

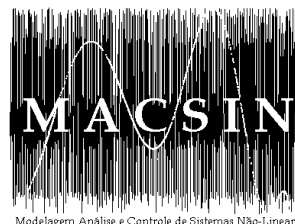
Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares

Departamento de Engenharia Eletrônica

Universidade Federal de Minas Gerais

Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG Brasil

Fone: +55 3499-4866 - Fax: +55 3499-4850



---

# Previsão de séries temporais com aplicações a séries de consumo de energia elétrica

---

**Roger Júnio Campos**

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

**Orientador:** Prof. Eduardo M.A.M. Mendes, Ph.D.

**Co-orientador:** Prof. Luis Antonio Aguirre, Ph.D.

Belo Horizonte, Abril de 2008



---

---

# Dedicatória

---

*Ao único DEUS que é digno de receber toda a honra e a glória, a força e o poder.*



---

# Agradecimentos

---

Agradeço às misericórdias de Deus, pois, sem essas eu nada seria.

À minha família: vó Lúcia, mãe (Lucinha), pai (Arnaldo), Flávia e Monique. Em especial à minha mãe, uma mulher a frente do seu tempo, com pensamentos retos e ousados, e ao mesmo tempo sempre com os pés no chão. Se não fosse o incentivo dela, eu nem estaria escrevendo estes agradecimentos.

À minha Lílian, que ao longo de cinco anos vêm sendo uma companheira inigualável, uma benção de Deus na minha vida. Te amo.

À turma de longa data, Gláucio, Rafael, Bonfim, Emerson e Demal, foram amigos nos momentos difíceis e excelentes companheiros nos momentos felizes.

Agradeço aos amigos do CPDEE, em especial, ao Marcão, Arlindo, Said, Eduardo, Luciana, Levi. Aos companheiros do MACSIN pelas importantes informações que foram de grande valia para a conclusão desta dissertação.

Ao meu orientador Prof. Eduardo Mazoni pela constante atuação e colaboração no desenvolvimento deste trabalho e também por acreditar e apostar em mim.

Ao Prof. Luis Aguirre pelo incentivo e confiança, e pelas sábias e serenas palavras.



---

# Epígrafe

---

A Palavra de Deus é alicerce para vida de todo cristão, é alimento para alma, direção e consolo a cada dia. Devemos sempre buscar comunhão com o Senhor através da leitura da Bíblia, para termos uma vida edificada e abençoada; muitas vezes passamos por tribulações querendo achar respostas em lugares diferentes, mas em Marcos 12:24 Jesus nos adverte: *“Por ventura não errais vós em razão de não saberdes as Escrituras nem o poder de Deus?”*

Nossa meta deve ser conhecer Jesus, *“até que todos cheguemos à unidade da fé, e ao conhecimento do Filho de Deus, a homem perfeito, à medida da estatura completa de Cristo”* (Efésios 4.13), alimentando-nos da Palavra de Deus, que é a Bíblia, que foi escrita por homens, mas inspirado pelo Espírito Santo de Deus para nossas vidas, *“Para que os seus corações sejam animados, estando unidos em amor, e enriquecidos da plenitude do entendimento para o pleno conhecimento do mistério de Deus-Cristo, no qual estão escondidos todos os tesouros da sabedoria e da ciência e digo isto, para que ninguém vos engane com palavras persuasivas”* (Colossenses.2:4).

*Então conheçamos, e prossigamos em conhecer ao SENHOR (Oséias 6.3a)!!!*





---

# Sumário

---

<b>Resumo</b>	<b>xiii</b>
<b>Abstract</b>	<b>xv</b>
<b>Lista de Figuras</b>	<b>xx</b>
<b>Lista de Tabelas</b>	<b>xxi</b>
<b>Lista de Abreviações</b>	<b>xxiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	2
1.2 Objetivos . . . . .	4
1.3 Apresentação do trabalho . . . . .	5
<b>2 Fundamentação Teórica</b>	<b>7</b>
2.1 Uma breve revisão sobre as técnicas de previsão . . . . .	7
2.1.1 Métodos de Previsão . . . . .	8
2.1.1.1 Métodos de Previsão de Médio e longo- prazos . . . . .	8
2.1.1.2 Métodos de Previsão de curto-prazo . . . . .	9
2.2 Identificação de Sistemas . . . . .	12

2.3	Representações Lineares e Não-Lineares . . . . .	13
2.3.1	Modelos ARIMA . . . . .	14
2.3.1.1	Previsão com Modelos ARIMA . . . . .	16
2.3.1.2	Atualizando as previsões . . . . .	18
2.3.2	Rede Neuro-Fuzzy . . . . .	18
2.3.3	Rede Neural Artificial . . . . .	24
2.3.3.1	Perceptron Multicamadas . . . . .	24
2.3.3.2	Controle de generalização . . . . .	26
2.4	Comentários finais . . . . .	27
<b>3</b>	<b>Metodologia</b>	<b>29</b>
3.1	Índices de desempenho . . . . .	29
3.2	Descrição dos dados . . . . .	32
3.2.1	Consumo de New England . . . . .	32
3.2.2	Consumo de Minas Gerais . . . . .	34
3.3	Componente Cíclica, sazonal e de tendência . . . . .	36
3.4	Procedimentos para separação das componentes da série temporal . . . . .	38
3.5	Procedimentos para os modelos ARIMA e NARIMA . . . . .	41
3.6	Procedimentos para a RNF . . . . .	43
3.6.1	Modelo de consumo de New England . . . . .	44
3.6.2	Modelo de consumo da CEMIG . . . . .	45
3.7	Procedimentos para a RNA . . . . .	47
3.7.1	Modelo de consumo de New England . . . . .	48
3.7.2	Modelo de consumo da Cemig . . . . .	49
3.8	Comentários finais . . . . .	49

<b>4</b>	<b>Estudo de caso 1: Série temporal de New England</b>	<b>51</b>
4.1	Modelo ARIMA . . . . .	51
4.2	Modelo NARIMA . . . . .	54
4.3	Rede Neuro-Fuzzy . . . . .	55
4.4	Rede Neural Artificial . . . . .	58
4.5	Comparação das Representações . . . . .	59
<b>5</b>	<b>Estudo de caso 2: Série temporal da Cemig</b>	<b>61</b>
5.1	Modelo ARIMA . . . . .	61
5.2	Modelo NARIMA . . . . .	64
5.3	Rede Neuro-Fuzzy . . . . .	64
5.4	Rede Neural Artificial . . . . .	68
5.5	Comparação das Representações . . . . .	70
<b>6</b>	<b>Discussão e Conclusão</b>	<b>75</b>
6.1	Proposta para trabalhos futuros . . . . .	78
	<b>Referências Bibliográficas</b>	<b>79</b>



---

# Resumo

---

O objetivo principal desta dissertação é estudar e aplicar métodos de previsão de longo-prazo para previsão de consumo de energia elétrica em séries com tendências e ciclos utilizando apenas os regressores da série histórica. Foram realizados dois estudos de caso, o primeiro usando a série de consumo de energia da cidade de New England (USA), e o outro a série de consumo do Estado de Minas Gerais (Brasil). O trabalho aborda métodos de previsão bastantes disseminados no meio acadêmico e científico, a saber são eles: o modelo *AutoRegressive Integrated Moving Average* (ARIMA), *Nonlinear AutoRegressive Integrated Moving Average* (NARIMA), a *Rede Neuro-Fuzzy* (RNF) e a *Rede Neural* (RNA). Usando essas quatro representações a previsão de 60 passos à frente do consumo de energia é estimado. Como primeiro passo, foi definida qual técnica seria usada para separar as componentes das séries temporais. A metodologia usada foi proposta por Mohr (2005). A componente sazonal é estimada usando os modelos NARIMA, RNF e RNA, e os modelos obtidos são somados à componente de tendência resultando as estimativas do consumo. O modelo ARIMA é estimado usando a série temporal sem separar suas componentes. Os modelos são comparados usando os índices de desempenho: MPE, MAPE e RMSE. No caso do consumo do Estado de Minas Gerais, como a série temporal é composta por poucas observações, na previsão de 60 passos à frente a componente de tendência é aproximada por uma reta, e como não existem amostras para realizar a comparação, é realizada uma análise espectral das previsões juntamente com a série original de consumo. Os resultados obtidos para os dois estudos de caso mostram que os modelos ARIMA, NARIMA, RNF e RNA são ferramentas eficientes que podem auxiliar no planejamento e tomadas de decisões no setor elétrico.

**Palavras-chave:** previsão longo-prazo, consumo de energia, série temporal, ARIMA, NARIMA, Neuro-Fuzzy, Rede Neural.



---

---

# Abstract

---

The main objective of this work is to study and to apply methods of long-term load forecasting to time series with trends and cycles just using regressors of the historical series. Two case studies were accomplished, the first using the series of load of the city of New England (USA), and the other the series of load of the state of Minas Gerais (Brazil). The work approaches many forecasting methods well know in the academic and scientific field, such as: the model ARIMA, NARIMA, the Neuro-Fuzzy Network (NFN) and the Artificial Neural Network (ANN). Using those four representations the forecast of 60 steps ahead of the load is estimated. First of all, it was defined which technique would be used to separate the components of the temporary series. The chosen methodology was proposed by Mohr (2005). The seasonal component was adjusted using the models NARIMA, RNF and RNA, and the obtained models are added to the tendency component resulting the estimates of the load. The ARIMA model is esteemed using the time series without separating their components. The models are compared using the performance indices: MPE, MAPE and RMSE. In the case of the load of the State of Minas Gerais, as the time series is composed by few observations, in the forecast of 60 steps ahead, the trend component is approximated by a straight line, and as there is no samples to accomplish the comparison, a spectral analysis of the forecasts is accomplished with the original series of load. The results to the two case studies show that the models ARIMA, NARIMA, RNF and RNA are efficient tools that can aid in the planning and sockets of decisions in the electric sector.

**Keywords:** long-term load forecasting, times series, ARIMA, NARIMA, Neuro-Fuzzy, Neural Network.





---

## Lista de Figuras

---

2.1	Modelo fuzzy Sugeno com duas entradas e e duas regras (Jang (1993)). . . . .	20
2.2	Arquitetura ANFIS equivalente para o modelo da Figura 2.1 (Jang (1993)) . . . . .	21
2.3	Arquitetura ANFIS para (a) modelo Sugeno e (b) espaço de entrada (Jang (1993)). . . . .	23
2.4	Arquitetura ANFIS para o modelo Sugeno, onde a normalização dos pesos é executada na última camada (Jang (1993)).	23
3.1	Consumo horário total da cidade de New England referente ao período que se estende de janeiro de 1980 a dezembro de 1998. Fonte: ISO-NE. . . . .	33
3.2	Consumo mensal total da cidade de New England referente ao período que se estende de janeiro de 1980 a dezembro de 1998. Fonte: ISO-NE . . . . .	34
3.3	Série de consumo mensal total do Estado de Minas Gerais referente ao período que se estende de dezembro de 1996 a julho de 2006. Fonte: CEMIG. . . . .	36
3.4	Componentes de tendência (- -), sazonalidade (-.-), ciclo (- +), e consumo total (-). . . . .	39
3.5	Componentes de “tendência + ciclo” (- -), sazonalidade (-.-), e consumo total (-). . . . .	40
3.6	Esquema das etapas do modelo ARIMA. . . . .	41
3.7	Representação esquemática de modelos (a) ARIMA e (b) NARIMA. . . . .	43
3.8	Representação esquemática de modelo Neuro-Fuzzy do consumo de energia de New England. . . . .	45

3.9	Representação esquemática de modelo Neuro-Fuzzy do consumo de energia do Estado de Minas Gerais. ( <i>ys</i> representa os dados da componente sazonal da série de consumo da Cemig.) . . . . .	46
3.10	Representação esquemática de modelo Neural do consumo de energia de New England. . . . .	48
3.11	Representação esquemática de modelo Neural do consumo de energia do Estado de Minas Gerais. ( <i>ys</i> representa os dados da componente sazonal da série de consumo da Cemig.)	50
4.1	Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: 0,996%; MAPE: 7,406%; RMSE: 1,096; $U_{Theil}$ : 0,892. . . . .	52
4.2	Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: 0,129%; MAPE: 5,920%; RMSE: 1,047; $U_{Theil}$ : 0,995. . . . .	53
4.3	Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: -0,298%; MAPE: 6,924%; RMSE: 0,961; $U_{Theil}$ : 0,782. . . . .	54
4.4	Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: 1,644%; MAPE: 6,810%; RMSE: 1,134; $U_{Theil}$ : 1,076. . . . .	55
4.5	Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: 0,801%; MAPE: 3,279%; RMSE: 0,518; $U_{Theil}$ : 0,873. . . . .	56

4.6	Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–) consumo previsto. Erros apresentados, MPE: 0,813%; MAPE: 2,672%; RMSE: 0,506; $U_{Theil}$ : 0,567. . . . .	57
4.7	Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–) consumo previsto. Erro apresentados, MPE: 0,204%; MAPE: 2,383%; RMSE: 0,364; $U_{Theil}$ : 1,054. . . . .	58
4.8	Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–) consumo previsto. Erro apresentados, MPE: 1,404%; MAPE: 3,142%; RMSE: 0,576; $U_{Theil}$ : 0,998. . . . .	59
5.1	Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–) consumo previsto. Erros apresentados, MPE: 1,956%; MAPE: 3,557%; RMSE: 1,732; $U_{Theil}$ : 1,425. . . . .	62
5.2	Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe do modelo ARIMA (–) consumo previsto. . . . .	63
5.3	Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo NARIMA. (–) consumo real, (–) consumo previsto. Erros apresentados, MPE: -0,075%; MAPE: 2,723%; RMSE: 1,354; $U_{Theil}$ : 1,114. . . . .	65
5.4	Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe do modelo NARIMA (–) consumo previsto. . . . .	66

5.5	Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–.) consumo previsto. Erros apresentados, MPE: 0,431%; MAPE: 1,883%; RMSE: 0,762; $U_{Theil}$ : 0,564. . . . .	67
5.6	Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe da saída da rede (–.) consumo previsto. . . . .	68
5.7	Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede (–) consumo real, (–.) consumo previsto. Erro apresentados, MPE: -0,361%; MAPE: 1,503%; RMSE: 0,590; $U_{Theil}$ : 0,313. . . . .	69
5.8	Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe da saída da rede. (–) consumo real, (–.) consumo previsto. . . . .	70
5.9	Comparação das simulações dos modelos: (a) ARIMA, (b) NARIMA, (c) RNF, (d) RNA. As previsões iniciam-se a partir da linha vertical contínua (observação 116). . . . .	71
5.10	Comparação dos espectros de frequência: (a) série temporal de consumo do Estado de Minas Gerais, (b) ARIMA, (c) NARIMA, (d) RNF, (e) RNA. . . . .	72

---

---

## Lista de Tabelas

---

3.1	Comparação dos modelos através do índice de desempenho: MPE, MAPE e RMSE para a série de consumo em Minas Gerais.	50
4.1	Comparação dos modelos através do índice de desempenho: MPE(%), MAPE(%) e RMSE para a série de consumo em New England. . . . .	60
5.1	Comparação dos modelos através do índice de desempenho: MPE(%), MAPE(%), RMSE e $U_{Theil}$ para a série de consumo em Minas Gerais. . . . .	71



---

## Lista de Abreviações

---

AIC	Critério de Informação de Akaike
ANFIS	Adaptive Neuro-Fuzzy Inference System
AR	Auto Regressive Model
ARIMA	Auto Regressive Integrated Moving Average Model
ARIMAX	Auto Regressive Integrated Moving Average Model with eXogenous input
ARMA	Auto Regressive Moving Average Model
ARMAX	Auto Regressive Moving Average Model with eXogenous input
CAKE	Current Accumulated Knowledge Element
CEMIG	Companhia Energética de Minas Gerais
CURNOLE	Current Residual Nonlinear Element
DAN2	Dynamic Artificial Neural Network Model
FAC	Função de autocorrelação
FACp	Função de autocorrelação parcial
FARMAX	FuzzyAuto Regressive Moving Average Model with eXogenous input
MA	Moving Average
MAE	Mean absolute error

MAPE	Mean absolute percentage error
ME	Mean error
MIMO	Multi input multi output
MISO	Multi input single output
MLP	Multi-layer Perceptron
MPE	Mean percentage error
MQO	Mínimos Quadrados Ordinários
MSE	Mean squared error
NARIMA	Nonlinear Auto Regressive Integrated Moving Average Model
NARMA	Nonlinear Auto Regressive Moving Average Model
RNA	Rede Neural Artificial
RNF	Rede Neuro-Fuzzy
SE	Sistema Especialista
SVM	Support Vector Machines
TCS	Trend-Cycle-Season Filter



# Introdução

---

O setor elétrico em todo o mundo vem sofrendo um processo de transformação, quer no campo político, quer no econômico. A privatização e a desregulamentação do sistema elétrico são os grandes responsáveis pela mudança de cenário deste sistema. No Brasil o processo de reestruturação foi iniciado com a promulgação da Lei de Concessões em 1995.

Nas últimas duas décadas, o setor de energia elétrica tem evoluído na direção do estabelecimento de uma estrutura de mercado. A chamada indústria de energia elétrica tendo como principal característica a livre concorrência na compra e venda de energia elétrica, para a qual constitui atributo indispensável a desverticalização, isto é, atividades de geração, transmissão e distribuição são desempenhadas de forma independente e autônoma.

Com tantas mudanças foi necessário toda uma reestruturação do sistema para que se tenha uma indústria de energia elétrica de alta qualidade, com fornecimento de energia de qualidade e segurança para o consumidor com a minimização dos custos. Logo, a *previsão de carga*<sup>1</sup> é um tema de grande interesse para este setor. Pois, os custos dos erros de previsão acabam por gerar um custo financeiro muito elevado (Oliveira (2004)). Seja ele por erro de previsão por excesso (*overforecast*) ou erro de previsão por falta (*underforecast*) (Serrão (2003)).

Como pode ser encontrado em Reis (1999) previsão de carga elétrica pode ser dividida em três categorias: previsão de carga de curto-prazo (e curtíssimo-prazo), médio-prazo e longo-prazo. Previsões de curto-prazo normalmente são de poucos minutos a uma semana à frente. Previsões de médio-prazo englobam horizontes que variam de uma semana a um ano, e por fim as *previsões de longo-prazo* que visam horizontes acima de um

---

<sup>1</sup>Nesta seção a palavra carga será utilizada de maneira ampla, de forma a contemplar qualquer uma das grandezas normalmente consideradas em problemas de previsão no setor elétrico, a saber: carga, demanda e consumo.

ano. Os diferentes horizontes de previsão são importantes para operações diferentes dentro de uma companhia de energia elétrica (Rodrigues (2002)).

Com o propósito de otimizar as tarefas de um sistema de grande escala, diversas técnicas estão sendo aplicadas com uma considerável expectativa de economia nos custos e de eficiência do sistema, com este objetivo muitas pesquisas vêm sendo desenvolvidas nesta área (De Gooijer e Hyndman (2006)).

Dando passos nesta direção este trabalho apresenta técnicas de modelagem matemática usadas para a previsão de carga. As representações ARIMA (linear), NARIMA (não-linear), a Rede Neuro-Fuzzy (não-linear), e a Rede Neural (não-linear) serão usadas a fim de alcançar este propósito.

Usando essas técnicas é realizada a previsão de 5 anos à frente do *consumo de energia do Estado de Minas Gerais*<sup>2</sup>. Para verificar-se a capacidade preditiva dos modelos serão realizados alguns testes preliminares com uma base de dados da companhia elétrica de New England (USA)<sup>3</sup>. Isto porque o histórico de dados da CEMIG-Distribuição é relativamente pequeno (do ano de 1996 a 2006), em especial por causa do racionamento em junho de 2001 que mudou toda a dinâmica da série temporal.

O trabalho foi desenvolvido dentro do escopo do projeto de pesquisa intitulado "*Pesquisa de Modelo de Previsão de Consumo de Energia para Curto e Médio Prazo, utilizando Inteligência Computacional*". O projeto, resultado da parceria entre a UFMG e a CEMIG-Distribuição, por meio de sua Gerência de Planejamento de Mercado, objetiva estudar o problema de previsão de consumo de energia e implementar um sistema computacional que sirva de suporte à tomada de decisões na área de comercialização de energia.

## 1.1 Motivação

Os anos de 2003 e 2004 foram marcantes para o setor elétrico brasileiro. A crise de abastecimento de energia do ano de 2001 forçou o Governo Federal a tomar uma série de medidas para tentar corrigir os erros do modelo em vigência e garantir a segurança do suprimento de energia elétrica no país. A Lei nº1040848/04 acompanhada do Decreto nº5.163/04

---

<sup>2</sup>A série de consumo do Estado de Minas Gerais é fornecida pela CEMIG-Distribuição ([www.cemig.com.br](http://www.cemig.com.br))

<sup>3</sup>Dados encontrados no site <http://www.iso-ne.com>

criou uma nova estrutura regulatória para o setor e instituiu uma série de mecanismos que regulamentam a comercialização de energia entre os agentes do mercado. Além de garantir a oferta de energia, o novo modelo tem como objetivo promover a universalização do acesso à eletricidade e a modicidade tarifária.

As novas regras definem o papel dos leilões de compra e venda de energia elétrica. O processo de comercialização passou a ser realizado dentro de um ambiente de contratação regulada em que a compra da energia pelas distribuidoras é realizada por meio de leilões, sendo observado o critério de menor tarifa. Uma característica importante é que os contratos negociados referem-se ao fornecimento de energia em horizontes futuros. Dependendo do tipo de leilão esse horizonte pode ser de 1, 3 ou 5 anos. O objetivo é reduzir o custo da aquisição da energia elétrica a ser repassada aos consumidores e viabilizar a expansão do setor. Ficou definido que, para negociar a energia, os geradores devem possuir lastro físico garantido de todo o montante comercializado nos leilões. Por sua vez, as distribuidoras de energia são obrigadas a remeter ao Ministério de Minas e Energia as previsões anuais do consumo de energia, em sua área de operação, em um horizonte de 60 meses, com um erro máximo de 3% para cima do montante verificado ao final do período.

A qualidade da projeção futura do consumo de energia a ser realizada pelas distribuidoras impacta de forma significativa em dois pontos distintos. O primeiro se refere ao próprio planejamento da expansão do sistema elétrico, uma vez que estas projeções são utilizadas pelo Ministério de Minas e Energia para dar suporte às decisões tomadas neste sentido. O segundo ponto está no contexto operacional e financeiro das próprias distribuidoras. Conforme citado por Oliveira (2004) todo o erro que ultrapasse os limites estabelecidos acarreta em perdas e influi diretamente nos resultados financeiros das distribuidoras. As perdas financeiras decorrem: (i) das multas aplicadas sobre os erros na previsão; (ii) dos custos referentes à devolução da energia sobre-contratada e (iii) da contratação de energia a preços mais altos em leilões de ajuste (curto-prazo) decorrente de previsões abaixo do consumo real verificado.

Tendo em vista o contexto criado pelas mudanças estruturais do setor, faz-se necessário por parte das distribuidoras e comercializadoras de energia, o investimento em técnicas e metodologias para gerenciar as incertezas inerentes ao mercado. A tomada de decisão relativa à contratação futura e

o próprio gerenciamento dos contratos (em escala anual ou mensal) devem estar embasados por previsões satisfatórias a cerca do consumo de energia. Percebe-se claramente a demanda, por parte das distribuidoras em: (i) realizar previsões anuais para a sua série de consumo em um horizonte de até 5 anos; (ii) realizar previsões mensais a cerca do comportamento do consumo de energia.

O desafio de prever o consumo de energia elétrica no contexto de uma economia estável, com séries temporais razoavelmente estacionárias, já seria um grande desafio. Esse desafio torna-se tanto maior no contexto de uma economia cujas características mudam significativamente em curtos intervalos de tempo, tudo isso agravado pela globalização mundial.

Conforme citado, a dinâmica do consumo de energia possui uma complexidade inerente ao processo. A influência de variáveis econômicas e climáticas, juntamente com a dificuldade de se conhecer as entradas do sistema e a relação destas com o consumo de energia, torna difícil o uso de modelos analíticos para tratar o problema. Por outro lado, o uso de ferramentas baseadas em *identificação de sistemas e inteligência computacional*<sup>4</sup> vem ganhando força e a literatura mostra diversas aplicações realizadas com sucesso. Alguns estudos podem ser encontrados em (Aslanargun, Mammadov, Yazici, e Yolacan (2007); Serrão (2003); Chow e Leung (1996); Haydari, Kavehnia, Askari, e Ganbariyan (2007)).

## 1.2 Objetivos

Este trabalho abordará o problema de previsão de série temporais não-estacionárias de consumo de energia elétrica por meio das seguintes metodologias de modelagem:

1. Modelo AutoRegressivo integrado de Média Móvel, ARIMA;
2. Modelo Não-linear AutoRegressivo integrado de Média Móvel, NARIMA;
3. Rede Neuro-Fuzzy, RNF;
4. Rede Neural Artificial, RNA.

---

<sup>4</sup>Tais como modelos ARIMA, Rede Neural, Lógica Fuzzy

Por meio de duas séries temporais: (i) consumo de energia da cidade de New England e (ii) consumo de energia do Estado de Minas Gerais, as metodologias numeradas acima serão aplicadas com intuito de realizar previsões de 60 passos (meses) à frente.

Deseja-se investigar até que ponto essas metodologias de fato mostram robustez à medida que o horizonte de previsão aumenta. Como também, estas metodologias se comportam quando são utilizadas para modelar séries temporais não-estacionárias e com irregularidades.

## 1.3 Apresentação do trabalho

Este trabalho apresenta mais 5 Capítulos. No Capítulo 2 será feita uma revisão bibliográfica contendo de forma resumida e geral os conceitos envolvidos no âmbito da previsão de séries temporais, em especial a previsão do consumo de energia elétrica.

No Capítulo 3 serão apresentadas as metodologias para obter os modelos de previsão. Será mostrada a metodologia para a decomposição de séries temporais, para o modelo ARIMA, NARIMA, Rede Neuro-Fuzzy e Rede Neural. A avaliação do desempenho da metodologia será feita nos Capítulos 4 e 5.

No Capítulo 4 serão mostrados os resultados, usando as diferentes representações matemáticas, para a série temporal de consumo de energia da cidade de New England, USA. E no Capítulo 5 serão mostrados os resultados da previsão para a série de consumo de energia do Estado de Minas Gérias, Brasil.

Uma breve discussão dos resultados encontrados, impressões gerais e conclusões estão disponíveis no Capítulo 6.



## Fundamentação Teórica

---

Neste Capítulo serão discutidos os principais trabalhos que aplicam técnicas para a previsão de séries temporais. Abordar-se-á também os fundamentos teóricos das técnicas que são usadas ao longo desta dissertação.

### 2.1 Uma breve revisão sobre as técnicas de previsão

Para previsão de carga de curto-prazo vários fatores devem ser considerados, tais como fatores de tempo, dados climáticos e as possíveis classes de consumidores. A previsão de médio e curto-prazos leva em conta a carga histórica e dados climáticos, o número de consumidores em cada classe, a economia e dados demográficos e suas respectivas previsões, dentre outros fatores (Serrão (2003); Rodrigues (2002)).

Os fatores de tempo incluem o mês de um ano, o dia de uma semana e uma hora de um dia. Existem importantes diferenças entre carga de dias úteis de uma semana e carga de um final de semana. A carga em diferentes dias úteis de uma semana também podem apresentar comportamentos diferentes. Por exemplo, segundas e sextas-feiras são adjacentes aos finais de semana, por isso podem ter estruturas diferentes do que terças a quintas-feiras. Isso ocorre particularmente durante o verão. Feriados são mais difíceis de prever que dias normais por causa de suas ocorrências infreqüentes.

Fatores climáticas influenciam na carga. De fato, os fatores climáticos são os mais importantes na previsão de carga de curto-prazo (Campos, Jesus, e Mendes (2007); Pandian, Duraiswamy, Rajan, e Kanagaraj (2006)). Várias variáveis climáticas devem ser consideradas para a previsão de carga. Temperatura e umidade do ar são as variáveis mais usadas para

a previsão da carga. Uma pesquisa de previsão curto-prazo de carga elétrica, (Hippert, Pedreira, e Souza (2001)), indica que dos 22 artigos que são considerados, 13 fazem o uso apenas da temperatura, 3 usam a temperatura e umidade, 3 utilizam parâmetros climáticos adicionais e 3 usam apenas parâmetros de carga.

### 2.1.1 Métodos de Previsão

Nas últimas décadas foram desenvolvidos vários métodos de previsão. Uma variedade de métodos, modelos de regressão, Redes Neurais, Lógica Fuzzy, sistemas especialistas e algoritmos de aprendizagem estatísticos, são usados normalmente para previsões de curto-prazo (De Gooijer e Hyndman (2006)). O desenvolvimento, melhorias, e investigação de ferramentas apropriadas têm levado o desenvolvimento de técnicas de previsão mais precisas.

Aproximações estatísticas usualmente requerem um modelo matemático que represente a carga elétrica como uma função de diferentes fatores, tais como, tempo, condições meteorológicas, e classes de consumidores. As duas categorias importantes de tal modelo matemático são: modelos aditivos e modelos multiplicativos. Eles diferem se a previsão de carga é um somatório (aditivo) de vários componentes ou o produto (multiplicativo) de vários fatores. Por exemplo, Chen, Canizares, e Singh (2001) apresentam um modelo aditivo que prediz a carga como uma função de quatro componentes, (i) a parte normal da carga (é um conjunto de cargas normalizadas para cada tipo de dia que foi identificado em um ano), (ii) a parte sensível da carga em relação as condições meteorológicas, (iii) uma componente especial que cria uma divergência do padrão de carga habitual e (iv) o termo aleatório (ruído).

#### 2.1.1.1 Métodos de Previsão de Médio e longo-prazos

A modelagem de uso final, modelagem econométrica, e suas combinações são os métodos de previsão de médio e longo-prazos usados com mais frequência (De Gooijer e Hyndman (2006)).

**Modelos Econométricos** - A aproximação econométrica combina teoria econômica e técnicas estatísticas para previsão da demanda elétrica. Esta



aproximação estima a relação entre consumo de energia (variável dependente) e fatores que influenciam no consumo. As relações são estimadas pelo método dos mínimos quadrados ou métodos de séries temporais De Gooijer e Hyndman (2006).

Uma das opções nessa estrutura é agregar a aproximação econométrica, quando o consumo em setores diferentes (residencial, comercial, industrial, etc) é calculado como uma função das condições meteorológicas, econômicas e outras variáveis, e então são agregadas as estimativas usando dados históricos. A integração da aproximação econométrica com a aproximação *condicional* se dá quando componentes comportamentais são inseridas nas equações da aproximação (*condicional*).

**Modelos de aprendizagem estatísticos** - Os métodos *condicional* e econométrico requerem grande informação dos consumidores, seu equipamentos, economia, etc. A aplicação desses métodos costuma ser complicada e requer participação humana. Além disso tal informação não está frequentemente disponível.

Para simplificar as previsões de médio-prazo e obter mais precisão, Feinberg, Hajagos, e Genethliou (2002, 2003) evitam o uso de informações de difícil disponibilidade, e desenvolvem um modelo estatístico para a carga por meio dos dados históricos. O modelo pode ser aplicado para previsões de médio e longo-prazos.

### 2.1.1.2 Métodos de Previsão de curto-prazo

Uma variedade de técnicas estatísticas e de inteligência computacional tem sido desenvolvida para previsão de carga de curto-prazo (De Gooijer e Hyndman (2006); Campos et al. (2007); Pandian et al. (2006); Ghiassi, Zimbra, e Saidane (2005b)). A seguir são, sucintamente, apresentados os métodos de previsão mais usados.

**Métodos de Regressão** - Engle, Mustafa, e Rice (1992) apresentaram vários modelos de regressão para a previsão de um dia à frente do pico de carga. Seus modelos incorporam influências determinísticas: feriados; influências estocásticas: média da carga; influência exógena: condições meteorológicas. Outras referências tais como, (Hyde e Hodnett (1997); Ruzic, Vuckovic, e Nikolic (1998); Charytoniuk, Chen, e Olinda (1998); Haida

e Muto (1994)) descrevem outras aplicações dos modelos de regressão para previsão de carga.

**Séries Temporais** - Esses métodos têm sido aplicados por décadas nas áreas de economia, processamento digital de sinais, como também na previsão da carga elétrica (Morettin e Tolo (2006); Box, Jenkins, e Reinsel (1994)). Em particular, ARMA, ARIMA, ARMAX, e ARIMAX são métodos clássicos mais usados em séries temporais.

Em Aslanargun et al. (2007) é realizada uma comparação entre os modelos ARIMA, Rede Neural e modelos híbridos usando a série temporal de turistas que chegam na Turquia. Seus resultados mostram que o modelo ARIMA apresenta bom desempenho, mas os modelos com componentes não-lineares apresentam melhores desempenho em relação ao ARIMA.

Chow e Leung (1996) apresentam uma rede neural não-linear autoregressiva integrada (NARI) para a previsão de curto-prazo (1 dia à frente) do consumo de energia de Hong Kong.

Fan e McDonald (1994) e Cho, Hwang, e Chen (1995) descrevem a implementação dos modelos ARIMAX (autoregressivo integrado de média móvel com entradas exógenas) para previsão de carga. Yang, Huang, e Huang (1996) usaram *programação evolucionária*<sup>1</sup> para identificar os parâmetros do modelo ARMAX (autoregressivo de média móvel com entradas exógenas) para previsão de uma semana à frente de dados horários de carga. Yang e Huang (1998) propuseram um modelo Fuzzy autoregressivo de média móvel com variáveis exógenas de entrada (FARMAX) para previsão de um ano à frente de dados horários de carga.

**Redes Neurais** - o uso de Redes Neurais artificiais (RNA) tem sido amplamente estudada na aplicação de previsão de carga desde 1990 (veja (Peng, Hubele, e Karady (1992))).

Bakirtzis, Petridis, Kiartzis, Alexiadis, e Maissis (1996) desenvolveram uma RNA baseada em um modelo de previsão de carga de curto-prazo para centro de controle de energia da Greek Public Power Corporation. As variáveis de entrada incluíam dados históricos de carga horária, temperatura, e o dia da semana. O modelo pôde prever perfis de carga de até sete dias à frente.

---

<sup>1</sup>Programação evolucionária é um método para simular a evolução e constitui um algoritmo de otimização estocástico (Fogel (1994)).

Papalexopoulos, Hao, e Peng (1994) também desenvolveram e implementaram uma RNA *feedforward* multi-camadas para previsão de curto-prazo do sistema de carga. No modelo três tipos de variáveis são usadas como entradas para a RNA: relação das estações do ano, condições climáticas, e dados históricos de carga.

Em Ghiassi, Saidane, e Zimbra (2005a) apresentam um modelo de Rede Neural dinâmico, nomeado de DAN2 (*dynamic artificial neural network model*). Ghiassi et al. (2005b) utilizaram essa metodologia para prever a carga elétrica em médio-prazo. A DAN2 foi usada para modelar e prever cargas, mensalmente, trimestralmente, e anualmente usando dados da Taiwan Power Company. Os resultados apresentados pelos autores apresentaram bom desempenho. Muitos estudos publicados usam RNAs em conjunto com outras técnicas de previsão, tais como, árvores de regressão (Mori e Kosemura (2001)), séries temporais (Chow e Leung (1996)) ou Lógica Fuzzy (Skarman e Georgiopoulos (1998)).

**Sistemas Especialistas (SE)** - Ho, Hsu, Chen, Lee, Liang, Lai, e Chen (1990) propuseram um SE baseado em conhecimento para previsão de carga de curto-prazo da Taiwan Power System. Conhecimentos do operador e observações horárias do sistema de carga foram empregadas para estabelecer onze tipos de dias diferentes. O algoritmo desenvolvido foi comparado com a metodologia de Box e Jenkins e apresentou bons resultados.

**Lógica Fuzzy** - Campos et al. (2007) desenvolveram um modelo Neuro-Fuzzy para previsão de curto-prazo do consumo de energia da cidade de Nova York. O modelo obtido apresentou desempenho satisfatório para previsões de até 48 passos (240 minutos) à frente. Os autores desse trabalho comparam os resultados obtidos com os resultados encontrados em Pandian et al. (2006), e concluem que a previsão usando o modelo Neuro-Fuzzy é mais eficiente. Referências (Kiartzis e Bakirtzis (2000); Miranda e Monteiro (2000); Skarman e Georgiopoulos (1998)) descrevem aplicações da Lógica Fuzzy para previsão de carga elétrica.

**Máquinas de Vetores de Suporte (SVM)**- Mohandes (2002) aplicou o método SVM para previsão de carga elétrica de curto-prazo. O autor

comparou seu desempenho com os modelos autoregressivos (AR). Os resultados indicam que a técnica SVM apresentou melhores resultados que os modelos AR. Chen, Chang, e Lin (2002) desenvolveram um modelo usando SVM para prever a demanda de carga diária. Li e Fang (2003) também usam um modelo SVM para previsão de carga de curto-prazo.

## 2.2 Identificação de Sistemas

A Identificação de Sistemas é a área do conhecimento que estuda técnicas alternativas de modelagem matemática. Nela há a necessidade de desenvolver formas de se obter modelos matemáticos a partir de dados observados e não exclusivamente partindo-se das equações que descrevem a física do processo (Ljung (1987)).

O problema de Identificação de Sistemas pode ser dividido em cinco etapas (Ljung (1987)):

- obtenção de dados de *experimentação* do sistema que se deseja modelar;
- aplicação de testes aos dados obtidos para *deteção de não-linearidades*;
- *deteção de estrutura* que será utilizada para representar o sistema;
- *estimação dos parâmetros* do modelo;
- *validação* do modelo obtido.

A experimentação consiste na determinação do sinal de entrada, na escolha da taxa de amostragem e na coleta dos dados para a estimação de parâmetros e ajuste do modelo (Söderström e Stoica (1989)). A existência de não-linearidades no sistema pode ser verificada a partir de testes não paramétricos, utilizando apenas os sinais coletados.

A deteção da estrutura para representação dos sistemas é de grande importância na Identificação de Sistemas. Nesta etapa termos desnecessários no modelo e que podem provocar instabilidade numérica e ocasionar comportamentos dinâmicos espúrios aos sistema são detectados e extraídos da estrutura do modelo (Billings, Chen, e Korenberg (1989); Aguirre e Billings (1995))

A estimação de parâmetros consiste em determinar os parâmetros de uma determinada estrutura a fim de que o modelo reproduza o comportamento do sistema original. Um critério comumente usado para estimação de parâmetros é o método dos Mínimos Quadrados que se resume na minimização da soma do quadrado das diferenças entre os dados e a saída estimada pelo modelo (Aguirre (2007); Corrêa (2001)).

Finalizando o processo de Identificação de Sistemas, o último passo é a validação. A validação consiste essencialmente em verificar se os modelos incorporam ou não as características de interesse do sistema original. Além disso, é interessante poder comparar os modelos entre si e decidir se há algum candidato significadamente melhor que os demais.

Todo o detalhamento do processo de Identificação de Sistemas pode ser visto em Aguirre (2007).

## 2.3 Representações Lineares e Não-Lineares

Existem diversas formas de representar o mesmo modelo matemático, ou seja, há várias formas em que as equações que descrevem o comportamento do sistema podem ser escritas. A forma em que um modelo matemático é expresso é às vezes chamada de *representação* (Aguirre (2007)).

Existem várias representações lineares e não lineares. Uma das mais importantes na modelagem de sistemas dinâmicos lineares é a função de transferência. Outra maneira útil de expressar modelos é usando representação em espaço de estados. Essas duas representações normalmente são apresentadas no contexto de sistemas contínuos no tempo. No caso de representações no contexto de sistemas lineares discretos pode-se citar os modelos ARX, ARMAX, ARMA, ARIMA. No caso de representações não-lineares citam-se os modelos de Hammerstein e de Wiener, modelo NARX, NARMAX, funções radiais de base, Redes Neurais Artificiais, Neuro-Fuzzy, dentre outras. As descrições dessas representações lineares e não-lineares são encontradas em Aguirre (2007).

Neste trabalho são usadas representações lineares e não lineares. O modelo ARIMA (linear), NARIMA (não-linear), Rede Neuro-Fuzzy (não-linear) e Rede Neural Artificial (não-linear) serão descritas no decorrer deste Capítulo.

### 2.3.1 Modelos ARIMA

Uma metodologia bastante utilizada na modelagem de séries temporais não-estacionárias foi proposta por Box e Jenkins (1970). Tal metodologia consiste em ajustar modelos auto-regressivos integrados de médias móveis, ARIMA (do inglês *AutoRegressive Integrated Moving Average*), uma extensão dos modelos lineares paramétricos ARMA (modelos auto-regressivos de médias móveis) amplamente divulgados na literatura científica (Aguirre (2007)).

Para a formulação do modelo ARIMA a partir de uma série temporal,  $y(k)$ , é preciso introduzir os seguintes operadores:

- Operador de atraso,  $q^{-1}$ , definido por

$$q^{-1}y(k) = y(k-1), \quad q^{-m}y(k) = y(k-m),$$

- Operador de avanço,  $q^{+1}$ , definido por

$$q^{+1}y(k) = y(k+1), \quad q^{+m}y(k) = y(k+m),$$

- Operador de primeira diferença assim definido

$$\Delta y(k) = y(k) - y(k-1) = (1 - q^{-1})y(k).$$

em que  $\Delta = 1 - q^{-1}$ .

- O operador soma,  $S$ , tem a seguinte expressão

$$\begin{aligned} Sy(k) &= \sum_{j=0}^{\infty} y(k-j) = y(k) + y(k-1) + \dots \\ &= (1 + q^{-1} + q^{-2} + \dots)y(k), \end{aligned}$$

Note que

$$Sy(k) = (1 - q^{-1})^{-1}y(k) = \Delta^{-1}y(k),$$

ou seja

$$S = \Delta^{-1}.$$

Várias séries temporais, especialmente econômicas e financeiras, são não-estacionárias (Morettin e Toloi (2006)), mas quando se toma a primeira diferença, normalmente a série resultante pode ser considerada estacionária<sup>2</sup>. Por exemplo, a série  $y(k)$  é não-estacionária, mas

$$W(k) = y(k) - y(k-1) = (1 - q^{-1})y(k) = \Delta y(k)$$

pode ser estacionária.

Se  $W(k) = \Delta^d y(k)$  for estacionária, pode-se representar  $W(k)$  por um modelo ARMA(p,q) da seguinte forma

$$\phi_p(q^{-p})W(k) = \theta_q(q^{-q})v(k), \quad (2.1)$$

sendo que  $v(k)$  é um processo aleatório branco (*ruído branco*).

Se  $W(k)$  é a primeira diferença de  $y(k)$ , então a série temporal  $y(k)$  é a soma (integral) de  $W(k)$  e pode ser modelada pelo seguinte modelo ARIMA

$$\phi_p(q^{-p})\Delta^d y(k) = \theta_0 + \theta_q(q^{-q})v(k), \quad (2.2)$$

de ordem  $(p,d,q)$ , normalmente referido como ARIMA(p,d,q).  $p$  e  $q$  são as ordens de  $\phi$  e  $\theta$ , respectivamente, isto é,  $\phi_p(q^{-p}) = 1 - \phi_1(q^{-1}) - \dots - \phi_p(q^{-p})$  e  $\theta_q(q^{-q}) = 1 - \theta_1(q^{-1}) - \dots - \theta_q(q^{-q})$ . O parâmetro  $\theta_0$  representa a constante no modelo.

Explicitando a Equação 2.2 em função de  $y(k)$  e considerando  $d = 1$ , tem-se o modelo ARIMA(p,1,q):

$$\begin{aligned} y(k) = & (1 + \phi_1)y(k-1) + (\phi_2 - \phi_1)y(k-2) + \dots \\ & + (\phi_p - \phi_{p-1})y(k-p) - \phi_p y(k-p-1) \\ & + \theta_0 - \theta_1 v(k-1) - \theta_2 v(k-2) \\ & - \dots - \theta_q v(k-q) + v(k) \end{aligned} \quad (2.3)$$

O modelo 2.3 é um modelo linear discreto no tempo que explica o valor da saída  $y(k)$  em função de valores prévios da própria saída. O modelo NARIMA (do inglês *Nonlinear Auto Regressive Integrated Moving Average Model*) é representado da seguinte forma:

<sup>2</sup>De acordo com Aguirre (2007) um processo é estacionário se as leis de probabilidade que o regem não variam com o tempo. Pode-se dizer que um processo estacionário está em equilíbrio estatístico. O conceito de estacionariedade está intimamente ligado ao de *invariância*.

$$\begin{aligned}
y(k) = & F^\ell[(1 + \phi_1)y(k-1) + (\phi_2 - \phi_1)y(k-2) + \dots \\
& + (\phi_p - \phi_{p-1})y(k-p) - \phi_p y(k-p-1) \\
& + \theta_0 - \theta_1 v(k-1) - \theta_2 v(k-2) \\
& - \dots - \theta_q v(k-q) + v(q)]
\end{aligned} \tag{2.4}$$

sendo que,  $\ell$  é o grau de não-linearidade, e  $F[.]$  é uma função não-linear que pode assumir uma variedade de formas, tais como racional e polinomial (Aguirre (2007)). Esta última é satisfatória em muitos casos (Corrêa (2001); Coelho (2002)). Uma das mais importantes vantagens do modelo polinomial NARIMA é que este é linear nos parâmetros, possibilitando assim, o uso do método dos mínimos quadrados na estimação dos mesmos.

Note que no modelo (2.1) todas as raízes de  $\phi_p(q^{-p})$  estão fora do círculo de raio unitário (Box et al. (1994); Morettin e Tolo (2006)). Portanto, pode-se escrever o modelo (2.2) da seguinte forma

$$\varphi(q^{-q})y(k) = \theta_q(q^{-q})v(k), \tag{2.5}$$

em que  $\varphi(q^{-q})$  é um operador autoregressivo não-estacionário, de ordem  $p+d$ , com  $d$  raízes iguais a um, sobre o círculo de raio unitário, e as restantes  $p$  fora do círculo, isto é,

$$\varphi(q^{-q})y(k) = \phi_p(q^{-p})\Delta^d = \phi_p(q^{-p})(1 - q^{-p}). \tag{2.6}$$

O modelo (2.2) supõe que a  $d$ -ésima diferença da série  $y(k)$  pode ser representada por um modelo ARMA. Na maioria dos casos, a primeira derivada,  $d = 1$ , é suficiente para tornar a série temporal estacionária, enquanto que se  $d > 2$  a relação sinal-ruído dos dados na fase de identificação pode ser alterada prejudicialmente.

Quando a série é corretamente diferenciada, a variância da série transformada diminui, ao passo que o excesso de diferenças aumentará a variância. Detalhes sobre a diferenciação de séries temporais quanto ao nível e a inclinação podem ser vistos em Morettin e Tolo (2006).

### 2.3.1.1 Previsão com Modelos ARIMA

Nesta seção será apresentado um método de previsão utilizando a representação do modelo ARIMA por Equação de diferenças (discutido an-



teriormente).

A idéia central é prever o valor de  $y(k+h)$  supondo que se conhece as observações passadas, isto é,  $\dots, y(k-2), y(k-1), y(k)$ . Nesse contexto, o instante  $k$  é chamado origem das previsões e  $h$  é o horizonte das previsões.

Esse método de previsão será demonstrado utilizando um exemplo com um modelo ARIMA(2,1,2). Suponha que o modelo construído para a série temporal  $y(k)$  seja um ARIMA(2,1,2) com a seguinte forma

$$(1 - \phi_1 q^{-1} - \phi_2 q^{-2})(1 - q^{-1})y(k) = (1 - \theta_1 q^{-1} - \theta_2 q^{-2})v(k),$$

então

$$(1 - \phi_1 q^{-1} - \phi_2 q^{-2})(1 - q^{-1})y(k+h) = (1 - \theta_1 q^{-1} - \theta_2 q^{-2})v(k+h).$$

Considerando que

$$\begin{aligned} q^{-1}y(k) &= y(k-1), \text{ e} \\ y(k+h-1) &= \hat{y}_k(h-1), \end{aligned}$$

então

$$\begin{aligned} y(k+h) &= (1 + \phi_1)y(k+h-1) - (\phi_1 - \phi_2)y(k+h-2) - \\ &\quad - \phi_3 y(k+h-3) + v(k+h) - \theta_1 v(k+h-1) - \\ &\quad - \theta_2 v(k+h-2). \end{aligned}$$

Logo,

$$\begin{aligned} \hat{y}_k(h) &= (1 + \phi_1)y_k(h-1) - (\phi_1 - \phi_2)y_k(h-2) - \phi_3 y_k(h-3) + \\ &\quad + v_k(h) - \theta_1 v_k(h-1) - \theta_2 v_k(h-2). \end{aligned}$$

Assim tem-se que

$$\begin{aligned} \hat{y}_k(1) &= (1 + \phi_1)y(k) - (\phi_1 - \phi_2)y(k-1) - \phi_3 y(k-2) - \theta_1 v(k) - \theta_2 v(k-1), \\ \hat{y}_k(2) &= (1 + \phi_1)\hat{y}_k(1) - (\phi_1 - \phi_2)y(k) - \phi_3 y(k-1) - \theta_2 v(k), \\ \hat{y}_k(3) &= (1 + \phi_1)\hat{y}_k(2) - (\phi_1 - \phi_2)\hat{y}_k(1) - \phi_3 y(k), \\ \hat{y}_k(h) &= (1 + \phi_1)\hat{y}_k(h-1) - (\phi_1 - \phi_2)\hat{y}_k(h-2) - \phi_3 \hat{y}_k(h-3), \quad h \geq 4 \end{aligned}$$

Das predições acima, observa-se que:

- os termos de médias móveis ( $v(k)$ ) desaparecem para  $h > q$  ( $q$  é a ordem da parte MA), pois, não existe este valor para  $h > q$ ;
- para calcular  $\hat{y}_k(h)$  é preciso de  $\hat{y}_k(h-1), \hat{y}_k(h-2), \dots$  que são calculados recursivamente.

### 2.3.1.2 Atualizando as previsões

Inicialmente, determina-se as previsões  $y_k(h+1)$  a partir de duas origens:

$$(a) \quad k+1: \hat{y}_{k+1}(h) = \psi_h v(k+1) + \psi_{h+1} v(k) + \psi_{h+2} v(k-1) + \dots, \quad (2.7)$$

$$(b) \quad k: \hat{y}_k(h+1) = \psi_{h+1} v(k) + \psi_{h+2} v(k-1) + \dots \quad (2.8)$$

Subtraindo (2.7) de (2.8) e após algumas manipulações obtém-se

$$\hat{y}_{k+1}(h) = \hat{y}_k(h+1) + \psi_h v_{t+1}. \quad (2.9)$$

Dessa maneira, a previsão de  $y_k(h+1)$ , feita no instante  $k$ , pode ser atualizada quando um novo dado,  $y(k+1)$ , é observado. Assim, a previsão  $y_k(h+1)$ , na origem  $k+1$ , pode ser realizada, adicionando-se à  $\hat{y}_k(h+1)$  um múltiplo do erro de previsão  $v(k+1) = y(k+1) - \hat{y}_k(1)$ .

## 2.3.2 Rede Neuro-Fuzzy

Uma rede neuro-fuzzy pode ser definida como sendo um *sistema fuzzy*<sup>3</sup> que é treinado por algum algoritmo derivado da teoria de aprendizado de máquina. Este tipo de rede é o resultado da união destes dois modelos. Com isto, tem-se a capacidade das redes neurais em reconhecimento e classificação, sem esquecer da robustez e habilidade de generalização.

Por outro lado, têm-se os sistemas fuzzy que, por meio de suas regras e conjuntos fuzzy, facilitam o entendimento do problema, porque modela o ambiente por meio de uma linguagem próxima da usada pelos especialistas. Por exemplo, os sistemas fuzzy têm demonstrado serem bastante

<sup>3</sup>A lógica Fuzzy, desenvolvida por Lofti A. Zadeh na Universidade da Califórnia em Berkeley na década de 60, combina teoria probabilística, inteligência artificial e redes neurais para que possa representar o pensamento humano, ou seja, ligar a lingüística e a inteligência humana, pois, muitos conceitos são melhores definidos por palavras do que pela matemática (Zadeh (1965)).

consistentes e confiáveis, quando aplicados a sistemas de controle, reconhecimento de padrões, interpolação de funções e previsão (Cox (1994); Jang (1993)).

O primeiro aspecto que torna possível essa integração é o fato de as redes neurais e os sistemas fuzzy serem aproximadores universais. Portanto, a rede neuro-fuzzy mantém essa característica fundamental para o desenvolvimento das aplicações. Outro aspecto é que ambos os sistemas manipulam dados incompletos, imprecisos, complexos e não-lineares, além de alto grau de abstração de informação irrelevante para o sistema.

As redes neuro-fuzzy herdam da rede neural sua capacidade de aprendizagem, generalização, classificação, robustez, adaptação e agrupamento de dados comuns em *clusters*. As redes neurais sempre foram vistas como uma caixa preta, ou seja, não é possível obter o conhecimento sobre as decisões da rede no processo de treinamento. O advento das redes neuro-fuzzy quebra este conceito de caixa preta das redes neurais, porque o comportamento deste modelo pode ser entendido através da observação das variáveis lingüísticas, das funções de pertinência, dos relacionamentos entrada/saída e das próprias regras fuzzy, as quais podem explicar facilmente o funcionamento do sistema, devido a simplicidade e proximidade com a linguagem humana.

A arquitetura ANFIS (*Adaptive Neuro-Fuzzy Inference System*) foi proposta por Jang (1993), e possivelmente seja o sistema neuro-fuzzy mais conhecido e empregado devido as suas qualidades de bom classificador e previsor<sup>4</sup>.

Estruturalmente, verifica-se que a única limitação na configuração da rede para os sistemas adaptativos como o ANFIS é que esta deve ser do tipo feedforward (Jang, Sun, e Mizutani (1997)). Apesar dessa mínima restrição, as redes adaptativas podem ser empregadas em uma ampla variedade de aplicações, como na tomada de decisão, no processamento de sinais e em controle. A seguir descreve-se a arquitetura do sistema de inferência ANFIS.

**Arquitetura** - Por simplicidade, considere um sistema de inferência fuzzy com duas entradas  $x$  e  $y$  e uma saída  $z$ . Para um modelo fuzzy

---

<sup>4</sup>Uma possível extensão ao ANFIS é o modelo CANFIS (*Coactive Neuro-Fuzzy Modelling*) proposto por Mizutani e Jang (1995). Uma das vantagens deste modelo é que ele aceita múltiplas entradas e múltiplas saídas.

Sugeno de primeira ordem, um conjunto usual de regras “se-então” é a seguinte:

Regra 1: **Se**  $x$  é  $A_1$  **e**  $y$  é  $B_1$  **então**  $f_1(x, y) = p_1x + q_1y + r_1$ ,

Regra 2: **Se**  $x$  é  $A_2$  **e**  $y$  é  $B_2$  **então**  $f_2(x, y) = p_2x + q_2y + r_2$ .

A Figura 2.1 ilustra o mecanismo de raciocínio para esse modelo fuzzy Sugeno de primeira ordem com duas entradas e duas regras; a arquitetura ANFIS equivalente é aquela mostrada na Figura 2.2, onde os nós da mesma camada têm funções similares. A seguir, passa-se a descrever cada camada do modelo.

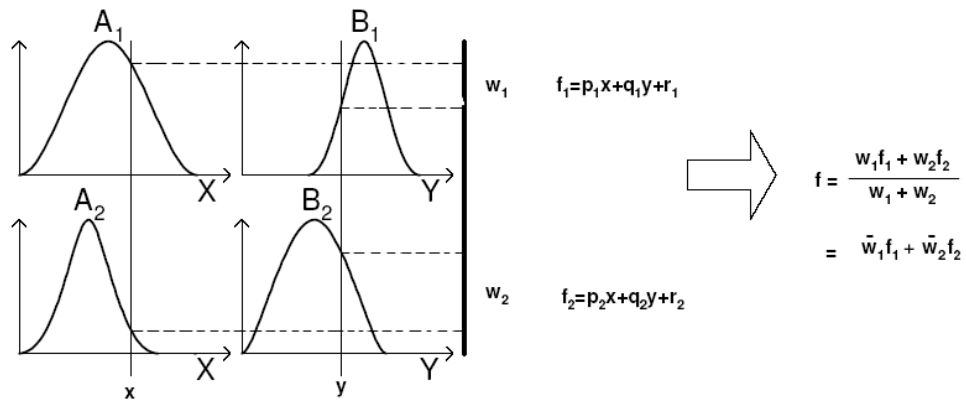


Figura 2.1: Modelo fuzzy Sugeno com duas entradas e e duas regras (Jang (1993)).

**Camada 1** - cada nó nesta camada é um nó adaptativo com uma função nó do tipo (denota-se a saída do  $i$ -ésimo nó na camada  $l$  como  $O_{l,i}$ ):

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x), \quad \text{para } i = 1, 2, \text{ ou} \\ O_{1,i} &= \mu_{B_{i-2}}(y), \quad \text{para } i = 3, 4, \end{aligned} \quad (2.10)$$

sendo  $x$  (ou  $y$ ) é a entrada para o nó  $i$  e  $A_i$  (ou  $B_{i-2}$ ) é um rótulo lingüístico (como “pequeno” ou “grande”) associado a este nó. Em outras palavras,  $O_{l,i}$  é a função de pertinência parametrizada apropriadamente, tal como a função sino (“bell”) generalizada:

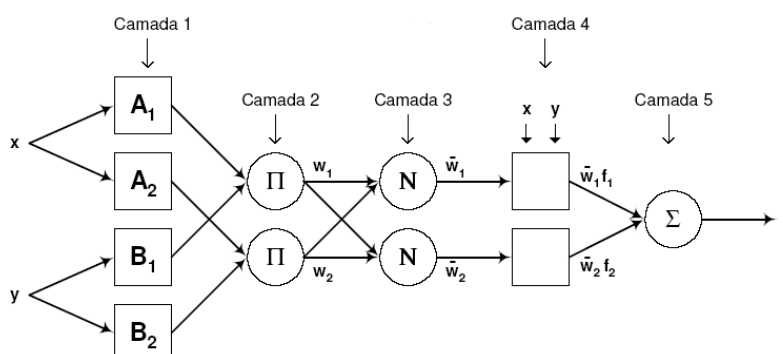


Figura 2.2: Arquitetura ANFIS equivalente para o modelo da Figura 2.1 (Jang (1993))

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x-c_i}{a_i} \right|^{2b}}, \quad (2.11)$$

$a_i, b_i, c_i$  é o conjunto de parâmetros que determinam função sino.

**Camada 2** - cada nó nesta camada é um nó fixo rotulado por  $\Pi$ , cuja saída é o produto de todos os sinais da entrada:

$$O_{2,i} = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i = 1, 2. \quad (2.12)$$

Cada nó de saída dessa camada representa o nível de disparo (*firing strength*),  $w_i$ , da regra. Em geral, qualquer outro operador T-norma pode ser usado como função nó nessa camada.

**Camada 3** - cada nó  $i$  nesta camada é um nó rotulado N. A partir dessa camada tem-se o processo de defuzzificação. A saída desse nó, chamada nível de disparo normalizada, é dada pela razão entre o  $i$ -ésimo nível de disparo da regra e a soma de todas os níveis de disparo:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (2.13)$$

**Camada 4** - Cada nó  $i$  nesta camada é um nó adaptativo com uma

função nó do tipo

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i), \quad (2.14)$$

em que  $\bar{w}_i$  é um nível de disparo normalizado da camada 3 e  $p_i$ ,  $q_i$ ,  $r_i$  é o conjunto de parâmetros deste nó. Tem-se, então, um produto entre os níveis de disparo normalizados e o valor do conseqüente da regra em si. Por isso, parâmetros nesta camada são denominados parâmetros conseqüentes.

**Camada 5** - Cada nó  $i$  nesta camada é um nó fixo rotulado  $\Sigma$ , que calcula a saída geral do sistema como a soma de todos os sinais de sua entrada:

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}. \quad (2.15)$$

Note-se que a arquitetura dessa rede adaptativa não é única; pode-se combinar as camadas 3 e 4 para se obter uma rede equivalente com somente quatro camadas. Da mesma forma, é possível também normalizar os pesos na última camada.

Finalmente, a Figura 2.3 (a) mostra uma arquitetura ANFIS que é equivalente a um modelo Sugeno de primeira ordem de duas entradas com nove regras, onde cada entrada tem três funções de pertinência associadas, e a Figura 2.3 (b) ilustra como o espaço de entrada bi-dimensional é particionado em nove regiões fuzzy sobrepostas (grid), cada uma delas governadas por uma regra fuzzy se-então. A identificação dos parâmetros do modelo ANFIS é realizada, em geral, empregando algoritmos de aprendizagem híbrida, isto é, combinando estimação de mínimos quadrados com retropropagação (Mathworks (2001)).

O treinamento do modelo ANFIS é realizado em duas etapas. Na primeira etapa os parâmetros dos antecedentes ficam fixos e os conseqüentes são estimados pelo método dos mínimos quadrados ordinários (MQO)<sup>5</sup>. Já na etapa 2 os parâmetros dos conseqüentes ficam fixos e os parâmetros dos antecedentes são ajustados pelo algoritmo Gradiente Descendente<sup>6</sup> (Jang

<sup>5</sup>Veja detalhes do método MQO em (Aguirre (2007)).

<sup>6</sup>Veja detalhes do algoritmo Gradiente Descendente em (Braga, Carvalho, e Ludemir (2000)).

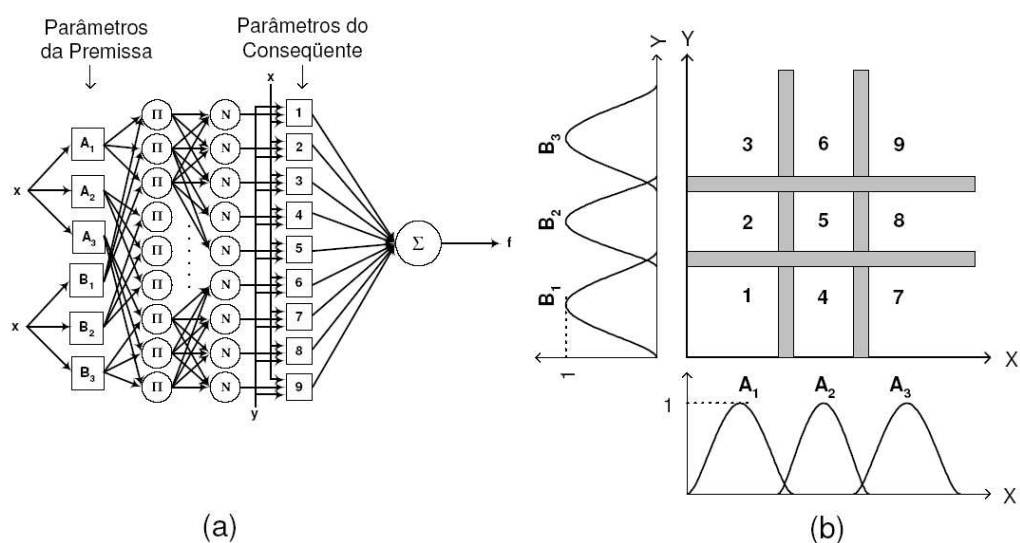


Figura 2.3: Arquitetura ANFIS para (a) modelo Sugeno e (b) espaço de entrada (Jang (1993)).

et al. (1997)).

Note que a arquitetura dessa rede adaptativa não é única; pode-se combinar as camadas 3 e 4 para se obter uma rede equivalente com somente quatro camadas. Da mesma forma é possível também normalizar os pesos na última camada, a Figura 2.3 ilustra uma ANFIS deste tipo. A estrutura mostrada na Figura 2.4 é utilizada neste trabalho.

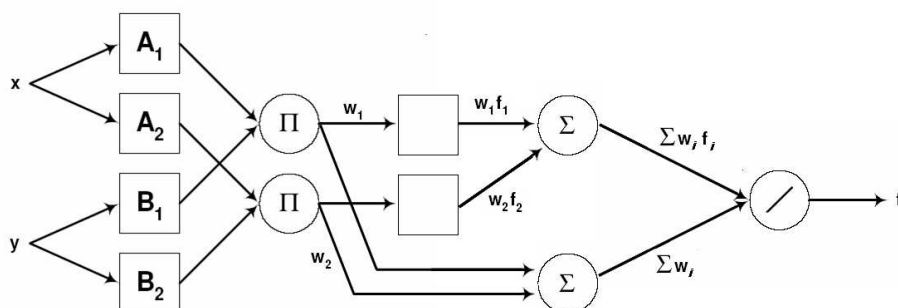


Figura 2.4: Arquitetura ANFIS para o modelo Sugeno, onde a normalização dos pesos é executada na última camada (Jang (1993)).

### 2.3.3 Rede Neural Artificial

Esta seção faz uma sucinta apresentação das Redes Neurais Artificiais (RNAs). Maiores informações a respeito do assunto podem ser encontradas nos textos (Braga et al. (2000); Haykin (2001)).

As redes neurais artificiais (RNAs) são estruturas matemáticas capazes de aprender, memorizar e generalizar determinadas situações e problemas a elas apresentadas. Uma Rede Neural é formada pela interconexão de um determinado número de nodos (neurônios). Braga et al. (2000) dizem que os neurônios são unidades de processamento não-linear. Toda fundamentação teórica referente às redes neurais artificiais pode ser vista com mais detalhes em Braga et al. (2000); Haykin (2001).

Os problemas, aos quais as RNAs são aplicáveis, consistem basicamente em situações onde existam dados, experimentais ou não, que são apresentados à rede em uma etapa denominada de “*treinamento*”, onde os pesos sinápticos são ajustados com intuito de que alguma tarefa seja realizada. Além da capacidade de aprendizado a partir de exemplos, as redes possuem a capacidade de generalização, que consiste na habilidade da mesma em apresentar soluções para dados distintos (dados de “*validação*”), diferentes dos dados apresentados na etapa de treinamento. Braga et al. (2000) citam que as principais características das RNAs são: (i) o aprendizado a partir de exemplos; (ii) adaptabilidade; (iii) capacidade de generalização; (iv) tolerância a falhas; e (v) rápida implementação.

No contexto deste presente trabalho, as RNAs serão aplicadas para a aproximação de funções e para a previsão de valores futuros da série de consumo de energia de New England e do Estado de Minas Gerais.

#### 2.3.3.1 Perceptron Multicamadas

A Perceptron Multicamada é mais conhecida por sua sigla MLP, do inglês *Multilayer Perceptron*. A característica marcante de uma rede MLP é que ela pode ter várias camadas ocultas. A MLP basicamente consiste de uma camada de nós, (fontes de entrada), uma ou mais camadas de nós processadores ou computacionais (neurônios) ocultos e uma camada de saída também composta por nós computacionais. A camada formada por neurônios ocultos (camada oculta) recebe este nome porque não há acesso da entrada e nem da saída sobre esta camada.

As redes MLP possuem capacidade de resolver problemas não-lineares.



As não-linearidades são inseridas nos modelos por meio das funções de ativação não-lineares de cada neurônio e da composição de sua estrutura em camadas sucessivas. A função de ativação, também chamada de função de transferência, é uma função matemática que, aplicada à combinação linear entre as variáveis de entrada e pesos que chegam a determinado neurônio, retorna ao seu valor de saída, ou seja, determina a relação entre entradas e saídas de cada neurônio da rede. Neste projeto é utilizada a função tangente hiperbólica, tanto para a camada de entrada quanto para a de saída, dada por

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (2.16)$$

A função de ativação tangente hiperbólica é geralmente usada para a previsão de séries temporais (Zhang, Patuwo, e Hu (1998)).

A Equação 2.16 se aproxima de uma função degrau, porém é passível de diferenciação tornando possível o treinamento por métodos que usam o gradiente. A saída  $y$  para uma rede MLP com uma camada escondida é dada por

$$y = \phi \left[ \sum_{j=1}^N W_j \cdot f \left( \sum_{i=1}^n x_i w_{ij} + b_i \right) \right] + B \quad (2.17)$$

Haykin (2001) cita que o modelo Neural multi-camadas, sendo um aproximador universal de funções, é capaz de aproximar qualquer função contínua desde que o número de neurônios na camada escondida seja suficientemente grande.

O algoritmo de treinamento de retropropagação do erro, mais conhecido como *backpropagation* (Rumelhart e McClelland (1986)), é o algoritmo mais popular para o treinamento de RNAs multicamadas. Basicamente o algoritmo consiste de dois passos: a propagação e a retropropagação. No primeiro passo, um vetor de entradas é aplicado à camada de entrada e o seu efeito se propaga pela rede (Equação 2.17) produzindo um conjunto de saídas. A resposta obtida pela rede é subtraída da resposta desejada para produzir um sinal de erro. O segundo passo consiste em propagar esse sinal de erro na direção contrária às conexões sinápticas, ajustando-as de

forma a aproximar as saídas da rede das saídas desejadas.

Em uma rede MLP, basicamente, tem-se como parâmetros de escolha: (i) o número de neurônios na camada de entrada; (ii) o número de camadas escondidas e o número de neurônios nestas camadas; (iii) o número de neurônios na camada de saída. Apesar da literatura citar alguns algoritmos para realizar a escolha destes parâmetros, não existe uma solução geral que atenda a todos os casos. Dessa forma, essa escolha torna-se parte do problema e a solução varia de acordo com cada caso e de acordo com a complexidade de cada processo.

Como o problema deste trabalho se resume a previsão de séries temporais, a escolha do número de neurônios na camada de entrada trata-se da escolha dos atrasos (lags) da série temporal. Novamente este é um fator que depende muito das características do processo em estudo.

### 2.3.3.2 Controle de generalização

Um ponto importante a ser observado no treinamento de redes MLP é o controle de generalização. Sabe-se que quando uma rede apresenta complexidade inferior à complexidade do problema ela não conseguirá se ajustar aos dados e irá apresentar erros elevados, tanto no treinamento, quanto na validação. Essa é uma questão que deve ser observada no momento do treinamento da rede. Uma forma de se tratar este problema é o uso do algoritmo chamado *Early Stopping* ou Parada Precoce (Braga et al. (2000)).

O processo de treinamento de uma RNA pode ser dividido em duas etapas distintas. Na primeira o erro de treinamento e o erro de validação diminuem mostrando que a rede está apreendendo as principais características dos dados apresentados. No segundo momento o erro de treinamento continua a diminuir enquanto que o erro de validação tende a aumentar. Nesse ponto a rede passa a se especializar aos dados do treinamento acarretando a perda da capacidade de generalização e no conseqüente aumento do erro de validação. O algoritmo sugere que o treinamento se encerre quando o erro de validação começa a aumentar.

## 2.4 Comentários finais

Este Capítulo teve como principal finalidade familiarizar o leitor com técnicas de aproximação de sistemas lineares e não-lineares. Foi discutido também fatores que afetam a precisão das previsões, tais como, fatores climáticos, classe de consumidores, assim como a economia, dentre outros. Métodos de previsão de carga usam modelagem matemática avançada.

Uma breve revisão bibliográfica foi apresentada mostrando as principais técnicas que estão sendo abordadas na literatura para a previsão de curto, médio e longo-prazos do consumo de energia elétrica.

Alguns aspectos abordados neste Capítulo serão usados nos Capítulos seguintes. No próximo Capítulo é apresentada a metodologia utilizada neste trabalho.



# Metodologia

---

*"Todos modelos são errados, mas alguns são úteis."*

G.E.P. Box

Este Capítulo apresenta a metodologia utilizada para a modelagem das séries temporais de consumo de energia elétrica da cidade de New England e do Estado de Minas Gérias.

Ao longo deste Capítulo serão abordados os principais índices de desempenho que são utilizados para quantificar o quão satisfatório são os resultados das previsões dos modelos. Em especial dar-se-á ênfase na literatura, que trata especificamente da previsão de longo-prazo no setor energético. A descrição das séries temporais de New England e de Minas Gerais serão mostradas. No caso de Minas Gerais, será mostrada a metodologia para se obter a série de consumo total.

A metodologia usada para separar as componentes de uma série temporal, o filtro TCS (Mohr (2005)), será vista. Logo após, os procedimentos para obtenção de modelos ARIMA, NARIMA, RNF e RNA serão descritos.

## 3.1 Índices de desempenho

Em problemas de previsão de séries temporais, uma importante tarefa é a de quantificar a qualidade da predição obtida. Isso permite, por exemplo, comparar diversos algoritmos e diversas estruturas de modelos utilizando índices de desempenho. A seguir são apresentados alguns dos índices de desempenho mais utilizados em problemas de previsão de séries temporais.

Antes de rever as definições de alguns índices, é necessário a seguinte nomenclatura: a série temporal a ser predita será indicada por  $y(k)$ , e

a previsão por  $\hat{y}(k)$ . A previsão usada neste trabalho é a predição de  $k$  passos-a-frente, em que  $k$  pode assumir valores de 1 até o horizonte de previsão. Os cinco erros mais usados na literatura são:

- O erro médio (ME do inglês *mean error*)

$$ME = \frac{1}{N} \sum_{k=1}^N y(k) - \hat{y}(k), \quad (3.1)$$

- O erro absoluto médio (MAE do inglês *mean absolute error*)

$$MAE = \frac{1}{N} \sum_{k=1}^N |y(k) - \hat{y}(k)|, \quad (3.2)$$

- O erro quadrático médio (MSE do inglês *mean squared error*)

$$MSE = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2, \quad (3.3)$$

- O erro percentual médio (MPE do inglês *mean percentage error*)

$$MPE = \frac{1}{N} \sum_{k=1}^N \frac{y(k) - \hat{y}(k)}{y(k)} \times 100, \quad (3.4)$$

- O erro percentual absoluto médio (MAPE do inglês *mean absolute percentage error*)

$$MAPE = \frac{1}{N} \sum_{k=1}^N \left| \frac{y(k) - \hat{y}(k)}{y(k)} \times 100 \right|, \quad (3.5)$$

Os índices de erro (3.1)-(3.5) servem, por exemplo, para comparar dois previsores. Entretanto, tais índices não indicam se o desempenho de um determinado previsor é "razoável". Dependendo do nível da parcela estocástica em uma série temporal, um  $MAPE = 10\%$  pode ser um ótimo desempenho. Semelhantemente, tomando-se uma série puramente determinística e estacionária, um  $MAPE = 1\%$  pode indicar um péssimo previsor. A fim de prover índices que permitam avaliar a qualidade de

um previsor relativa aos dados, define-se a raiz do erro quadrático médio normalizado (RMSE do inglês *root mean square error*)

$$RMSE = \frac{\sqrt{\sum_{k=1}^N (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^N (y(k) - \bar{y})^2}}, \quad (3.6)$$

em que  $\bar{y}$  é o valor médio de  $y(k)$ , sendo que a média é calculada na janela de identificação (estimação de parâmetros ou treinamento). A expressão (3.6) mostra que o erro de previsão do modelo sendo avaliado (numerador na Equação (3.6)) é normalizado pelo erro cometido por um *preditor trivial*, que nesse caso nada mais é do que o valor médio dos dados. Assim, valores  $RMSE \approx 1$  indicam que a ação do previsor sendo avaliado não é significativamente melhor do que simplesmente tomar o valor médio dos dados,  $\bar{y}$ , e considerá-lo como previsão. Deve ser notado que a utilidade do índice RMSE aumenta à medida que a componente estocástica da série temporal  $y(k)$  se torna dominante.

Para dados em que a(s) componente(s) determinística(s) é(são) dominante(s), outros índices devem ser usados. Uma alternativa interessante é

$$U_{Theil} = \frac{\sqrt{\sum_{k=1}^N (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^N (y(k) - y(k-1))^2}}, \quad (3.7)$$

que é conhecido como estatística *U de Theil* (Theil (1958)). A interpretação das Equações 3.6 e 3.7 é simples. No primeiro caso, descrito acima, o índice compara as previsões do modelo com a média temporal do sinal, que é usada como *preditor trivial*. Semelhantemente, na estatística *U de Theil* o preditor trivial é o valor anterior da série temporal. Em ambos os casos, valores menores do que a unidade indicam um melhor desempenho em relação ao preditor padrão considerado (média ou um passo à frente). Uma boa comparação de vários índices pode ser encontrada em (Fair (1986)).

Deve ser notado que o previsor trivial no *U de Theil* é apenas uma revisão de um passo à frente. Isso torna o previsor trivial difícil de ser superado quando se deseja fazer a previsão vários passos a frente.

Vale observar que para avaliar corretamente o desempenho de um

determinado previsor, qualquer que seja o índice usado, o mesmo precisa ser calculado em um trecho de dados de *validação*, ou seja dados que não foram usados na construção do modelo.

Na literatura especializada, artigos que trataram especificamente da predição de longo-prazo no setor energético utilizaram os seguintes critérios de erro: MPE (Al-Saba e El-Amin (1999)), max(MPE) (Al-Saba e El-Amin (1999)), erros nas unidades de medição (MWh) (Da, Jiangyan, e Jilai (2000)), MSE (Ang e Goh (1991)), MAPE (Fung e Tummala (1993)). Aqui vale a pena ressaltar que em (Kandil, El-Debeiky, e Hasanien (2001)) seis diferentes índices de erro (os cinco das eqs. 3.1-3.5 e mais o desvio padrão dos erros) foram usados na avaliação de previsões de longo-prazo. Em Makridakis, Wheelwright, e Myndman (1998) a estatística *U de Theil* é utilizada para avaliar o desempenho de diversos previsores em problemas de predição de longo-prazo de diversas grandezas.

Como pode ser constatado, ao contrário do que acontece em trabalhos na área de previsão de curtíssimo prazo, onde o índice MAPE domina, não há preferência entre os autores sobre qual índice de erro utilizar. Isso sugere que mais de um índice deva ser levado em conta na avaliação de diferentes métodos de previsão. Neste trabalho os índices MPE (3.4), MAPE (3.5), RMSE (3.6) e *U de Theil* (3.7) são considerados para avaliar o desempenho das previsões.

## 3.2 Descrição dos dados

Neste trabalho são realizados dois estudos de caso distintos. No primeiro estudo, a previsão do consumo de energia elétrica usando dados da companhia de elétrica de New England, no USA, é realizada. E no segundo estudo é realizada a previsão do consumo de energia elétrica no Estado de Minas Gerais, Brasil. A primeira série temporal é usada para verificar a capacidade preditiva dos modelos, uma vez que o histórico de dados da segunda série é muito pequeno, dificultando, assim, a validação. As séries de consumo são apresentadas nas próximas seções.

### 3.2.1 Consumo de New England

Os dados de consumo da companhia energética de New England estão disponibilizados na internet ([www.iso-ne.com](http://www.iso-ne.com)), e compreendem o con-



sumo horário de energia do ano de 1980 a 1998. Os arquivos disponíveis estão no formato Excel e cada arquivo contém 8760 amostras (alguns contêm 8784 amostras, referentes aos anos bissextos no período de 1980 a 1998). Assim, as amostras de consumo horário relativas ao ano de 1980 são encontradas no arquivo `HourlyData1980.xls`, e sucessivamente até o ano de 1998 (`HourlyData1998.xls`), totalizando dezoito arquivos. Estes arquivos foram unidos no Matlab e salvos no arquivo `NE_hourly.mat`, com 166560 amostras horárias (dezoito anos). Como os dados de carga estão agrupados por hora, o valor de demanda e de consumo de cada ponto é numericamente o mesmo. Sendo assim, os dados podem ser agrupados por dia e por mês<sup>1</sup>.

A série temporal de consumo de energia horária total da companhia elétrica de New England é mostrada na Figura (3.1) e a Figura (3.2) exhibe o consumo total de energia agrupado mensalmente.

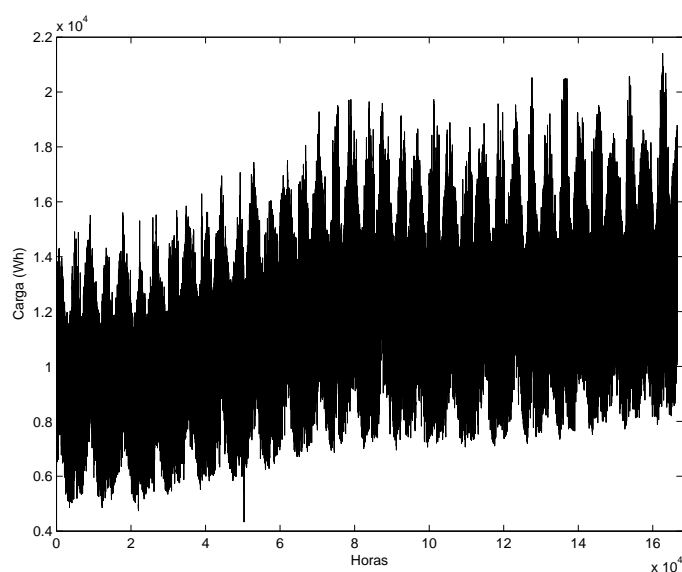


Figura 3.1: Consumo horário total da cidade de New England referente ao período que se estende de janeiro de 1980 a dezembro de 1998. Fonte: ISO-NE.

O agrupamento mensal dos dados da companhia de New England totaliza 228 amostras.

<sup>1</sup>Todas as análises que são realizadas neste trabalho levam em consideração os dados agrupados mensalmente.

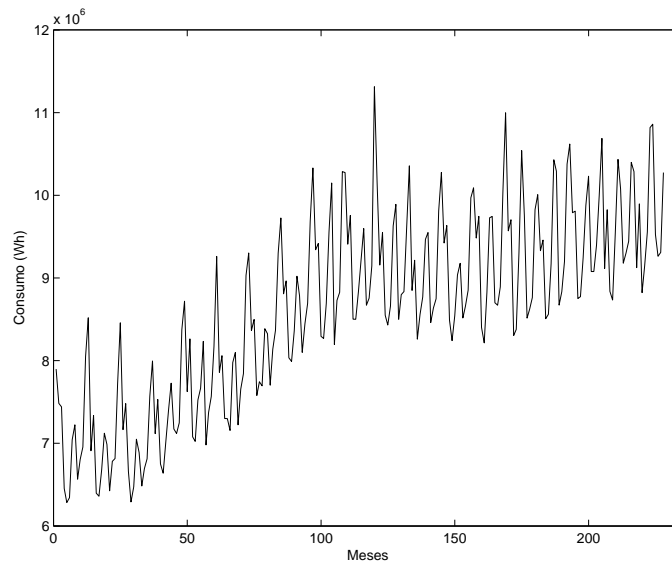


Figura 3.2: Consumo mensal total da cidade de New England referente ao período que se estende de janeiro de 1980 a dezembro de 1998. Fonte: ISO-NE

### 3.2.2 Consumo de Minas Gerais

No segundo estudo de caso são usados dados da Companhia Energética de Minas Gerais, a CEMIG-Distribuição<sup>2</sup>. A base de dados disponibilizada pela CEMIG contempla os valores mensais de consumo de energia em watt-hora (*wh*) entre dezembro de 1996 e julho de 2006. Para cada classe de consumidores, determinadas internamente pela CEMIG, existe uma série contendo os valores mensais no horizonte temporal citado. As séries disponíveis são mostrados abaixo:

- residencial ( $y_{re}$ );
- comercial ( $y_{co}$ );
- industrial ( $y_{in}$ );
- rural ( $y_{ru}$ );
- poder público ( $y_{pp}$ );

<sup>2</sup>A partir desse ponto CEMIG-Distribuição será citada apenas como CEMIG.

- iluminação pública ( $y_{ip}$ );
- serviço público ( $y_{sp}$ );
- consumo próprio ( $y_{cp}$ ).

A série de consumo total de energia consiste no somatório do consumo referente às demais séries e é definida por

$$y_{to} = y_{re} + y_{co} + y_{in} + y_{ru} + y_{pp} + y_{ip} + y_{sp} + y_{cp}.$$

Em uma etapa anterior os dados passaram por um processo de filtragem em que as amostras despadronizadas *outliers* foram retiradas e substituídas por valores obtidos por meio de interpolação polinomial. Para identificação de *outliers*, foi aplicado o teste de Grubbs (Grubbs (1969)). Esse algoritmo considera que os dados obedecem razoavelmente a uma distribuição normal, ou seja, aqueles pontos que parecem fugir do padrão estatístico de uma distribuição gaussiana são caracterizados como *outliers*. Retirados os *outliers*, a interpolação das séries foi realizada por meio de uma função que utiliza um polinômio de 3ª ordem.

O procedimento é descrito em um relatório interno (Aguirre, Aguirre, Mendes, e Martinez (2006)). Por questões relacionadas a confidencialidade, os dados são apresentados com escala normalizada, omitindo-se os valores reais em (*wh*) do consumo de energia da companhia. A normalização dos dados foi feita tal que os valores da série se situassem dentro do range entre 0 e 1. Para isso tem-se:

$$y_n(k) = \frac{y(k)}{y_{\max}}, \quad (3.8)$$

em que  $y_n(k)$  é o vetor de dados normalizado e  $y_{\max}$  é o maior valor da série  $y(k)$ . A série normalizada é apresentada na Figura 3.3. Esta série da corresponde a 1/8 da energia total distribuída pela CEMIG e contempla um nível de tensão de distribuição.

A Figura 3.3 mostra o comportamento da série de consumo total que será utilizada como base para obtenção dos modelos, já que o objetivo é realizar a previsão do consumo total da companhia. Alguns pontos podem ser percebidos claramente por uma simples inspeção visual:

- A série apresenta uma quebra na tendência no ano de 2001 causada pela crise de abastecimento;

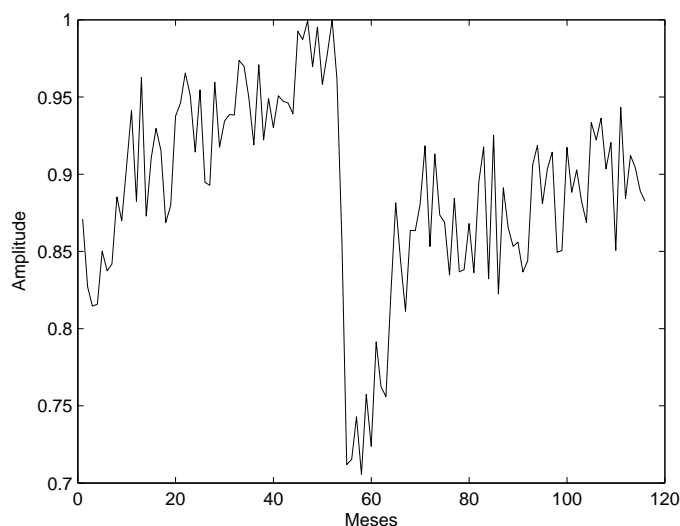


Figura 3.3: Série de consumo mensal total do Estado de Minas Gerais referente ao período que se estende de dezembro de 1996 a julho de 2006. Fonte: CEMIG.

- ii) A componente de tendência sofre alterações e apresenta comportamentos distintos no pré e pós-acionamento;
- iii) É possível verificar visualmente a existência de certa sazonalidade nos dados, possivelmente causada pela influência de fatores sócio-comerciais. É importante chamar a atenção para o fato de esta série ser menos irregular do que aquela mostrada na Figura 3.2.

A próxima seção detalha uma metodologia que é usada neste trabalho para filtrar as componentes das séries temporais: componente de tendência, componente cíclica e componente sazonal.

### 3.3 Componente Cíclica, sazonal e de tendência

A abordagem de análise de séries temporais baseados em tendências parte do pressuposto de que uma série temporal pode ser entendida como a composição de comportamentos de tendência, fatores cíclicos, variações sazonais, além de fatores aleatórios representados de maneira genérica como ruído branco.

*Tendência* em uma série temporal é a mudança gradual observada por meio da variação dos valores da série ao longo do tempo e que se mantém ao se remover os componentes de ciclos, sazonalidades e fatores aleatórios. Quando aplicado ao consumo de energia elétrica, esse conceito normalmente está relacionado ao comportamento do consumo ao longo do tempo devido à alteração correspondente na quantidade de consumidores ou às mudanças de perfil de clientes percebidas a longo-prazo.

*Ciclos e sazonalidades* são comportamentos estocásticos que acontecem de maneira recorrente ao longo de um período definido.

Morettin e Toloi (2006) conceituam que os comportamentos sazonais são flutuações ocasionadas na série temporal devido à influência de algum fator externo de sazonalidade.

Os componentes de ciclo apresentam um comportamento similar, no entanto, normalmente apresentam comprimento maior que os componentes sazonais e não apresentam duração uniforme.

No caso de séries temporais de energia elétrica é comum observar a existência de comportamentos recorrentes que caracterizam a sazonalidade devido a influência de fatores exógenos, em diferentes horizontes de observação. Tais comportamentos podem ser identificados em situações como as descritas a seguir:

- Variações anuais, normalmente associadas a influência de políticas governamentais;
- Variações regulares ao longo dos meses do ano, devido às mudanças de temperatura características de cada época que geram necessidades de comportamentos particulares quanto a utilização de equipamentos eletro-intensivos, sobretudo para a classe residencial;
- Comportamentos característicos em cada um dos dias da semana, principalmente devido a intensidade das atividades comerciais e industriais;
- Comportamentos característicos para cada hora do dia, de acordo com o perfil das classes consumidoras.

Com a remoção das componentes de tendência, ciclos e sazonalidades, têm-se os componentes *residuais* que podem, eventualmente, representar fatores aleatórios para o entendimento do comportamento do consumo.

Considerando que toda série temporal de energia elétrica possui um componente gerado pela influência de fatores aleatórios, mesmo se o comportamento exato dos demais componentes da série forem identificados com precisão, ainda existirá divergência entre os valores previstos pelo modelo e valores observados. Uma vez que a componente residual é isolada, a magnitude deste componente pode ser utilizada para se dimensionar o percentual de incerteza que se deve considerar ao se realizar a previsão de valores futuros da série.

No presente trabalho foi utilizada uma abordagem para identificação de tendências, ciclos e sazonalidades proposta por Mohr (2005) é chamada filtro TCS (*Trend Cycle Season*). O filtro TCS pode ser entendido como um método para a decomposição de séries temporais univariadas nas componentes de tendência, ciclo e sazonalidade, baseada no filtro HP de Hodrick e Prescott (Hodrick e Prescott (1997)). É fundamentada em modelos estocásticos explícitos tanto para a identificação de tendência quanto de ciclos e sazonalidades, permitindo a extração simultânea dos três componentes da série.

No filtro HP, a tendência estocástica é restrita a um modelo de segunda ordem. O filtro TCS é uma extensão do filtro HP na medida em que cria um modelo de extração de tendência de qualquer ordem e adiciona um modelo estocástico para a extração das componentes de ciclo e sazonalidades. Os processos cíclicos e sazonais definidos no algoritmo TCS assumem que os componentes seguem a definição de processos estacionários ARMA.

Em (Mohr (2005)) são discutidas as características dos filtros TCS e é apresentada a demonstração matemática do algoritmo. Uma outra técnica, encontrada em Coelho (2006), verifica se há algum padrão determinístico na série temporal além do comportamento quase-periódico.

### **3.4 Procedimentos para separação das componentes da série temporal**

A série de consumo do Estado de Minas Gerais apresenta apenas 116 amostras. Como pode ser visto na Figura 3.3 a série de consumo do Estado de Minas Gerais apresenta mudança de comportamento, principalmente a partir da amostra 55, onde o racionamento de energia elétrica no país ocorreu devido ao baixo nível de água dos reservatórios. Devido a essa

mudança de dinâmica da série, e também por apresentar a componente de tendência com crescimento bastante elevado, mesmo tendo ocorrido o racionamento, foi preciso modelar o consumo através da componente sazonal e “tendência + ciclos” separadas. As componentes de tendência, sazonalidade e ciclo dos dados da série temporal de consumo, Figura 3.3, foram separados usando a metodologia citada na se 3.3. Na Figura 3.4 pode observar as componentes de tendência, sazonalidade e ciclo, separadas pelo filtro TCS. A Figura 3.5 apresenta a componente de sazonalidade e a soma das componentes de tendência e ciclo.

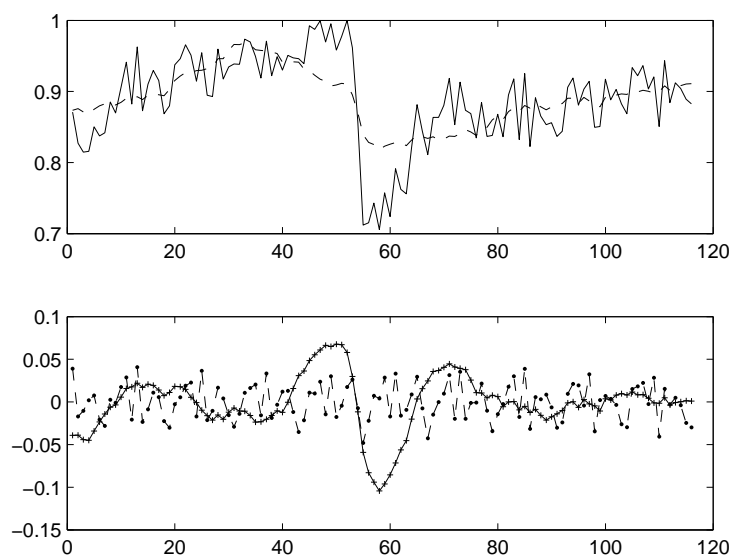


Figura 3.4: Componentes de tendência (- -), sazonalidade (-.-), ciclo (-+), e consumo total (-).

A componente sazonalidade foi mantida separada para servir como entrada para as representações não-lineares, pois esta componente apresenta ciclos anuais bem definidos. As componentes restantes foram somados (tendência + ciclo). Percebe-se na Figura 3.5 que quando a componente cíclica é somada a componente de tendência o resultado é uma componente que acompanha melhor a série temporal. A partir daqui, quando for citado no texto "componente tendência", deve-se entender que corresponde à soma das componentes de tendência e ciclo.

Para as previsões livre de 60 passos à frente, para a componente de

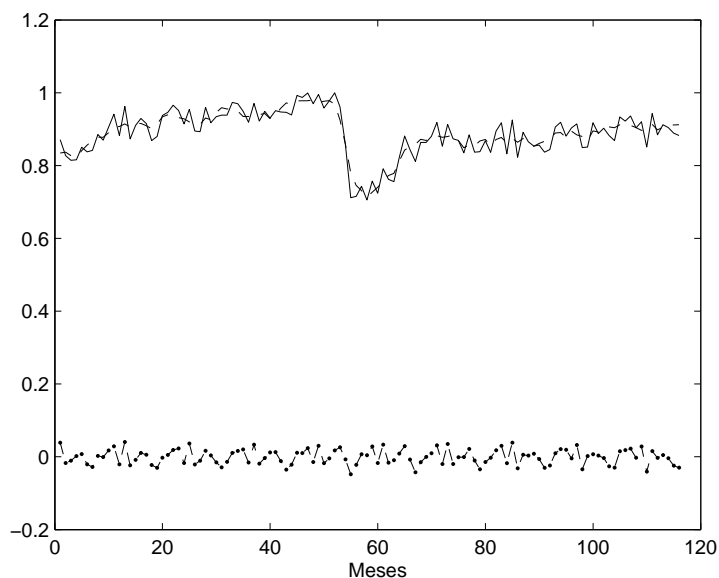


Figura 3.5: Componentes de “tendência + ciclo” (- -), sazonalidade (-.-), e consumo total (-).

tendência foi ajustada uma reta com o mesmo coeficiente angular do trecho referente às observações 80 a 116, pois é a partir dessas amostras que tendência do consumo se mostra crescente, desconsiderando assim a tendência pré-acionamento. Assim considerando-se que a tendência das observações 80 a 116 é uma reta a previsão de 60 passos à frente é obtida.

Para a previsão livre de 60 passos à frente da série de New England não foi necessário aproximar a tendência por uma reta, pois, esta série de consumo possui bastante amostras, uma vez que as previsões não extrapolam o tamanho da série.

Para obtenção de modelos NARIMA também foi necessário retirar a componente da tendência da série de New England. No entanto, os resultados da modelagem usando as representações Neural e Neuro-Fuzzy mostraram-se satisfatórias sem a extração da componente tendência.

Para a série de consumo do Estado de Minas Gerais o procedimento de separar a componente de tendência da série foi aplicado para obtenção de modelos NARIMA, Neurais e Neuro-Fuzzy.



## 3.5 Procedimentos para os modelos ARIMA e NARIMA

Os procedimentos utilizados para obtenção de modelos ARIMA são embasados em (Box e Jenkins (1970); Aguirre (2007)).

Em Box e Jenkins (1970) é detalhado toda metodologia dos modelos ARIMA. Os autores mostram que a estratégia para a escolha do modelo é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados. A Figura 3.6 mostra esse ciclo.

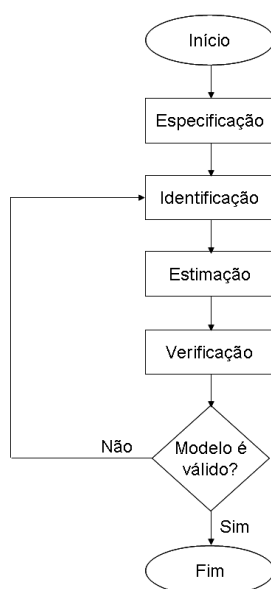


Figura 3.6: Esquema das etapas do modelo ARIMA.

Os estágios do ciclo iterativo são:

- Na *especificação*, uma classe de modelos é considerada para a análise.
- Na *identificação*, um modelo é escolhido com base na análise de autocorrelações, autocorrelações parciais e outros critérios. Nessa etapa, é identificada a necessidade de diferenciar os dados para tornar a série estacionária e calculada a ordem do processo autoregressivo ( $p$ ) e do processo de médias móveis ( $q$ ). Tradicionalmente, as principais ferramentas nessa fase são as funções de autocorrelação e autocorrelações parciais.

- A seguir vem a fase de *estimação*, na qual os parâmetros do modelo identificado são estimados, ou seja, os coeficientes AR e MA.
- Finalmente, na fase *verificação* ou *diagnóstico*, o modelo encontrado é determinado adequado ou não usando-se, por exemplo, a análise de resíduos.

No entanto, neste trabalho, é proposta uma estrutura alternativa para os modelos ARIMA. Nas fase de identificação as ordens do processo AR e do MA são escolhidas usando-se a *taxa de redução do erro* (Aguirre (2007)) e na fase da estimação dos parâmetros é realizada usando-se os *mínimos quadrados* (Aguirre (2007)). Estas técnicas podem, em princípio, retirar os termos desnecessários do modelo que proporcionam instabilidade numérica e ocasionam comportamentos dinâmicos e espúrios ao sistema. A Figura 3.7 (a) mostra a representação esquemática para o modelo ARIMA, onde os regressores são os máximos atrasos da série temporal, no caso doze (12).

Inicialmente a obtenção de modelos NARIMA mostrou-se difícil, pois, usando a metodologia empregada pelos modelos ARIMA a obtenção de modelos não-lineares não mostrou-se eficiente, resultando sempre em modelos instáveis.

Uma possível solução, para se obter modelos NARIMA<sup>3</sup>, foi retirar a componente de tendência das séries temporais de New England e do Estado de Minas Gerais e após isso obter a primeira diferença. Esse procedimento apresentou modelos estáveis. Uma outra alternativa, que não é mostrada neste trabalho, foi retirar a componente de tendência e modelar as séries temporais sem obter a primeira diferença, resultando num modelo NARMA (não-linear autoregressivo de média móvel). Este procedimento também resultou em modelos estáveis.

A seção 3.4 detalha como a componente de tendência foi filtrada da série original de consumo de Minas Gerias. A Figura 3.7 (b) exhibe a topologia para o modelo não-linear (NARIMA).

Na topologia usada para o modelo NARIMA (Figura 3.7) são passados como entradas os doze primeiros regressores da componente sazonal que

---

<sup>3</sup>Por definição os modelos ARIMA e NARIMA a tendência é removida pela primeira diferença. No entanto, simulações dos modelos NARIMA com a tendência mostraram instabilidade, e a saída encontrada foi retirar a tendência e após isso realizar a primeira diferença

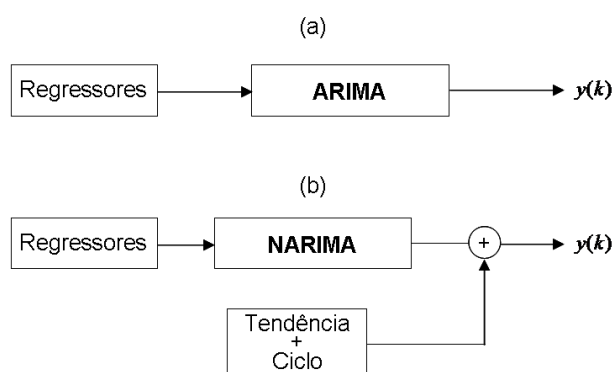


Figura 3.7: Representação esquemática de modelos (a) ARIMA e (b) NARIMA.

são modelados pela representação NARIMA, e na saída do modelo são somadas as componentes cíclica e a tendência, resultando no modelo. O aumento do número de regressores da série temporal não acarretou melhora no desempenho do modelo.

### 3.6 Procedimentos para a RNF

Os procedimentos utilizados para a rede Neuro-Fuzzy são embasados de acordo com a seção 2.3.2, onde a fundamentação teórica foi detalhada. Os procedimentos são apresentados para os dois casos de estudo, a série de consumo de New England e a da CEMIG.

Os dados apresentados nas Figuras 3.2 e 3.3 foram normalizados. Para a normalização os dados foram divididos pelo máximo valor de consumo de acordo com Equação (3.8). A normalização dos dados, tanto para a rede Neuro-Fuzzy quanto para a Rede Neural, proporciona a obtenção de resultados melhores, (Braga et al. (2000); Haykin (2001)). Os resultados das previsões de série de New England são apresentados desnormalizados, isto é, multiplicados pelos máximos das série temporal.

A RNF, a princípio, não é uma estrutura dinâmica. Para que a RNF incorporasse a dinâmica das séries temporais foi preciso definir os regressores (atrasos) para a entrada da rede. Como os dados de consumo são mensais, sabe-se que, em condições normais, o consumo de um certo mês do ano é bastante parecido com o consumo do mesmo mês do ano interior

devido as variações regulares ao longo dos meses do ano, e as mudanças características de cada época, caracterizando a sazonalidade de período 12. Assim, o décimo segundo regressor contém bastante informação dinâmica das séries. Outros regressores também foram usados como entrada para a rede. No decorrer da descrição dos procedimentos os outros regressores para rede serão definidos.

Como mencionado anteriormente, neste trabalho são apresentados dois estudos de caso. O primeiro leva em conta a série temporal de consumo de energia elétrica de New England (USA), e o segundo a série temporal de consumo de Minas Gerais, fornecida pela Cemig. Para os dois casos inicialmente é configurada uma RNF que realiza a previsão de 12 passos à frente para avaliar a eficiência da técnica usada. Após isso, a topologia de 12 passos é reajustada para realizar a previsão livre de 60 passos à frente. Como cada amostra das séries representa o consumo total de um mês, a previsão livre de 12 passos à frente é equivalente a previsão de 1 ano à frente, e a de 60 passos à frente é equivalente a 5 anos à frente, ou 60 meses a frente. Neste trabalho, usa-se a metodologia de realimentar a topologia do modelo para obter a previsão de  $k$  passos à frente, sendo  $k$  o horizonte de previsão.

O desempenho das previsões das redes é avaliado por meio da análise dos erros MPE, MAPE, RMSE e  $U$  de Theil, Equações 3.4, 3.5, 3.6, 3.7, respectivamente. Várias redes foram treinadas e as que apresentaram os menores índices citados foram selecionadas.

### 3.6.1 Modelo de consumo de New England

A série de consumo de New England tem 228 observações. Destas 228 amostras, as primeiras 168 amostras foram separadas (janeiro de 1980 – dezembro de 1993). A previsão livre de 12 passos à frente foi efetuada, usando-se para o treinamento 156 observações das 168. As 12 observações que restaram foram usadas para a validação. Com isso, a eficiência da rede foi avaliada. A partir daí, a mesma rede é usada para treinar as 168 primeiras observações, e a previsão livre de 60 passos à frente é avaliada. As 60 observações finais (169 – 228) foram usadas como dados de validação. Os índices MPE, MAPE, RMSE e  $U$  de Theil são avaliados. A Figura 3.8 mostra a topologia do modelo Neuro-Fuzzy para o consumo de energia de New England.

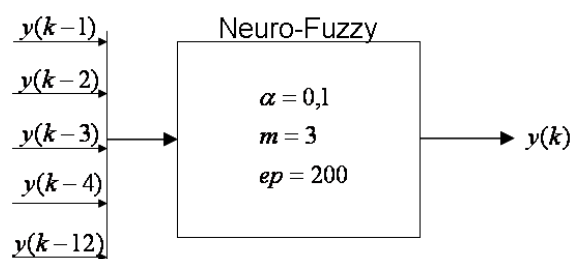


Figura 3.8: Representação esquemática de modelo Neuro-Fuzzy do consumo de energia de New England.

A saída da rede é desnormalizada multiplicando o vetor de saída pelo valor máximo da série temporal.

A escolha dos atrasos (lags) é um fator que depende muito das características do processo em estudo. Para as séries de consumo mensais de New England e do Estado de Minas Gerais foram usados os mesmos atrasos de cada série para compor as entradas dos modelos neurais. Foram feitas simulações para os atrasos [1-4,12]. Também tentou-se simular com outras entradas, como exemplo [1-12], [1-4,12,23-24], dentre outras, entretanto os melhores resultados foram obtidos com os atrasos [1-4,12]. Estes mesmos testes foram realizados usando-se a RNF. Aguirre, Rodrigues, Lima, e Martinez (2008) apresentam uma discussão que pode servir para justificar os atrasos escolhidos para a camada de entrada do modelo Neuro-Fuzzy e Neural.

### 3.6.2 Modelo de consumo da CEMIG

Para obtenção do modelo Neuro-Fuzzy do consumo de Minas Gerais usou-se a metodologia para extração das componentes da série, citada na seção 3.4.

A RNF é treinada usando-se as 104 primeiras observações (de dezembro de 96 a julho de 2005) da componente sazonalidade. E as 12 últimas amostras (de agosto de 2005 a julho de 2006) são usadas para a validação livre de 12 passos (1 ano) à frente das redes. Os regressores usados na entrada da RNF são mostrados na Figura 3.9.

Após a avaliação do comportamento da rede para a previsão de 12 passos à frente, treinou-se a mesma rede considerando todas as 116 ob-

servações da série, e a previsão de 60 passos à frente (de agosto de 2006 a julho de 2011) é realizada. Neste caso, no entanto, não existem amostras para validação.

Para a previsão livre de 60 meses à frente, como não existem amostras para comparação, também não existe componente de tendência para os 60 passos futuros. Sendo assim, os valores futuros da tendência foram aproximados por uma reta, que por sua vez, incorpora o coeficiente angular da tendência após o racionamento de energia, ou seja, tende a acompanhar o mesmo crescimento do consumo de energia dos dados passados (a partir da observação 80).

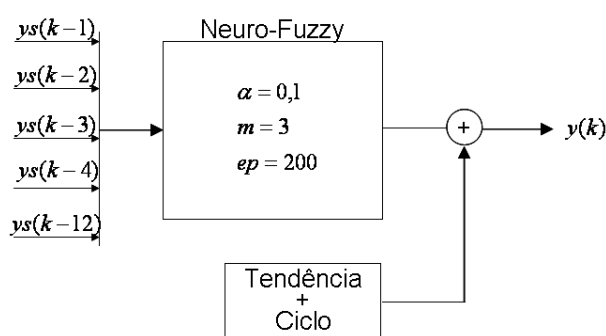


Figura 3.9: Representação esquemática de modelo Neuro-Fuzzy do consumo de energia do Estado de Minas Gerais. ( $y_s$  representa os dados da componente sazonal da série de consumo da Cemig.)

Como pode ser visto na Figura 3.9 na fase de treinamento da rede Neuro-Fuzzy o critério de parada é definido pelo número de épocas. Tipicamente uma época do treinamento é definida como uma única apresentação de todos os vetores da entrada a rede. Então a rede é atualizada de acordo com os resultados de todas estas apresentações. A taxa de aprendizagem especifica a velocidade com que os pesos são atualizados, e é definida como sendo 0,1. E outro parâmetro que deve ser definido para o treinamento da rede é o número de regras do tipo (if-then) para cada entrada associada, que é igual a 3.

Os parâmetros de treinamento foram os mesmos para os estudos de caso apresentados neste trabalho. Essas configurações apresentaram os melhores resultados.

A saída da RNF é somada à componente de tendência, resultando no consumo de energia normalizado do Estado de Minas Gerais (Figura 3.9).

## 3.7 Procedimentos para a RNA

Para os procedimentos da RNA devem ser levados em conta a escolha das entradas do modelo, escolha do número de neurônios na camada escondida, escolha das saídas, porcentagem de dados para treinamento e validação. Algumas considerações sobre a escolha da topologia da rede são feitas a seguir. Vale citar uma frase de Zhang et al. (1998): "*O projeto de uma RNA é mais uma arte que uma ciência*".

Nesta seção são apresentados os procedimentos para obtenção do modelo Neural. Mauro (2007) apresenta detalhes desta metodologia de modelos Neurais que realizam a previsão a longo-prazo de consumo de energia elétrica. No trabalho o autor aborda a previsão de 60 meses à frente do consumo de energia da CEMIG.

A escolha do número de neurônios na camada de entrada, no problema de modelagem de séries temporais, trata-se basicamente da escolha dos atrasos utilizados como entradas no modelo. Na seção 3.6.1 foi descrito como a escolha dos regressores para a entrada da Rede Neural e da Neuro-Fuzzy foram escolhidos.

O neurônio da camada de saída corresponde, no problema de modelagem de séries temporais, ao horizonte de previsão. Existem duas formas de se construir um modelo de previsão. A primeira é a previsão de 1 passo à frente, em que apenas um neurônio é utilizado na camada de saída. A segunda forma é utilizada quando se deseja ter como saída a previsão em um horizonte maior do que a amostragem da série. Por exemplo, deseja-se obter a previsão para 6 meses à frente, para uma série de amostragem mensal. O que se pode fazer é utilizar as previsões de um passo à frente e realimentá-las no modelo para obter as previsões no horizonte maior, ou utilizar mais neurônios na camada de saída obtendo a saída diretamente sem a necessidade de realimentar previsões. Neste trabalho usa-se a metodologia de realimentar a topologia do modelo para obter a previsão de  $k$  passos à frente, sendo  $k$  o horizonte de previsão, conhecida como simulação livre. As funções de ativação tanto da camada escondida, como da camada de saída foram a função tangente hiperbólica (Equação 2.16).

O neurônio da camada escondida normalmente é obtido por tentativas. Na literatura, até onde foi pesquisado, não se encontrou uma metodologia para determinar o número de neurônios da camada escondida. Neste trabalho, foram realizadas várias simulações com o número de neurônios

na camada escondida variando de 2 até 20. Para cada neurônio foram realizadas 50 simulações, inicializando a rede em condições iniciais diferentes para tentar evitar que o algoritmo de treinamento alcançasse um mínimo local. A simulação correspondente ao número de neurônios da camada que apresentou melhor desempenho, em função dos índices de desempenho, foi escolhida.

Uma questão importante a ser definida é a composição dos dados para treinamento e validação, isto é, quais os dados que serão usados para treinar os modelos, e quais serão utilizados para validá-los. A literatura sugere que pelo menos 20% da base de dados seja separada para validar os modelos obtidos na fase de treinamento (Braga et al. (2000)). A seguir são detalhados as topologias obtidas para a série de New England e da Cemig.

### 3.7.1 Modelo de consumo de New England

Os mesmos procedimentos usados na seção 3.6.1 para separar os dados de treinamento e validação são utilizados para a RNA. Das 228 amostras da série de consumo de energia elétrica de New England, separou-se 168 amostras para a previsão de 12 meses à frente. Uma topologia de 8 neurônios na camada escondida foi determinada, (Figura 3.10).

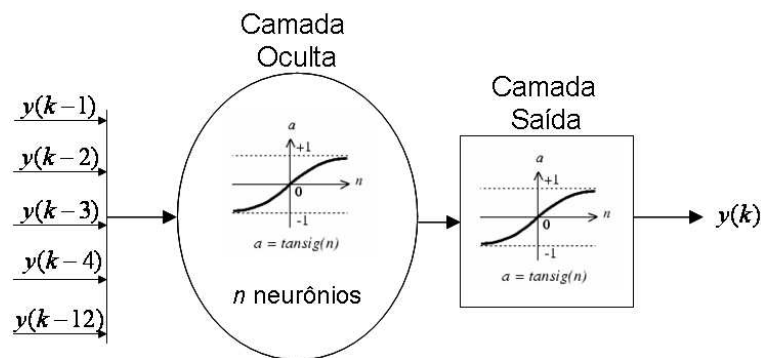


Figura 3.10: Representação esquemática de modelo Neural do consumo de energia de New England.

A metodologia utilizada no treinamento da rede MLP é

- dados normalizados para o intervalo  $[0,1]$ ;



- as redes foram treinadas utilizando o algoritmo de *Levenberg-Marquardt* (Braga et al. (2000)), com controle de generalização do tipo parada antecipada;
- os regressores para o modelo e o número de neurônios ( $n$ ) na camada de escondida são obtidos por tentativa e erro. São realizadas simulações para vários atrasos da série temporal. Os atrasos mais significativos são [1-4,12]. E também é realizada uma busca para os valores de  $n$  variando de [2:2:20], sendo escolhido 8 neurônios na camada intermediária. Tanto a escolha dos regressores, quanto a determinação de  $n$  foram feitas em função dos índices de desempenho;
- dentre os modelos gerados é escolhido aquele com menor erro de validação.

### 3.7.2 Modelo de consumo da Cemig

A metodologia empregada para estimar o modelo Neuro-Fuzzy em 3.6.2 é a mesma aqui empregada para encontrar o modelo Neural do consumo de energia do Estado de Minas Gerais.

Os detalhes de toda escolha e procedimentos da estrutura da Rede Neural podem ser vistos em (Mauro (2007)). O autor tratou, com sucesso, da modelagem da série temporal de consumo da CEMIG utilizando Redes Neurais.

Inicialmente os dados foram separados em componentes de tendência, sazonalidade e ciclos, como mostrado na Figura 3.5. A Figura 3.11 mostra a topologia usada para o modelo Neural do consumo da Cemig.

Estimada a rede para a previsão de 12 passos à frente, esta mesma rede é atualizada para realizar as previsões de 60 passos à frente. A topologia da rede apresenta 8 neurônios na camada escondida. A saída da RNA é somada à componente de tendência e ciclo, resultando no consumo de energia normalizado do Estado de Minas Gerais.

## 3.8 Comentários finais

Neste capítulo foram descritos todos procedimentos necessários que conduzem na obtenção de modelos de previsão para o consumo de energia

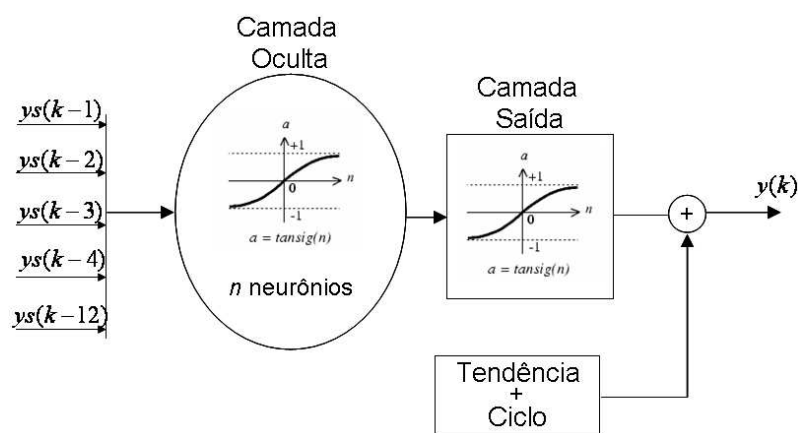


Figura 3.11: Representação esquemática de modelo Neural do consumo de energia do Estado de Minas Gerais. ( $y_s$  representa os dados da componente sazonal da série de consumo da Cemig.)

da cidade de New England e para o Estado de Minas Gerais. Os próximos capítulos apresentarão os resultados da aplicação destes procedimentos mostrados aqui.

A Tabela 3.1 apresenta uma síntese da forma que os dados são apresentados às representações em cada estudo de caso. Por exemplo, no caso de New England, o modelo NARIMA é obtido através da componente sazonalidade e à sua saída são somadas a tendência e ciclo.

Tabela 3.1: Comparação dos modelos através do índice de desempenho: MPE, MAPE e RMSE para a série de consumo em Minas Gerais.

Estudo de caso/ Representação	New England	CEMIG
ARIMA	Série total	Série total
NARIMA	Sazonalidade	Sazonalidade
RNF	Série total	Sazonalidade
RNA	Série total	Sazonalidade

O Capítulo 4, mostrará os resultados referentes a cidade de New England, e o Capítulo 5 os resultados do Estado de Minas Gerais.

# Estudo de caso 1: Série temporal de New England

---

*"Você tem que se apaixonar com seus dados, mas nem sempre com seus modelos."*

G.M. Jenkins

Serão utilizados neste capítulo os procedimentos apresentados no capítulo anterior. Os resultados e algumas discussões da obtenção dos modelos ARIMA, NARIMA, Neuro-Fuzzy e Rede Neural para a previsão do consumo de energia da cidade de New England são apresentados. Todos os modelos deste estudo de caso são obtidos a partir da série temporal sem sua extração da componente de tendência. O único modelo em que houve a extração desta componente foi o modelo NARIMA, como se verá abaixo.

As representações RNF e RNA, para este caso, não foi necessário extrair a componente de tendência, pois, estes modelos conseguiram bom funcionamento geral em face à não estacionariedade.

## 4.1 Modelo ARIMA

O modelo ARIMA estimado a partir da série de consumo de New England, da observação 1 até a 168<sup>1</sup>, sendo que os 12 meses<sup>2</sup> seguintes foram utilizados para escolher o seguinte modelo:

$$\begin{aligned} y(k) = & +0,804 y(k-12) + 0,888 y(k-1) \\ & -0,0891 y(k-4) - 0,804 y(k-13) \quad . \\ & +0,112 y(k-2) + 0,0891 y(k-5) \end{aligned} \quad (4.1)$$

---

<sup>1</sup>Isso corresponde a 14 anos, ou seja, de janeiro de 1980 a dezembro de 1993.

<sup>2</sup>De janeiro a novembro de 1994.

O modelo (4.1) foi simulado livremente 12 passos à frente. Veja Figura 4.1. O erro percentual médio (MPE) para a simulação livre de 12 passos (1 ano) à frente do modelo (4.1) foi de 0,996%.

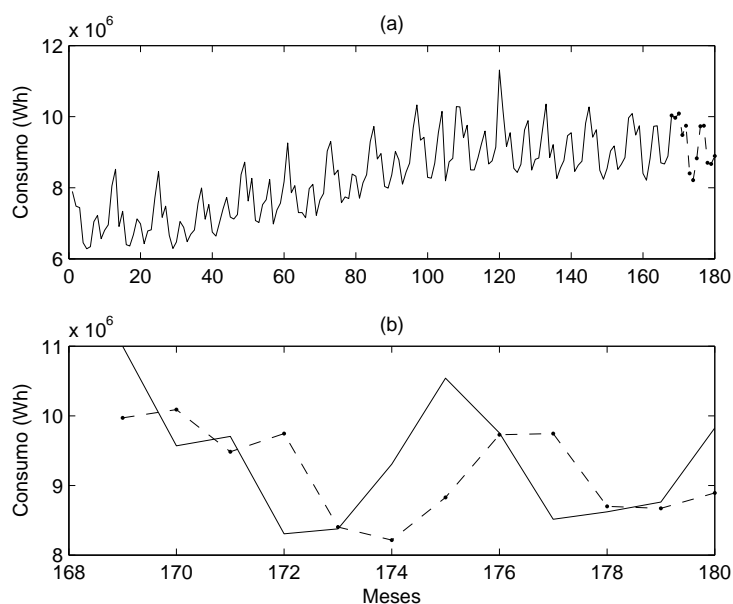


Figura 4.1: Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (—) consumo real, (-.-) consumo previsto. Erros apresentados, MPE: 0,996%; MAPE: 7,406%; RMSE: 1,096;  $U_{Theil}$ : 0,892.

O modelo (4.1) agora é simulado para 60 meses (5 anos)<sup>3</sup> à frente.

A Figura 4.2 mostra a simulação livre de 60 passos à frente do modelo (4.1). Nessa simulação observa-se que o modelo apresenta oscilações apreciáveis que acompanham as oscilações reais do consumo. Nesta mesma figura percebe-se também que o modelo generaliza de forma convincente a tendência da série. O erro percentual médio para a simulação livre de 60 passos foi de 0,129%. Entretanto, percebe-se que a medida que o horizonte de previsão aumenta, a resposta do modelo tende a amortecer para um valor de consumo médio.

<sup>3</sup>De janeiro de 1994 a dezembro de 1998.

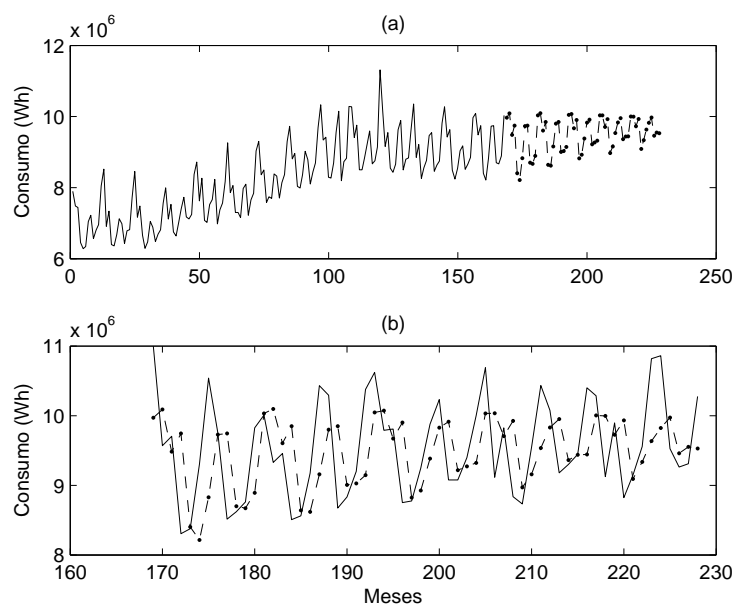


Figura 4.2: Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (—) consumo real, (-.-) consumo previsto. Erros apresentados, MPE: 0,129%; MAPE: 5,920%; RMSE: 1,047;  $U_{Theil}$ : 0,995.

## 4.2 Modelo NARIMA

O modelo NARIMA foi estimado a partir da série de consumo de New England sem a componente de tendência (Figura 3.7), da observação 1 até a 168, sendo que os 12 meses seguintes foram utilizados para escolher o seguinte modelo:

$$\begin{aligned}
 y(k) = & +0,986 y(k-12) + 0,0205 y(k-7) \\
 & + y(k-1) - 0,986 y(k-13) \\
 & +0,987 y(k-11)y(k-12) \quad . \quad (4.2) \\
 & -0,987 y(k-12)y(k-13) \\
 & -0,0205 y(k-8)
 \end{aligned}$$

A Figura 4.3 mostra a previsão livre de 12 passos à frente do modelo (4.2) já adicionada a componente de tendência.

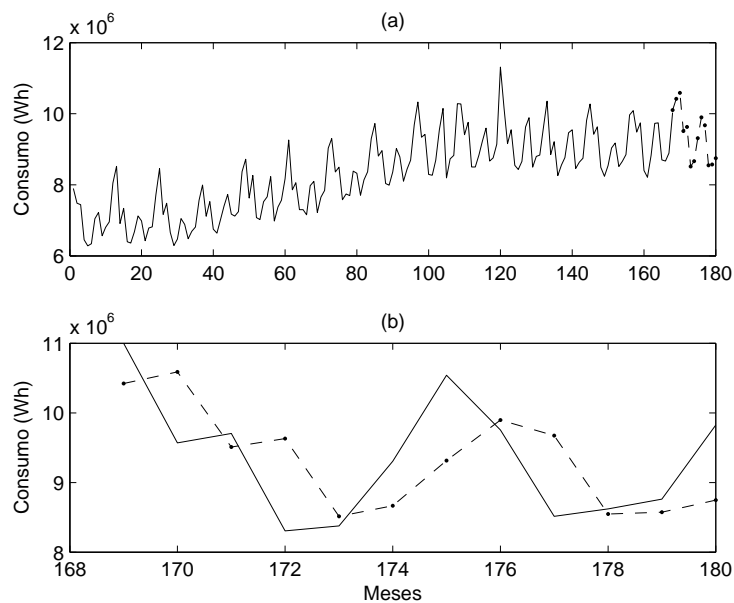


Figura 4.3: Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (– –) consumo previsto. Erros apresentados, MPE: -0,298%; MAPE: 6,924%; RMSE: 0,961;  $U_{Theil}$ : 0,782.

A simulação livre de 12 passos à frente, de acordo com a Figura 4.1,

apresenta erro percentual médio de -0,298%.

Da mesma forma que foi apresentado para o modelo ARIMA na seção anterior, o modelo (4.2) agora é simulado para 60 meses (5 anos) à frente.

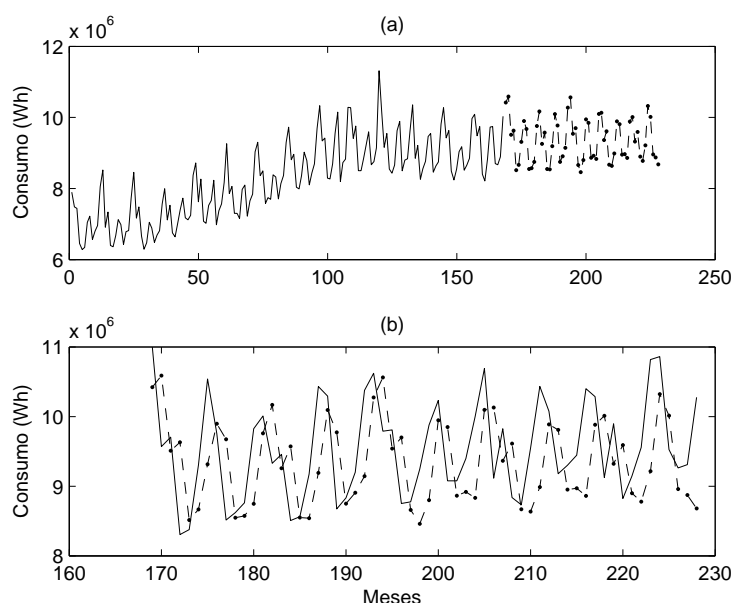


Figura 4.4: Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (—) consumo real, (-.-) consumo previsto. Erros apresentados, MPE: 1,644%; MAPE: 6,810%; RMSE: 1,134;  $U_{Theil}$ : 1,076.

A previsão livre de 60 meses à frente, Figura 4.4, apresenta índices de desempenho relativamente pequenos, e também observa-se que o modelo consegue absorver a dinâmica da série de consumo representando adequadamente as oscilações e mudanças de tendência que acontecem a partir da observação 168 (janeiro de 1994) até a 228 (dezembro de 1998).

### 4.3 Rede Neuro-Fuzzy

Como foi apresentado na Seção 3.6 as entradas para o modelo Neuro-Fuzzy foram os regressores [1-4,12] do consumo de energia elétrica (Figura 3.8). A Figura 4.5 mostra a simulação livre de 12 passos a frente do modelo.

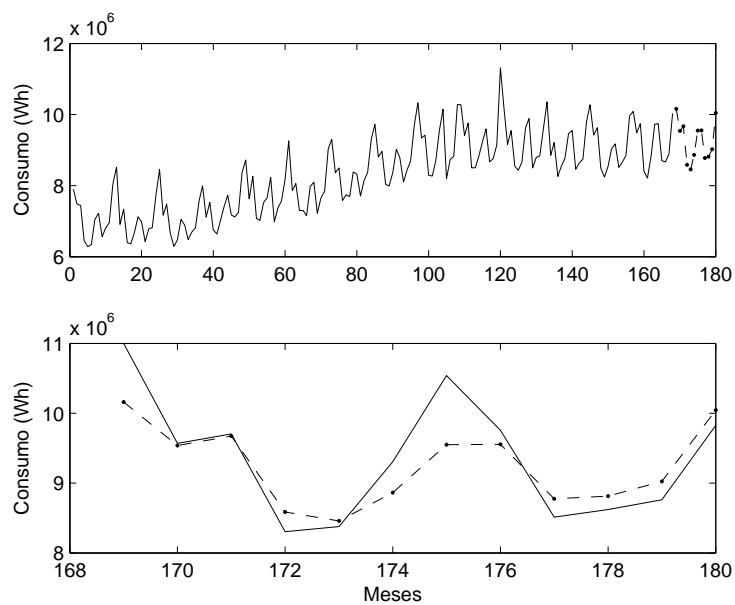


Figura 4.5: Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–.–) consumo previsto. Erros apresentados, MPE: 0,801%; MAPE: 3,279%; RMSE: 0,518;  $U_{Theil}$ : 0,873.



A previsão de 12 passos do consumo de energia da cidade de New England usando a rede Neuro-Fuzzy apresentou erro MPE de 0,801%, MAPE: 3,279%, RMSE: 0,518.

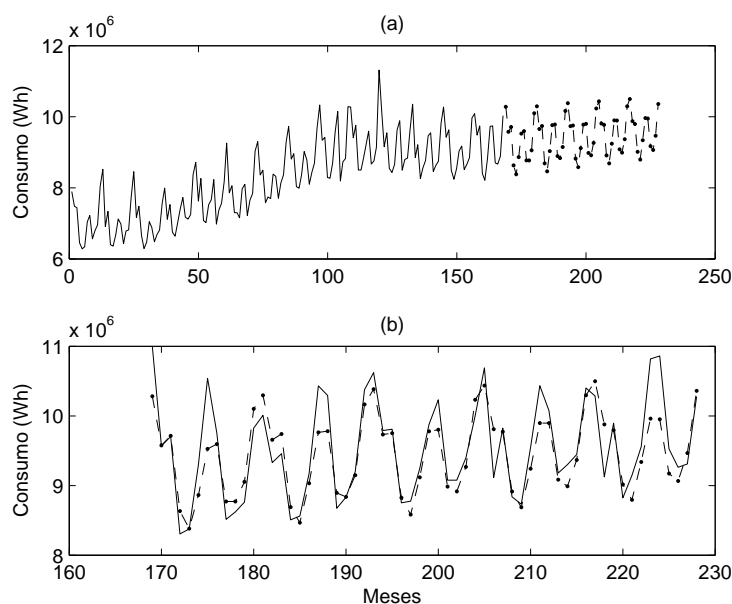


Figura 4.6: Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–) consumo previsto. Erros apresentados, MPE: 0,813%; MAPE: 2,672%; RMSE: 0,506;  $U_{Theil}$ : 0,567.

A Figura 4.6 (a) apresenta as observações usadas para o treinamento da rede e a saída de 60 passos à frente da rede. E na Figura 4.6 (b) é apresentado o detalhe das observações usadas para a validação, com as respectivas amostras previstas pelo modelo Neuro-Fuzzy.

A previsão livre de 60 passos à frente apresentou índices MPE: 0,813, MAPE: 2,672, RMSE: 0,506. A Figura 4.6 mostra que a RNF apresenta flutuações apreciáveis. Isto é, na previsão de 60 de passos o modelo Neuro-Fuzzy consegue modelar a sazonalidade anual contida nos dados de consumo e também consegue acompanhar a tendência positiva dos dados de validação.

## 4.4 Rede Neural Artificial

Nesta seção é usada uma Rede Neural Artificial para realizar a previsão do consumo de energia de New England. A descrição da série temporal de New England é detalhada em (3.2.1).

A série de consumo possui 228 observações mensais. De acordo com a Seção 3.6.1 foram separadas as primeiras 156 observações para o treinamento do modelo Neural e as 12 próximas amostras para a validação livre de 12 passos à frente. A Figura 4.7 mostra o desempenho visual do modelo Neural para a validação de 12 passos à frente.

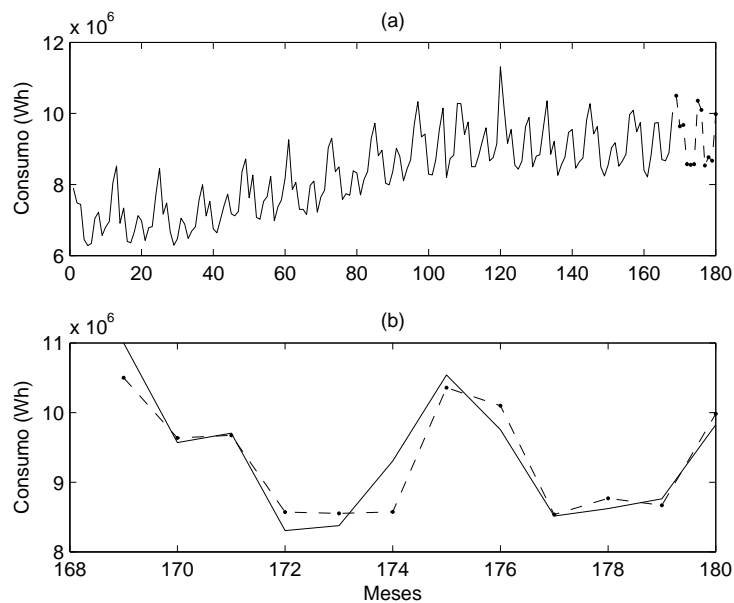


Figura 4.7: Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (—) consumo real, (-.-) consumo previsto. Erro apresentados, MPE: 0,204%; MAPE: 2,383%; RMSE: 0,364;  $U_{Theil}$ : 1,054.

A validação livre de 12 passos à frente, apresentada na Figura 4.7, apresenta índices de desempenho, MPE = 0,204, MAPE = 2,383 e RMSE = 0,364, que indicam bom desempenho do modelo Neural. A topologia do modelo Neural apresenta 8 neurônios na camada escondida e na camada de entrada os regressores usados são [1-4,12]. Ambas camadas usam funções

de ativação tangente hiperbólica.

A Figura 4.8 mostra o modelo Neural avaliado para a previsão de 60 meses à frente.

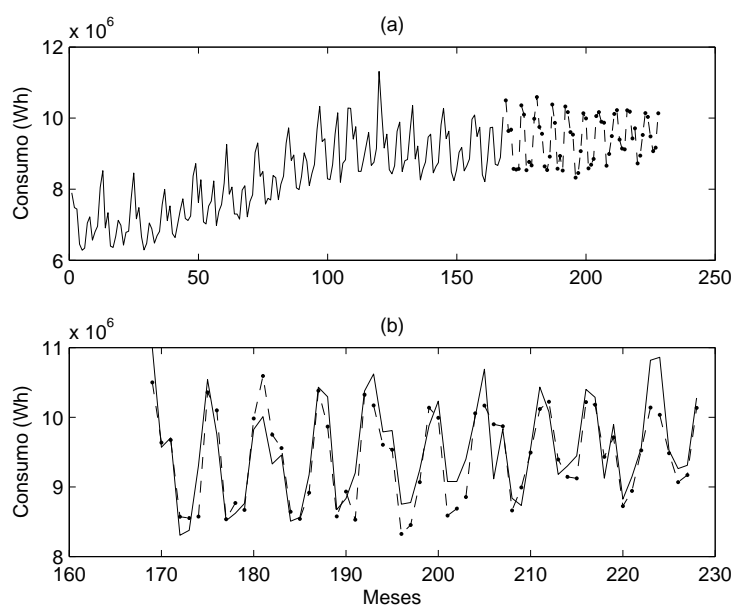


Figura 4.8: Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–) consumo previsto. Erro apresentados, MPE: 1,404%; MAPE: 3,142%; RMSE: 0,576;  $U_{Theil}$ : 0,998.

Visualmente percebe-se que o modelo Neural, Figura 4.8 (b), consegue generalizar a dinâmica da série de consumo de New England. Tanto a tendência da série, quanto as oscilações sazonais são absorvidas pelo modelo. Os índices de desempenho revelam que o modelo apresenta bom desempenho para a previsão livre de 60 meses à frente.

## 4.5 Comparação das Representações

Todas as representações usadas neste capítulo, a saber: ARIMA, NARIMA, RNF e RNA, mostraram grande robustez na previsão de longo-prazo do consumo de energia de New England. A Tabela 4.1 abaixo

detalha os erros MPE, MAPE e RMSE das representações para a previsão livre de 12 e 60 passos à frente.

Tabela 4.1: Comparação dos modelos através do índice de desempenho: MPE(%), MAPE(%) e RMSE para a série de consumo em New England.

Previsão	MPE		MAPE		RMSE		$U_{Theil}$	
	12	60	12	60	12	60	12	60
ARIMA	0,996	<b>0,129</b>	7,406	5,920	1,096	1,047	0,892	0,995
NARIMA	-0,298	1,644	6,924	6,810	0,961	1,133	<b>0,782</b>	1,076
RNF	0,801	0,813	3,279	<b>2,672</b>	0,518	<b>0,506</b>	0,873	<b>0,567</b>
RNA	<b>0,204</b>	1,404	<b>2,383</b>	3,142	<b>0,364</b>	0,576	1,054	0,998

Pela análise dos índices apresentados na Tabela 4.1 e nas Figuras deste capítulo, percebe-se que todas as representações usadas demonstram convergência e boa precisão sobre os dados de consumo de energia na cidade de New England. O modelo ARIMA apresentou 0,129% de erro percentual médio, o menor para a o horizonte de 60 meses à frente. O modelo Neuro-Fuzzy apresentou erro MAPE de 2,672% e RMSE de 0,506 para previsões de 60 meses à frente, os menores entre os modelos. De acordo com esses resultados o modelo RNF apresenta os melhores resultados para a previsão livre de 60 meses à frente. Em relação à previsão de 12 passos à frente a RNA apresenta índices menores em relação às outras apresentações.

É relevante lembrar que de todas as representações mostradas na Tabela 4.1 a única que teve a retirada da tendência para a modelagem foi a NARIMA, pois, este mostrou-se instável quando simulado considerando a série temporal com todas suas componentes.

## Estudo de caso 2: Série temporal da Cemig

---

Neste capítulo são apresentados os resultados da previsão do consumo de energia elétrico do Estado de Minas Gerais. As simulações com respectivos índices de desempenho dos modelos ARIMA, NARIMA, RNF e RNA são mostrados.

Como foi visto no Capítulo 3 para os modelos NARIMA, RNF e RNA foi necessário extrair a componente de tendência e somá-la na saída do modelo.

### 5.1 Modelo ARIMA

A Equação 5.1 abaixo representa o modelo ARIMA estimado a partir da série temporal de consumo do Estado de Minas Gerais. A amostra 1 a amostra 104<sup>1</sup> foram usadas para identificação do modelo, e as 12 amostras (meses)<sup>2</sup> seguintes foram utilizados para validar o seguinte modelo:

$$\begin{aligned} y(k) = & +0,646 y(k - 12) + 0,71 y(k - 1) \\ & -0,0341 y(k - 11) - 0,225 y(k - 10) \quad . \quad (5.1) \\ & -0,387 y(k - 13) + 0,29 y(k - 2) \end{aligned}$$

O modelo 5.1 foi simulado livremente 12 passos à frente e a Figura 5.1 mostra a previsão livre de 12 passos à frente.

A Figura 5.2 mostra a simulação livre de 60 passos à frente<sup>3</sup> do modelo (5.1). Nesta simulação observa-se que o modelo tende acompanhar as oscilações e a tendência reais do consumo. Percebe-se também, que a

---

<sup>1</sup>Isso corresponde a 9 anos e oito meses, ou seja, de dezembro de 1996 a julho de 2005.

<sup>2</sup>De agosto de 2005 a julho de 2006.

<sup>3</sup>De agosto de 2006 a julho de 2011.

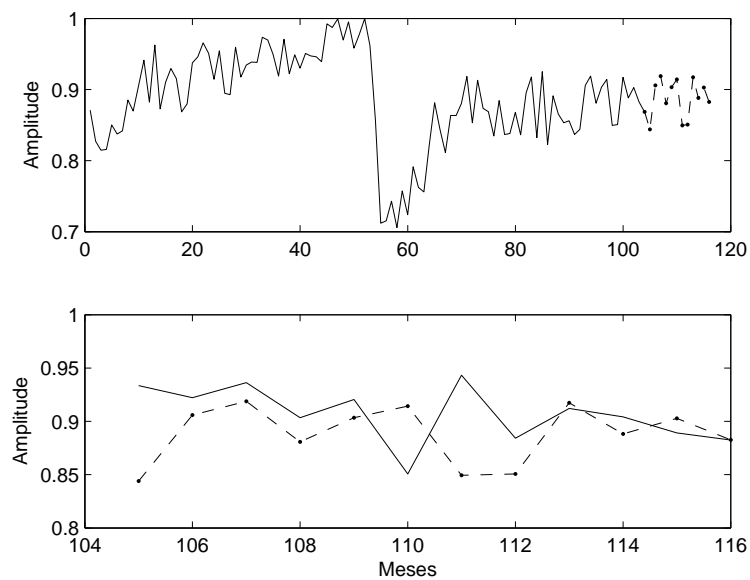


Figura 5.1: Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo ARIMA. (–) consumo real, (–.–) consumo previsto. Erros apresentados, MPE: 1,956%; MAPE: 3,557%; RMSE: 1,732;  $U_{Theil}$ : 1,425.

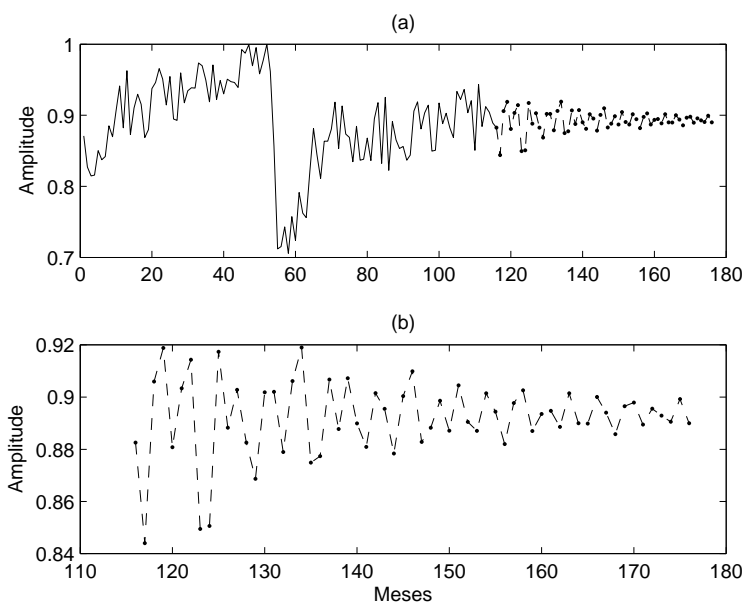


Figura 5.2: Simulação do modelo ARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe do modelo ARIMA (-.-) consumo previsto.

medida que o horizonte de previsão aumenta o modelo tende estabilizar para um valor de consumo médio. O erro percentual médio para a simulação livre de 60 passos não pôde ser calculado por não existirem amostras de validação.

## 5.2 Modelo NARIMA

O modelo NARIMA foi estimado a partir da série de consumo da CEMIG sem a componente de tendência (Figura 3.7), da observação 1 até a 104, sendo que os 12 meses seguintes foram utilizados para escolher o seguinte modelo:

$$\begin{aligned}
 y(k) = & +0,895 y(k-12) + 0,885 y(k-1) \\
 & -0,895 y(k-13) + 0,115 y(k-2) \\
 & -2,290 y(k-5)y(k-7) \\
 & +2,290 y(k-6)y(k-8)
 \end{aligned} \quad (5.2)$$

Este modelo foi obtido a partir da componente sazonal, e a Figura 5.3 mostra a previsão livre de 12 passos à frente do modelo 5.2 já adicionada a componente tendência.

A simulação livre de 12 passos à frente, de acordo com a Figura 5.3 apresenta erro percentual médio de -0,075%.

Da mesma forma que foi apresentado para o modelo ARIMA, na Seção 5.1, o modelo (5.2) agora é simulado 60 meses (5 anos) à frente e na sua saída é adicionada a componente de tendência, veja o esquema na Figura 3.7 (b).

A previsão livre de 60 meses à frente, apresentada na Figura 5.4, consegue absorver a dinâmica da série de consumo representado as oscilações sazonais e mudanças de tendência que acontecem a partir da observação 104 (agosto de 2005) até a observação 176 (julho de 2011).

## 5.3 Rede Neuro-Fuzzy

De acordo com a seção 3.6.2 são apresentados os resultados. Para obtenção do modelo Neuro-Fuzzy a tendência foi extraída. A Figura 5.5 apresenta a simulação da rede Neuro-Fuzzy com previsão livre de 12 passos à frente usando os dados de consumo de energia do Estado de Minas



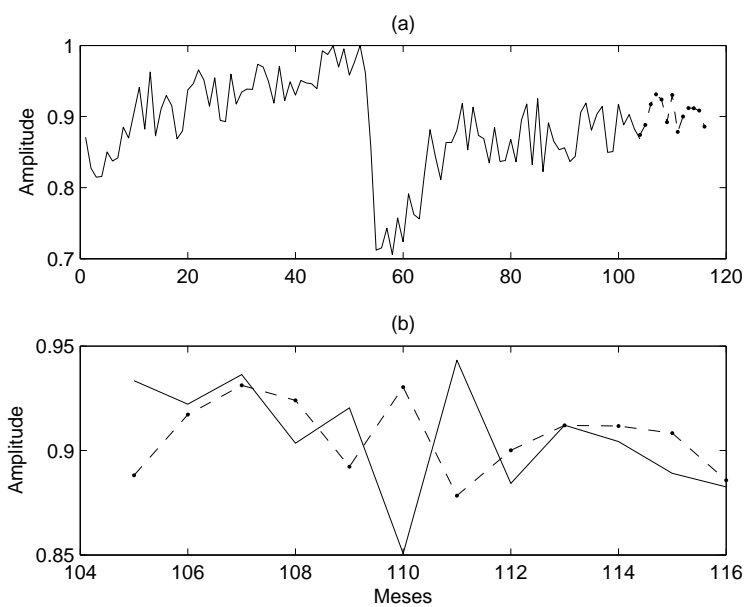


Figura 5.3: Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 12 passos à frente, e (b) apresenta o detalhe dos dados de validação com a saída do modelo NARIMA. (—) consumo real, (-.-) consumo previsto. Erros apresentados, MPE: -0,075%; MAPE: 2,723%; RMSE: 1,354;  $U_{Theil}$ : 1,114.

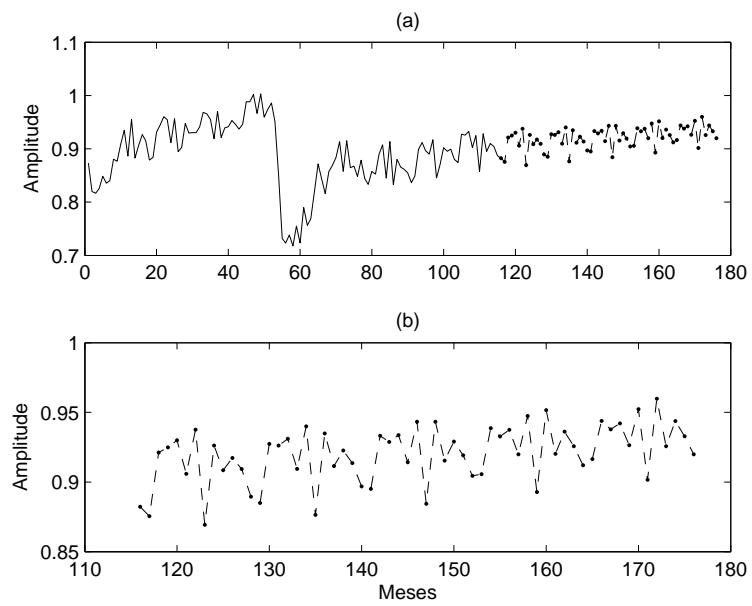


Figura 5.4: Simulação do modelo NARIMA. (a) apresenta os dados de consumo de energia com a previsão de 60 passos à frente, e (b) apresenta o detalhe do modelo NARIMA (-.-) consumo previsto.

Gerais. Nesta simulação o modelo Neuro-Fuzzy é apresentado já somado à tendência. A rede foi escolhida em função do índice de desempenho MPE.

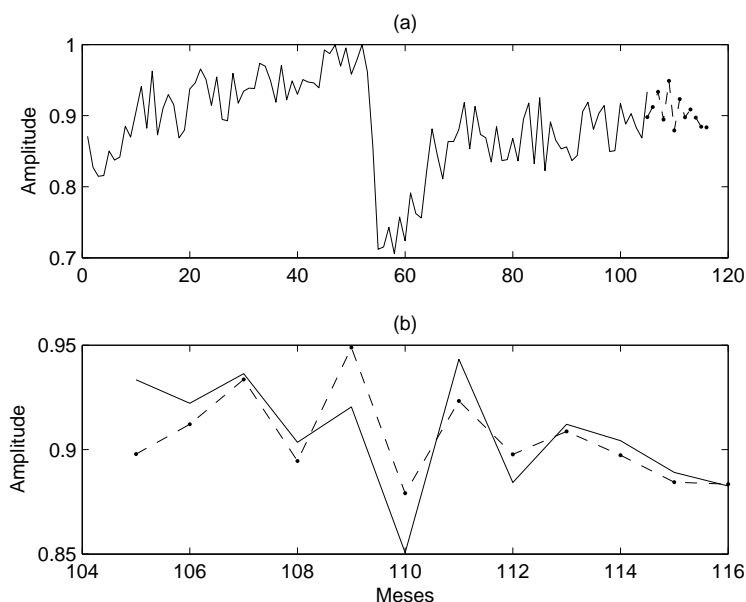


Figura 5.5: Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede. (–) consumo real, (–.–) consumo previsto. Erros apresentados, MPE: 0,431%; MAPE: 1,883%; RMSE: 0,762;  $U_{Theil}$ : 0,564.

A previsão de 12 passos do consumo de energia do Estado de Minas Gerais apresentou erro MPE de 0,431%, MAPE de 1,883%, RMSE de 0,762 e  $U_{Theil}$  de 0,564.

Na validação de 12 passos à frente, Figura 5.5 (b), nota-se que a RNF consegue absorver a dinâmica dos dados reais de consumo de energia elétrica. As oscilações da série temporal foi capturada nestes 12 passos, como também a tendência positiva, que indica crescimento no consumo de energia.

A Figura 5.6 mostra as previsões de consumo a partir de janeiro de 2007 até dezembro de 2011, ou seja, 60 meses à frente.

Claro que, como ainda não estão disponíveis os dados de consumo de energia a partir de 2007, a previsão de 60 meses à frente a partir de janeiro

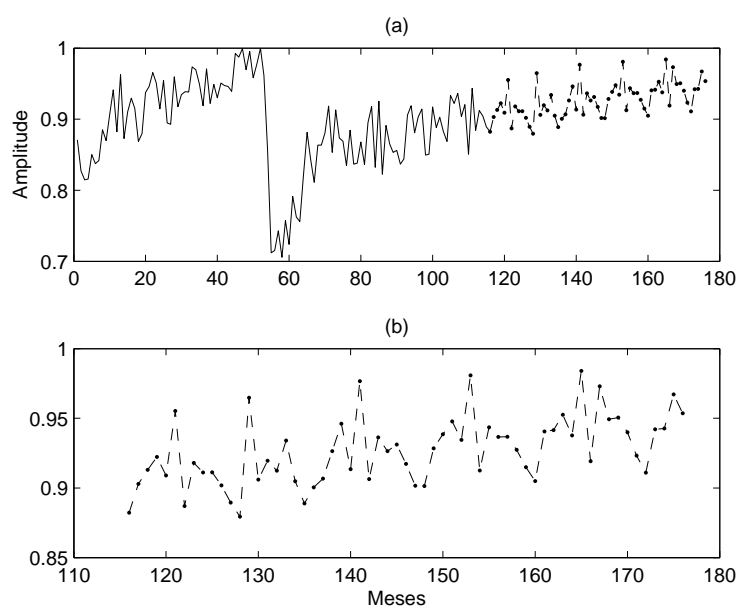


Figura 5.6: Simulação da rede Neuro-Fuzzy. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe da saída da rede (-.-) consumo previsto.

de 2007 é uma estimativa e por isso não tem como validar estas previsões através de índices de erro, como exemplo o MPE, MAPE, RMSE e o  $U$  de Theil. No entanto, o estudo de caso 1, seção 4, apresentou resultados apreciáveis, com índices de erros muito pequenos, tanto para 12 passos, quanto para 60 passos à frente.

Sendo assim, a estimativa do consumo de energia para o Estado de Minas Gerais mostra-se válida, mesmo porque a resposta da rede para a previsão de 60 meses à frente (Figura 5.6) apresenta flutuações sazonais parecidas com o último ciclo da série de consumo, e também a rede consegue generalizar de maneira satisfatória a tendência de consumo.

## 5.4 Rede Neural Artificial

Nesta seção é apresentado os resultados obtidos com uso da Rede Neural Artificial. Toda a metodologia para obtenção desses resultados foi mostrada em 3.7.2. O modelo Neural aqui apresentado foi obtido sem

a componente de tendência, e na saída dele a componente é adicionada (Figura 3.11).

A Figura 5.7 mostra os dados de treinamento e a validação de 12 meses à frente do modelo obtido, com atrasos [1-4,12] e número de neurônios na camada escondida,  $n = 8$ .

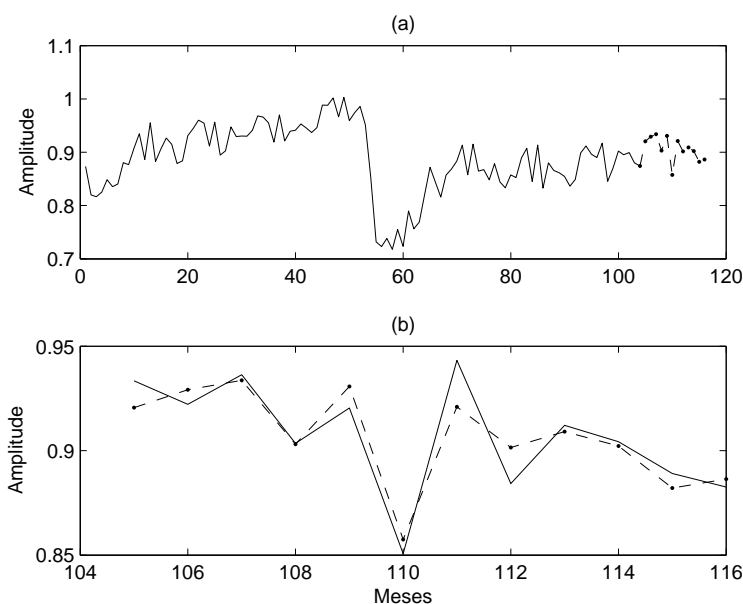


Figura 5.7: Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 12 passos à frente, e a (b) apresenta o detalhe dos dados de validação com a saída da rede (—) consumo real, (-.-) consumo previsto. Erro apresentados, MPE: -0,361%; MAPE: 1,503%; RMSE: 0,590;  $U_{Theil}$ : 0,313.

Vale lembrar que os resultados apresentados referem-se à previsão livre de 12 meses à frente. De acordo com os procedimentos (seção 3.7.2) a componente de tendência foi extraída dos dados de consumo e modelo neural foi obtido a partir da componente sazonal. À saída da rede foi somada a tendência. A Figura 5.7 (b) mostra a comparação da previsão com os dados reais de consumo de energia.

Para a previsão livre de 60 meses à frente, Figura 5.8, não existem dados para comparar a eficiência da rede. No entanto, a confiabilidade das previsões apresentadas aqui tomam como base os resultados obtidos com a série de consumo de New England (seção 4.4). A Figura 5.8 mostra

a previsão de 60 meses à frente, a partir de julho de 2006 até o horizonte de previsão, agosto de 2011.

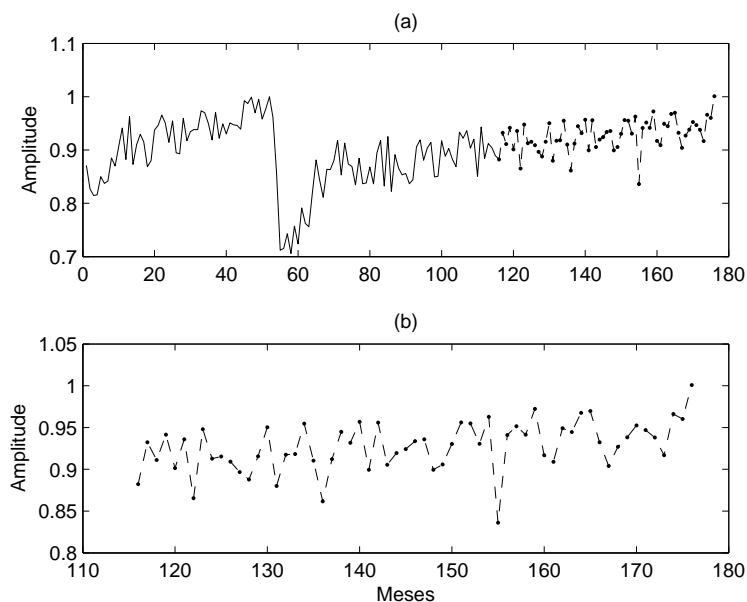


Figura 5.8: Simulação da Rede Neural Artificial. (a) apresenta a comparação do consumo de energia com os dados de validação de 60 passos à frente, e a (b) apresenta o detalhe da saída da rede. (–) consumo real, (–) consumo previsto.

Pode-se perceber na Figura 5.8 que as flutuações previstas pelo modelo mantêm-se aparentemente constantes durante a janela de previsão.

## 5.5 Comparação das Representações

Neste capítulo a comparação dos modelos para a previsão do consumo de energia do Estado de Minas Gerais em função dos índices de desempenho só pode ser realizada para a previsão livre de 12 passos à frente (Tabela 5.1). Para a de 60 os índices não podem ser calculados, pois, as amostras não estão disponíveis para a validação.

Para análise das previsões livre a partir de julho de 2006 é realizada uma comparação entre as previsões geradas por cada modelo. Esta análise consiste em comparar visualmente as previsões livre (Figura 5.9), e

Tabela 5.1: Comparação dos modelos através do índice de desempenho: MPE(%), MAPE(%), RMSE e  $U_{Theil}$  para a série de consumo em Minas Gerais.

Previsão	MPE		MAPE		RMSE		$U_{Theil}$	
	12	60	12	60	12	60	12	60
ARIMA	1,956	–	3,557	–	1,732	–	1,425	–
NARIMA	<b>-0,075</b>	–	2,723	–	1,354	–	1,114	–
RNF	0,431	–	1,883	–	0,762	–	0,564	–
RNA	-0,361	–	<b>1,503</b>	–	<b>0,509</b>	–	<b>0,313</b>	–

comparar a resposta espectral dessas previsões.

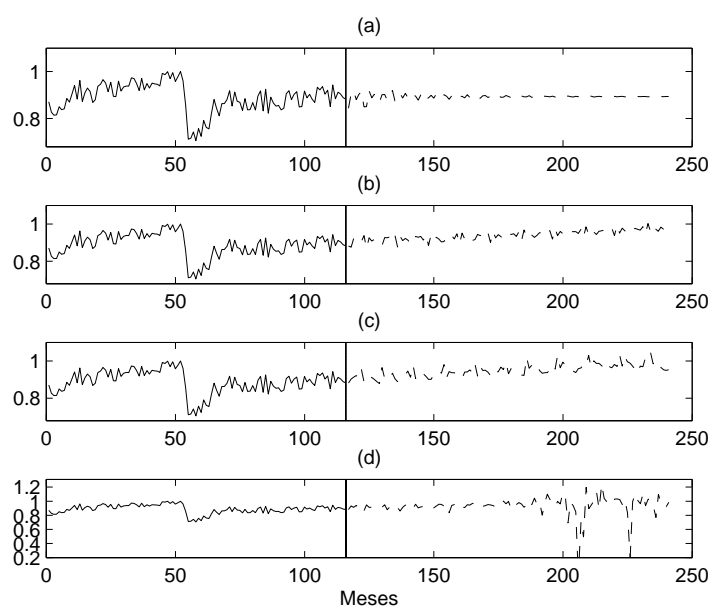


Figura 5.9: Comparação das simulações dos modelos: (a) ARIMA, (b) NARIMA, (c) RNF, (d) RNA. As previsões iniciam-se a partir da linha vertical contínua (observação 116).

Nesta comparação as previsões livres foram geradas para um horizonte de 125 passos à frente, isto é 10 anos e cinco meses, até dezembro de 2016. O horizonte de previsão foi aumentado para poder observar se no decorrer do meses o modelo conseguiria manter a tendência e oscilações dos dados

reais, claro que, considerando que nenhum distúrbio<sup>4</sup> afetasse o consumo de energia.

Comparando visualmente as previsões (Figura 5.9) nota-se que à medida que se aumenta o horizonte de previsão para o modelo ARIMA ele tende a amortecer as oscilações e estabilizar em um valor por volta de 0,9 (normalizado). Já as outras estruturas conseguem prolongar os ciclos sazonais e manter a tendência da série de consumo real. A previsão usando a RNA (Figura 5.9 (d)) mostra que o modelo diverge à medida que o horizonte de previsão aumenta, aumentando gradativamente os ciclos sazonais, o que acarreta aumento da variância das previsões.

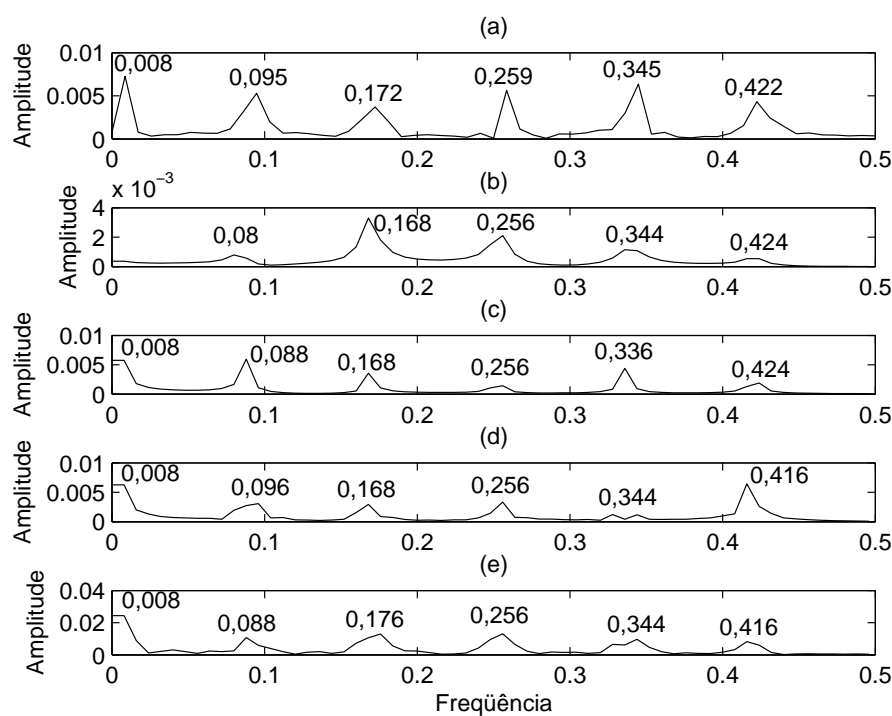


Figura 5.10: Comparação dos espectros de frequência: (a) série temporal de consumo do Estado de Minas Gerais, (b) ARIMA, (c) NARIMA, (d) RNF, (e) RNA.

De acordo com a Figura 5.9 e a Tabela 5.1 percebe-se que a RNA apresenta índices menores para previsões de 12 passos, porém, à medida que o

<sup>4</sup>Como exemplo cita-se o racionamento de energia que provoca mudanças consideráveis no consumo de energia.



horizonte de previsão aumenta a RNA tende à instabilidade, aumentando a variância das suas oscilações. O modelo ARIMA tende a estabilizar para um valor médio de consumo à longo-prazo. No entanto, os modelos NARIMA e RNF apresentam, a longo-prazo, oscilações bem parecidas com a da série histórica e conseguem manter a tendência crescente. Este resultado mostra que os modelos Neuro-Fuzzy e NARIMA podem ser ferramentas importantes quando o problema é realizar a previsões a longo-prazo de séries de consumo.

A Figura 5.10 mostra a comparação dos espectros de frequência<sup>5</sup> das previsões (de agosto de 2006 até dezembro de 2016) dos diferentes modelos obtidos com a série de consumo de Minas Gerais. De acordo com essa figura nota-se que o espectro do consumo de Minas Gerais (5.10 (a)) apresenta basicamente seis frequências. Enquanto que para o espectro das previsões das diferentes representações usadas nesse trabalho (5.10 (b), (c), (d) e (e)) observa-se que estas conseguiram absorver a dinâmica da série temporal de consumo de Minas Gerais, pois, conseguem generalizar as principais frequências da série temporal. Contudo, é preciso tomar o seguinte cuidado ao analisar a Figura 5.10. A única informação que o espectro pode dizer é a forma de onda da série temporal. Uma forma de onda idêntica, mas deslocada, seria uma má previsão, mas daria exatamente o mesmo espectro.

---

<sup>5</sup>Análise e obtenção de espectro de frequência a partir da Transformada de Fourier podem ser encontrados em Oppenheim e Schaffer (1989).



## Discussão e Conclusão

---

Aplicou-se técnicas de identificação e inteligência computacional linear e não-linear na obtenção de modelos para estimar o consumo de energia elétrica na cidade de New England e no Estado de Minas Gerais. Para a modelagem usou-se quatro representações, ARIMA, NARIMA, Rede Neural e Rede Neuro-Fuzzy, sendo a primeira linear e as três últimas não-lineares.

A primeira série apresenta um conjunto maior de amostras históricas de consumo, e foi usada para avaliar a robustez das representações citadas. A segunda série temporal apresenta um conjunto de dados com menor número de observações e ainda mudança na dinâmica da mesma como consequência ao racionamento que aconteceu em 2001 devido ao baixo nível de água nos reservatórios. Os quatro tipos de representações foram aplicadas e a previsão de 5 anos à frente foi estimada.

No Capítulo 3, toda metodologia usada para obtenção dos modelos foi detalhada. É importante ressaltar que a série de consumo do Estado de Minas Gerais, para as representações não-lineares, foi necessário separar a componente de tendência dos dados e estimar o modelo a partir da sazonalidade da série temporal. A tendência foi somada na saída do modelo. O mesmo procedimento foi realizado para a série de New England, usando a representação NARIMA.

Outros procedimentos foram analisados, entretanto, não produziram resultados melhores do que os apresentados nesta dissertação. Por exemplo, retirou-se os dados pré racionamento da série de Minas Gerais e os modelos não-lineares não demonstraram convergência rápida e precisão sobre os dados. A vantagem dos procedimentos usados no Capítulo 3 é que quando se decompõe a série temporal em suas componentes e modela-se apenas a sazonalidade, as representações conseguem interpretar os ciclos

sazonais e estimá-los, extrapolando-os para as previsões de 1 ano e 5 anos à frente.

Nos Capítulos 4 e 5 foram apresentados os resultados dos modelos para a série de New England e de Minas Gerais, respectivamente.

No Capítulo 4 todas representações apresentaram desempenhos favoráveis, como visto na Tabela 4.1. Se for levado em consideração o erro médio quadrático (MPE) apresentado nesta Tabela, a representação RNA atingiu 0,204% de erro para a previsão livre de 12 passos, e para 60 passos o ARIMA apresentou 0,129% de erro, e as outras representações conseguiram atingir índices de erro bem próximo destes citados. Este resultado comprova como as representações ARIMA, NARIMA, RNA, RNF, podem ser importantes ferramentas para a previsão do consumo de energia da cidade de New England. Neste Capítulo apenas a representação NARIMA foi obtida sem a componente de tendência, e esta foi somada na saída do modelo.

Na previsão de 60 meses à frente o ARIMA mostrou que a medida que o horizonte de previsão aumenta, as oscilações sazonais do modelo tendem a ficar amortecidas (Figura 4.2), subentendendo-se que, extrapolando para infinitos meses à frente o modelo irá convergir para um valor fixo de consumo. Este fato é comprovado na Seção 5.5, onde a previsão foi extrapolada para 125 meses à frente e o consumo convergiu para um valor fixo. Isso mostra que o modelo ARIMA é bastante eficiente para previsões de curto e médio-prazos, pois quando o horizonte de previsão é muito longo esta representação tende a capturar a média extrapolada, ou seja, o modelo ARIMA tende a modelar a tendência da série temporal, sendo um ótimo previsor de tendência a longo-prazo.

No Capítulo 5 os desempenhos foram medidos apenas para as previsões livre de 12 meses à frente, pois, a série de consumo do Estado de Minas tem apenas 116 amostras. A Tabela 5.1 detalha os índices de desempenho dos modelos. O modelo NARIMA apresentou erro percentual médio de 0,075% para previsão livre de 12 meses, a RNA de 0,361%, a RNF de 0,431% e o ARIMA de 1,956%. O NARIMA apresentou o menor índice entre as representações. Todas as representações não-lineares (NARIMA, RNF e RNA) foram obtidas extraindo-se a tendência, e após a obtenção esta foi adicionada ao modelo.

Como não existem amostras para a validação das previsões de 60 meses à frente, uma alternativa encontrada foi de primeiro extrapolar as

previsões em cada modelo (Figura 5.9), e depois comparar o espectro de frequência dessas previsões (Figura 5.10). Quando aumenta-se o horizonte de previsões, percebe-se que o modelo ARIMA converge para um valor de consumo fixo, como já foi citado acima. Na Figura 5.9 (b) o modelo NARIMA conseguiu manter oscilações para as previsões livre de 125 meses à frente, demonstrando robustez e eficiência na previsão a longo-prazo do consumo de energia de Minas Gerais. O mesmo aconteceu com o modelo Neuro-Fuzzy que conseguiu generalizar as oscilações sazonais que existem na série para o horizonte de previsão sem que houvesse convergência para um valor.

Em relação à previsão usando a RNA notou-se que esta representação diverge à medida que o horizonte de previsão aumenta. A divergência se dá devido às oscilações sazonais tenderem a aumentar, aumentando-se assim a variância, como foi mostrado na Figura 5.9 (d). De acordo com os índices de desempenho apresentados, a RNA pode ser considerada um modelo eficiente para previsões de médio-prazo (1 ano à frente). Quando o horizonte é maior, longo-prazo, a RNF mostra eficiente, pois, este modelo consegue mapear os ciclos sazonais das séries de consumo e “copiá-los” para o horizonte de previsão, como pode ser notado na Figura 5.9.

Em relação à comparação espectral das previsões pode-se concluir que os espectros (Figura 5.10) apresentam basicamente as mesmas frequências. Este fato leva a concluir que as previsões, estimadas a partir das diferentes representações, conseguem capturar as principais oscilações da série histórica e as replicam para o horizonte previsto.

Em relação à retirada ou não da tendência das séries temporais para a modelagem pode-se concluir que o modelo NARIMA converge apenas quando a componente é extraída, isto para qualquer um dos estudos de caso aqui abordados. Quanto à retirada da tendência para os modelos RNF e RNA, isso foi preciso para a série temporal da CEMIG. Acredita-se que este fato deve-se à dinâmica diferente entre as duas séries temporais apresentadas neste trabalho. Na série de New England os ciclos sazonais são mais bem definidos e, a partir da amostra 100, as mudanças de níveis (tendência) são menos bruscas, o que já não acontece na série da CEMIG. A série da CEMIG não apresenta ciclos sazonais bem definidos e apresenta uma mudança brusca de nível devido ao racionamento em 2001.

Dessas discussões apresentadas e dos resultados obtidos conclui-se que as representações ARIMA, NARIMA, RNF e RNA, usadas neste trabalho,

são eficientes para a previsão de 5 anos à frente do consumo de energia da cidade de New England e do Estado de Minas Gerais.

## **6.1 Proposta para trabalhos futuros**

A seguir sugerem-se alguns aspectos que ainda podem ser explorados:

- pesquisar / utilizar outras ferramentas para a previsão a longo-prazo do consumo de energia;
- aplicar o procedimento apresentado a outros sistemas reais;
- estudo mais aprofundado dos modelos NARIMA. Obter respostas sobre a instabilidade desses modelos quando aplicada à metodologia de Box e Jenkins. Procurar responder o por quê esta metodologia é instável quando aplicada em séries temporais com tendências.

---

## Referências Bibliográficas

---

- Aguirre, L.A., Aguirre, L., Mendes, E.M.A.M., e Martinez, C.B. Pesquisa de modelo de previsão de energia, para curto e médio prazo, utilizando inteligência computacional. Technical report, Universidade Federal de Minas Gerais, 2006. Relatório Técnico Interno.
- Aguirre, L.A. e Billings, S.A. Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3): 569–587, 1995.
- Aguirre, L.A., Rodrigues, D.D., Lima, S.T., e Martinez, C.B. Dynamical prediction and pattern mapping in short-term load forecasting. *Int. J. Electrical Power and Energy Systems*, 30:73–78, 2008.
- Aguirre, Luis Antonio. *Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais*. UFMG, 2007. ISBN 85-7041-220-7. 3ª Edição Revista e Ampliada.
- Al-Saba, T. e El-Amin, I. Artificial neural networks as applied to long-term forecasting. *Artificial Intelligence in Engineering*, 13:189–197, 1999.
- Ang, X.Q. Liu and B.W. e Goh, T.N. Forecasting of electricity consumption: a comparison between an econometric model and a neural network. In *Proceeding of the IEEE International Joint Conference on Neural Networks*, pages 1254–1259, 1991.
- Aslanargun, A., Mammadov, M., Yazici, B., e Yolacan, S. Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting. *Journal of Statistical Computation and Simulation*, 77 (1):29–53, Janeiro 2007.

- Bakirtzis, A.G., Petridis, V., Kiartzis, S.J., Alexiadis, M.C., e Maissis, A.H. A neural network short-term load forecasting model for the greek power system. *IEEE Transactions on Power Systems*, 11:858–863, 1996.
- Billings, S.A., Chen, S., e Korenberg, M.J. Identification of mimo nonlinear systems using a forward-regression estimator. *International Journal Control*, 49(6):2157–2189, 1989.
- Box, G.E.P. e Jenkins, G.M. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden Day, 1970.
- Box, G.E.P., Jenkins, G.M., e Reinsel, G. *Time Series Analysis: Forecasting and Control*. Englewood Cliffs: Prentice Hall, third edition, 1994.
- Braga, A.P., Carvalho, A.C., e Ludemir, T.B. *Redes Neurais Artificiais: teoria e aplicações*. Livraria Técnica Científica, 2000.
- Campos, R.J., Jesus, T.A., e Mendes, E.M.A.M. Uma abordagem fuzzy para a previsão de curto-prazo do consumo de energia elétrica. In *XXX Congresso Nacional de Matemática Aplicada - CNMAC*, Brasil, Florianópolis, Setembro 2007.
- Charytoniuk, W., Chen, M.S., e Olinda, P. Van. Nonparametric regression based short-term load forecasting. *IEEE Transactions on Power Systems*, 13:725–730, 1998.
- Chen, B.J., Chang, M.W., e Lin, C.J. Load forecasting using support vector machines: A study on eunite competition 2001. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, 2002.
- Chen, H., Canizares, C.A., e Singh, A. Ann-based short-term load forecasting in electricity markets. *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, 2:411–415, 2001.
- Cho, M.Y., Hwang, J.C., e Chen, C.S. Customer short-term load forecasting by using ARIMA transfer function model. In 1., editor, *Proceedings of the International Conference on Energy Management and Power Delivery*, pages 317–322, 1995.



- Chow, T.W.S. e Leung, C.-T. Nonlinear autoregressive integrated neural network model for short-term load forecasting. In 143, Vol., editor, *IEE Proceedings Gener. Trasm. Distrib.* No 5, 1996.
- Coelho, M.C.S. Modelos de Hammerstein e de Wiener: conexões com modelos narx e sua aplicação em identificação de sistemas não-lineares. Master's thesis, Universidade Federal de Minas Gerais, Brasil, Belo Horizonte, 2002.
- Coelho, M.C.S. *Análise de Determinismo em Séries Temporais Quase-Periódicas*. PhD thesis, Universidade Federal de Minas Gerais, Brasil, Belo Horizonte, Dezembro 2006.
- Corrêa, Marcelo Vieira. *Identificação caixa-cinza de sistemas não-lineares utilizando representações NARMAX racionais e polinomiais*. PhD thesis, Universidade Federal de Minas Gerais, Brasil, Belo Horizonte, 2001.
- Cox, E. *The Fuzzy Systems Handbook*. Academy Press Inc., Estados Unidos, 1994.
- Da, X., Jiangyan, Y., e Jilai, Y. The physical series algorithm of mid-long term load forecasting of power systems. *Electric Power Systems Research*, 53:31–37, 2000.
- De Gooijer, Jan G. e Hyndman, Rob J. 25 years of time series forecasting. *International Journal of Forecasting*, 22(3):443–473, 2006. available at <http://ideas.repec.org/a/eee/intfor/v22y2006i3p443-473.html>.
- Engle, R.F., Mustafa, C., e Rice, J. Modeling peak electricity demand. *Journal of Forecasting*, 11:241–251, 1992.
- Fair, R.C. *Evaluating the predictive accuracy of models*. In Griliches, Z. and Intriligator, volume III, chapter 33. Elsevier Science Publishers, Amsterdam, 1986.
- Fan, J.Y. e McDonald, J.D. A real-time implementation of short-term load forecasting for distribution power systems. *IEEE Transactions on Power Systems*, 9:988–994, 1994.
- Feinberg, E.A., Hajagos, J.T., e Genethliou, D. Load pocket modeling. In *Proceedings of the 2nd IASTED International Conference: Power and Energy Systems*, pages 50–54, Crete, 2002. Palm Springs.

- Feinberg, E.A., Hajagos, J.T., e Genethliou, D. Statistical load modeling. In *Proceedings of the 7th IASTED International Multi-Conference: Power and Energy Systems*, pages 88–91, CA, 2003. Palm Springs.
- Fogel, D.B. An introduction to simulated evolutionary optimization. *IEEE Transactions on Neural Networks*, 5:3–14, 1994.
- Fung, Y.H. e Tummala, V.M.R. Forecasting of electricity consumption: a comparative analysis of regression and artificial neural networks models. In *IEE 2nd International Conferenc on Advances in Power Systems Control, Operation and Management*, pages 782–787, 1993.
- Ghiassi, M., Saidane, H., e Zimbra, D.K. A dynamic artificial neural network model for forecasting time series events. *International Journal of Forecasting*, 21:341–362, 2005a.
- Ghiassi, M., Zimbra, D.K., e Saidane, H. Medium term system load forecasting with a dynamic artificial neural network model. *Electric Power Systems Rechearch*, 76:302–316, October 2005b.
- Grubbs, Frank. Procedures for detecting outlying observations in samples. *Technometrics*, 11(1):1–21, 1969.
- Haida, T. e Muto, S. Regression based peak load forecasting using a transformation technique. *IEEE Transactions on Power Systems*, 9:1788–1794, 1994.
- Haydari, Z., Kavehnia, F., Askari, M., e Ganbariyan, M. Time series load modelling and load forecasting using neuro-fuzzy techniques. In *9th International Conference - Electrical Power Quality and Utilization*, Espanha, Barcelona, Outubro 2007.
- Haykin, Simon. *Redes Neurais: princípios e prática*. Bookman, Porto Alegre, segunda edition, 2001.
- Hippert, H.S., Pedreira, C.E., e Souza, R.C. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16:44–55, 2001.
- Ho, K.L., Hsu, Y.Y., Chen, F.F., Lee, T.E., Liang, C.C., Lai, T.S., e Chen, K.K. Short-term load forecasting of taiwan power system using a knowledge

- based expert system. *IEEE Transactions on Power Systems*, 5:1214–1221, 1990.
- Hodrick, Robert e Prescott, Edward C. Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, 1997.
- Hyde, O. e Hodnett, P.F. An adaptable automated procedure for short-term electricity load forecasting. *IEEE Transactions on Power Systems*, 12:84–93, 1997.
- Jang, J.S.Roger, Sun, C.T., e Mizutani, E. *Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence*. NJ: Prentice Hall, 1997.
- Jang, Jyh-Shing Roger. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665–685, 1993.
- Kandil, M.S., El-Debeiky, S.M., e Hasanien, N.E. The implementation of long-term forecasting strategies using a knowledge-based expert system: part ii. *Electric Power Systems Research*, 58:19–25, 2001.
- Kiartzis, S.J. e Bakirtzis, A.G. A fuzzy expert system for peak load forecasting: Application to the greek power system. In 3., editor, *Proceedings of the 10th Mediterranean Electrotechnical Conference*, pages 1097–1100, 2000.
- Li, Y. e Fang, T. Wavelet and support vector machines for short-term electrical load forecasting. In *Proceedings of International Conference on Wavelet Analysis and its Applications*, pages 399–404, 2003.
- Ljung, L. *System Identification: Theory for the user*. Prentice Hall, 1987.
- Makridakis, S., Wheelwright, S.C., e Myrdman, R.J. *Forecasting: Methods and Applications*. John Wiley and Sons, Hoboken, NJ, 3rd edition, 1998.
- Mathworks, The. Fuzzy logic toolbox. [www.mathworks.com](http://www.mathworks.com), 2001.
- Mauro, Daniel Fonseca. Desenvolvimento de um módulo baseado em redes MLP de um sistema de previsão de consumo de energia. Technical report, Universidade Federal de Minas Gerais, Junho 2007. Monografia.

- Miranda, V. e Monteiro, C. Fuzzy inference in spatial load forecasting. In *Proceedings of IEEE Power Engineering Winter Meeting*, pages 1063–1068, 2000.
- Mizutani, E. e Jang, J.S.R. Coactive neural fuzzy modeling. *IEEE Proceedings of the International Conference on Neural Networks*, pages 760–765, 1995.
- Mohandes, M. Support vector machines for short-term electrical load forecasting. *International Journal of Energy Research*, 26:335–345, 2002.
- Mohr, M. A trend-cycle-season filter. *ECB Working Paper Series*, (499), Julho 2005.
- Morettin, Pedro A. e Toloí, Clélia M.C. *Análise de Séries Temporais*. Edgard Blücher, São Paulo, segunda edição revista e ampliada edition, 2006.
- Mori, H. e Kosemura, N. Optimal regression tree based rule discovery for short-term load forecasting. In *Proceedings of IEEE Power Engineering Society Transmission and Distribution Conference*, pages 421–426, 2001.
- Oliveira, A. Novo modelo elétrico - papel ampliado para geração distribuída. In *Seminário INEE - Instituto Nacional de Eficiência Energética*, Setembro 2004.
- Oppenheim, A.V. e Schaffer, R.W. *Discrete-Time Signal Processing*. Prentice-Hall, 1989.
- Pandian, S.C., Duraiswamy, K., Rajan, C.C.A., e Kanagaraj, N. Fuzzy approach for short term load forecasting. *Electric Power Systems Research*, 76:541–548, 2006.
- Papalexopoulos, A.D., Hao, S., e Peng, T.M. An implementation of a neural network based load forecasting model for the ems. *IEEE Transactions on Power Systems*, 9:1956–1962, 1994.
- Peng, M., Hubele, N.F., e Karady, G.G. Advancement in the application of neural networks for short-term load forecasting. *IEEE Transactions on Power Systems*, 7:250–257, 1992.
- Reis, A.J.R. Pré-processamento de séries temporais para um previsor de carga elétrica a curto prazo via redes neurais artificiais. In *XIII Congresso Brasileiro de Automática*, pages 2036–2041, 1999.

- Rodrigues, D.D. Previsão de fornecimento de energia elétrica: Um estudo de caso do ponto de vista do barramento. Master's thesis, Universidade Federal de Minas Gerais, Brasil, Belo Horizonte, Julho 2002.
- Rumelhart, D. e McClelland, J. *Parallel Distributed Processing*, volume 1. MIT Press, Cambridge, MA, EUA, 1986.
- Ruzic, S., Vuckovic, A., e Nikolic, N. Weather sensitive method for short-term load forecasting in electric power utility of serbia. *IEEE Transactions on Power Systems*, 18:1581–1586, 1998.
- Söderström, T. e Stoica, P. *System Identification*. Prentice-Hall, 1989.
- Serrão, F.C.C. Modelo de previsão de carga de curto prazo utilizando redes neurais e lógica fuzzy. Master's thesis, Pontifícia Universidade Católica, Brasil, Rio de Janeiro, 2003.
- Skarman, S.E. e Georgiopoulos, M. Short-term electrical load forecasting using a fuzzy artmap neural network. In *Proceedings of SPIE*, pages 181–191, 1998.
- Theil, H. *Economic forecasts and policy*. North-Holland Publishing Co., 1958.
- Yang, H.T. e Huang, C.M. A new short-term load forecasting approach using self-organizing fuzzy armax models. *IEEE Transactions on Power Systems*, 13:217–225, 1998.
- Yang, H.T., Huang, C.M., e Huang, C.L. Identification of ARMAX model for short-term load forecasting: An evolutionary programming approach. *IEEE Transactions on Power Systems*, 11:403–408, 1996.
- Zadeh, L.A. Fuzzy sets. *Fuzzy Sets, Information and Control*, 8:338–353, 1965.
- Zhang, G., Patuwo, E., e Hu, M. Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, (14):35–62, 1998.

