



**MAÍRA RODRIGUES VILLAMAGNA**

**SELEÇÃO DE MODELOS DE SÉRIES  
TEMPORAIS E REDES NEURAIS ARTIFICIAIS  
NA PREVISÃO DE CONSUMO E DEMANDA DE  
ENERGIA ELÉTRICA**

**LAVRAS-MG**

**2013**

**MAÍRA RODRIGUES VILLAMAGNA**

**SELEÇÃO DE MODELOS DE SÉRIES TEMPORAIS E REDES  
NEURAS ARTIFICIAIS NA PREVISÃO DE CONSUMO E DEMANDA  
DE ENERGIA ELÉTRICA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do Título de Mestre.

Orientadora

Dra. Thelma Sáfadi

Coorientador

Dr. Fabiano Lemes Ribeiro

**LAVRAS-MG**

**2013**

**Ficha Catalográfica Elaborada pela Divisão de Processos Técnicos da  
Biblioteca da UFLA**

Villamagna, Máira Rodrigues.

Seleção de modelos de séries temporais e redes neurais artificiais na previsão de consumo e demanda de energia elétrica / Máira Rodrigues Villamagna. – Lavras : UFLA, 2013.

113 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2013.

Orientador: Thelma Sáfadi.

Bibliografia.

1. Universidade Federal de Lavras. 2. Inteligência artificial. 3. Tendência. 4. Sazonalidade. 5. Modelos de Box e Jenkins. I. Universidade Federal de Lavras. II. Título.

CDD – 519.55

**MAÍRA RODRIGUES VILLAMAGNA**

**SELEÇÃO DE MODELOS DE SÉRIES TEMPORAIS E REDES  
NEURAS ARTIFICIAIS NA PREVISÃO DE CONSUMO E DEMANDA  
DE ENERGIA ELÉTRICA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do Título de Mestre.

APROVADA em 27 de fevereiro de 2013.

Dr. Joaquim Paulo da Silva                      UFLA

Dr. Wilian Soares Lacerda                      UFLA

Dra. Thelma Sáfadi  
Orientadora

Dr. Fabiano Lemes Ribeiro  
Coorientador

**LAVRAS-MG**

**2013**

À minha filha Rara, com todo amor e carinho; à minha  
família e aos educadores que contribuíram para a  
minha formação.

**DEDICO.**

## **AGRADECIMENTOS**

Muitas pessoas passam por nossas vidas para deixar um pouco do que sabem e cabe a nós aproveitar cada ensinamento e agradecer humildemente por esta contribuição.

Portanto, agradeço à minha família: meus pais, Salete e David, pelos ensinamentos diários; minhas irmãs, Sara e Ada, pelo companheirismo e minha filha, Rara, por me mostrar que a felicidade está nas coisas simples da vida.

Agradeço a minha orientadora, Thelma e ao meu coorientador, Fabiano, pelo apoio, confiança e por disponibilizarem seu tempo em orientações e debates, esclarecendo e motivando-me.

Agradeço aos professores Wilian e Joaquim, por serem tão atenciosos e contribuírem de forma significativa com meu trabalho.

Agradeço ao grupo IC-Complex, pela amizade e bate papos que são tão ou mais instrutivos que muitas aulas teóricas.

Aos colegas de curso que, com a convivência, tornaram-se verdadeiros amigos.

Aos professores e funcionários do Departamento de Ciências Exatas (DEX).

À Universidade Federal de Lavras e ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão de bolsa de estudos.

Sou eternamente grata àqueles que contribuíram tanto para a minha formação profissional quanto para o meu crescimento pessoal.

## RESUMO

Neste trabalho foram analisadas as séries mensais de consumo e demanda de energia elétrica da Universidade Federal de Lavras (UFLA), no período de janeiro de 1995 a dezembro de 2011, com o objetivo de prever os valores futuros através dos métodos de Box e Jenkins e Redes Neurais Artificiais (RNA). Foram realizadas comparações entre as metodologias sugeridas, por meio do erro quadrático médio de previsão (EQMP) e erro percentual absoluto médio de previsão (MAPE), para verificar o desempenho das mesmas. A metodologia Box e Jenkins apresentou-se eficiente para representar a série em estudo, captando os efeitos de tendência e sazonalidade ocorridos. Os resultados mostram que as Redes Neurais Artificiais obtiveram melhores valores nos índices de desempenho testados e se apresentam como um método satisfatório para a previsão um passo à frente do consumo e da demanda de energia elétrica.

Palavras chave: Séries Temporais. Redes Neurais Artificiais. Box-Jenkins. Previsão. Energia Elétrica.

## **ABSTRACT**

This study analyzed the series of consumption and electricity demand for University Federal of Lavras (UFLA), covering the period from January 1995 to December 2011, with the objective of predicting future values by the methods of Box and Jenkins and Artificial Neural Networks (ANN). Comparisons were made between the methodologies suggested to check the performance of the same, in order to generate more accurate figures for next year. The Box-Jenkins methodology presented efficient to represent the series under study, capturing the effects of seasonality and trend occurred. The results show that the artificial neural network had a better forecasting performance of the demand and consumption of electricity.

Keywords: Time Series. Artificial Neural Network. Box-Jenkins. Forecasting.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>10</b>
1.1	Objetivo Geral.....	12
1.2	Objetivo Específico.....	12
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>13</b>
2.1	Energia elétrica.....	13
2.1.1	Histórico da política de energia elétrica no Brasil.....	15
2.1.2	Sistema tarifário do setor elétrico.....	16
2.1.3	Tarifação da Universidade Federal de Lavras.....	19
2.2	Séries temporais.....	20
2.2.1	Componente tendência.....	21
2.2.2	Componente sazonal.....	22
2.2.3	Modelos Box e Jenkins.....	23
2.2.3.1	Modelos ARIMA sazonal-SARIMA.....	25
2.2.3.2	Ciclo iterativo.....	26
2.2.4	Análise de intervenção.....	28
2.3	Redes neurais.....	29
2.3.1	Neurônio biológico.....	31
2.3.2	Neurônio artificial.....	32
2.3.3	Funções de ativação.....	34
2.3.4	Arquitetura da rede.....	37
2.3.5	Perceptron de camada única.....	39
2.3.6	Perceptron de múltiplas camadas.....	41
2.3.7	Processamento temporal.....	45
<b>3</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>50</b>
3.1	Material.....	50
3.2	Métodos.....	51
3.2.1	Metodologia Box e Jenkins.....	52
3.2.2	Metodologia RNA.....	53
3.3	Avaliação do desempenho dos modelos.....	55
<b>4</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>57</b>
4.1	Modelagem por Box e Jenkins.....	57
4.1.1	Série de consumo de energia elétrica em horário de ponta.....	57
4.1.2	Série de consumo de energia elétrica em horário fora de ponta.....	65
4.1.3	Série da demanda de energia elétrica em horário de ponta.....	70

4.1.4	Série da demanda de energia elétrica em horário fora de ponta.....	75
4.2	Modelagem por redes neurais artificiais.....	79
4.2.1	Série de consumo de energia elétrica em horário de ponta.....	81
4.2.2	Série de consumo de energia elétrica em horário fora de ponta.....	85
4.2.3	Série da demanda de energia elétrica em horário de ponta.....	89
4.2.4	Série da demanda de energia elétrica em horário fora de ponta.....	92
4.3	Análise comparativa.....	96
4.3.1	Análise comparativa do consumo de energia elétrica nos Horários de ponta e fora de ponta.....	96
4.3.2	Análise comparativa da demanda de energia elétrica nos horários de ponta e fora de ponta.....	98
5	CONCLUSÃO.....	101
	REFERÊNCIAS.....	103
	APÊNDICE.....	107

## 1 INTRODUÇÃO

A energia elétrica é um recurso indispensável e estratégico para o desenvolvimento socioeconômico global. O setor elétrico brasileiro é caracterizado pela predominância de usinas hidrelétricas no seu parque gerador. De acordo com o Ministério de Minas e Energia (BRASIL, 2012), aproximadamente 81,7% da capacidade de geração instalada são de origem hidráulica, devido, em grande parte, às condições geográficas favoráveis. Outra característica é a presença de grandes extensões de linhas de transmissão.

Um dos grandes fatores que fundamentam o uso e o aproveitamento correto dessa modalidade de energia são as formas de economia e a previsão de ações concernentes ao consumo de eletricidade. A busca por um constante aprimoramento de técnicas que levem à identificação das tendências principais de sua evolução, visa, primordialmente, reduzir os desvios entre o mercado previsto e o realizado, notadamente por ser o setor elétrico caracterizado por investimentos de grande porte e lenta maturação (CAIO; BERMAN, 1998).

Diante do exposto, a aplicação de teorias que usam como base as séries temporais é fator de grande viabilidade para a previsão de ações de consumo de energia elétrica, expresso em quilowatt hora (kWh), bem como demanda de energia elétrica, expresso em potência instalada (kW), dentre as quais se destacam a metodologia de Box e Jenkins (1979) e a de Redes Neurais Artificiais (RNA).

Na literatura, têm sido propostas estas e outras abordagens para o estudo de consumo e demanda de energia, tais como o trabalho de Moura, Valiani e Galvão Neto (2011), que previram a demanda nacional de energia elétrica utilizando a metodologia Box e Jenkins, Prado (2011) realizou previsões utilizando esta metodologia, da série de consumo e demanda de energia da Universidade Federal de Lavras (UFLA).

Campos (2008) abordou modelos autoregressivos integrados de médias móveis (ARIMA), modelos não lineares autorregressivos integrados de médias móveis (NARIMA), redes neuro-fuzzy (RNF) e redes neurais artificiais (RNA) para a previsão do consumo de energia do estado de Minas Gerais, afirmando que são ferramentas eficientes que podem auxiliar no planejamento e na tomadas de decisões no setor elétrico. Romera, Morán e Fernández (2007) utilizam as Redes Neurais Artificiais para prever o consumo de energia elétrica na Espanha.

Santos, Valiani e Galvão Neto (2011) incluíram a temperatura na previsão da demanda de energia elétrica. Araújo (2005) incluiu, além da temperatura ambiente, a umidade, o PIB anual e a população anual dos 125 municípios do Rio Grande do Sul e apresentou bons resultados ao utilizar estes fatores.

No presente trabalho, integrado às ideias apresentadas por Prado (2011), analisa-se a aplicação de uma nova metodologia para a previsão do consumo e da demanda de energia da Universidade Federal de Lavras (UFLA), as Redes Neurais Artificiais (RNA). A importância desta abordagem se deve, principalmente, ao fato de as RNAs apresentarem grande habilidade em aprender padrões complexos e pela capacidade de lidar com a não linearidade e não estacionariedade presentes em algumas séries temporais.

Neste contexto, é apresentada análise das séries do consumo e demanda de energia elétrica da UFLA, referente ao período de janeiro de 1995 a dezembro de 2011, mensalmente, com o objetivo de realizar previsões, um passo a frente, utilizando os métodos de Box e Jenkins e Redes Neurais Artificiais (RNA). Os resultados foram comparados, no intuito de verificar quais modelos apresentam melhor desempenho, ou seja, menor erro de previsão, fornecendo, desse modo, um cenário mais seguro para os gestores responsáveis pela tomada de decisão.

Este trabalho está estruturado da seguinte forma: no capítulo inicial, encontram-se as definições gerais de energia, bem como o histórico desta em nosso país. Ressaltam-se também sistemas de tarifação de energia no Brasil e para a Universidade Federal de Lavras (UFLA). As metodologias de Box e Jenkins e Redes Neurais Artificiais (RNA), explicitadas descritivamente, são alicerce de pesquisa e desenvolvimento deste trabalho, junto à previsão de consumo e de demanda da energia elétrica da UFLA. No capítulo final apresentam-se resultados do desenvolvimento desta pesquisa, em que, a partir das metodologias citadas são ressaltados resultados e discussões sobre a previsão de consumo e de demanda da energia elétrica na UFLA, finalizando com as conclusões.

### **1.1 Objetivo geral**

Identificar modelos de previsão para o consumo e a demanda de energia elétrica da Universidade Federal de Lavras (UFLA).

### **1.2 Objetivos específicos**

Aplicar as técnicas de séries temporais para ajustar modelo de previsão às séries de consumo e demanda de energia elétrica.

Aplicar as técnicas de Redes Neurais Artificiais às séries de consumo e demanda de energia elétrica para identificar modelos e realizar previsões.

Comparar os resultados provenientes da previsão dos métodos supracitados.

Verificar quais modelos apresentam menor erro de previsão, fornecendo um cenário mais seguro para os gestores responsáveis pela tomada de decisão.

## **2 REFERENCIAL TEÓRICO**

### **2.1 Energia elétrica**

A partir da Revolução Industrial, ao final do século XVIII, os processos produtivos e de consumo passaram a requerer cada vez mais energia, matéria-prima e mão de obra. A necessidade de energia elétrica para atender a uma população mundial em crescimento apresentou seus primeiros impactos na segunda metade do século XX. A globalização da economia, com forte apelo para um consumo desenfreado das sociedades, colocou em discussão a geração, a disponibilidade e o consumo de energia elétrica.

O desenvolvimento social e econômico brasileiro registrado nos últimos anos aumentou o poder de compra da população. Esse crescimento melhorou a qualidade de vida e o acesso a bens de consumo e serviços. Com as pessoas consumindo mais, todo o sistema de abastecimento do país precisa sustentar esse desenvolvimento, ou seja, produzir cada vez mais. Isso significa maior consumo de energia, predominantemente elétrica.

Para continuar a crescer, o Brasil precisa investir na geração de energia elétrica, ou corre o risco dos apagões – a falta de luz. Atualmente, 81,7% da energia gerada no país vêm das hidrelétricas e o resto é dividido entre fontes termelétricas, eólicas e usinas nucleares, como mostrado na Figura 2.1 (BRASIL, 2012).

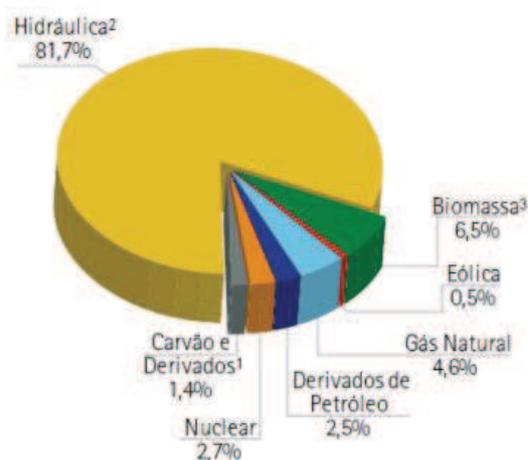


Figura 1 Matriz elétrica brasileira, 2011.<sup>1</sup>Inclui gás de coqueria, <sup>2</sup>Inclui importação, <sup>3</sup> Inclui lenha, bagaço de cana, lixívia e outras recuperações

Fonte: Brasil (2012)

No caso da energia elétrica gerada em hidrelétricas, a maior parte da produção é proveniente de usinas de grande porte, tais como Furnas, Itaipu e Três Marias. Esses projetos requerem altos investimentos e apresentam forte impacto ambiental. Atualmente, pequenas centrais hidrelétricas (PCH), assim como termoelétricas (UTE), estão recebendo estímulos governamentais, inclusive para o aproveitamento de energia renovável, como ocorre em plantas de produção de açúcar, álcool e bioenergia. De acordo com o Ministério de Minas e Energia, prevê-se que a participação das hidrelétricas deverá cair para 67%, enquanto a das PCHs deverá dobrar, chegando a quase 15% nos próximos anos.

### **2.1.1 Histórico da política de energia elétrica no Brasil**

Os serviços de energia elétrica no planejamento urbano das principais cidades brasileiras iniciaram-se nas duas últimas décadas do século XIX, com iluminação de trechos de ferrovias, iluminação de praças públicas, implantação e início de operação de pequenas usinas termoelétricas e hidroelétricas. Empresas estrangeiras vislumbraram o mercado da geração e transmissão de eletricidades no início do século XX, contudo, a falta de legislação e de regulamentação desses serviços impedia o desenvolvimento do setor. A primeira versão de uma lei específica sobre energia elétrica foi feita em 15 de junho de 1906, mas a matéria somente seria aprofundada depois da Primeira Guerra Mundial, com a legislação sobre quedas d'água de 1919 (SAES, 2009).

Com capital estrangeiro, as empresas São Paulo Light e Rio Light dominavam o mercado. Para regulamentar o setor energético, o poder legislativo se inspirava nas legislações da Europa e dos Estados Unidos. Em 1934, após inúmeras leis já em vigor, foi promulgado o Código de Águas, concedendo à União exclusividade na exploração de hidrelétricas para serviços públicos. As estatais e as federais do setor, como Furnas, Cemig, Escelsa, RGS, CEEE e outras, foram criadas em 1943. Frente à contínua e incessante demanda por energia elétrica no desenvolvimento da economia nacional, foi construída a maior hidroelétrica do país, a binacional Itaipu, em 1984 (ESPÍRITO SANTO CENTRAIS ELÉTRICAS, ESCELSA, 2012).

A Lei nº 9.427, de 26 de dezembro de 1996, criou a Agência Nacional de Energia Elétrica (ANEEL), vinculada ao Ministério de Minas e Energia, que determina a política energética no Brasil. O artigo 2º desta lei traça os objetivos principais da organização: “A Agência Nacional de Energia Elétrica - ANEEL tem por finalidade regular e fiscalizar a produção, transmissão, distribuição e

comercialização de energia elétrica, em conformidade com as políticas e diretrizes do governo federal” (BRASIL, 1996).

Atualmente, outros órgãos de assessoria, como o Conselho Nacional de Política Energética (CNPE), do Ministério de Minas e Energia, e a Empresa de Pesquisas Energéticas (EPE), além de outras, fortalecem as políticas adotadas para a organização do setor energético no Brasil. Embora o setor produtivo brasileiro tenha passado por apagões elétricos, o maior deles em 2001, que resultou em racionamento de energia elétrica por um ano, há grande preocupação, por parte do poder público, que investe, atualmente, em hidroelétricas como a de Belo Monte (PA), a de Santo Antônio do Jari (PA/AP), prevista para início em 2014 e a de Santo Antônio (RO), entre outras.

### **2.1.2 Sistema tarifário do setor elétrico**

Como citado anteriormente, no Brasil, as taxas cobradas, atualmente, pelas concessionárias de energia elétrica seguem o que é estabelecido pela ANEEL, que adota mecanismos para impedir atos de concentração de mercado e abuso tarifário. Situação essa diferente das tarifas cobradas nas décadas de 70 a 90 do século XX, quando os consumidores pagavam uma única tarifa de energia elétrica, independente da classificação ou da localidade.

Ocorre, portanto, a regulação, pela ANEEL, dos preços pagos pelos consumidores, a qual elaborou um plano de revisão tarifária, garantindo tarifas justas, cobradas para todas as categorias consumidoras: residencial, industrial e rural. Com isso, aumentam a eficiência e a qualidade do serviço de distribuição de energia elétrica.

Para o estudo do planejamento energético é necessário o conhecimento da tarifação da energia elétrica. As concessionárias, de acordo com a legislação

brasileira, calculam as faturas em função do consumo de energia, da demanda de potência e do fator de potência.

Consumo e demanda são definidos da seguinte forma:

A demanda de potência é medida em quilowatt (kW) e corresponde à média da potência elétrica solicitada pelo consumidor à empresa distribuidora, durante um intervalo de tempo especificado normalmente em 15 minutos e é faturada pelo valor medido durante o período de fornecimento, normalmente de 30 dias. O consumo de energia é medido em quilowatt-hora (kWh) e corresponde ao valor acumulado pelo uso de potência elétrica disponibilizada ao consumidor ao longo de um período de consumo, normalmente de 30 dias (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, ANEEL, 2005, p. 10).

Em circuitos elétricos, geralmente, têm-se potências ativas, reativas e aparentes. Define-se como fator de potência a razão entre potência ativa e potência aparente, e seu valor varia entre 0 e 1 (ELETROBRAS, 2011).

Para a elaboração das faturas, os consumidores são classificados em dois grupos, A e B, de acordo com o nível de tensão em que são atendidos, conforme apresentado na Tabela 1.

Tabela 1 Grupos de consumidores de acordo com o nível de tensão.

Grupo A - Alta tensão	Grupo B - Baixa tensão
A-1 - 230 kV ou mais	B-1 - Residencial
A-2 - 88 a 138 kV	B-1 - Residencial baixa renda
A-3 - 69 kV	B-2 - Rural
A-3a - 30 a 44 kV	B-3 - Não residencial, nem rural
A-4 - 2,3 a 13,8 kV	B-4 - Iluminação pública
A.S. - 2,3 a 13,8 kV (subterrâneo)	

Fonte: Eletrobras (2011).

Conforme a Resolução 456 da ANEEL (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, ANEEL, 2000), existem duas modalidades tarifárias. Dos consumidores do grupo B é cobrada apenas a energia que consomem (monômia) e, dos consumidores do grupo A, cobra-se tanto pela demanda quanto pela energia que consomem (binômia).

Para os consumidores do grupo A existem três alternativas tarifárias: convencional, horo-sazonal azul e horo-sazonal verde.

O cálculo das faturas no sistema convencional considera apenas os parâmetros de tributação, enquanto, no sistema horo-sazonal, são considerados os parâmetros de tributação e as variações horo-sazonais, ou seja, valores diferenciados segundo horários do dia e períodos do ano. As variações são:

- a) horário de ponta: corresponde ao intervalo de 3 horas consecutivas, ajustado de comum acordo entre a concessionária e o cliente, situado no período compreendido entre as 17h e 22h e, durante o horário de verão, das 18h à 23h;
- b) horário fora de ponta: corresponde às horas complementares ao horário de ponta;
- c) período seco: compreende os meses de maio a novembro de cada ano (sete meses);
- d) período úmido: compreende os meses de dezembro de um ano a abril do ano seguinte (cinco meses).

A Tarifa Azul aplica-se às unidades consumidoras que têm processo produtivo contínuo e é obrigatória para os consumidores dos tipos A-1, A-2 e A-3, sendo opcional para os demais. Enquanto a Tarifa Verde aplica-se aos consumidores com capacidade de modulação do processo produtivo, a Azul é opcional para os consumidores do Grupo A dos tipos A-3a, A-4 e A-S.

### 2.1.3 Tarifação da Universidade Federal de Lavras

A Universidade Federal de Lavras (UFLA), atualmente, se enquadra nos consumidores do grupo A4 (2,3 a 13,8 kV), na tarifação horo-sazonal azul. O horário de ponta definido entre a concessionária e esta instituição é das 19h às 22h e, no horário de verão, das 20h às 23h.

Sua fatura é calculada da seguinte forma:

$$Fatura = P_{consumo} + P_{demanda}$$

em que:

$P_{consumo}$  = tarifa de consumo no horário de ponta  $\times$  consumo medido no horário de ponta + tarifa de consumo no horário fora de ponta  $\times$  consumo medido no horário fora de ponta;

$P_{demanda}$  = tarifa de demanda no horário de ponta  $\times$  demanda medida no horário de ponta + tarifa de demanda no horário fora de ponta  $\times$  demanda medida no horário fora de ponta.

Quando a demanda medida ultrapassa a demanda contratada, acima do limite de 5% de tolerância, é cobrada uma parcela de ultrapassagem, dada por:

$P_{ultrapassagem}$  = tarifa de ultrapassagem no horário de ponta  $\times$  (demanda medida no horário de ponta – demanda contratada no horário de ponta) + tarifa de ultrapassagem no horário fora de ponta  $\times$  (demanda medida no horário fora de ponta – demanda contratada no horário fora de ponta).

As tarifas de consumo são diferenciadas por período do ano, sendo mais caras no período seco (maio a novembro). As tarifas de ultrapassagem são diferenciadas por horário, sendo mais caras no horário de ponta e as tarifas de demanda não são diferenciadas.

Esta seção foi dedicada à introdução dos conceitos básicos sobre a energia elétrica. Apresentou-se um breve histórico da política energética

brasileira e as melhorias no sistema tarifário do setor elétrico, ressaltando a importância de conhecer cada parâmetro da tarifação para poder inferir sobre ele. Nas próximas seções apresentam-se revisões dos métodos de previsão abordados neste trabalho. Será definido o conceito de séries temporais, exposto o método de modelagem de Box e Jenkins e apresentam-se discussões sobre as Redes Neurais Artificiais, seus principais componentes e sua aplicação em séries temporais.

## 2.2 Séries temporais

Uma série temporal é um conjunto de valores de uma variável, ordenados no tempo. A análise de séries temporais se distingue das demais análises estatísticas pelo fato de as observações serem registradas em instantes distintos e sucessivos, sendo dependentes entre si.

Segundo Morettin e Toloi (2006), os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas.

Com base nas componentes da série temporal são utilizados frequentemente os seguintes modelos:

(i) modelo aditivo

$$Z_t = T_t + S_t + a_t \quad (1)$$

(ii) modelo multiplicativo

$$Z_t = T_t \times S_t \times a_t \quad (2)$$

em que  $Z_t$  é a variável dependente,  $T_t$  é a componente tendência,  $S_t$  é a componente de sazonalidade e  $a_t$  é a componente aleatória.

Removendo-se as componentes  $T_t$  e  $S_t$ , o que sobra é a componente aleatória residual ou irregular,  $a_t$ . A suposição usual é a de que  $a_t$  seja uma série

puramente aleatória ou ruído branco, ou seja, independente e identicamente distribuída com média zero e variância constante.

O modelo aditivo é mais adequado quando a força do movimento sazonal e aleatório opera de forma absoluta, ou seja, não depende das outras componentes, como a tendência. Mas, se as amplitudes sazonais variam com a tendência, o modelo mais adequado é o multiplicativo.

As transformações são utilizadas para estabilizar a variância e tornar o efeito sazonal aditivo. Em algumas séries que têm tendência, pode ocorrer um acréscimo da variância à medida que o tempo passa. Neste caso, uma transformação logarítmica pode ser adequada.

A análise das séries temporais busca isolar e interpretar tais componentes. Em qualquer caso, o isolamento objetiva identificar componentes e levantar informações úteis ao estudo e a previsões de valores. No entanto, as várias componentes podem atuar de modo isolado ou inter-relacionado e as previsões com base em séries temporais só são, de fato, válidas se as propriedades de suas componentes permanecerem relativamente estáveis durante o tempo ao qual a previsão diz respeito (MILONE, 2004).

### 2.2.1 Componente tendência

A componente tendência pode ser entendida como um aumento ou uma diminuição gradual das observações ao longo de um período. Os métodos mais utilizados para estimar  $T_t$  são tendência polinomial, suavização e o método das diferenças.

Estimando-se a tendência por meio de  $\hat{T}_t$  obtém-se a série ajustada para tendência ou livre de tendência (supondo que a componente sazonal não esteja presente).

$$Y_t = Z_t - \hat{T}_t \quad (3)$$

Os testes utilizados para identificação da componente tendência são o teste de sequências (Wald-Wolfowitz) e o teste do sinal (Cox-Stuart). Mais detalhes podem ser vistos em Morettin e Tolo (2006).

### 2.2.2 Componente sazonal

A componente sazonal mostra as flutuações ocorridas em subperíodos  $s$ , que podem ser medidos mensalmente, trimestralmente ou diariamente, por exemplo.

Para estimar a componente sazonal  $S_t$ , é preciso obter estimativas  $\hat{S}_t$  de  $S_t$  e subtrair a série estimada de  $Z_t$ .

$$Z_t^{SA} = Z_t - \hat{S}_t \quad (4)$$

Se o modelo for multiplicativo, da forma

$$Z_t = T_t S_t a_t \quad (5)$$

tomando-se o logaritmo, obtém-se o modelo aditivo. A série sazonalmente ajustada será

$$Z_t^{SA} = Z_t / \hat{S}_t \quad (6)$$

Ao se estimar  $S_t$  ocorre um erro de ajustamento sazonal que será ótimo minimizando-se  $E(\delta_t^2)$ .

$$\delta_t = S_t - \hat{S}_t \quad (7)$$

Existem vários procedimentos para estimar  $S_t$ . Os mais utilizados são o método de regressão e o método de médias móveis.

O método de regressão, segundo Morettin e Tolo (2006), é ótimo para séries que apresentam sazonalidade determinística, ou seja, que podem ser previstas perfeitamente a partir de meses anteriores. Já o método de médias móveis é apropriado quando se tem uma série temporal cuja componente sazonal varia com o tempo, ou seja, para séries cuja sazonalidade é estocástica.

O teste utilizado para identificação da componente sazonal é o teste de Fisher, descrito por Morettin e Toloí (2006), que utiliza as ordenadas do periodograma baseado nas estatísticas  $g$  e  $z$ , para testar a existência de periodicidade na série.

### 2.2.3 Modelos Box e Jenkins

A metodologia de Box e Jenkins (1976) consiste em ajustar modelos autorregressivos integrados de médias móveis, ARMA(p,d,q), a um conjunto de dados, em que  $p$  é a ordem do modelo autorregressivo AR(p),  $q$  a ordem do modelo de médias móveis MA(q) e  $d$  o número de diferenças necessárias para a série ser estacionária.

O processo autorregressivo AR(p) modela uma autorregressão da variável  $Z_t$  com essa mesma variável defasada ( $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ ), para os  $p$  períodos de defasagem em que a autocorrelação parcial entre as variáveis é significativa e cujo modelo escreve-se

$$\phi(B)Z_t = a_t \quad (8)$$

em que  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  é o operador autorregressivo de ordem  $p$ .

Em um processo de médias móveis MA(q), cada observação  $Z_t$  é gerada por uma média ponderada dos erros aleatórios  $q$  períodos no passado, cuja equação é

$$Z_t = \theta(B)a_t \quad (9)$$

em que  $\theta(B)a_t = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$  é o operador de médias móveis de ordem  $q$ .

Assume-se que os erros aleatórios são gerados por um processo de ruído branco, ou seja,  $a_t \sim N(0, \sigma_a^2)$  e covariância  $\gamma_k = 0$  para defasagens  $k \neq 0$ .

Os modelos autorregressivos e de médias móveis ARMA(p,q) vão modelar o grau de autocorrelação e de observações defasadas. Em sua forma geral, se escrevem

$$\phi(B)Z_t = \theta(B)a_t \quad (10)$$

Para modelar a série temporal, é necessário que as variáveis sejam estacionárias. Caso contrário, devem ser aplicadas diferenças, até que a condição de estacionariedade se torne válida. O número de diferenças necessárias é denominado ordem de integração (I).

Uma série temporal é estacionária quando se desenvolve no tempo, aleatoriamente, ao redor de uma média constante, refletindo uma forma de equilíbrio estável.

Para tornar a série temporal estacionária é necessário transformar os dados originais, e o procedimento mais comum consiste em tomar diferenças sucessivas da série original.

A primeira diferença de  $Z_t$  é definida por

$$\Delta Z_t = Z_t - Z_{t-1}, \quad (11)$$

a segunda diferença é

$$\Delta^2 Z_t = \Delta[\Delta Z_t] = \Delta[Z_t - Z_{t-1}] \quad (12)$$

ou seja,

$$\Delta^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2} \quad (13)$$

logo, a  $n$ -ésima diferença de  $Z_t$  é

$$\Delta^n Z_t = \Delta[\Delta^{n-1} Z_t]. \quad (14)$$

Normalmente, será necessário tomar uma ou duas diferenças para que a série se torne estacionária.

Portanto, o modelo para a série diferenciada ARIMA(p,d,q) é

$$\phi(B)(1 - B)^d Z_t = \theta(B)a_t \quad (15)$$

em que  $\phi(B)$  e  $\theta(B)$  são operadores autorregressivos e de médias móveis descritos anteriormente;  $(1 - B)$  é o operador diferença e  $d$  é o número de diferenças.

### 2.2.3.1 Modelos ARIMA sazonal – SARIMA

Geralmente, séries de consumo e demanda de energia apresentam comportamentos periódicos e, mesmo após eliminar a sazonalidade determinística, pode haver autocorrelação significativa em “lags” de baixa ordem ou em “lags” sazonais, múltiplos de período  $s$ .

Portanto, há necessidade de se considerar uma sazonalidade estocástica, ou seja, ajustar à série original um modelo ARIMA sazonal (SARIMA).

Um modelo SARIMA  $(p,d,q)(P,D,Q)_s$  pode ser escrito como

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D Z_t = \theta(B)\theta(B^s)a_t \quad (16)$$

em que

$s$  é o período;

$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}$  é o operador autorregressivo de ordem  $P$ ;

$\theta(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_Q B^{Qs}$  é o operador de médias móveis sazonal de ordem  $Q$ ;

$(1 - B^s)^D$  é o operador diferença sazonal e  $D$  indica o número de diferenças sazonais.

### 2.2.3.2 Ciclo iterativo

Box e Jenkins (1979) propõem um ciclo iterativo para a construção de modelos, no qual a escolha do modelo é feita com base nos próprios dados. As etapas do ciclo são:

- (1) especificação: uma classe geral do modelo é considerada para análise;
- (2) identificação de um modelo, com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (3) fase de estimação, na qual os parâmetros do modelo identificado são estimados;
- (4) fase de verificação ou diagnóstico do modelo ajustado, por meio de uma análise de resíduos, para saber se este é adequado para fins em vista, neste caso, previsão.

Na fase de identificação do modelo, o objetivo é encontrar valores parcimoniosos de  $p$ ,  $d$  e  $q$  do modelo especificado ARIMA( $p,d,q$ ).

O procedimento de identificação é feito, basicamente, em três etapas, que são:

- a) verificar a necessidade de transformação para estabilizar a variância, por meio do gráfico amplitude  $\times$  média, o qual consiste em dividir o conjunto de dados em subconjuntos de mesmo tamanho, encontrar a média ( $\bar{Z}$ ) e a amplitude ( $\omega$ ) de cada um desses subconjuntos. No gráfico, o eixo das abscissas traz as médias e o eixo das ordenadas, as amplitudes. Se  $\omega$  depender de  $\bar{Z}$ , obtém-se uma reta paralela ao eixo das abscissas e, neste caso, não há necessidade de transformação. Se  $\omega$  for diretamente proporcional a  $\bar{Z}$ , a transformação logarítmica é apropriada;
- b) fazer diferenças para obter uma série estacionária;
- c) identificar as ordens do modelo por meio das funções de autocorrelação e autocorrelação parcial.

Para se identificar um modelo SARIMA  $(p,d,q) \times (P,D,Q)_s$ , deve-se, inicialmente, determinar os valores de  $d$  e  $D$  por meio da análise do gráfico das funções de autocorrelação e de autocorrelação parcial. Se valores elevados e persistentes nos “lags” com defasagens próximas aos múltiplos do período sazonal (para  $s = 12, 24$  e  $36$ , por exemplo) pode-se inferir também a necessidade de diferenciação sazonal da série.

Na etapa de estimação é necessário estimar os parâmetros do modelo AR, os parâmetros do modelo MA e a variância do erro. A estimação pode ser feita utilizando-se o método dos mínimos quadrados ou máxima verossimilhança.

Após estimar o modelo, é feita uma análise dos resíduos para verificar a adequação dos modelos ajustados. Box e Jenkins (1979) sugerem o uso do teste de Box e Pierce (1970) para verificar se os resíduos apresentam um processo de ruído branco.

O teste é baseado nas primeiras  $k$  autocorrelações  $(r_k)$  dos resíduos e pode indicar se os valores no comportamento de ruído branco são muito altos.

Diversos modelos podem ser identificados para representar uma série. Dentre os avaliadores da qualidade de ajuste de modelos de séries temporais, tem-se o Critério de Informação de Akaike (AIC):

$$AIC(k, l) = \ln(\sigma_{k,l}^2) + \frac{2(k+l)}{N} \quad (17)$$

em que,  $\sigma_{k,l}^2$  é o estimador de máxima verossimilhança da variância dos resíduos do modelo ARMA(k,l) ajustado às  $N$  observações da série. Dependendo do valor de  $k$  e de  $l$ , muitos modelos têm que ser ajustados, a fim de se obter o mínimo de AIC (MORETTIN; TOLOI, 2006).

Após a construção de um modelo ARIMA(p,d,q) ou SARIMA (p,d,q) x (P,D,Q)<sub>s</sub>, por meio do ciclo iterativo citado anteriormente, utiliza-se o modelo identificado e estimado para fazer previsões.

A previsão de  $Y_{t+h}$ , para  $h = 1, 2, \dots$ , é denotada por  $\hat{Y}_t(h)$  e é definida como a esperança condicional de  $Y_{t+h}$ , dados todos os valores passados, ou seja,

$$\hat{Y}_t(h) = E[Y_{t+h} | Y_t, Y_{t-1}, \dots] \quad (18)$$

O erro de previsão é dado por

$$e_t(h) = Y_{t+h} - \hat{Y}_t(h) \quad (19)$$

em que  $Y_{t+h}$  é o valor real da série e  $\hat{Y}_t(h)$  é o valor previsto.

#### 2.2.4 Análise de intervenção

Em algumas séries temporais podem ocorrer intervenções que afetam a série, temporariamente ou permanentemente. Segundo Morettin e Toloí (2006), uma intervenção é a ocorrência de algum tipo de evento, em um instante T, conhecido a priori. A análise de intervenção tem por objetivo avaliar o impacto de um determinado evento no comportamento da série.

Uma intervenção pode afetar uma série temporal de várias maneiras. Ela pode ser abrupta ou gradual na sua manifestação e permanente ou temporária na sua duração.

Uma classe geral de modelos, que leva em conta a ocorrência de múltiplas intervenções, é dada por

$$Z_t = \sum_{j=1}^k v_j(B) X_{j,t} + N_t, \quad (20)$$

em que

$X_{j,t}$  é uma variável binária de intervenção. No caso do presente trabalho, observou-se a do tipo impulso, em que o efeito de intervenção é temporário, afetando a série apenas no instante T.

$$X_{j,t} = I_t^{(T)} = \begin{cases} 0, & t \neq T; \\ 1, & t = T, \end{cases} \quad (21)$$

$v_j(B)$  são funções racionais da forma  $\frac{\omega_j(B)B_j^b}{\delta_j(B)}$ , em que  $\omega_j(B) = \omega_{j,0} - \omega_{j,1}B - \dots - \omega_{j,s}B^s$  e  $\delta_j(B) = 1 - \delta_{j,1}B - \dots - \delta_{j,r}B^r$  são polinômios em  $B$ ,  $b_j$  é a defasagem no tempo para o início do efeito da  $j$ -ésima intervenção e  $N_t$  é a série temporal livre do efeito de intervenções e é denominada série residual.

Para maiores detalhes, podem ser consultadas, dentre outras, as referências de Morettin e Toloí (2006) e Box e Jenkins (1979).

### 2.3 Redes neurais artificiais

Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no cérebro humano e têm a capacidade de reconhecer, classificar, converter e aprender padrões. Em vista disso, são empregadas em diversas áreas do conhecimento, tais como reconhecimento de padrões, processamento de sinais digitais e imagens, otimização, visão computacional, robótica, previsão de séries temporais e outros (BRAGA et al., 1998).

Estudos realizados por McCulloch e Pitts (1943) foram os pioneiros na área de inteligência artificial. Inspirados no neurônio biológico, estes pesquisadores dedicaram sua carreira à tentativa de representar e modelar eventos no sistema nervoso. Eles propuseram uma discussão sobre redes lógicas de neurônios artificiais, cujas características foram inspiradas no neurônio biológico. Mas, se concentraram apenas em descrever um modelo artificial do neurônio e apresentar suas capacidades computacionais, e não apresentaram técnicas de aprendizado.

Um desenvolvimento significativo para as redes neurais foi realizado por Hebb (1949), que propôs uma regra de aprendizagem fisiológica para a modificação sináptica quando houver sincronismo entre os níveis de atividades de entrada e saída. Seguindo esta linha de pesquisa, Rosenblatt (1958) desenvolveu o perceptron de aprendizado supervisionado de camada única e Widrow e Hoff (1960) sugeriram um modelo parecido, o adaline, sobre padrões linearmente separáveis por hiper-superfícies, com o auxílio da regra delta de aprendizagem.

Após um artigo de Minsky e Papert (1969) sobre as limitações do perceptron, as redes neurais passaram a ser vistas de outra forma e, por um longo tempo, seu estudo foi reduzido. A partir dos anos 1980, devido aos avanços computacionais, Hopfield, Kohonen, dentre outros, retomaram o estudo de redes neurais com novas propostas e soluções. Hopfield (1982) explicitou o princípio do armazenamento de informação em redes dinamicamente estáveis e Kohonen publicou um artigo sobre os mapas auto-organizáveis. Rumelhart, Hinton e Williams (1986) também influenciaram a volta do interesse em redes neurais artificiais e revolucionaram o estudo sobre o treinamento do perceptron múltiplas camadas, com a demonstração do algoritmo de aprendizagem por retropropagação.

Nas últimas duas décadas, houve um progresso notável em relação ao estudo das redes neurais artificiais e espera-se que, com o avanço da tecnologia, um conhecimento mais sofisticado sobre este tema seja adquirido.

Para a operação de uma rede neural artificial, o neurônio artificial é uma unidade de processamento de informação representada por grafos. Inspirados no funcionamento do neurônio biológico, os neurônios artificiais conectados em redes têm a capacidade de aprender e generalizar, mas é importante reconhecer que são bastante primitivos, em relação à complexidade do nosso cérebro. Nas

seções seguintes, apresenta-se, de forma simplificada, o funcionamento de um neurônio biológico e de um neurônio artificial e suas possíveis semelhanças.

### **2.3.1 Neurônio biológico**

O neurônio biológico é composto, de forma simplificada, por um corpo celular ou soma, dendritos e axônio. A informação do neurônio é transmitida sob a forma de impulsos nervosos em sentido unidirecional: corpo celular, axônio, dendritos e neurônio seguinte. Esses sinais são transmitidos de neurônio pré-sináptico para outro pós-sináptico, por meio de junções interneurais chamadas de sinapses.

Nas sinapses químicas, o primeiro neurônio secreta, na fenda sináptica, uma substância química chamada neurotransmissor, que atua sobre proteínas receptoras localizadas na membrana do neurônio seguinte para excitá-lo ou inibi-lo.

O terminal pré-sináptico tem estruturas internas para a função excitatória ou inibitória da sinapse: as vesículas sinápticas, que contêm substâncias transmissoras, as quais são liberadas na fenda sináptica. Essas substâncias se excitam se a membrana neuronal do pós-sináptico contiver receptores excitatórios e o inibem se ela tiver receptores inibitórios (GUYTON; HALL, 1997).

Na Figura 2 estão representados os elementos principais de um neurônio biológico.

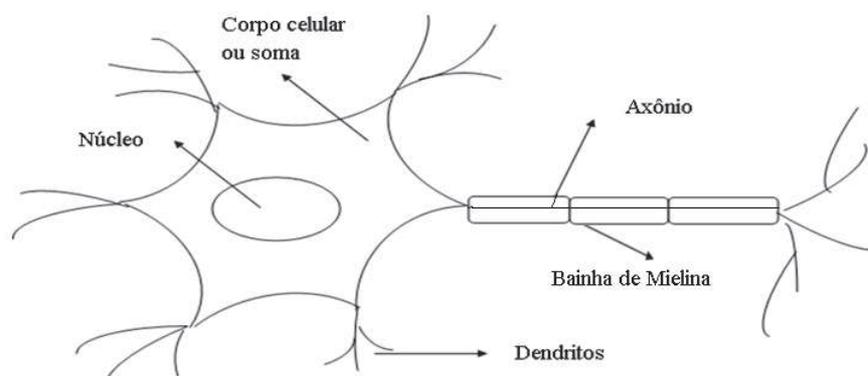


Figura 2 Componentes principais do neurônio biológico

### 2.3.2 Neurônio artificial

Na Figura 3 observa-se um modelo no qual podem ser vistos sinapses (representadas pelas entradas) e pesos sinápticos, somatório e função de ativação. Cada sinapse é caracterizada por um estímulo de entrada multiplicado pelo seu peso sináptico correspondente. Depois desta multiplicação, cada sinal de entrada é somado e o resultado é, então, aplicado a uma função de ativação que restringe a saída do neurônio.

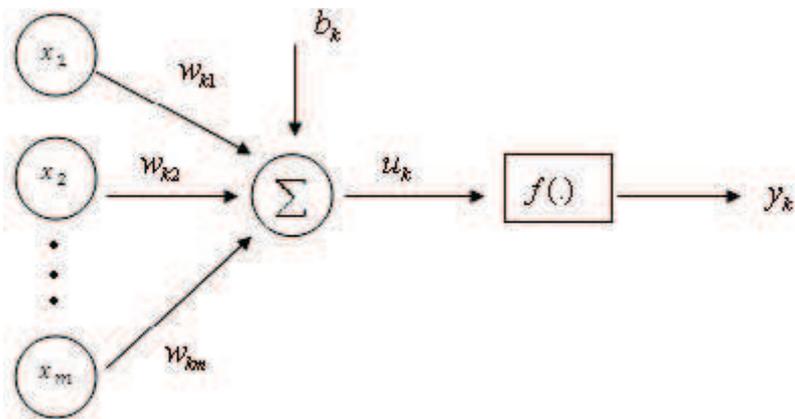


Figura 3 Modelo geral do neurônio artificial

Em termos matemáticos, um neurônio  $k$  é escrito da seguinte forma:

$$u_k = \sum_{j=1}^m w_{kj} x_j + b_k. \quad (22)$$

A função é calculada pela combinação linear dos pesos e das entradas ou, equivalentemente, ao produto escalar do vetor de pesos  $w^T = [w_{k1} \ w_{k2} \ \dots \ w_{km}]$  e vetor de entradas  $x^T = [x_1 \ x_2 \ \dots \ x_m]$

$$u_k = w^T x + b_k. \quad (23)$$

A saída é dada aplicando-se uma função de ativação

$$y_k = f(u_k) \quad (24)$$

em que

$x_1, x_2, \dots, x_m$  são os sinais de entradas (números reais ou binários);

$w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos que indicam a força de entrada para a resolução de um problema (valores positivos ou negativos);

$u_k$  é o resultado da junção somadora dos sinais de entrada, ponderados pelas respectivas sinapses;

$f(.)$  é a função de ativação que restringe a amplitude da saída do neurônio a um valor finito;

$b_k$  é o viés, que possui o efeito de aumentar ou diminuir a entrada da função de ativação;

$y_k$  é o sinal de saída do neurônio.

Segundo Haykyn (2001), o conhecimento é adquirido a partir de seu ambiente, por meio de um processo de aprendizagem e os pesos sinápticos são utilizados para armazenar o conhecimento adquirido.

O processo de aprendizado consiste em ajustar os pesos sinápticos e bias, de forma que a aplicação de um conjunto de entradas produza um conjunto de saídas desejadas. O tipo de aprendizado é determinado pela maneira como os parâmetros são modificados.

O processo em que a rede é treinada para fornecer a saída desejada a um estímulo de entrada específico é denominado aprendizado supervisionado, no qual não há uma saída específica em relação aos estímulos de entrada. A rede se auto-organiza em relação às particularidades do conjunto de entrada.

No presente trabalho, utilizou-se o aprendizado supervisionado para treinamento da rede. Nas seções seguintes, descrevem-se o algoritmo de aprendizagem utilizado e os principais tipos de funções de ativação e arquiteturas da rede mais utilizadas na literatura.

### **2.3.3 Funções de ativação**

A função de ativação define a saída do neurônio a partir dos valores dos pesos e de entradas. Os principais tipos de função de ativação são a função de

limiar ou degrau, a função sigmoide e a função linear, cujas equações e gráficos são apresentadas a seguir.

A saída do neurônio  $k$  que emprega a função de limiar é expressa como

$$f(u_k) = y_k = \begin{cases} 1 & \text{se } u_k \geq 0 \\ 0 & \text{se } u_k < 0 \end{cases} \quad (25)$$

Nesta função, a saída de um neurônio assume o valor 1 se o campo local induzido daquele neurônio é não negativo e 0, caso contrário, ou seja, as sinapses correspondentes são inibitórias ou excitatórias. Seu gráfico é apresentado na Figura 4.

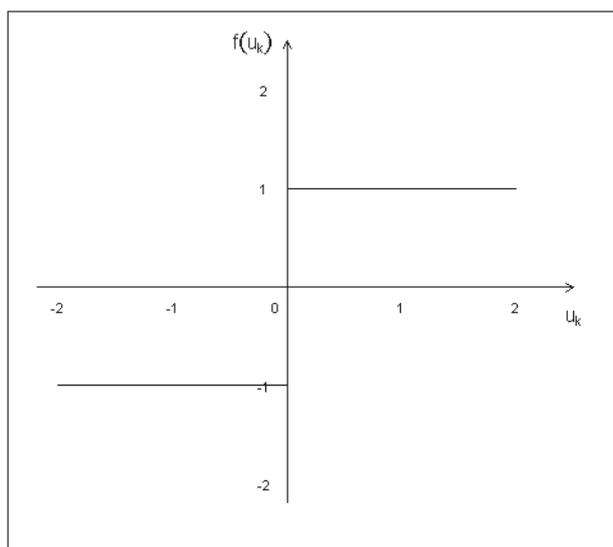


Figura 4 Gráfico da função limiar ou degrau

Uma aproximação contínua da função degrau é a função de ativação sigmoideal que, segundo Haykin (2001), é uma das mais utilizadas no processo de redes neurais, pois possui um comportamento estritamente crescente que exibe um balanceamento entre comportamento linear e não linear. Um exemplo de função sigmoide é a função logística, definida por

$$f(u_k) = \frac{1}{1+e^{-au_k}}, \quad (26)$$

em que  $a$  é o parâmetro de inclinação da função sigmoide. Ela assume um intervalo contínuo entre valores de 0 a 1 (Figura 5) e é diferenciável em todos os pontos, fato que auxilia na aproximação de funções contínuas.

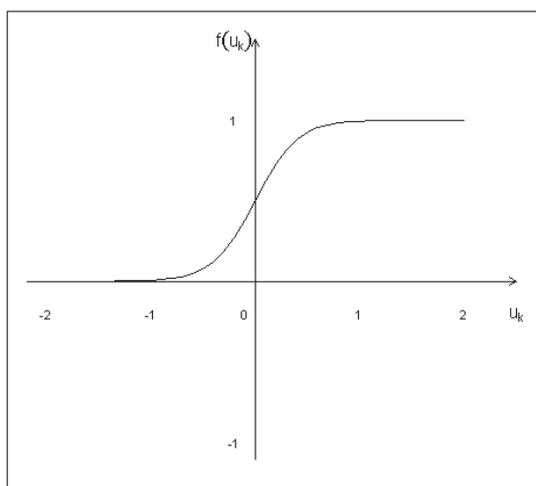


Figura 5 Gráfico da função sigmoide

A função linear também pode ser utilizada, dependendo do tipo de problema a ser abordado.

$$f(u_k) = au_k \quad (27)$$

Nesta equação,  $a$  é o parâmetro de inclinação que define a saída linear  $y$  para os valores de entrada  $u_k$ . Na Figura 6 está representado o gráfico da função.

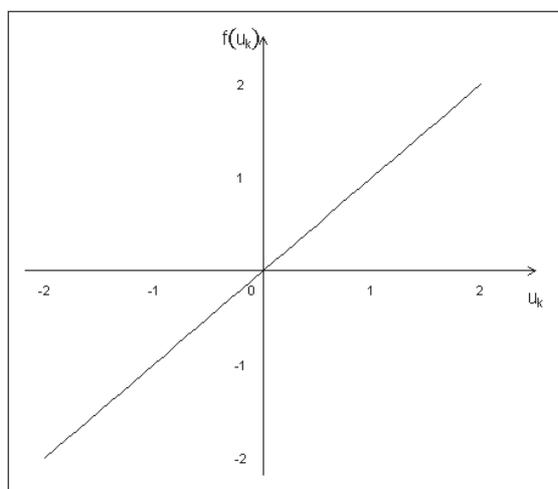


Figura 6 Gráfico da função linear

### 2.3.4 Arquitetura da rede

Em uma rede neural, os neurônios estão dispostos em camadas. Quanto ao número de camadas, eles podem ser classificados em redes alimentadas de uma única camada e redes alimentadas com múltiplas camadas.

A arquitetura da rede (maneira como os neurônios estão arranjados) está fortemente relacionada com o algoritmo utilizado para treinar a rede.

Nas redes com camada única, representadas na Figura 7, tem-se uma camada de entrada que se projeta sobre a camada de saída de neurônios. São aplicadas em reconhecimento de padrões e memória associativa.

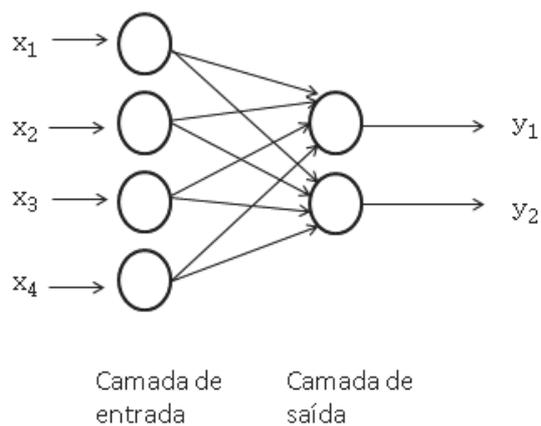


Figura 7 Exemplo de rede alimentada com 4 neurônios na camada de entrada e 2 neurônios na camada de saída

Redes com camadas múltiplas se distinguem pela presença de uma ou mais camadas ocultas de neurônios. As camadas ocultas intervêm entre a entrada externa e a saída, tornando a rede capaz de extrair estatísticas de ordem elevada (Haykin, 2001). São utilizadas na aproximação de funções, reconhecimento de padrões, identificação e controle.

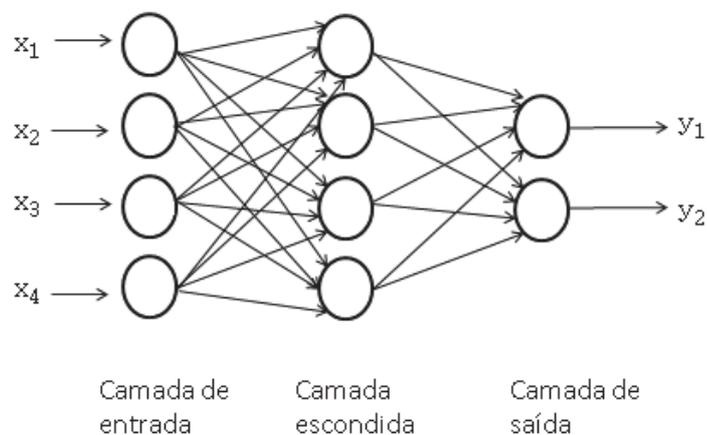


Figura 8 Exemplo de rede alimentada diretamente com 4 neurônios na camada de entrada, 4 neurônios na camada escondida e 2 neurônios na camada de saída

### 2.3.5 Perceptron de camada única

Redes de camada única com função de ativação limiar foram estudadas por Rosenblatt (1958), que as chamou de perceptron. Essas redes foram aplicadas a problemas de classificação de padrões linearmente separáveis, em que as entradas eram, geralmente, imagens binárias de caracteres.

Problemas linearmente separáveis são aqueles em que a solução pode ser obtida por meio de uma reta ou hiperplano, que divide o espaço de soluções em regiões ou classes (BISHOP, 1997).

O objetivo do perceptron é classificar corretamente o conjunto de padrões da entrada em uma de duas classes  $C_1$  e  $C_2$ , cuja regra de decisão é

$$\begin{cases} +1, & y \geq 0 \text{ para todo vetor de entrada } x \text{ pertencente à classe } C_1 \\ -1, & y < 0 \text{ para todo vetor de entrada } x \text{ pertencente à classe } C_2 \end{cases} \quad (28)$$

Na Figura 9 observa-se um exemplo de problema linearmente separável, em que a classe  $C_1$  é representada pelo sinal “+” e a classe  $C_2$  é representada por “o”. Uma reta separa o espaço de soluções, pois se trata de um problema com apenas duas dimensões.

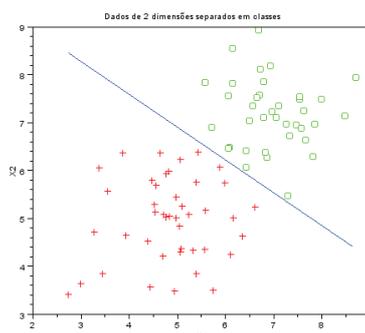


Figura 9 Ilustração de um problema linearmente separável, com uma linha reta como fronteira de decisão na separação de duas classes

O treinamento do perceptron é baseado no princípio de aprendizado de Hebb (1949), cujo objetivo é encontrar um vetor de peso  $w$  que satisfaça à regra de decisão, referente à equação 28. Para cada exemplo de treinamento apresentado à rede, subtrai-se o valor desejado  $d(n)$  pelo valor obtido no treinamento  $y(n)$ . Se a saída desejada for diferente da saída do neurônio, é feito um ajuste dos pesos, de acordo com a seguinte regra:

$$w(n + 1) = w(n) + \eta (d(n) - y(n))x(n) \quad (29)$$

em que  $\eta$  é a taxa de aprendizado e  $d(n) - y(n)$  é o erro  $e(n)$  obtido no processo de treinamento.

Em alguns problemas mais complexos, existem conjuntos de dados que não são separáveis apenas por uma reta e, sim, por um conjunto de retas ou

regiões espaciais mais complexas. Neste caso, o algoritmo de treinamento do Perceptron não converge. Na Figura 10 apresenta-se um exemplo de duas classes separadas por duas retas. Neste caso, foi utilizada uma rede neural de múltiplas camadas.

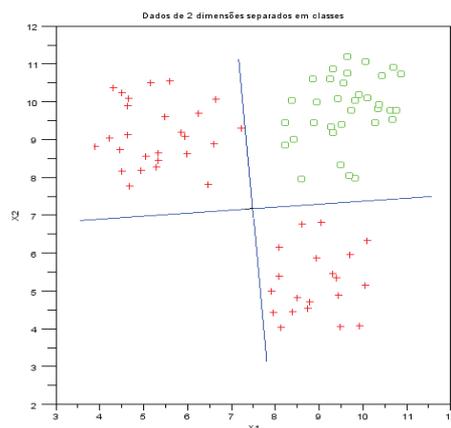


Figura 10 Ilustração de um problema com duas classes que são linearmente separáveis por duas retas

### 2.3.6 Perceptron múltiplas camadas

A rede de perceptron múltiplas camadas é constituída, basicamente, de camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A característica que o diferencia do perceptron simples é que cada neurônio da rede inclui uma função de ativação não linear, como, por exemplo, a função sigmoide. A rede contém uma ou mais camadas ocultas e entre elas existe um alto grau de conectividade, determinado pelas sinapses da rede.

Elas têm sido aplicadas em problemas complexos, por meio do treinamento supervisionado com um algoritmo conhecido como algoritmo de

retropropagação de erro (*error backpropagation*), desenvolvido por Rumelhart, Hinton e Willians (1986).

O perceptron de múltiplas camadas, juntamente com o algoritmo de treinamento por retropropagação, soluciona problemas não linearmente separáveis. Portanto, é aplicável a um número maior de problemas reais, como reconhecimento de caracteres, diagnóstico médico e previsão de séries temporais. Um exemplo no qual as classes não são linearmente separáveis é mostrado na Figura 11.

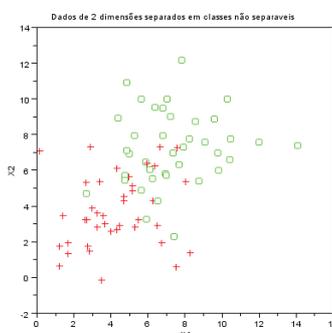


Figura 11 Ilustração de um problema com duas classes que não são linearmente separáveis

O algoritmo de retropropagação é um algoritmo de aprendizado supervisionado, baseado na correção de erros. Pode ser resumido em dois tipos de processamento, de acordo com Haykin (2001), que são:

- 1) processamento para frente: o sinal é propagado e aparece na camada de saída;
- 2) processamento de retroprogramação: o erro é corrigido camada a camada, começando pela camada de saída até a primeira camada oculta.

Seu objetivo é encontrar na superfície de erro um mínimo global, ou seja, valores para os pesos sinápticos que minimizam os erros da rede.

O desempenho de aprendizagem da rede é medido por meio da função de custo, para um dado conjunto de treinamento. A função custo adotada é o erro quadrático médio (EQM), que é obtido através do erro.

A saída desejada para as entradas  $d_j(n)$  é fornecida para a rede. O sinal de saída do neurônio  $j$ , representado por  $y_j(n)$ , é comparado com a saída desejada. Logo, o erro na saída, para o exemplo  $n$ , é definido por

$$e_j(n) = d_j(n) - y_j(n) \quad (30)$$

O valor instantâneo da energia do erro é definido por

$$\frac{1}{2}(e_j(n))^2 \quad (31)$$

e o valor instantâneo da energia total do erro é dado pela soma das energias dos neurônios

$$\varepsilon(n) = \frac{1}{2} \sum_{j=1}^c (e_j(n))^2, \quad (32)$$

em que  $c$  é o número total de neurônios na camada de saída da rede.

O erro quadrático médio é obtido calculando-se a média das energias instantâneas totais para todos os  $n$  padrões de entrada da rede, dada por

$$EQM = \frac{1}{N} \sum_{n=1}^N \varepsilon(n). \quad (33)$$

O objetivo do processo de aprendizagem é ajustar os parâmetros livres da rede (os pesos e bias) para minimizar o EQM.

O algoritmo de retropropagação aplica uma correção  $\Delta w_{ji}(n)$  ao peso sináptico  $w_{ji}(n)$ , que é proporcional à derivada parcial  $\partial \varepsilon(n) / \partial w_{ji}(n)$ . Como o vetor gradiente tem a mesma direção da maior variação do erro, o ajuste dos pesos deve se dar na direção contrária a este,  $\Delta w_{ji}(n) \propto -\nabla \varepsilon$ , justificando o sinal negativo da correção.

A correção do erro para cada camada é dada por

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \quad (34)$$

em que  $\eta$  é a taxa de aprendizado que define a velocidade com que os pesos são modificados. Ou, ainda,

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_j(n), \quad (35)$$

em que  $\delta_j(n)$  é o gradiente local.

Dessa forma, a atualização dos pesos é dada por

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n) \quad (36)$$

O gradiente local  $\delta_j(n)$  varia de acordo com a fase de processamento. Na fase de propagação para frente, o sinal de entrada  $x_j(n)$  é propagado camada a camada e a saída  $y_j(n)$  é comparada à saída desejada  $d_j(n)$  na camada de saída.

O gradiente local da saída  $\delta_j(n)$  é dado por

$$\delta_j(n) = e_j(n) f'(u_j(n)) \quad (37)$$

em que  $f'(\cdot)$  é a derivada da função de ativação e  $u_j(n)$  é o resultado da junção somadora dos sinais de entrada, ponderados pelo vetor de pesos  $w_{ji}(n)$ .

Quando o neurônio está localizado na camada oculta, não existe uma resposta esperada. Então, o sinal de erro é determinado em função dos neurônios da camada seguinte ao qual o neurônio está conectado. O gradiente local para neurônios da camada oculta  $l$  é dado por

$$\delta_j(n) = f'(u_j^{(l)}(n)) \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n), \quad (38)$$

em que o índice  $k$  se refere a um neurônio que se encontra em uma camada à direita do neurônio  $j$ , quando este pertence a uma camada oculta.

Além da taxa de aprendizado, pode ser incluído o termo momento ( $\alpha$ ) no algoritmo de retropropagação, que representa uma pequena modificação na atualização dos pesos, mas que, segundo Haykin (2001), pode evitar que o

processo de aprendizagem termine em um mínimo local raso na superfície do erro.

O processo de aprendizado é repetido para cada exemplo de treinamento apresentado à rede, até que um critério de parada seja satisfeito. Um dos critérios mais utilizados é definir o número máximo de épocas de treinamento ou um erro de tolerância, ou seja, quando a variação do erro quadrático médio (EQM) de uma época para outra atingir um valor suficientemente pequeno.

Após definir alguns conceitos básicos para a modelagem de Redes Neurais Artificiais, como o tipo de função de ativação e de arquiteturas da rede e o processo de aprendizado, será introduzido um tipo de processamento dinâmico, responsável por prever séries temporais.

### 2.3.7 Processamento temporal

Uma das principais aplicações das Redes Neurais Artificiais (RNA) é o reconhecimento de padrões temporais e previsão. O reconhecimento temporal de padrões requer um processamento que evolui no tempo. Para incorporar uma estrutura temporal na rede neural, que é estática, deve-se incorporar uma memória de curto prazo. Com isso, a saída da rede se torna uma função do tempo. Segundo Haykin (2001), a rede estática é responsável pela não linearidade e a rede dinâmica (memória) é responsável pelo tempo.

Especificamente, dado um sinal de entrada de valores presentes  $x(n)$  e de  $p$  valores passados  $x(n-1) \dots x(n-p)$  armazenados em uma memória de linha de atraso de tempo de ordem  $p$ , os parâmetros livres da rede neural são ajustados para minimizar o erro quadrático médio entre a saída da rede  $y(n)$  e a resposta desejada  $d(n)$ .

Na Figura 12 tem-se um exemplo de uma rede alimentada adiante atrasada no tempo, com atraso de ordem  $p$  e um perceptron de múltiplas

camadas. Para treinar o filtro, pode-se utilizar o algoritmo de retropropagação, descrito anteriormente.

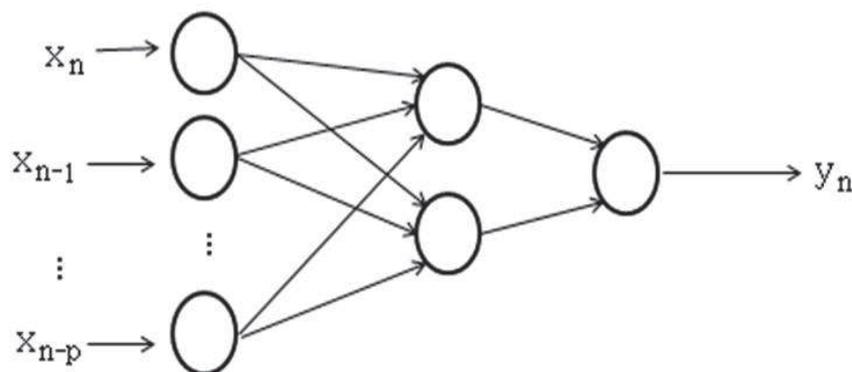


Figura 12 Rede utilizando a técnica de janelas de tempo para processamento temporal

Na Figura 13 está representado o exemplo de uma rede, com ordem  $p=12$ , que mostra como o atraso de tempo é processado. A primeira amostra de dados (janeiro a dezembro de 1995) representa a janela de ordem 12 e é utilizada para treinar a rede, gerando uma saída desejada (janeiro de 1996). Como o valor de janeiro de 1996 é conhecido, o erro de previsão para este um passo a frente pode ser calculado.

Na próxima amostra de dados a janela é atualizada, eliminando-se a amostra de dado mais antiga (janeiro de 1995) e deslocam-se as amostras restantes para frente. Utiliza-se a janela de dados atualizados para treinar novamente a rede e gerar outra saída desejada (fevereiro de 1996). Este processo é repetido até o último valor da série escolhida para treinamento.

Assim, teremos uma matriz com dados de entrada e um vetor com a saída desejada, conforme Figura 14. A matriz representa os dados de treinamento, cada coluna representa uma amostra de dados e cada linha é uma

entrada da rede neural. O vetor apresenta a saída desejada que se forma, à medida que os dados são atualizados.

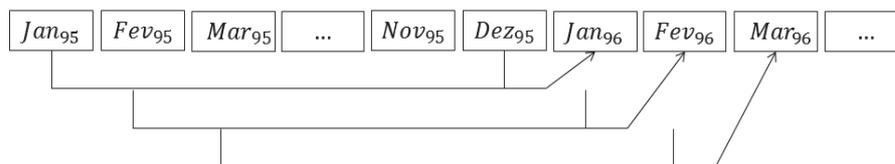


Figura 13 Exemplo de uma rede dinâmica com atraso de ordem 12 para previsão um passo a frente de uma série temporal, no período de janeiro de 1995 em diante

Jan <sub>95</sub>	Fev <sub>95</sub>	Mar <sub>95</sub>	Abr <sub>95</sub>	Mai <sub>95</sub>	Jun <sub>95</sub>	Jul <sub>95</sub>	Ago <sub>95</sub>	Set <sub>95</sub>	Out <sub>95</sub>	Nov <sub>95</sub>	Dez <sub>95</sub>	...
Fev <sub>95</sub>	Mar <sub>95</sub>	Abr <sub>95</sub>	...	...	...	...	...	...	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...
Mar <sub>95</sub>	Abr <sub>95</sub>	...	...	...	...	...	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...	Fev <sub>96</sub>	...
Abr <sub>95</sub>	Mai <sub>95</sub>	...	...	...	...	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	Mar <sub>96</sub>	...
Mai <sub>95</sub>	Jun <sub>95</sub>	...	...	...	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	...	Abr <sub>96</sub>	...
Jun <sub>95</sub>	Jul <sub>95</sub>	...	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	...	...	...	Mai <sub>96</sub>	...
Jul <sub>95</sub>	Ago <sub>95</sub>	...	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	...	...	...	...	Jun <sub>96</sub>	...
Ago <sub>95</sub>	Set <sub>95</sub>	...	Jan <sub>96</sub>	...	...	...	...	...	...	...	Jul <sub>96</sub>	...
Set <sub>95</sub>	Out <sub>95</sub>	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	...	...	...	...	...	Ago <sub>96</sub>	...
Out <sub>95</sub>	Nov <sub>95</sub>	Jan <sub>96</sub>	...	...	...	...	...	...	...	...	Set <sub>96</sub>	...
Nov <sub>95</sub>	Dez <sub>95</sub>	Jan <sub>96</sub>	...	...	...	...	...	...	Set <sub>96</sub>	Out <sub>96</sub>	Nov <sub>96</sub>	...
Dez <sub>95</sub>	Jan <sub>96</sub>	Fev <sub>96</sub>	Mar <sub>96</sub>	Abr <sub>96</sub>	Mai <sub>96</sub>	Jun <sub>96</sub>	Jul <sub>96</sub>	Ago <sub>96</sub>	Set <sub>96</sub>	Out <sub>96</sub>	Nov <sub>96</sub>	...
Jan <sub>96</sub>	Fev <sub>96</sub>	Mar <sub>96</sub>	Abr <sub>96</sub>	Mai <sub>96</sub>	Jun <sub>96</sub>	Jul <sub>96</sub>	Ago <sub>96</sub>	Set <sub>96</sub>	Out <sub>96</sub>	Nov <sub>96</sub>	Dez <sub>96</sub>	...

Figura 14 Matriz de uma rede dinâmica utilizada no processamento temporal de uma série temporal, no período de janeiro de 1995 em diante

O uso de RNA na previsão de séries temporais é cada vez mais frequente em pesquisas, pois elas possuem grande habilidade em efetuar previsão em um conjunto de dados com forte componente de não linearidade. Estudos comparativos mostram que elas são capazes de estabelecer resultados com o mesmo nível de precisão ou, em alguns casos, até maior que os modelos de séries temporais mais usuais.

Oliveira et al. (2010) descreveram a aplicação de RNA na tarefa de previsão da produção de álcool, no Brasil, no ano de 2006. Estes autores fizeram

uma comparação entre os resultados obtidos por meio da rede neural com os obtidos utilizando técnicas de séries temporais, tendo a rede neural obtido melhores resultados.

Angelo et al. (2011) compararam as abordagens que utilizam modelos de Box e Jenkins e redes neurais para a previsão de vendas agregadas do mercado varejista brasileiro. Desempenho semelhante entre os modelos foi obtido e os autores afirmam que o desempenho das redes neurais é comparável ao obtido com o método de Box e Jenkins (1976).

Calôba (2002) apresenta o uso de RNA como ferramenta auxiliar na modelagem de sistemas dinâmicos e de séries temporais.

De acordo com o autor, a modelagem clássica de séries temporais auxilia na escolha de parâmetros para a modelagem de redes neurais artificiais.

Existem algumas dificuldades na utilização de RNA na previsão de séries temporais referentes à determinação da arquitetura ótima da rede, tais como o número de camadas, o número de neurônios, a taxa de aprendizagem e o tempo de treinamento.

Frank, Davey e Hunt (2001) discutem a necessidade de encontrar uma janela de tamanho adequado e a taxa de aprendizagem. Foi demonstrado que a estratégia heurística pode fornecer informações úteis para definir a arquitetura da rede apropriada. Cigizoglu (2005) concluiu que os modelos ARMA têm informações úteis para determinar o número de camadas de entrada de redes neurais.

De acordo com Haykyn (2001), a escolha da quantidade de dados de treinamento deve ser criteriosa, pois, quando se dispõe de poucos dados, pode surgir um problema conhecido como *overfitting*, que pode levar a uma superespecialização da rede. Já o uso de muitos neurônios pode introduzir uma complexidade que o problema não possui, fazendo com que a rede fique muito restrita e propensa a *underfitting*.

Portanto, o tratamento adequado dos dados, bem como uma análise detalhada da série a ser prevista, é um dos fatores fundamentais na obtenção de resultados satisfatórios utilizando RNA.

### **3 MATERIAL E MÉTODOS**

#### **3.1 Material**

Os dados foram fornecidos pela Universidade Federal de Lavras, por meio das contas de energia elétrica emitidas pela Companhia Energética de Minas Gerais (CEMIG). A leitura da conta de luz abrange o período do dia 15 do mês anterior ao dia 15 do mês de pagamento.

A série em estudo consta de 204 observações mensais, abrangendo o período de janeiro de 1995 a dezembro de 2011, referentes ao consumo de energia elétrica registrados em horários de ponta e fora de ponta e demanda mensal de potência registrada em horários de ponta e fora de ponta, como mostrado na Figura 15.

O horário de ponta refere-se a um período de 3 horas consecutivas, das 19h às 22h (durante o horário de verão, das 20h às 23h) e o horário fora de ponta, dentro das 21 horas restantes. A série foi separada em horários distintos, devido à tarifação de energia paga pela UFLA, mencionada no capítulo 2, seção 2.1.3.

Cabe ressaltar que o período de maior consumo de energia nesta instituição é das 13h às 16h, exceto em finais de semanas e feriados, o que não coincide com o horário de ponta adotado pelas concessionárias, das 19h às 22h, quando o maior consumo de energia elétrica ocorre nas residências.

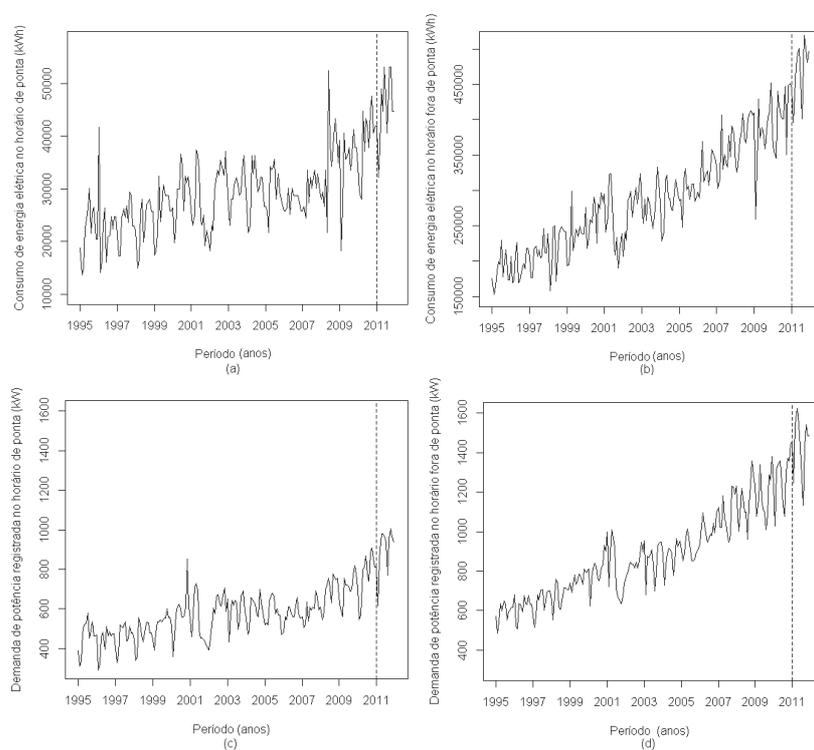


Figura 15 Gráficos referentes aos dados de consumo de energia elétrica (kWh) registrados em horários de ponta (a) e fora de ponta (b) e demanda mensal de potência registrada (kW) em horários de ponta (c) e fora de ponta (d)

### 3.2 Métodos

A estratégia metodológica proposta para o presente trabalho consiste em determinar procedimentos a fim de descrever e analisar os dados de consumo e demanda de energia elétrica da Universidade Federal de Lavras (UFLA).

Inicialmente, foram ajustados modelos por meio da metodologia de Box e Jenkins e, em seguida, modelos de Redes Neurais Artificiais (RNA), para fins

de previsão um passo a frente. As metodologias são apresentadas na próxima seção.

Para a modelagem de Box e Jenkins foi utilizado o programa R Project versão 2.15.1, R DEVELOPMENT CORE TEAM (2011), utilizando as funções do pacote *tseries* e *TSA*. Para implementação das RNA, foi utilizado o software livre SciLab, na versão 5.3.3, com o toolbox ANN, SCILAB ENTERPRISES (2011).

O trabalho foi desenvolvido no ambiente Windows XP 32 bits, em uma máquina com processador Intel® Core™2 Dual-Core E7500 (2,93 GHz), com 2 GB de memória RAM.

### 3.2.1 Metodologia Box e Jenkins

A estratégia metodológica utilizada para ajustar consumo e demanda de energia elétrica da UFLA, de acordo com os modelos de Box e Jenkins, consiste em utilizar as primeiras 192 observações para ajuste dos modelos e as 12 últimas observações servirão como base para comparar as previsões. A seguir, cumpriram-se os seguintes passos:

- a) construção do gráfico da série original para uma cuidadosa inspeção, a fim de caracterizar a presença de algum comportamento relevante, como tendência e/ou sazonalidade ou comportamentos atípicos;
- b) verificar a necessidade de transformação para estabilizar a variância, por meio do gráfico amplitude  $\times$  média;
- c) construção e análise dos gráficos das funções de autocorrelação (*fac*) e autocorrelação parcial (*facp*) da série original, para observar características de tendência e/ou sazonalidade;

- d) verificar a presença de tendência e sazonalidade por meio dos testes de sequência de Wald-Wolfowitz ou do sinal (Cox-Stuart) e o teste de Fisher, respectivamente;
- e) estimar os componentes de tendência e/ou sazonalidade, a fim eliminá-las da série;
- f) fazer diferenças, se necessário, para tornar a série estacionária;
- g) construir gráficos das funções de autocorrelação (fac) e autocorrelação parcial (facp) da série livre das componentes, para estimar as ordens do modelo;
- h) utilizar as estimativas da fac e facp para testar a adequabilidade do modelo, ou seja, se os resíduos constituíam um ruído branco, por meio do teste Box-Pierce;
- i) escolher os modelos que possuem melhor ajuste, por meio dos critérios de Informação Akaike (AIC) e, por fim,
- j) a previsão, a ser realizada com base no modelo verificado nos passos anteriores e o teste de seu desempenho por meio do erro quadrático médio de previsão (EQMP) e do erro percentual absoluto médio de previsão (MAPE).

### **3.2.2 Metodologia RNA**

Para ajustar as séries de consumo e demanda de energia elétrica da UFLA de acordo com a metodologia de Redes Neurais Artificiais (RNA), é necessário pré-estabelecer alguns parâmetros e critérios, tais como:

- a) separar os dados para treinamento e teste;
- b) arquitetura da rede (escolha do número de neurônios na camada de entrada, camada escondida e camada de saída, número de camadas escondidas);
- c) escolha do algoritmo de treinamento;

- d) função de ativação;
- e) taxa de aprendizado ( $\eta$ );
- f) termo momento ( $\alpha$ );
- g) normalizar os dados, se necessário;
- h) estabelecer critérios de parada.

Para utilizar o método de RNA, primeiramente, é preciso separar os dados a serem analisados, em duas partes. Os dados são divididos em treinamento para calibrar a rede e em teste para validar a rede.

Para uma melhor previsão, os dados devem se encontrar em determinados limites, pois valores de uma determinada magnitude podem inviabilizar alguns tipos de modelos. Devido a este fato, os dados de entrada devem ser normalizados, a fim de permanecer dentro dos limites. No presente trabalho, a normalização foi feita dividindo-se os dados de entrada pelo seu valor máximo, abrangendo um intervalo entre 0 e 1.

A função de ativação escolhida é a sigmoide, que fornece a saída neste intervalo. Tal escolha deveu-se ao fato de ser necessária uma previsão com valores positivos, visto que consumo e demanda de energia elétrica sempre terão valores positivos, devido à sua natureza.

Serão utilizadas as redes de múltiplas camadas, cuja arquitetura é formada pela entrada, camada oculta e camada de saída, com janela de tempo, pois se trata de um processamento temporal. O número de neurônios na camada de entrada varia de acordo com o tamanho da janela adotada. Consideram-se uma ou duas camadas ocultas, como na maioria das aplicações encontradas na literatura e a saída gera a previsão um passo a frente. Portanto, tem-se apenas uma saída.

O algoritmo utilizado para treinamento no presente trabalho é o retropropagação. O algoritmo de aprendizagem modifica os pesos da rede de

forma orientada para encontrar na superfície de erro um mínimo global, ou seja, valores para os pesos sinápticos que minimizam os erros da rede.

A escolha da taxa de aprendizado ( $\eta$ ) depende da função a aproximar. De acordo com Haykin (2001), valores muito pequenos de  $\eta$  tornam o treinamento lento, enquanto valores muito grandes podem provocar divergência do processo de treinamento. Foram testados modelos com  $\eta=0,01$  e  $\eta=0,1$ .

A constante momento ( $\alpha$ ) pode ser incluída, ou não, no algoritmo de retropropagação. Seu valor deve ser restrito ao intervalo  $0 < |\alpha| < 1$ . Os valores  $\alpha=0,5$  e  $\alpha=0,9$  foram escolhidos, pois a constante momento pode evitar que o processo de aprendizagem termine em um mínimo local na superfície do erro. O critério de parada escolhido foi definir a tolerância do erro e o número de épocas.

Os desempenhos dos modelos foram testados pelo erro quadrático médio do treinamento (EQM) e a qualidade de previsão, pelos avaliadores citados na próxima seção.

### 3.3 Avaliação do desempenho dos modelos

O poder preditivo dos modelos é verificado por meio do erro quadrático médio de previsão (EQMP), calculado pela média dos quadrados dos erros de previsão, representado pela fórmula

$$EQMP = \frac{1}{N} \sum_{h=1}^N [Y_{t+h} - \hat{Y}_t(h)]^2 \quad (39)$$

e pelo erro percentual absoluto médio de previsão (MAPE), que representa a média percentual da divisão entre erro absoluto de previsão e o valor real

$$MAPE = \frac{100}{N} \sum_{h=1}^N \left| \frac{Y_{t+h} - \hat{Y}_t(h)}{Y_{t+h}} \right| \quad (40)$$

em que  $Y_{t+h}$  é o valor real da série,  $\hat{Y}_t(h)$  é o valor preditivo,  $N$  é o número de observações da previsão e  $h = 1, \dots, N$ .

Estas medidas são úteis para a comparação entre modelos, quando o critério de decisão for escolher aquele que fornecer melhores previsões. Os modelos com menores valores de EQMP e MAPE são os fornecem previsões mais acuradas.

## **4 RESULTADOS E DISCUSSÃO**

Os resultados obtidos foram dispostos da seguinte forma: foram realizados os procedimentos da metodologia de Box e Jenkins para se fazer previsões com as séries de consumo e demanda de energia elétrica da UFLA em horários de ponta e fora de ponta. Em seguida, realizaram-se a análise e as previsões dessas séries com a metodologia de redes neurais artificiais. Finalmente, os resultados obtidos nas análises anteriores foram comparados de acordo com os critérios sugeridos na metodologia.

### **4.1 Modelagem por Box e Jenkins**

#### **4.1.1 Série do consumo de energia elétrica em horário de ponta**

Na Figura 16 é possível verificar o comportamento da série consumo de energia elétrica em horário de ponta da UFLA, no período de janeiro de 1995 a dezembro de 2011. Os valores de 2011 foram utilizados para fins de previsão.

De acordo com o gráfico, vê-se, claramente, que a série é não estacionária, pois esta não se desenvolve no tempo aleatoriamente, ao redor de uma média constante.

Observa-se que o consumo de energia no período de ponta aumentou ao longo do tempo e, dessa forma, existe um indício de tendência. A presença da componente sazonal também pode ser observada. Alguns fatores podem influenciar a oscilação do gráfico. Um deles é o clima, pois, em períodos secos, há mais consumo de energia e, em períodos úmidos, este gasto diminui. Outro fator se refere ao período de férias, geralmente dezembro, janeiro e fevereiro, em que diminui o consumo de energia.

Um suposto efeito de intervenção abrupta e temporária pode ser visto na observação 80, referente ao mês de junho de 2001, que corresponde ao período em que houve uma crise no abastecimento de energia elétrica. Na observação 170, referente ao mês de fevereiro de 2009, o consumo foi baixo, pois se trata do período de recesso escolar. As observações circuladas mostram os pontos de intervenção.

O correlograma da série original (Figura 16) apresenta as funções de autocorrelação (fac) e autocorrelação parcial (facp). Pelo comportamento da fac nota-se que esta não decresce rapidamente para zero, indicando a não estacionariedade da série.

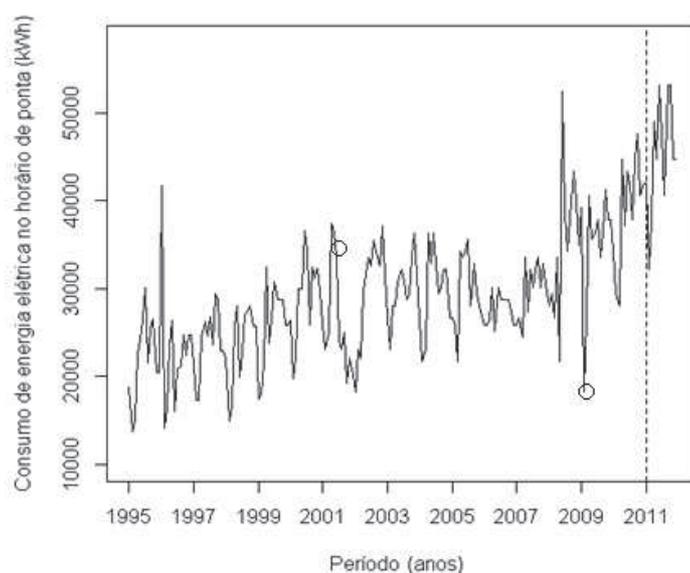


Figura 16 Série do consumo de energia elétrica, em horário de ponta, da UFPA, no período de janeiro de 1995 a dezembro de 2011

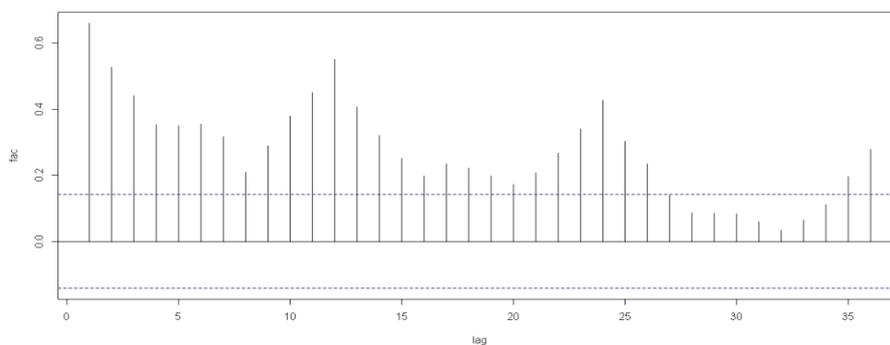


Figura 17 Correlograma da série de consumo de energia elétrica, em horário de ponta, da UFPA, no período de janeiro de 1995 a dezembro de 2010

Para verificar se há necessidade de transformação, a série foi dividida em grupos de 12 observações consecutivas, calculando-se, para cada grupo, a média e a amplitude, como sugerido na metodologia. O comportamento da Figura 18 sugere que uma transformação logarítmica é necessária para estabilizar a variância da série de consumo de energia elétrica em horário de ponta.

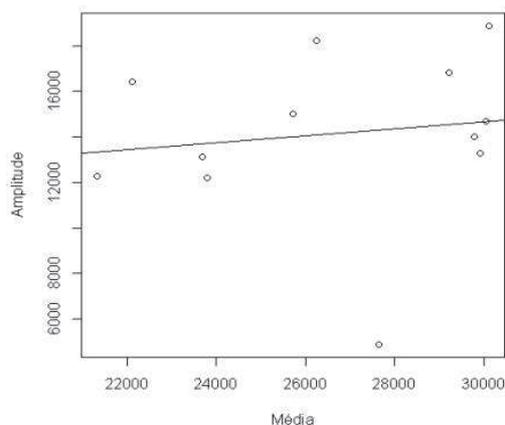


Figura 18 Gráfico amplitude × média da série de consumo de energia elétrica em horário de ponta

Por meio do periodograma da série de consumo de energia no horário de ponta representado na Figura 19 observa-se um maior pico no período de 12 meses. Este efeito sazonal no período 12 pode ser confirmado pelo teste de Fisher; os valores das estatísticas  $g = 6,299$  e  $z_{0,05} = 0,072$  mostram que  $g > z$ , portanto, a série apresenta sazonalidade.

Para verificar a existência de tendência, foi aplicado o teste de sequência Wald-Wolfowitz, cujas hipóteses testadas são  $H_0$ : não existe tendência e  $H_1$ : existe tendência. O valor da estatística foi de  $-7,3806$  com valor  $p = 0,00004$ , em que se pode rejeitar a hipótese nula e concluir a presença da componente tendência.

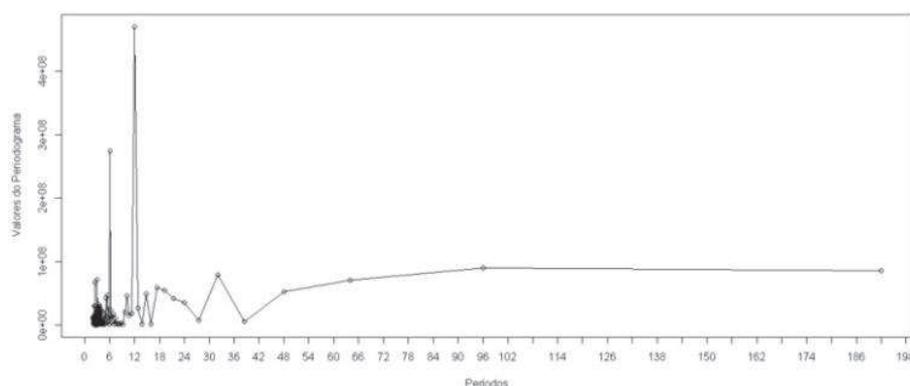


Figura 19 Periodograma da série de consumo de energia elétrica, em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Assim, para tornar a série estacionária, aplicaram-se a primeira diferença e uma diferença sazonal de ordem 12, nos dados transformados. Na Figura 4.5 apresenta-se a série diferenciada, agora estacionária.

As possíveis ordens do modelo a ser ajustado são sugeridas pela  $fac$  e  $facp$  da série diferenciada, representadas na Figura 4.6. A análise do

correlograma revela a presença de correlação alta no lag 12, sugerindo um modelo SARIMA.

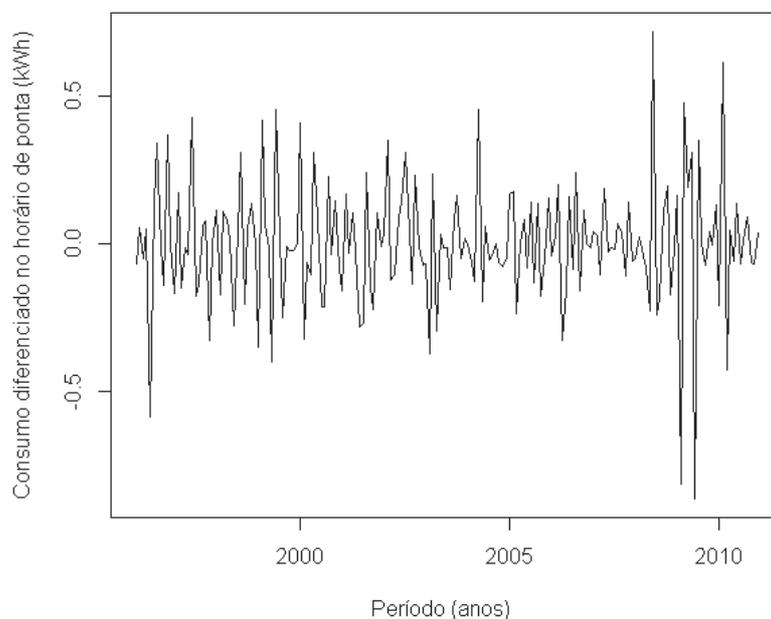


Figura 20 Série diferenciada do consumo de energia elétrica, em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

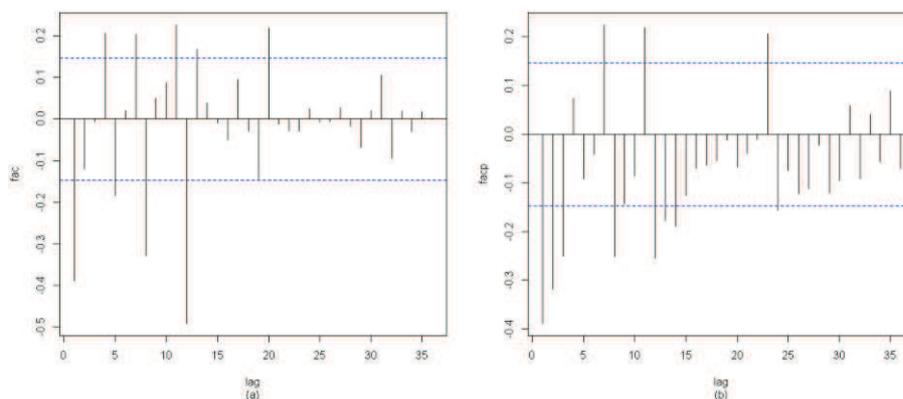


Figura 21 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada do consumo de energia elétrica, em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Os modelos sugeridos foram:

SARIMA(1,1,0)(0,1,1)<sub>12</sub>

$$Z_t = \frac{(1 - \theta_1 B)(1 - \theta_1 B^{12})}{(1 - \phi_1 B)(1 - B^{12})(1 - B)} a_t$$

SARIMA(1,1,1)(0,1,1)<sub>12</sub> com intervenção

$$Z_t = \frac{\omega_1}{1 - \delta_1 B} X_{1,t} + \frac{\omega_2}{1 - \delta_2 B} X_{2,t} + \frac{(1 - \theta_1 B)(1 - \theta_1 B^{12})}{(1 - \phi_1 B)(1 - B^{12})(1 - B)} a_t$$

em que as variáveis de intervenção  $X_{1,t}$  e  $X_{2,t}$  são

$$X_{1,t} = \begin{cases} 0, & t = 80 \\ 1, & t \neq 80 \end{cases}$$

$$X_{2,t} = \begin{cases} 0, & t = 170 \\ 1, & t \neq 170 \end{cases}$$

Na Tabela 2 são apresentadas as estimativas dos modelos sugeridos.

Tabela 2 Estimativas dos parâmetros dos modelos sugeridos para a série de consumo de energia elétrica, em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro padrão
SARIMA(1,1,1)(0,1,1) <sub>12</sub>	$\phi_1$	0,34	0,12
	$\theta_1$	0,89	0,08
	$\theta_1$	0,75	0,05
SARIMA(1,1,1)(0,1,1) <sub>12</sub> - CI	$\theta_1$	0,77	0,04
	$\theta_1$	0,69	0,05
	$\omega_1$	-0,47	0,09
	$\delta_1$	0,89	0,04
	$\omega_2$	-0,15	0,04
	$\delta_2$	0,97	0,00

Pelo teste Box-Pierce, tem-se que  $Q_{24}=24,42$  e  $\chi^2_{(21)} = 32,67$ , para o modelo SARIMA(1,1,1)(0,1,1)<sub>12</sub> e  $Q_{24}=18,91$  e  $\chi^2_{(22)} = 33,92$ , para SARIMA(0,1,1)(0,1,1)<sub>12</sub>-CI (com intervenção), o que confirma que os resíduos de ambos os modelos são ruído branco.

Os dois modelos se adequam bem aos dados. Para a escolha do melhor modelo utilizou-se o critério de AIC e, para fins de previsão, o EQMP e o MAPE. Na Tabela 3 apresentam-se os resultados.

Tabela 3 Estimativas dos valores do critério de AIC, EQMP e MAPE dos modelos sugeridos.

<b>Modelo</b>	<b>AIC</b>	<b>EQMP</b>	<b>MAPE</b>
<b>SARIMA(1,1,1)(0,1,1)<sub>12</sub></b>	-201,02	22970351,00	8,75%
<b>SARIMA(1,1,1)(0,1,1)<sub>12</sub>- CI</b>	-205,25	12420867,24	6,50%

De acordo com os resultados apresentados, o modelo com intervenção apresentou menor AIC e menores EQMP e MAPE. Substituindo-se os valores das estimativas apresentadas na Tabela 1, tem-se a seguinte equação para o modelo SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI:

$$Z_t = \frac{-0,47}{1 - 0,89B} X_{1,t} + \frac{-0,15}{1 - 0,97B} X_{2,t} + \frac{(1 - 0,77B)(1 - 0,69B^{12})}{(1 - B^{12})(1 - B)} a_t$$

As observações de janeiro a dezembro de 2011 foram utilizadas para a verificação da previsão. Os valores reais e estimados são apresentados na Tabela

4 e a comparação entre os valores reais e os previstos pelo modelo SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI está representada na Figura 22.

Tabela 4 Valores reais e previstos do consumo de energia elétrica, em horário de ponta, para o ano de 2011 do modelo SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI.

Período	Valor real	Valor previsto
jan/11	42000	38617,31
fev/11	32200	31584,06
mar/11	37800	32994,99
abr/11	49000	47201,51
mai/11	44800	39673,64
jun/11	53200	48852,89
jul/11	47600	46239,97
ago/11	40600	42412,46
set/11	53200	47023,29
out/11	53200	48711,27
nov/11	44800	45670,28
dez/11	44800	43560,53

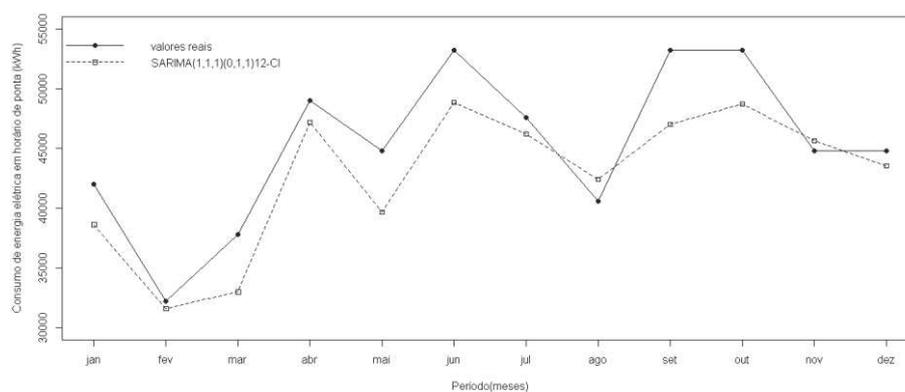


Figura 22 Comparação entre os valores reais e os previstos pelo modelo SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI

#### 4.1.2 Série do consumo de energia elétrica em horário fora de ponta

A série original mensal do consumo de energia elétrica em horário fora de ponta, no período de janeiro de 1995 a dezembro de 2011, está representada na Figura 23. A análise para ajuste de modelos foi feita de janeiro de 1995 a dezembro de 2010, tendo os valores do ano de 2011 sido utilizados para fins de previsão.

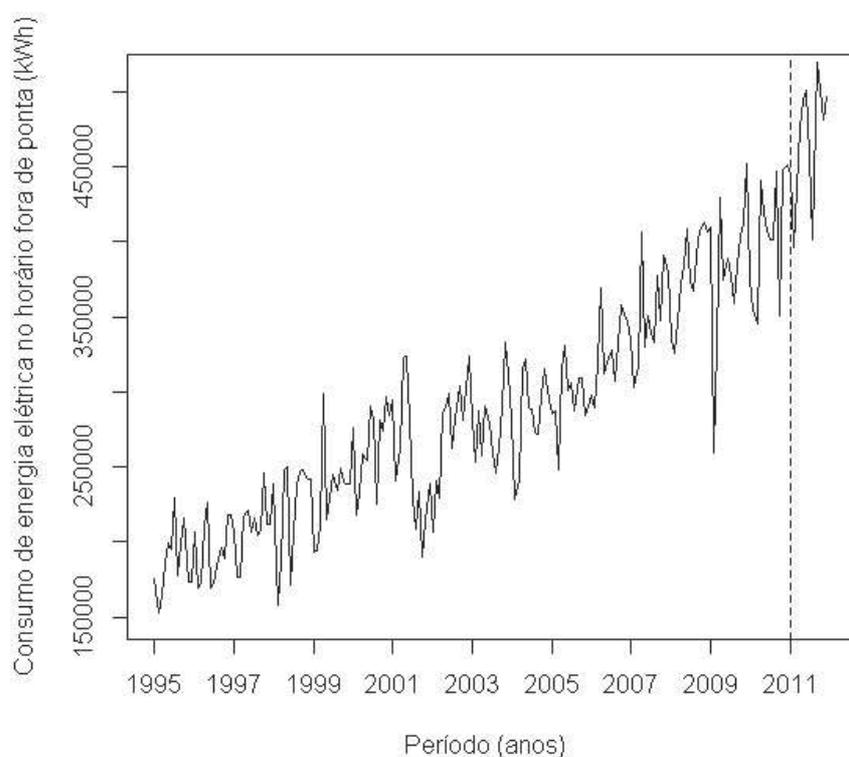


Figura 23 Série do consumo de energia elétrica, em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2011

Por meio de uma análise visual, verifica-se uma tendência crescente na série. Os procedimentos para tornar a série estacionária foram realizados de forma análoga ao da série de consumo de energia elétrica no horário de ponta e estão disponíveis no Apêndice. Na Figura 24 tem-se a série estacionária.

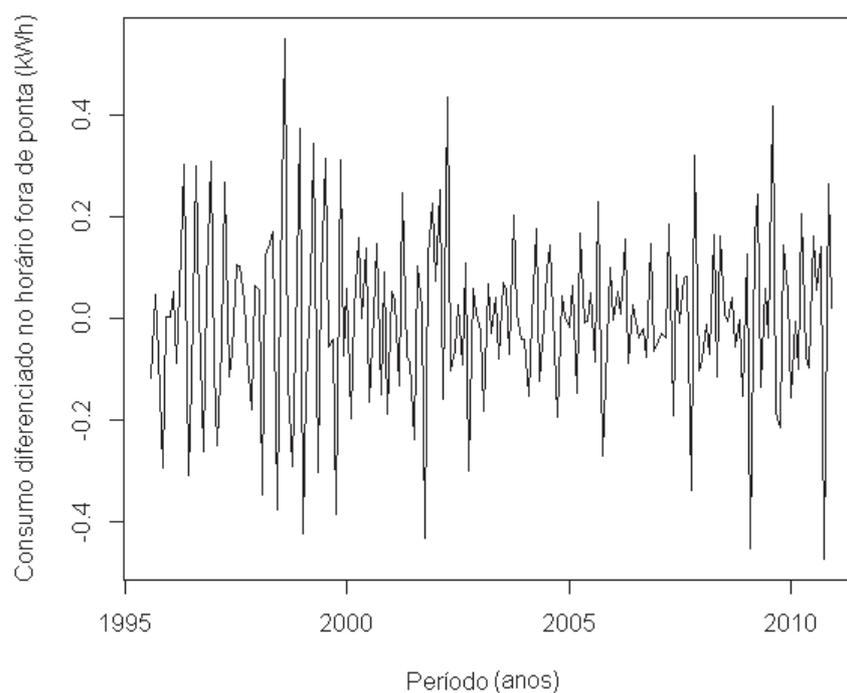


Figura 24 Série diferenciada de consumo de energia elétrica, em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Para a identificação do modelo SARIMA e a estimação dos seus parâmetros é preciso analisar as funções de autocorrelação e autocorrelação parcial da série estacionária, apresentadas na Figura 25. Os picos presentes nas defasagens 1, 2 e 6 da função de autocorrelação indicam a presença de termos de média móvel e média móvel sazonal. A função de autocorrelação parcial é

significativa para defasagem de ordem 3, indicando, possivelmente, um processo autorregressivo de ordem elevada.

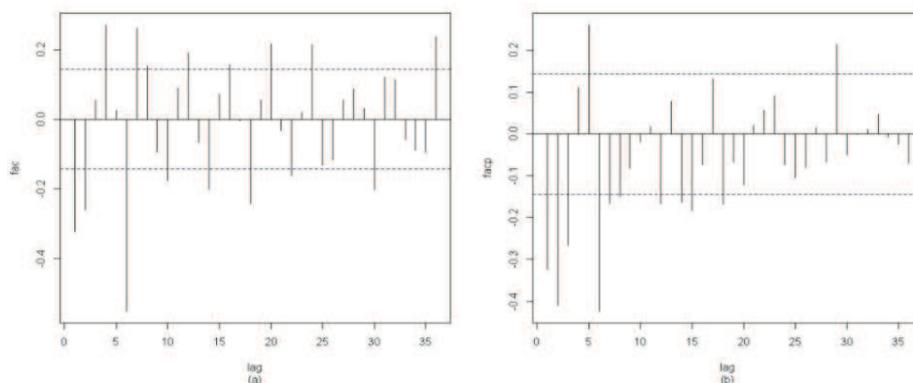


Figura 25 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de consumo de energia elétrica, em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Os modelos sugeridos foram  $SARIMA(2,1,0)(0,1,1)_6$  e  $SARIMA(1,1,0)(0,1,1)_6$ . Na Tabela 5 apresentam-se as estimativas dos parâmetros dos modelos.

Tabela 5 Estimativas dos parâmetros dos modelos sugeridos para a série de consumo de energia elétrica, em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro padrão
$SARIMA(2,1,0)(0,1,1)_6$	$\phi_1$	0,41	0,06
	$\phi_2$	-0,47	0,08
	$\theta_1$	0,58	0,09
$SARIMA(1,1,0)(0,1,1)_6$	$\phi_1$	-0,33	0,069
	$\theta_1$	0,87	0,032

Pelo teste Box-Pierce, tem-se que  $Q_{24}=23,12 < \chi^2_{(21)} = 32,67$ , para o modelo SARIMA(2,1,0)(0,1,1)<sub>6</sub> e  $Q_{24}=21,54 < \chi^2_{(22)} = 33,92$ , para SARIMA(1,1,0)(0,1,1)<sub>6</sub>, comprovando, portanto, que os resíduos de ambos os modelos são ruído branco.

Para a escolha do melhor modelo, utilizou-se o critério de AIC e, para fins de previsão, EQMP e MAPE. Na Tabela 6 apresentam-se os resultados.

Tabela 6 Estimativa dos valores dos critérios de AIC e BIC e o EQMP dos modelos sugeridos.

<b>Modelo</b>	<b>AIC</b>	<b>EQMP</b>	<b>MAPE</b>
<b>SARIMA(2,1,0)(0,1,1)<sub>6</sub></b>	-205,25	1707265154,00	7,28%
<b>SARIMA(1,1,0)(0,1,1)<sub>6</sub></b>	-350,13	813226848,7	4,16%

De acordo com os resultados, o modelo SARIMA(1,1,0)(0,1,1)<sub>6</sub> forneceu melhores previsões, com AIC, EQMP e MAPE menores. A equação do modelo é descrita por

SARIMA(1,1,0)(0,1,1)<sub>6</sub>

$$Z_t = \frac{(1 - \theta_1 B^6)}{(1 - \phi_1 B)(1 - B^6)(1 - B)} a_t$$

O modelo com as estimativas dos parâmetros é dado por

$$Z_t = \frac{(1 - 0,87B^6)}{(1 + 0,33B)(1 - B^6)(1 - B)} a_t$$

As previsões realizadas no período de janeiro a dezembro de 2011 estão apresentadas na Tabela 7 e a comparação entre os valores reais e os previstos pelo modelo SARIMA(1,1,0)(0,1,1)<sub>6</sub>, na Figura 26.

Tabela 7 Valores reais e valores preditos da série de consumo de energia elétrica, em horário fora de ponta, da UFLA, para o modelo SARIMA(1,1,0)(0,1,1)<sub>6</sub>.

Período	Valor real	SARIMA
jan/11	445200	415596,9
fev/11	396200	383172,0
mar/11	435400	414081,3
abr/11	476000	439762,4
mai/11	494200	448845,5
jun/11	501200	453439,8
jul/11	457800	427694,9
ago/11	401800	394326,0
set/11	519400	426135,2
out/11	501200	452563,8
nov/11	481600	461911,3
dez/11	497000	466639,4

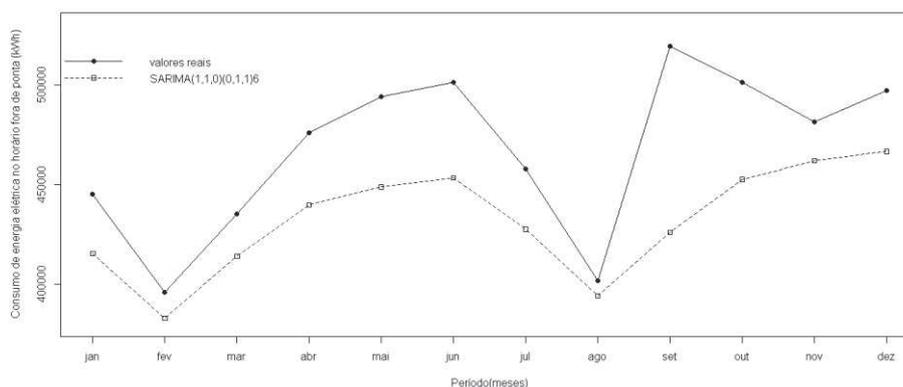


Figura 26 Comparação entre os valores reais e os previstos pelo modelo SARIMA(1,1,0)(0,1,1)<sub>6</sub>

### 4.1.3 Série da demanda de energia elétrica em horário de ponta

A Figura 27 é a representação gráfica da série de demanda de energia elétrica registrada em horário de ponta. Pela análise visual, verifica-se a presença de tendência e indícios de sazonalidade. Para tornar a série estacionária, utilizou-se o mesmo procedimento realizado na série do consumo de energia no horário de ponta e pode ser visto no Apêndice.

Na Figura 28 apresenta-se a série estacionária e, na Figura 29, o correlograma da série estacionária.

As ordens do modelo são identificadas pelas funções de autocorrelação e autocorrelação parcial da série diferenciada presentes na Figura 29.

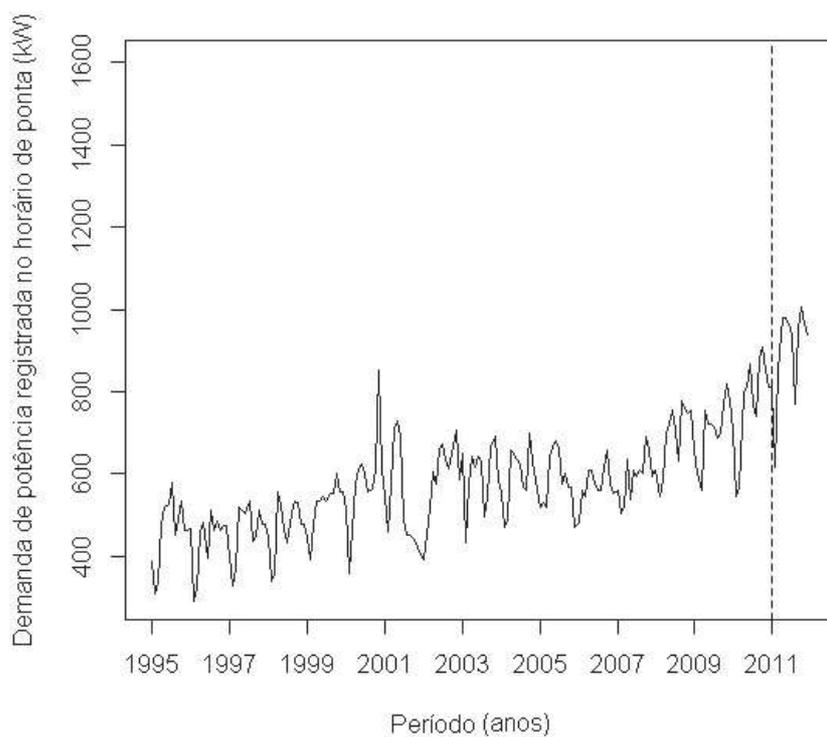


Figura 27 Série de demanda de energia elétrica registrada em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2011

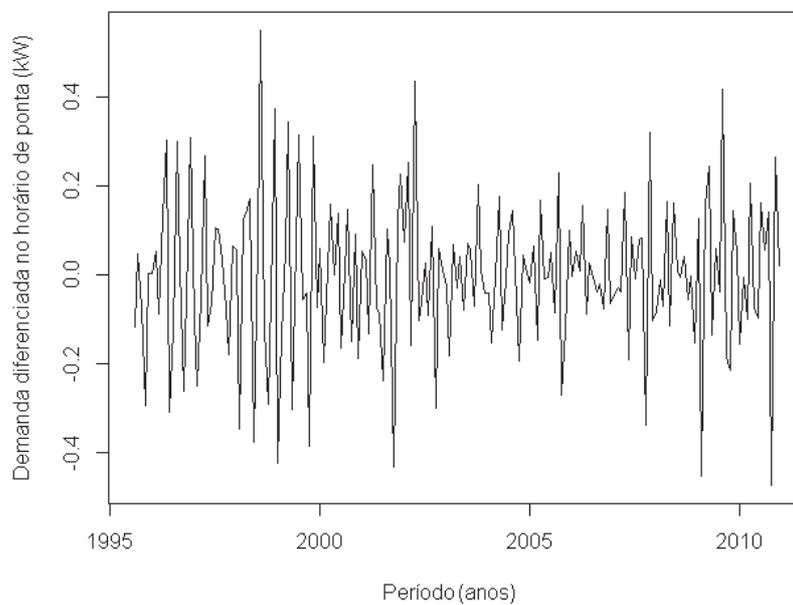


Figura 28 Série diferenciada da demanda registrada em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

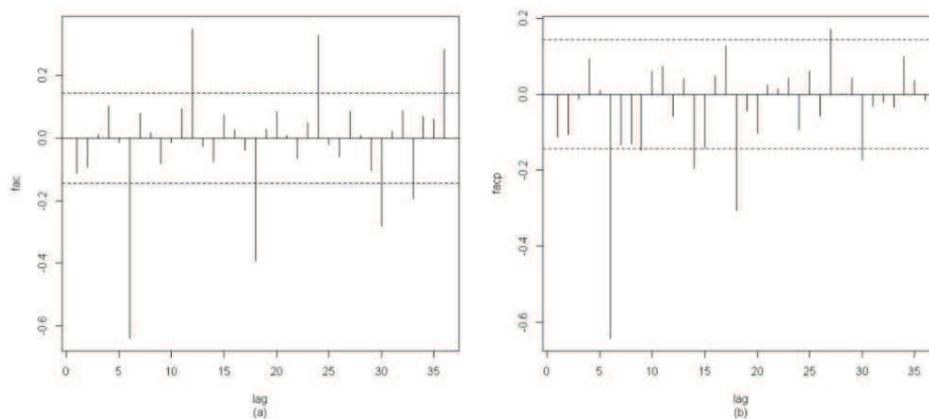


Figura 29 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de demanda de energia elétrica em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Os modelos sugeridos, a princípio, são o SARIMA(0,1,0)(3,1,2)<sub>6</sub> e o SARIMA(0,1,1)(3,1,1)<sub>6</sub>. As estimativas dos parâmetros dos modelos se encontram na Tabela 8.

Tabela 8 Estimativas dos parâmetros dos modelos sugeridos para a série de demanda de energia elétrica, em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro padrão
<b>SARIMA(0,1,0)(3,1,2)<sub>6</sub></b>	$\Phi_1$	0,9202	0,1277
	$\Phi_2$	-0,7882	0,1798
	$\Phi_3$	0,3327	0,1012
	$\theta_1$	0,6718	0,1057
	$\theta_2$	-0,8625	0,1231
<b>SARIMA(0,1,1)(3,1,1)<sub>6</sub></b>	$\theta_1$	-0,3432	0,0854
	$\Phi_1$	-0,1466	0,0841
	$\Phi_2$	0,1540	0,0890
	$\Phi_3$	-0,3043	0,0792
	$\theta_1$	-0,8966	0,0735

O teste de Box-Pierce foi aplicado e obtiveram-se as estatísticas  $Q_{24}=22,32 < \chi^2_{(19)} = 30,14$  e  $Q_{24}=28,54 < \chi^2_{(19)} = 30,14$ , respectivamente, o que confirma a hipótese de os resíduos serem um processo de ruído branco. A escolha do melhor modelo se deu pelo critério de AIC e para escolha do modelo de previsão pelo EQMP e MAPE, cujos valores estão na Tabela 9.

Tabela 9 Estimativa dos valores dos critérios de AIC, EQMP e MAPE dos modelos sugeridos.

<b>Modelo</b>	<b>AIC</b>	<b>EQMP</b>	<b>MAPE</b>
<b>SARIMA(0,1,0)(3,1,2)<sub>6</sub></b>	-286,65	94667,44	10,99%
<b>SARIMA(0,1,1)(3,1,1)<sub>6</sub></b>	-295,51	76098,82	8,02%

O modelo escolhido para realizar previsões é o SARIMA(0,1,1)(3,1,1)<sub>6</sub>, pois apresenta menores AIC, EQMP e MAPE. O modelo tem a seguinte forma:

$$Z_t = \frac{(1 - \theta_1 B)(1 - \theta_1 B^6)}{(1 - \Phi_1 B^6)(1 - \Phi_2 B^{12})(1 - \Phi_3 B^{18})(1 - B^6)(1 - B)} a_t$$

Substituindo-se os valores encontrados na Tabela 4.7, tem-se

$$Z_t = \frac{(1 + 0,3432B)(1 + 0,8966B^6)}{(1 + 0,1466B^6)(1 - 0,1540B^{12})(1 + 0,3043B^{18})(1 - B^6)(1 - B)} a_t$$

Os valores reais, as estimativas dos valores preditos, para o período de janeiro a dezembro de 2011, do modelo SARIMA(0,1,1)(3,1,1)<sub>6</sub>, encontram-se na Tabela 10 e na Figura 30 está representada a comparação entre os valores reais e previstos.

Tabela 10 Valores reais e valores preditos, no período de janeiro a dezembro de 2011, da demanda registrada no horário de ponta, da UFLA, para o modelo SARIMA(0,1,1)(3,1,1)<sub>6</sub>.

Período	Valor real	Valor predito
jan/11	812	774,2821
fev/11	616	669,7238
mar/11	854	707,0360
abr/11	980	875,9105
mai/11	980	846,7495
jun/11	966	848,8051
jul/11	938	812,2513
ago/11	770	789,3004
set/11	952	861,1143
out/11	1006	918,5172
nov/11	966	884,9885
dez/11	938	833,7137

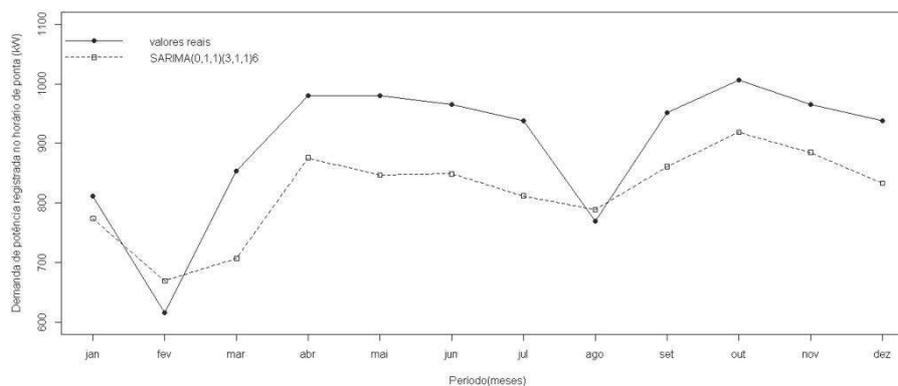


Figura 30 Comparação entre os valores reais e os previstos pelo modelo SARIMA(0,1,1)(3,1,1)<sub>6</sub>

#### 4.1.4 Série da demanda de energia elétrica em horário fora de ponta

Na Figura 31 é mostrado o gráfico da série mensal de demanda registrada em horário fora de ponta da UFLA. É possível visualizar a presença de tendência crescente na série e uma possível sazonalidade.

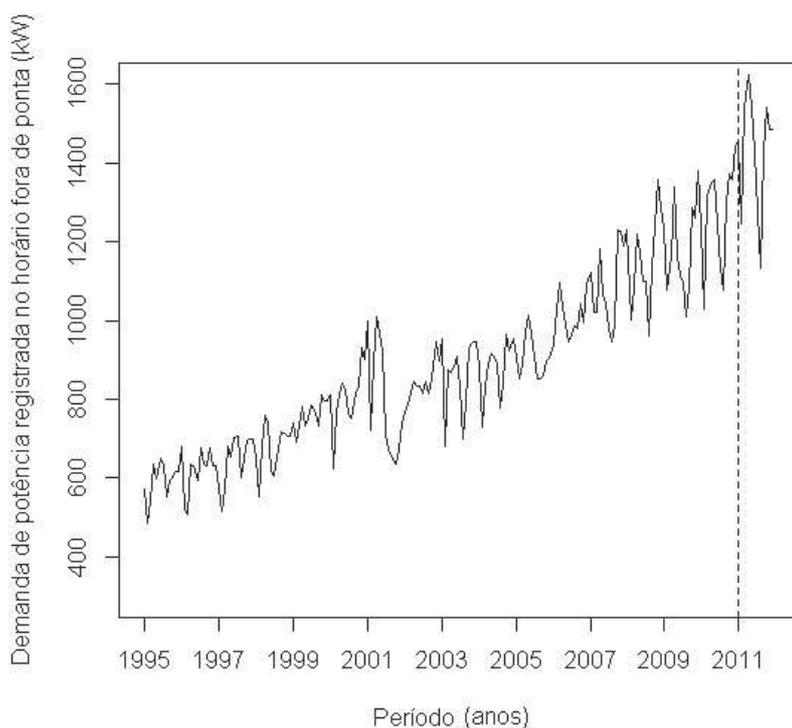


Figura 31 Série de demanda de energia elétrica registrada em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2011

Na Figura 32 é apresentada a série diferenciada de acordo com os procedimentos realizados na série de consumo de energia elétrica no horário de ponta. As funções de autocorrelação e autocorrelação parcial da série diferenciada são apresentadas na Figura 33.

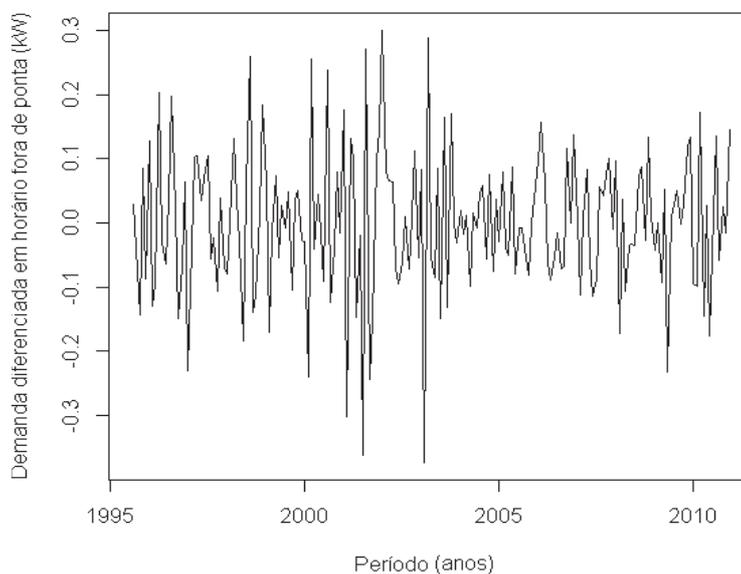


Figura 32 Série diferenciada de demanda de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

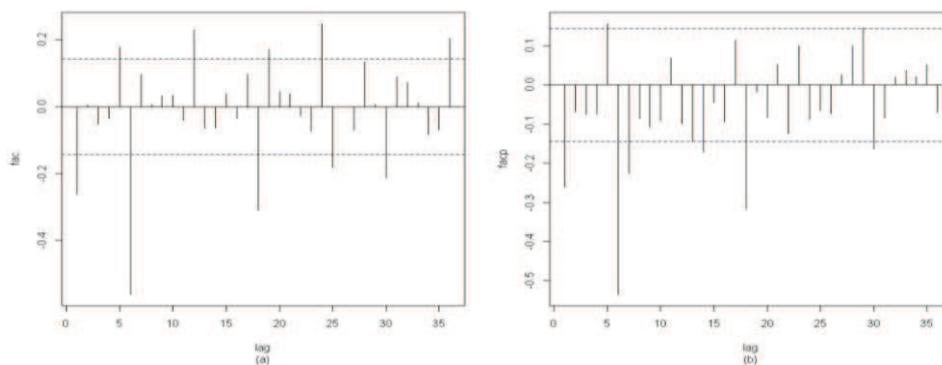


Figura 33 Funções de autocorrelação (a) e autocorrelação parcial (b) da série diferenciada de demanda de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Os modelos sugeridos são SARIMA(1,1,1)(1,1,2)<sub>6</sub> e SARIMA(1,1,1)(0,1,1)<sub>6</sub>. As estimativas dos parâmetros dos modelos se encontram na Tabela 11.

Tabela 11 Estimativas dos parâmetros dos modelos sugeridos para a série de demanda registrada em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010.

Modelo	Parâmetro	Estimativa	Erro padrão
SARIMA(1,1,1)(1,1,2) <sub>6</sub>	$\phi_1$	0,5093	0,0750
	$\theta_1$	0,9520	0,0363
	$\Phi_1$	0,9441	0,0561
	$\theta_1$	0,1307	0,1057
	$\theta_2$	-0,8693	0,1007
SARIMA(1,1,1)(0,1,1) <sub>6</sub>	$\phi_1$	-0,5056	0,0735
	$\theta_1$	0,9585	0,0356
	$\theta_1$	1,0000	0,0576

O teste de Box-Pierce foi aplicado, obtendo-se o p-valor = 0,82, para o primeiro modelo e p-valor = 0,60, para o segundo. Logo, em ambos os modelos, os resíduos são ruído branco. A escolha do melhor modelo se deu pelo critério de AIC e, para escolha do modelo de previsão, por EQMP e MAPE, cujos valores estão na Tabela 12.

Tabela 12 Estimativa dos valores dos critérios de AIC, EQMP e MAPE dos modelos.

Modelo	AIC	EQMP	MAPE
SARIMA(1,1,1)(1,1,2) <sub>6</sub>	-437,86	195587,7	7,83%
SARIMA(1,1,1)(0,1,1) <sub>6</sub>	-433,08	17115,32	7,49%

O modelo escolhido para realizar previsões é o SARIMA(1,1,1)(0,1,1)<sub>6</sub>, pois apresenta menor EQMP e MAPE, apesar de possuir maior AIC. O modelo possui a seguinte forma:

$$Z_t = \frac{(1 - \theta_1 B)(1 - \theta_1 B^6)at}{(1 - \phi_1 B)(1 - B^6)(1 - B)}$$

Substituindo-se os valores encontrados na Tabela 4.10, tem-se

$$Z_t = \frac{(1 - 0,9585B)(1 - 1,00B^6)at}{(1 + 0,5056 B)(1 - B^6)(1 - B)}$$

Os valores reais e as estimativas dos valores preditos, para o período de janeiro a dezembro de 2011, encontram-se na Tabela 13. Na Figura 34 apresenta-se a comparação entre os valores reais e previstos.

Tabela 13 Valores reais e valores preditos do modelo SARIMA(1,1,1)(0,1,1)<sub>6</sub>, no período de janeiro a dezembro de 2011, da demanda registrada no horário fora de ponta, na UFLA.

<b>Período</b>	<b>Valor original</b>	<b>Valor predito</b>
jan/11	1456	1385,444
fev/11	1246	1200,841
mar/11	1540	1301,862
abr/11	1624	1398,454
mai/11	1540	1365,643
jun/11	1428	1325,530
jul/11	1274	1296,991
ago/11	1134	1189,234
set/11	1428	1284,921
out/11	1540	1404,978
nov/11	1484	1413,021
dez/11	1484	1433,351

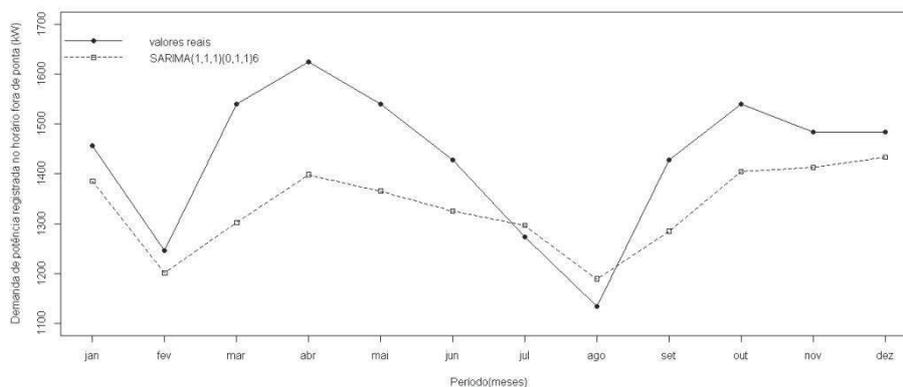


Figura 34 Comparação entre os valores reais e os previstos pelo modelo SARIMA(1,1,1)(0,1,1)<sub>6</sub>

#### 4.2 Modelagem por redes neurais artificiais

Nesta seção é utilizada a metodologia das Redes Neurais Artificiais (RNA) para prever as séries em estudo. A rede utilizada é do tipo perceptron de múltiplas camadas para a predição de um passo a frente. O algoritmo para treinamento é o de retropropagação com aprendizagem supervisionada.

A arquitetura da rede foi representada de acordo com o número de neurônios em cada camada por RNA (a, b, c, d), onde a representa o número de neurônios da camada de entrada, b da primeira camada escondida, c da segunda camada escondida e d da camada de saída. Diversas arquiteturas foram simuladas, alternando a quantidade de camadas escondidas e o número de neurônios utilizados.

Cada série analisada (consumo de energia elétrica no horário de ponta e no horário fora de ponta e demanda registrada de energia elétrica no horário de ponta e no horário fora de ponta) tem 204 observações, tendo as 192 primeiras sido separados para o treinamento, correspondendo ao intervalo de janeiro de

1995 a dezembro de 2010 e as observações restantes foram reservadas para teste. O conjunto de teste é utilizado após o treinamento para validar a rede e verificar se ela aprendeu.

Como critério de parada de treinamento da rede foi definida a tolerância do erro em 0,0001 ou 10.000 iterações. O poder preditivo dos modelos encontrados foi verificado por meio de EQMP e MAPE.

Além do tipo de rede e treinamento, alguns parâmetros devem ser pré-definidos, como número de neurônios na camada de entrada, nas camadas escondidas e na saída; função de ativação; taxa de aprendizado; termo de momento e quantidade de épocas para treinamento e teste. A escolha desses valores é feita empiricamente, até que se encontre uma melhor arquitetura, a fim de minimizar o erro.

Os atrasos de série temporal foram variados de 6 a 36, em razão da presença de sazonalidade de 6 e 12 meses. Assim, cada valor de entrada corresponde a um mês. A saída tem apenas um neurônio, pois se trata de um processamento temporal e a saída representa a previsão um passo a frente (um mês adiante). Mas, após inúmeros testes, foi verificado que arquiteturas com 12 e 24 neurônios na camada de entrada apresentaram melhor desempenho e, portanto, testaram-se três tipos de arquitetura RNA(12,12,0,1), RNA(12,24,12,1) e RNA(24,48,12,1) para cada série.

Para todas as séries em estudo foi utilizada a mesma composição de rede, em que foram fixados alguns parâmetros variáveis, apresentados na Tabela 14.

Tabela 14 Parâmetros pré-definidos e fixos da rede neural para a modelagem das séries.

Parâmetros	Valores
Função de ativação para todos os neurônios	sigmoide
Taxa de aprendizagem	0,01 e 0,1
Termo momento	0,5 e 0,9
Épocas	10000
Tolerância do erro	0,001

#### 4.2.1 Série do consumo de energia elétrica em horário de ponta

Para o consumo de energia elétrica no horário de ponta, na Tabela 15 são apresentadas as arquiteturas testadas, o erro quadrático médio de treinamento, bem como o tempo gasto para treinar a rede. Dentre os modelos gerados, foi escolhido aquele com menor erro quadrático médio (EQM). Na Figura 35 é apresentada a evolução do EQM de treinamento em função das arquiteturas testadas, o que facilita a visualização do modelo a ser escolhido.

Tabela 15 Configuração do perceptron múltiplas camadas para a série de consumo de energia no horário de ponta.

Arquitetura	Taxa de aprendizado	Momento	Tempo segundos	EQM
RNA(12,12,0,1)	0,01	0,5	575,078	0,0182091
	0,01	0,9	582,453	0,0161871
	0,1	0,5	577,063	0,0121518
	0,1	0,9	562,688	0,0116967
RNA(12,24,12,1)	0,01	0,5	850,969	0,0161698
	0,01	0,9	863,75	0,0126338
	0,1	0,5	838,703	0,0112031
	0,1	0,9	835,359	0,0265497
RNA(24,48,12,1)	0,01	0,5	1048,313	0,0715607
	0,01	0,9	1039,594	0,0661332
	0,1	0,5	1055,14	0,1079610
	0,1	0,9	1088,203	0,0898517

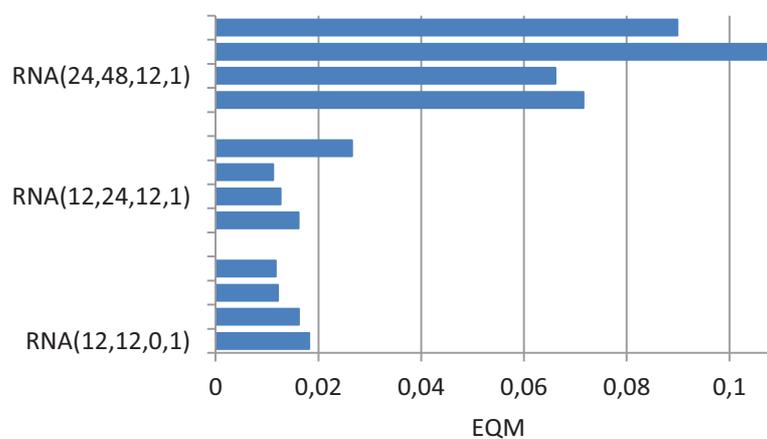


Figura 35 Valores do erro quadrático médio de treinamento das arquiteturas testadas para o consumo de energia elétrica em horário de ponta

A arquitetura RNA(12,12,0,1) apresentou um tempo de treinamento menor que as demais, pois possui uma camada de neurônios a menos. A RNA(12,24,12,1) com taxa de aprendizado ( $\eta=0,1$ ) e termo momento ( $\alpha=0,5$ ) apresentou o menor EQM. Na Figura 36 é possível verificar o erro quadrático médio em função do número de épocas, durante o treinamento.

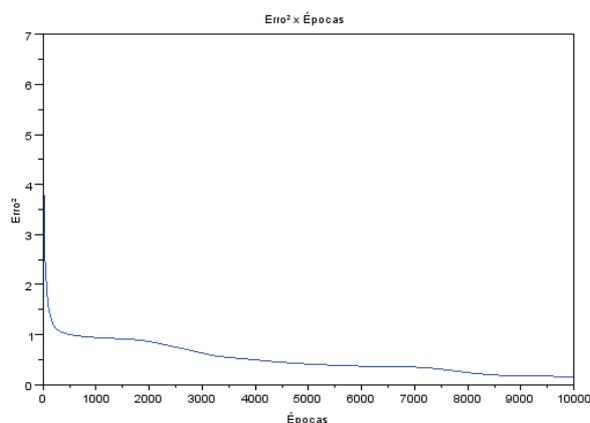


Figura 36 Gráfico do número de épocas em função do erro quadrático médio do modelo com menor erro de previsão RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,5$

Após a escolha do modelo com melhor desempenho preditivo, apresentam-se, na Tabela 16, os resultados da previsão de uma passo a frente da RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,5$ , referente ao período de janeiro a dezembro de 2011. Na Figura 37 ilustram-se os valores reais comparados com a saída da rede. Com os valores originais e os valores previstos é possível calcular os erros de previsão e apresentar os valores dos EQMP e MAPE presentes na Tabela 17 para o modelo de redes neurais do consumo de energia elétrica, em horário de ponta, da UFLA.

Tabela 16 Valores previstos da série do consumo de energia elétrica, em horário de ponta, da UFLA, com o modelo RNA(12,24,12,1)  $\eta=0,1$  e  $\alpha=0,5$ .

Período	Valor original	Valor previsto
jan/11	42000	39298,8
fev/11	32200	30939,7
mar/11	37800	35452,9
abr/11	49000	47381,6
mai/11	44800	40743,6
jun/11	53200	55927,2
jul/11	47600	49144
ago/11	40600	39231,5
set/11	53200	50452,3
out/11	53200	43117,8
nov/11	44800	42496,7
dez/11	44800	41770,2

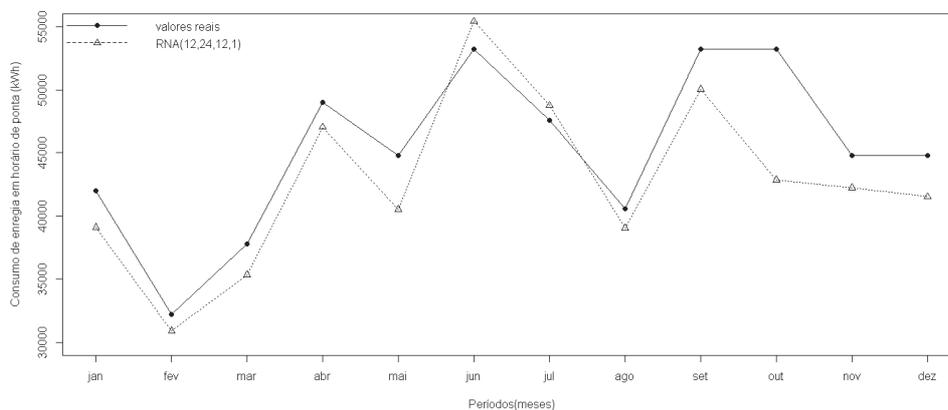


Figura 37 Comparação entre os valores reais e os previstos pelo modelo RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,5$

Tabela 17 Valores do EQMP e MAPE do modelo com menor erro de treinamento  $\eta=0,1$  e  $\alpha=0,5$ .

<b>Modelo</b>	<b>EQMP</b>	<b>MAPE</b>
<b>RNA(12,24,12,1) <math>\eta=0,1</math> e <math>\alpha=0,5</math></b>	14070594,62	6,38%

De modo análogo à modelagem realizada nesta seção, serão apresentados os resultados do desempenho dos modelos testados, o gráfico do EQM de treinamento em função das arquiteturas testadas, a evolução do erro de treinamento com o número de iterações do modelo escolhido como melhor preditor, os valores de previsão, bem como o EQMP e o MAPE do modelo de redes neurais para as séries de consumo de energia elétrica no horário fora de ponta (seção 4.2.2), demanda de potência registrada no horário de ponta (seção 4.2.3) e demanda de potência registrada no horário fora de ponta (seção 4.2.4).

#### **4.2.2 Série do consumo de energia elétrica em horário fora de ponta**

Na Tabela 18 apresenta-se o desempenho dos modelos testados em relação aos parâmetros pré-definidos.

Tabela 18 Configuração do perceptron múltiplas camadas para a série de consumo de energia no horário fora de ponta.

Arquitetura	Taxa de aprendizado	Momento	Tempo (segundos)	EQM
RNA(12,12,0,1)	0,01	0,5	595,313	0,0072774
	0,01	0,9	623,594	0,0056040
	0,1	0,5	758,344	0,0064860
	0,1	0,9	741,969	0,0087365
RNA(12,24,24,1)	0,01	0,5	895,047	0,0057681
	0,01	0,9	894,937	0,0067935
	0,1	0,5	898,906	0,0087903
	0,1	0,9	890,359	0,0264402
RNA(24,48,12,1)	0,01	0,5	1130,391	0,0078916
	0,01	0,9	1129,109	0,0095167
	0,1	0,5	1134,813	0,0116589
	0,1	0,9	1137,844	0,0280467

Na Figura 38 apresenta-se o gráfico do erro quadrático médio do treinamento e observa-se que o modelo RNA(12,12,0,1), com  $\eta=0,01$  e  $\alpha=0,9$ , apresenta melhor desempenho, com o menor EQM.

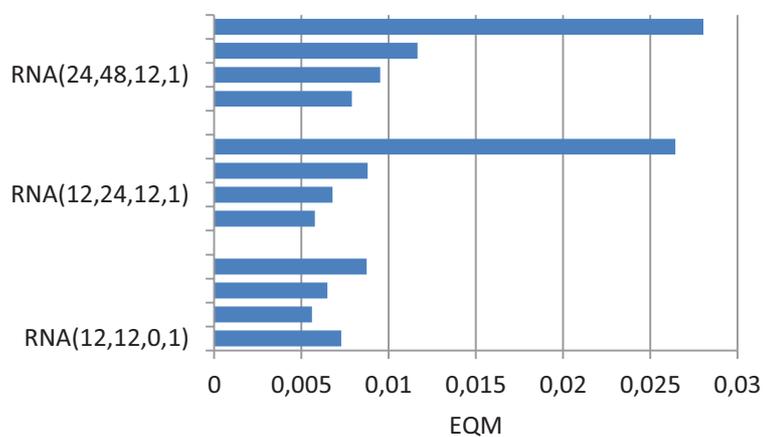


Figura 38 Valores do erro quadrático médio de treinamento das arquiteturas testadas, para o consumo de energia elétrica em horário fora de ponta

A evolução do erro de treinamento com o número de iterações é mostrado na Figura 39, que comprova a eficácia da arquitetura escolhida.

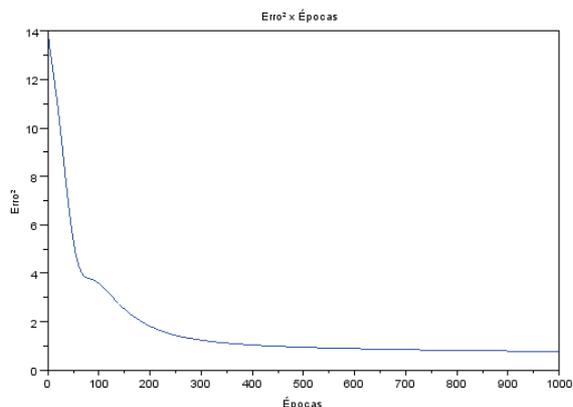


Figura 39 Gráfico do número de épocas, em função do erro quadrático médio do modelo RNA(12,12,1), com  $\eta=0,1$  e  $\alpha=0,5$

No intuito de verificar a capacidade de previsão e adequação a valores futuros, foi comparada a saída da rede ao conjunto dos valores de janeiro a dezembro de 2011. Esta diferença produziu um erro e com ele foram calculados o erro quadrático médio de previsão e o erro percentual médio de previsão. Para o modelo sugerido, tem-se que EQMP=1118942341,00 e MAPE=5,29%. Na Tabela 19 apresentam-se os valores de previsão e, na Figura 40, ilustram-se as previsões dos 12 meses de 2011.

Tabela 19 Previsão da série de consumo de energia elétrica em horário fora de ponta, da UFLA, para modelo SARIMA e RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,5$ .

Período	Valor original	Valor Previsto
jan/11	445200	411964,17
fev/11	396200	407492,97
mar/11	435400	405149,07
abr/11	476000	480159,62
mai/11	494200	460295,37
jun/11	501200	458490,39
jul/11	457800	437914,73
ago/11	401800	400157,89
set/11	519400	434560,72
out/11	501200	487129,7
nov/11	481600	461495,95
dez/11	497000	487453,77

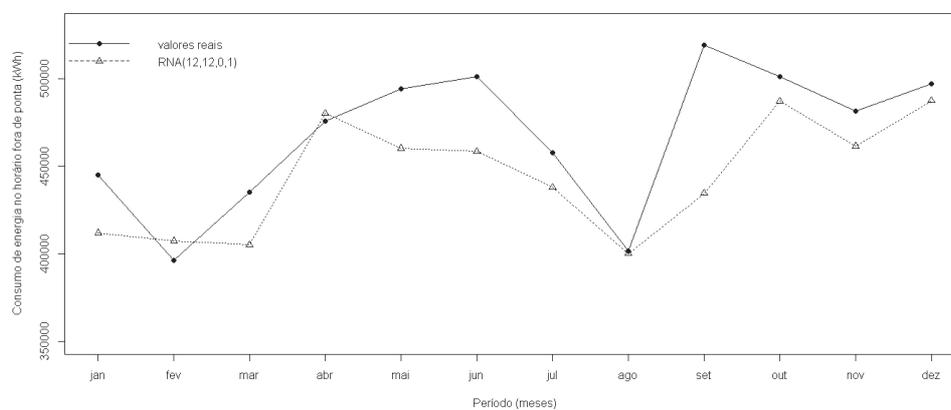


Figura 40 Comparação entre os valores reais e os previstos pelo modelo RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,5$

#### 4.2.3 Série da demanda de energia elétrica em horário de ponta

As configurações de rede, o tempo gasto para treinamento e teste e o EQM das principais redes testadas para a demanda registrada em horário de ponta podem ser observados na Tabela 20. Na Figura 41 apresenta-se o gráfico das arquiteturas em função do EQM.

Tabela 20 Configuração do perceptron múltiplas camadas para a série da demanda registrada no horário de ponta.

Arquitetura	Taxa de aprendizado	Momento	Tempo segundos	EQM
MLP(12,12,0,1)	0,01	0,5	565,641	0,0091682
	0,01	0,9	578,531	0,0095823
	0,1	0,5	566,922	0,0083196
	0,1	0,9	569,00	0,0077717
MLP(12,24,12,1)	0,01	0,5	838,219	0,0084338
	0,01	0,9	836,906	0,0089734
	0,1	0,5	835,047	0,0088914
	0,1	0,9	824,344	0,0088814
MLP(24,48,12,1)	0,01	0,5	1045,484	0,0070072
	0,01	0,9	1040,25	0,0076136
	0,1	0,5	1043,015	0,0119521
	0,1	0,9	1045,578	0,0200876

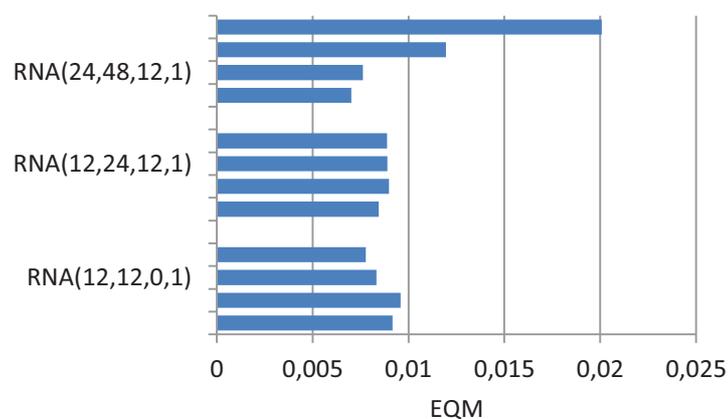


Figura 41 Valores do erro quadrático médio de treinamento das arquiteturas testadas para demanda registrada de energia elétrica em horário de ponta

Pelos resultados apresentados, observa-se que os valores do EQM das arquiteturas com um número maior de neurônios em sua composição são maiores do que arquiteturas com 12 neurônios na camada de entrada.

A arquitetura RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,9$ , foi selecionada para realizar previsões, pois apresentou menor EQM. A evolução do erro de treinamento com o número de iterações é mostrado na Figura 42.

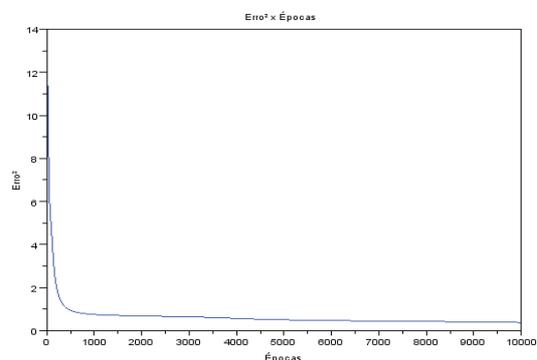


Figura 42 Gráfico do número de épocas, em função do erro quadrático médio do modelo RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,9$

Na Tabela 21 apresentam-se os valores previstos da demanda para o ano de 2011. O valor do EQMP obtido foi de 5.450,85 e o MAPE, de 7,38%. Na Figura 43 ilustra-se a comparação entre os valores reais e os previstos pelo modelo.

Tabela 21 Valores previstos da série de demanda de energia elétrica em horário de ponta, da UFLA, com o modelo RNA(12,12,0,1),  $\eta = 0,1$  e  $\alpha = 0,9$ .

<b>Período</b>	<b>Valor real</b>	<b>Valor previsto</b>
jan/11	812	788,224
fev/11	616	667,011
mar/11	854	791,384
abr/11	980	864,803
mai/11	980	895,414
jun/11	966	870,19
jul/11	938	883,77
ago/11	770	762,672
set/11	952	858,61
out/11	1006	927,486
nov/11	966	886,371
dez/11	938	870,949

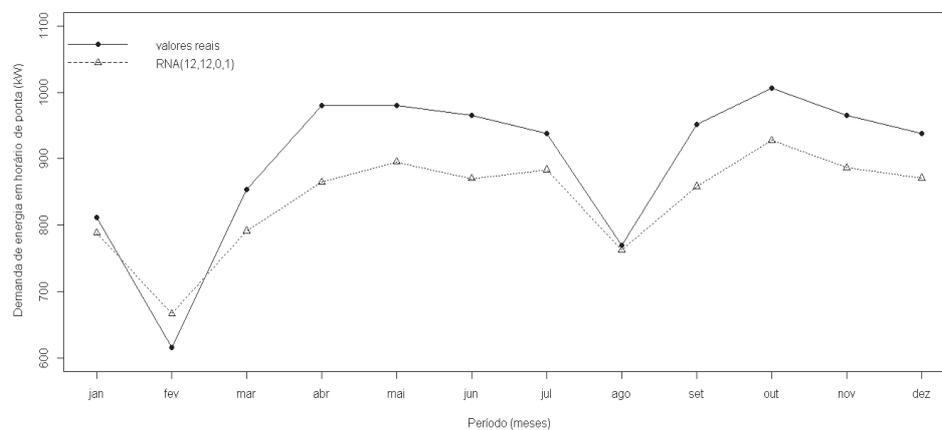


Figura 43 Comparação entre os valores reais e os previstos pelo modelo RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,9$

#### 4.2.4 Série da demanda registrada em horário fora de ponta

Para a demanda de energia elétrica registrada no horário de fora de ponta, na Tabela 22 apresentam-se as arquiteturas testadas, o erro quadrático médio de treinamento e o tempo gasto para treinar a rede. Na Figura 44 observa-se a evolução do EQM de treinamento em função das arquiteturas testadas, o que facilita a visualização do modelo a ser escolhido.

Tabela 22 Configuração do perceptron múltiplas camadas para a série da demanda registrada no horário fora de ponta.

Arquitetura	Taxa de aprendizado	Momento	Tempo, em segundos	EQM
MLP(12,12,0,1)	0,01	0,5	568,766	0,0043217
	0,01	0,9	562,437	0,0030599
	0,1	0,5	558,719	0,0028351
	0,1	0,9	592,281	0,0026596
MLP(12,24,12,1)	0, 01	0,5	842,484	0,0034310
	0, 01	0,9	854,89	0,0026120
	0,1	0,5	844,594	0,0027997
	0,1	0,9	851,406	0,0055960
MLP(24,48,12,1)	0, 01	0,5	1111,00	0,0032922
	0, 01	0,9	1071,015	0,0040308
	0,1	0,5	1076,406	0,0053723
	0,1	0,9	1072,36	0,0075194

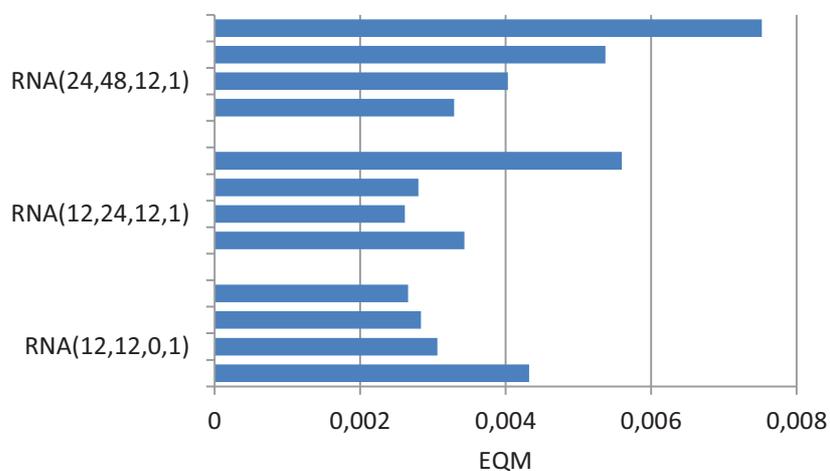


Figura 44 Valores do erro quadrático médio de treinamento das arquiteturas testadas para demanda registrada de energia elétrica em horário fora de ponta

Dentre os modelos gerados, o RNA (12,24,12,1), com  $\eta=0,01$  e  $\alpha=0,9$ , foi o que apresentou menor erro de generalização.

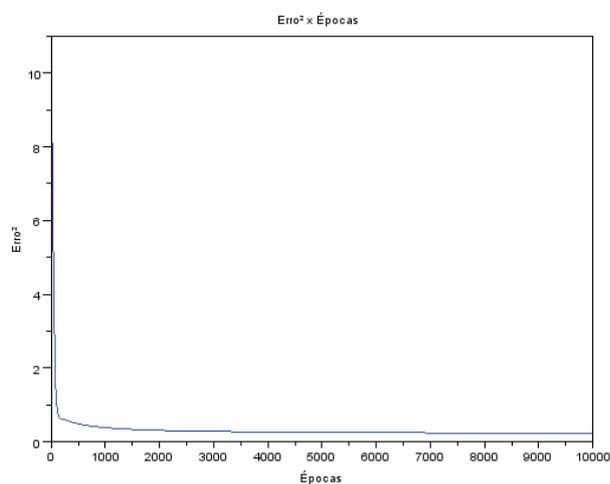


Figura 45 Gráfico do número de épocas em função do erro quadrático médio do modelo RNA(12,24,12,1), com  $\eta=0,01$  e  $\alpha=0,9$

As previsões para este modelo se encontram na Tabela 23 e, na Figura 46, apresenta-se a comparação entre os valores reais e previstos para o ano de 2011. Por meio do erro de previsão tem-se que  $EQMP=8.009,33$  e  $MAPE=4,99\%$ .

Tabela 23 Valores previstos da série de demanda de energia elétrica em horário fora de ponta, da UFLA, com o modelo RNA(12,24,12,1),  $\eta=0,01$  e  $\alpha=0,9$ .

Período	Valor original	RNA
jan/11	1456	1391,06
fev/11	1246	1274,51
mar/11	1540	1376,92
abr/11	1624	1455,96
mai/11	1540	1404,83
jun/11	1428	1395,15
jul/11	1274	1255,92
ago/11	1134	1067,62
set/11	1428	1354,84
out/11	1540	1470,51
nov/11	1484	1452,84
dez/11	1484	1452,13

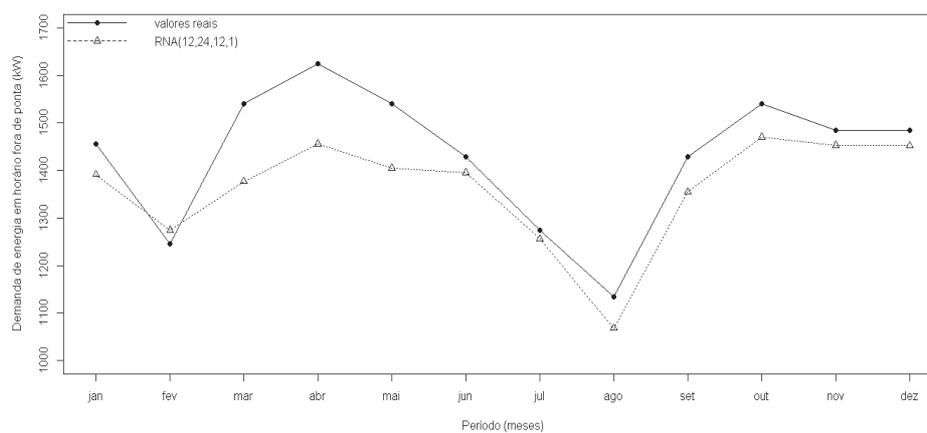


Figura 46 Comparação entre os valores reais e os previstos pelo modelo RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,5$

### 4.3 Análise comparativa

Conforme proposto neste trabalho, são apresentadas as comparações entre as metodologias para avaliar o desempenho das mesmas.

Todas as séries apresentaram as componentes tendência e sazonalidade de acordo com a análise de séries temporais. A sazonalidade se justifica pelo fato de o consumo de energia elétrica em uma instituição de ensino diminuir no período de férias, sendo diretamente influenciado pelo período de aulas.

A princípio, são comparados os resultados de previsão dos modelos sugeridos para a série do consumo de energia elétrica (kWh) em horário de ponta e fora de ponta e, em seguida, para as séries de demanda registrada de energia elétrica (kW) nos mesmos horários.

#### 4.3.1 Análise comparativa do consumo de energia elétrica nos horários de ponta e fora de ponta

Na Figura 47 (a) apresentam-se as previsões fornecidas pelos modelos SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI e RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,5$  e (b) para os modelos SARIMA(1,1,0)(0,1,1)<sub>6</sub> e RNA(12,12,0,1), com  $\eta=0,01$  e  $\alpha=0,9$ , com seus respectivos erros de previsão, de forma detalhada. Na Tabela 24 mostram-se os resultados dos erros quadráticos médio de previsão e do erro percentual médio de previsão.

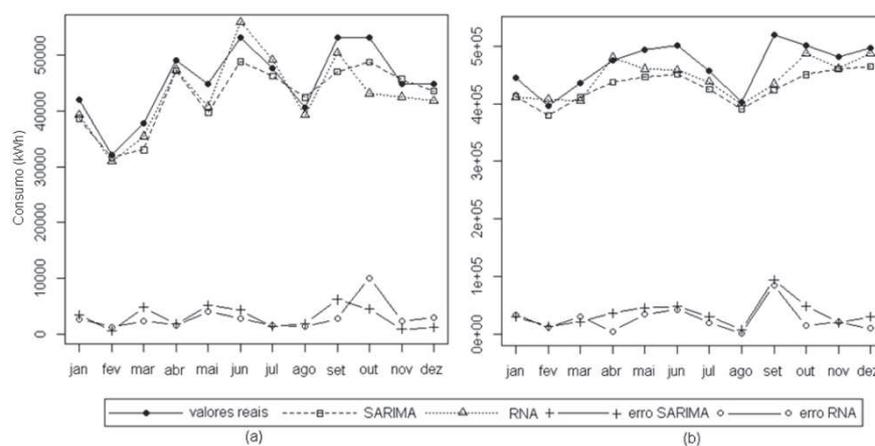


Figura 47 (a) Previsão para o consumo de energia elétrica em horário de ponta; (b) previsão para o consumo de energia elétrica em horário fora de ponta, por Box e Jenkins e Redes Neurais Artificiais, no período de janeiro a dezembro de 2011 e análise comparativa dos erros entre valor previsto e valor real

Tabela 24 Resultados dos erros de previsão da série de consumo de energia elétrica no horário de ponta e fora de ponta, para cada modelagem proposta.

Consumo de energia elétrica	Modelos	EQMP	MAPE
Horário de ponta	SARIMA(1,1,1)(0,1,1) <sub>12</sub> - <i>CI</i>	12420867,24	6,50%
	RNA(12,24,12,1) $\eta=0,1$ e $\alpha=0,5$	14070594,62	6,38%
Horário fora de ponta	SARIMA(1,1,0)(0,1,1) <sub>6</sub>	1707265154,00	7,28%
	RNA(12,12,0,1) $\eta=0,01$ e $\alpha=0,9$	1118942341,00	5,29%

No horário de ponta, o consumo, no ano de 2011, variou em torno de 32.200 kWh e 53.200 kWh e, no horário fora de ponta, de 396.200 kWh, a 519.400 kWh. Devido à dimensão dos dados, o erro absoluto apresentou altos valores, mas, em percentual, pode-se observar que os modelos apresentaram um bom desempenho.

No horário de ponta, o modelo de séries temporais, no mês de abril, ultrapassou, em 3%, o valor real do consumo. No mês de outubro, o modelo de redes neurais apresentou um erro de previsão de 18% em relação ao valor real, o que compromete a qualidade do desempenho do modelo sugerido. Neste horário, o modelo SARIMA apresentou menor EQMP, mas, em porcentagem, seus erros foram maiores.

Para o horário fora de ponta, de acordo com o gráfico dos erros, o mês de setembro apresentou uma defasagem significativa, para ambos os modelos. As previsões foram satisfatórias, mas o modelo SARIMA(1,1,0)(0,1,1)<sub>6</sub> apresentou suas estimativas bem abaixo dos valores reais, o que pode ser comprovado pelos valores altos dos erros e do EQMP e MAPE. Portanto, as redes neurais foram mais eficientes neste caso específico.

#### **4.3.2 Análise comparativa da demanda registrada de energia elétrica nos horários de ponta e fora de ponta**

Na Figura 48 apresenta-se o desempenho da previsão dos modelos selecionados. No gráfico (a) são apresentadas as previsões comparativas de demanda de energia elétrica no horário de ponta, dos modelos SARIMA(0,1,1)(3,1,1)<sub>6</sub> e RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,9$  e seus respectivos erros absolutos de previsões. Em (b) estão as previsões comparativas SARIMA(1,1,1)(0,1,1)<sub>6</sub> e RNA(12,24,12,1), com  $\eta=0,01$  e  $\alpha=0,9$  e seus respectivos erros absolutos para o horário fora de ponta.

A demanda contratada no ano 2011 também foi incluída nos gráficos para verificar o comportamento dos valores mensais observados e estimados em relação à tolerância de ultrapassagem de 5%. Para o horário de ponta, foram contratados 1.000 kW, nos meses de janeiro e fevereiro e 1.100 kW, no restante do ano. Para o horário de fora ponta, foram contratados 1.400 kW, nos meses de janeiro e fevereiro e 1.600 kW, no restante do ano.

Nota-se que, nos meses de janeiro, fevereiro e julho de 2011, a demanda de energia diminui e, nos outros meses, aumenta, devido ao período de aulas. As previsões também acompanharam a sazonalidade de seis meses. Na Tabela 25 apresentam-se os resultados do EQMP e MAPE dos modelos SARIMA e RNA sugeridos.

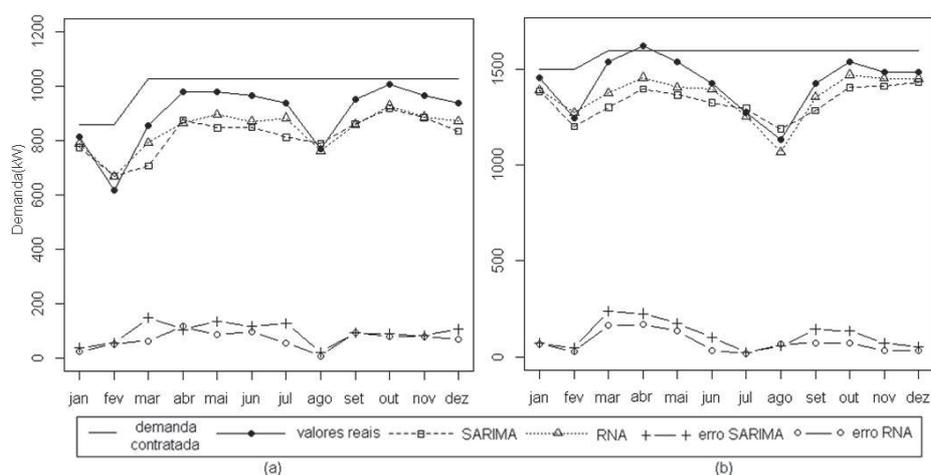


Figura 48 (a) Previsões para a demanda de energia elétrica em horário de ponta, por Box e Jenkins e Redes Neurais Artificiais; (b) previsões para a demanda de energia elétrica em horário fora de ponta, por Box e Jenkins e Redes Neurais Artificiais, no período de janeiro a dezembro de 2011 e análise comparativa dos erros entre valor previsto e valor real

Tabela 25 Resultados dos erros de previsão da série de demanda de energia elétrica no horário de ponta e fora de ponta, para cada modelagem proposta.

Demanda de energia elétrica	Modelos	EQMP	MAPE
Horário de ponta	SARIMA(0,1,1)(3,1,1) <sub>6</sub>	76098,12	8,02%
	RNA(12,12,0,1) $\eta=0,1$ e $\alpha=0,9$	5450,85	7,38%
Horário fora de ponta	SARIMA(1,1,1)(0,1,1) <sub>6</sub>	17115,32	7,49%
	RNA(12,24,12,1) $\eta=0,1$ e $\alpha=0,9$	8009,33	4,99%

Conforme a Figura 49 (a), nas previsões para o horário de ponta, ambos os modelos apresentaram-se abaixo da demanda contratada para este período. O modelo SARIMA(0,1,1)(3,1,1)<sub>6</sub> apresentou maiores valores dos erros absolutos em relação ao modelo proposto por redes neurais. De acordo com a Tabela 25, EQMP e MAPE para o modelo RNA apresentaram menores erros de previsão.

Para a série de demanda no horário fora de ponta, apenas o valor observado no mês de abril, de 1.624 kW, ultrapassou a demanda contratada, mas não excedeu ao limite de tolerância, que era de 1.680 kW.

O erro absoluto do modelo SARIMA, durante todo o ano, apresentou-se acima do modelo RNA. Com isso, o MAPE foi maior para este modelo. Nesse tipo de série, as redes neurais apresentaram um poder preditivo melhor.

De acordo com os resultados, a modelagem baseada na metodologia Box e Jenkins apresentou-se eficiente para representar as séries em estudo, captando os efeitos de tendência e sazonalidade ocorridos e descrevendo satisfatoriamente o comportamento destas. Mas, em relação à estimação de valores futuros, a aplicação de RNA obteve melhores resultados para as séries de consumo e demanda de energia elétrica, nos horários de ponta e fora de ponta da UFLA.

## 5 CONCLUSÃO

Foi realizado um estudo sobre o consumo e a demanda de energia elétrica da Universidade Federal de Lavras (UFLA), referente ao período de janeiro de 1995 a dezembro de 2011.

Utilizando-se os métodos de Box e Jenkins e Redes Neurais Artificiais (RNA), foram selecionados modelos para fins de previsão um passo a frente, de acordo com a metodologia sugerida.

Para a série de consumo de energia elétrica (kWh), os modelos de Box e Jenkins que apresentaram melhor desempenho foram SARIMA(1,1,1)(0,1,1)<sub>12</sub>-CI, para o horário de ponta e SARIMA(1,1,0)(0,1,1)<sub>6</sub>, para fora de ponta. Os modelos de redes neurais artificiais foram RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,5$  e RNA(12,12,0,1), com  $\eta=0,01$  e  $\alpha=0,9$ , para os horários de ponta e fora de ponta, respectivamente.

Pela análise comparativa, os erros de previsão foram menores para o modelo RNA, em ambos os horários. Portanto, as redes neurais foram mais eficientes neste caso específico.

Na demanda de energia registrada (kW), os modelos encontrados nos horários de ponta e fora de ponta foram SARIMA(0,1,1)(3,1,1)<sub>6</sub> e SARIMA(1,1,1)(0,1,1)<sub>6</sub>, para a modelagem de Box e Jenkins e, para as RNA, os modelos foram RNA(12,12,0,1), com  $\eta=0,1$  e  $\alpha=0,9$  e RNA(12,24,12,1), com  $\eta=0,1$  e  $\alpha=0,9$ .

De acordo com os resultados de previsão, os valores do MAPE e EQMP foram menores para as Redes Neurais Artificiais, comprovando a habilidade deste método em aprender padrões complexos e em lidar com a não linearidade e a não estacionariedade presentes nas séries temporais em estudo.

A metodologia Box e Jenkins apresentou-se eficiente para representar a série em estudo, captando os efeitos de tendência e sazonalidade ocorridos. As

redes neurais artificiais apresentam um bom recurso para a previsão de séries temporais de consumo e potência de demanda de energia elétrica.

É importante ressaltar que os modelos de redes neurais artificiais foram modelados com o auxílio das funções de autocorrelação e autocorrelação parcial da série e do periodograma, que forneceram a ordem de sazonalidade para escolher as entradas das RNAs testadas.

Portanto, ambos os métodos são válidos para serem utilizados como preditores e podem ser empregados conjuntamente para auxiliar na escolha do melhor modelo e, com isso, enriquecer ainda mais os trabalhos que tomam como base a previsão para o auxílio em tomadas de decisões.

Para trabalhos futuros, para os modelos de redes neurais, podem ser testadas outras arquiteturas, variando a taxa de aprendizagem e o termo de momento, bem como o número de épocas de treinamento, para verificar se ocorrem previsões mais precisas.

Outros fatores que influenciam a demanda de energia podem ser incluídos, tais como, mês, o número de alunos, professores e funcionários da instituição, índice pluviométrico, velocidade do vento, umidade, temperatura, dentre outros.

## REFERÊNCIAS

- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução N.º 456, de 29 de novembro de 2000**. Disponível em: <<http://www.aneel.gov.br/cedoc/res2000456.pdf>>. Acesso em: 23 nov. 2012.
- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Tarifas de fornecimento de energia elétrica**. 2005. 30 p. (Cadernos Temáticos ANEEL, 4).
- ANGELO, C. F. et al. Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro. **BBR - Brazilian Business Review**, p. 1-21, abr./jun. 2011.
- ARAÚJO, M. A. G. **Previsão de demanda de energia elétrica por meio de redes neurais artificiais**. 2005. 110 f. Dissertação (Mestrado em Administração)-Universidade do Rio Grande do Sul, Porto Alegre, 2005.
- BISHOP, C. M. **Neural networks for pattern recognition**. Oxford: Oxford University, 1997. 482 p.
- BOX, G.; JENKINS, G. **Time series analysis, forecasting and control**. San Francisco: Holden-Day, 1979. 575 p.
- BOX, G.; PIERCE, D. Distribution of residual correlations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, Alexandria, v. 65, p. 1509–1526, 1970.
- BRAGA, A.; CARVALHO, A.; LUDERMIR, T. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: Livros Técnicos e Científicos, 1998.
- BRASIL. Ministério de Minas e Energia. Empresa de Pesquisa Energética. **Balço energético nacional 2012**. Brasília, 2012. 57 p. Disponível em: <<https://ben.epe.gov.br/BENRelatorioFinal2012.aspx>>. Acesso em: 18 dez. 2012.
- BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. **Lei N.º 9.427, de 26 de Dezembro de 1996. Agência Nacional de Energia Elétrica – ANEEL**. Brasília, 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Leis/L9427cons.htm](http://www.planalto.gov.br/ccivil_03/Leis/L9427cons.htm)>. Acesso em: 27dez. 2012.

CAIO, L. S.; BERMANN, C. Análise das metodologias de previsão de mercado de energia elétrica face ao novo perfil de planejamento no ambiente pós-privatização. In: CONGRESSO BRASILEIRO DE PLANEJAMENTO ENERGÉTICO, 1998, São Paulo. **Anais...** São Paulo: 1998.

CALÔBA, L. P. Introdução ao uso de redes neurais na modelagem de sistemas dinâmicos e séries temporais. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 14., 2002, Natal, **Anais...** Natal, 2002.p. 52.

CAMPOS, R. J. **Previsão de séries temporais com aplicações a séries de energia elétrica**. 2008. 100 p. Dissertação (Mestrado em Engenharia Elétrica)- Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

CIGIZOGLU, H. Application of generalized regression neural networks to intermittent flow forecasting and estimation. **Journal of Hydrologic Engineering**, v. 10, n.4, p. 336–341, 2005.

ELETROBRAS. **Manual de tarifação de energia elétrica**. Rio de Janeiro, 2011. 58 p.

ESPÍRITO SANTO CENTRAIS ELÉTRICAS. **Energias do Brasil**. Disponível em: <<http://www.escelsa.com.br/aescelsa/historia-ee-brasil.asp>>. Acesso em: 27 dez. 2012.

FRANK, R. J.; DAVEY, N.; HUNT, S.P. Time series prediction and neural networks. **Journal of Intelligent and Robotic Systems**, v. 31, n. 1, p. 91-103, 2001.

GUYTON, A. C., HALL, J. E. **Tratado de fisiologia médica**. 9. ed. Rio de Janeiro: Guanabara Koogan, 1997. 864 p.

HAYKIN, S. **Neural networks, a comprehensive foundation**. New Jersey: Englewood Cliffs. 2001. 842 p.

HEBB, D. **The organization of behavior**. New York: J. Wiley, 1949. 378 p.

HOPFIELD, J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 79, p.2554-2558, 1982.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas imanente in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, p. 115-133, 1943.

- MILONE, G. **Estatística: geral e aplicada**. São Paulo: Pioneira Thomson Learning, 2004.
- MINSKY, M. L.; PAPERT, S. A. **Perceptrons**. Cambridge: MIT, 1969.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: E. Blücher, 2006. 538 p.
- MOURA, F. A. **Previsão do consumo de energia elétrica por setores através do modelo SARMAX**. 2011, 138f. Dissertação (Mestrado em Administração) - Universidade de São Paulo, São Paulo, 2011.
- OLIVEIRA, A. C. S. et al. Aplicação de Redes Neurais Artificiais na previsão da produção de álcool. **Ciência e Agrotecnologia**, Lavras, v. 34, n. 2, p. 279-284, mar./abr. 2010.
- PRADO, J. R. **Modelos para demanda e consumo de energia elétrica utilizando séries temporais na Universidade Federal de Lavras**. 2011. 114 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária)- Universidade Federal de Lavras, Lavras, 2011.
- R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna, 2011. Disponível em <<http://www.r-project.org>>. Acesso em: 3 ago. 2011.
- ROMERA, E.; MORÁN, M.; FERNÁNDEZ, D. Forecasting of the electric energy demand trend and monthly fluctuation with neural networks. **Elsevier Computers and Industrial Engineering**, v. 52, n. 3, p. 336-343, 2007.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, Washington, v. 65, p. 386-408, 1958.
- RUMELHART, D.; HINTON, G.; WILLIAMS, R. Learning representation sbyback-propagations erros. **Nature**, London, v. 323, p. 533-536,1986.
- SAES, A. M. **Energia elétrica e legislação no Brasil no início do século XX**. In: ANPUH- SIMPÓSIO NACIONAL DE HISTÓRIA, 25., 2009, Fortaleza. Disponível em: <<http://anpuh.org/anais/wp-content/uploads/mp/pdf/ANPUH.S25.0005.pdf>>. Acesso em: 27 dez. 2012.

SANTOS, J. V. C. dos; VALIANI, J. F.; GALVÃO NETO, W. P. Previsão de demanda por energia elétrica: inclusão do atributo temperatura. **PODES**. Pesquisa operacional para o desenvolvimento, Rio de Janeiro, v. 3, n. 2, p. 123-131, maio/ago. 2011.

SCILAB ENTERPRISES. **Scilab: free and open source software for numerical computation**. Versailles, 2011. Disponível em <<http://www.scilab.org>>. Acesso em 13 mai. 2011.

WIDROW, B.; HOFF, M. E. Adaptatives witching circuits. **IRE WESCON Convention Record**, p. 96-104, 1960.

## APÊNDICE

Apresentam-se as análises iniciais das séries de consumo de energia elétrica no horário fora de ponta e demanda registrada no horário de ponta e fora de ponta, cujo procedimento foi feito de forma análoga ao da série de consumo de energia elétrica no horário de ponta apresentado na seção 4.1.1.

### **Série do consumo de energia elétrica no horário fora de ponta**

Na Figura 1A apresenta-se o correlograma da série original, sugerindo a presença de tendência devido ao decaimento lento dos lags. Pelo teste de sequências Wald-Wolfowitz, rejeitou-se a hipótese nula de que a série não possui tendência, pois se tem que valor- $p = 0,000002$ , com a estatística de  $-9,9855$ . Pelo Teste de Fisher  $g = 0,1733 > z = 0,0726$ , logo, a série apresenta presença de periodicidade. Este fato é confirmado por meio do periodograma na Figura 2A, que apresenta um maior pico no período de seis meses, o que indica sazonalidade de ordem 6.

Foi feito o gráfico da amplitude *versus* média para verificar se há necessidade de uma transformação logarítmica na série de consumo de energia. Pela Figura 3A verifica-se que é preciso fazer a transformação devido à inclinação da reta. Portanto, toma-se uma primeira diferença para retirar a tendência e uma diferença de 6 para retirar a sazonalidade na série transformada e, finalmente, é obtida uma série estacionária.

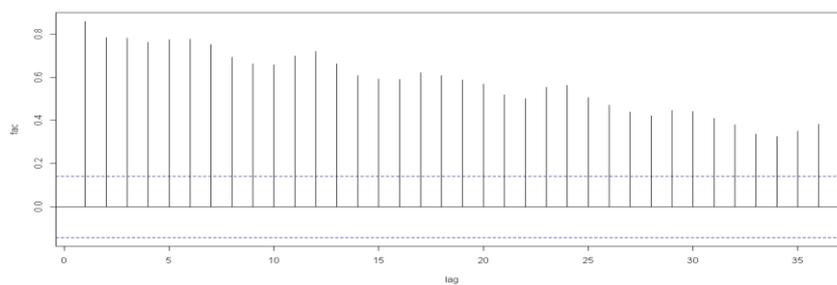


Figura 1A Correlograma da série de consumo de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

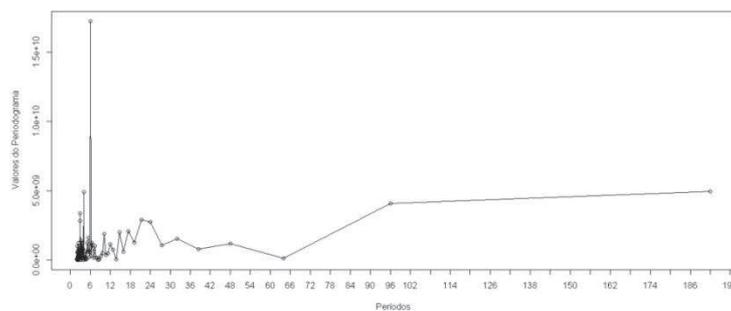


Figura 2A Periodograma da série de consumo de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

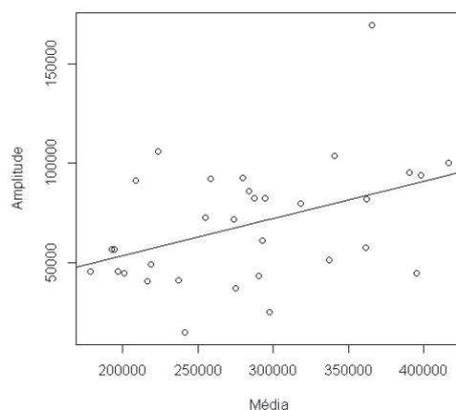


Figura 3A Gráfico média  $\times$  amplitude da série de consumo de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

#### Série da demanda registrada em horário de ponta

O periodograma representado na Figura 4 é utilizado para testar a sazonalidade e apresenta um pico no período 6. Pelo teste de Fisher, têm-se  $g = 0,2134788$  e  $z = 0,07267625$ , verificando-se que a série em estudo tem periodicidade 6, pois a estatística  $g$  é maior que  $z$ .

Aplicando-se o teste de sequências Wald-Wolfowitz, tem-se que o valor- $p < 2,2e^{-16}$  para a estatística  $-9,1043$ , ou seja, a série possui tendência. Na Figura 5 apresenta-se o correlograma da série original, sugerindo a presença de tendência devido ao decaimento lento dos lags.

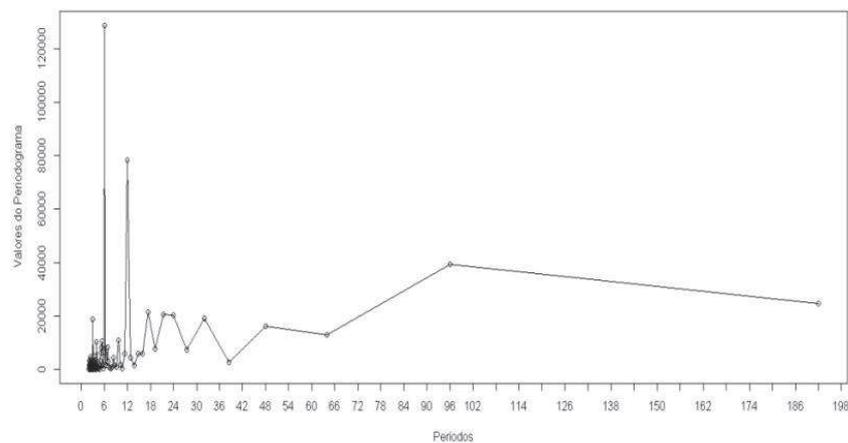


Figura 4A Periodograma da série de demanda de energia elétrica em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

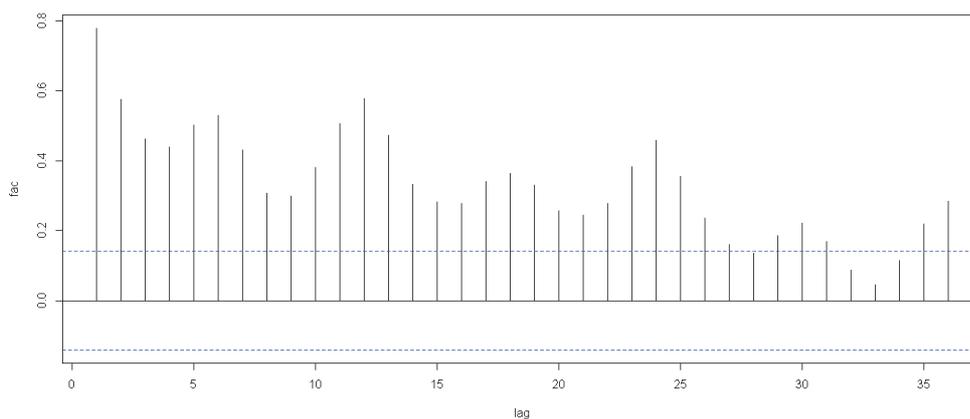


Figura 5A Correlograma da série de demanda de energia elétrica em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

Por meio do gráfico da média *versus* amplitude, verificou-se a necessidade de uma transformação logarítmica. Após a transformação, a retirada da tendência e sazonalidade por meio de diferenças, tem-se a série estacionária, como mostrado na Figura 6.

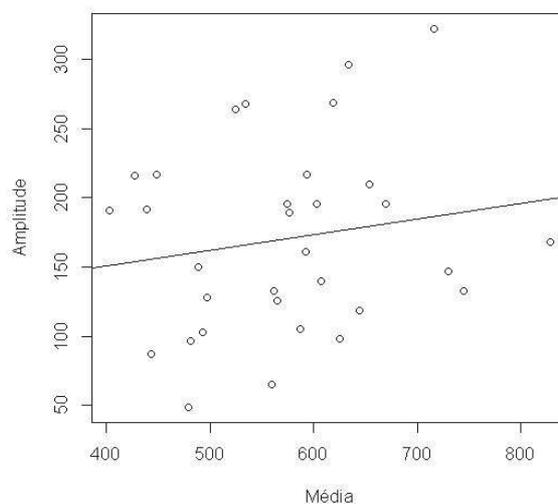


Figura 6A Gráfico média  $\times$  amplitude da série de demanda de energia elétrica em horário de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

### Série da demanda registrada em horário fora de ponta

Na Figura 7A apresenta-se o periodograma, que apresenta um pico no período 6, indicando uma componente periódica de 6 meses e, por meio do correlograma da série original (Figura 8A), verifica-se que há indícios de tendência e sazonalidade.

A série foi dividida em grupos de seis observações consecutivas e, para cada grupo, calcularam-se a média e a amplitude. O comportamento do gráfico da média  $\times$  amplitude, apresentado na Figura 9A, sugere uma transformação logarítmica.

O teste de Fisher foi aplicado para verificar a presença da componente sazonal, obtendo-se o valor das estatísticas  $g = 0,1885895$  e  $z = 0,07267625$ . Como  $g > z$ , a série apresenta sazonalidade de 6 meses. Com isso fizemos uma diferença sazonal de ordem 6, com o objetivo de eliminar esta componente.

Em seguida, aplicou-se o teste de sequências Wald-Wolfowitz, com p-valor  $< 2,2e-16$  e estatística  $-12,0107$ , confirmando a presença de tendência. Uma diferença foi feita para eliminar esta componente e, com isso, a série torna-se estacionária.

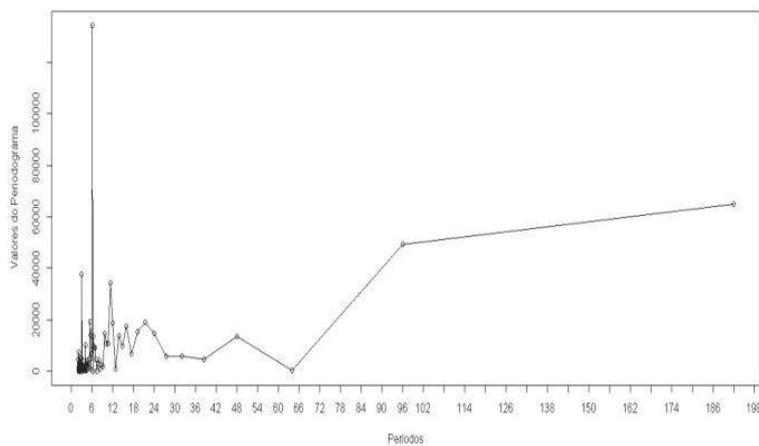


Figura 7A Periodograma da série de demanda de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

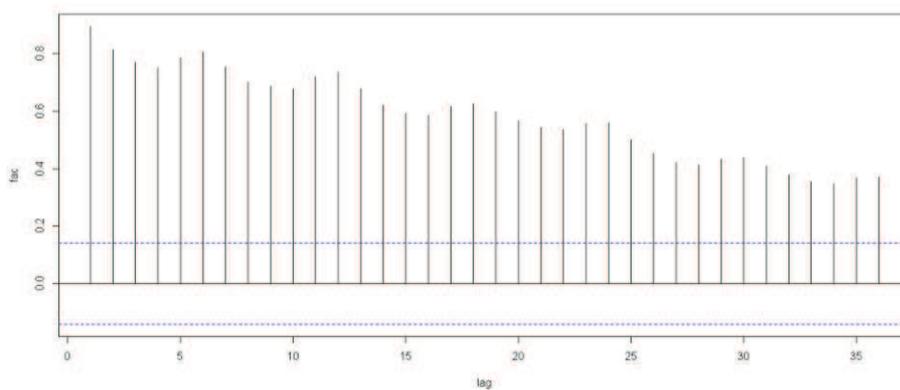


Figura 8A Correlograma da série de demanda de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

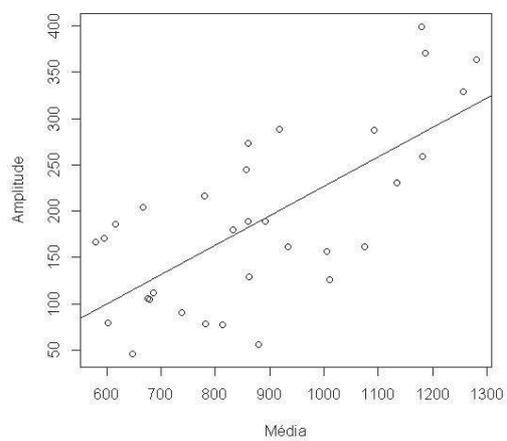


Figura 9A Gráfico média  $\times$  amplitude da série de demanda de energia elétrica em horário fora de ponta, da UFLA, no período de janeiro de 1995 a dezembro de 2010

### Algoritmo de retropropagação para treinamento da RNA

```

//mlp com toolbox ANN
//consumo total

clear
clc
// DECLARAÇÃO DAS FUNÇÕES

// PROGRAMA PRINCIPAL
// Arquitetura da rede
// #entradas, # neurônios na camada oculta e 1 neurônio de saída
Entradas = 12;
CE1 = 24;
CE2 = 12;
Saida = 1;
N=[Entradas, CE1, CE2, Saida]

// Amostras (colunas) das 2 entradas (linhas)
X_original = [];//conjunto de dados

//leitura do arquivo
arq=fopen('C:\maira\soma_consumo.txt','r');

while (feof(arq) == 0)//encontrar o fim do arquivo
    str = fscanf(arq, "%s");//leitura de uma string
    num = strtod(str);//converter string para número
    X_original = [X_original num]//conjunto de dados
end
fclose (arq);//fecha arquivo

X_original = X_original'

X_treina = [];
Y_treina = [];
inicio=1;
fim=Entradas; //atraso
n=2; //numero de anos para teste (2 anos) e perda(1 ano)

for i=1:(length(X_original)-(n*fim));//tamanho de dados apenas para treinamento
    X_treina(:,i)=X_original(inicio:fim);
    inicio=inicio+1;
    fim=fim+1;
    Y_treina(i)=X_original(fim);// Saída desejada para cada amostra

```

```

end
Y_treina=Y_treina';

//Normalizar os dados de entrada
X_norm=X_treina/max(X_original);

//Normalizar saida desejada
Y_norm=Y_treina/ max(X_original );

// Parâmetros de treinamento da rede
// 0.1 é a taxa de aprendizado; 0 é o valor mínimo para a atualização do pesos
durante o treinamento, 0.5 é a constante momento
lp = [0.1 0 0.5 0];

// Número de épocas de treinamento
T = 10000;

printf(' Épocas: %d\n',T)

EEpocas = [];

/////treinamento da rede/////
// X são os dados de entrada
// Y é a saída desejada
// N é a arquitetura da rede
// W são os pesos iniciais
// lp são os parâmetros
// T é o número de épocas

// Inicializa a MPL
disp('INICIANDO A REDE...')
W =ann_FF_init(N);
//inicializa matriz de peso
Delta_W_old=hypermat(size(W));

disp('TREINANDO A REDE...')
tic();
for i= 1:T

[W,Delta_W_old]=ann_FF_Mom_online(X_norm,Y_norm,N,W,lp,1,Delta_W_old);
// Execução da rede
Y_saida_treinamento = ann_FF_run(X_norm,N,W);
E = Y_norm - Y_saida_treinamento;
EEpoca= sum(E.^2);//erro por época

```

```

    printf("erro: %f\nepoca: %d\n",EEpoca,i)
    EEpocas= [EEpocas EEpoca];
end
tempo=toc();
disp("TEMPO DE TREINAMENTO: ")

//resultados desnormalizados
Y_resultado=Y_saida_treinamento*max(X_original);

scf(0);
plot(Y_treina,'k-')//plotar os dados originais
plot(Y_resultado,'r-')//plotar os dados classificados pela rede
title("CLASSIFICAÇÃO DOS DADOS DE TREINAMENTO");
legend('Dados originais','Classificação da rede')
xtitle('período','consumo')
//isoview(-1,2,-1,2);
scf(1);
title("Erro² x Épocas ");
xlabel("Épocas");
ylabel("Erro² ");
plot(EEpocas);

//*****//
////teste da rede com os anos 2010 e 2011////////

disp("TESTANDO A REDE...")
n = 2;
inicio2=1
fim2=Entradas
indice = (length(X_original)-(n*fim2));
X_aux = X_original((indice+1):$);

for i=1:10;//tamanho de dados apenas para teste
    X_teste(:,i) = X_aux(inicio2:fim2);
    inicio2=inicio2+1;
    fim2=fim2+1;
    Y_teste(i)=X_aux(fim2);// Saída desejada para cada amostra
end
Y_teste=Y_teste';//valores reais

//normalização dos dados de teste

X_teste_norm=X_teste/ max(X_original);

```

```
//normalização da saída desejada
Y_teste_norm=Y_teste/ max(X_original);

//execução da rede treinada com dados de teste
Y_saida_teste=ann_FF_run(X_teste_norm,N,W);

Y_teste_resultado= Y_saida_teste* max(X_original);

scf(2)
plot(Y_teste_resultado,'k.-')
plot(Y_teste,'r.-')
title('CLASSIFICAÇÃO DOS DADOS DE TESTE')
legend('valores previstos', 'valores originais',2)
xtitle('período','consumo')
```