

**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE – UFCG  
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA – CEEI  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA – DEE**

THAINÁ SANTOS XAVIER

**PREVISÃO DE SÉRIES TEMPORAIS UTILIZANDO MODELOS CLÁSSICOS E  
REDES NEURAS ARTIFICIAIS**

Campina Grande, PB

Agosto de 2018

THAINÁ SANTOS XAVIER

**PREVISÃO DE SÉRIES TEMPORAIS UTILIZANDO MODELOS CLÁSSICOS E  
REDES NEURAS ARTIFICIAIS**

*Trabalho de Conclusão de Curso submetido à Unidade  
Acadêmica de Engenharia Elétrica da Universidade  
Federal de Campina Grande como parte dos requisitos  
necessários para a obtenção do grau de Bacharel em  
Ciências no Domínio da Engenharia Elétrica.*

Área de Concentração: Processamento de Energia

Orientador:  
George Rossany Soares de Lira, D. Sc.

Campina Grande, Paraíba  
Agosto, 2018

THAINÁ SANTOS XAVIER

**PREVISÃO DE SÉRIES TEMPORAIS UTILIZANDO MODELOS CLÁSSICOS E  
REDES NEURAIS ARTIFICIAIS**

*Trabalho de Conclusão de Curso submetido à Unidade  
Acadêmica de Engenharia Elétrica da Universidade  
Federal de Campina Grande como parte dos requisitos  
necessários para a obtenção do grau de Bacharel em  
Ciências no Domínio da Engenharia Elétrica.*

Área de Concentração: Processamento de Energia

Aprovado em        /        /

**Professor Edson Guedes da Costa, D. Sc.**  
Universidade Federal de Campina Grande  
Avaliador

**Professor George Rossany Soares de Lira, D. Sc.**  
Universidade Federal de Campina Grande  
Orientador

Dedico este trabalho à minha avó Amparo (*in memorian*), que me deu tanto amor e tanto desejou, em suas orações, que eu estivesse aqui hoje.

## Agradecimentos

Agradeço, em primeiro lugar, ao bom Deus, pelas infinitas graças derramadas em minha vida e por me conceder o dom da fé, a qual me levantou em tantos momentos difíceis ao longo da caminhada da graduação.

À minha família, em especial, aos meus pais, por todo amor, carinho, cuidado e por todo esforço para criar e educar a mim e à minha irmã, nos proporcionando sempre o melhor que podiam. Eles foram e são minha maior motivação para buscar um futuro melhor.

A todos os professores do Departamento de Engenharia Elétrica da UFCG, que contribuíram direta ou indiretamente para minha formação.

Aos colegas de curso, agradeço por todo apoio, por toda ajuda recebida e pelos conhecimentos compartilhados.

E a George, meu orientador neste trabalho, agradeço pela paciência, atenção e compreensão em tantos momentos, não só aqui, mas em todos os projetos de pesquisa nos quais foi meu orientador.

## Resumo

Este trabalho apresenta os principais conceitos que envolvem a análise de séries temporais, como também os principais modelos utilizados na previsão de séries temporais. Ao final do trabalho, é realizado um estudo de caso, no qual foi feita a previsão de carga da região Nordeste utilizando um modelo clássico de previsão, especificamente, o modelo ARIMA sazonal, e uma rede neural artificial (RNA). A metodologia utilizada para identificar o modelo ARIMA é baseada na análise das funções de autocorrelação e autocorrelação parcial da série temporal. A RNA foi implementada no *software* Matlab®, por meio da *toolbox Neural Network Time Series Tool*. A qualidade do ajuste dos modelos e das previsões foram medidas por meio de índices de desempenho. Chegou-se à conclusão de que ambos os métodos de previsão são satisfatórios, por apresentarem bons índices de desempenho. Porém, as redes neurais artificiais apresentam algumas vantagens sobre o modelo ARIMA, como a possibilidade de incorporar variáveis de regressão ao modelo. Este trabalho tem como principal contribuição a de dar início a uma linha de pesquisa no Laboratório de Alta Tensão da UFCG, servindo como um guia para os alunos de graduação que irão trabalhar com o tema daqui por diante.

**Palavras-chave:** previsão de séries temporais, processos estocásticos, previsão de carga, ARIMA, redes neurais artificiais.

## Abstract

This work presents the main concepts that involve the analysis of time series, as well as the main models used in the time series forecasting. At the end of the work, a case study was carried out, in which the load Northeast region was predicted using a classical forecasting model, specifically the seasonal ARIMA model, and an artificial neural network (ANN). The methodology used to identify the ARIMA model is based on the analysis of autocorrelation and partial autocorrelation functions of the time series. The ANN was implemented in the Matlab® software through the Neural Network Time Series Tool toolbox. The quality of adjustment of models and forecasts was measured by performance indexes. It was concluded that both forecasting methods are satisfactory, since they have good performance indices. However, artificial neural networks have some advantages over the ARIMA model, such as the possibility of incorporating regression variables into the model. This work has as main contribution to initiate a research line in the High Voltage Laboratory of the UFCG, serving as a guide for the graduating students who will work on the subject from now on.

**Keywords:** time series forecasting, stochastic processes, load forecasting, ARIMA, artificial neural networks.

## Lista de Ilustrações

Figura 2.1 Série temporal do consumo mensal de energia elétrica da região Nordeste.....	4
Figura 2.2 Tendência da série de consumo de energia elétrica da região Nordeste. ....	6
Figura 2.3 Componente sazonal da série de consumo de energia elétrica da região Nordeste. ....	7
Figura 2.4 Componente aleatória da série de consumo de energia elétrica da região Nordeste. ....	8
Figura 2.5 FAC amostral da série de consumo de energia elétrica da região Nordeste. ....	12
Figura 3.1 Processo de suavização de dados (Adaptado de: MONTGOMERY <i>et al.</i> , 2008). ....	15
Figura 3.2 Modelo de um neurônio artificial, chamado k (Adaptado de: HAYKIN, 2009).....	22
Figura 3.3 Efeito do <i>bias</i> (Adaptado de: HAYKIN, 2009). ....	22
Figura 3.4 Função degrau (Adaptado de: KRIESEL, 2007). ....	23
Figura 3.5 Formas da função logística para diferentes valores de a (Adaptado de: KRIESEL, 2007). ....	24
Figura 3.6 Função tangente hiperbólica (Adaptado de: KRIESEL, 2007). ....	24
Figura 3.7 Rede neural recorrente (Adaptado de: HAYKIN, 2009).....	25
Figura 3.8 Rede neural não recorrente com uma camada escondida (Adaptado de: HAYKIN, 2009). ....	25
Figura 4.1 Etapas do processo de obtenção do modelo ARIMA (Fonte: CAMPOS, 2008). ....	29
Figura 4.2 Diagrama de blocos da RNA simulada no MATLAB. ....	31
Figura 4.3 Estrutura da RNA para obter previsões de um passo à frente. ....	32
Figura 5.1 FAC amostral da série derivada. ....	36
Figura 5.2 FACP amostral da série derivada. ....	36
Figura 5.3 Modelo ARIMA sazonal ajustado à série temporal de consumo de energia elétrica da região Nordeste. ....	38
Figura 5.4 Previsões com modelo ARIMA sazonal com horizonte de 12 meses. ....	38
Figura 5.5 Previsões com modelo ARIMA sazonal com horizonte de 24 meses. ....	39
Figura 5.6 Previsões com modelo ARIMA sazonal com horizonte de 60 meses. ....	40
Figura 5.7 Resposta da RNA ajustada aos dados de consumo de energia elétrica da região Nordeste. ....	41
Figura 5.8 Previsões com RNA com horizonte de 12 meses. ....	42
Figura 5.9 Previsões com RNA com horizonte de 24 meses. ....	42
Figura 5.10 Previsões com RNA com horizonte de 60 meses. ....	43

## Lista de símbolos e abreviaturas

AR – *Autorregressive*

MA – *Moving average*

ARMA – *Autoregressive moving average*

ARIMA - *Autoregressive integrated moving average*

CPTEC – Centro de Previsão de Tempo e Estudos Climáticos

EQM – Erro quadrático médio

FAC – Função de autocorrelação

FACP – Função de autocorrelação parcial

MAE – *Mean absolute error*

MAPE – *Mean absolute percentage error*

ME – *Mean error*

MPE - *Mean percentage error*

MSE – *Mean squared error*

NE - Nordeste

ONS – Operador Nacional do Sistema Elétrico

PIB – Produto interno bruto

RMSE – *Root mean squared error*

RNA – Rede neural artificial

SSE – *Sum of squared errors*

SST – *Total sum of squares*

## Sumário

1.	Introdução.....	1
1.1	Objetivo.....	2
1.2	Estrutura do trabalho.....	2
2.	Fundamentação teórica.....	4
2.1	Séries temporais: definição e análise clássica.....	4
2.2	Conceitos básicos de processos estocásticos.....	9
2.2.1	Estacionariedade.....	9
2.2.2	Funções de autocovariância e autocorrelação.....	10
3.	Modelos para previsão de séries temporais.....	13
3.1	Modelos regressivos.....	13
3.2	Modelos de suavização exponencial.....	14
3.2.1	Modelo de suavização exponencial simples.....	15
3.3	Modelos lineares para séries temporais estacionárias.....	16
3.3.1	Modelos autorregressivos – AR(p).....	17
3.3.2	Modelos de médias móveis - MA(q).....	17
3.3.3	Modelos mistos autorregressivos de médias móveis - ARMA(p, q).....	18
3.4	Modelos autorregressivos integrados de médias móveis - ARIMA(p, q, d).....	18
3.4.1	Modelos ARIMA sazonais.....	19
3.5	Redes neurais artificiais.....	21
3.5.1	Funções de ativação.....	23
3.5.2	Arquiteturas de redes neurais artificiais.....	24
3.5.3	Aprendizagem de redes neurais artificiais.....	26
4.	Metodologia.....	28
4.1	Previsões com modelo ARIMA sazonal.....	28
4.2	Previsões com redes neurais artificiais.....	30
4.3	Índices de desempenho.....	32
5.	Estudo de caso: Previsão de carga da região Nordeste.....	35
5.1	Resultados obtidos com modelo ARIMA sazonal.....	35
5.2	Resultados obtidos com a RNA.....	40
6.	Conclusões.....	44
	Referências	45

## 1. Introdução

O trabalho de realizar previsões envolve incertezas, pois, devido a aleatoriedade dos fenômenos que são previstos, sempre existirão erros inerentes ao processo de previsão. Por isso, a previsão de séries temporais é um campo de estudo da Estatística e está associada a teoria dos processos estocásticos. Prever acontecimentos futuros é uma atividade de interesse dos mais diversos setores, como a indústria, comércio, economia e política.

O exemplo mais comum de previsão de séries temporais é a previsão do tempo, a qual é noticiada todos os dias nos telejornais. Atualmente, ela é realizada em tempo real por supercomputadores, que resolvem modelos matemáticos complexos, e são interpretadas pelos meteorologistas. Apesar de anos de estudos e sofisticação de equipamentos de medição, este tipo de previsão pode falhar. Segundo o Centro de Previsão de Tempo e Estudos Climáticos - CPTEC/INPE, as previsões são geradas para até 15 dias, com 98% de acerto para os dois primeiros, chegando a 70% com cinco dias.

A previsão do tempo é um exemplo de previsão que pode ser utilizada no processo de tomada de decisão de empresas e até mesmo no dia-a-dia das pessoas. Prever o futuro faz parte da rotina de grandes organizações que desejam fazer um planejamento estratégico eficiente, por meio da análise de demandas futuras, tendências de preços e da expansão ou retração da economia, inflação e outras variáveis econômicas. Essas análises norteiam o uso dos recursos financeiros, apontando de que forma o capital deve ser investido nos horizontes de curto, médio e longo prazo.

No setor elétrico, a previsão de séries temporais é bastante utilizada na previsão de demanda de energia elétrica. Segundo o Operador Nacional do Sistema Elétrico (ONS), as atividades de programação da operação têm como insumo as informações sobre a expansão da geração e transmissão, o estado atual de armazenamento dos reservatórios, previsões de carga de energia por patamar, a análise das condições meteorológicas verificadas e previstas nas principais bacias do SIN e previsões de aflúncias aos aproveitamentos hidrelétricos.

Para que seja possível fazer uma previsão a partir de uma série temporal, é necessário construir um modelo matemático a partir do qual seja possível obter

valores futuros da série. Os modelos de Box e Jenkins, também conhecidos como ARIMA (*Autoregressive Integrated Moving Average*), são amplamente utilizados. Mas novos métodos baseados em inteligência artificial, como as redes neurais artificiais (RNA), também vêm ganhando espaço.

A escolha do modelo ou método é uma etapa importante no processo de previsão, visto que cada metodologia se adequa melhor a determinadas séries temporais, com características bem definidas, como estacionariedade, sazonalidade e tendência. A escolha do método correto minimiza os erros de previsão, tornando o processo mais preciso e confiável.

Diante do exposto, a proposta deste trabalho é a realização de uma revisão bibliográfica acerca dos principais conceitos relacionados à previsão de séries temporais, abordando também os modelos mais utilizados nas previsões. Além disso, pretende-se realizar um estudo de caso, no qual será feita a previsão de carga de médio/longo prazo da região Nordeste, com base nas séries temporais de consumo da região fornecidas pelo Operador Nacional do Sistema (ONS).

## 1.1 Objetivo

O objetivo geral deste trabalho é estudar os principais conceitos relacionados à análise e previsão de séries temporais e apresentar os modelos mais utilizados nas previsões.

Os objetivos específicos são:

- Realizar previsão de carga de médio/longo prazo da região Nordeste por meio do modelo ARIMA sazonal e por uma rede neural artificial (RNA).
- Analisar, por meio de técnicas adequadas, o desempenho dos modelos utilizados na previsão de carga, de modo a identificar qual o mais adequado para o problema em questão.

## 1.2 Estrutura do trabalho

Este trabalho está dividido em seis capítulos, incluindo esta introdução. No Capítulo 2, serão apresentados as ferramentas e os conceitos básicos de processos

estocásticos utilizados na análise de séries temporais, com ênfase no conceito de estacionariedade e definição das funções de autocovariância e autocorrelação.

O Capítulo 3 trata a respeito dos principais modelos utilizados na modelagem e previsão de séries temporais: modelos regressivos, modelos de suavização exponencial, modelos para séries estacionárias (autorregressivos – AR, médias móveis – MA, e mistos autorregressivos de médias móveis – ARMA), modelos para séries não estacionárias (autorregressivo integrado de médias móveis – ARIMA) e redes neurais artificiais – RNA.

A metodologia utilizada para realizar as previsões do consumo de energia elétrica da região Nordeste será apresentada no Capítulo 4. Ainda neste capítulo, serão definidos alguns índices de desempenho utilizados na análise de desempenho de modelos e previsões de séries temporais.

O Capítulo 5 apresenta os resultados obtidos na modelagem e previsão da série de consumo em estudo, fazendo uma comparação entre o desempenho de um modelo clássico (ARIMA) e de uma RNA. O trabalho é finalizado com as conclusões no Capítulo 6.

## 2. Fundamentação teórica

Neste capítulo, serão apresentados os principais conceitos que envolvem a análise de séries temporais. Na seção 2.1, define-se o que é uma série temporal e como se dá sua análise clássica. Já na seção 2.2, serão abordados os conceitos básicos de processos estocásticos, cuja utilidade será justificada ao longo do capítulo.

### 2.1 Séries temporais: definição e análise clássica

Uma série temporal é um conjunto de observações, em função do tempo, de um fenômeno ou uma variável qualquer. A exemplo, a Figura 2.1 apresenta a série temporal do consumo mensal de energia elétrica da região Nordeste, de janeiro de 2000 a dezembro de 2017. Segundo Ehlers (2009), a característica mais importante das séries temporais é que as observações vizinhas são dependentes. Seu estudo consiste em analisar e modelar esta dependência por meio de conceitos e modelos probabilísticos.

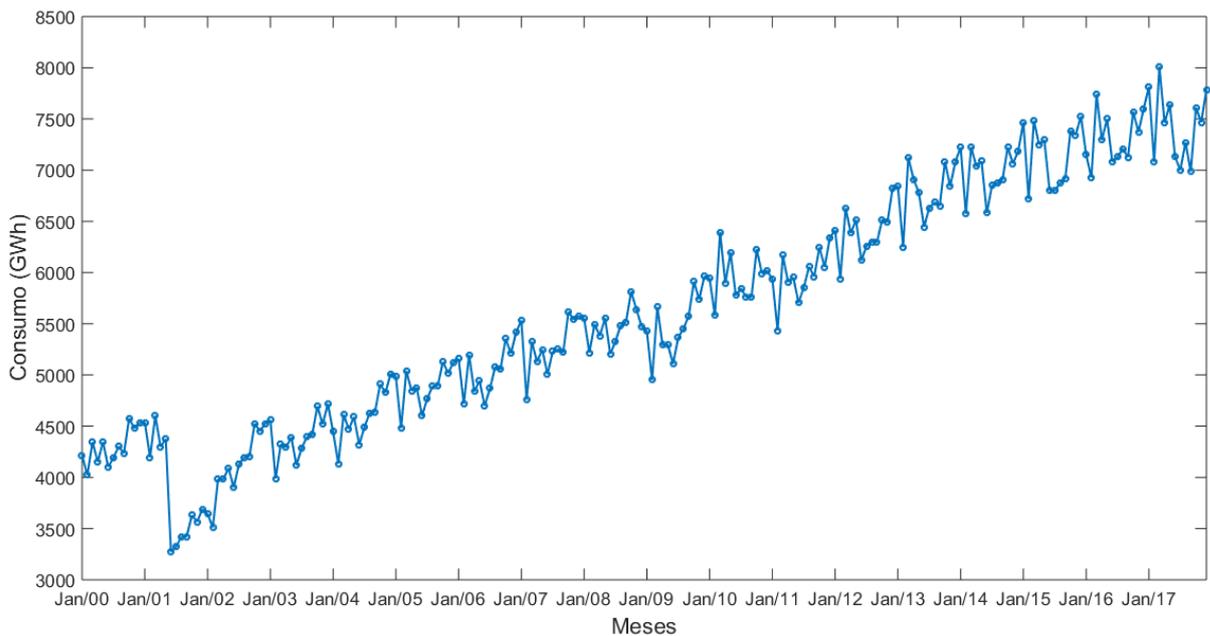


Figura 2.1 Série temporal do consumo mensal de energia elétrica da região Nordeste.

Uma série temporal também pode ser vista como uma realização ou trajetória de um processo estocástico, ou seja, uma observação, dentre infinitas possíveis, de um experimento aleatório. Neste sentido, o fenômeno que gerou a série é descrito por uma distribuição de probabilidade, da qual se deseja conhecer as propriedades. Os conceitos básicos de processos estocásticos serão apresentados na seção 2.2.

As séries temporais podem ser discretas ou contínuas. São discretas quando as observações são realizadas em instantes de tempo específicos, que, em geral, estão igualmente espaçados no tempo por determinado período. No caso da Figura 2.1, as observações são mensais. Quando as observações são realizadas ininterruptamente, então a série temporal resultante é dita contínua.

Um dos principais objetivos da análise de séries temporais é identificar um modelo que se adeque aos dados, o qual poderá ser utilizado para gerar previsões. Esta análise pode ser feita no domínio do tempo (análise temporal) ou no domínio da frequência (análise espectral).

A análise no domínio do tempo tem como objetivo observar a magnitude do evento que ocorre em determinado instante e na relação entre as observações. A ferramenta utilizada é a função de autocorrelação (FAC). Na análise espectral, o objetivo é observar a frequência com que os eventos ocorrem. Neste caso, uma das ferramentas utilizadas é a transformada de Fourier.

Segundo Ballini (2000), as duas formas de análise não são alternativas, mas complementares, sendo que a análise no domínio da frequência é conveniente na procura de características determinísticas, e a análise no domínio do tempo, na de processos não determinísticos.

Os objetivos da análise de séries temporais, segundo Morettin e Tolo (1981), são:

- a) Investigar o mecanismo gerador da série temporal;
- b) Fazer previsões de valores futuros da série, que podem ser de curto, médio ou longo prazo;
- c) Descrever apenas o comportamento da série, por meio da construção de gráficos, verificação da existência de tendências, ciclos e variações sazonais, construção de histogramas e diagramas de dispersão.
- d) Procurar periodicidades relevantes nos dados, o que pode ser feito por meio de análise espectral.

A análise clássica de séries temporais consiste, basicamente, em decompor a série em três componentes: de tendência, sazonal e aleatória. Esta decomposição sugere a representação de uma série  $Y_t$  de acordo com a equação:

$$Y_t = T_t + S_t + a_t, \quad (2.1)$$

em que  $T_t$  é a componente de tendência,  $S_t$  é a componente sazonal e  $a_t$  é a componente aleatória.

A tendência é a componente da série que descreve seu comportamento crescente ou decrescente. Para obter  $T_t$ , deve-se ajustar aos dados uma função do tempo, que pode ser polinomial, exponencial ou outra função suave do tempo (MORETTIN E TOLOI, 1981). A Figura 2.2 apresenta a série de consumo de energia elétrica da região Nordeste com sua linha de tendência do tipo linear.

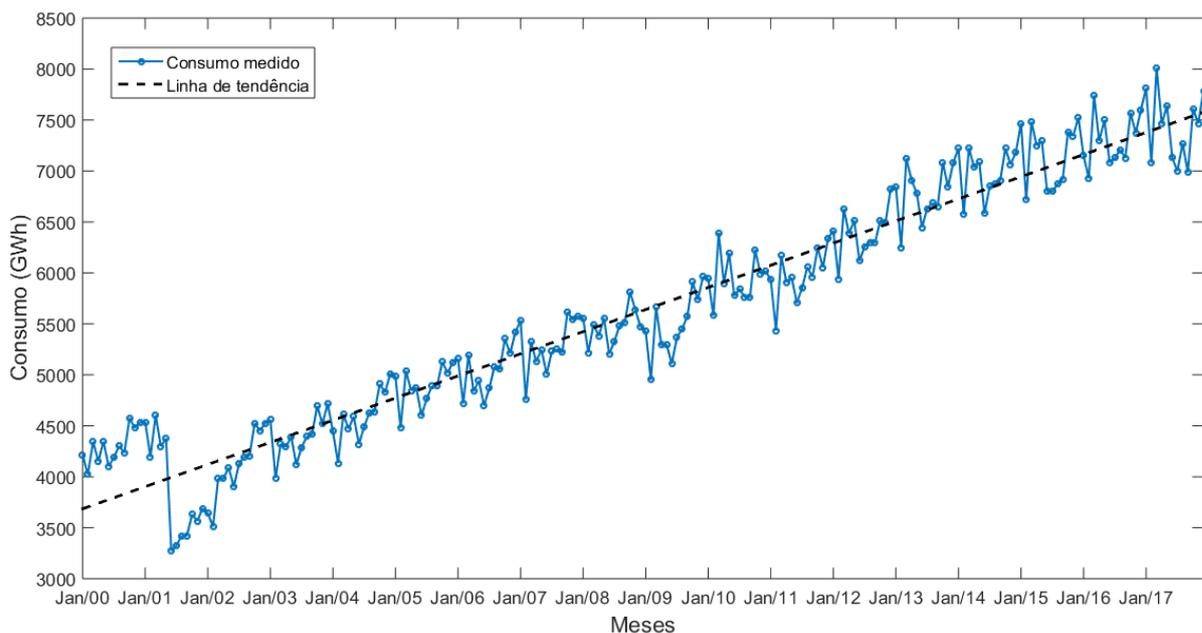


Figura 2.2 Tendência da série de consumo de energia elétrica da região Nordeste.

A componente sazonal descreve a característica periódica da série. Brockwell e Davis (2002) apresentam a expressão 2.2 a seguir para obtenção desta componente, dado que a série tem período  $N = 2q + 1$ , se  $N$  for ímpar, ou  $N = 2q$ , se  $N$  for par, e o número de observações da série é  $n$ :

$$S_k = w_k - \frac{1}{N} \sum_{i=1}^N w_i, \quad k = 1, \dots, N, \quad (2.2)$$

em que, para cada  $k$ ,  $w_k$  é a média dos desvios, calculados pela expressão:

$$Y_{k+jN} - T_{k+jN}, \quad q < k + jN \leq n - q. \quad (2.3)$$

Para qualquer  $k > N$ ,  $S_k = S_{k-N}$ . Desta forma, obteve-se a componente sazonal da série de consumo de energia elétrica da região Nordeste, a qual apresenta periodicidade de 12 meses ( $N = 12$ ), conforme pode ser observado na Figura 2.3 a seguir.

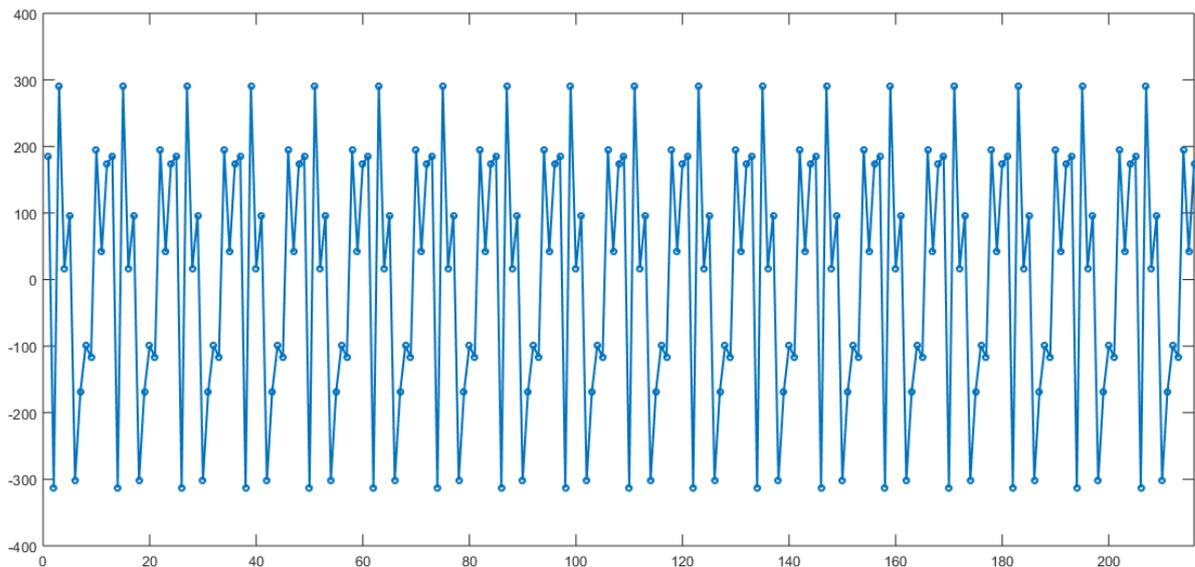


Figura 2.3 Componente sazonal da série de consumo de energia elétrica da região Nordeste.

Por fim, a componente aleatória  $a_t$  é obtida subtraindo-se as componentes  $T_t$  e  $S_t$  da série original. A suposição é que  $a_t$  seja um processo estocástico puramente aleatório, ou seja, um ruído branco, ou um processo estacionário, com média zero e variância constante (MORETTIN E TOLOI, 1981). Para a série temporal em estudo, a componente aleatória pode ser observada na Figura 2.4.

Deste modo, remover as componentes de tendência e sazonalidade é uma forma de tornar uma série estacionária. Essa transformação é necessária devido ao fato de que a maioria dos procedimentos de análise estatística de séries temporais supõem que elas sejam estacionárias (MORETTIN E TOLOI, 1981). Um procedimento

mais comum e útil consiste em diferenciar a série sucessivamente até que se torne estacionária. A primeira diferença é definida por:

$$\nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t, \quad (2.4)$$

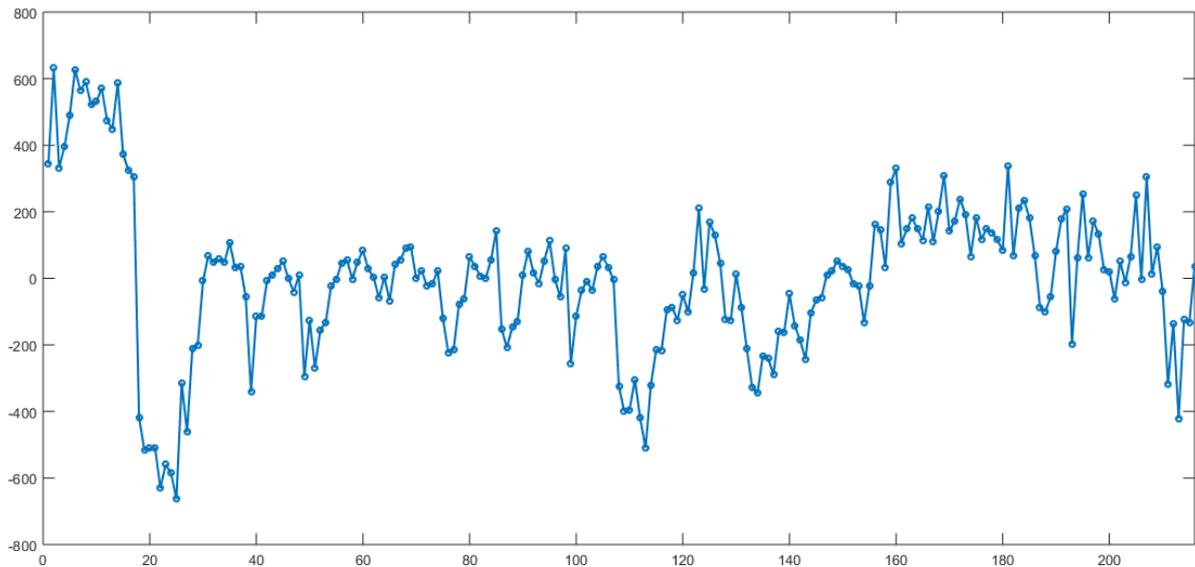


Figura 2.4 Componente aleatória da série de consumo de energia elétrica da região Nordeste.

em que  $\nabla$  é chamado de operador diferença e  $B$  é o operador atraso. Assim, a  $k$ -ésima diferença é dada por:

$$\nabla^k Y_t = \nabla(\nabla^{k-1} Y_t) = (1 - B)^k Y_t. \quad (2.5)$$

Para  $k = 2$ , tem-se:

$$\nabla^2 Y_t = \nabla(\nabla Y_t) = (1 - B)(1 - B)Y_t = (1 - 2B + B^2)Y_t = Y_t - 2Y_{t-1} + Y_{t-2}. \quad (2.6)$$

Para uma série não sazonal, tomar a primeira ou a segunda diferença, em geral, é suficiente para tornar a série estacionária (MORETTIN E TOLOI, 1981). Esta afirmação reside no fato de que muitas séries podem ser bem aproximadas por uma função polinomial de grau razoavelmente baixo (BROCKWELL, 2002). Já para uma série sazonal de período  $N$ , deverá ser aplicado à série o operador diferença de atraso  $N$ ,  $\nabla_N$ , definido por:

$$\nabla_N Y_t = Y_t - Y_{t-N}. \quad (2.7)$$

Aplicando o operador  $\nabla_N$  ao modelo da equação 2.1, obtém-se:

$$\nabla_N Y_t = T_t - T_{t-N} + S_t - S_{t-N} + a_t - a_{t-N}. \quad (2.8)$$

Como  $S_t = S_{t-N}$ , a equação 2.8 fica:

$$\nabla_N Y_t = T_t - T_{t-N} + a_t - a_{t-N}. \quad (2.9)$$

A série  $\nabla_N Y_t$  resultante ainda contém a componente de tendência, que pode ser removida aplicando-se o operador diferença apresentado anteriormente (equações 2.4 e 2.5).

## 2.2 Conceitos básicos de processos estocásticos

Matematicamente, um processo estocástico é definido como uma família  $Y = \{Y_t, t \in T\}$ , tal que para cada  $t \in T$ ,  $Y_t$  é uma variável aleatória. O conjunto  $T$  pode ser o conjunto dos números inteiros positivos,  $\mathbb{Z}_+$ , se a série for discreta, ou o conjunto dos números reais positivos,  $\mathbb{R}_+$ , se a série for contínua.

Na prática, no caso da análise de séries temporais, é impossível fazer mais do que uma observação em cada tempo  $t$ . Desta forma, tem-se apenas uma realização do processo estocástico e, para determinado  $t \in T$ , apenas uma observação da variável aleatória  $Y_t$  (EHLERS, 2009).

### 2.2.1 Estacionariedade

O conceito mais abordado na literatura a respeito de análise de séries temporais é o de estacionariedade. De modo geral, um processo estocástico é estacionário se suas propriedades estatísticas não se alteram no tempo, de modo que a escolha de uma origem dos tempos não é importante (BALLINI, 2000; MORETTIN E TOLOI, 1981). Ou seja, as características de  $Y_{t+\tau}$ , para todo  $\tau$ , são as mesmas de

$Y_t$ . Há dois tipos de estacionariedade: estrita (ou forte) e fraca (ou ampla, ou de segunda ordem).

**Definição 2.1** Um processo estocástico é *estritamente estacionário* se a *distribuição de probabilidade conjunta* associada a  $n$  observações  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$  é a mesma que a associada a  $n$  observações  $Y_{t_1+\tau}, Y_{t_2+\tau}, \dots, Y_{t_n+\tau}$ , para quaisquer  $\tau$  e  $n \geq 1$ .

Ou seja, um deslocamento no tempo de  $\tau$  com relação a origem não tem efeito sobre a distribuição de probabilidade conjunta (EHLERS, 2009). Isto significa, em particular, que todas as distribuições unidimensionais, isto é, quando  $n = 1$ , são invariantes no tempo. Deste modo, o valor esperado,  $\mu_t = E(Y_t)$ , e a variância,  $\sigma_t^2 = E[(Y_t - \mu_t)^2]$ , são constantes, para todo  $t \in T$ :

$$\begin{aligned}\mu_t &= \mu, \\ \sigma_t^2 &= \sigma^2.\end{aligned}\tag{2.10}$$

**Definição 2.2** Um processo estocástico é *estacionário no sentido amplo* ou *fracamente estacionário* se:

1.  $\mu_t = E(Y_t) = \mu$  é constante para todo  $t \in T$ ;
2.  $\gamma(t_1, t_2) = Cov(Y_{t_1}, Y_{t_2})$  é uma função que só depende de  $t_2 - t_1$ .

### 2.2.2 Funções de autocovariância e autocorrelação

Dado uma série estacionária com valor esperado  $\mu$ , a covariância entre uma observação  $Y_t$  e outra observação  $Y_{t+h}$  é chamada de autocovariância *no atraso*  $h$ , definida por:

$$\gamma_h = Cov(Y_t, Y_{t+h}) = E[(Y_t - \mu)(Y_{t+h} - \mu)].\tag{2.11}$$

O conjunto dos valores  $\gamma_h$ ,  $h = 0, 1, 2, \dots$ , em função de  $h$ , é chamado de função de autocovariância. A autocovariância no atraso  $h = 0$  é a variância da série temporal.

O coeficiente de autocorrelação no atraso  $h$  é definido por:

$$\rho_h = \frac{Cov(Y_t, Y_{t+h})}{Var(Y_t)} = \frac{\gamma_h}{\gamma_0}. \quad (2.12)$$

O conjunto dos valores  $\rho_h$ ,  $h = 0, 1, 2, \dots$ , em função de  $h$ , é chamado de função de autocorrelação (FAC). Da equação 2.12, nota-se que a função de autocorrelação é adimensional, ou seja, não depende da escala de medida da série temporal. Além disso,  $\rho_0 = 1$ .

Para uma série temporal estacionária com  $n$  observações,  $Y_1, Y_2, \dots, Y_n$ , a função de autocorrelação amostral (FAC amostral) é definida por:

$$r_h = \frac{c_h}{c_0}, \quad h = 0, 1, 2, \dots, n-1, \quad (2.13)$$

em que,

$$c_h = \frac{1}{n} \sum_{t=1}^{n-h} (Y_t - \bar{Y})(Y_{t+h} - \bar{Y}), \quad h = 0, 1, 2, \dots, n-1, \quad (2.14)$$

sendo que  $c_h$  é a estimativa do coeficiente de autocovariância,  $\gamma_h$ , e  $\bar{Y}$  é a média das observações.

A FAC é uma importante ferramenta na análise de séries temporais, especialmente no que diz respeito a sua estacionariedade. Uma forte FAC com decaimento lento sugere desvios da estacionariedade (MONTGOMERY *et al.*, 2008). Logo, uma série temporal estacionária deve apresentar uma FAC com decaimento rápido.

Na Figura 2.5 a seguir, pode-se observar a FAC amostral da série de consumo de energia elétrica da região Nordeste. O comportamento da FAC indica que a série não é estacionária.

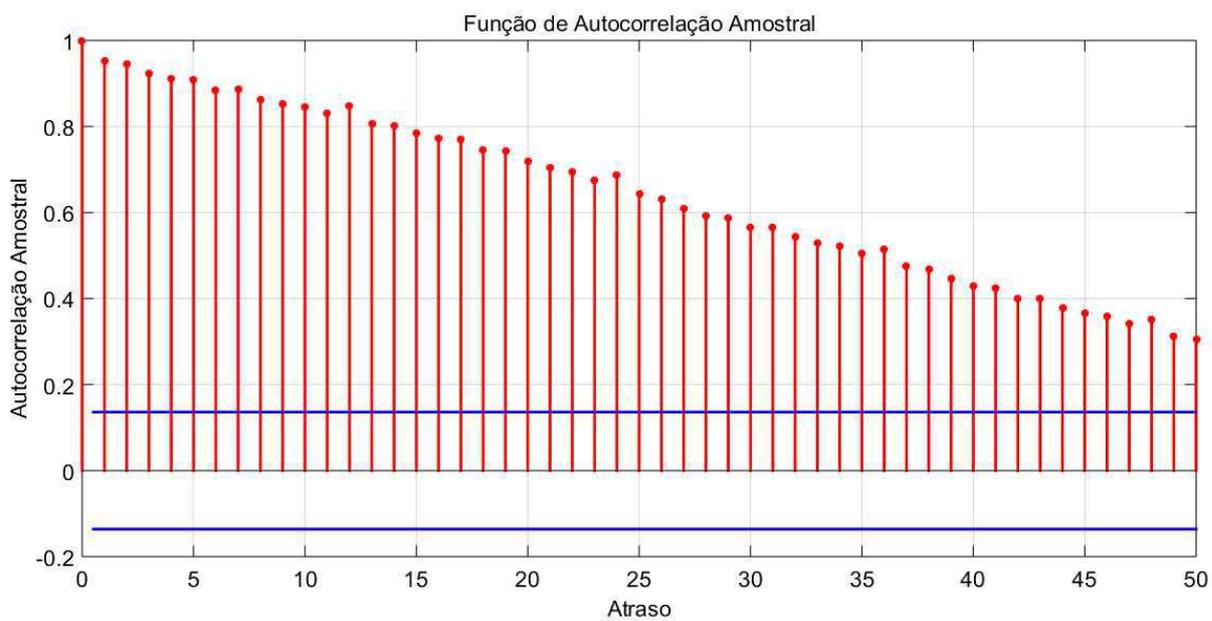


Figura 2.5 FAC amostral da série de consumo de energia elétrica da região Nordeste.

### 3. Modelos para previsão de séries temporais

Os principais objetivos da modelagem de séries temporais são descrever o comportamento da série e realizar previsões de valores futuros. De modo geral, um modelo que descreve uma série temporal é aquele que minimiza os erros entre os valores estimados e os dados observados. Entretanto, segundo Morettin e Tolo (1981), um modelo que descreve uma série não conduz, necessariamente, a um procedimento de previsão, pois, além do modelo, é necessário especificar uma função perda (como o erro quadrático médio).

Os métodos de previsão são classificados, basicamente, em dois tipos: qualitativos e quantitativos. Os métodos qualitativos são subjetivos, baseados na intuição e experiência do analista ou estudioso que esteja realizando a previsão. São utilizados quando não se dispõe de dados históricos, ou quando os dados existentes são insuficientes ou aparentam estar fora da realidade.

Já os métodos quantitativos são baseados em modelos matemáticos ou estatísticos, e requerem o uso de dados históricos. Segundo Montgomery *et al.* (2008), o modelo resume formalmente os padrões inerentes aos dados e expressa uma relação estatística entre os valores anteriores e atuais da série. Em seguida, o modelo é utilizado para projetar os padrões dos dados no futuro, ou seja, para extrapolar o comportamento passado e atual para o futuro.

A seguir, serão apresentados os principais modelos utilizados na previsão de séries temporais, além das Redes Neurais Artificiais (RNA), que vêm sendo amplamente utilizadas na realização de previsões em diversos setores.

#### 3.1 Modelos regressivos

Os modelos regressivos relacionam os dados observados de uma série temporal com uma ou mais variáveis de regressão,  $x_1, \dots, x_n$ , em que uma dessas variáveis pode ser o tempo, ou as próprias variáveis de regressão são funções do tempo. De modo geral, um modelo regressivo é representado por:

$$y_t = f(x_{1,t}, \dots, x_{n,t}). \quad (3.1)$$

Esta relação pode ser linear, polinomial, senoidal, etc. Por exemplo, o modelo de regressão linear é dado por:

$$y_t = \alpha_0 + \alpha_1 x_{1,t} + \dots + \alpha_n x_{n,t}. \quad (3.2)$$

Os parâmetros  $\alpha_0, \alpha_1, \dots, \alpha_n$  podem ser estimados por meio do método dos mínimos quadrados, o qual será apresentado sucintamente no capítulo 4 deste trabalho. Para realizar previsões com este tipo de modelo, basta conhecer os valores futuros das variáveis de regressão.

### 3.2 Modelos de suavização exponencial

Uma série temporal pode ser vista como um conjunto de dados que possui duas componentes: o sinal, que apresenta um comportamento padrão, e o ruído. A suavização ou alisamento exponencial é uma técnica utilizada para separar estas componentes, atuando como um filtro. A Figura 3.1 representa como se dá o processo de suavização.

O modelo mais comum de suavização é o de média móvel simples, o qual consiste em calcular a média aritmética das  $r$  observações mais recentes, conforme a equação:

$$M_t = \frac{Y_t + Y_{t-1} + \dots + Y_{t-r+1}}{r}. \quad (3.3)$$

A escolha da constante  $r$  é a maior problemática deste modelo. Se  $r$  for um valor alto, o modelo reagirá lentamente às variações da série, tendendo para um valor constante e igual a média aritmética de todas as observações, quando  $r = n$ . Já para  $r$  pequeno, o modelo reage rapidamente às variações da série, sendo que para  $r = 1$ , os valores do modelo coincidem com as próprias observações. Um procedimento sugerido por Morettin e Toloí (1981) é selecionar o valor de  $r$  que forneça a melhor previsão de um passo à frente das observações obtidas anteriormente.

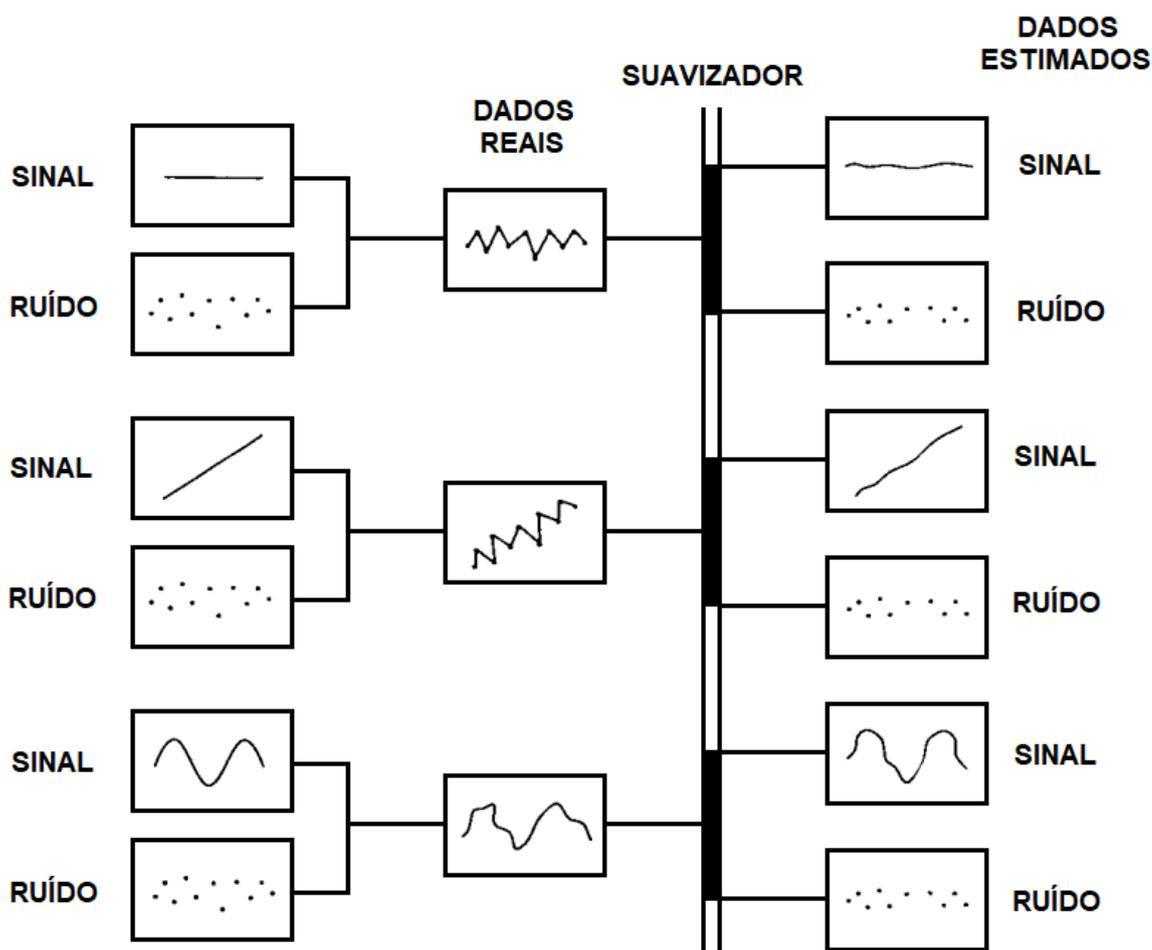


Figura 3.1 Processo de suavização de dados (Adaptado de: MONTGOMERY *et al.*, 2008).

Este modelo é adequado para séries localmente constantes. A previsão dos valores futuros da série pode ser feita utilizando-se a última média móvel calculada (MORETTIN E TOLOI, 1981).

### 3.2.1 Modelo de suavização exponencial simples

O modelo de suavização exponencial simples é definido por:

$$\bar{Y}_t = \alpha Y_t + (1 - \alpha)\bar{Y}_{t-1}, \quad \bar{Y}_0 = Y_1, \quad t = 1, \dots, n, \quad (3.4)$$

ou

$$\begin{aligned}\bar{Y}_t &= \alpha \sum_{k=0}^{t-1} (1-\alpha)^k Y_{t-k} + (1-\alpha)^t \bar{Y}_0, \quad t = 1, \dots, n, \\ \bar{Y}_t &= \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \dots,\end{aligned}\tag{3.5}$$

em que  $\bar{Y}_t$  é denominado valor exponencialmente suavizado e  $\alpha$  é a constante de suavização. Se as observações anteriores devem ser descontadas de forma geometricamente decrescente, então deve-se ter  $|\alpha| \leq 1$  (MONTGOMERY *et al.*, 2008). Para a escolha do  $\alpha$ , Morettin e Toloí (1981) sugerem o mesmo procedimento de escolha do  $r$ .

O método para previsão é o mesmo utilizado para o modelo de média móvel simples, ou seja, a previsão dos valores futuros da série é o último valor exponencialmente suavizado calculado. Outros modelos de suavização exponencial podem ser encontrados na literatura (MORETTIN e TOLOI, 1981; MONTGOMERY *et al.*, 2008), com aplicação às séries com tendência e séries sazonais.

### 3.3 Modelos lineares para séries temporais estacionárias

Segundo o teorema de decomposição de Wold (1938), qualquer processo estacionário de segunda ordem pode ser representado por:

$$\begin{aligned}Y_t &= \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \\ &= \mu + \sum_{i=0}^{\infty} \psi_i B^i \varepsilon_t \\ &= \mu + \Psi(B) \varepsilon_t,\end{aligned}\tag{3.6}$$

em que  $B$  é o operador atraso, definido na seção 2.1,  $\varepsilon_t$  é um ruído branco, com  $E(\varepsilon_t) = 0$ ,  $\sigma_t^2 = \sigma^2$ , e  $\psi_i$  é uma sequência de constantes com  $\sum_{i=0}^{\infty} \psi_i^2 < \infty$ . Sendo  $Y_t$  uma série temporal estacionária,  $\mu$  é a média da série. Se  $Y_t$  não for estacionária,  $\mu$  não tem significado e representa apenas o nível da série (MORETTIN e TOLOI, 1981).

Este modelo é chamado de média móvel infinita e serve como uma classe geral de modelos para qualquer série temporal estacionária (MONTGOMERY *et al.*, 2008).

O operador  $\Psi(B)$  pode ser pensado como um filtro linear, que ao ser aplicado à sua entrada um ruído branco  $\varepsilon_t$ , produz a saída  $Y_t$ . Desta forma, sendo a entrada do filtro linear um processo estacionário, então sua saída também será um processo estacionário (BROCKWELL, 2002).

Existem três casos especiais para o modelo da equação 3.6:

1. Autorregressivo de ordem finita,  $AR(p)$ , do inglês *autoregressive*, modelo em que os valores passados da série são considerados.
2. Média móvel de ordem finita,  $MA(q)$ , do inglês *moving average*, no qual  $\psi_i = 0$  para  $i > q$ .
3. Misto autorregressivo de média móvel,  $ARMA(p, q)$ .

### 3.3.1 Modelos autorregressivos – $AR(p)$

Um modelo autorregressivo,  $AR(q)$ , é expresso por:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t, \quad (3.7)$$

ou

$$\Phi(B)Y_t = \varepsilon_t, \quad (3.8)$$

em que  $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ . O termo  $\varepsilon_t$  representa as quantidades que não são explicadas pelos valores passados da série. O modelo é chamado de autorregressivo porque pode ser visto como um modelo de regressão, em que as “variáveis de regressão” são valores passados da variável de interesse,  $Y_t$ .

Para  $p = 1$ , tem-se o modelo  $AR(1)$ , expresso por:

$$Y_t = \phi Y_{t-1} + \varepsilon_t. \quad (3.9)$$

### 3.3.2 Modelos de médias móveis - $MA(q)$

No modelo  $MA(q)$ , apenas os  $q$  últimos valores de  $\varepsilon_t$  na equação 3.6 são considerados, ou seja:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \quad (3.10)$$

ou,

$$Y_t = \Theta(B)\varepsilon_t, \quad (3.11)$$

em que  $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ . Para  $q = 1$ , tem-se o modelo  $MA(1)$ , expresso por:

$$Y_t = \varepsilon_t - \theta \varepsilon_{t-1}. \quad (3.12)$$

### 3.3.3 Modelos mistos autorregressivos de médias móveis - $ARMA(p, q)$

O modelo  $ARMA(p, q)$  é uma combinação dos modelos  $AR(p)$  e  $MA(q)$ . É definido por:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \quad (3.13)$$

ou,

$$\Phi(B)Y_t = \Theta(B)\varepsilon_t. \quad (3.14)$$

O modelo  $ARMA(1,1)$ , com  $p = 1$ ,  $q = 1$ , é dado por:

$$Y_t = \phi Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (3.15)$$

### 3.4 Modelos autorregressivos integrados de médias móveis - $ARIMA(p, q, d)$

Os modelos  $ARIMA(p, q, d)$ , do inglês *autoregressive integrated moving average*, foram propostos por Box e Jenkins (1970), e são utilizados para modelar

séries não estacionárias.

Seja  $Y_t$  uma série não estacionária, tal que se diferenciada por  $d$  vezes consecutivas torna-se estacionária, então a série  $Z_t$ , definida por:

$$Z_t = \nabla^d Y_t, \quad (3.16)$$

é estacionária. Deste modo,  $Z_t$  pode ser representado por um modelo  $ARMA(p, q)$ , ou seja,

$$\Phi(B)Z_t = \Theta(B)\varepsilon_t. \quad (3.17)$$

Se  $Z_t$  é uma diferença de  $Y_t$ , então  $Y_t$  é uma integral de  $Z_t$ . Logo,  $Y_t$  pode ser modelado por um modelo  $ARIMA(p, q, d)$ . Substituindo a equação 3.16 na 3.17, tem-se:

$$\begin{aligned} \Phi(B)\nabla^d Y_t &= \Theta(B)\varepsilon_t, \\ \Phi(B)(1 - B)^d Y_t &= \Theta(B)\varepsilon_t. \end{aligned} \quad (3.18)$$

Um exemplo desta classe de modelos é o  $ARIMA(0,1,1)$ , com  $p = 0$ ,  $q = 1$ ,  $d = 1$ . Tem-se que:

$$\begin{aligned} \Phi(B) &= 1, \\ \Theta(B) &= 1 - \theta B, \\ (1 - B)Y_t &= (1 - \theta B)\varepsilon_t, \\ Y_t &= Y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}. \end{aligned} \quad (3.19)$$

### 3.4.1 Modelos ARIMA sazonais

Os modelos ARIMA sazonais, também chamados de SARIMA, são utilizados para modelar séries que apresentam sazonalidade. Seja  $Y_t$  uma série temporal que possui uma componente sazonal. Podemos escrever  $Y_t$  como:

$$Y_t = S_t + N_t, \quad (3.20)$$

em que  $S_t$  é uma componente determinística com periodicidade  $s$ , e  $N_t$  é um processo estacionário que pode ser representado por um modelo  $ARMA(p, q)$ . Tem-se que:

$$S_t - S_{t-s} = (1 - B^s)S_t = 0. \quad (3.21)$$

Multiplicando a equação 3.21 por  $(1 - B^s)$ , obtém-se:

$$\begin{aligned} (1 - B^s)Y_t &= (1 - B^s)S_t + (1 - B^s)N_t, \\ (1 - B^s)Y_t &= (1 - B^s)N_t. \end{aligned} \quad (3.22)$$

Como  $N_t$  pode ser modelado como um processo  $ARMA(p, q)$ , então:

$$\begin{aligned} \Phi(B)N_t &= \Theta(B)\varepsilon_t, \\ N_t &= \frac{\Theta(B)}{\Phi(B)}\varepsilon_t. \end{aligned} \quad (3.23)$$

Substituindo a equação 3.23 na equação 3.22, obtém-se:

$$\begin{aligned} (1 - B^s)Y_t &= (1 - B^s)\frac{\Theta(B)}{\Phi(B)}\varepsilon_t, \\ \Phi(B)(1 - B^s)Y_t &= (1 - B^s)\Theta(B)\varepsilon_t. \end{aligned} \quad (3.24)$$

O modelo da equação 3.24 considera apenas a sazonalidade de  $Y_t$ . A forma geral do modelo sazonal ARIMA, chamado de modelo sazonal multiplicativo de ordem  $(p, q, d) \times (P, Q, D)_s$ , é dado por (MORETTIN e TOLOI, 1981):

$$\Phi^*(B^s)\Phi(B)(1 - B)^d(1 - B^s)^D Y_t = \Theta^*(B^s)\Theta(B)\varepsilon_t, \quad (3.25)$$

em que:

- $\Phi^*(B^s) = 1 - \phi_1^*B^s - \phi_2^*B^{2s} - \dots - \phi_p^*B^{ps}$ ;
- $\Theta^*(B^s) = 1 - \theta_1^*B^s - \theta_2^*B^{2s} - \dots - \theta_Q^*B^{Qs}$ ;
- $\Phi(B)$  e  $\Theta(B)$  seguem as definições dadas anteriormente.

A exemplo, para o modelo  $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$ , tem-se:

$$\begin{aligned}
 \Phi^*(B^{12}) &= 1, \\
 \Phi(B) &= 1, \\
 \Theta^*(B^{12}) &= 1 - \theta^*B^{12}, \\
 \Theta(B) &= 1 - \theta B, \\
 (1 - B)(1 - B^{12})Y_t &= (1 - \theta^*B^{12})(1 - \theta B)\varepsilon_t, \\
 (1 - B - B^{12} - B^{13})Y_t &= (1 - \theta^*B^{12} - \theta B + \theta^*\theta B^{13})\varepsilon_t, \\
 Y_t &= Y_{t-1} + Y_{t-12} + Y_{t-13} + \varepsilon_t - \theta\varepsilon_{t-1} - \theta^*\varepsilon_{t-12} - \theta^*\theta\varepsilon_{t-13}.
 \end{aligned} \tag{3.26}$$

### 3.5 Redes neurais artificiais

Uma rede neural artificial (RNA) é uma estrutura computacional cuja filosofia é baseada no sistema nervoso biológico. Pode ser definida como um processador distribuído em paralelo que é composto de unidades de processamento simples, chamados de neurônios, os quais são capazes de armazenar conhecimento pela experiência (BALLINI, 2000; HAYKIN, 2009). A Figura 3.2 a seguir apresenta o modelo de um neurônio artificial na forma de um diagrama de blocos.

Os elementos básicos do neurônio são (HAYKIN, 2009):

1. Conexões, chamadas de sinapses, cada uma caracterizada por um peso  $\omega_k$ . Estas conexões ligam as entradas  $x_1, \dots, x_m$  ao neurônio, multiplicando-as pelos pesos  $\omega_{k1}, \dots, \omega_{km}$ , respectivamente.
2. Um somador, para somar os sinais de entrada ponderados pelos pesos sinápticos. Ou seja, esta operação é uma combinação linear.
3. Uma função de ativação, também chamada de função de esmagamento, para limitar a amplitude da saída.
4. Um parâmetro externo, chamado de *bias*, o qual tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, como mostra a Figura 3.3 (os valores  $u_k$  e  $v_k$  serão definidos na equação 3.27).

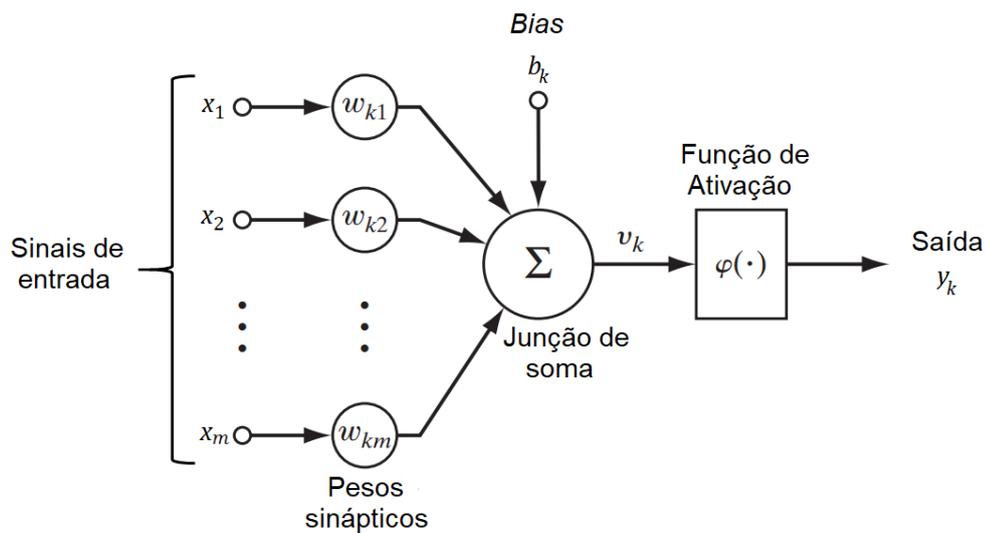


Figura 3.2 Modelo de um neurônio artificial, chamado  $k$  (Adaptado de: HAYKIN, 2009).

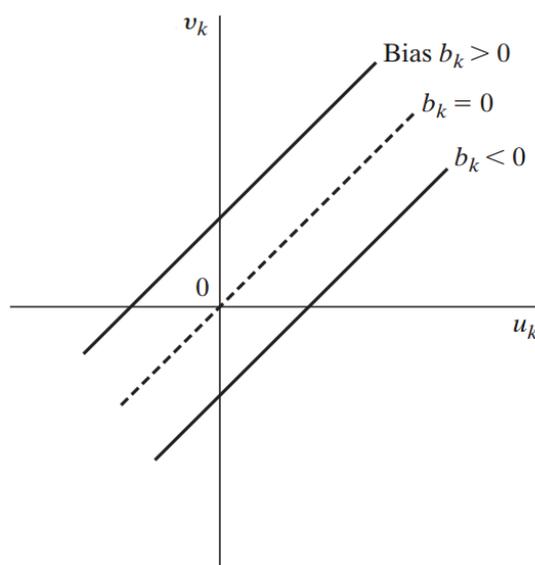


Figura 3.3 Efeito do *bias* (Adaptado de: HAYKIN, 2009).

Matematicamente, o neurônio  $k$  da Figura 3.2 pode ser descrito por:

$$\begin{aligned}
 u_k &= \sum_i^m \omega_{ki} x_i, \\
 v_k &= u_k + b_k, \\
 y_k &= \varphi(v_k),
 \end{aligned}
 \tag{3.27}$$

em que  $u_k$  é a combinação linear das entradas.

### 3.5.1 Funções de ativação

A função de ativação determina a saída  $y_k$  do neurônio em termos de  $v_k$ . Ela pode ser de dois tipos:

1. Função degrau (Figura 3.4), também chamada de função de Heaviside, definida por:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (3.28)$$

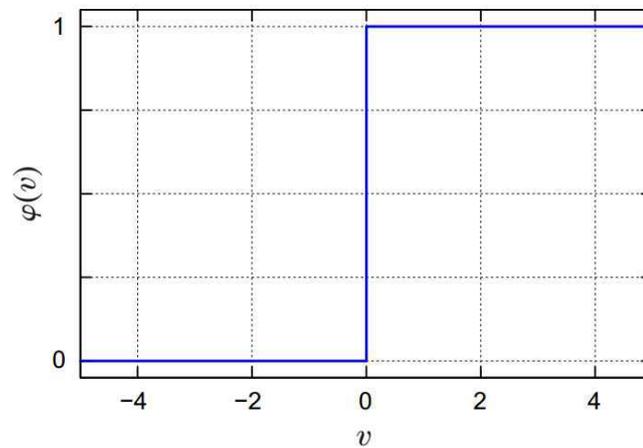


Figura 3.4 Função degrau (Adaptado de: KRIESEL, 2007).

2. Funções sigmóides, que são funções em forma de “S”. Exemplos de função sigmoide são a função logística (Figura 3.5), definida por:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (3.29)$$

e a função tangente hiperbólica (Figura 3.6), definida por:

$$\varphi(v) = \tanh(v) = \frac{\exp(v) - \exp(-v)}{\exp(v) + \exp(-v)} \quad (3.30)$$

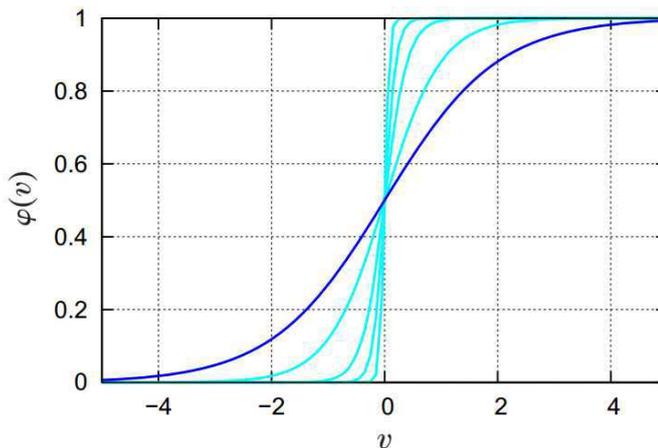


Figura 3.5 Formas da função logística para diferentes valores de  $a$  (Adaptado de: KRIESEL, 2007).

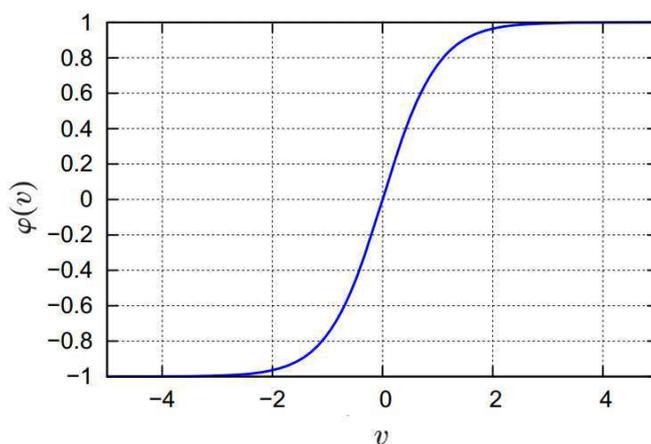


Figura 3.6 Função tangente hiperbólica (Adaptado de: KRIESEL, 2007).

### 3.5.2 Arquiteturas de redes neurais artificiais

As arquiteturas ou topologias de RNA's são divididas em dois tipos: *feedback* (ou recorrente) e *feedforward* (ou não recorrente). Uma RNA do tipo *feedback* possui, pelo menos, um circuito fechado de realimentação. Essa realimentação pode ser direta (também chamada de auto-recorrência), quando neurônios são ligados a si mesmos, ou indireta, quando neurônios tem suas saídas ligadas a neurônios de camadas anteriores (HAYKIN, 2009). A Figura 3.7 mostra um exemplo desta arquitetura.

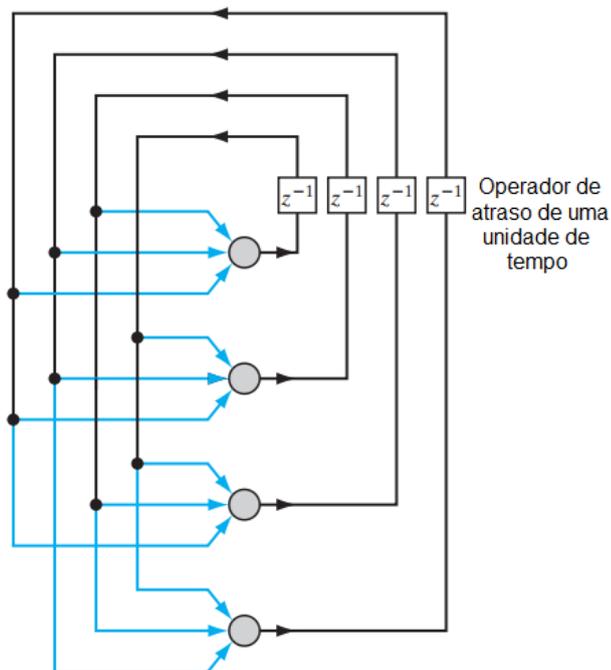


Figura 3.7 Rede neural recorrente (Adaptado de: HAYKIN, 2009).

Já nas RNA's do tipo *feedforward*, as camadas de neurônios são claramente separadas: uma camada de entrada, uma camada de saída e uma ou mais camadas de processamento que são invisíveis externamente (também chamadas de camadas escondidas). Conexões só são permitidas para os neurônios da camada seguinte (KRIESEL, 2007). Um exemplo desta topologia pode ser visto na Figura 3.8.

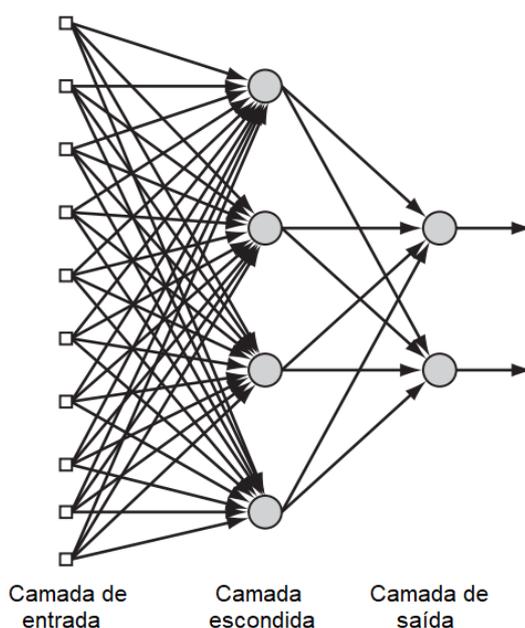


Figura 3.8 Rede neural não recorrente com uma camada escondida (Adaptado de: HAYKIN, 2009).

### 3.5.3 Aprendizagem de redes neurais artificiais

A característica mais importante das redes neurais é sua capacidade de generalização, ou seja, uma RNA é capaz de se adaptar a qualquer problema por meio de aprendizagem. O procedimento mais comum de aprendizagem é a realização de mudanças nos pesos sinápticos de acordo com regras, as quais podem ser formuladas como algoritmos (KRIESEL, 2007).

A aprendizagem pode ser de dois tipos: supervisionada e não supervisionada. Na aprendizagem supervisionada, existe um “professor” que fornece o conhecimento para a RNA, sendo este conhecimento representado por um conjunto de padrões de entrada-saída, também referido como conjunto de treinamento. Para cada entrada, o professor responde com um vetor de treinamento.

Os parâmetros da rede são ajustados sob a influência combinada do vetor de treinamento e do sinal de erro. O sinal de erro é definido como a diferença entre a resposta desejada e a resposta real da rede (HAYKIN, 2009). O treinamento é concluído quando o erro mínimo é atingido, e esta minimização se dá pelo ajuste dos pesos sinápticos.

Na aprendizagem não supervisionada, não existe um professor. O conjunto de treinamento consiste apenas em padrões de entrada, e a rede tenta, por si só, detectar semelhanças e gerar padrões (KRIESEL, 2007).

A tabela a seguir apresenta um resumo das características e utilização de cada um dos modelos apresentados neste capítulo, incluindo as RNA's.

Tabela 1. Principais características e utilização dos modelos de séries temporais.

<b>Método de previsão</b>	<b>Característica e utilização</b>
Suavização exponencial	Pesos maiores às demandas mais recentes (RITZMAN e KRAJEWSKI, 2004) Possui variações para inclusão de padrões de tendência e sazonalidade (RITZMAN e KRAJEWSKI, 2004; MAKRIDAKIS e WHEELWRIGHT, 1977)
Média móvel	Modelo simples (CORREA et al, 2001) Reduz o impacto da aleatoriedade dos dados (MAKRIDAKIS e WHEELWRIGHT, 1977)
Regressão Linear	Identifica o efeito da variável de previsão sobre a demanda do produto em análise (TUBINO, 2000) Usa-se quando os dados históricos são disponíveis e as relações entre o fator a ser previsto e outros fatores são conhecidos (RITZMAN e KRAJEWSKI, 2004)
Box-Jenkins	Considera que a série é gerada por processo estocástico e o método busca identificar este gerador. (MONTANES et al, 2002) Considera tendência e sazonalidade (MONTANES et al, 2002)
Redes Neurais	Consegue representar modelos não lineares (RAO e ALI, 2002; SILVA, 2003) Aproximador universal de função (HILL <i>et al.</i> , 1996) Necessita menor extensão de dados observados (HILL <i>et al.</i> , 1996)

Fonte: Adaptado de Branco e Sampaio (2008).

## 4. Metodologia

A proposta deste trabalho, além de realizar uma revisão bibliográfica acerca dos conceitos que envolvem a análise e a previsão de séries temporais, é a realização de um estudo de caso para previsão de carga da região Nordeste. Os dados utilizados para esta previsão foram o consumo mensal de energia elétrica da região, de janeiro de 2000 a dezembro de 2017 (Figura 2.1), disponibilizados pelo Operador Nacional do Sistema Elétrico (ONS) em sua página na Internet<sup>1</sup>.

Decidiu-se fazer uma comparação entre as previsões obtidas por um modelo clássico, especificamente, o modelo ARIMA sazonal, e as obtidas por uma rede neural artificial (RNA). Para fazer esta comparação, é necessário utilizar algum índice de desempenho de modo que seja possível identificar qual a melhor metodologia para este tipo de problema. Neste capítulo, será apresentado o passo-a-passo para a obtenção das previsões, como também os índices de desempenho utilizados.

### 4.1 Previsões com modelo ARIMA sazonal

A forma geral do modelo ARIMA sazonal é dado pela equação 3.25. A metodologia consiste, basicamente, em ajustar o modelo ao conjunto de dados disponível. A série temporal em estudo contém o histórico do consumo de energia elétrica da região Nordeste de 216 meses (18 anos). O ajuste do modelo foi feito com os dados dos 13 primeiros anos, e os dados dos 5 anos seguintes foram utilizados para validar as previsões, com horizontes de 1 ano, 2 anos e 5 anos.

A estratégia para a obtenção do modelo, sugerida por Box e Jenkins (1970), é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados. Os estágios do ciclo iterativo são (MORETTIN e TOLOI, 1981; CAMPOS, 2008):

- Especificação: uma classe de modelos é considerada para a análise.
- Identificação: um modelo é escolhido com base na análise de autocorrelações e autocorrelações parciais. Nessa etapa, é identificada a necessidade de diferenciar os dados para tornar a série estacionária e são calculadas as ordens dos

---

<sup>1</sup> <http://ons.org.br/paginas/resultados-da-operacao/historico-da-operacao>

processos autorregressivo ( $p$ ) e do processo de médias móveis ( $q$ ).

- Estimação: os parâmetros do modelo identificado são estimados, ou seja, os coeficientes AR e MA.
- Verificação: o modelo ajustado é verificado através de uma análise de resíduos.

Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação, como mostra o fluxograma a seguir.

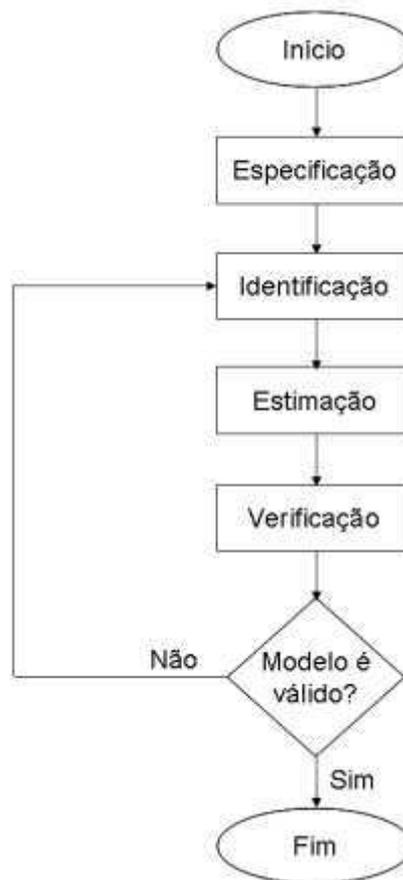


Figura 4.1 Etapas do processo de obtenção do modelo ARIMA (Fonte: CAMPOS, 2008).

Na etapa de identificação, é necessário analisar as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) da série. A tabela a seguir apresenta um resumo de como identificar as ordens  $p$  e  $q$  dos modelos AR, MA e ARMA por meio da FAC e da FACP. As demonstrações podem ser vistas no trabalho de Box e Jenkins (1970).

Tabela 2. Comportamento da FAC e FACP dos processos AR, MA e ARMA.

Modelo	FAC	FACP
$AR(p)$	Decai exponencialmente e/ou como uma senoide amortecida	Corta após o atraso $p$
$MA(q)$	Corta após o atraso $q$	Decai exponencialmente e/ou como uma senoide amortecida
$ARMA(p, q)$	Decai exponencialmente e/ou como uma senoide amortecida	Decai exponencialmente e/ou como uma senoide amortecida

Fonte: Adaptado de MONTGOMERY *et al.* (2008).

Na etapa de estimação, utilizou-se o método dos mínimos quadrados, que consiste em minimizar a soma dos quadrados das diferenças entre os dados medidos e a curva do modelo. Seja o erro (ou resíduo) associado a cada valor estimado dado por:

$$e_t = \hat{y}_t - y_t, \quad t = 1, 2, \dots, n, \quad (4.1)$$

em que  $\hat{y}_t$  é o valor estimado para  $y_t$  no tempo  $t$ , e  $n$  é o número de observações da série. Na forma vetorial:

$$\mathbf{e} = \hat{\mathbf{y}} - \mathbf{y}. \quad (4.2)$$

Deste modo, a função que se deseja minimizar é dada por:

$$f = \sum_{t=1}^n e_t^2 = \mathbf{e}'\mathbf{e}. \quad (4.3)$$

## 4.2 Previsões com redes neurais artificiais

Para implementação da rede neural artificial, utilizou-se o *software* Matlab®<sup>1</sup>, que possui uma aplicação (*toolbox*) específica para solução de problemas de modelagem e previsão de séries temporais utilizando RNA's, chamada de *Neural*

<sup>1</sup> Marca registrada de The MathWorks, Inc.

*Network Time Series Tool*. A Figura 4.2 a seguir apresenta, em forma de diagrama de blocos, a RNA utilizada neste trabalho.

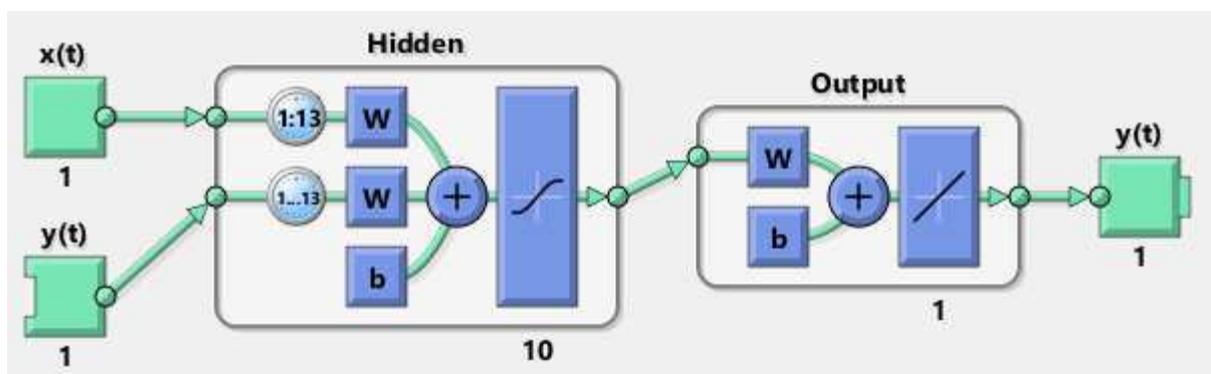


Figura 4.2 Diagrama de blocos da RNA simulada no MATLAB.

Aqui, a metodologia consiste, basicamente, em determinar os parâmetros da RNA, a saber: as entradas do modelo, número de neurônios na camada escondida, atrasos, distribuição dos dados entre treinamento e validação, e algoritmo de treinamento. As funções de ativação são pré-definidas no Matlab, e correspondem à tangente hiperbólica na camada escondida, e à uma função linear na camada de saída.

As entradas do modelo, observadas na Figura 4.2, são:  $y(t)$ , que é a série temporal a ser modelada, e  $x(t)$ , que representa uma variável de regressão para o modelo. Neste trabalho,  $y(t)$  é a série de consumo de energia elétrica da região Nordeste (Figura 2.1), considerando os dados dos 13 primeiros anos, e  $x(t)$  é uma série que indica a que mês se refere cada dado de  $y(t)$  (janeiro corresponde a  $x(t) = 1$ , fevereiro, a  $x(t) = 2$ , e assim por diante, até dezembro, correspondendo a  $x(t) = 12$ ).

O número de neurônios na camada escondida normalmente é obtido por tentativa e erro. Em alguns trabalhos, como o de Campos (2008), a metodologia utilizada para a escolha deste número é a realização de sucessivas simulações da RNA, variando-se o número de neurônios dentro de um intervalo, e o valor escolhido é aquele que apresentou o melhor desempenho, em termos de índices de desempenho específicos. Neste trabalho, utilizou-se o valor padrão do Matlab, que é de 10 neurônios na camada escondida.

Os atrasos definem a equação do modelo. Segundo Campos (2008), a escolha dos atrasos depende muito das características do processo em estudo. Como a série

em estudo neste trabalho tem sazonalidade de 12 meses, é essencial que um atraso igual a 12 seja incluído no modelo. Os demais atrasos foram escolhidos baseado na equação do modelo ARIMA sazonal obtido para a série em estudo.

Quanto a distribuição dos dados entre treinamento e validação, a literatura sugere que pelo menos 20% da base de dados seja separado para validar os modelos obtidos na fase de treinamento (BRAGA *et al.*, 2000). Com relação ao algoritmo de treinamento, entre os disponíveis no Matlab, utilizou-se o de Levenberg-Marquardt.

Para realizar as previsões, é necessário fazer uma modificação na estrutura da RNA, de modo que ela possa gerar previsões de um passo à frente, como mostra a Figura 4.3. Para obter as demais previsões, a estrutura da RNA foi realimentada sucessivamente com a previsão de um passo à frente até se obter o total de previsões desejado.

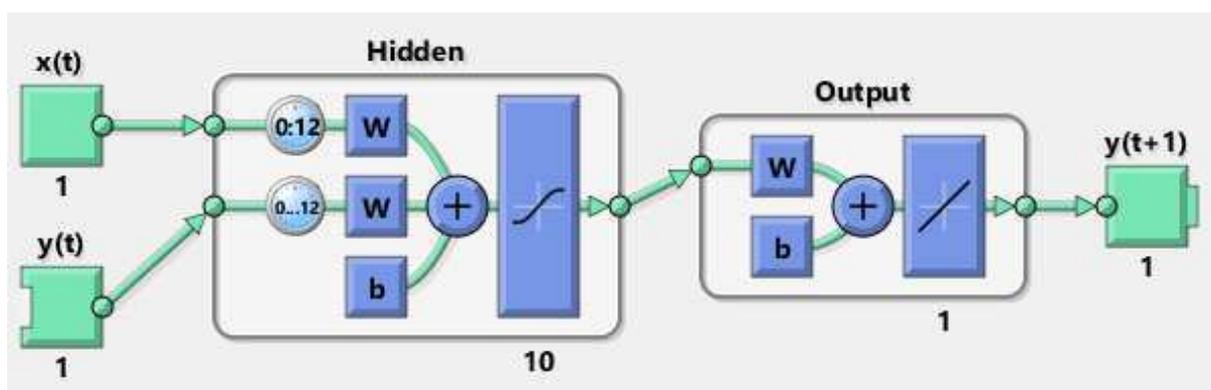


Figura 4.3 Estrutura da RNA para obter previsões de um passo à frente.

### 4.3 Índices de desempenho

Como dito anteriormente, um modelo que descreve uma série temporal é aquele que minimiza os erros entre os valores estimados pelo modelo e os dados medidos. A seguir, serão definidos os principais índices utilizados para medir os erros na modelagem de séries temporais. Nas equações,  $n$  é o número total de observações da série temporal,  $\hat{y}_t$  representa os valores estimados,  $y_t$ , os dados reais medidos e  $\bar{y}_t$ , o valor médio de  $y_t$ .

- Erro médio ( $ME$ , do inglês *mean error*)

$$ME = \frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t). \quad (4.4)$$

- Erro absoluto médio ( $MAE$ , do inglês *mean absolute error*)

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{y}_t - y_t|. \quad (4.5)$$

- Erro quadrático médio ( $MSE$ , do inglês *mean squared error*)

$$MSE = \frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2. \quad (4.6)$$

- Erro percentual médio ( $MPE$ , do inglês *mean percentage error*)

$$MPE = \frac{1}{n} \sum_{t=1}^n \frac{\hat{y}_t - y_t}{y_t} \times 100. \quad (4.7)$$

- Erro percentual absoluto médio ( $MAPE$ , do inglês *mean absolute percentage error*)

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \times 100 \right|. \quad (4.8)$$

Estes índices são medidas de erro, mas não indicam se um modelo se ajustou bem ou não a um conjunto de dados. Por isso, neste trabalho, utilizou-se o coeficiente de determinação,  $R^2$ , que dá uma medida da qualidade do ajuste de um modelo. Quando  $R^2 = 1$ , significa que o modelo conseguiu explicar 100% dos dados. O  $R^2$  é definido por:

$$R^2 = 1 - \frac{SSE}{SST},$$
$$SSE = \sum_{i=1}^m (\hat{y}_t - y_t)^2, \quad (4.9)$$
$$SST = \sum_{i=1}^m (y_t - \bar{y}_t)^2,$$

em quem  $SSE$  é a soma dos quadrados dos erros (do inglês *sum of squared errors*), que representa a variação total do que o modelo não consegue explicar, e  $SST$  é a soma dos quadrados totais (do inglês *total sum of squares*), que representa a variação total de  $y_t$  com relação a sua média.

## 5. Estudo de caso: Previsão de carga da região Nordeste

Para realizar a previsão de carga da região Nordeste, foram utilizados os procedimentos descritos no capítulo anterior. A série temporal considerada para realizar o ajuste dos modelos foi o consumo de energia elétrica da região Nordeste de janeiro de 2000 a dezembro de 2012, totalizando 156 meses. Os dados dos meses de janeiro de 2013 a dezembro de 2017 foram utilizados para validar as previsões.

### 5.1 Resultados obtidos com modelo ARIMA sazonal

O modelo ARIMA sazonal é capaz de representar séries que apresentam tendências e sazonalidade, como é o caso da série temporal em estudo. A forma geral do modelo foi dada na equação 3.25, no Capítulo 3. Antes de estimar os parâmetros do modelo, é necessário identificar a ordem das componentes AR e MA do modelo.

Pela análise da FAC amostral da série de consumo da região Nordeste (Figura 2.5), verificou-se que ela é não estacionária, justificando o uso da classe de modelos ARIMA. Além disso, a partir das análises expostas no Capítulo 2, observou-se que a série apresenta uma tendência linear, sugerindo que  $d = 1$  (componente de integração do modelo ARIMA). A periodicidade da série é de 12 meses, ou seja, no modelo multiplicativo, tem-se que  $s = 12$ .

Sendo a série não estacionária, é necessário derivá-la para remover suas componentes de tendência e sazonalidade. Após a derivação, espera-se que a análise das FAC e FACP da série resultante permita identificar a ordem dos processos AR e MA, conforme foi discutido na seção 4.1. A série de consumo de energia elétrica da região Nordeste foi derivada, e as funções de autocorrelação e de autocorrelação parcial amostrais da série derivada foram plotadas, conforme pode ser observado nas Figuras 5.1 e 5.2 a seguir.

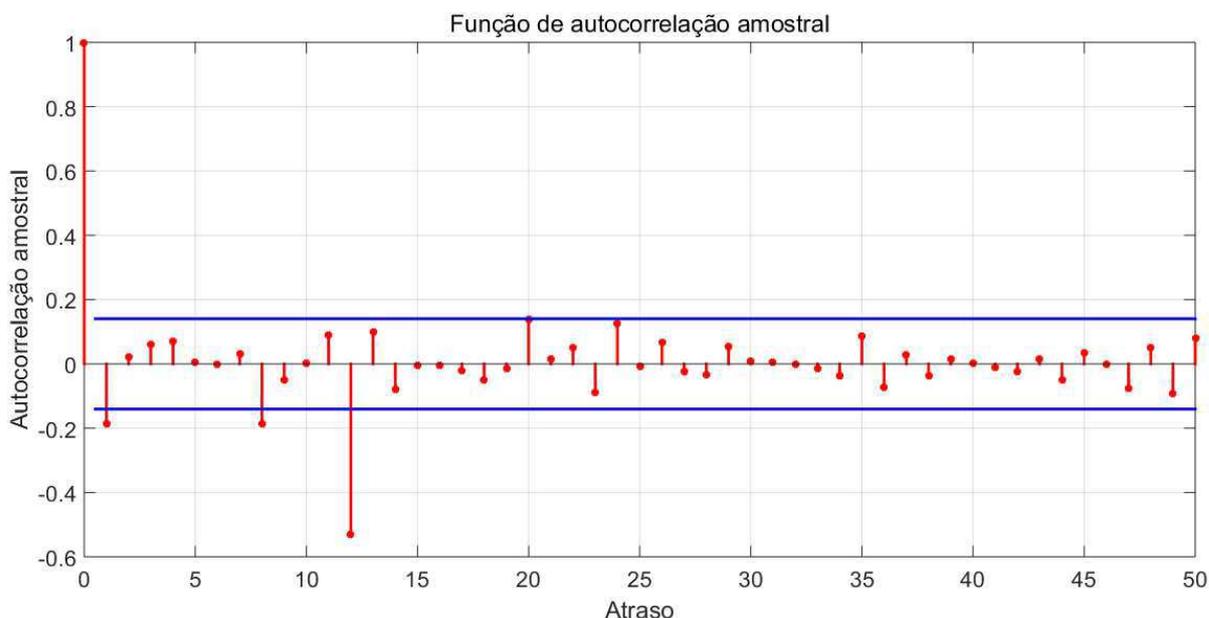


Figura 5.1 FAC amostral da série derivada.

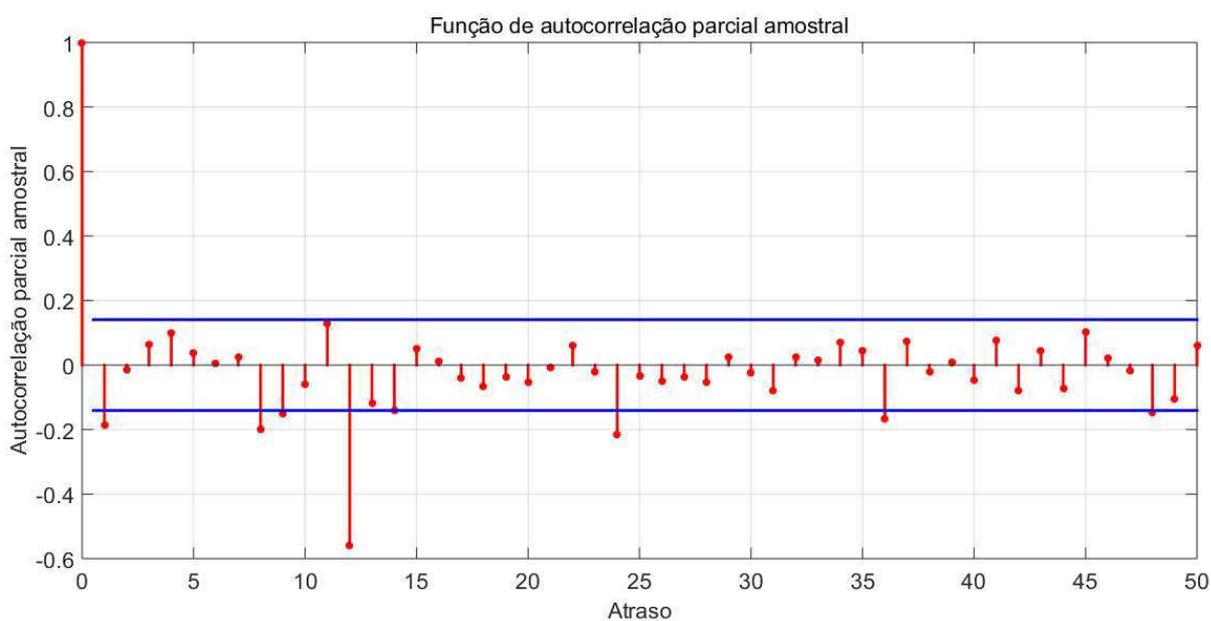


Figura 5.2 FACP amostral da série derivada.

Os comportamentos esperados para a FAC e FACP para os processos AR, MA e ARMA foram apresentados na Tabela 2. Buscou-se identificar esses comportamentos nos gráficos das Figuras 5.1 e 5.2, mas observou-se que os picos ocorrem em diferentes atrasos, tanto na FAC quanto na FACP, dificultando a identificação do modelo. Por isso, optou-se por utilizar o modelo sugerido na literatura (MONTGOMERY *et al.*, 2008; MORETTIN e TOLOI, 1981) para séries com as

mesmas características da série temporal em estudo. O modelo é o  $ARIMA(0,1,1) \times (0,1,1)_{12}$ , cuja forma foi dada na equação 3.26 e é apresentada na equação 5.1 a seguir:

$$Y_t = Y_{t-1} + Y_{t-12} + Y_{t-13} + \varepsilon_t - \theta\varepsilon_{t-1} - \theta^*\varepsilon_{t-12} - \theta^*\theta\varepsilon_{t-13}. \quad (5.1)$$

Os parâmetros da equação 5.1 a serem estimados são  $\theta$  e  $\theta^*$ . Utilizando-se o método dos mínimos quadrados, obteve-se os valores apresentados na equação 5.2:

$$\begin{aligned} \theta &= 0,09719; \\ \theta^* &= 0,76803. \end{aligned} \quad (5.2)$$

Na Figura 5.3 a seguir, pode-se observar a resposta do modelo ARIMA dado pela equação 5.1. A qualidade do ajuste do modelo foi verificada por meio do cálculo do coeficiente de determinação,  $R^2$ , o qual foi definido na seção 4.3. O valor do  $R^2$  obtido com o ajuste do modelo ARIMA sazonal aos dados da série de consumo de energia elétrica da região NE foi de:

$$R^2 = 0,9731, \quad (5.3)$$

o que significa que o modelo consegue explicar 97,31% dos dados. Deste modo, pode-se afirmar que o modelo conseguiu se ajustar bem aos dados e, portanto, é um modelo adequado para realizar previsões.

Foram realizadas previsões livres com horizontes de um ano, dois anos e cinco anos, utilizando o modelo obtido. Nas Figuras 5.4, 5.5 e 5.6, encontram-se os resultados das previsões. Para verificar a qualidade das previsões, calculou-se o erro percentual absoluto médio, o  $MAPE$ , conforme equação 4.8. Analisando a Figura 5.4, observa-se que as previsões de até um ano à frente foram bastante satisfatórias. O erro percentual da 12<sup>a</sup> previsão foi de 0%. O  $MAPE$  resultante das 12 primeiras previsões foi de:

$$MAPE = 1,08\%. \quad (5.4)$$

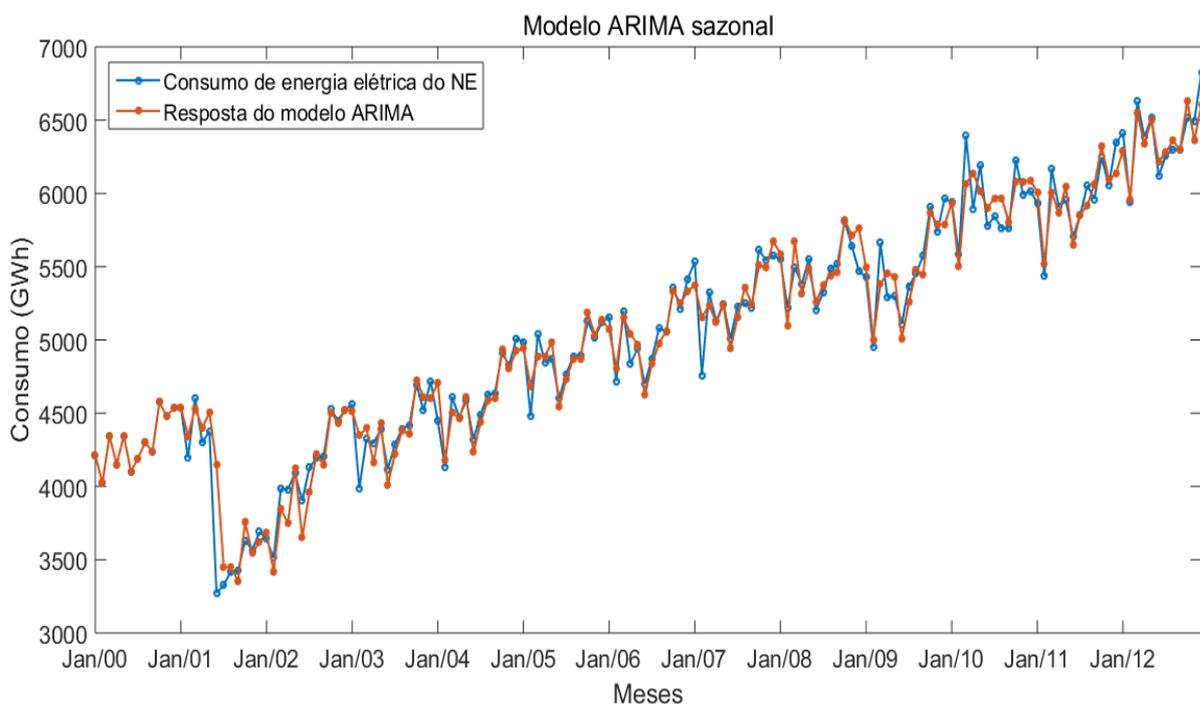


Figura 5.3 Modelo ARIMA sazonal ajustado à série temporal de consumo de energia elétrica da região Nordeste.

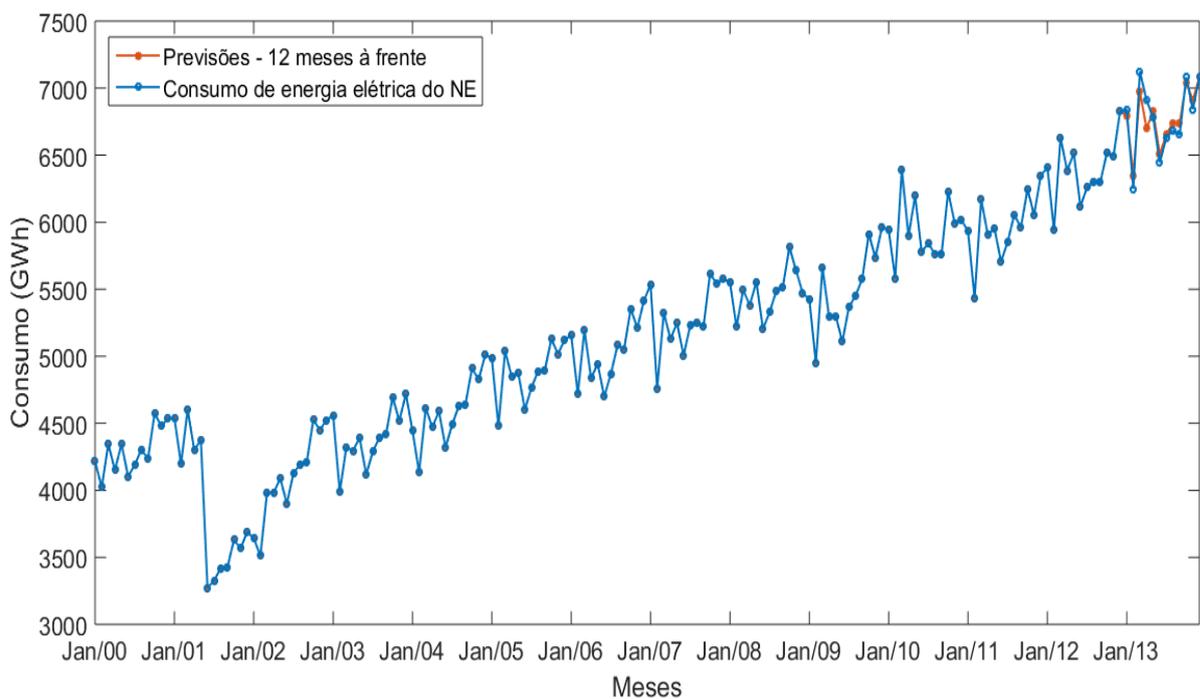


Figura 5.4 Previsões com modelo ARIMA sazonal com horizonte de 12 meses.

A previsão com horizonte de 24 meses, apresentada na Figura 5.5, também foi satisfatória. O *MAPE* calculado foi de:

$$MAPE = 1,26\%. \quad (5.5)$$

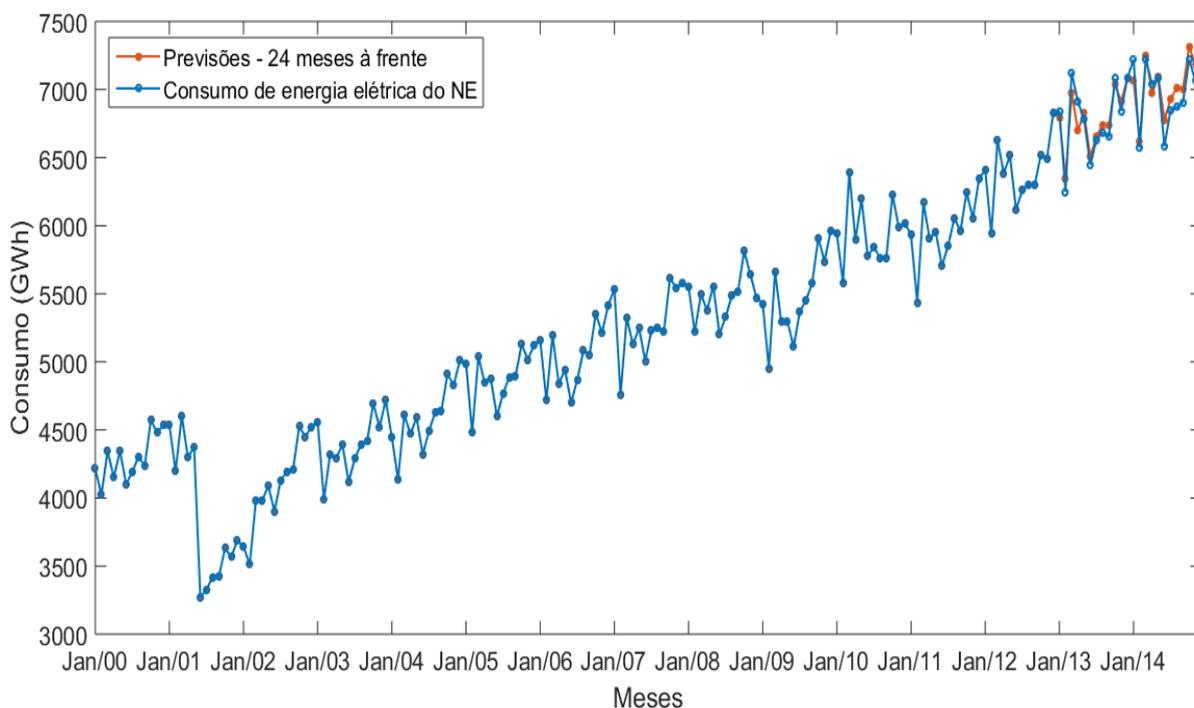


Figura 5.5 Previsões com modelo ARIMA sazonal com horizonte de 24 meses.

Já as previsões para um horizonte de cinco anos à frente apresentaram maior erro percentual. Observa-se na Figura 5.6 que as previsões realmente não conseguem acompanhar tão bem os dados reais como nos dois casos anteriores. O  $MAPE$  calculado foi de:

$$MAPE = 3,01\%. \quad (5.6)$$

Destes resultados, pode-se concluir que o erro percentual aumenta com o aumento do horizonte de previsão. Isto ocorre porque, como a previsão é livre, ou seja, os valores realizados não são atualizados a cada iteração, a componente de médias móveis do modelo, representada por  $\Theta^*(B^S)\Theta(B)\varepsilon_t$ , passa a ser zero a partir de um horizonte de previsão maior que 13 meses. Para chegar a esta conclusão, basta inspecionar a forma do modelo utilizado, dada na equação 5.1.

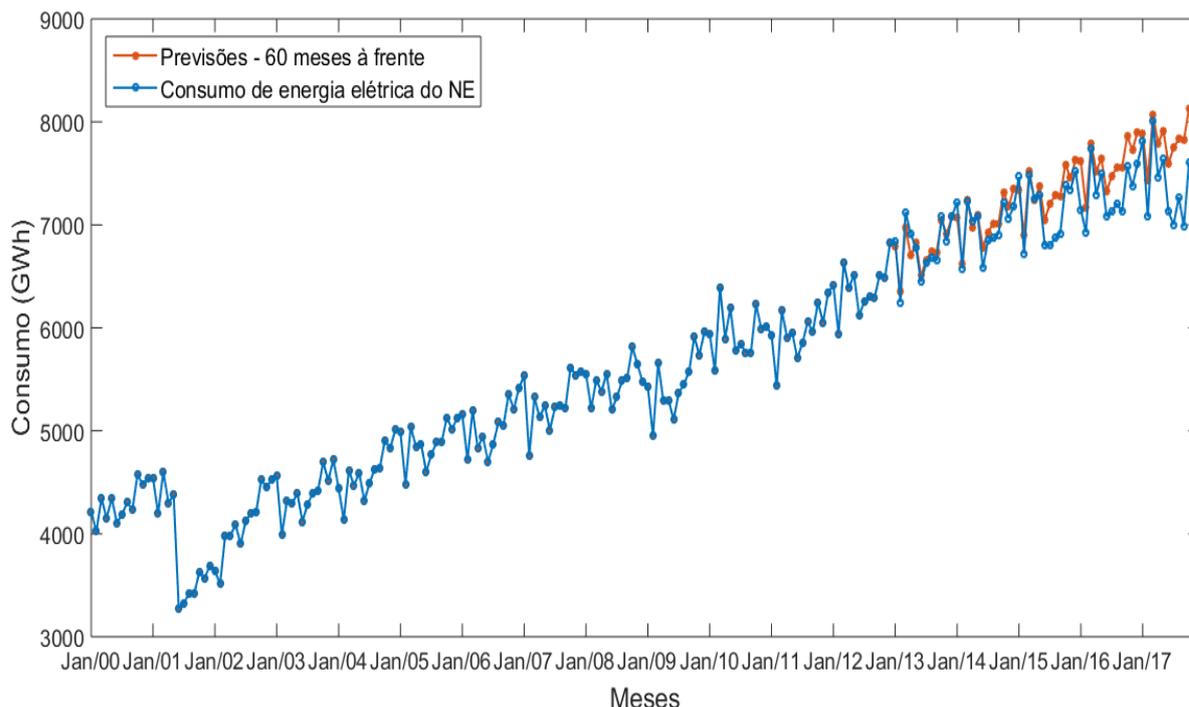


Figura 5.6 Previsões com modelo ARIMA sazonal com horizonte de 60 meses.

## 5.2 Resultados obtidos com a RNA

A rede neural apresentada na Figura 4.2 foi especificada e treinada conforme os procedimentos descritos na seção 4.2. Assim como no caso do modelo ARIMA, os dados utilizados para treinamento e validação da RNA foram os referentes aos 13 primeiros anos da série temporal de consumo de energia elétrica da região Nordeste, e os dados dos 5 anos seguintes foram utilizados para validar as previsões, com horizontes de 1 ano, 2 anos e 5 anos.

A Figura 5.7 a seguir apresenta a resposta da RNA. A qualidade do ajuste da resposta foi medida por meio do coeficiente de determinação,  $R^2$ , que foi igual a:

$$R^2 = 0,9838, \quad (5.7)$$

o que significa que o modelo consegue explicar 98,38% dos dados. Deste modo, pode-se afirmar que a resposta da RNA conseguiu se ajustar bem aos dados e, portanto, a rede neural obtida pode ser utilizada para realizar previsões.

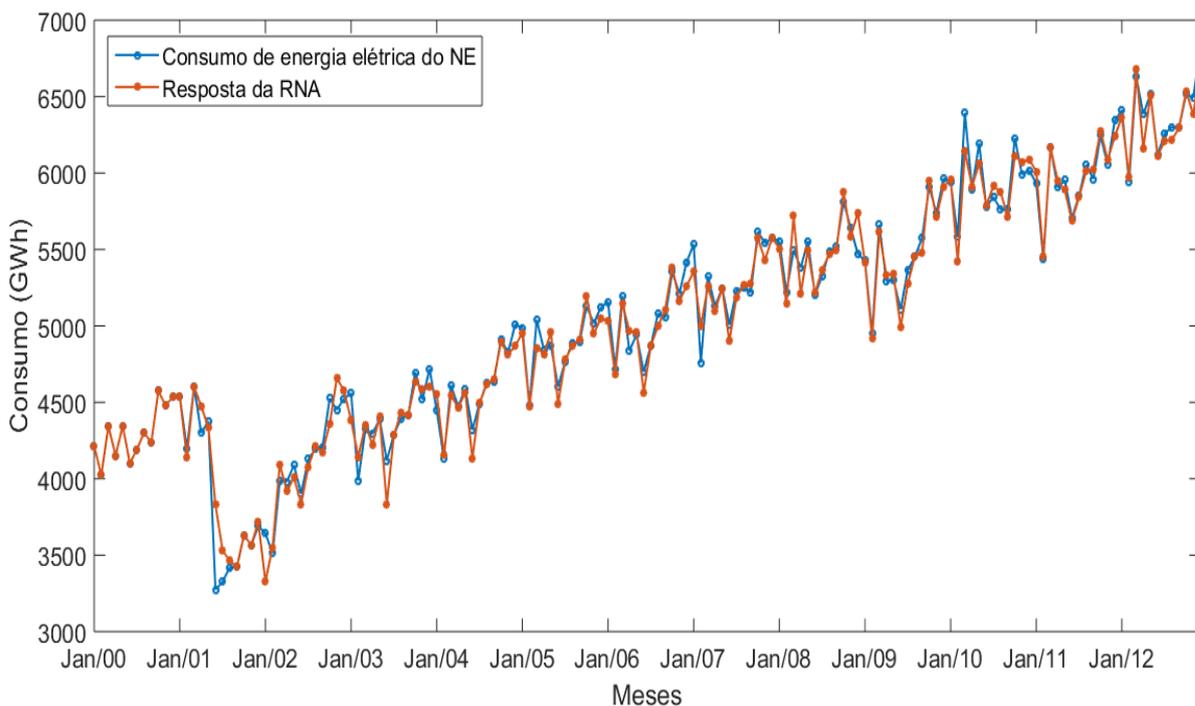


Figura 5.7 Resposta da RNA ajustada aos dados de consumo de energia elétrica da região Nordeste.

Nas Figuras 5.8, 5.9 e 5.10, encontram-se os resultados das previsões. Para verificar a qualidade das previsões, também foi calculado o erro percentual absoluto médio, o *MAPE*. A previsão livre de 12 meses à frente é apresentada na Figura 5.8. Observou-se que as previsões obtidas com a RNA não acompanham tão bem o comportamento da série quanto o modelo ARIMA sazonal. O valor do *MAPE*, dado abaixo, reflete esse resultado:

$$MAPE = 2,80\%. \quad (5.8)$$

Analisando as Figuras 5.9 e 5.10, observou-se que, apesar dos erros percentuais obtidos com as previsões da RNA serem maiores comparados com os erros obtidos com o modelo ARIMA sazonal, as previsões da RNA acompanharam melhor a tendência da série. Visualmente, a previsão da RNA de 60 meses à frente parece até ser melhor que a previsão de 60 meses do modelo ARIMA. Os erros *MAPE* calculados para as previsões de 24 e 60 meses à frente, foram, respectivamente:

$$\begin{aligned} MAPE &= 2,84\% \text{ (previsões de 24 meses à frente),} \\ MAPE &= 3,05\% \text{ (previsões de 60 meses à frente).} \end{aligned} \quad (5.9)$$

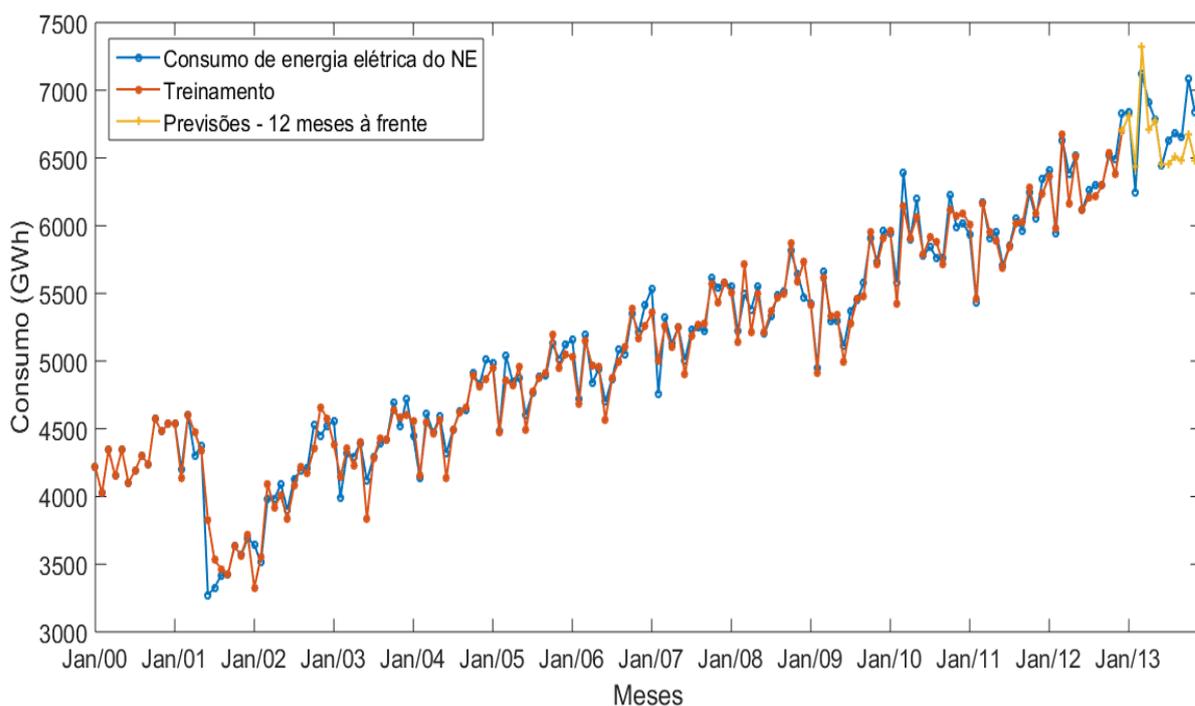


Figura 5.8 Previsões com RNA com horizonte de 12 meses.

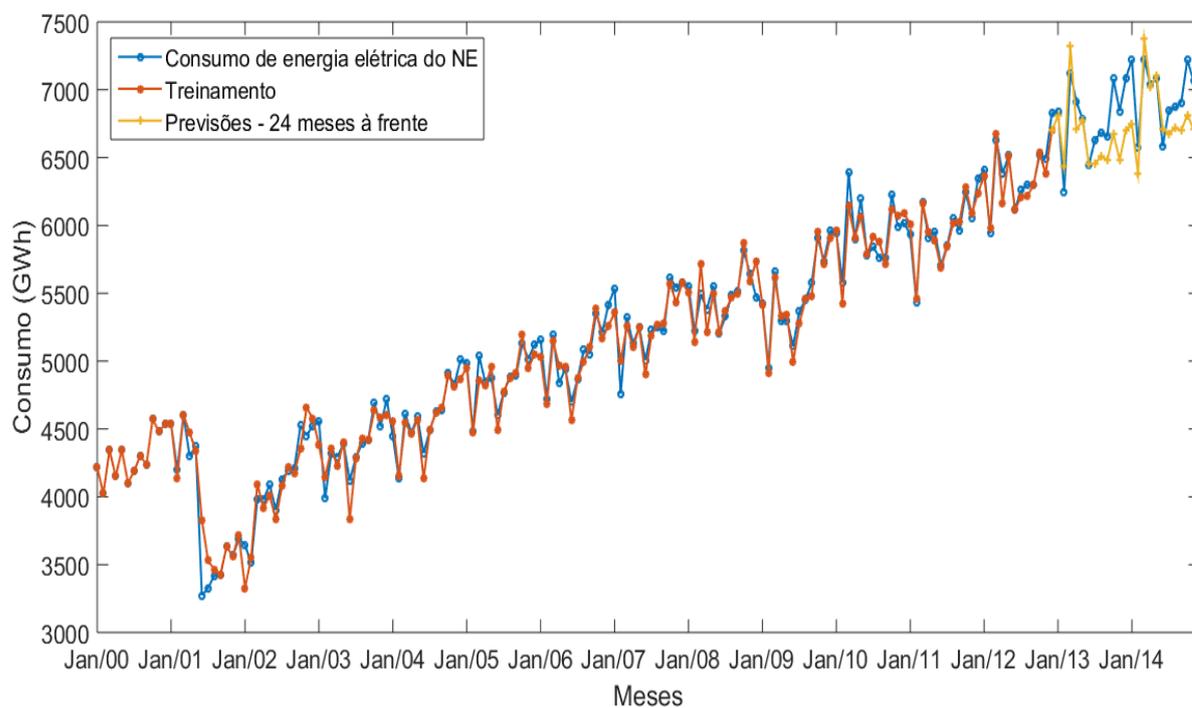


Figura 5.9 Previsões com RNA com horizonte de 24 meses.

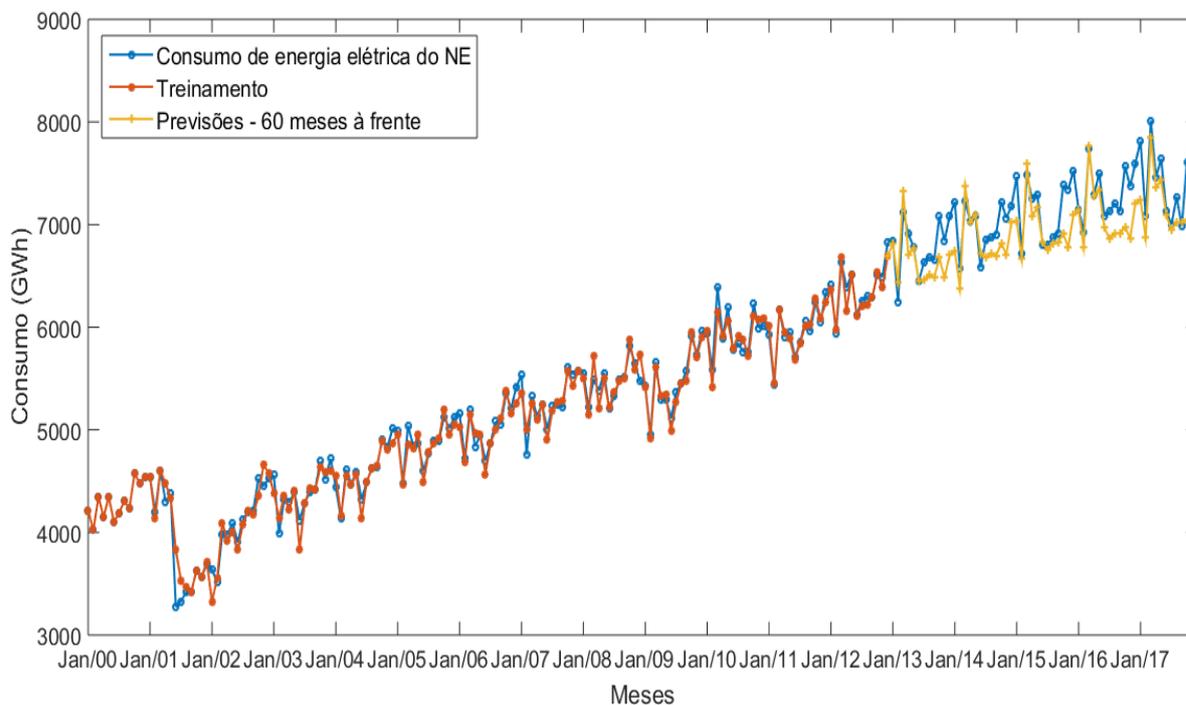


Figura 5.10 Previsões com RNA com horizonte de 60 meses.

A tabela a seguir apresenta uma comparação entre o desempenho dos dois modelos utilizados neste estudo de caso. Apesar da resposta da RNA resultar em um melhor coeficiente de determinação, os erros de previsão menores foram obtidos com o modelo ARIMA sazonal.

Tabela 2. Comparação entre o desempenho do modelo ARIMA sazonal e da RNA.

<b>Modelo</b>	<b><math>R^2</math></b>	<b>MAPE para previsão de 12 meses à frente</b>	<b>MAPE para previsão de 24 meses à frente</b>	<b>MAPE para previsão de 60 meses à frente</b>
ARIMA sazonal	0,9731	1,08%	1,26%	3,01%
RNA	0,9838	2,80%	2,84%	3,05%

## 6. Conclusões

As previsões de séries temporais podem ser utilizadas nas mais diversas áreas para tomada de decisões importantes, realização de planejamento estratégico de empresas ou até mesmo de um país. Neste trabalho, foram apresentados os principais conceitos que envolvem a análise de séries temporais, além dos modelos mais utilizados na previsão de séries temporais.

Ao fim do trabalho, foi apresentado um estudo de caso de previsão de séries temporais utilizando um modelo clássico, bastante empregado em previsão de carga, que é o modelo ARIMA sazonal (ou SARIMA), e uma rede neural artificial (RNA). Os resultados foram analisados por meio de índices de desempenho, cujos valores obtidos foram apresentados na Tabela 2.

Pode-se concluir que ambos os modelos ou metodologias apresentaram resultados satisfatórios, sendo que o modelo ARIMA sazonal apresentou menores erros com relação aos valores previstos, nos três horizontes de previsão avaliados.

Observou-se também que os erros associados às previsões aumentam com o aumento do horizonte de previsão. Isto porque, sendo a previsão livre, e os modelos não são realimentados com os valores realizados a cada iteração, então não se tem uma medida do erro associado à cada previsão.

A vantagem das redes neurais artificiais sobre o modelo ARIMA é que elas podem incorporar variáveis de entrada, ou variáveis de regressão, que podem tornar as previsões mais confiáveis. No caso da previsão de carga, essas variáveis podem ser: crescimento populacional, produto interno bruto (PIB), fatores climáticos, e outros fatores que influenciam no consumo de energia elétrica. Além disso, a utilização de uma RNA não requer uma análise estatística prévia da série temporal a ser modelada, como acontece com a classe de modelos ARIMA.

Este trabalho tem como principal contribuição a de dar início a uma linha de pesquisa no Laboratório de Alta Tensão da UFCG, servindo com um guia para os alunos de graduação que irão trabalhar com o tema daqui por diante. Fica como sugestão para a continuação deste trabalho que as metodologias aqui apresentadas sejam mais bem implementadas, de modo a obter resultados mais consistentes.

## Referências

- BALLINI, R. **Análise e Previsão de Vazões Utilizando Modelos de Séries Temporais, Redes Neurais e Redes Neurais Nebulosas**. Tese (Doutorado em Engenharia Elétrica) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Elétrica e Computação. São Paulo, 2000.
- BOX, G.E.P.; JENKINS, G.M. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden Day, 1970.
- BRAGA, A.P.; CARVALHO, A.C.; LUDEMIR, T.B. **Redes Neurais Artificiais: teoria e aplicações**. Livraria Técnica Científica, 2000
- BRANCO, S.T., SAMPAIO, R.J.B. **Aplicação de Redes Neurais Artificiais em Modelos de Previsão de Demanda Para Equipamentos de Infra-Estrutura de Telecomunicações**. XXVIII Encontro Nacional de Engenharia de Produção, Rio de Janeiro-RJ, 13 a 16 de outubro de 2008.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. 2nd Ed. Springer, 2002.
- CAMPOS, R. J. **Previsão de séries temporais com aplicações a séries de consumo de energia elétrica**. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, 2008.
- EHLERS, R. S. **Análise de Séries Temporais**. 2009. Disponível em: <http://www.icmc.usp.br/~ehlers/stemp/stemp.pdf>. Acessado em 3 de março de 2018.
- HAYKIN, S. **Neural Networks and Learning Machines**. 3rd Ed. Pearson, 2009.
- HILL, T.; O CONNOR, M.; REMUS, W. **Neural Network Models for Time Series Forecasts**. Management Science, v.42, n.7, p.1082-1092, Jul 1996.
- KRIESEL, D. **A Brief Introduction to Neural Networks**. 2007. Disponível em: <http://www.dkriesel.com>. Acessado em 23 de abril de 2018.
- LEVENBERG, K. **A method for the solution of certain problems in least squares**. Quart. Ap. Math. 2, 1944, pp. 164–168.
- MAKRIDAKIS, S.G.; WHEELWRIGHT, S.C. **Forecasting: Issues & Challenges for Marketing Management**. Journal of Marketing, v.41, n.4, p.24-38, 1977.
- MARQUARDT, D. **An algorithm for least squares estimation of nonlinear parameters**. J. Soc. Indust. Ap. Math. 11, 1963, pp. 431–441.
- MONTGOMERY, D. C.; JENNINGS, C. L; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting**. John Wiley & Sons, 2008.

MORETTIN, P. A.; TOLOI, C. M. **Modelos para previsão de séries temporais**. Instituto de Matemática Pura e Aplicada. Rio de Janeiro, 1981.

RAO, C., ALI, J. Neural **Network model for database marketing in the new global economy**. Marketing Intelligence and Planning, v.20, n.1, p. 35-43, 2002

RITZMAN, L.; KRAJEWSKI, L. **Administração da Produção e Operações**. Practice Hall. São Paulo, 2004.

SILVA, C.S. **Previsão Multivariada da Demanda Horária de Água em Sistemas Urbanos de Abastecimento**. Tese (Doutorado em Engenharia Civil) – Faculdade de Engenharia Civil, Universidade Estadual de Campinas (UNICAMP). São Paulo, 2003.

WOLD, H. O. **A Study in the Analysis of Stationary Time Series**. Almqvist & Wiksell, Uppsala, Sweden, 1938.