Universidade Federal
do Rio de Janeiro
Escola Politécnica

PREVISÃO DE CARGA ATRAVÉS DE MODELOS NEURO-FUZZY

Victor Andrade de Almeida

Projeto de Graduação apresentado ao curso de Engenharia Elétrica da Escola Politécnica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do grau de Engenheiro Eletricista.

Orientador: Tatiana Mariano Lessa de Assis, D.Sc.

Coorientador: José Francisco Moreira Pessanha, D.Sc.

Rio de Janeiro

Agosto de 2013

PREVISÃO DE CARGA ATRAVÉS DE MODELOS NEURO-FUZZY

Victor Andrade de Almeida

PROJETO SUBMETIDO AO CORPO DOCENTE DO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA ESCOLA POLITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE ENGENHEIRO ELETRICISTA.

Examinada por:

Prof. Tatiana Mariano Lessa de Assis, D.Sc.
(Orientador)

Prof. José Francisco Moreira Pessanha, D.Sc.
(Coorientador)

Prof. Oumar Diene, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

Agosto de 2013

Almeida, Victor Andrade de.

Previsão de Carga através de Modelos Neuro-Fuzzy / Victor Andrade de Almeida – Rio de Janeiro: UFRJ/ESCOLA POLITÉCNICA, 2013.

XI, 80 p.: il.; 29,7 cm.

Orientadora: Tatiana Mariano Lessa de Assis.

Projeto de Graduação – UFRJ/ POLI/ Engenharia Elétrica, 2013.

Referências Bibliográficas: p. 60-61.

1. Previsão de Carga. 2. Redes Neurais. 3. Lógica Fuzzy. I. Assis, Tatiana Mariano Lessa de. II. Universidade Federal do Rio de Janeiro, UFRJ, Engenharia Elétrica. III. Título.

Agradecimentos

Agradeço primeiramente a Deus, por sempre me guiar pelos longos caminhos da vida.

Aos meus pais, Nizeuma e Fernando, por acreditarem no meu potencial e me apoiarem de maneira incondicional em todas as fases da minha formação como profissional e como ser humano.

À minha orientadora, Tatiana Mariano Lessa de Assis, pela colaboração e paciência em todos os momentos da elaboração deste trabalho.

Ao meu coorientador, José Francisco Moreira Pessanha, pelos conhecimentos transmitidos durante os 2 anos de convivência direta e grande colaboração para a elaboração deste trabalho.

Aos professores do Departamento de Engenharia Elétrica por sempre tentarem, em cada disciplina ministrada, extrair de nós alunos o melhor.

A todos os meus amigos do curso de graduação, em especial ao Renato Silva, Renato Valadão e Rodrigo Leonel. A amizade de vocês tornou tudo muito menos complicado.

Por fim, agradeço à minha namorada, Aline, que esteve ao meu lado em todos os momentos do vestibular à formatura, me dando forças e acreditando em mim quando eu mesmo duvidava que fosse capaz.

Resumo do Projeto de Graduação apresentado a Escola Politécnica/UFRJ

como parte dos requisitos necessários para a obtenção do grau de Engenheiro

Eletricista.

PREVISÃO DE CARGA ATRAVÉS DE MODELOS NEURO-FUZZY

Victor Andrade de Almeida

Agosto de 2013

Orientadora: Tatiana Mariano Lessa de Assis

Coorientador: José Francisco Moreira Pessanha

Curso: Engenharia Elétrica

Este projeto de graduação apresenta um modelo de previsão de carga de sistemas de energia elétrica utilizando de forma combinada Redes Neurais Artificiais e Lógica Fuzzy. Ao longo do trabalho são apresentados os fundamentos dessas duas técnicas de inteligência computacional, assim como a utilização de ambas na metodologia proposta. A partir de dados de carga históricos e outras informações como as datas de feriados e finais de semana, são realizadas separadamente previsões de ponta de carga e perfil de carga. São descritos de forma detalhada os dados de entrada, a estrutura do modelo proposto e também os ajustes realizados nos parâmetros do modelo a fim de obter-se uma maior qualidade na previsão. A programação do modelo, o tratamento de dados e avaliação do resultado foram realizados através do software MATLAB.

Palavras-chave: Previsão de Carga, Redes Neurais, Lógica Fuzzy.

Abstract of Undergraduate Project presented to POLI/UFRJ as a partial fulfillment of the requirements for the degree of Engineer.

LOAD FORECAST USING NEURO-FUZZY MODELS

Victor Andrade de Almeida

August 2013

Advisor: Tatiana Mariano Lessa de Assis

Co-Advisor: José Francisco Moreira Pessanha

Course: Electrical Engineering

The present project relates to a charge forecast model of electric power systems by the combined use of neural networks and fuzzy logic theories. Throughout the paper we present the foundations of these two computational intelligence techniques, as well as its use in the proposed methodology. Considering historical load data and other information such as holidays and weekend days, predictions are made separately for peak load and load profile. This paper describes in detail the type of input data, the structure of the proposed model and also the adjustments made on model parameters in order to obtain a higher quality prediction. Model programming, data processing and evaluation of the results were performed using MATLAB software.

Keywords: Load Forecasting, Neural Networks, Fuzzy Logic.

Sumário

1 INTRODUÇÃO 1

1.1 OBJETIVOS 1

1.2 ESTRUTURA DO TRABALHO 2

2 PREVISÃO DE CARGA 3

2.1 DESCRIÇÃO DO PROBLEMA 3

2.2 REVISÃO DA LITERATURA 4

2.2.1 Métodos Estatísticos de Previsão 5

2.2.2 Métodos Baseados em Inteligência Artificial 9

3 MODELOS INTELIGENTES 11

3.1 REDES NEURAIS ARTIFICIAIS 11

3.1.1 Tipos de Treinamento 14

3.1.2 Redes Neurais Supervisionadas 15

3.1.3 Redes Neurais Não-Supervisionadas 25

3.2 LÓGICA FUZZY 27

3.2.1 Sistema de Inferência Fuzzy (FIS) 29

3.2.2 Método de Wang e Mendel 29

3.2.3 Cálculo das Pertinências utilizando Fuzzy C-Means 33

3.3 MODELOS HÍBRIDOS NEURO-FUZZY 35

4 MODELO PROPOSTO PARA A PREVISÃO DE CARGA 36

4.1 DADOS DE ENTRADA 37

4.2 MODELO DE PREVISÃO 39

4.2.1 Tratamento de Dados 41

4.2.2 MLP 44

4.2.3 Mapa de Kohonen 46

4.2.4 Sistema de Inferência Fuzzy 48

4.2.5 Combinação dos Resultados 52

4.3 ÍNDICES DE DESEMPENHO 52

4.4 AJUSTE DO MODELO 54

4.4.1 MLP 54

4.4.2 Mapa de Kohonen 54

4.4.3 Sistema de Inferência Fuzzy (FIS) 54

4.5 PREVISÃO DE VALORES FUTUROS 55

5 IMPLEMENTAÇÃO DO MODELO E RESULTADOS 56

5.1 DADOS UTILIZADOS 56

5.2 PREVISÃO DA SÉRIE DE PONTA 56
5.3 PREVISÃO DA SÉRIE DO PERFIL DE CARGA 58
5.4 TESTES 60
5.4.1 Número Camadas Intermediárias 60
5.4.2 Número de Neurônios nas Camadas Intermediárias 60
5.4.3 Dias Anteriores à previsão (MLP) 61
5.4.4 Número de Neurônios do Mapa de Kohonen 61
5.4.5 Constante de Fuzzyficação 62
5.4.6 Dias antecedentes à Previsão (FIS) 62
5.5 RESULTADOS OBTIDOS 63
6 CONCLUSÕES 68
6.1 CONCLUSÕES 68
6.2 TRABALHOS FUTUROS 69
REFERÊNCIAS BIBLIOGRÁFICAS 70
ANEXO I 73

Lista de Figuras

Figura 3.1 – Neurônio Artificial. 11
Figura 3.2 – Neurônio Artificial com Bias incorporado. 13
Figura 3.3 – Funções de ativação : (a) Função Linear. (b) Função Degrau. (c) Função Sigmoidal. (d) Função Gaussiana 13
Figura 3.4 – Rede Neural Multicamadas 23
Figura 3.5 - Fases *Forward* e *Backward* do algoritmo de retropropagação do erro 24
Figura 3.6 – Mapa de Kohonen Bidimensional 25
Figura 3.7 – Redução gradual da região de vizinhança desde o instante t=1 até o instante t=3 27
Figura 3.8 - Conjuntos fuzzy e funções de pertinência para a variável temperatura 28
Figura 3.9 – Funções de Pertinência. (a) Função Trapezoidal. (b) Função Triangular. (c) Função Gaussiana. 28
Figura 3.10 – Série Temporal Genérica 31
Figura 3.11 – Ativação dos Conjuntos Fuzzy 32
Figura 3.12 – Curvas diárias (a), (b), (c) e (d). 34
Figura 4.1 - Diagrama das etapas da previsão 36
Figura 4.2 - Exemplo de série de entrada com 3 dias e discretização horária 38
Figura 4.3 – Fluxograma do modelo proposto. 40
Figura 4.4 - Vetor de Ponta de Carga 43
Figura 4.5 - Perfil da Carga 43
Figura 4.6 – Conjuntos de treinamento e validação 44
Figura 4.7 – Clusterização do dia 1 47
Figura 4.8 - Clusterização do dia 2 47
Figura 4.9 - Clusterização do dia 3 48
Figura 5.1 – Padrões de Treinamento 57
Figura 5.2 – Treinamento da rede supervisionada através da ferramenta *Neural Network Toolbox* 57
Figura 5.3 - Treinamento da rede não-supervisionada através da ferramenta *Neural Network Toolbox* 58
Figura 5.4 – Curvas horárias de entrada classificadas no Mapa de Kohonen com 36 neurônios 59
Figura 5.5 – Regras fuzzy 59

Figura 5.6 - Previsão de Carga Janeiro 2008 – Semana 1 64
Figura 5.7 - Previsão de Carga Janeiro 2008 – Semana 2 64
Figura 5.8 - Previsão de Carga Janeiro 2008 – Semana 3 65
Figura 5.9 - Previsão de Carga Janeiro 2008 – Semana 4 65
Figura 5.10 - Previsão de Carga Janeiro 2008 – Semana 5 66

Lista de Tabelas

Tabela 3.1 - Exemplos para treinamento do Perceptron 17
Tabela 3.2 – Exemplo de treinamento 1 18
Tabela 3.3 – Exemplo de Treinamento 2 19
Tabela 3.4 – Exemplo de Treinamento 3 19
Tabela 3.5 – Exemplo de Treinamento 4 20
Tabela 3.6 – Verificação do Exemplo de Treinamento 1 21
Tabela 3.7 - Verificação do Exemplo de Treinamento 2 21
Tabela 3.8 - Verificação do Exemplo de Treinamento 3 22
Tabela 3.9 - Verificação do Exemplo de Treinamento 4 22
Tabela 3.10 - Grau de Pertinência de cada Conjunto 32
Tabela 3.11 – Grau de Pertinência da regra 32
Tabela 4.1 - Exemplo de série de entrada com 3 dias e discretização horária 37
Tabela 4.2 – Vetores das variáveis explicativas 38
Tabela 4.3 – Vetor de Ponta de Carga 42
Tabela 4.4 – Matriz do Perfil da Carga 42
Tabela 4.5 – Padrão de treinamento 45
Tabela 4.6 – Padrões de entrada 45
Tabela 4.7 – Clusterização de Curvas 46
Tabela 4.8 – Exemplo de regra fuzzy 48
Tabela 5.1 – Dias anteriores à previsão - MLP 61
Tabela 5.2 - Dias anteriores à previsão - FIS 62
Tabela 5.3 – Índices de desempenho obtidos 63
Tabela 5.4 – MAPE diário da previsão de Janeiro de 2008 66
Tabela 5.5 – MAPE médio para cada dia da semana 67

Capítulo 1

1 Introdução

A previsão da carga dos sistemas de energia elétrica é uma tarefa fundamental. Seja na geração, transmissão, distribuição ou comercialização; no planejamento ou na operação, a previsão da carga é amplamente utilizada como base para tomadas de decisões que vão do tempo real ao longo prazo.

A partir de uma boa previsão de demanda, é possível decidir com segurança pelo aumento de geração de energia em um determinado horário do dia, ou até determinar quantas usinas geradoras precisam ser construídas em um horizonte de 10 anos.

Com a reestruturação do sistema elétrico, que implicou a desverticalização das empresas do setor elétrico, ocorrido no início da década de 2000, que separou o produto – a eletricidade – e os serviços de transmissão, distribuição e comercialização [1], a previsão de carga tornou-se ainda mais importante. Isso ocorre, pois, no atual modelo brasileiro, o custo de energia elétrica no futuro está diretamente ligado, entre outros fatores, ao preço da energia no presente.

Contudo, são muitas as variáveis que afetam a demanda de energia em um sistema elétrico, o que torna a previsão bastante complexa. A temperatura, os dias da semana, as estações do ano, e até algumas variáveis econômicas são exemplos desses fatores.

Vários modelos são capazes de fazer uma previsão para a carga, com maiores ou menores índices de erro. Entre eles, estão os métodos baseados em Inteligência Computacional, como as Redes Neurais Artificiais e a Lógica Fuzzy. Existem ainda os chamados métodos híbridos, que combinam diferentes técnicas em um só modelo. O presente trabalho apresenta em especial um desses modelos combinacionais, o modelo Neuro-Fuzzy.

1.1 Objetivos

O objetivo do trabalho é apresentar um modelo Neuro-Fuzzy, que consiste em uma metodologia baseada no uso combinado da Lógica Fuzzy e Redes Neurais Artificiais

supervisionadas e não supervisionadas. E, a partir desse modelo, obter uma previsão do perfil de carga diário, com discretização horária, em um total de 24 previsões por dia.

1.2 Estrutura do trabalho

Este trabalho divide-se em 6 capítulos. No capítulo 2 será feita uma apresentação do problema da previsão de carga em um sistema elétrico e uma breve revisão sobre os métodos mais utilizados para a previsão de séries temporais, incluindo suas vantagens e desvantagens.

No capítulo 3 serão apresentadas, de maneira mais aprofundada, as teorias dos modelos inteligentes que fazem parte do modelo Neuro-Fuzzy apresentado neste trabalho. Será mostrada a metodologia das Redes Neurais Artificiais, da Lógica Fuzzy e ainda a dos modelos híbridos, que combinam dois ou mais tipos de modelos inteligentes.

No capítulo 4 será feita a apresentação do modelo proposto para a previsão da carga. Serão mostradas a estrutura do modelo, os dados utilizados na previsão e o processo adotado para o ajuste do modelo.

No capítulo 5 serão mostrados os resultados da previsão de carga obtidos através do modelo proposto.

No capitulo 6, são desenvolvidas as conclusões a respeito dos resultados obtidos e as possíveis melhorias aplicáveis ao modelo em futuros trabalhos.

Capítulo 2

2 Previsão de Carga

Neste capítulo será feita uma apresentação do que é a previsão de carga, dos diferentes tipos de previsão, além das principais metodologias utilizadas na previsão de séries temporais em geral.

2.1 Descrição do Problema

Uma série temporal pode ser definida como fenômenos cujo processo observacional e consequente quantificação numérica geram uma sequência de dados distribuídos no tempo [2]. No caso do consumo de energia elétrica, a previsão pode ser dividida em quatro tipos: curtíssimo prazo, curto prazo, médio prazo e longo prazo [3]. Cada uma dessas previsões tem a sua importância e o seu objetivo.

No curtíssimo prazo, que tem como horizonte de previsão até uma hora à frente da previsão, são utilizadas como entradas apenas os valores de consumo de instantes anteriores à previsão.

Na previsão de curto prazo, o interesse é de algumas horas ou dias adiante. Nesse tipo de previsão são mais utilizados como dados de entrada as informações de carga, temperatura da área de estudo e o preço da energia.

Nas previsões de demanda de médio prazo, é analisado um horizonte de algumas semanas ou meses. Este tipo de previsão é muito utilizada para o planejamento de manutenções das redes elétricas, planejamento na compra de combustíveis e negociações de contratos.

No longo prazo, trabalha-se com previsões anuais. Este tipo de previsão serve para fornecer informações para produtores e distribuidores sobre a evolução da demanda, o que possibilita o estudo de estratégias de expansão das linhas de transmissão, das usinas geradoras ou dos vários componentes da rede elétrica. Normalmente, são utilizadas para a

previsão a longo prazo as informações de carga do sistema e também informações socioeconômicas da região de interesse.

A previsão de séries temporais está sempre associada a possíveis erros causados pela complexidade das variáveis envolvidas no processo e pelos próprios métodos utilizados. Por isso, os valores obtidos com as previsões serão aceitáveis ou não, dependendo da finalidade do estudo em particular. No caso da energia elétrica, erros menores contribuem para um melhor planejamento e operação do sistema e também um melhor uso dos recursos econômicos. Por outro lado, caso esta seja mal executada, corre-se o risco de termos esses erros amplificados. Portanto, são cada vez mais necessários os estudos de aperfeiçoamento e desenvolvimento de técnicas de previsão.

2.2 Revisão da literatura

Para realizar a previsão de uma determinada série temporal é possível utilizar diferentes métodos. Pode-se classificá-los basicamente entre métodos estatísticos e baseados em inteligência artificial.

Dentre os métodos estatísticos destacam-se [4]:

- Regressão Linear Múltipla;
- Alisamento Exponencial;
- Média Móvel Integrada Auto-Regressiva (ARIMA)

Entre os métodos baseados em inteligência artificial tem-se [4]:

- Redes Neurais Artificiais;
- Lógica Fuzzy.

Há ainda aqueles chamados de modelos híbridos, que combinam dois ou mais tipos de previsão. Segue uma pequena introdução sobre cada método citado, assim como suas vantagens e desvantagens.

2.2.1 Métodos Estatísticos de Previsão

Os métodos estatísticos de previsão podem ser subdivididos em métodos de regressão e métodos baseados em séries temporais. Os modelos de regressão avaliam a carga a ser prevista como uma combinação linear da temperatura, de valores passados da carga entre outras variáveis . São exemplos de métodos regressivos o método de regressão linear e o alisamento exponencial. Entre os métodos baseados em séries temporais destaca-se a Média Móvel Integrada Auto-Regressiva (ARIMA).

A principal vantagem dos métodos estatísticos é o fato de serem técnicas bem desenvolvidas e amplamente estudadas. Por outro lado, estes métodos apresentam dificuldades quando são expostos a variações rápidas dos fatores que influenciam a previsão, como o clima, feriados, fins de semana, etc.

2.2.1.1 Regressão Linear Múltipla

Na Regressão Linear Múltipla, modela-se a carga como função linear de múltiplas variáveis, como mostrado em (2.1), [4]:

$$y(t) = a_0 + a_1\, x_1(t) + a_2\, x_2(t) + \cdots + a_n\, x_n(t) + \mathcal{E} \qquad (2.1)$$

Sendo:

$y(t)$ – carga no instante t;

a_0, a_1, a_2, a_n – coeficientes de regressão;

$x_1(t), x_2(t), x_n(t)$ – variáveis explicativas;

$\mathcal{E}$ - erro aleatório.

Para encontrar os coeficientes de regressão pode-se utilizar o método dos mínimos quadrados e encontrar o coeficiente de cada variável explicativas na série, que é o nome

dado a cada fator que influencia a carga do sistema, como por exemplo, a temperatura do sistema.

2.2.1.2 Alisamento Exponencial

No Alisamento Exponencial, calcula-se a previsão através de (2.2), [5]:

$$F_{t+1} = \alpha\, x_t + (1 - \alpha)\, F_t \qquad (2.2)$$

Sendo:

F_{t+1} – previsão no instante t+1;

α – peso atribuído à observação mais recente, conhecido como constante de alisamento (0< α < 1);

x_t– observação no instante t (último valor observado).

Este método utiliza um valor constante para *α*, o ultimo valor observado e a previsão do instante anterior. É possível perceber que pequenos valores de *α* geram previsões mais dependentes das observações passadas. Já valores de *α* próximos de 1 geram previsões mais dependentes das observações mais recentes.

2.2.1.3 Modelo ARIMA – Abordagem de Box & Jenkis

O modelo ARIMA (Média Móvel Integrada Auto- Regressiva) [6] é uma extensão do modelo ARMA (Média Móvel Auto- Regressiva) [7], que, por sua vez, é a combinação de dois outros modelos mais simples, o modelo AR (Auto- Regressivo) e o modelo MA (Média Móvel).

2.2.1.3.1 Modelo Auto Regressivo

O modelo Auto Regressivo admite que o valor da carga y(t) é função linear das cargas nos instantes anteriores, somado a um ruído aleatório r(t). E pode ser calculado por (2.3), [4]:

$$y(t) = \alpha_1\, y(t-1) + \alpha_2\, y(t-2) + \cdots + \alpha_n\, y(t-n) + r(t) \qquad (2.3)$$

Sendo:

α – constantes da função auto regressiva;

n – numero de instantes anteriores usados na regressão.

2.2.1.3.2 Modelo da Média Móvel

O modelo da média móvel [7] admite que o valor da carga y(t) é função linear dos valores de uma série ruído branco r(t), r(t-1), ..., que é definida pelos erros de previsões atuais e passados. Um ruído branco é uma sequência de erros aleatórios que, entre outras características, pode-se considerar de média zero. A previsão é dada por (2.4), [4]:

$$y(t) = r(t) - \beta_1\, r(t-1) - \beta_2\, r(t-2) - \ldots - \beta_m\, r(t-m) \qquad (2.4)$$

Sendo:

β - constantes da função de média móvel

m – numero de instantes anteriores usados no modelo

2.2.1.3.3 Modelo ARMA

O modelo ARMA [7] combina o modelo Auto Regressivo com o modelo de Média Móvel. Portanto, neste processo, admite-se que a carga é função linear de valores passados da carga e de uma série de ruído branco, que conta com valores atuais e passados, segundo (2.5):

$$y(t) = \alpha_1\, y(t-1) + \cdots + \alpha_n\, y(t-n) - \beta_1\, r(t-1) - \ldots - \beta_m\, r(t-m) \quad (2.5)$$

Sendo:

α – constantes da função auto regressiva;

n – numero de instantes anteriores usados na regressão;

β - constantes da função de média móvel;

m – numero de instantes anteriores usados no modelo.

A ordem do modelo será dada pelo valor da carga n e ruído branco mais antigos m .

2.2.1.3.4 Modelo ARIMA

O modelo ARIMA [6] é uma adaptação do modelo ARMA para processos ditos não estacionários.. Um processo é estacionário se todas as características do comportamento do processo não são alteradas no tempo. Por exemplo, uma série pode apresentar uma tendência de crescimento. Neste caso, esta série não será estacionária.

Uma forma de lidar com processos não estacionários é remover tendências ou sazonalidades presentes na série temporal. Para remover tendências de séries temporais é comum tomar sucessivas diferenças da serie original a fim de encontrar uma série estacionária. Portanto, se tomarmos a primeira diferença da série de y(t), como em (2.6) temos [4]:

$$d(t) = y(t) - y(t-1) \qquad (2.6)$$

Se d(t) for estacionária pode-se então aplicar o método ARMA na série resultante.

2.2.2 Métodos Baseados em Inteligência Artificial

A Inteligência Artificial (IA) pode ser entendida como o conjunto de técnicas e metodologias de programação usadas para tentar resolver problemas de forma mais próxima possível de um ser humano. As principais vantagens destes métodos é o fato de que não é necessária a elaboração de modelos matemáticos complexos, já que a inteligência artificial utiliza os dados históricos de um sistema para identificar e determinar o comportamento do mesmo. Uma grande desvantagem dos métodos de inteligência artificial é a dificuldade em codificar o conhecimento e o pensamento humano. É complexo determinar quais variáveis são importantes para um sistema de previsão, por exemplo. Contudo, os métodos de IA vêm sendo largamente utilizados, pois seus resultados são geralmente superiores a outros tipos de técnicas. Dentre os métodos de Inteligência Artificial destacam-se os Sistemas Especialistas, as Redes Neurais Artificiais e a Lógica Fuzzy.

2.2.2.1 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) pode ser definida como um sistema constituído por elementos de processamento interconectados, chamados neurônios, os quais estão dispostos em camadas e são responsáveis pela não linearidade e pela memória da rede [8].

Elas possuem técnicas de aprendizado semelhantes às do cérebro humano e utilizam estas técnicas para a resolução de problemas.

Entre as aplicações das RNAs, destacam-se: reconhecimento e classificações de padrões, *clustering* ou agrupamento, previsão de séries temporais, aproximação de funções, otimização, processamento de sinais, analise de imagens e controle de processos [9].

Para realizar a previsão de carga é necessário treinar a rede através de um conjunto de dados de entrada e de saída. Os dados de entrada são os valores de carga que são conhecidos e os dados de saída são os valores de carga que se deseja prever. A partir da rede treinada com os valores de carga passados é possível prever valores futuros.

2.2.2.2 Lógica Fuzzy

O conceito da Lógica Fuzzy está na possibilidade de uma variável assumir qualquer valor entre 0 e 1. Diferentemente da lógica convencional, a Lógica Fuzzy não apresenta apenas valores absolutos, o que permite ao sistema incorporar incertezas e imperfeições do mundo real.

No início do seu desenvolvimento, a teoria da Lógica Fuzzy foi empregada na área de controle. Contudo, houve um aumento significativo da utilização de sistemas fuzzy em outras áreas como a mineração de dados, planejamento, otimização e previsão de séries temporais [3].É uma teoria bastante utilizada na previsão de carga devido a sua capacidade de aproximar qualquer função não linear com uma exatidão arbitrária e de encontrar padrões em grandes conjuntos de dados [4].

Capítulo 3

3 Modelos Inteligentes

Neste capítulo serão apresentadas, de forma mais aprofundada, as teorias de Redes Neurais Artificiais e Lógica Fuzzy. Inicialmente, tais técnicas são tratadas separadamente e, em seguida, um modelo híbrido, denominado Neuro-Fuzzy, é descrito.

3.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são modelos matemáticos inspirados nos princípios de funcionamento dos neurônios biológicos e na estrutura do cérebro [10]. Ou seja, são técnicas computacionais que se assemelham a estrutura neural dos organismos inteligentes e adquirem conhecimento através da experiência [11]. Este conhecimento é armazenado em pesos sinápticos que interligam os neurônios. A Figura 3.1 apresenta um diagrama esquemático de um neurônio artificial proposto por McCulloch e Pitts, na década de 40.

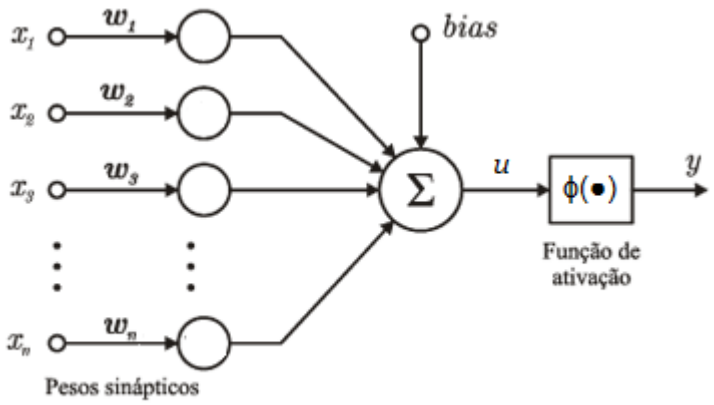

Figura 3.1 – Neurônio Artificial.

No diagrama da Figura 3.1, pode-se identificar os principais componentes de um neurônio artificial:

- As sinapses ou elos de conexão, que são caracterizadas por um peso próprio. Basicamente, um sinal x_i na entrada é multiplicado pelo seu respectivo peso w_i;
- Um somador que serve para adicionar os sinais de entrada multiplicados pelos seus respectivos pesos;
- Uma função de ativação que pode limitar a amplitude de saída do neurônio, restringindo o intervalo de amplitude do sinal de saída a um valor finito.

O neurônio artificial apresentado na Figura 3.1 possui uma saída y que pode ser escrita segundo (3.1):

$$y = \phi(\sum_{i=1}^{n} x_i w_i + b) \qquad (3.1)$$

Sendo:

x_i – sinais de entrada do neurônio;

w_i – pesos sinápticos do neurônio;

b – *bias*;

ϕ(●) – função de ativação;

y – sinal de saída do neurônio.

O *bias* é um parâmetro externo que serve basicamente para aumentar o grau de liberdade dos ajustes dos pesos. No entanto, ele pode ser incorporado ao neurônio como se fosse uma das entradas do sistema. Isto é possível, adaptando (3.1) para (3.2):

$$y = \phi(\sum_{i=0}^{n} x_i w_i) \qquad (3.2)$$

Fixando x_0=1 e w_0 = b, pode-se reformular o modelo do neurônio artificial para o da Figura 3.2:

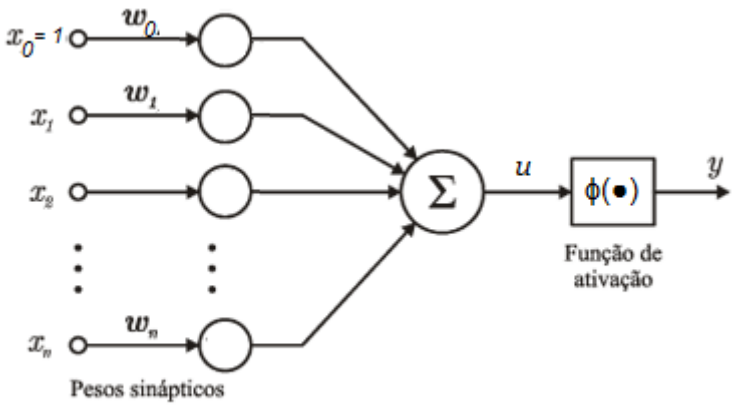

Figura 3.2 – Neurônio Artificial com Bias incorporado.

A função de ativação pode apresentar características lineares ou não lineares, de forma a determinar a saída de um neurônio a partir do seu potencial de ativação [12]. Dentre os tipos de função de ativação pode-se destacar a função linear, a função degrau, a função sigmoide e a função gaussiana. Estes quatros tipos de função de ativação são vistos na Figura 3.3:

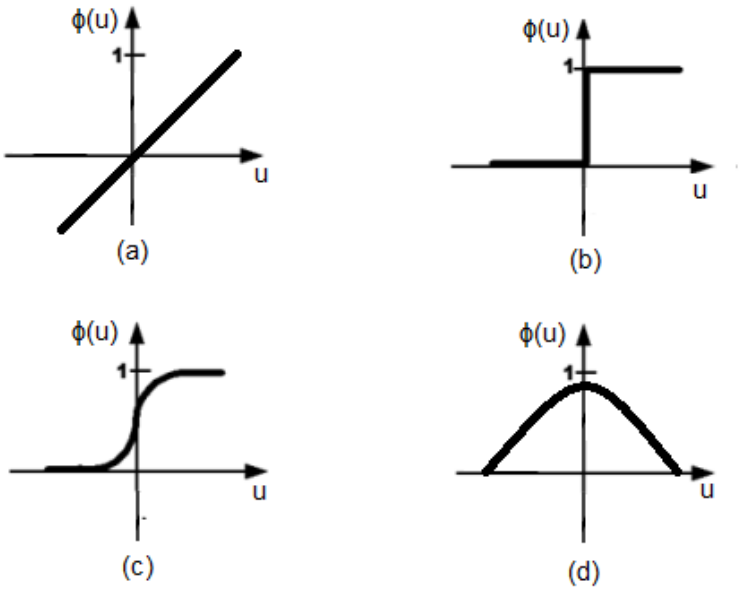

Figura 3.3 – Funções de ativação : (a) Função Linear. (b) Função Degrau. (c) Função Sigmoidal. (d) Função Gaussiana

As redes neurais são, graças a estas várias funções de ativação, capazes de lidar com problemas lineares ou não lineares. As RNAs também possuem grande capacidade de

improvisação e adaptabilidade, isto é, elas são capazes de resolver problemas que não fazem parte da base de conhecimento fornecida durante o seu treinamento.

As grandes dificuldades do uso das redes neurais são a escolha da topologia da rede, que não segue uma teoria exata e acontece de forma empírica, e a dificuldade na interpretação dos conceitos aprendidos pelas RNAs, pois estes são codificados em seus pesos formando uma espécie de caixa preta.

Basicamente, existem dois tipos de estruturas de redes neurais, as redes do tipo *feedforward e* as redes do tipo *feedback*. Nas redes *feedforward* o fluxo de sinal acontece da camada de entrada para a camada de saída em sentido único, isto é, cada neurônio recebe sinais apenas das camadas anteriores. Já nas redes *feedback* existe a conexão das saídas de uma camada para a entrada da mesma camada ou de camadas anteriores, ou seja, existe um tipo realimentação na rede e por isso, são ditas redes com memória. Também é possível diferenciar as RNAs quanto à maneira com que elas adquirem conhecimento: redes supervisionadas e redes não supervisionadas.

3.1.1 Tipos de Treinamento

O treinamento de redes neurais artificiais é a maneira pela qual uma rede adquire conhecimento. Este treinamento pode ser realizado de várias maneiras diferentes. Os algoritmos de aprendizagem diferem um dos outros pela forma em que os pesos sinápticos são modificados. Contudo, eles podem ser classificados em dois grupos distintos:

- Aprendizado supervisionado – A rede é treinada utilizando pares de treinamento que contém vetores com valores de entrada e saída. Um vetor de entrada é inserido à rede e a saída correspondente é calculada e comparada com o respectivo sinal de saída desejado, gerando um sinal de erro. O objetivo do algoritmo de treinamento consiste em ajustar os pesos sinápticos de tal forma a minimizar este erro entre a saída gerada pela rede e a saída que se deseja;
- Aprendizado não supervisionado – É um tipo de aprendizado onde não existe um vetor de saída desejado. São utilizados vetores de entrada para que a rede identifique padrões e agrupe esses padrões em classes(Cluster Analysis).

3.1.2 Redes Neurais Supervisionadas

3.1.2.1 O Perceptron

O modelo de neurônio *Perceptron* é o modelo mais simples de Rede Neural, no qual várias unidades de processamento estão conectadas unicamente a uma unidade de saída, através dos pesos sinápticos [8]. O perceptron é composto pelo modelo de neurônio proposto por McCulloch e Pitts; ou seja, possui uma camada de entrada e uma camada de saída, estando cada entrada ligada a saída através da conexão representada pelos pesos sinápticos. A saída utiliza como função de ativação a função degrau, que define quando o neurônio ligado (emite sinal 1) ou desligado (emite sinal 0). Portando, pode-se definir a função de ativação do perceptron segundo (3.3):

$$\phi = \begin{cases} 1, se\ u \geq 0 \\ 0, se\ u < 0 \end{cases} \qquad (3.3)$$

Sendo:

$$u = \sum_{i=0}^{n} x_i w_i \qquad (3.4)$$

O perceptron será utilizado, a partir de agora, para introduzir os conceitos básicos de funcionamento de uma RNA, que são o processo de aprendizagem da rede e o cálculo de saída da rede.

3.1.2.1.1 Cálculo da Saída de um Perceptron

A saída y de um perceptron é calculada através do sinal u e da função de ativação ϕ. Por exemplo, podemos admitir que um perceptron possua entradas x_1= 1 e x_2 = -1, e

pesos sinápticos w_0= 1, w_1 =2, w_2 =-2. A saída y desta rede neural será calculada da seguinte maneira:

$$u = x_0 w_0 + x_1 w_1 + x_2 w_2$$

$$u = 1.(1) + 1.(2) + -1.(-2) = 5$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 1$$

3.1.2.1.2 Aprendizagem de um Perceptron

O processo de aprendizagem do perceptron é feito através do que chamamos de treinamento da rede. Esse treinamento consiste em ajustar os pesos através de uma regra de aprendizagem até que a rede neural forneça respostas satisfatórias ao problema proposto. Para realizar o treinamento é necessário um conjunto de entradas associadas ao seu desejado padrão de saída. Este tipo de treinamento é chamado de supervisionado e obedece a uma regra de aprendizagem dada por (3.5)

$$w_i\,(novo) = w_i\,(antigo) + \alpha\,(d - y)\,.x_i \qquad (3.5)$$

Sendo:

w_i = pesos sinápticos;

α = taxa de aprendizagem(valor entre 0 e 1);

d = valor desejado para a saída;

y = valor calculado para a saída;

$(d - y)$ = erro apresentado na saída, representado por e_i

A partir das regras de aprendizado e do conjunto de entradas e saídas desejadas, pode-se estabelecer um algoritmo para o treinamento do perceptron [8]:

1. Inicialização dos pesos sinápticos w_i - são geralmente inicializados aleatoriamente com valores pequenos;
2. O conjunto de entradas e saídas é apresentado à rede;
3. Estipula-se um valor para a taxa de aprendizagem e calcula-se o sinal u e o sinal de saída y;
4. Calcula-se o erro e_i;
5. Ajustam-se os pesos sinápticos a partir do erro e_i;
6. Realiza-se o teste dos critérios de parada, que podem ser um erro satisfatório ou um número máximo de ciclos de treinamento. Caso os critérios de parada não sejam atendidos, repete-se os passos 2 ao 6 até que os resultados sejam satisfatórios.

O algoritmo de treinamento da rede pode ser mais facilmente compreendido com o auxílio do Exemplo 3.1, adaptado de [8].

Exemplo 3.1 – Deseja-se treinar uma rede neural do tipo perceptron de duas entradas de forma que a saída obedeça a Tabela 3.1:

Tabela 3.1 - Exemplos para treinamento do Perceptron

ENTRADA 1	ENTRADA 2	SAÍDA
0	0	0
0	1	0
1	0	0
1	1	1

Para resolver este exemplo, basta seguir o algoritmo de treinamento já apresentado, ou seja:

- Inicialização aleatória dos pesos sinápticos:

  $w_0 = 0;$

  $w_1 = 3;$

  $w_2 = 3;$

- Apresentação à rede do primeiro exemplo entrada/saída, mostrado Tabela 3.2:

Tabela 3.2 – Exemplo de treinamento 1

ENTRADA 1	ENTRADA 2	SAÍDA
0	0	0

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(0) + \; 0\,.(3) + \; 0\,.(3) = 0$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 1$$

Como o valor de saída do perceptron não é o desejado para este exemplo, os pesos sinápticos devem ser ajustados:

Definindo a taxa de aprendizagem α = 1, temos:

$$w_0\,(novo) = w_0\,(antigo) + \; \alpha\,(\,d - y)\,.x_0 = \; 0 + 1.(0 - 1).1 = -1$$

$$w_1\,(novo) = w_1\,(antigo) + \; \alpha\,(\,d - y)\,.x_1 = \; 3 + 1.(0 - 1).0 = 3$$

$$w_2\,(novo) = w_2\,(antigo) + \; \alpha\,(\,d - y)\,.x_2 = \; 3 + 1.(0 - 1).0 = 3$$

- Apresentação à rede do segundo exemplo entrada/saída, como mostrado na Tabela 3.3:

Tabela 3.3 – Exemplo de Treinamento 2

ENTRADA 1	ENTRADA 2	SAÍDA
0	1	0

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(-1) + \; 0\,.(3) + \; 1\,.(3) = 0$$

Aplicando a função de ativação degrau, temos:

$$y = \; \phi(u) = 1$$

Como o valor de saída do perceptron não é o desejado para este exemplo, os pesos sinápticos devem ser ajustados:

Definindo a taxa de aprendizagem α = 1, temos:

$$w_0\,(novo) = \; w_0\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_0 = \; -1 + 1.(0 - 1).1 = \; -2$$

$$w_1\,(novo) = \; w_1\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_1 = \; 3 + 1.(0 - 1).0 = \; 3$$

$$w_2\,(novo) = \; w_2\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_2 = \; 3 + 1.(0 - 1).1 = \; 2$$

- Apresentação à rede do terceiro exemplo entrada/saída, como mostrado na Tabela 3.4:

Tabela 3.4 – Exemplo de Treinamento 3

ENTRADA 1	ENTRADA 2	SAÍDA
1	0	0

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(-2) + \; 1\,.(3) + \; 0\,.(2) =$$

Aplicando a função de ativação degrau, temos:

$$y = \; \phi(u) = 1$$

Como o valor de saída do perceptron não é o desejado para este exemplo, os pesos sinápticos devem ser ajustados:

Definindo a taxa de aprendizagem α = 1, temos:

$$w_0\,(novo) = \; w_0\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_0 = \; -2 + 1.(0 - 1).1 = \; -3$$

$$w_1\,(novo) = \; w_1\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_1 = \; 3 + 1.(0 - 1).1 = \; 2$$

$$w_2\,(novo) = \; w_2\,(antigo) + \; \alpha\,(\,d - \; y)\,.x_2 = \; 2 + 1.(0 - 1).0 = \; 2$$

- Apresentação à rede do quarto exemplo entrada/saída, como mostrado na Tabela 3.5:

Tabela 3.5 – Exemplo de Treinamento 4

ENTRADA 1	ENTRADA 2	SAÍDA
1	1	1

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(-3) + \; 1\,.(2) + \; 1\,.(2) = 1$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 1$$

Como o valor de saída do perceptron é o desejado para este exemplo, os pesos sinápticos estão corretos.

Após a apresentação de todo o conjunto de exemplos à rede, é necessário realizar um novo ciclo para verificar se o perceptron responde corretamente aos primeiros exemplos:

- Verificação das saídas do Perceptron

Exemplo 1,Tabela 3.6 :

Tabela 3.6 – Verificação do Exemplo de Treinamento 1

ENTRADA 1	ENTRADA 2	SAÍDA
0	0	0

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(-3) + \; 0\,.(2) + \; 0\,.(2) = -3$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 0$$

Logo, a saída responde corretamente ao primeiro exemplo.

Exemplo 2, Tabela 3.7:

Tabela 3.7 - Verificação do Exemplo de Treinamento 2

ENTRADA 1	ENTRADA 2	SAÍDA
0	1	0

$$u = x_0\, w_0 + \; x_1\, w_1 + \; x_2\, w_2$$

$$u = 1\,.(-3) + \ \ 0\,.(2) + 1\,.(2) = -1$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 0$$

Logo, a saída responde corretamente ao segundo exemplo.

Exemplo 3, Tabela 3.8:

Tabela 3.8 - Verificação do Exemplo de Treinamento 3

ENTRADA 1	ENTRADA 2	SAÍDA
1	0	0

$$u = x_0\, w_0 + \ x_1\, w_1 + \ x_2\, w_2$$

$$u = 1\,.(-3) + \ \ 1\,.(2) + 0\,.(2) = -1$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 0$$

Logo, a saída responde corretamente ao terceiro exemplo.

Exemplo 4,Tabela 3.9:

Tabela 3.9 - Verificação do Exemplo de Treinamento 4

ENTRADA 1	ENTRADA 2	SAÍDA
1	1	1

$$u = x_0\, w_0 + \ x_1\, w_1 + \ x_2\, w_2$$

$$u = 1\,.(-3) + \;\; 1\,.(2) + 1\,.(2) = 1$$

Aplicando a função de ativação degrau, temos:

$$y = \phi(u) = 1$$

Logo, a saída responde corretamente ao quarto exemplo e, portanto, a todo o conjunto de entradas e saídas desejadas. Quando isto ocorre, dizemos que a rede está treinada.

3.1.2.2 O Perceptron de Várias Camadas (Multilayer Perceptron - MLP)

Como visto anteriormente, o perceptron é o tipo mais básico de rede neural existente. Contudo, existem vários tipos de configurações utilizadas para a construção de outras redes neurais. Geralmente, as redes neurais possuem apenas uma camada de entrada e uma camada de saída, porém entre as duas podem existir um número indeterminado de camadas intermediárias. Pode-se ainda variar o número de neurônios em cada uma das camadas da RNA, possibilitando que cada saída de neurônio seja uma entrada para os neurônios de outras camadas. Um exemplo de rede neural com duas camadas intermediária e vários neurônios pode ser visualizado na Figura 3.4:

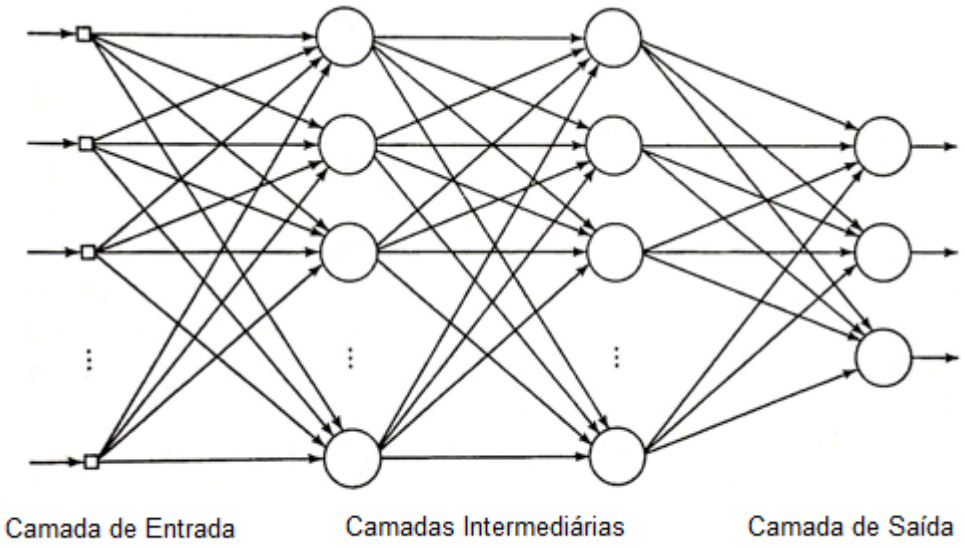

Figura 3.4 – Rede Neural Multicamadas

Um tipo muito comum de rede neural com várias camadas é o MLP (*Multilayer Perceptron*), que possui as seguintes características principais [8]:

- Uma camada de entrada – nela cada neurônio representa uma variável considerada como entrada para o problema;
- Pelo menos uma camada intermediária – os neurônios desta camada possuem geralmente função de ativação sigmoidal (logística ou tangente hiperbólica);
- Uma camada de saída – é responsável pela resposta da rede, determinado uma ou mais variáveis que se deseja prever.

Para o treinamento de uma rede MLP o algoritmo mais utilizado é a retropropagação de erro ou *error backpropagation*. Neste processo, o ajuste dos pesos se dá através de duas fases: fase *forward* e *backward*. Na fase *forward* o sinal é propagado a partir da camada de entrada em direção à camada de saída, gerando um sinal de saída que é comparado com o sinal é desejado. Calcula-se então o erro da saída. Na fase *backward* o erro da camada de saída é propagado no sentido contrário, isto é, em direção à camada de entrada. Isto permite que os pesos de todas as camadas intermediárias e da camada de entrada sejam ajustados. Este processo é visualizado na Figura 3.5.

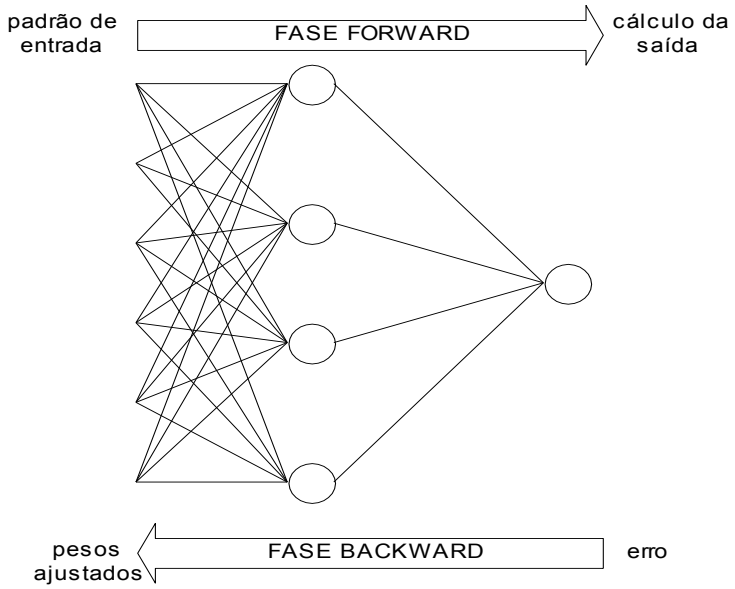

Figura 3.5 - Fases *Forward* e *Backward* do algoritmo de retropropagação do erro

3.1.3 Redes Neurais Não-Supervisionadas

Nos modelos de redes neurais não supervisionados, não há comparação da saída gerada pela rede com uma saída desejada. Este tipo de rede tem como objetivo identificar padrões e características dos dados de entrada e organizá-los em uma saída topologicamente ordenada. Uma grande aplicação para este método é o agrupamento ou *clustering*, que consiste em agrupar dados de características semelhantes. Um dos modelos de rede neural capaz de realizar tarefas deste tipo são os chamados mapas auto organizáveis ou mapas de Kohonen.

3.1.3.1 Mapas de Kohonen

Uma rede neural de Kohonen [13] é uma estrutura onde os neurônios de agrupam em uma rede ou reticulado, normalmente unidimensional ou bidimensional, como mostrado na Figura 3.6.

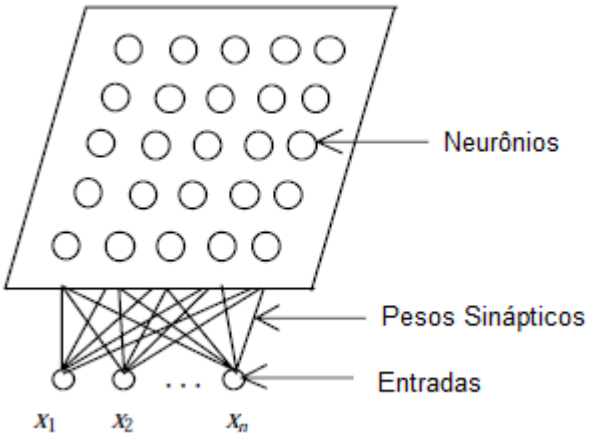

Figura 3.6 – Mapa de Kohonen Bidimensional

Os mapas de Kohonen fazem parte de um conjunto de redes neurais que são baseadas em modelos de competição. Este nome ocorre devido ao fato dos neurônios competirem entre si para saber quem gera o maior valor na saída, ou seja, qual o vetor de pesos mais semelhante com o padrão de entrada. Esta medida é geralmente calculada pela distancia euclidiana entre os exemplos de entrada e os pesos do neurônios [8] e é dada por (3.6):

$$d_{xy} = \sqrt{\sum_{j=1}^{n}(x_j - w_{ij})^2} \qquad (3.6)$$

Sendo:

d_{xy} – distância euclidiana;

x_j - exemplos de entrada;

w_{ij} – pesos sinápticos que ligam as entradas aos neurônios da grade de saída;

n – numero de entradas;

Durante o processo de treinamento da rede, dado um exemplo de entrada, o neurônio vencedor será aquele que apresentar a menor distância euclidiana. Uma vez encontrado o neurônio vencedor, a atualização dos pesos sinápticos pode se acontecer de duas maneiras: modelo “vencedor leva tudo” ou modelo “vencedor leva parte”.

No modelo “vencedor leva tudo” apenas os pesos sinápticos do neurônio vencedor serão atualizados. A regra de atualização do neurônio é dada por (3.7):

$$w_{ij}(novo) = w_{ij}(antigo) + \alpha \left(x_j - w_{ij}(antigo)\right) \qquad (3.7)$$

Sendo:

w_{ij} – pesos sinápticos que ligam as entradas ao neurônio vencedor;

x_j – exemplos de entrada;

α – taxa de aprendizagem;

No modelo “vencedor leva parte”, além do neurônio vencedor, os neurônios vizinhos a ele também tem seus pesos sinápticos ajustados. Neste caso, é necessário estabelecer um critério para definir o limite desta vizinhança. Uma opção é definir este limite em função da distancia de um neurônio até o neurônio vencedor através de uma função gaussiana, por exemplo.

No algoritmo de treinamento é necessário reduzir a região de vizinhança à medida que aumentam o número de iterações, como observado na Figura 3.7. Com isso, espera-se obter uma organização do mapa topológico, ou seja, os padrões detectados por um determinado neurônio estarão relacionados com a posição do neurônio na grade. Há ainda

uma fase de ajuste fino do mapa topológico, onde utiliza-se uma taxa de aprendizagem baixa e uma região de vizinhança pequena.

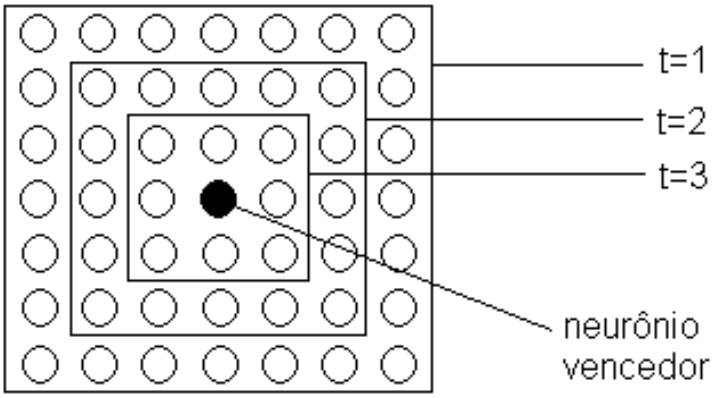

Figura 3.7 – Redução gradual da região de vizinhança desde o instante t=1 até o instante t=3

3.2 Lógica Fuzzy

A Lógica Fuzzy [14] ou lógica nebulosa é uma teoria desenvolvida com a finalidade de permitir o tratamento matemático de situações de imprecisões e incertezas. Na Lógica Fuzzy o grau de verdade de uma declaração é representado por um número em um intervalo de 0 a 1, diferentemente do que ocorre com a lógica clássica onde o grau de verdade pode assumir apenas dois valores distintos: 0 (declaração totalmente falsa) ou 1 (declaração totalmente verdadeira). Esta propriedade da Lógica Fuzzy é muito útil em situações práticas onde a intensidade de um fenômeno é descrito por meio de variáveis linguísticas como baixo, alto, médio ou normal; e ainda conceitos como “quase verdadeiro”. Um exemplo de classificação de uma variável fuzzy é mostrado na Figura 3.8.

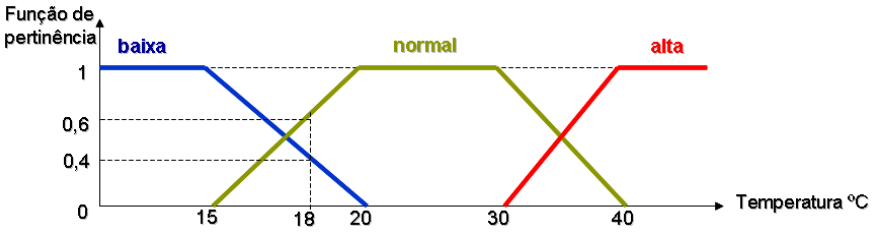

Figura 3.8 - Conjuntos fuzzy e funções de pertinência para a variável temperatura

Na Figura 3.8 a variável temperatura foi modelada arbitrariamente por três conjuntos fuzzy: temperatura baixa, temperatura normal e temperatura alta. Estes três conceitos não são delimitados de forma precisa, pois estão relacionados com a sensação térmica. Esta característica de imprecisão é representada pela sobreposição entre os conjuntos fuzzy. Por exemplo, a temperatura de 18 ºC é baixa, normal ou alta? A resposta desta pergunta é feita atribuindo graus de pertinência a cada um destes conjuntos. A variável temperatura pertence ao conjunto baixa com grau de 0,4; ao conjunto normal com grau 0,6 e ao conjunto alta com grau de 0.

No exemplo da Figura 3.8 foi utilizada a função trapezoidal para relacionar a variável temperatura ao seu conjunto fuzzy e ao seu grau de pertinência. Porém, outras funções podem ser utilizadas, entre elas tem-se a função gaussiana e a função triangular. Essas funções recebem o nome de funções de pertinência. Na Figura 3.9 são mostrados exemplos das funções de pertinência.

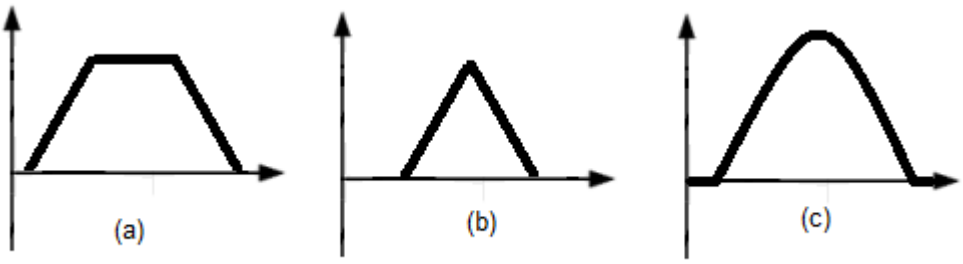

Figura 3.9 – Funções de Pertinência. (a) Função Trapezoidal. (b) Função Triangular. (c) Função Gaussiana.

A flexibilidade de lidar com problemas que envolvem incertezas, raciocínios aproximados e imprecisões é a principal vantagem do uso da Lógica Fuzzy. Um sistema fuzzy é capaz de utilizar o conhecimento já existente armazenando as informações em sua base de regras, porém nem sempre é simples definir as funções de pertinência e as regras de ativação. Os sistemas fuzzy também não possuem capacidade de aprendizagem, uma

vez que são sistemas estáticos, ou seja, não se adaptam a variações muito dinâmicas dos problemas.

3.2.1 Sistema de Inferência Fuzzy (FIS)

O sistema de inferência da Lógica Fuzzy está associado a regras do tipo: *SE antecedente ENTÃO consequência.* Por exemplo, *SE temperatura é alta ENTÃO carga é alta*. Durante o processo, o sistema executa as seguintes etapas [15]:

- "Fuzzifycação" – transformação das variáveis do problema em valores fuzzy através das funções de pertinência;
- Cálculo dos antecedentes das regras;
- Cálculo das consequências das regras, isto é, calcula a resposta e uma determinada regra a um conjunto de entradas;
- Combinação das saídas de cada regra para obter-se um resultado global;
- "Defuzzyficação" – converte a saída fuzzy em um valor preciso.

3.2.2 Método de Wang e Mendel

A partir do conceito da Lógica Fuzzy, Wang e Mendel [16] propõe um algoritmo para a extração de regras fuzzy a partir de dados numéricos para, entre outros objetivos, previsão de séries temporais. O método sugere um algoritmo de 5 passos para a geração das regras fuzzy [16]:

1. Dividir os dados de entrada e saída em conjuntos fuzzy. Os conjuntos devem ser sobrepostos e cada conjunto será associado a uma função de pertinência, que tem a finalidade de mapear os vetores de entrada e definir os pares de entrada (janela de previsão) e saída(alvo).
2. Construir pares de entrada e saída e gerar as regras fuzzy. Considerar uma janela de entrada formada por valores passados e uma janela de saída com valores futuros. Por exemplo, uma janela de entrada de tamanho 4(t-3, t-2, t-1, t) e uma janela de saída de tamanho 1(previsão de um passo à frente, t+1).

3. Atribuir um grau de pertinência a cada regra: O grau será definido pelo produto das pertinências de todos os antecedentes e dos respectivos consequentes.
4. Construir uma base de regras armazenando todas as regras geradas.

   Prever valores futuros (deffuzyficação): Para a previsão de valores futuros, os dados da janela de entrada são adquiridos e a base de regras é aplicada. O cálculo do valor previsto se dá pela combinação das saídas das regras ativadas e utilizando o método do centro de gravidade, utilizando (3.8):

$$y(t+1) = \frac{\sum_{i=1}^{k}[\mu(antecedentes_i)\,.y_i^c]}{\sum_{i=1}^{k}\mu(antecedentes_i)} \qquad (3.8)$$

Sendo:

$\mu(antecedentes_i)$ - produto das pertinências das variáveis que formam o antecedente regra i;

k – número de regras ativadas pelo padrão de entrada;

y_i^c – centro de gravidade do conjunto fuzzy do consequente ativado pela regra i.

Como um exemplo da aplicação deste algoritmo para a criação de uma regra, pode-se observar a série temporal genérica da Figura 3.10, que apresenta os valores [10;25;46;28;24;36;37;43;50;26;20;15;30].

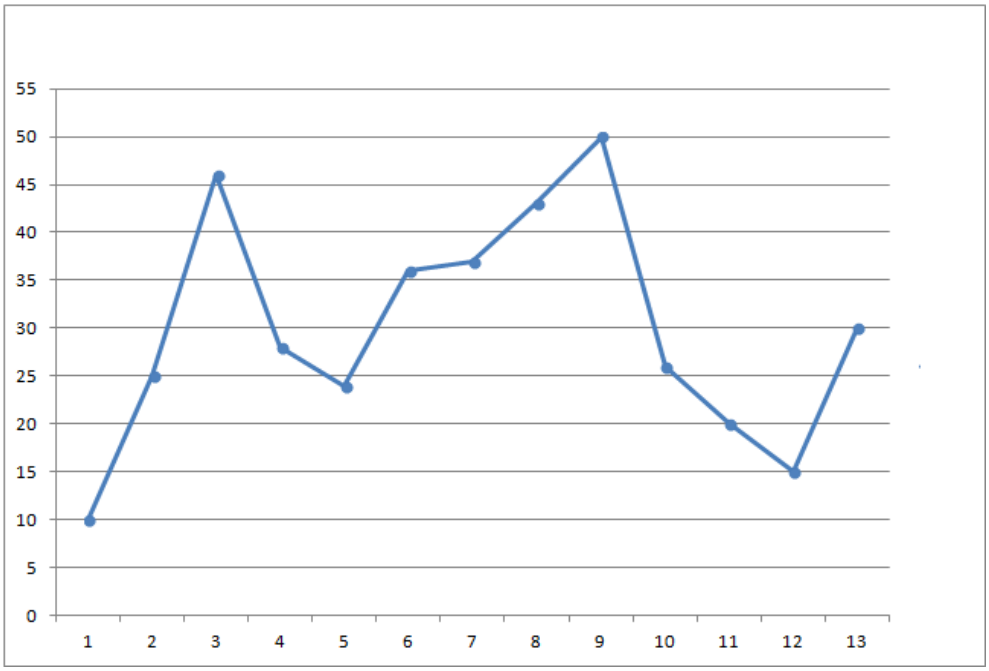

Figura 3.10 – Série Temporal Genérica

Como primeiro passo, define-se os conjuntos fuzzy sobrepostos A, B, C, D e E. O conjunto A foi definido no intervalo [40,60], o conjunto B no intervalo [30,50]; o conjunto C no intervalo [20,40]; o conjunto D no intervalo [10,30] e o conjunto E no intervalo [0,20]. A função de pertinência escolhida para este exemplo foi a triangular. Desta forma, cada variável pode ativar até dois conjuntos fuzzy. Como mostrado na Figura 3.11, o conjunto A é ativado pela variável no instante t5; o conjunto B é ativado pela variável nos instantes t3, t4 e t5; o conjunto C é ativado pela variável nos instantes t1, t2, t3 e t4 e o conjunto D é ativado pela variável nos instantes t1 e t2, e apresentam os graus de pertinência exibidos na Tabela 3.10:

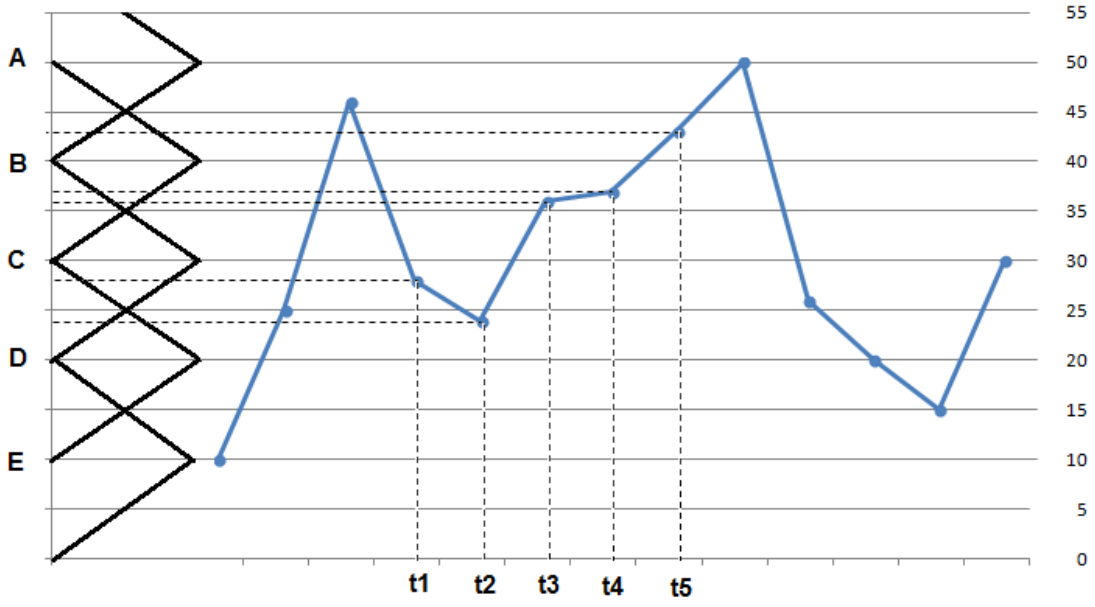

Figura 3.11 – Ativação dos Conjuntos Fuzzy

Tabela 3.10 - Grau de Pertinência de cada Conjunto

t1	C=0,8	D=0,2
t2	C=0,4	D=0,6
t3	A=0,4	B=0,6
t4	A=0,3	B=0,7
t5	A=0,2	B=0,8

Para a definição do conjunto fuzzy de cada instante, seleciona-se o maior grau de ativação em cada um deles. Portanto, a sequência que forma a conjunto antecedente é C,D,B,B e o conjunto consequente é definido por B.

Como já visto, a regra fuzzy será gerada da forma “SE antecedente ENTÃO consequente”. Logo, a regra gerada para este exemplo é “SE t1=C e t2=D e t3=B e t4=B ENTÃO t5=B”. O grau de pertinência desta regra será dado pelo produto das pertinências de todos os antecedentes e consequentes. Neste exemplo, teremos o grau de pertiência exibido na Tabela 3.11:

Tabela 3.11 – Grau de Pertinência da regra

t1	t2	t3	t4	t5	Grau de Pertinência
0,8	0,6	0,6	0,7	0,8	0,16128

Para a previsão de um valor futuro é necessário a geração de várias outras regras, que são criadas movendo a janela de antecedentes/consequentes por toda a série temporal.

3.2.3 Cálculo das Pertinências utilizando Fuzzy C-Means

Uma outra forma de calcular as pertinências fuzzy e gerar regras para a criação de um sistema de inferência fuzzy é utilizando o método Fuzzy C-Means [17]. Neste método, os conjuntos fuzzy não são definidos arbitrariamente, mas sim calculados através das distâncias relativas entre as curvas diárias da série temporal. Este cálculo se dá através de (3.9)e (3.10):

$$dxy = \sqrt{\sum_{i=1}^{k}(x(i) - y(i))^2} \qquad (3.9)$$

Sendo:

d_{xy} - distância da curva x em relação à curva y;

k - número de pontos das curvas x e y;

$$\mu_{xy} = \frac{1}{\sum_{i=1}^{n}\left(\frac{dxy}{dxi}\right)^{\frac{2}{m-1}}} \qquad (3.10)$$

Sendo:

μ_{xy} - Pertinência da curva x em relação à curva y;

d_{xy} - distância da curva x em relação à curva y;

d_{xi} - distância da curva x em relação à curva i;

n - número total de curvas;

m - constante de fuzzyficação. Responsável por regular a sensibilidade da pertinência em relação às distâncias.

A aplicação deste modelo pode ser visualizada no exemplo que utiliza as curvas exibidas na Figura 3.12. Neste exemplo, calcula-se a pertinência da curva d em relação à curva b utilizando (3.9) e (3.10), o que possibilita quantificar o quanto a curva d assemelha-se à curva b.

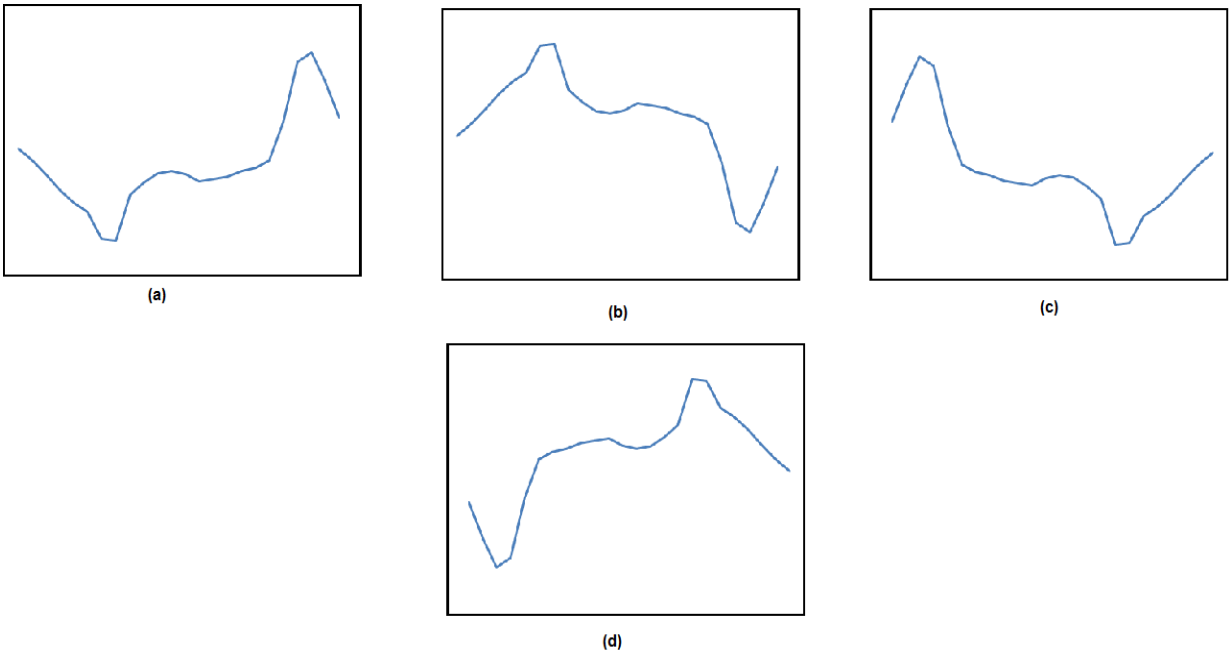

Figura 3.12 – Curvas diárias (a), (b), (c) e (d).

As distâncias d_{ad}, d_{bd} e d_{cd} são calculadas segundo (3.9) e a pertinência μ_{bd} é calculada por (3.11):

$$\mu_{bd} = \frac{1}{\left(\frac{d_{bd}}{d_{ad}} + \frac{d_{bd}}{d_{cd}} + \frac{d_{bd}}{d_{cd}}\right)^{\frac{2}{m-1}}} \qquad (3.11)$$

Desta forma, pode-se calcular as pertinências de cada antecedente e consequente para gerar a base de regras fuzzy e depois realizar a defuzzyficação e obter a previsão da variável desejada através de $y(t+1) = \frac{\sum_{i=1}^{k}[\mu(antecedentes_i)\,.y_i^c]}{\sum_{i=1}^{k}\mu(antecedentes_i)}$).

3.3 Modelos Híbridos Neuro-Fuzzy

Os modelos híbridos são aqueles que combinam dois ou mais tipos de métodos de previsão.. Entre as vantagens de se utilizar mais de um modelo em um sistema de previsão pode–se destacar [18]:

- A integração de técnicas complementares possibilita que uma técnica complete as deficiências das outras e com isso obtenha-se um melhor desempenho;
- Em alguns casos não é possível resolver alguns problemas complexos utilizando apenas uma única técnica. O problema então é dividido em problemas menores, os quais são tratados de forma individual, de acordo com a técnica mais adequada para cada tipo de problema.

Nesse contexto, uma rede Neuro-Fuzzy é uma associação dos modelos de Redes Neurais Artificiais com os modelos que utilizam a Lógica Fuzzy. Como resultado desta combinação tem-se um modelo que potencializa as características destes dois sistemas. Com a utilização destas duas técnicas, pode-se beneficiar com aproveitamento do conhecimento prévio do problema e do estilo de raciocínio humano dos sistemas fuzzy e a capacidade de generalização e aprendizagem das redes neurais.

No capítulo 4, será proposto um modelo Neuro-Fuzzy para a previsão de séries temporais de carga em sistemas de energia elétrica.

Capítulo 4

4 Modelo Proposto para a Previsão de Carga

Este capítulo apresenta o modelo proposto para a previsão de séries temporais de carga em sistemas de energia elétrica. Serão apresentados a estrutura do modelo, os índices de desempenho utilizados para avaliar a qualidade da previsão e as formas de ajuste do modelo.

As etapas do processo de previsão podem ser visualizados no fluxograma da Figura 4.1.

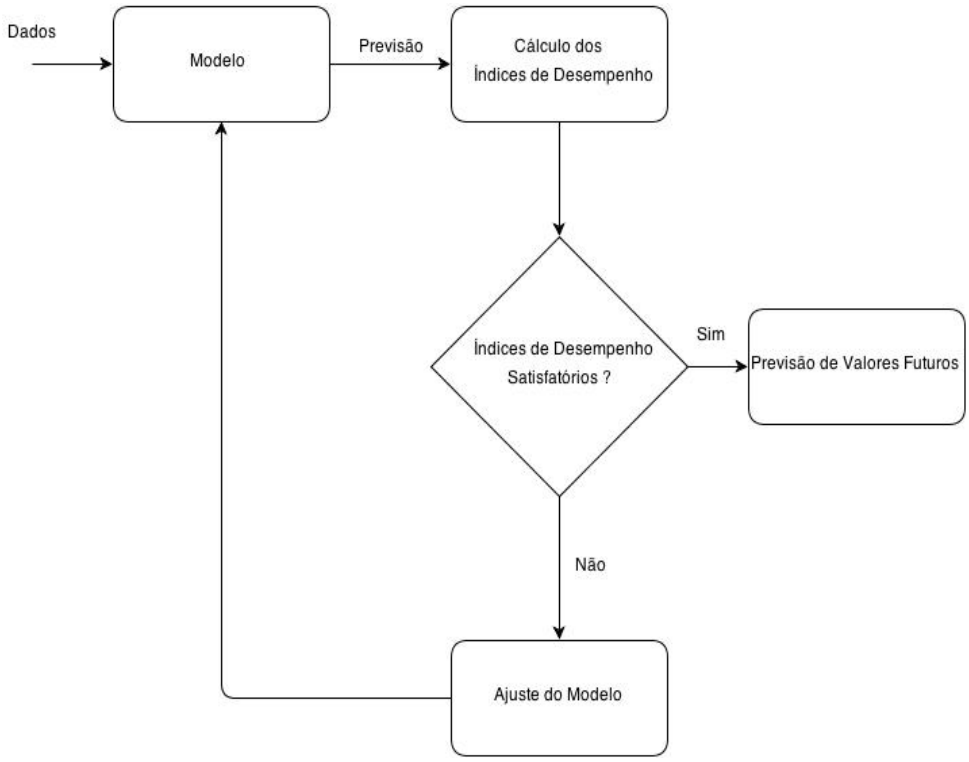

Figura 4.1 - Diagrama das etapas da previsão

Os dados de entrada do modelo são séries de carga históricas obtidas através de medições realizadas no sistema elétrico. O modelo processa os dados e gera, a partir dos parâmetros selecionados, uma série de carga estimada. A série de carga gerada pelo modelo é então comparada com a série histórica observada através de índices de

desempenho. Caso os índices de desempenho sejam considerados satisfatórios, o modelo está pronto para gerar previsões de séries de carga futuras. Caso contrário, os parâmetros utilizados são modificados na fase de ajuste do modelo, uma nova estimativa é gerada e novamente comparada com série histórica observada. Este processo se repete até que os índices de desempenho apresentem valores dentro de uma margem satisfatória. Também pode-se determinar um número máximo de tentativas de ajuste dos parâmetros para o caso do modelo não apresente valores aceitáveis.As seções deste capítulo analisarão separadamente cada etapa do processo de previsão.

4.1 Dados de Entrada

A série temporal utilizada como entrada de dados do modelo deve ter um formato de matriz de dimensão 24 x n, sendo n o número de dias contidos na série. Portanto, é necessário que os dados de carga estejam discretizados de forma horária. A Tabela 4.1 exemplifica um série de entrada com discretização horária, que é visualizada na Figura 4.2

Tabela 4.1 - Exemplo de série de entrada com 3 dias e discretização horária

	Carga(MW)		
Horário	Dia 1	Dia 2	Dia 3
00:00	43067	41051	44117
01:00	42123	40107	43173
02:00	40840	38824	41890
03:00	39444	37428	40494
04:00	38329	36313	39379
05:00	37561	35545	38611
06:00	35188	33172	36238
07:00	35027	33011	36077
08:00	39041	37025	40091
09:00	40106	38090	41156
10:00	40884	38868	41934
11:00	41135	39119	42185
12:00	40816	38800	41866
13:00	40209	38193	41259
14:00	40418	38402	41468
15:00	40624	38608	41674
16:00	41116	39100	42166
17:00	41383	39367	42433
18:00	42063	40047	43113

19:00	45406	43390	46456
20:00	50655	48639	51705
21:00	51430	49414	52480
22:00	49033	47017	50083
23:00	45806	43790	46856

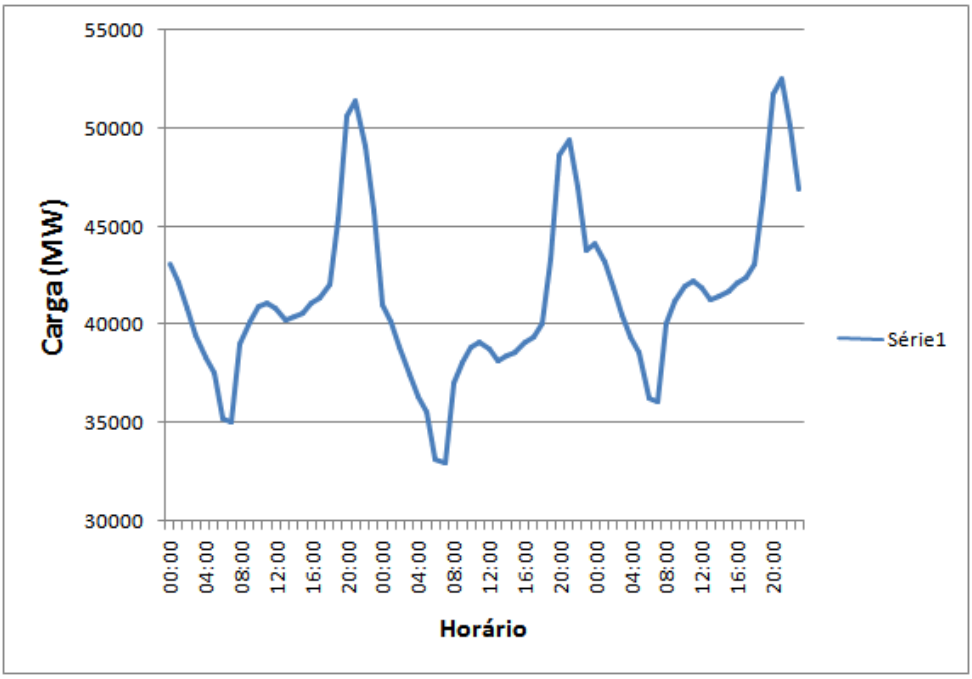

Figura 4.2 - Exemplo de série de entrada com 3 dias e discretização horária

Além da série de carga, também são entradas do modelo Neuro-Fuzzy outras informações (variáveis explicativas) que influenciam diretamente o consumo de energia elétrica [19] e, consequentemente, a previsão. Para o modelo proposto foram escolhidas as informações que indicam se cada dia da série é feriado, final de semana ou se pertence ao horário de verão brasileiro. Cada variável explicativa deve possuir um formato de vetor de dimensão n, sendo n é o número de dias da série histórica observada, onde o valor 1 indica que o dia é feriado, final de semana ou faz parte do horário de verão. Para a série de entrada genérica da Tabela 4.1, os vetores contendo as variáveis explicativas serão, por exemplo, os da

Tabela 4.2:

Tabela 4.2 – Vetores das variáveis explicativas

Vetor	Dia 1	Dia 2	Dia 3
Feriado	0	0	1
Final de Semana	1	1	0

Horário de Verão	0	0	0

Portanto, a partir dos vetores, o modelo reconhece que, para a série da Tabela 4.1, o dia 3 é feriado, os dias 1 e 2 são finais de semana e nenhum dos 3 dias fazem parte do horário de verão.

É importante ressaltar que uma série de carga proveniente de um sistema elétrico está sujeita a ruídos e erros de medição. Sendo assim, é necessária uma filtragem dos dados antes que estes sejam inseridos no modelo, uma vez que este não conta com nenhum tipo de filtro. Existem várias maneiras apropriadas para realizar a filtragem de séries temporais, mas estas não serão abordadas por não ser o objetivo do trabalho.

4.2 Modelo de Previsão

O modelo proposto para a previsão da série de carga consiste na divisão do problema em dois problemas menores: a previsão da ponta diária e previsão do perfil de curva diário da carga. A previsão da ponta diária da carga, que é o valor máximo de consumo em um dia, é feita utilizando as redes neurais supervisionadas do tipo Multilayer Perceptron (MLP) e a previsão do perfil de curva é realizada através da combinação de redes neurais não-supervisionadas, do tipo Mapa de Kohonen, e um Sistema de Inferência Fuzzy (FIS) baseado no método das distâncias. Após realizar cada previsão separadamente, o modelo une as duas respostas em uma única saída, que é a previsão de carga. A estrutura do modelo pode ser verificada no fluxograma da Figura 4.3, cujos blocos principais são descritos nas Seções 4.2.1 a 4.2.5.

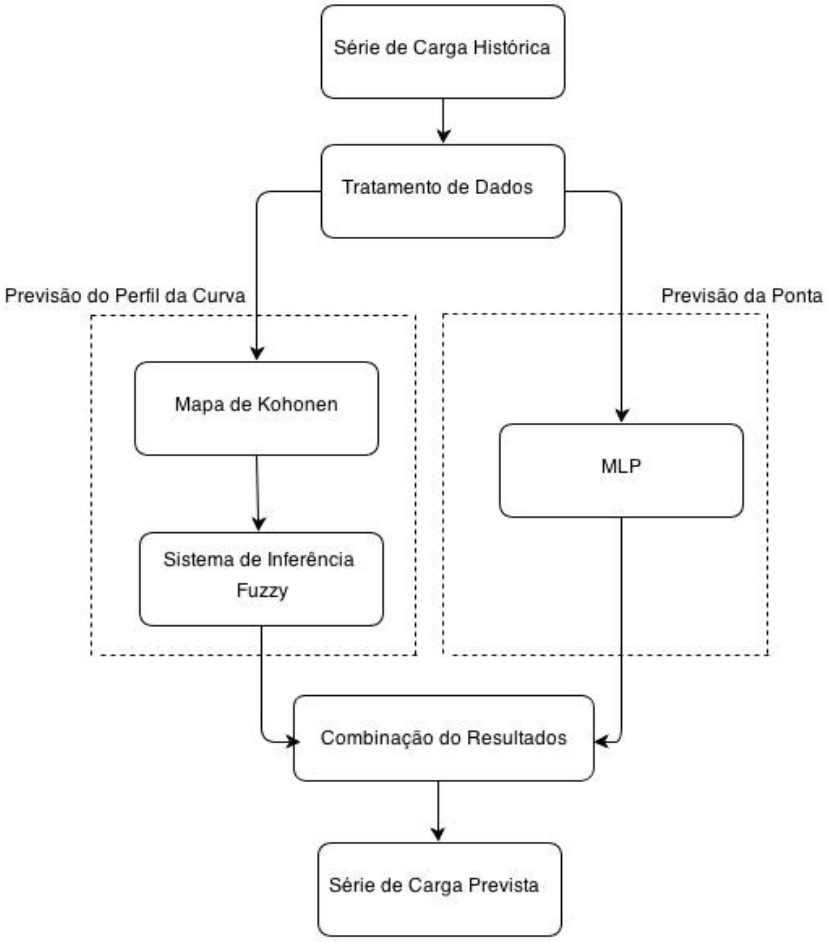

Figura 4.3 – Fluxograma do modelo proposto.

4.2.1 Tratamento de Dados

A etapa de tratamento de dados consiste em separar a série de carga de entrada em duas outras séries: a série de ponta de carga e a série do perfil de carga. A ponta ou pico de carga é dado pelo maior valor de consumo registrado em um determinado período. Para encontrar a série de ponta basta calcular o máximo valor de carga em cada dia da série horária, como mostra (4.1). A série de ponta calculada terá o formato de um vetor de dimensão n, onde n é o número de dias da série de entrada.

$$p(i) = \max\ [dia(i)] \tag{4.1}$$

Sendo:

$p(i)$- ponta de carga do dia i;

$dia(i)$ – vetor da carga do dia i com discretização horária.

Em seguida, a série do perfil de carga é obtida normalizando cada dia da série de entrada pelo respectivo valor da ponta da carga diária, segundo (4.2). A série do perfil da carga será uma matriz com a mesma dimensão da série de entrada.

$$c(i) = \frac{dia(i)}{p(i)} \tag{4.2}$$

Sendo:

$c(i)$- vetor do perfil de carga do dia i;

$dia(i)$ – vetor de carga do dia i com discretização horária;

$p(i)$- ponta de carga do dia i.

Como exemplo, pode-se separar o valor de ponta e o perfil de carga das curvas horárias da Tabela 4.1 - Exemplo de série de entrada com 3 dias e discretização horária,

resultando no vetor de ponta da Tabela 4.3 e na matriz do perfil de carga da Tabela 4.4, que são visualizados na Figura 4.4 e na Figura 4.5, respectivamente.

Tabela 4.3 – Vetor de Ponta de Carga

	Dia 1	Dia 2	Dia 3
Ponta (MW)	51430	49414	52480

Tabela 4.4 – Matriz do Perfil da Carga

	Carga (pu)		
Horário	Dia 1	Dia 2	Dia 3
00:00	0,8373	0,8307	0,8406
01:00	0,8190	0,8116	0,8226
02:00	0,7940	0,7856	0,7982
03:00	0,7669	0,7574	0,7716
04:00	0,7452	0,7348	0,7503
05:00	0,7303	0,7193	0,7357
06:00	0,6841	0,6713	0,6905
07:00	0,6810	0,6680	0,6874
08:00	0,7591	0,7492	0,7639
09:00	0,7798	0,7708	0,7842
10:00	0,7949	0,7865	0,7990
11:00	0,7998	0,7916	0,8035
12:00	0,7936	0,7852	0,7977
13:00	0,7818	0,7729	0,7861
14:00	0,7858	0,7771	0,7901
15:00	0,7898	0,7813	0,7940
16:00	0,7994	0,7912	0,8034
17:00	0,8046	0,7966	0,8085
18:00	0,8178	0,8104	0,8215
19:00	0,8828	0,8780	0,8852
20:00	0,9849	0,9843	0,9852
21:00	1,0000	1,0000	1,0000
22:00	0,9533	0,9514	0,9543
23:00	0,8906	0,8861	0,8928

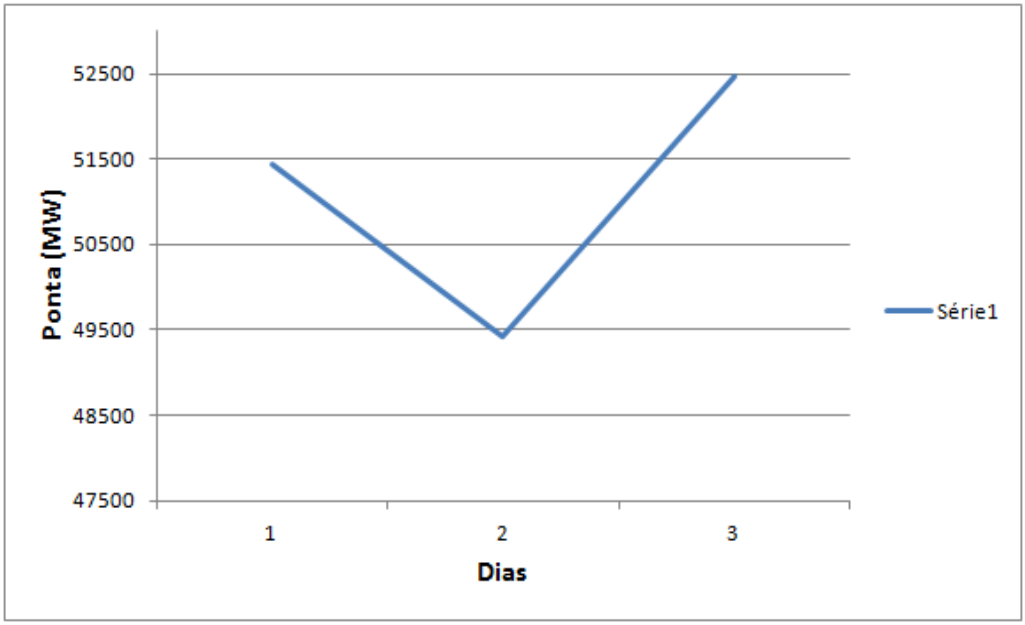

Figura 4.4 - Vetor de Ponta de Carga

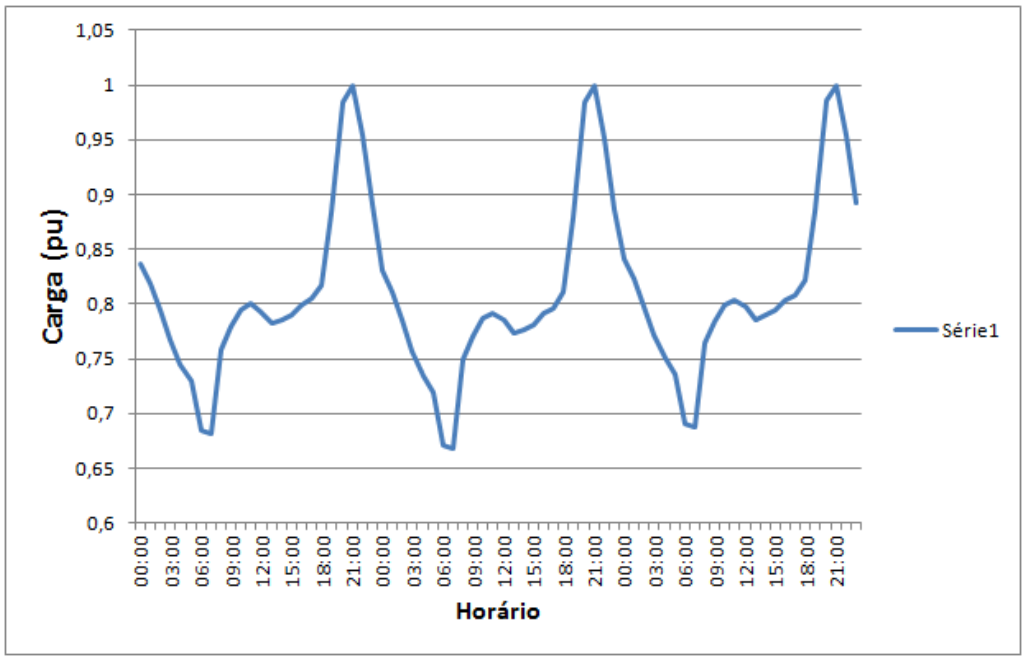

Figura 4.5 - Perfil da Carga

É importante notar que a série de perfil de carga não carrega consigo as informações absolutas de variação da carga. Ela possui apenas informações relativas; ou seja, não é possível, somente através do perfil da carga, saber o quanto a carga variou em um determinado período, mas sim como ela variou. Por outro lado, observando apenas a série de ponta tem-se apenas a informação do pico de consumo do dia, não havendo a

discriminação do momento em que ele ocorre. A série de ponta de carga é processada pela Rede Neural Artificial MLP, descrito na Seção 4.2.2, enquanto a série do perfil de carga é processada pelo Mapa de Kohonen e pelo Sistema de Inferência Fuzzy, descritos nas Seções 4.2.3 e 4.2.4.

4.2.2 MLP

A previsão da série de ponta é realiza através de uma Rede Neural Artificial do tipo MLP – Multilayer Perceptron, com algoritmo de treinamento de retropropagação de erro (*backpropagation error*). As regras para a escolha dos parâmetros da rede como o número de entradas, o número de camadas intermediárias e o número de neurônios em cada camada, não são bem definidas, de forma que a seleção destes parâmetros se dá por meio da realização de testes durante a fase de ajuste do modelo.

De posse da série de ponta histórica do sistema, o passo a seguir é separar a série de ponta em dois conjuntos: conjunto de treinamento e conjunto de validação, como mostra a Figura 4.6, que separa série de ponta da Tabela 4.3. Não há uma regra em relação à dimensão dos conjuntos, porém quanto maior for o conjunto de treinamento, maior será o aprendizado da rede neural e consequentemente melhor será a previsão.

Dia 1	Dia 2
51430	49414

Conjunto de Treinamento

Dia 3
52480

Conjunto de Validação

Figura 4.6 – Conjuntos de treinamento e validação

No conjunto de treinamento são fornecidos pares de entradas/saída desejada para o treinamento da rede, ou seja, para o ajuste dos pesos sinápticos da rede neural. Já no conjunto de validação, somente os sinais de entrada são aplicados à rede para que esta calcule os valores de saída.

A próxima tarefa é gerar pares de entrada e saída desejada para o treinamento da rede. Esses pares também são chamados de padrões de treinamento. Para estimar a ponta de carga de um dia qualquer, são selecionados como entrada valores de ponta observados nos dias anteriores e as varáveis explicativas relativas ao dia estimado. A saída desejada

será o valor de ponta observada no dia que se está prevendo. Como exemplo, pode-se gerar o padrão de treinamento do dia 2 da Tabela 4.1. Este padrão de treinamento é mostrado na Tabela 4.5.

Tabela 4.5 – Padrão de treinamento

Saída desejada	49414
Ponta do dia anterior	51430
Feriado	0
Final de Semana	1
Horário de Verão	0

A rede neural recebe o padrão de treinamento, faz uma estimativa do sinal de saída e a compara com o valor de saída desejada, ajustando seus pesos sinápticos. O processo de treinamento consiste em apresentar todos os padrões de treinamento à rede. Ao final deste procedimento, diz-se que a rede está treinada.

Após o treinamento da rede, são apresentados à rede dados de entrada do período de validação. Estes dados possuem o mesmo formato dos padrões de treinamento, porém sem a informação de saída desejada. O objetivo é justamente avaliar o quão bem a rede está executando a previsão. Um exemplo de padrões de entrada da rede no conjunto de validação é o do dia 3 da Tabela 4.1, como mostra a Tabela 4.6.

Tabela 4.6 – Padrões de entrada

Ponta do dia anterior	49414
Feriado	1
Final de Semana	1
Horário de Verão	0

Ao aplicar um padrão de entrada à rede, obtém-se a previsão da ponta para o dia referente aquele padrão. Para a previsão de uma série de ponta é necessário a criação e apresentação a rede dos padrões de treinamento referentes todos os dias da série. Sendo apresentado um padrão por vez e consequentemente, obtendo a previsão de um dia por vez até totalizar todos os dias da série. Esta série de previsões de ponta gerada pela rede no conjunto de validação passa então para outra etapa do modelo, a combinação dos resultados, descrita na Seção 4.2.5.

4.2.3 Mapa de Kohonen

No Mapa de Kohonen, o objetivo é classificar as curvas diárias em padrões semelhantes em processo conhecido como clusterização. No caso de curvas de carga, cada neurônio presente no reticulado ou grade representa um tipo de curva diferente. Por exemplo, curvas de finais de semanas possuem características diferentes de curvas de dias de semana e serão, portanto, representadas por neurônios distintos. Cada neurônio possui um cluster associado a ele, de forma que todas as curvas representadas por um neurônio são armazenadas no cluster correspondente, sendo que o número de neurônios/clusters é definido através de testes realizados na etapa de ajustes do modelo. Desta forma, o Mapa de Kohonen analisa uma curva diária por vez, e a classifica em um determinado cluster. A média de todas as curvas armazenadas em um mesmo cluster recebe o nome de centro de gravidade do cluster. Um exemplo do processo de clusterização das curvas da Tabela 4.1 é visualizado na Tabela 4.7 e na Figura 4.7, na Figura 4.8 e na Figura 4.9. Neste exemplo, a curva do Dia 1 é agrupada no cluster 2, pois neste cluster estão as curvas que possuem padrões(identificados pelo Mapa de Kohonen) mais semelhantes aos da curva do Dia 1. As curvas do Dia 2 e do Dia 3 são agrupadas, respectivamente, nos clusters 1 e 3 pela mesma razão.

Tabela 4.7 – Clusterização de Curvas

Dia	Cluster
Dia 1	2
Dia 2	1
Dia 3	3

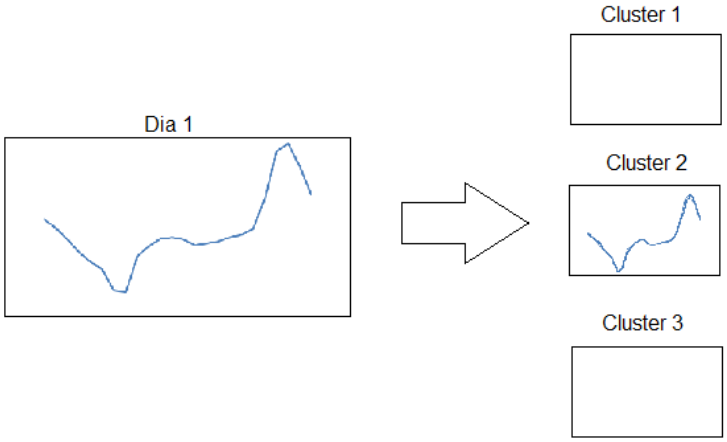

Figura 4.7 – Clusterização do dia 1

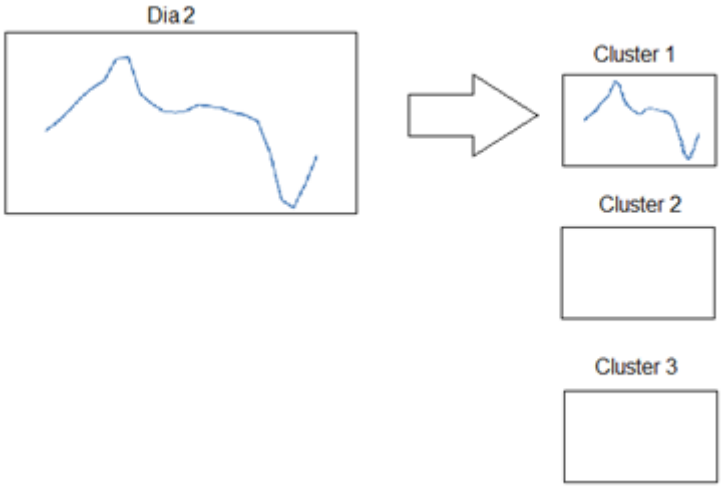

Figura 4.8 - Clusterização do dia 2

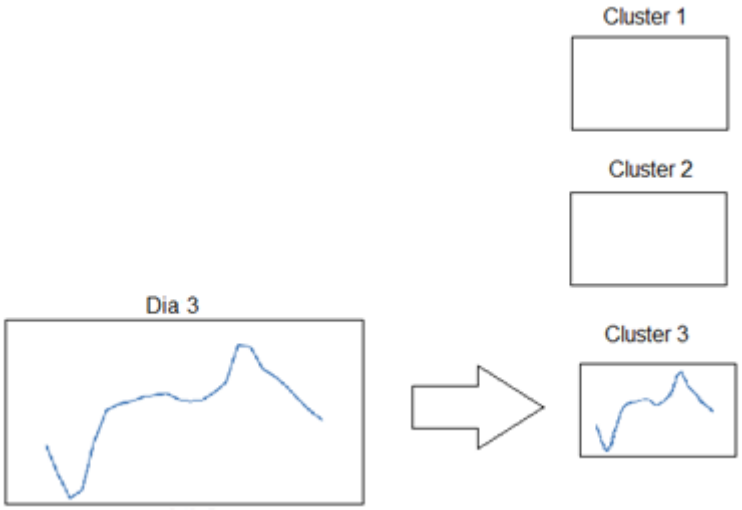

Figura 4.9 - Clusterização do dia 3

4.2.4 Sistema de Inferência Fuzzy

Analogamente à previsão de ponta, na previsão de perfil de carga também é necessário dividir a série de perfil de carga em dois conjuntos distintos: conjunto para a criação da base de regras e conjunto de validação de dados.

No conjunto de criação da base de regras, os dados de entrada são analisados pelo sistema de inferência e as regras fuzzy são geradas e armazenadas. Cada regra fuzzy é formada a partir dos dias antecedentes, variáveis explicativas e dia consequente. Os antecedentes e o consequente são representados pelos clusters em que a curva de cada dia foi classificada, como mostra o exemplo de uma regra fuzzy da Tabela 4.8, que monta a regra corresponde ao dia 2 da Tabela 4.1.

Tabela 4.8 – Exemplo de regra fuzzy

Regra	Antecedente	Feriado	Final de Semana	Horário de Verão	Consequente
Dia 2	Cluster 2	0	1	0	Cluster 1

A regra da Tabela 4.8 segue o padrão da Lógica Fuzzy e significa "*SE* antecedente = Cluster 2 e feriado=0 e final de semana=1 e horário de verão=0 *ENTÃO* consequente= Cluster 1".

Neste caso, o antecedente significa que no dia anterior ao dia 2, ou seja, no dia 1, houve um perfil de carga que foi classificado segundo os padrões do cluster 2. Por sua vez,

o consequente significa que no próprio dia 2, houve um perfil de carga que foi classificado segundo os padrões do cluster 1.

Para criar a base de regras, todos os dias do conjunto são analisados de forma que se tenha o maior número de regras possível. Quanto mais regras forem geradas, maior o número de situações memorizadas pelo modelo e com isso espera-se que melhor seja a previsão.

Após a geração das regras fuzzy a partir de todos os dias pertencentes ao conjunto de criação de regras, o sistema de inferência está apto a estimar o perfil de carga de dias que não fazem parte deste conjunto. Para cada dia que se deseja prever o perfil de carga é necessário calcular a pertinência deste dia em relação à cada regra. Esta pertinência é calculada através das distâncias entre as curvas dos dias que antecedem o dia de previsão e os centros de gravidade dos clusters associados aos antecedentes das regras, de acordo com (3.9) e (3.10). Como exemplo, pode-se calcular a pertinência da regra da Tabela 4.8, gerada a partir do dia 2, em relação ao dia 3, que é o dia do conjunto de validação. Este processo é visualizado em (4.3):

$$\mu_{32} = \frac{1}{\left(\frac{d_{32}}{d_{31}} + \frac{d_{32}}{d_{32}} + \frac{d_{32}}{d_{33}}\right)^{\frac{2}{m-1}}} \qquad (4.3)$$

Sendo:

μ_{32} – pertinência da curva do dia 3 em relação ao centro de gravidade do cluster do antecedente 2;

d_{31} – distância da curva do dia 3 em relação ao centro de gravidade do cluster 1;

d_{32} – distância da curva do dia 3 em relação ao centro de gravidade do cluster 2;

d_{33} – distância da curva do dia 3 em relação ao centro de gravidade do cluster 3.

Onde cada distância calculada através de (4.4):

$$d_{xy} = \sqrt{\sum_{i=1}^{24}(x(i) - y(i))^2} \qquad (4.4)$$

Sendo:

d_{xy}- distância da curva do dia x ao centro de gravidade do cluster y;

x – curva horária do dia x;

y – curva horária do dia y;

A pertinência da regra será igual a μ_{32}. É importante ressaltar que a pertinência de uma regra é o produto de todas as pertinências dos antecedentes que formam aquela regra, como mostra em (4.5). Como neste exemplo temos apenas 1 antecedente, a pertinência deste antecedente será também a pertinência de toda a regra.

$$\mu_x^k = \prod_{i=1}^{n} \mu_{xn} \qquad (4.5)$$

Sendo:

μ_x^k – pertinência da curva do dia x em relação à regra k;

μ_{xn} - pertinência da curva do dia x em relação ao centro de gravidade do cluster do antecedente n;

n – número de antecedentes de cada regra.

Ao multiplicar-se a pertinência de uma regra pelo centro de gravidade do cluster classificado como consequente desta mesma regra, tem-se a chamada contribuição da regra, como mostra (4.6):

$$contribuição_x^k = \mu_x^k . consequente^k \tag{4.6}$$

Sendo:

$contribuição_x^k$ – contribuição da regra k em relação ao dia x;

μ_x^k – pertinência da curva do dia x em relação à regra k;

$consequente^k$ – centro de gravidade do cluster consequente da regra k.

O perfil de carga previsto para um dia é encontrado somando-se as contribuições de cada regra ativada e dividindo o resultado pela soma de todas as pertinências ativadas, como mostra (4.7). As regras ativadas são aquelas que pertencem ao mesmo conjunto de variáveis explicativas do dia que se deseja prever. Por exemplo, para um dia que é feriado, não é final de semana e não faz parte do horário de verão, e portando forma a sequência de variáveis explicativas [1;0;0], as regras ativadas serão somente aquelas que possuem a mesma sequência de variáveis explicativas.

$$c_x^p = \frac{\sum_{k=1}^{n} contribuição_x^k}{\sum_{k=1}^{n} \mu_x^k} \tag{4.7}$$

Sendo:

c_x^p - curva prevista para dia x;

$contribuição_x^k$ – contribuição da regra ativada k em relação ao dia x;

μ_x^k – pertinência da curva do dia x em relação à regra ativada k;

n – número de regras ativadas.

A previsão de uma série do perfil de carga é feita realizando a previsão de um dia de cada vez, através do mesmo processo; isto é, para cada dia que se deseja prever, calcula-se a pertinência de cada regra da base de dados, as contribuições de cada regra, e finalmente a previsão para este dia segundo (4.7).

Esta série prevista é a saída do sistema de inferência fuzzy. A série então passa a etapa de combinação dos resultados, descrita na Seção 4.2.5.

4.2.5 Combinação dos Resultados

A previsão de carga diária realizada pelo modelo é dada pelo produto da previsão de ponta pelo perfil da curva de um mesmo dia, segundo a (4.8):

$$carga_x^p = p_x^p \, . \, c_x^p \qquad (4.8)$$

Sendo:

$carga_x^p$ - carga prevista para o dia x;

p_x^p - ponta prevista para o dia x;

c_x^p – perfil da curva previsto para o dia x.

4.3 Índices de Desempenho

Os índices de desempenho utilizados para avaliação do modelo proposto são divididos em três tipos: MAD, MAPE e RMSE [20].

- MAD (Erro Médio Absoluto) - média dos desvios absolutos individuais gerados em cada período. O MAD não considera a direção do erro (positivo ou negativo), o que é útil quando se trabalha com séries temporais, pois estas contêm variações positivas e negativas, o que poderia anular uma variação. O MAD é dado por (4.9):

$$MAD = \frac{\sum_{i=1}^{n} |e_i - o_i|}{n} \qquad (4.9)$$

- MAPE (Erro Percentual Absoluto Médio) – soma dos percentuais absolutos dos erros divididos pelo número de observações. Identifica o percentual médio que a previsão desviou do valor desejado. O MAPE é dado por (4.10):

$$MAPE = \frac{\sum_{i=1}^{n} \frac{|e_i - o_i|}{o_i}}{n} \, x \, 100\% \qquad (4.10)$$

- RMSE – É a raiz quadrada do erro quadrático médio, conforme (4.11):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (e_i - o_i)^2}{n}} \qquad (4.11)$$

Sendo:

e_i - valor estimado pelo modelo;

o_i - valor observado na série histórica;

n - número de observações.

Para a avaliação do modelo, considerou-se uma boa previsão aquela em que o índice MAPE é inferior a 5% [21].

4.4 Ajuste do Modelo

Na fase de ajuste do modelo os parâmetros de entrada de cada modelo são alterados manualmente, isto é, o usuário é quem modifica os parâmetros. O objetivo destas alterações é obter qual combinação entre estes parâmetros gera os melhores índices de desempenho. Como mencionado anteriormente, não há nenhuma regra que define como escolher de maneira ótima estes parâmetros. Portanto, este tipo de teste é uma boa alternativa para identificar a combinação dos parâmetros a rede se adapta melhor.

Cada etapa do modelo proposto – MLP, Mapa de Kohonen e Sistema de Inferência Fuzzy – possui parâmetros que podem ser alterados visando uma melhor previsão. As Seções 4.4.1, 4.4.2 e 4.4.3 apresentam os parâmetros de cada modelo separadamente.

4.4.1 MLP

Os parâmetros a serem ajustados na rede MLP do modelo são:

- Número de camadas intermediárias;
- Número de neurônios nas camadas intermediárias;
- Número de dias anteriores à previsão usados como padrões de treinamento e de entrada.

4.4.2 Mapa de Kohonen

O parâmetro a ser ajustado no Mapa de Kohonen do modelo é:

- Número de neurônios/clusters do mapa.

4.4.3 Sistema de Inferência Fuzzy (FIS)

Os parâmetros a serem ajustados no FIS do modelo são:

- Constante de fuzzyficação;
- Dias antecedentes à previsão utilizados para a criação das regras fuzzy e cálculo das pertinências.

4.5 Previsão de Valores Futuros

Após a realização do ajuste do modelo, onde os parâmetros de entrada são definidos de maneira se obter os melhores índices de desempenho, o modelo está apto a realizar previsões de valores futuros, ou seja, valores que não fazem parte do período utilizado como treinamento, criação da base de regras fuzzy e validação de dados.

O modelo de previsão descrito nesse capítulo foi implementado utilizando o software MATLAB.

Capítulo 5

5 Implementação do Modelo e Resultados

Neste capítulo serão apresentados os resultados obtidos a partir do modelo de previsão de carga proposto. São apresentados de maneira detalhada os dados utilizados como entrada do modelo, as etapas da previsão e os testes realizados na fase de ajuste do modelo e os resultados obtidos para a série prevista. O software utilizado para a programação da metodologia proposta, realização do ajuste do modelo e análise dos resultados foi o MATLAB.

5.1 Dados Utilizados

Os dados utilizados como entradas do modelo de previsão proposto são baseados em uma série de carga real de parte do sistema brasileiro e possui as informações de consumo dos anos de 2004 a 2008. O período de janeiro de 2004 a dezembro de 2007 foi selecionado como conjunto de treinamento na etapa MLP e como conjunto para criação da base de regras do sistema de inferência fuzzy. O período de janeiro de 2008 foi escolhido como período de validação de dados, tanto no MLP quanto no FIS.

Os dados apresentam uma discretização horária e foram previamente filtrados a fim de minimizar erros provocados pela medição dos dados utilizando [22]

5.2 Previsão da Série de Ponta

A previsão da série de ponta é separada nas etapas de treinamento da rede neural e validação de dados, como descrito na Seção 4.2.2 . O treinamento da rede MLP foi realizado através da ferramenta *Neural Network Toolbox* do programa MATLAB [23]. Nesta ferramenta são introduzidas matrizes que contêm os vetores com padrões de treinamento da rede gerados a partir do processo descrito na Seção 4.2.2. A Figura 5.1 mostra alguns

dos padrões de treinamento utilizados e a Figura 5.2 mostra a ferramenta *Neural Network Toolbox* durante o treinamento da rede MLP. Onde P são os padrões de entrada da rede e Td são respectivas saídas desejadas.

```
P =

  Columns 1 through 5

    0.0323    0.0322    0.0327    0.0333    0.0327
    0.0327    0.0327    0.0326    0.0330    0.0337
    0.0268    0.0327    0.0327    0.0326    0.0330
         0         0         0         0         0
    0.0637    0.0637    0.0637    0.0637    0.0637

Td =

  Columns 1 through 5

    0.0327    0.0326    0.0330    0.0337    0.0333
```

Figura 5.1 – Padrões de Treinamento

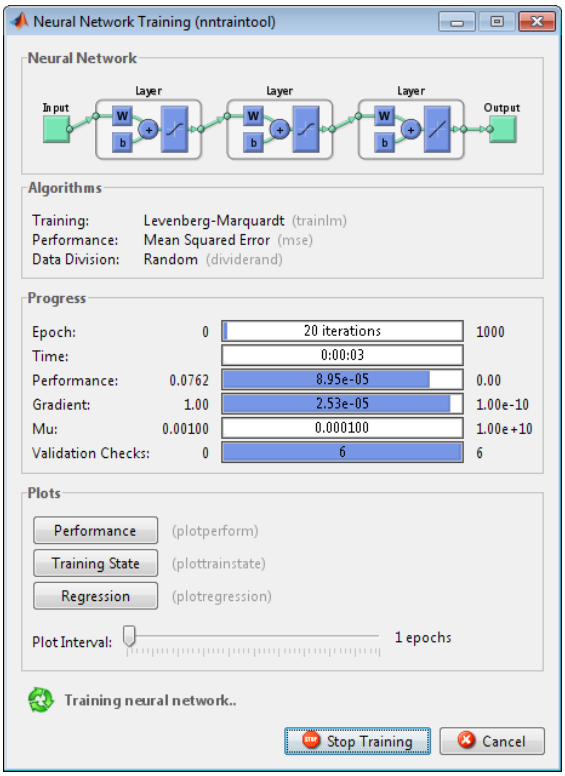

Figura 5.2 – Treinamento da rede supervisionada através da ferramenta *Neural Network Toolbox*

Após o treinamento da rede, o MLP é capaz de prever valores de saída a partir de entradas no mesmo formato daquelas fornecidas à rede no período de treinamento, como descrito na Seção 4.2.2. Ao apresentar os padrões de entrada à rede treinada, obtém-se a previsão da série de carga para o período de validação.

5.3 Previsão da Série do Perfil de Carga

Na previsão do perfil de carga, os dados primeiramente são classificados em padrões semelhantes e armazenados em clusters, como descrito na Seção 4.4.2. A Figura 5.3 mostra o treinamento da rede neural não supervisionada através da ferramenta *Neural Network Toolbox* do programa MATLAB. A Figura 5.4 mostra as curvas diárias classificadas no Mapa de Kohonen com 36 neurônios no reticulado.

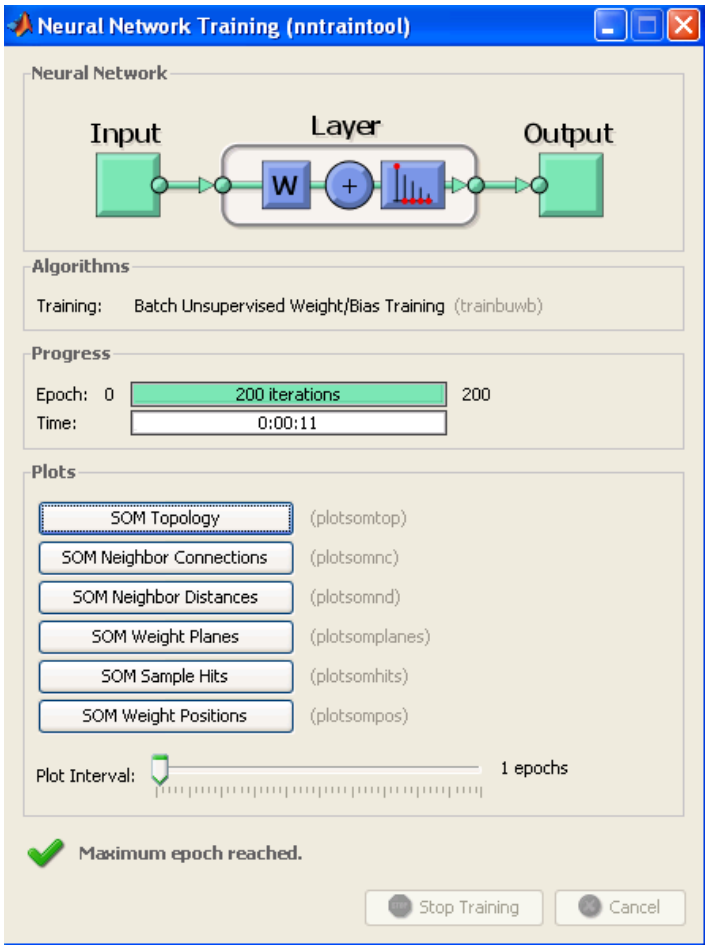

Figura 5.3 - Treinamento da rede não-supervisionada através da ferramenta *Neural Network Toolbox*

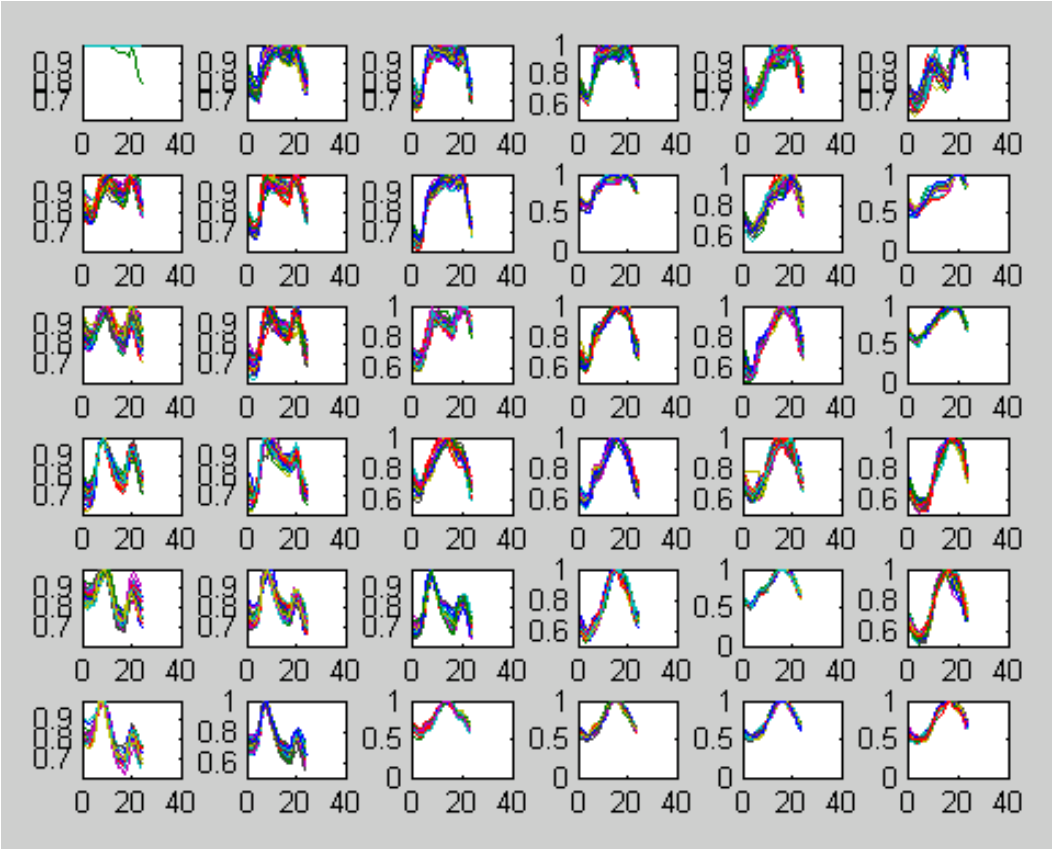

Figura 5.4 – Curvas horárias de entrada classificadas no Mapa de Kohonen com 36 neurônios

Após a clusterização de todas as curvas diárias, há a geração da base de regras fuzzy, como descrito na Seção 4.2.4. A Figura 5.5 mostra algumas das regras fuzzy geradas nesse processo.

```
regra =

   16.0000   16.0000   16.0000   16.0000   16.0000    4.0000         0    1.0000         0   16.0000
   16.0000   16.0000   16.0000   16.0000   16.0000   12.0000         0    1.0000         0   16.0000
   16.0000   16.0000   16.0000    8.0000    8.0000    4.0000         0    1.0000    1.0000    8.0000
    8.0000   16.0000   16.0000    4.0000    4.0000    4.0000         0    1.0000    1.0000    4.0000
    4.0000    8.0000   16.0000   16.0000   12.0000   12.0000         0    1.0000         0   16.0000
   16.0000    4.0000    8.0000   16.0000   16.0000   16.0000         0    1.0000         0   16.0000
   16.0000   16.0000    4.0000   16.0000   16.0000   16.0000         0    1.0000         0   16.0000
   16.0000   16.0000   16.0000   16.0000   16.0000   16.0000         0    1.0000         0   16.0000
   16.0000   16.0000   16.0000    8.0000    8.0000    8.0000         0    1.0000    1.0000    8.0000
    4.0000    8.0000   16.0000   16.0000   16.0000   16.0000         0    1.0000         0   12.0000
```

Figura 5.5 – Regras fuzzy

Com a base de regras fuzzy, é possível prever perfis de carga a partir do cálculo da pertinência de cada regra, como descrito na Seção 4.2.4.

5.4 Testes

Como descrito na Seção 4.4, os parâmetros de entrada de cada modelo são escolhidos a partir de testes que avaliam qual combinação destes parâmetros gera os melhores índices de desempenho. Os testes foram realizados variando cada parâmetro de cada vez, fixando-se todos os outros. As Seções 5.4.1 a 5.4.6 descrevem os testes realizados para cada parâmetro.

5.4.1 Número Camadas Intermediárias

O número de camadas intermediárias foi definido baseando-se no fato de que a rede neural MLP *backpropagation* não demonstra vantagem significativa no uso de duas camadas escondidas ao invés de uma para problemas menores. Por isso, para a grande maioria dos problemas utiliza-se apenas uma camada escondida quando muito duas e não mais que isso [24]. Sendo assim, para a previsão da série de ponta foram utilizadas duas camadas intermediárias.

5.4.2 Número de Neurônios nas Camadas Intermediárias

Para a definição do número de neurônios nas camadas intermediárias foram realizados testes variando o número de neurônios em cada camada entre 1 e 10. Assim temos uma combinação de 100 resultados possíveis. Estes resultados são visualizados na Tabela 1 do Anexo I, e apresentam os menores índices de desempenho quando configurados para 5 neurônios na primeira camada intermediária e 6 neurônios na segunda camada escondida, com um índice MAPE de 2,07%. Assim, esses foram os valores adotados.

5.4.3 Dias Anteriores à previsão (MLP)

Os testes para a definição dos dias anteriores à previsão utilizados para a previsão de ponta são semelhantes aos realizados para a definição do número de neurônios nas camadas intermediárias, isto é, fixam-se todos os outros parâmetros e varia-se a combinação de dias anteriores e analisa-se os resultados a fim de obter qual delas apresenta os menores índices de desempenho. Foram avaliados cenários com 1 até 6 dias anteriores, onde estes dias não são necessariamente consecutivos. Pode-se, por exemplo, utilizar o primeiro o quarto e o sétimo dias anteriores à previsão, na forma de um vetor [1, 4,7]. A Tabela 5.1 apresenta as combinações que apresentaram os melhores resultados, sendo o melhor desempenho obtido utilizando os dias anteriores 1, 2, 7,14 e 28, nesta ordem. Neste caso, obteve-se um índice MAPE de 1,65%.

Tabela 5.1 – Dias anteriores à previsão - MLP

Dias Anteriores	MAD[MW]	MAPE[%]	RMSE[MW]
1-2-7-14-28	52,19	1,65	65,07
1-7-4	55,18	1,74	67,75
1-7-14-28	57,50	1,85	73,61
1-2-3-7-14-28	56,61	1,85	77,18
1-2-7-14	61,24	1,95	77,92
1-2-3-4-5-14-28	61,54	1,98	88,01
1-2-7-14-28-35	63,26	1,99	89,03
1-2-3-4-5-6-7	61,77	2,00	88,92
7-14-28	75,87	2,38	96,78

5.4.4 Número de Neurônios do Mapa de Kohonen

Os testes para o ajuste do número de neurônios e clusters do Mapa de Kohonen consistiram em variar o número de neurônios de 1 a 100, variando-se o número de linhas e de colunas do reticulado, e fixando-se todos os outros parâmetros e analisando qual desse valor gera os menores índices de desempenho. Os resultados são visualizados na Tabela 2 do Anexo I e apresentam como o melhor índice a quantidade de 16 neurônios no reticulado, com 4 linhas de neurônios com 4 colunas cada, obtendo um índice MAPE de 1,48%. Assim, esses foram os valores adotados.

5.4.5 Constante de Fuzzyficação

Para o ajuste da constante de fuzzyficação, os testes foram realizados fixando-se os outros parâmetros e variando a constante de fuzzyficação de 0.1 a 10 e observados os menores valores dos índices de desempenho. Os resultados são mostrados na Tabela 3 do Anexo I e os melhores resultados encontrados para a constante de fuzzyficação igual a 2.2, com um índice MAPE de 1,51%. Assim, foi este o valor adotado.

5.4.6 Dias antecedentes à Previsão (FIS)

Os testes para a definição dos dias anteriores à previsão utilizados para a previsão de ponta foram realizados fixando-se todos os outros parâmetros, variando a combinação de dias antecedentes e analisando os resultados para obter qual das combinações apresenta os menores índices de desempenho. Foram avaliados cenários com 1 até 10 dias anteriores,onde estes dias não são necessariamente consecutivos. A Tabela 5.2 apresenta as combinações que apresentaram os melhores resultados, sendo o melhor desempenho obtido utilizando os dias anteriores 1, 2, 3,7,14 e 28, nesta ordem. Com um índice MAPE de 1,78%.

Tabela 5.2 - Dias anteriores à previsão - FIS

Dias Anteriores	MAD[MW]	MAPE[%]	RMSE[MW]
1-2-3-7-14-28	55,84	1,78	71,80
1-3-7-14-28	58,99	1,86	80,33
1-2-7-14-28	60,52	1,94	99,91
1-7-14-28	63,19	2,02	88,86
1-7-14	66,44	2,06	117,11
1-2-3-4-7-14	69,25	2,17	105,36
1-3-7-14	70,60	2,22	116,84
1-2-7-14	79,48	2,49	121,88
1-4-7-14	92,70	2,91	142,58

5.5 Resultados Obtidos

Após os ajustes de todos os parâmetros de entrada do modelo, foi realizada uma nova previsão do período de validação, utilizando os parâmetros selecionados na fase de ajuste do modelo, que são:

- Número de camadas intermediárias – 2;
- Número de neurônios nas camadas intermediárias – 5 e 6;
- Dias anteriores à previsão (MLP) – 1, 2,7,14 e 28;
- Numero de neurônios do Mapa de Kohonen – 16;
- Constante de fuzzyficação – 2.2
- Dias antecedentes à previsão (FIS) – 1, 2 ,3,7,14,28.

A partir destes parâmetros, obteve-se a previsão para janeiro de 2008. O resultado da previsão apresenta os índices de desempenho da Tabela 5.3 e pode ser visualizado graficamente nas Figuras Figura 5.6 à Figura 5.10, que apresentam as cargas previstas separadas por semana.

Tabela 5.3 – Índices de desempenho obtidos

MAD[MW]	MAPE[%]	RMSE[MW]
48,40	1,52	62,12

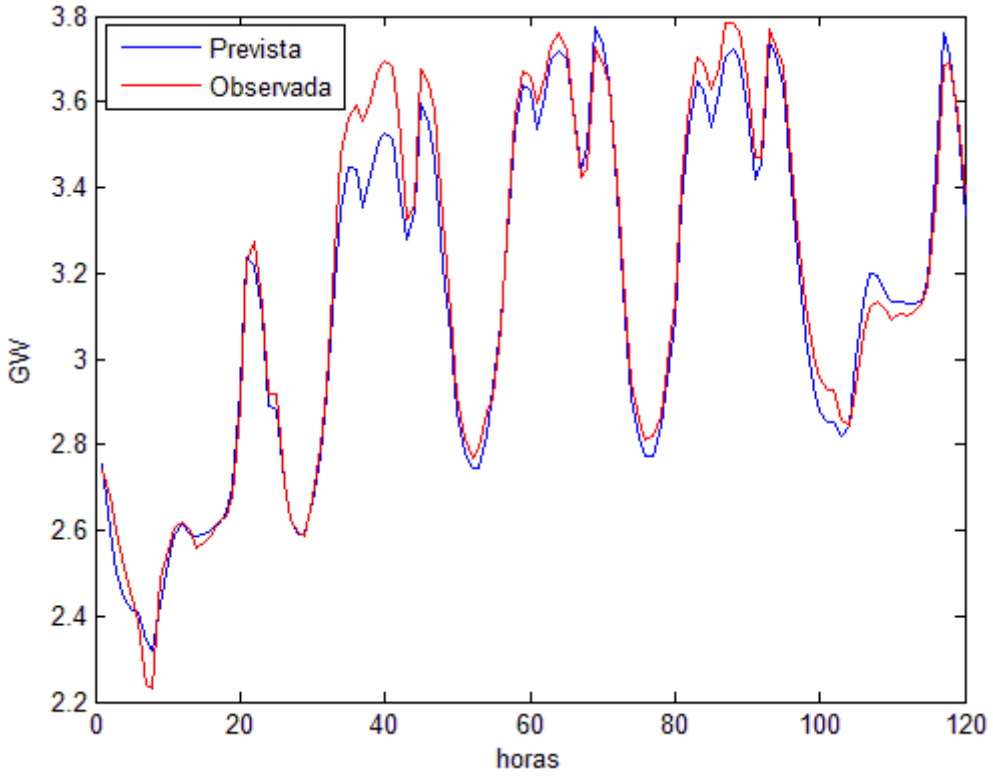

Figura 5.6 - Previsão de Carga Janeiro 2008 – Semana 1

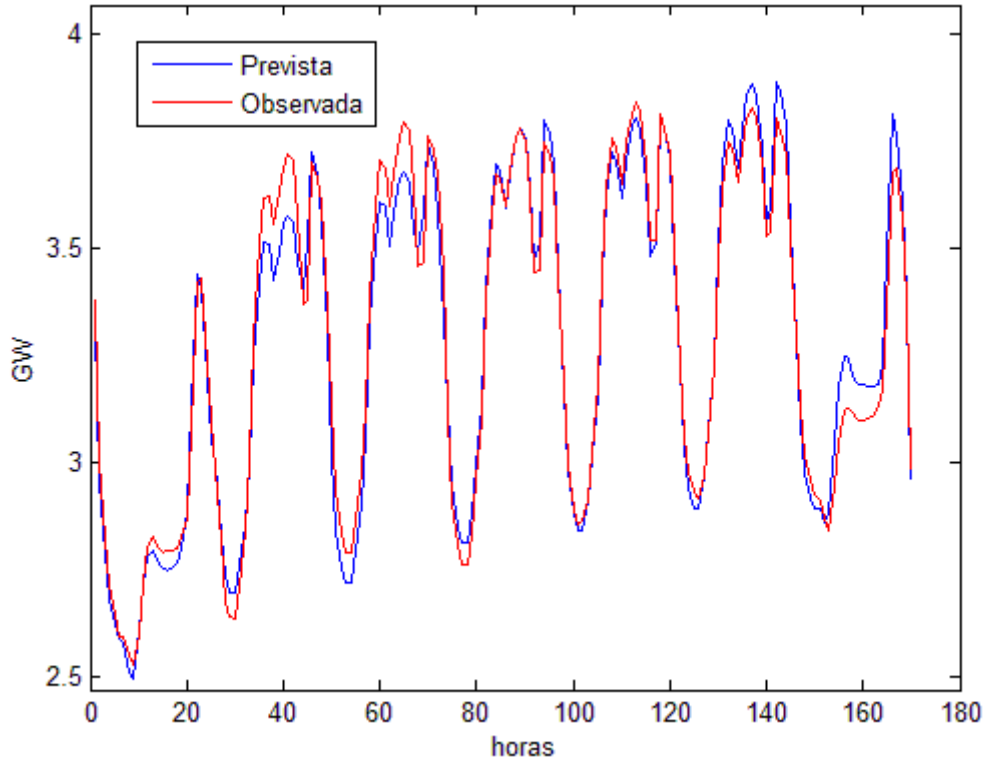

Figura 5.7 - Previsão de Carga Janeiro 2008 – Semana 2

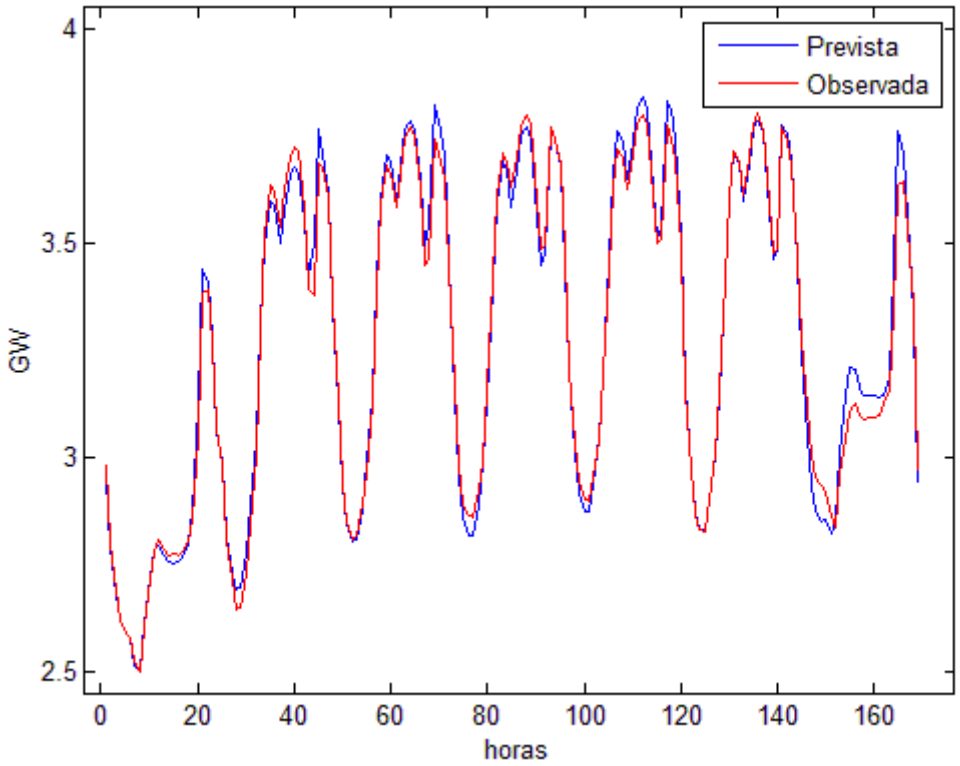

Figura 5.8 - Previsão de Carga Janeiro 2008 – Semana 3

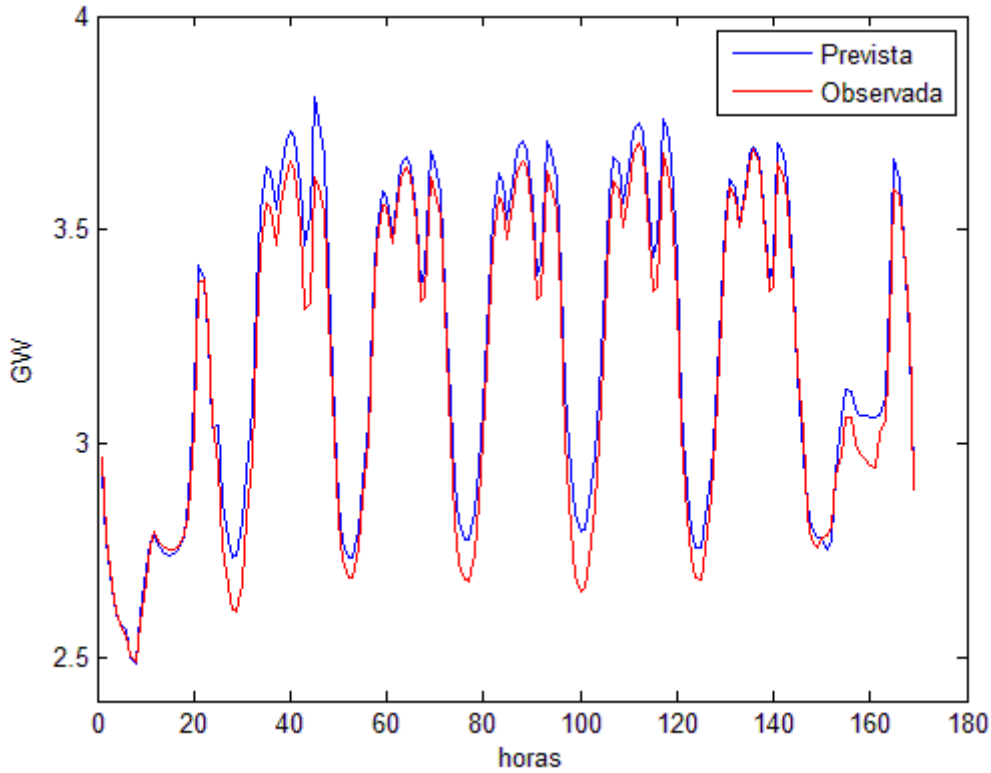

Figura 5.9 - Previsão de Carga Janeiro 2008 – Semana 4

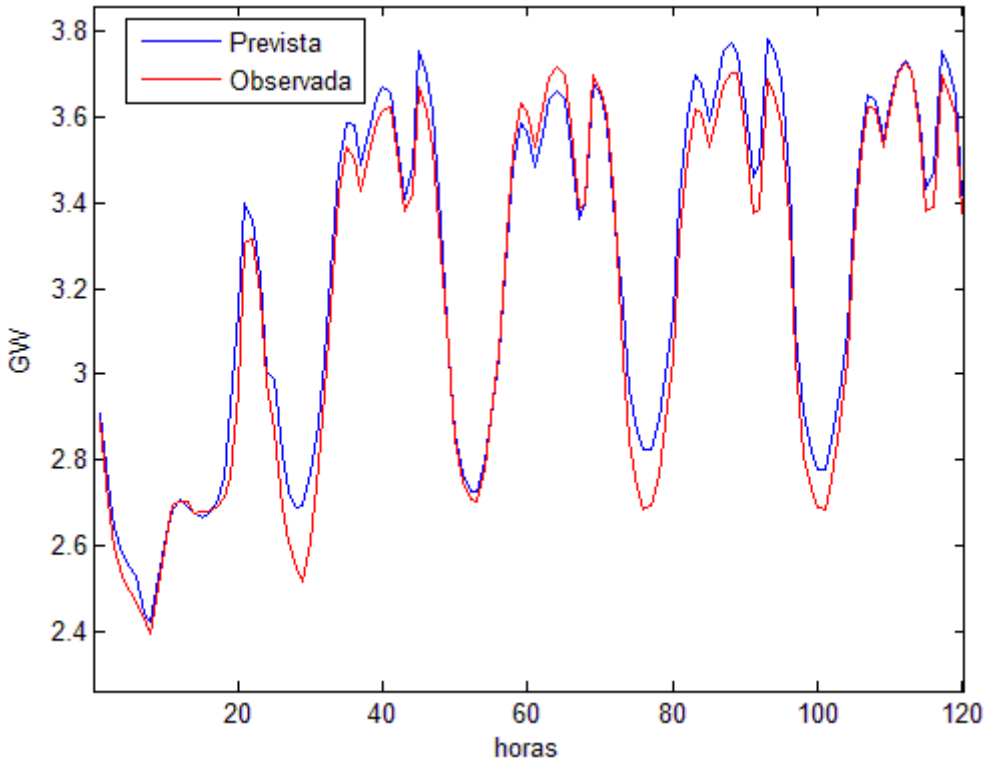

Figura 5.10 - Previsão de Carga Janeiro 2008 – Semana 5

A Tabela 5.4 apresenta o índice de desempenho MAPE para cada dia da previsão.

Tabela 5.4 – MAPE diário da previsão de Janeiro de 2008

Dia	Dia da Semana	MAPE [%]
1	Terça-Feira	1,39
2	Quarta-Feira	2,43
3	Quinta-Feira	0,96
4	Sexta-Feira	1,38
5	Sábado	1,54
6	Domingo	1,04
7	Segunda-Feira	2,10
8	Terça-Feira	2,44
9	Quarta-Feira	0,92
10	Quinta-Feira	0,63
11	Sexta-Feira	1,09
12	Sábado	2,09
13	Domingo	0,51
14	Segunda-Feira	1,13
15	Terça-Feira	0,73
16	Quarta-Feira	0,77
17	Quinta-Feira	0,81

18	Sexta-Feira	0,22
19	Sábado	1,91
20	Domingo	0,56
21	Segunda-Feira	3,54
22	Terça-Feira	1,14
23	Quarta-Feira	2,18
24	Quinta-Feira	2,87
25	Sexta-Feira	1,24
26	Sábado	1,70
27	Domingo	1,48
28	Segunda-Feira	2,75
29	Terça-Feira	0,83
30	Quarta-Feira	2,99
31	Quinta-Feira	1,61

Analisando a Tabela 5.4 e especificamente o dia 1 de janeiro, que é feriado, percebe-se que o modelo conseguiu estimar de forma satisfatória esse tipo de dia, já que o índice MAPE do dia 1 é menor do que aquele referente a todo o período previsto. A Tabela 5.5 apresenta o MAPE médio para cada dia da semana no período previsto.

Tabela 5.5 – MAPE médio para cada dia da semana

Dia da Semana	MAPE [%]
Domingo	0,90
Segunda-Feira	2,38
Terça-Feira	1,31
Quarta-Feira	1,86
Quinta-Feira	1,38
Sexta-Feira	0,98
Sábado	1,81

Analisando a Tabela 5.5 podemos perceber que o dia da semana com maior dificuldade de previsão foi a segunda-feira. Isto ocorre, provavelmente, pelo fato da carga deste dia ser fortemente influenciada pela carga de domingo, que é quase sempre menor em relação aos dias úteis. Contudo, o modelo não conseguiu ajustar uma forma de prever de maneira mais precisa a carga das segundas-feiras. Isto pode ser ajustado alterando-se os dias antecedentes das previsões para obter uma previsão melhor, porém isso prejudicaria a previsão dos demais dias, já que os parâmetros foram selecionados de forma a minimizar o erro global.

Capítulo 6

6 Conclusões

6.1 Conclusões

Neste trabalho apresentou-se um modelo para a previsão de séries de carga em sistemas de energia elétrica. O método desenvolvido foi motivado pela grande importância da previsão para o planejamento e operação dos diversos agentes do setor elétrico e a consequente necessidade de aperfeiçoamento das ferramentas utilizadas para a previsão. Em comparação aos modelos de previsão estatísticos ou modelos que utilizam simplesmente um método de inteligência artificial, a combinação Neuro-Fuzzy apresenta grande vantagem pois associa a capacidade de aprendizado das Redes Neurais Artificiais com a habilidade extrair conhecimento de informações através de regras da Lógica Fuzzy, além da sua aptidão em lidar com incertezas e imprecisão de dados. Esta combinação resulta em um modelo capaz de melhor improvisar diante de situações não conhecidas e consequentemente realizar melhores previsões.

A partir do modelo desenvolvido, utilizando-se dados de 2004 a 2007 foram estimados os consumos de carga relativos ao mês de janeiro de 2008, obtendo índices de desempenho dentro de uma margem satisfatória, o que comprova a validade do uso combinado das Redes Neurais Artificiais e da Lógica Fuzzy quando aplicados à previsão de séries temporais.

O modelo proposto realizou previsões bem próximas ao valor registrado em todo o período analisado, mesmo em dias atípicos como feriados e finais de semana. O maior erro MAPE encontrado na previsão de uma curva diária foi de 3,54%, enquanto que o maior MAPE médio entre os dias da semana foi de 2,38%(segunda-feira), ambos dentro do limite considerado aceitável de 5%.Portanto, para uma previsão inicial, sem considerar muitos das variáveis que podem influenciar o consumo de carga elétrica em um sistema como temperatura local ou índices sócio-econômicos, o modelo proposto mostra-se eficaz e apto a ser utilizado como uma alternativa para a previsão séries de consumo de carga elétrica.

6.2 Trabalhos futuros

As ferramentas e modelos utilizados para a previsão de carga são objetos de estudos em todo o mundo e estão em constante evolução. Assim, são listadas algumas sugestões para o aperfeiçoamento do modelo proposto em possíveis trabalhos futuros:

- Inclusão da temperatura local como variável do modelo que pode influenciar o consumo de energia elétrica;
- Desenvolvimento de uma metodologia específica para a definição dos parâmetros das redes neurais artificiais e da Lógica Fuzzy;
- Aumento da quantidade de dados utilizados no treinamento da rede e criação da base de regras fuzzy para uma maior capacidade de previsão do modelo.

Referências Bibliográficas

1. OLIVEIRA, B. N. D. "Modelo de Comercialização de Energia pela Opção de Disponibilidade na Geração Termelétrica". Tese de Mestrado, COPPE/UFRJ. Rio de Janeiro. 2008.

2. SOUZA, R. C. "Modelos Estruturais para Previsão de Séries Temporais : Abordagens Clássica e Bayesiana". Instituto de Matemática Pura e Aplicada. Rio de Janeiro, p. 171. 1989.

3. MOREIRA DE ANDRADE, L. C.; FLAUZINO, R. A.; DA SILVA, I. N. "Sistemas Fuzzy Para Previsão De Demanda De Energia Elétrica No Curtíssimo Prazo". Brazilian Conference On Dynamics Control And Their Applications, Junho 2010.

4. GUIRELLI, C. R. "Previsao de Carga de Curto Prazo de Áreas Elétricas Através de Técnicas de Inteligência Artificial". Tese de Doutorado, USP. SÃO PAULO. 2006.

5. BROWN, R. G. "Smoothing, Forecasting and Prediction of Discrete Time Series". 12. ed. New Jersey: Prentice-Hall, 1963. 468 p.

6. CALDWELL, J. G. "Box-Jenkins Forecasting System - Reference Manual", 1, 1971. Disponivel em: <http://www.foundationwebsite.org/TIMESVol1TechnicalBackground.pdf>.

7. WHITTLE, P. "Hypothesis Testing in Time Series Analysis". 1. ed. Uppsala: Almqvist & Wiksells boktr, 1951. 120 p.

8. VALENÇA, M. "Fundamentos das redes neurais: exemplos em Java". 1. ed. Olinda, PE.: Livro Rápido, 2007.

9. DIAS, R. D. S. "Metodologia de Previsão de Carga de Longo Prazo de Energia Elétrica". Tese de Mestrado, Universidade Federal do Ceará. Fortaleza. 2010.

10. GOLDSCHMIDT, R.; PASSOS, E. "Data Mining: Um Guia Prático". 4. ed. Rio de Janeiro:

Campos, 2005.

11. ALTRAN, A. B. "Sistema Inteligente Para Previsão de Carga Multinodal em Sistemas Elétricos de Potência". Tese de Doutorado, Universidade Estadual Paulista Júlio de Mesquita Filho. Ilha Solteira. 2010.

12. RIBEIRO, C. V.; GOLDSCHMIDT, R. R.; CHOREN, R. "Métodos para Previsão de Séries Temporais e suas Tendências de Desenvolvimento". Tese de Mestrado, IME. Rio de Janeiro. 2009.

13. KOHONEN, T. Self-Organized Formation of Topologically Correct Feature Maps. Biological Cybernetics, Espoo, Finland., 1982. pp.59-69.

14. ZADEH, L. A. Fuzzy Sets. Information and Control, 1965. pp. 338-353.

15. JANG, J.-S. R. "Anfis: Adaptive-Network-Based Fuzzy Inference System". IEEE Transactions On Systems, Man, And Cybernetics, Vol 23, May/June 1993.

16. WANG, L.-X.; MENDEL, J. M. "Generating Fuzzy Rules By Learning From Examples". IEEE Transactions On Systems Man, And Cybernectics, Vol 22, 1992.

17. DUNN, J. C. "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". Journal of Cybernetic, 3, n. 3, 1973. pp. 32-57.

18. JACINTHO, L. F. D. O. "Redes Neuro-Fuzzy: Um Estudo de Caso em Diagnóstico de Alzheimer". Tese de Mestrado, Universidade Federal do ABC. Santo André. 2010.

19. SILVA, T. A. D. A. "Previsão De Cargas Elétricas Através De Um Modelo Híbrido De Regressão Com Redes Neurais". Tese de Mestrado, Universidade Estadual Paulista Júlio de Mesquita Filho. Ilha Solteira. 2012.

20. SILVA, I. V. G. D. "Previsão de Vento para Geração de Energia Elétrica". Tese de Mestrado, PUC-RJ. Rio de Janeiro. 2010.

21. MARA C. S. COELHO; AGUIRRE, L. A. "Previsão on-line de Carga a curto prazo utilizando Redes NeoFuzzy – Redes Yamakawa". Centro de Pesquisas Hidráulicas e Recursos Hídricos- UFMG. Belo Horizonte, p. 6. 2011.

22. PESSANHA, J. F. M.; JUSTINO, T. C.; MACEIRA, M. E. P. "Metodologia para filtragem

de registros de carga". XII SIMPÓSIO DE ESPECIALISTAS EM PLANEJAMENTO DA OPERAÇÃO E EXPANSÃO ELÉTRICA. Rio de Janeiro. 2012.

23. BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. "Neural Network Toolbox - User's Guide". MathWorks. Natick, MA. 2013.

24. SILVA, E.; OLIVEIRA, A. C. D. "Dicas Para A Configuração De Redes Neurais". NCE-UFRJ. Rio de Janeiro. 2010.

Anexo I

Este anexo apresenta os resultados obtidos durante os testes realizados para a definição dos parâmetros de entrada utilizados no modelo de previsão. A Tabela 1 apresenta os resultados obtidos durante a escolha do número de neurônios nas camadas intermediárias da rede neural MLP.

Tabela 1 - Índices de desempenho/Número de neurônios MLP

Número de Neurônios	MAD(MW)	MAPE(%)	RMSE(MW)
1-1	69,87	2,29	125,12
1-2	69,97	2,30	126,76
1-3	93,94	3,09	150,09
1-4	72,93	2,39	128,00
1-5	82,01	2,67	137,45
1-6	71,89	2,35	126,30
1-7	70,97	2,34	128,94
1-8	72,33	2,35	126,98
1-9	77,05	2,53	134,88
1-10	72,41	2,35	126,46
2-1	73,75	2,43	133,15
2-2	75,98	2,48	132,63
2-3	71,38	2,34	127,08
2-4	75,58	2,43	128,63
2-5	69,60	2,29	127,32
2-6	89,33	2,93	155,75
2-7	83,60	2,73	139,50
2-8	73,23	2,38	127,97
2-9	81,38	2,66	135,84
2-10	74,68	2,46	132,08
3-1	67,48	2,21	124,58
3-2	75,06	2,49	138,81
3-3	73,47	2,40	126,58
3-4	87,84	2,85	150,58
3-5	72,35	2,37	126,28
3-6	78,96	2,59	138,48
3-7	75,80	2,47	129,86
3-8	72,36	2,39	130,63
3-9	69,65	2,25	123,04
3-10	81,51	2,68	145,99

4-1	71,87	2,37	129,51
4-2	73,82	2,42	131,86
4-3	73,31	2,39	128,85
4-4	67,82	2,24	125,23
4-5	73,23	2,40	125,04
4-6	76,24	2,47	129,50
4-7	75,48	2,48	132,86
4-8	62,76	2,07	120,50
4-9	78,37	2,53	144,03
4-10	75,55	2,42	129,23
5-1	73,78	2,44	135,74
5-2	73,25	2,40	130,29
5-3	89,01	2,86	140,87
5-4	79,79	2,57	136,64
5-5	77,31	2,52	133,10
5-6	63,96	2,07	118,53
5-7	66,99	2,18	123,08
5-8	71,04	2,36	134,01
5-9	81,39	2,64	142,87
5-10	71,31	2,36	129,74
6-1	68,90	2,27	125,13
6-2	66,24	2,18	124,90
6-3	74,15	2,43	131,09
6-4	74,36	2,46	134,24
6-5	72,94	2,40	134,65
6-6	64,57	2,10	115,55
6-7	75,34	2,44	129,58
6-8	77,11	2,53	141,37
6-9	98,48	3,28	184,40
6-10	71,84	2,37	130,25
7-1	71,57	2,32	124,47
7-2	72,82	2,34	128,03
7-3	75,87	2,49	133,39
7-4	77,20	2,55	137,59
7-5	85,61	2,76	144,91
7-6	69,35	2,26	126,91
7-7	62,87	2,07	124,34
7-8	88,06	2,92	158,64
7-9	79,16	2,54	134,07
7-10	73,66	2,41	129,87
8-1	72,91	2,39	125,09
8-2	67,42	2,20	123,95
8-3	67,70	2,19	121,44
8-4	72,88	2,38	130,48

8-5	91,73	2,96	170,93
8-6	68,11	2,22	120,61
8-7	86,60	2,81	151,53
8-8	75,36	2,46	128,80
8-9	96,48	3,09	171,74
8-10	73,49	2,40	129,87
9-1	74,67	2,42	127,12
9-2	75,71	2,46	133,42
9-3	72,39	2,36	127,20
9-4	74,88	2,45	131,92
9-5	78,99	2,56	136,20
9-6	80,05	2,58	131,88
9-7	71,92	2,34	128,34
9-8	76,52	2,53	137,02
9-9	72,78	2,40	134,82
9-10	91,00	2,93	147,17
10-1	68,80	2,27	125,62
10-2	84,40	2,73	142,57
10-3	77,35	2,54	133,14
10-4	71,22	2,34	128,44
10-5	85,50	2,82	146,73
10-6	80,62	2,62	133,97
10-7	75,85	2,53	145,37
10-8	63,08	2,07	119,73
10-9	76,33	2,46	127,60
10-10	72,67	2,36	131,80

A Tabela 2 apresenta os resultados obtidos durante os testes para a definição do número de neurônios no Mapa de Kohonen do modelo de previsão.

Tabela 2 - Índices de desempenho/Número de neurônios do Mapa de Kohonen

Número de Neurônios	MAD(MW)	MAPE(%)	RMSE(MW)
1-1	226,37	6,95	251,54
1-2	160,42	4,97	184,24
1-3	89,67	2,85	119,84
1-4	83,79	2,69	114,29
1-5	58,91	1,88	74,20
1-6	70,58	2,27	87,79
1-7	71,99	2,33	88,23
1-8	65,27	2,09	80,58
1-9	72,35	2,34	88,74

1-10	57,69	1,84	73,37
2-1	59,61	1,91	74,19
2-2	58,08	1,86	72,13
2-3	53,56	1,72	68,55
2-4	54,50	1,75	69,13
2-5	54,29	1,75	68,33
2-6	53,66	1,72	67,51
2-7	53,71	1,73	67,59
2-8	51,29	1,62	63,32
2-9	50,86	1,62	65,53
2-10	53,21	1,69	66,70
3-1	53,87	1,72	70,81
3-2	51,42	1,63	63,58
3-3	51,44	1,63	63,42
3-4	56,03	1,80	69,75
3-5	54,86	1,77	68,75
3-6	51,22	1,63	64,49
3-7	50,45	1,61	62,96
3-8	51,43	1,64	64,94
3-9	51,67	1,65	64,59
3-10	51,30	1,64	63,97
4-1	54,52	1,77	70,27
4-2	50,27	1,64	63,38
4-3	51,33	1,49	64,21
4-4	50,54	1,48	63,18
4-5	50,52	1,62	63,10
4-6	54,57	1,77	67,02
4-7	51,83	1,66	64,07
4-8	52,79	1,70	64,10
4-9	49,12	1,56	60,25
4-10	56,14	1,81	69,62
5-1	54,61	1,60	67,12
5-2	51,24	1,63	62,34
5-3	47,28	1,62	58,29
5-4	46,89	1,48	58,43
5-5	56,48	1,83	69,34
5-6	52,81	1,70	65,29
5-7	53,89	1,73	65,29
5-8	51,48	1,65	62,24
5-9	54,71	1,76	65,57
5-10	55,21	1,77	65,47
6-1	56,43	1,82	69,13
6-2	54,93	1,77	66,01
6-3	54,72	1,77	65,89

6-4	55,86	1,80	67,88
6-5	53,48	1,71	70,31
6-6	55,33	1,79	67,44
6-7	54,64	1,76	66,21
6-8	57,24	1,85	69,48
6-9	57,77	1,87	70,24
6-10	49,31	1,56	62,28
7-1	49,27	1,56	64,60
7-2	49,15	1,56	64,17
7-3	61,38	1,98	90,91
7-4	49,22	1,56	64,17
7-5	50,01	1,58	65,58
7-6	51,55	1,63	66,78
7-7	50,01	1,58	65,71
7-8	50,79	1,62	65,27
7-9	52,47	1,67	65,44
7-10	61,59	1,97	93,44
78-1	49,97	1,58	67,57
8-2	48,98	1,55	64,17
8-3	52,36	1,67	68,20
8-4	56,02	1,81	79,11
8-5	57,59	1,85	80,92
8-6	57,05	1,85	80,12
8-7	49,27	1,56	67,13
8-8	58,60	1,91	86,76
8-9	58,18	1,89	86,51
8-10	264,79	8,19	842,18
9-1	57,77	1,88	85,52
9-2	264,86	8,19	842,50
9-3	59,53	1,93	84,99
9-4	50,23	1,60	63,40
9-5	53,97	1,73	67,39
9-6	57,85	1,88	84,63
9-7	375,75	11,19	1063,65
9-8	49,98	1,59	67,90
9-9	58,19	1,88	78,67
9-10	374,76	11,14	1063,55
10-1	48,90	1,55	66,11
10-2	56,13	1,82	81,00
10-3	698,90	20,99	1485,87
10-4	49,49	1,56	67,11
10-5	373,76	11,11	1063,49
10-6	59,59	1,94	86,25
10-7	374,56	11,15	1063,37

10-8	50,33	1,60	62,99
10-9	49,31	1,57	61,10
10-10	52,00	1,66	64,62

A Tabela 3 apresenta os resultados obtidos durante os testes para a definição da constante de fuzzyficação m.

Tabela 3 - Índices de desempenho/Constante de fuzzyficação m

m	MAD(MW)	MAPE(%)	RMSE(MW)
0,1	377,15	11,40	434,26
0,2	388,04	11,71	449,40
0,3	378,19	11,41	440,35
0,4	410,22	12,33	483,10
0,5	391,95	11,79	465,48
0,6	398,33	11,96	477,03
0,7	430,98	12,92	515,00
0,8	407,18	12,21	494,05
0,9	407,18	12,22	496,26
1,0	-	-	-
1,1	58,48	1,89	74,06
1,2	59,40	1,92	76,24
1,3	56,19	1,82	71,28
1,4	57,05	1,84	71,49
1,5	57,24	1,84	71,28
1,6	52,58	1,69	68,77
1,7	54,81	1,77	72,11
1,8	51,15	1,65	67,27
1,9	50,76	1,62	64,74
2,0	49,98	1,60	64,81
2,1	56,64	1,79	70,28
2,2	47,42	1,51	60,78
2,3	50,28	1,60	63,58
2,4	51,75	1,67	66,55
2,5	50,28	1,60	63,67
2,6	52,55	1,69	67,13
2,7	52,74	1,70	67,44
2,8	58,26	1,84	73,41
2,9	48,80	1,56	63,60
3,0	55,37	1,79	71,17
3,1	58,83	1,77	74,83
3,2	55,59	1,64	71,25

3,3	55,66	1,49	71,59
3,4	53,41	1,48	69,67
3,5	60,14	1,88	77,68
3,6	55,10	1,77	72,14
3,7	61,37	1,91	79,88
3,8	63,82	1,99	81,79
3,9	64,57	2,03	82,38
4,0	58,19	1,86	76,35
4,1	68,85	1,60	87,48
4,2	61,64	1,63	80,76
4,3	64,37	1,62	83,74
4,4	63,80	2,05	83,75
4,5	65,09	2,08	85,05
4,6	68,68	2,21	89,18
4,7	68,76	2,20	89,66
4,8	70,32	2,26	91,40
4,9	73,04	2,34	94,35
5,0	72,03	2,30	93,33
5,1	90,61	2,83	110,43
5,2	78,03	2,50	100,18
5,3	77,73	2,48	100,56
5,4	80,00	2,56	102,88
5,5	82,40	2,63	105,47
5,6	83,59	2,66	106,62
5,7	103,84	3,24	124,77
5,8	82,04	2,59	102,29
5,9	84,70	2,66	104,83
6,0	90,79	2,88	114,48
6,1	90,85	2,89	115,22
6,2	93,60	2,97	117,86
6,3	95,29	3,02	119,67
6,4	97,18	3,09	121,46
6,5	120,72	3,75	142,32
6,6	99,73	3,16	124,28
6,7	102,21	3,24	126,76
6,8	103,47	3,28	128,00
6,9	103,80	3,29	128,56
7,0	106,63	3,37	131,33
7,1	107,39	3,39	131,81
7,2	107,60	3,40	132,28
7,3	109,68	3,47	134,75
7,4	110,73	3,49	135,51
7,5	113,78	3,59	138,81
7,6	113,20	3,57	138,06

7,7	115,31	3,64	140,54
7,8	117,19	3,69	141,81
7,9	117,14	3,69	142,01
8,0	120,38	3,79	145,36
8,1	121,58	3,83	146,48
8,2	122,83	3,86	147,75
8,3	123,76	3,89	148,63
8,4	125,31	3,94	150,31
8,5	126,50	3,97	151,47
8,6	130,01	4,08	154,74
8,7	128,39	4,03	153,29
8,8	128,99	4,05	153,76
8,9	129,79	4,07	154,89
9,0	149,37	4,62	171,52
9,1	131,67	4,10	154,41
9,2	133,52	4,18	158,08
9,3	145,28	4,53	168,78
9,4	134,78	4,23	159,87
9,5	159,95	4,96	182,06
9,6	154,71	4,79	176,86
9,7	150,57	4,69	173,75
9,8	139,08	4,36	163,95
9,9	140,41	4,40	165,12
10,0	141,06	4,41	165,73