiobjectivo - ning, *Electr. Power Syst. Res.* 141 (2016) 392–406.
- [28] K.B. Song, Y.S. Baek, D.H. Hong, G. Jang, Short-term load forecasting for the holidays using fuzzy linear regression method, *IEEE Trans. Power Syst.* 20 (1) (2005) 96–101.
- [29] P.E. McSharry, S. Bouwman, G. Bloemhof, Probabilistic forecast of the magnitude and timing of peak electricity demand, *IEEE Trans. Power Syst.* 20 (2) (2005) 1166–1172.
- [30] B.J. Chen, M.W. Chang, C.-J. Lin, Load forecasting using support vector machines: a study on EUNITE competition 2001, *IEEE Trans. Power Syst.* 19 (4) (2004) 1821–1830.
- [31] A. Pardo, V. Meneu, E. Valor, Temperatura e sazonalidade influences on Spanish electricity load, *IEEE Energy Econ.* 24 (2002) 55–1830.
- [32] A.P. Douglas, A.M. Breipohl, F.N. Lee, R. Adapa, Load forecasting using support vector machines: a study on EUNITE competition 2001, *IEEE Trans. Power Syst.* 13 (4) (1998) 1507–1513.
- [33] M.S. Kandil, S.M. El-Debeiky, N.E. Hasanien, Overview e comparação de técnicas de previsão a longo prazo para uma utilidade em rápido desenvolvimento: Parte I, *Electr. Power Syst. Res.* 58 (1) (2001) 11–17.
- [34] H. Morita, T. Kase, Y. Tamura, S. Iwamoto, Intervalo de previsão da procura anual da maxi-múmia usando modelo dinâmico cinzento, *Int. J. Electr. Power Energy Syst.* 18 (7) (1996) 409–413.
- [35] D.C. Park, et al., Electric load forecasting using an artificial neural network, *IEEE Trans. Power Syst.* 6 (2) (1991) 442–449.
- [36] K.Y. Lee, Y.T. Cha, J.H. Park, Short-term load forecasting using an artificial neural network, *IEEE Trans. Power Syst.* 16 (1) (2001) 124–132.
- [37] H.S. Hippert, C.E. Pedreira, R.C. Souza, Neural networks for short-term load forecasting: a review and evaluation, *IEEE Trans. Power Syst.* 16 (1) (2001) 44–55.
- [38] A.I. Sarwat, M.H. Amini, A. Domijan Jr., A. Damjanovic, F. Kaleem, Weather-based interruption prediction in the smart grid using chronological data, *J. Mod. Power Syst. Clean Energy* 4 (2) (2016) 308–315.
- [39] N. Amjadi, Previsão de carga horária de curto prazo utilizando modelos de série temporal com capacidade de estimativa de pico de carga, *IEEE Trans. Power Syst.* 16 (3) (2001) 498–505.
- [40] W. Charytoniuk, M.-S. Chen, Previsão de carga a muito curto prazo utilizando redes neurais artificial, *IEEE Trans. Power Syst.* 15 (1) (2002) 263–268.
- [41] Y. Wang, Q. Xia, C. Kang, Secondary forecasting based on deviation analysis for short-term load forecasting, *IEEE Trans. Power Syst.* 26 (2) (2011) 500–507.
- [42] C. Garca-Ascanio, C. Mate, Electric power demand forecasting using interval time series: a comparison between VAR and iMLP, *Energy Policy* 38 (2) (2010) 715–725.
- [43] D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E. Atlas, M.J. Damborg, Electric load forecasting using an artificial neural network, *IEEE Trans. Power Syst.* 6 (2) (1991) 442–449.
- [44] G. Dudek, Modelos de regressão linear local baseados em padrões para previsão de carga a curto prazo, *Electr. Power Syst. Res.* 130 (2016) 139–147.
- [45] T.H.D. Ngo, The Box-Jenkins methodology for time series models, *SAS Globo. Fórum* 54 (2013).
- [46] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung, *Time-Series Analysis: Forecasting and Control*, Holden-Day, CA, 1976.
- [47] S.S. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafidas, Previsão da procura de electricidade do sistema eléctrico helénico usando um modelo ARMA, *Electr. Power Syst. Res.* 80 (3) (2010) 256–264.
- [48] A.A. El Desouky, M.M. El Kateb, Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA, *IEE Proc. Gener. Transm. Distrib.* 147 (4) (2000) 213–217.
- [49] S.J. Huang, K.R. Shih, Short-term load forecasting via ARMA model identification including considerações sobre processos não-Gaussianos, *IEEE Trans. Power Syst.* 18 (2) (2003) 673–679.
- [50] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, ARIMA modelos para pre-ditar os preços da electricidade no dia seguinte, *IEEE Trans. Power Syst.* 18 (3) (2003) 1014–1020.
- [51] J.W. Taylor, Triple seasonal methods for short-term load forecasting, *Eur. J. Oper. Res.* 204 (1) (2010) 139–152.
- [52] G.E.P. Box, D.R. Cox, An analysis of transformations, *J. R. Stat. Soc. Ser. B* 26 (1964) 211–252.
- [53] <http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx>.
- [54] H. Shaker, H. Chitsaz, H. Zareipour, D. Wood, Em comparação de duas estratégias na previsão da procura líquida usando a rede neural wavelet, in: *North American Power Symposium (NAPS)*, 2014, pp. 1–6.
- [55] Disponível Online: <http://pjm.com/about-pjm/who-we-are.aspx>.
- [56] Disponível online: <http://pjm.com//media/about-pjm/pjm-zones.ashx>.