



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Campus de Ilha Solteira

ALINE JARDIM AMORIM



**Desenvolvimento de Uma Rede Neuro-*Fuzzy* Para a Previsão da
Carga**

Ilha Solteira
2015



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Campus de Ilha Solteira

ALINE JARDIM AMORIM

A large, semi-circular graphic composed of several overlapping, light blue, semi-transparent geometric shapes (triangles and quadrilaterals) arranged in a pattern that resembles a network or a stylized globe. The shapes are outlined in white and have a slight gradient, giving them a three-dimensional appearance.

**Desenvolvimento de Uma Rede Neuro-*Fuzzy* Para a Previsão da
Carga**

Carlos Roberto Minussi

Orientador

Dissertação submetida à Faculdade de
Engenharia - UNESP – Campus de Ilha
Solteira, para obtenção do título de
Mestre em Engenharia Elétrica.
Área de Conhecimento: Automação.

Ilha Solteira
2015

FICHA CATALOGRÁFICA

Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

A524d Amorim, Aline Jardim.
Desenvolvimento de uma rede *neuro-fuzzy* para a previsão da carga / Aline Jardim Amorim. -- Ilha Solteira: [s.n.], 2015
98 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2015

Orientador: Carlos Roberto Minussi
Inclui bibliografia

1. Redes neurais. 2. Lógica fuzzy. 3. Sistema elétricos de potência.
4. Previsão de carga.



UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE ILHA SOLTEIRA
FACULDADE DE ENGENHARIA DE ILHA SOLTEIRA



CERTIFICADO DE APROVAÇÃO


TÍTULO: Desenvolvimento de Uma Rede Neuro-Fuzzy Para a Previsão da Carga


AUTORA: ALINE JARDIM AMORIM

ORIENTADOR: Prof. Dr. CARLOS ROBERTO MINUSSI

Aprovada como parte das exigências para obtenção do Título de Mestre em Engenharia Elétrica ,
Área: AUTOMAÇÃO, pela Comissão Examinadora:


Prof. Dr. CARLOS ROBERTO MINUSSI
Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira


Profa. Dra. MARA LÚCIA MARTINS LOPES
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira


Prof. Dr. JOSÉ GUILHERME MAGALINI SANTOS DECANINI
Câmpus de Presidente Epitácio / Instituto Federal de São Paulo - Ifsp

Data da realização: 13 de abril de 2015.

“Escrever é sacudir o sentido do mundo.”
Roland Barthes

Agradecimentos

Venho agradecer primeiramente a Deus por ter me dado o dom da vida. Agradeço aos meus pais Izabel Jardim da Silva e Odorico Santos da Silva pela dedicação e a educação que me foi dada; ao meu esposo Estelio pela compreensão e apoio nessa jornada; aos meus queridos e amados filhos Júlia e Gabriel os quais são a minha inspiração.

Agradeço ao meu Orientador Carlos Roberto Minussi e a colaboradora Mara Lúcia Lopes pelo apoio, incentivo, e por ter dado as ferramentas necessárias para eu poder prosseguir.

Agradeço a minha companheira de Laboratório Thays e ao Marcos pela mão amiga nas horas de dificuldade.

Agradeço ao meu amigo Uiliam pela sua disponibilidade em me ajudar em momentos de completa escuridão, pela paciência e compreensão.

Agradeço os funcionários do Departamento de Engenharia Elétrica, em especial ao técnico do Laboratório de Computação Deoclécio Mitsuiti Kosaka. Aos funcionários da Biblioteca da FEIS pela orientação e paciência na elaboração da dissertação de acordo com as normas.

À CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo apoio financeiro, que viabilizou a realização deste trabalho.

Resumo

O estudo do processo de previsão constitui-se numa rotina de extrema importância para o funcionamento e a operação dos sistemas elétricos de potência. Para isso, há a necessidade de se prever os acontecimentos futuros visando o fornecimento de energia aos consumidores, de forma segura, eficaz e econômica. Neste sentido, nesta pesquisa, apresenta-se uma metodologia, baseada nas redes neurais, para resolver o problema de previsão de carga. São propostas melhorias na rede neural, via o emprego do algoritmo retropropagação (*backpropagation*), por meio da adaptação dos parâmetros de inclinação e de translação da função sigmoide (função ativação da rede neural). Além disto, é desenvolvido um controlador *fuzzy* com o propósito de ajustar a taxa de treinamento durante o processo de convergência. Visando testar a metodologia proposta, apresenta-se um estudo de previsão da carga considerando-se uma base de dados de uma companhia do setor elétrico.

Palavras-Chave: Redes neurais artificiais. Lógica *Fuzzy*. Sistemas elétricos de potência. Previsão de carga.

Abstract

The prediction process study is a very important task for the operation of the electrical power systems. Therefore, it is necessary to predict the future behavior to provide energy with security, efficiency and economy to the users. This work presents a methodology based on neural networks to solve the problem of load forecasting. The proposal is to improve the neural network using the backpropagation algorithm adapting the inclination and translation parameters of the sigmoid function (activation function of the neural network). Besides, it is developed a *fuzzy* controller to adjust the training during the convergence process. To test the proposed methodology, it is presented a load forecasting study considering a database from an electrical energy company.

keywords: Neural networks. *Fuzzy* logic. Electric power systems. Load forecasting.

LISTA DE FIGURAS

Figura 1-	Série temporal estacionária	22
Figura 2-	Série temporal não-estacionária	23
Figura 3-	Componentes de um neurônio biológico	35
Figura 4-	Componentes de um neurônio artificial	37
Figura 5-	Representação gráfica da função degrau	38
Figura 6-	Representação gráfica da função degrau bipolar ou função sinal	39
Figura 7-	Representação gráfica da função rampa simétrica	39
Figura 8-	Representação gráfica da função logística	40
Figura 9-	Representação gráfica da função tangente hiperbólica	41
Figura 10-	Representação gráfica da função gaussiana	41
Figura 11-	Representação gráfica da função linear	42
Figura 12-	Disposição das redes neurais artificiais	43
Figura 13-	Funções de pertinência para o erro global eg	58
Figura 14-	Funções de pertinência para a variação do erro global Δeg	58
Figura 15-	Funções de pertinência para a ação de controle $\Delta \psi$	59
Figura 16-	Modelo de previsão de carga global via redes neurais	62
Figura 17-	Estratégia de previsão via deslocamento de janela	63
Figura 18-	Modelo de previsão de carga modular via redes neurais	64
Figura 19-	Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema previsor Global (d) treinados com o algoritmo RP convencional	68
Figura 20-	Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema previsor Global (d) treinados com o algoritmo RP convencional via controlador <i>fuzzy</i>	70

Figura 21- Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema previsor Global (d) treinados com o algoritmo RP convencional via controlador *fuzzy* e as adaptações dos parâmetros de inclinação e de translação da função de ativação 72

Figura 22- Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema previsor Global (d) treinados com o algoritmo RP convencional via controlador *fuzzy* e as adaptações dos parâmetros de inclinação e de translação da função de ativação 74

LISTA DE TABELAS

Tabela 1-	Regras do controlador nebuloso	59
Tabela 2-	Parâmetros utilizados no treinamento com o algoritmo RP convencional	67
Tabela 3-	Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema previsor global treinado com o algoritmo RP convencional	68
Tabela 4-	Parâmetros utilizados no treinamento com o algoritmo RP via controlador <i>fuzzy</i>	69
Tabela 5-	Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema previsor global treinado com o algoritmo RP convencional via controlador <i>fuzzy</i>	70
Tabela 6-	Parâmetros utilizados no treinamento com o algoritmo RP com a adaptação da inclinação e da translação da função sigmoide	71
Tabela 7-	Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema previsor global treinado com o algoritmo RP com as adaptações da inclinação e da translação da função sigmoide	72
Tabela 8-	Parâmetros utilizados no treinamento com o algoritmo RP com a adaptação da inclinação e da translação da função sigmoide e o controlador <i>fuzzy</i>	73
Tabela 9-	Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema previsor global treinado com o algoritmo RP convencional via controlador <i>fuzzy</i> e as adaptações da inclinação e da translação da função sigmoide	74
Tabela 10-	Comparando os valores de MAPE do algoritmo RP convencional com os do algoritmo com o controlador <i>fuzzy</i> e os do algoritmo com o controlador <i>fuzzy</i> e as adaptações da inclinação e da translação da função de ativação	75
Tabela 11-	Comparando os valores de Erro Máximo do algoritmo RP convencional com os do algoritmo com o controlador <i>fuzzy</i> e os do algoritmo com o controlador <i>fuzzy</i> e as adaptações da inclinação e da translação da função de ativação	76
Tabela 12-	Comparando os valores do número de épocas do algoritmo RP convencional com os do algoritmo com o controlador <i>fuzzy</i> e os do algoritmo com o controlador <i>fuzzy</i> e as adaptações da inclinação e da	77

	translação da função de ativação	
Tabela 13-	Comparando os valores de tempo de processamento (s) do algoritmo RP convencional com os do algoritmo com o controlador <i>fuzzy</i> e os do algoritmo com o controlador <i>fuzzy</i> e as adaptações da inclinação e da translação da função de ativação	77
Tabela 14-	Comparação entre metodologias de Altran (2010) e a presente pesquisa	79
Tabela 15-	Comparando os resultados da tese de Altran (2010) utilizando a função sigmoide e o resultado que foi obtido nesse presente trabalho.	79

LISTA DE SÍMBOLOS

$y(t)$	carga elétrica
a_0, a_1, \dots, a_n	coeficientes de regressão
q_1, q_2, \dots, q_n	variáveis explanatórias
$a(t)$	variável aleatória com média zero e variância constante
$\varphi(t)$	vetor de coeficiente
$f(t)$	vetor da função de ajuste para o processo
$\varepsilon(t)$	ruído branco
T	operador transposto
$A(t)$	matriz de transição
$w(t)$	ruído branco com matriz de covariância $Q(t)$
$z(t)$	vetor de medição t
$B(t)$	matriz que correlaciona $y(t)$ e $z(t)$
$v(t)$	erro medido (ruído com a matriz de covariância $S(t)$)
$k(t)$	fator de combinação
θ_0	termo constante
$a(t)$	séries aleatórias
ϕ_1, \dots, ϕ_p	coeficientes <i>auto-regressive</i>
r	ordem do modelo
$\theta_1, \dots, \theta_q$	coeficientes de média móvel
w_{ij}	peso do neurônio
x_i	entrada do neurônio
s_i	saída intermediária do neurônio
m	agente modificador
A	conjunto nebuloso
i	valor do conjunto que tem um valor de pertinência μ_i
n	número de regras nebulosas
ε_i	erro instantâneo de cada neurônio alocado na última camada
d_i	saída desejada do i -ésimo elemento da última camada da rede
y_i	saída do i -ésimo elemento da última camada da rede
ns	número de neurônios da última camada da rede
γ	taxa de treinamento da rede neural

h	índice de iteração
$\nabla_i(h)$	gradiente do erro quadrático com relação aos pesos do neurônio avaliado em h
V_i	vetor contendo os pesos do neurônio i
ρ	constante que determina a translação da curva y_i
λ	constante que determina a inclinação da curva y_i
$sgm'(s_i)$	derivada da função sigmoide com relação a s_i
$R(j)$	conjunto dos índices dos elementos que se encontram na fileira seguinte à fileira do elemento i e que estão interligados ao elemento i
$\nabla_i^\lambda(h)$	gradiente do erro quadrático com relação a inclinação do neurônio i
$\nabla_i^\rho(h)$	gradiente do erro quadrático com relação a translação do neurônio i
λ_i	vetor contendo a inclinação do neurônio i
σ_i^λ	derivada da função sigmoide com relação a λ_i
σ_i^ρ	derivada da função sigmoide com relação a ρ_i
w_{ij}	peso correspondente à interligação entre o i -ésimo e j -ésimo neurônio
εg	erro global da rede neural
np	número de vetores padrão da rede
$\Delta \varepsilon g$	variação do erro global
k	ação de controle exercida sobre o sistema
α^q	número positivo arbitrário
$\Delta \psi^q$	acréscimo procedente do controlador nebuloso
$L(h)$	valor da carga real referente à hora h
$\underline{L}(h)$	valor da carga estimada pela rede referente à hora h
N	número de horas total
m	dimensão do vetor X
$L(h - p)$	valor da carga p horas anterior à hora corrente h
$L(h + 1)$	valor da carga elétrica da hora subsequente à hora corrente h
t	vetor referente ao tempo dos dados históricos (dia da semana, hora, etc.) representados entre $[0,1]$
$X(h)$	vetor de entrada da rede neural correspondente à hora h
$Y(h)$	estímulo aplicado à saída da rede neural correspondente à hora $h + 1$

Sumário

1	INTRODUÇÃO	16
1.1	ORGANIZAÇÃO DO TRABALHO	18
2	PREVISÕES DE CARGAS ELÉTRICAS.....	20
2.1	IMPORTÂNCIA DA PREVISÃO DE CARGA.....	20
2.2	CARGAS ELÉTRICAS	20
2.3	HORIZONTES DA PREVISÃO.....	24
2.4	TÉCNICAS PARA A PREVISÃO DE CARGAS ELÉTRICAS	25
2.4.1	Métodos estatísticos	25
2.4.1.1	<i>Método de regressão linear ou múltipla</i>	<i>26</i>
2.4.1.2	<i>Alisamento exponencial</i>	<i>26</i>
2.4.1.3	<i>Espaço de estado e Filtro de Kalman.....</i>	<i>27</i>
2.4.1.4	<i>Série temporal de Box-Jenkins</i>	<i>28</i>
2.5	SISTEMAS ESPECIALISTAS	31
2.6	COMENTÁRIOS	31
3	REDES NEURAIS ARTIFICIAIS.....	33
3.1	NOTA HISTÓRICA.....	33
3.2	ÁREAS DE APLICAÇÃO	34
3.3	NEURÔNIO BIOLÓGICO	35
3.4	MODELO NEURAL ARTIFICIAL.....	36
3.4.1	Funções de ativação parcialmente diferenciáveis.....	38
3.4.2	Funções de ativação totalmente diferenciáveis	40
3.5	ESTRUTURA DAS REDES NEURAIS ARTIFICIAIS	42
3.6	TREINAMENTO DA REDE NEURAL.....	44
3.6.1	Treinamento supervisionado	44
3.6.2	Treinamento não supervisionado	45
3.6.3	Treinamento com reforço	45
3.7	TIPOS DE REDES NEURAIS ARTIFICIAIS	45
3.8	COMENTÁRIOS	47
4	ALGORITMO RETROPROPAGAÇÃO	48
4.1	PROBLEMAS GERADOS PELO ALGORITMO RETROPROGAÇÃO	52
4.2	MELHORIAS NO ALGORITMO RETROPROPAGAÇÃO.....	53
4.2.1	Adaptação dos parâmetros da rede	53

4.3	ANÁLISE DE PRECISÃO	60
4.4	COMENTÁRIOS	60
5	METODOLOGIA DESENVOLVIDA PARA PREVISÃO DE CARGA ELÉTRICA	61
5.1	METODOLOGIA DESENVOLVIDA.....	61
5.2	COMENTÁRIOS	65
6	RESULTADOS OBTIDOS.....	66
6.1	PARÂMETROS UTILIZADOS NA REDE	66
6.2	RESULTADOS	66
6.2.1	Análise numérica e gráfica para os sistemas previsores com o algoritmo RP convencional.....	67
6.2.2	Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação via controlador fuzzy	69
6.2.3	Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação com a adaptação da inclinação e da translação da função sigmoide.....	71
6.2.4	Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação via controlador fuzzy e adaptação da inclinação e da translação da função sigmoide	73
6.3	COMPARANDO RESULTADOS.....	75
6.4	COMPARANDO METODOLOGIAS	78
6.5	COMENTÁRIOS	80
7	CONSIDERAÇÕES FINAIS.....	82
7.1	SUGESTÕES PARA TRABALHOS FUTUROS.....	83
	REFERÊNCIAS	83
	APÊNDICE A - LÓGICA FUZZY.....	88
	APÊNDICES B- DADOS HISTÓRICOS DE CARGA	93
	APÊNDICE B1 – Sistema Previsor Global.....	94
	APÊNDICE B2 – Sistema Previsor Local – Módulo 1	95
	APÊNDICE B3 – Sistema Previsor Local – Módulo 2.....	96
	APÊNDICE B4 – Sistema Previsor Local – Módulo 3.....	97
	APÊNDICE C - TRABALHO SUBMETIDO.....	98

Capítulo 1

1 INTRODUÇÃO

A previsão de carga caracteriza-se como uma ferramenta de grande relevância no ambiente de sistemas elétricos de potência, seja no planejamento, seja na operação. Para tanto, tem-se procurado cada vez mais por metodologias que fornecem ferramentas computacionais e tecnológicas que auxiliem na realização da previsão de forma eficaz e precisa. Na literatura encontram-se inúmeras técnicas para a obtenção da previsão de carga, tais como: regressão linear múltipla ou simples, alisamento exponencial, estimação de estado, filtro de Kalman. Modelo ARIMA de Box & Jenkins (BOX; JENKINS, 1976). Esses métodos são técnicas que precisam de uma modelagem prévia da carga para posteriormente serem aplicadas. Para modelagem da carga é necessário se ter à disposição informações climáticas, tais como: dias nublados, velocidade do vento, variações de temperatura, enfim os efeitos atmosféricos que podem influenciar na previsão. Além disso, outro fator importante a ser considerado é a identificação dos dias atípicos, tais como feriados, sábados e domingos. Assim, quanto maior for o número de informações, mais chance de se ter uma previsão precisa.

Além dessas técnicas de previsão apresentadas, existe outra metodologia que tem obtido espaço nesse horizonte que são as Redes Neurais Artificiais (RNAs). Trata-se de um recurso alternativo para a resolução de problemas de previsão de cargas elétricas. As RNAs também se destacam pela sua interdisciplinaridade, sendo utilizada por diversos pesquisadores da psicologia, física, computação, neurofisiologia, engenharia, entre outras áreas do conhecimento humano. O princípio básico de uma RNA é uma estrutura chamada neurônio artificial, que é baseado no neurônio biológico, no qual cada elemento do neurônio artificial é comparado com os elementos do neurônio biológico. As RNAs são computacionalmente concepções de *software* e/ou *hardware* para emularem o funcionamento do cérebro.

O surgimento dos estudos relacionados à RNA foi, em 1943, com a proposta apresentada por McCulloch & Pitts. Posteriormente, passou por um processo de evolução e estagnação por um período, retomando-se o interesse no final da década de 70 e início da década de 80. Alguns pesquisadores que contribuíram para essa retomada foram, principalmente, Werbos (WERBOS, 1974), Grossberg (GROSSBERG, 1976), Kohonen (KOHONEN, 1987) e Hopfield (HOPFIELD, 1982). O avanço nas tecnologias e o

melhoramento da capacidade dos computadores, também contribuíram para o atual sucesso das redes neurais artificiais.

Dentre esses fatores que impulsionaram o retorno do estudo de RNA foi a proposição de novos algoritmos. Neste caso, destaca-se o algoritmo Retropropagação (*Backpropagation*) (WERBOS, 1974). Esse algoritmo tem por funcionalidade o aprendizado supervisionado de RNAs dispostas em fileiras de neurônios, em paralelismo, e funções de ativação contínuas. Esse algoritmo é considerado um dos melhores em relação à precisão, porém o tempo de execução deixa-se a desejar pelo fato de ter uma convergência muito lenta e chegando, às vezes, ao ponto de não haver convergência caindo em “morros e vales” chamados mínimos locais.

Pesquisadores têm buscado soluções para esse problema (lentidão e ausência de convergência). Por meio de proposições baseadas na adaptação da taxa de treinamento dos pesos, desenvolvimento de controlador *fuzzy* (para realizar o ajuste ótimo da taxa de treinamento) (LOPES, 2000) e na utilização do termo momento que promove movimentos em direção fixa fazendo com que o mínimo seja alcançado rapidamente diminuindo assim o tempo de convergência da rede e possibilitando que a solução não caia em um mínimo local. Utiliza-se também uma função exponencial decrescente para estabilizar a atuação do regulador no tempo de treinamento inicial e evitar a instabilidade no processo de convergência (LOPES, 2005), tornando assim o método mais competitivo do que outros métodos convencionais.

Outra melhoria na metodologia foi proposta por Stamatis (STAMATIS et al., 1999) que corresponde a implementação de um método para adaptar os parâmetros da inclinação e da translação da função de ativação (função sigmoide), por meio do método do gradiente descendente, fazendo com que a cada iteração tenha capacidade de encontrar uma melhor solução possibilitando que não caia em um mínimo local acelerando, assim, o processo de convergência da rede.

Essa ferramenta de apoio, que auxilia na minimização do tempo de execução do processo de convergência, é o controlador *fuzzy*. Trata-se de um processo baseado na lógica *fuzzy* que é uma teoria para se tratar de incertezas e imprecisões (ZADEH, 1965).

Esta pesquisa de mestrado tem por objetivo realizar a previsão global e multimodal definida como um sistema inteligente que considera vários pontos da rede elétrica (tais como:

subestação, barramentos, etc.) durante a realização da previsão. Neste sentido, utilizaram-se como ferramentas para realizar essa previsão as redes neurais *feedforward* compostas por vários módulos, via o emprego do algoritmo retropropagação, no qual foram introduzidas importantes melhorias por meio da adaptação dos parâmetros de inclinação e de translação da função sigmoide por meio do método do gradiente descendente, do implemento de um controlador *fuzzy* para ajustar a taxa de treinamento durante o processo de convergência e da junção das duas proposta, ou seja, as adaptações da inclinação e da translação e o controlador *fuzzy*. Visando testar a metodologia proposta, apresenta-se um estudo de previsão da carga multimodal de curto prazo (24 horas) considerando-se uma base de dados de uma companhia do setor elétrico da Nova Zelândia (Electricity Commission da Nova Zelândia). Os dados de cargas foram considerados em um intervalo de 15 dias, no período de 15 de outubro a 29 de outubro de 2007, sendo a carga utilizada em hora em hora, ou seja, 24 horas de cada dia, totalizando assim uma entrada de 360 dados para cada um dos sistemas.

1.1 ORGANIZAÇÃO DO TRABALHO

Este texto está dividido em sete capítulos: previsão de cargas elétricas, redes neurais artificiais, algoritmo Retropropagação, metodologia desenvolvida para previsão de carga elétrica, resultados obtido e finalizando com as considerações finais e sugestões para futuros trabalho.

No capítulo 2 é abordada a teoria de previsão de carga elétrica no qual é descrito as características das cargas elétricas, os fatores que influenciam na modelagem, a classificação dos tipos de carga elétrica e finalizando com as técnicas de previsão de carga existentes na literatura.

No capítulo 3 apresenta-se um breve histórico do surgimento do estudo de redes neurais. São abordados, também, os tipos de redes, o funcionamento de um neurônio biológico comparado com o neurônio artificial, sua estrutura, os algoritmos e os tipos de treinamento envolvidos na teoria de redes neurais.

No capítulo 4 é elaborado um minucioso estudo sobre o algoritmo Retropropagação sendo retratadas as suas duas formulações: (1) algoritmo retropropagação com o controlador *fuzzy* para adaptar a taxa de treinamento, (2) e a adaptação da inclinação e da translação da função sigmoide, usando-se o método do gradiente descendente. São apresentados também os

tipos de convergência da rede, os problemas gerados pelo algoritmo e são propostas as melhorias via adaptação da translação e da inclinação da função sigmoide.

No Capítulo 5 é apresentada a metodologia desenvolvida para previsão de carga elétrica em que são abordados os tipos de dados de entrada utilizados pela rede neural, tanto no sistema previsor global quanto no local.

No capítulo 6 são apresentados os resultados obtidos pela metodologia desenvolvida, por meio de gráficos e de cálculo de erros para medir a precisão dos cálculos: a primeira conhecida por sua abreviação MAPE (erro percentual médio absoluto) e a segunda a do erro máximo.

No capítulo 7 finaliza-se o trabalho apresentando os comentários e as considerações finais sobre o trabalho desenvolvido e as sugestões para futuras pesquisas.

No apêndice A é apresentada a teoria dos conjuntos *fuzzy*, as operações dos conjuntos, a estrutura do controlador, o processo de controle e finalizando com algumas possibilidades de áreas de aplicação.

No apêndice B apresentam-se os dados de cargas elétricas utilizadas como banco de dados sendo composto pela carga global (Apêndice B1), módulo 1 (Apêndice B2), módulo 2 (Apêndice B3) e módulo 3 (Apêndice B4). E finalizando, no apêndice C é relacionado o artigo submetido a um evento científico.

Capítulo 2

2 PREVISÕES DE CARGAS ELÉTRICAS

Nesse capítulo serão abordados os principais conceitos envolvidos nas previsões de carga elétrica, fatores que influenciam na previsão (climáticos, econômico e temporal), horizontes de previsão (longo, médio, curto e curtíssimo prazo) e os tipos de métodos utilizados para realização dessas previsões.

2.1 IMPORTÂNCIA DA PREVISÃO DE CARGA

A sociedade moderna enfrenta um momento muito delicado em termos de consumo de energia elétrica, passando por longos períodos sem chuva fazendo com que os rios fiquem abaixo do volume esperado para a geração de energia elétrica. Logo, o consumo de energia deve ser consciente, racional e sem desperdícios. Com o intuito de reduzir desperdícios, houve um grande esforço visando realizar estudos sobre a previsão de carga no contexto da geração, da transmissão e da distribuição. Outros fatores são levados em conta já que a energia elétrica é vendida e o mercado de eletricidade quer obter lucros ao invés de prejuízos. Fazendo-se uma previsão com o menor erro possível, o mercado de eletricidade irá maximizar seus lucros e minimizar os prejuízos. Existem diversas áreas de aplicação de previsão de carga, tais como: planejamento de manutenção, segurança de sistemas, geração econômica de potência, controle de operação e distribuição, entre outras.

2.2 CARGAS ELÉTRICAS

Para realizar a previsão de carga elétrica é necessário conhecer as características do consumidor, pois é ele que caracteriza a demanda de energia elétrica a ser gerada. A demanda de energia elétrica refere-se à quantidade total requerida pelos consumidores e estes consumidores são divididos em três categorias: residencial, industrial e comercial. O consumidor residencial é expresso como o conjunto de residências ligadas às redes de distribuição, nesta categoria existem fatores que influenciam diretamente na demanda de energia, tais como: taxa de natalidade, mortalidade, rendas das famílias, grau de educação, consumo médio por residência e temperatura, que com o aumento desta o consumo também aumenta devido à utilização dos equipamentos eletrodomésticos tais como: ar-condicionado,

ventiladores, abertura da porta da geladeira. O consumidor industrial compõe o setor produtivo e este é totalmente influenciado pela economia nacional, este tipo de consumidor é conhecido pelo seu alto consumo de energia elétrica, pois a energia elétrica é fundamental no processo de produção. O consumidor comercial está relacionado ao conjunto de unidades comerciais fornecedoras de serviços que, por sua vez, também é influenciado pela economia. Possui um consumo razoável de energia elétrica pelo fato de possuir refrigeradores, ar-condicionados, iluminações, etc. E a composição desses diferentes tipos de consumidores caracterizará a demanda de carga total a ser gerada pelo sistema e esta apresenta uma tendência que pode ser prevista estatisticamente.

Na literatura encontram-se vários métodos de previsão no intuito de realizar uma previsão com certo grau de confiabilidade e precisão. Os métodos necessitam, basicamente, considerar (ALMEIDA; FISHWICH; TANG, 1991):

- 1- informações sobre os dados passados;
- 2- que o ocorrido no passado irá se repetir no futuro.

Os métodos aplicados na previsão de cargas elétricas podem ser divididos em três classes (O'DONOVAN, 1983):

1-Método Intuitivo: Nessa classe os métodos previsores descrevem o comportamento da carga utilizando um modelo matemático formal.

2-Método Autônomo: A extrapolação de observações de cargas passadas é o que caracteriza esse método de previsão onde é analisada a evolução da carga no tempo e dependendo do seu comportamento é classificada em horizontal, sazonal, tendenciosa e cíclicas. A característica do padrão horizontal consiste na flutuação dos dados em torno de um determinado valor médio constante no tempo. Já no padrão tendencioso, os dados flutuam em torno de uma curva de tendência (média variável no tempo). No padrão sazonal, a carga é influenciada por fatores temporais, tais como: estação do ano e dias da semana. No padrão cíclico as flutuações ocorrem variando a amplitude e a duração.

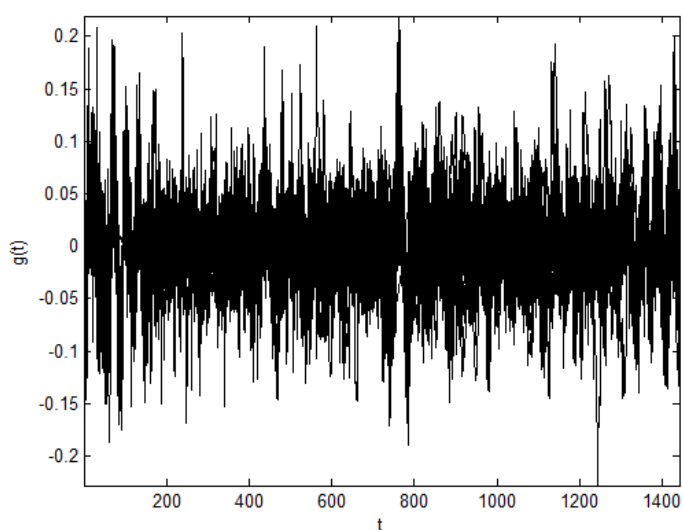
3-Método Casual: Neste método as variáveis independentes descrevem seu comportamento dependendo totalmente de sua relação com a carga. Várias variáveis

justificam a tendência de crescimento da carga tais como: índices socioeconômicos, crescimento populacional, etc.

Para escolher o modelo adequado é necessário determinar o padrão de comportamento da carga, e visando descrever a evolução da carga no tempo e por meio dessa evolução pode-se dividir a série temporal em duas classes (O'DONAVAN, 1983):

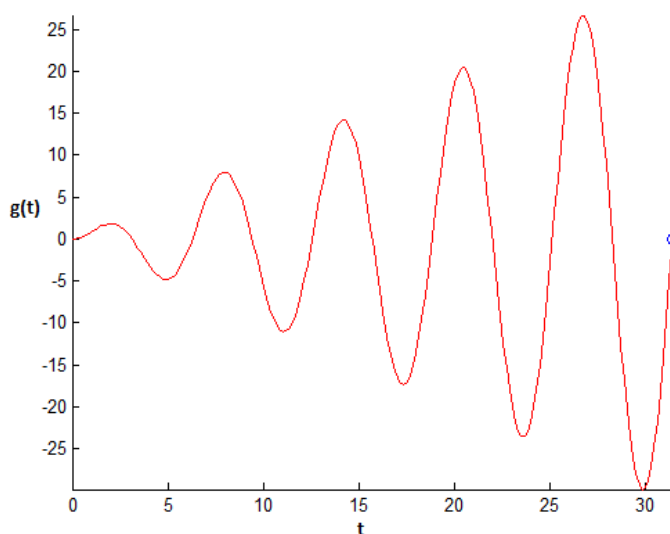
1-Estacionária: Uma série temporal é dita estacionária se os dados da carga tiverem flutuações em torno de uma média constante de tempo. Na figura 1, observa-se que os dados possuem uma média constante no tempo, o que indica que a série temporal é estacionária.

Figura 1 - Série Temporal Estacionária.



Fonte: Elaboração da própria autora.

2-Não-Estacionária: A série temporal é não-estacionária se os dados da carga variam no tempo, ou seja, se os dados (série temporal) apresentam padrões tendenciosos, cíclicos, sazonais ou aleatórios. Na figura 2, nota-se um padrão sazonal com aparente tendência para cima caracterizando assim uma série temporal não estacionária. Trata-se, portanto, de um comportamento que não se observa uma tendência estatística.

Figura 2 - Série temporal não estacionária.

Fonte: Elaboração da própria autora.

O sistema elétrico de potência atende diferentes tipos de consumidores, tais como: residencial, comercial e industrial, os quais influenciam e dificultam de certa forma a previsão de carga, por exemplo, no residencial tem-se um consumo particular variando de residência a residência. O consumidor industrial, apesar de ser previsível por causa do fato de sua produção ser praticamente constante, tem-se a possibilidade da ocorrência de eventos inesperados, tais como, por exemplo, quebra de maquinários, greves dos funcionários que venham ocasionar distúrbios inesperados, etc.. Esse tipo de evento afeta a previsão de carga já que a indústria é um dos maiores consumidores de energia elétrica. Assim, dentre vários fatores que influenciam na previsão de carga estão:

- **Fatores Climáticos:** a característica desse fator é de que sua ocorrência é de curto prazo. Alguns exemplos de fatores climáticos são: temperatura, neblina, velocidade do vento, umidade relativa do ar, temporal com relâmpagos e trovões, chuva, neve, etc. A temperatura é um fator climático que é influenciada praticamente pelas outras citadas tais como: nebulosidade, chuva, neve, entre outros fatores. E esta, por sua vez, influencia diretamente no consumo de energia elétrica delimitando o uso de ar-condicionado, aquecedores, refrigeradores, etc. (SWARUP; SATISH, 2002; HIPPERT; PEDREIRA; SOUZA, 2001, MURTO, 1998);

- **Fatores econômicos e demográficos:** são fatores que ocorrem em longo prazo e que possuem importância fundamental na evolução da previsão de carga. Tem-se, por exemplo, área demográfica (rural, residencial e industrial), crescimento industrial, variação na agricultura, tendências econômicas, variação nos preços de eletricidade, administração da demanda, etc. (MURTO, 1998);
- **Fatores de tempo:** são desempenhados pelos comportamentos cíclicos e sazonais. No comportamento sazonal, ocorre que em certas épocas do ano há uma variação da carga devido a determinados período do ano, tais como: férias, feriados, início do ano escolar, natal, ano novo, semana típica na qual ocorre o trabalho e o descanso, etc.. No comportamento cíclico, ocorre um aumento da carga nos dias da semana, e aos sábados e domingos ocorre uma redução da carga devido ao fato de que muitas pessoas descansam nesses dias (MURTO, 1998);
- **Fatores esporádicos:** São acontecimentos eventuais que podem causar perturbações nas cargas, tais como: eventos esportivos, programa de televisão, paralisação das indústrias, início ou parada de eventos onde a demanda de carga elétrica é grande (usina siderúrgica, fornalhas, fábricas, etc.) (MURTO, 1998).

2.3 HORIZONTES DA PREVISÃO

A previsão de carga temporal consiste na determinação de procedimentos efetuados para o fornecimento da carga e pode ser realizada entre horas, dias, meses ou anos. Desta forma, pode-se classificar a previsão de carga de acordo com o período em que a carga é prevista. Existem quatro tipos de previsão de cargas, são elas (AL-KANDARI; SOLIMAN; EL-HAWARY, 2004; METAXIOTIS et al., 2003; MURTO, 1998; SWARUP; SATISH, 2002; YALCINOZ; EMINOGLU, 2004):

1-Previsão de Carga a longo prazo: esse tipo de previsão é realizada principalmente na operação de planejamento de sistemas elétricos de potência. Essa previsão envolve um período de previsão de 10 a 20 anos. É realizada no intuito de efetuar um planejamento a longo prazo. Nesse tipo de previsão, não se considera os dados de carga diárias ou semanais, somente a tendência de consumo de demanda de carga a longo prazo.

2-Previsão de Carga a médio prazo: usam-se intervalos mensais; sua principal aplicação é na otimização da alocação das máquinas geradoras, com o objetivo de se ter uma operação econômica. Existe também o emprego da previsão de médio prazo para executar

estudos elétricos de situações de instabilidade da rede elétrica, objetivando manter a segurança do sistema.

3-Previsão de Carga a curto prazo: utilizam-se intervalos semanais; tem por finalidade a alocação econômica dos geradores, observando restrições operativas, de maneira que o custo da produção de energia elétrica seja menor. Com uma boa previsão de carga de curto prazo também se pode otimizar o uso de geradores térmicos, objetivando o despacho econômico.

4-Previsão de Carga a curtíssimo prazo: são essencialmente usadas para o planejamento de produção *on-line* e o controle. A previsão ocorre entre poucos segundos e pouco menos de 15 minutos.

Neste trabalho, somente será focado o problema de previsão a curto prazo e a previsão realizada apenas para 24 horas à frente.

2.4 TÉCNICAS PARA A PREVISÃO DE CARGAS ELÉTRICAS

Na literatura, existem inúmeras técnicas ou métodos para efetuar uma previsão de carga elétrica. Essas podem ser caracterizadas em duas classes fundamentais que são: (GROSS; GALIANA, 1987; BUNN; FARMER, 1985):

- **Modelo dinâmico:** a carga é identificada não apenas como função de tempo do dia, mas também de seus comportamentos recentes. Por essa razão, a previsão para certa hora precisa de previsões feitas em horas anteriores.
- **Modelo não-dinâmico:** a carga é expressa como uma série temporal discreta consistindo de valores previstos para cada hora durante um período.

Os métodos encontrados na literatura especializada estão divididos em duas classes: os métodos estatísticos e os sistemas especialistas.

2.4.1 Métodos estatísticos

Os métodos estatísticos são aqueles baseados em técnicas estatísticas onde a formulação matemática da modelagem da carga é de extrema importância. Dentre os métodos

estatísticos, podem-se destacar: regressão linear ou múltipla, alisamento exponencial, filtro de Kalman, espaço de estado e série temporal de Box-Jenkins.

2.4.1.1 Método de regressão linear ou múltipla

A modelagem feita por esse método descreve o comportamento estocástico dos padrões de carga horária em um sistema de potência. No modelo assume-se que a carga, em uma determinada hora, pode ser estimada a partir da relação da carga padrão e algumas variáveis explanatórias, que, por exemplo, podem ser variáveis climáticas, tipo do dia, classe de clientes, etc. O método pode ser descrito da seguinte forma: (MOGHRAM; RAHMAN, 1989; SWARUP; SATISH, 2002; SFETSOS, 2003; MURTO, 1998):

$$y(t) = a_0 + \sum_{i=1}^n a_i q_i(t) + a(t) \quad (1)$$

em que:

$y(t)$: carga elétrica;

a_0, a_1, \dots, a_n : coeficientes de regressão;

q_1, q_2, \dots, q_n : variáveis explanatórias;

$a(t)$: variável aleatória com média zero e variância constante.

Os coeficientes são encontrados usando-se de métodos dos mínimos quadrados. São realizados testes estatísticos para determinar a significância desses coeficientes de regressão.

2.4.1.2 Alisamento exponencial

A modelagem da carga $y(t)$, para o tempo t , utilizando a função de ajuste (MOGHRAM; RAHMAN, 1989):

$$f(t) = A(t)f(t + 1) \quad (2)$$

É dada pela equação (3):

$$y(t) = \varphi(t)^T f(t) + \varepsilon(t) \quad (3)$$

sendo:

$y(t)$: carga elétrica;

$\varphi(t)$: vetor de coeficiente;

$f(t)$: vetor da função de ajuste para o processo;

$\varepsilon(t)$: ruído branco;

$A(t)$: matriz de transição;

T : operador transposto.

2.4.1.3 Espaço de estado e Filtro de Kalman

A carga é modelada como uma variável de estado usando a formulação de espaço de estado. A descrição desse modelo de espaço de estado, para a carga, no tempo t , é definida da seguinte forma (MOGHRAM; RAHMAN, 1989; VILLALBA; BEL, 2000; MURTO, 1998):

Equação do espaço de estado:

$$y(t + 1) = A(t)y(t) + w(t) \quad (4)$$

Equação de medida:

$$z(t) = B(t)y(t) + v(t) \quad (5)$$

sendo:

$A(t)$: matriz de transição;

$y(t)$: carga elétrica;

$w(t)$: ruído branco com matriz de covariância $Q(t)$;

$z(t)$: vetor de medição de t ;

$B(t)$: matriz que correlaciona $y(t)$ e $z(t)$;

$v(t)$: erro medido (ruído com a matriz de covariância $S(t)$).

O filtro de Kalman tem por objetivo utilizar medições, ao longo do tempo, que possuem ruídos e incertezas e gerar resultados que tendam a se aproximar dos valores mais reais possíveis.

O esquema de implementação do filtro de Kalman é descrito a seguir:

- 1- Encontra-se a estimativa $y(t/t - 1)$ e o erro da matriz de covariância $p(t/t - 1)$;
- 2- Calcular o ganho de Kalman:

$$k(t) = p \left(\frac{t}{t-1} \right) B(t)^T [B(t)p \left(\frac{t}{t-1} \right) B(t)^T - S(t)^{-1}] \quad (6)$$

3- Calcular a estimativa atualizada do erro da matriz de covariância:

$$p(t/t) = [1 - k(t)B(t)]p(t/t - 1) \quad (7)$$

sendo:

$k(t)$: fator de combinação.

4- Projetar, à frente, a estimativa $y(t + 1/t)$ e o erro da matriz de covariância $p(t + 1/t)$ associada com:

$$y(t + 1/t) = A(t)y(t/t) \quad (8)$$

$$p(t + 1/t) = A(t)p(t/t)A(t)^T + Q(t) \quad (9)$$

5- Retornar ao passo 2 mudando para o próximo passo de tempo.

O modelo de espaço de estado é baseado no modelo de filtro de Kalman. A maior dificuldade está no processo de identificação.

2.4.1.4 *Série temporal de Box-Jenkins*

Essa metodologia foi proposta na década de 70. Tem por objetivo realizar previsões, permitindo que valores futuros de uma série sejam previstos tomando por base apenas seus valores presentes ou passados e isso é feito por meio da correlação temporal existente entre os valores.

O modelo de Box-Jenkins é aplicado aos processos estocásticos que sejam estacionários consistindo em uma classe de modelos lineares conhecidos como ARIMA (O'DONOVAN, 1983; MOGHARAM; RAHMAN, 1989; MURTO, 1998). Dentre essa classe pode-se citar os seguintes modelos: o modelo ruído branco, o modelo *Auto-Regressive* (AR), o modelo *Moving Average* (MA) e o modelo *Auto-Regressive Moving Average* (ARMA).

- Modelo de ruído branco

O modelo de ruído branco é definido pela equação (10) (O'DONOVAN,1983):

$$y(t) = \theta_0 + a(t) \quad (10)$$

em que:

$y(t)$: carga elétrica;

θ_0 : termo constante;

$a(t)$: séries aleatórias.

- Modelo Auto-Regressive (AR(p))

A equação (11) define o modelo *Auto-Regressive* (O'DONOVAN,1983, MOGHRAM; RAHMAN, 1989):

$$y(t) = \theta_0 + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_r y(t-r) + a(t) \quad (11)$$

sendo:

ϕ_1, \dots, ϕ_r : coeficientes *auto-regressive*;

r : ordem do modelo.

- *Modelo Moving Average (MA(q))*

A equação que define o modelo *Moving Average* é descrita pela equação (12) (O'DONOVAN, 1983; MOGHRAM; RAHMAN, 1989):

$$y(t) = \theta_0 + a(t) - \theta_1 a(t-1) - \theta_2 a(t-2) - \dots - \theta_r a(t-r) \quad (12)$$

sendo:

$\theta_1, \dots, \theta_r$: são coeficientes de média móvel;

r : ordem do modelo.

- Modelo Auto Regressive Moving Average (ARMA(p,q))

Este modelo é a junção dos modelos *Auto-Regressive* e *Moving Average* e está representado pela equação (13) (O'DONOVAN, 1983; MOGHRAM; RAHMAN, 1989; MURTO, 1998):

$$y(t) = \theta_0 + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + a(t) + \theta_0 + a(t) - \theta_1 a(t-1) - \theta_2 a(t-2) - \dots - \theta_r a(t-r) \quad (13)$$

Os modelos apresentados, até agora, são adequados para séries estacionárias, ou seja, aquelas em que a média é constante em todo tempo. Porém, na prática as séries são geralmente não-estacionárias. Os modelos ARIMA são a classe de modelos mais comuns que representam uma série temporal não estacionária.

Para tornar uma série não estacionária em uma série estacionária deve-se tomar a diferença quantas vezes for necessárias até alcançar a estacionariedade da série.

O modelo que será apresentado agora são modelos para séries temporais não estacionárias. Dentre eles são: Modelo Walk Aleatório e os modelos *Auto-Regressive Integrated Moving Average* (ARIMA).

- Modelo Walk Aleatório

O modelo é definido pela equação (14) (O'DONOVAN, 1983):

$$y(t) = y(t-1) + \theta_0 + a(t) \quad (14)$$

A característica desse modelo é que as primeiras diferenças regulares desse modelo forma o modelo de ruído branco e que a previsão para todos os períodos de tempo futuro é exatamente o último valor observado da série temporal.

- Modelos Auto-Regressive Integrated Moving Average (ARIMA)

Esses modelos possuem a característica de que pelas primeiras diferenças regulares formam o modelo ARMA, tais como AR(1), AR(2), MA(1) e MA(2) (O'DONOVAN, 1983):

O modelo *Auto-Regressive Integrated* ARI(1,1) é definido como (O'DONOVAN, 1983):

$$y(t) = \theta_0 + (1 + \phi_1)y(t - 1) - \phi_1y(t - 2) + a(t) \quad (15)$$

em que as primeiras diferenças formam o AR(1).

O modelo ARI (2,1) é encontrado de forma similar. As primeiras diferenças regulares formam o AR (2). O modelo IMA (1,1) é definido pela equação (16) (O'DONOVAN, 1983):

$$y(t) = \theta_0 + y(t - 1) - \theta_1a(t - 1) + a(t) \quad (16)$$

na qual, as primeiras diferenças formam o modelo MA(1).

Por fim, tem-se o modelo ARIMA (1,1,1), sendo as primeiras diferenças regulares formam o ARMA(1,1). Este modelo é descrito pela equação (17) (O'DONOVAN, 1983):

$$y(t) = \theta_0 + (1 + \phi_1)y(t - 1) - \phi_1y(t - 2) + a(t) - \theta_1a(t - 1) \quad (17)$$

Pode-se observar que os modelos ARIMA são aqueles cujas diferenças regulares formam o modelo ARMA.

2.5 SISTEMAS ESPECIALISTAS

Os sistemas especialistas são modelos heurísticos baseados na inteligência artificial. Tentam imitar o raciocínio humano, no qual possui uma base de conhecimento em um domínio restrito e utiliza o raciocínio complexo e conclusivo na execução de suas tarefas (MOGHAM; RAHMAN, 1989; METAXIOTIS et al., 2003; MURTO, 1998).

Ao realizar a previsão de carga baseiam-se em um banco de dados históricos que contém informações que auxiliam na previsão tais como: dia da semana, hora, fatores sociais e temporais.

2.6 COMENTÁRIOS

Neste capítulo foi focado a importância da previsão de carga no sistema elétrico e os fatores que influenciam direta e indiretamente na previsão. Foram apresentados o horizonte de

previsão (de longo, médio, curto ou curtíssimo prazo) e os modelos matemáticos encontrados na literatura para a previsão de carga: modelos estacionários e não-estacionários.

3 REDES NEURAIS ARTIFICIAIS

Neste capítulo serão abordados um breve histórico do surgimento do estudo de redes neurais, as potenciais áreas de aplicações, os principais conceitos envolvidos, a estrutura de uma rede, os tipos de treinamento e, também, os principais tipos de rede neurais artificiais.

3.1 NOTA HISTÓRICA

A curiosidade pelo conhecimento impulsionou o homem a estudar o funcionamento do cérebro humano. Por volta de 1943 os desbravadores McCulloch & Pitts publicaram seu primeiro trabalho relacionado à neurocomputação. Neste trabalho, os autores propuseram o primeiro modelo matemático inspirado em um neurônio biológico. Em 1949, Donald Hebb publicou um livro cujo título é “*The Organization of Behavior*” e introduziu o conceito básico da regra de aprendizado de Hebb, sendo este trabalho baseado em hipóteses e observações de caráter neurofisiológico.

Em 1951, com Minsky surgiu as redes neurais baseadas em reforço com formas simples de aprendizado. Em 1958, o cérebro foi descrito por Von Neumann no seu livro “*The Computer and the Brain*” (VON NEUMANN, 1958) como uma modelagem do desempenho do cérebro por meio de *hardware*. Outros pesquisadores continuaram desenvolvendo modelos matemáticos baseados na estrutura de um neurônio biológico. Rosenblatt publicou um livro intitulado “*Principles of Neurodynamics*” (ROSENBLATT, 1962), no qual introduziu o conceito de redes neurais denominadas *perceptron* que despertou interesse de outros pesquisadores por causa da sua capacidade de reconhecer padrões simples. Minsky & Papert descobriram que o modelo proposto por Rosenblatt não resolvia problema do tipo ou-exclusivo. Com isso houve um desinteresse por essa linha de pesquisa.

Somente em no final da década de 70 e início da década de 80 é que foi retomado o interesse pelo estudo de redes neurais. Grossberg, com a implementação da rede ART (*Adaptive Resonance Theory*), Kohonen com a formulação de mapas auto-organizáveis, Hopfield propondo as redes recorrentes baseadas em função de energia, etc., contribuíram para essa retomada. O interesse em voltar a pesquisar essa área foi por causa destas contribuições e também pelo fato de que os computadores tiveram um aumento na capacidade

de processamento, da criação de algoritmos de otimização mais eficientes e robustos e, finalmente, das novas descobertas sobre o sistema nervoso biológico (SILVA; SPATTI; FLAUZINO, 2010).

Outro marco importante foi o desenvolvimento do algoritmo *Backpropagation* que motivou e reacendeu o interesse dos pesquisadores. Esse algoritmo foi proposto por Werbos, em 1974, e em 1986, foi redescoberto por Rumerhart, Hinton e Willians (RUMELHART; HINTON; WILLIAMS, 1986). Recentemente há inúmeras áreas de aplicação desse algoritmo e pesquisas relacionadas para o melhoramento dessa rede.

3.2 ÁREAS DE APLICAÇÃO

O interesse dessa linha de pesquisa é muito vasto pelo fato de que se aplica em diversos problemas relacionados às engenharias e às ciências. Alguns exemplos encontrados na literatura podem ser citados, tais como (SILVA; SPATTI; FLAUZINO, 2010):

Aproximação universal de funções: é feito um mapeamento entre o relacionamento funcional e as variáveis de um sistema a partir de um conjunto conhecido de seus valores;

Controle de processos: cujo objetivo é identificar ações de controle que permitam o alcance dos requisitos de qualidade, eficiência e segurança do processo. Exemplos: controles empregados em robótica, aeronaves, elevadores, etc.;

Reconhecimento e classificações de padrão: associa um padrão de entrada para umas das classes previamente definidas como, por exemplo, no reconhecimento de imagens, voz, escrita, etc.;

Agrupamento de dados: identifica e detecta a similaridade e a particularidade entre os diversos padrões de entrada e os agrupa, por exemplo, podemos citar identificação automática de classes e garimpagem de dados;

Sistema de previsão: estimam-se valores futuros de um processo por meio de observações anteriores, por exemplo, tem-se a previsão de séries temporais, previsão de mercado financeiro, previsões climáticas, etc.;

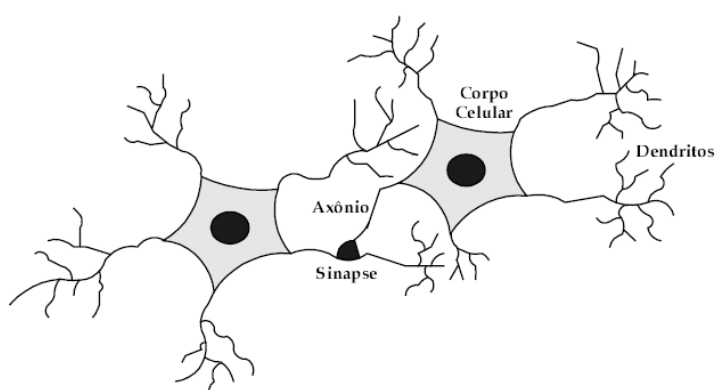
Otimização de Sistemas: o objetivo é minimizar ou maximizar uma função custo obedecendo também eventuais restrições que são impostas para o correto mapeamento do problema.

Memórias associativas: recuperam-se padrões corretos mesmo se os seus elementos constituintes forem apresentados de forma incerta ou imprecisa, por exemplo, tem-se o processamento de imagens, a transmissão de sinais, a identificação de caracteres manuscritos, etc.

3.3 NEURÔNIO BIOLÓGICO

As redes neurais artificiais são baseadas nas estruturas e no comportamento do cérebro humano. No cérebro humano, o processamento de informações é feito por processadores biológicos que operam em paralelo, tendo como objetivo produzir ações no corpo humano como pensar, raciocinar e memorizar. O neurônio é uma célula fundamental do sistema nervoso e tem o papel fundamental que é conduzir os estímulos elétricos vindo de reações físico-químicas. Essa célula pode ser dividida em três partes principais, isto é, os dendritos, o corpo celular e o axônio. A figura 3 ilustra os componentes de um neurônio biológico.

Figura 3 - Componentes de um neurônio biológico.



Fonte: Mendes Filho e Carvalho (1997).

Os dendritos (terminais de entrada) são finos prolongamentos que possuem a função de captar, de forma contínua os estímulos vindos dos outros neurônios ou do próprio meio externo e enviam por um longo e fino suporte conhecido como axônio (terminais de saída). O corpo celular é o responsável por processar as informações advindas dos dendritos e produzir um potencial de ativação que indicará se o neurônio poderá ou não disparar um sinal de impulso elétrico ao longo de seu axônio (MENDES; CARVALHO, 1997; STERGIOU, 1996, SILVA; SPATTI; FLAUZINO, 2010).

Os neurônios são interligados por meio de sinapses, região na qual o axônio de um neurônio se conecta com o dendrito do outro neurônio. Um neurônio somente é ativo quando os sinais eletroquímicos são disparados ao longo do seu axônio, que por meio das sinapses, é enviado para outro neurônio. Esse sinal eletroquímico do neurônio só é disparado se o sinal total recebido dos dendritos exceder certo nível (*threshold*) e a eficácia das sinapses é fundamental nessa ativação. As sinapses possuem uma abertura com um neurotransmissor químico para transmitir um impulso nervoso, por meio dessa conectividade é atribuída à inteligência.

3.4 MODELO NEURAL ARTIFICIAL

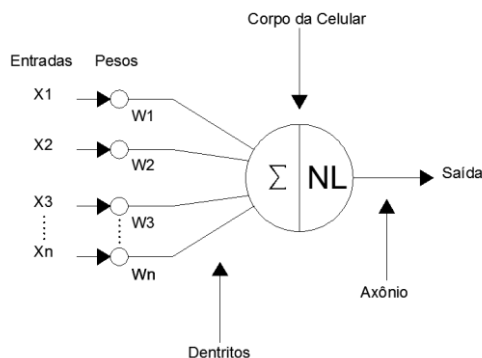
As redes neurais artificiais foram desenvolvidas a partir do modelo de um neurônio biológico do próprio cérebro humano. Tais modelos foram inspirados a partir da análise da geração e propagação de impulsos elétricos pela membrana celular dos neurônios (HODGKIN; HUXLEY, 1952).

A rede neural artificial possui a capacidade de armazenar conhecimento por causa do fato de ser um processador em paralelo composto por unidades de processamento. Muito similar ao cérebro humano, o neurônio artificial possui características semelhantes ao neurônio biológico. Essas características são tais (WASSERMAN, 1989):

- Processamento paralelo maciço: Armazenamento do conhecimento via treinamento (STORB; WAZLAWICK, 1999).
- Interligação entre neurônios feita por conexões sinápticas: Utilizadas para armazenamento de informações (STORB; WAZLAWICK, 1999).
- Capacidade de generalização: Permite que a rede neural artificial forneça, em tempo real, saídas com uma razoável precisão para entradas que não participaram da fase de treinamento (STORB; WAZLAWICK, 1999).

Na figura 4 ilustram-se os componentes de um neurônio artificial proposto por McCulloch e Pitts (MCCULLOCH; PITTS, 1943; STERGIU, 1996).

Figura 4 - Componentes de um neurônio artificial.



Fonte: McCulloch e Pitts (1943).

O neurônio artificial possui um processo de treinamento da seguinte forma (WASSERMAN, 1989):

1. É apresentado à rede um conjunto de entrada X, x_1, x_2, \dots, x_n . Essas entradas correspondem aos sinais dentro das sinapses de um neurônio biológico e são normalizadas visando aumentar a eficiência computacional.
2. Existe um peso associado, V, v_1, v_2, \dots, v_n que multiplica cada sinal x_i , que está associada com a “resistência” de uma ligação sináptica biológica.
3. Após a entrada ser multiplicada por um peso associado é aplicados ao um bloco somatório Σ , denominado soma, que corresponde ao corpo celular biológico, cuja função é adicionar todos os sinais de entrada que foram ponderados a fim de produzir um valor de potencial de ativação.
4. Se a soma ponderada das entradas for maior que o valor do “*bias*” o neurônio é ativado provocando um pulso na saída. Caso contrário, o neurônio não é ativado. Os pesos *bias* ponderam a entrada (arbitrária), $x_0 = +1$, que atua como controle do nível de saída do neurônio (LOPES, 2005).

A saída intermediária do neurônio é, então, dada por:

$$s_i = \sum_j^n w_{ij} x_i \quad (18)$$

5. A ativação do neurônio artificial se dá pela função de ativação, cujo seu objetivo é limitar a saída do neurônio dentro de um intervalo de valores.

As funções de ativação podem ser divididas em dois grupos: (1) das funções parcialmente diferenciáveis e o (2) das funções totalmente diferenciáveis (SILVA; SPATTI; FLAUZINO, 2010).

3.4.1 Funções de ativação parcialmente diferenciáveis

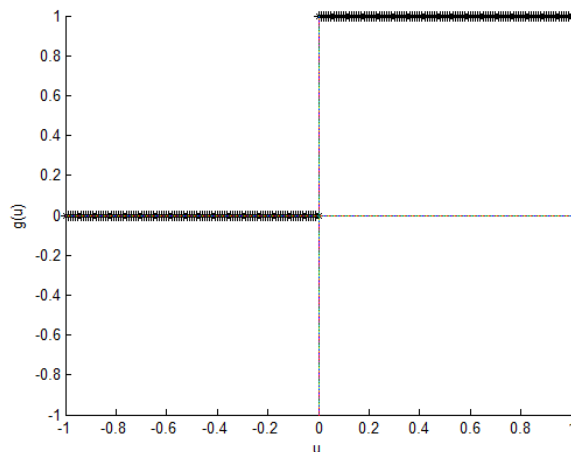
São aquelas que possuem pontos cuja derivada de primeira ordem são inexistentes. As principais são: função degrau, função degrau bipolar e função rampa simétrica.

3.4.1.1 Função degrau

O resultado da aplicação dessa função será valores unitários ou iguais a zero quando o potencial de ativação for maior que zero; caso contrário, o resultado assumirá valores nulos:

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases} \quad (19)$$

Figura 5 - Representação gráfica da função degrau.



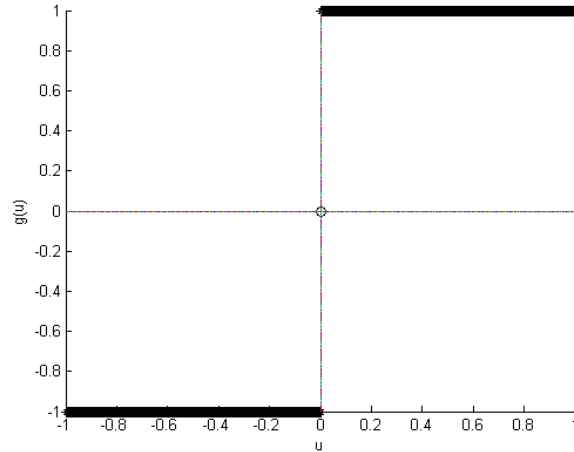
Fonte: Elaboração da própria autora.

3.4.1.2 Função degrau bipolar ou função sinal

Os valores assumidos pela aplicação dessa função serão unitários positivos quando o potencial de ativação for maior que zero; valor nulo quando o potencial também for nulo, e valores unitários negativos quando o potencial for menor que zero.

$$g(u) = \begin{cases} 1, & \text{se } u > 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u < 0 \end{cases} \quad (20)$$

Figura 6 - Representação gráfica da função degrau bipolar ou função sinal.



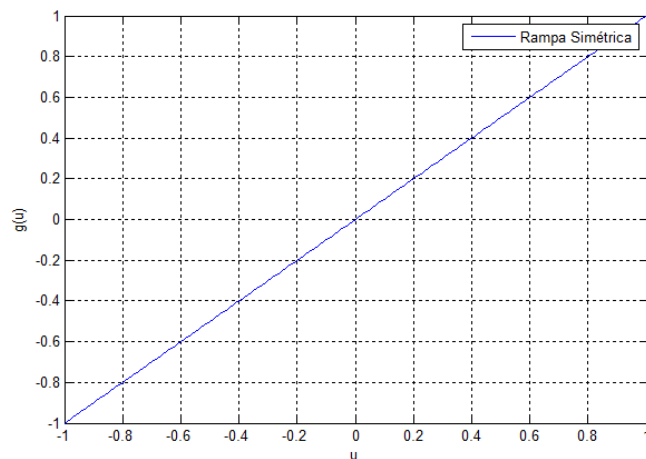
Fonte: Elaboração da própria autora.

3.4.1.3 Função rampa simétrica

Os valores retornados são iguais aos próprios valores dos potenciais de ativação quando estão definidos no intervalo $[-a, a]$, restringindo-se aos valores limites, em caso contrário.

$$g(u) = \begin{cases} a, & \text{se } u > a \\ u, & \text{se } -a \leq u \leq a \\ -a, & \text{se } u < -a \end{cases} \quad (21)$$

Figura 7 - Representação gráfica da função rampa simétrica.



Fonte: Elaboração da própria autora.

3.4.2 Funções de ativação totalmente diferenciáveis

São aquelas cujas derivadas de primeira ordem existem e são conhecidas em todos os pontos de seu domínio. As quatro principais são: (1) a função logística, (2) tangente hiperbólica, (3) gaussiana e (4) a função linear (SILVA; SPATTI; FLAUZINO, 2010).

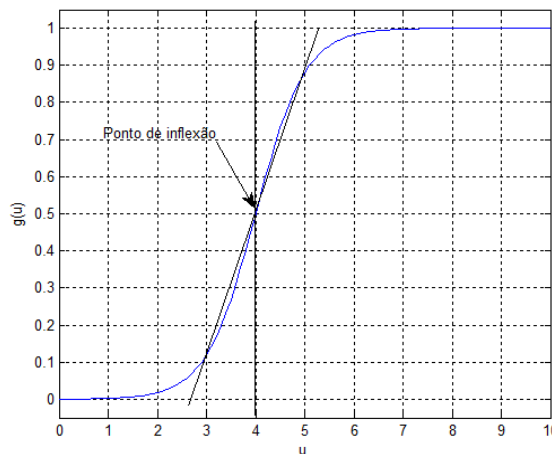
3.4.2.1 Função Logística

As saídas produzidas pela função logística sempre assumirão valores reais pertencentes ao intervalo zero e um. A expressão matemática que descreve essa função é dada por:

$$g(u) = \frac{1}{1+e^{-\lambda u}} \quad (22)$$

sendo λ uma constante real associada ao nível de inclinação da função logística frente ao seu ponto de inflexão.

Figura 8 - Representação gráfica da Função Logística.



Fonte: Elaboração da própria autora.

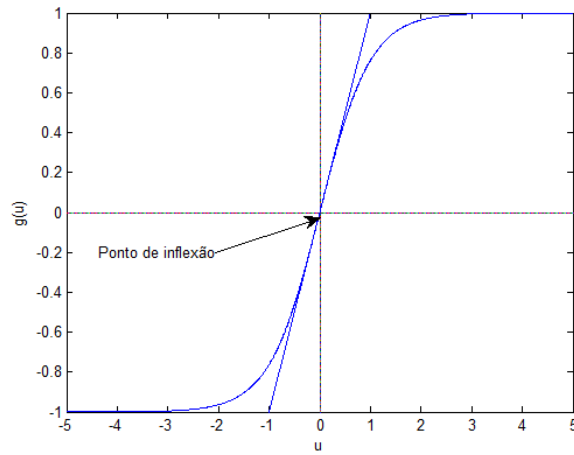
3.4.2.2 Função tangente hiperbólica

O resultado de saída sempre assume valores reais compreendidos entre -1 e 1. A equação que descreve essa função é definida pela equação (23):

$$g(u) = \frac{1-e^{-\lambda u}}{1+e^{-\lambda u}} \quad (23)$$

sendo que λ está associado ao nível de inclinação da função tangente hiperbólica em relação ao seu ponto de inflexão.

Figura 9 - Representação gráfica da Função Tangente Hiperbólica.



Fonte: Elaboração da própria autora.

As funções tangentes hiperbólicas e a função logística pertencem à família das funções denominadas sigmoidais.

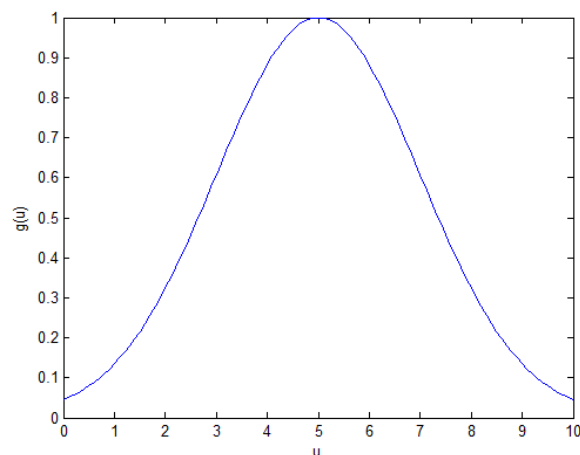
3.4.2.3 Função gaussiana

A saída do neurônio produzirá resultados iguais para aqueles valores de potencial de ativação $\{u\}$ que estejam posicionados a uma mesma distância de seu centro (média), sendo a curva simétrica em relação a este. A função gaussiana é dada pela equação (24):

$$g(u) = e^{\frac{-(u-c)^2}{2\sigma^2}} \quad (24)$$

sendo que c é o parâmetro que define o centro da função gaussiana e σ denota o desvio padrão associado à mesma.

Figura 10 - Representação gráfica da Função Gaussiana.



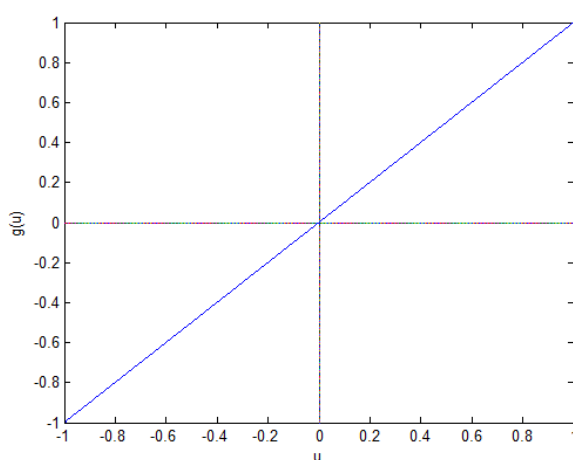
Fonte: Elaboração da própria autora.

3.4.2.4. Função linear

Esta função produz resultados de saída idênticos aos valores do potencial de ativação $\{u\}$, tendo sua expressão matemática definida pela equação (25):

$$g(u) = u \quad (25)$$

Figura 11 - Representação gráfica da Função Linear.



Fonte: Elaboração da própria autora.

A função linear é aplicada nas redes neurais como aproximadoras universais de funções visando mapear o comportamento entre as variáveis de entrada e saída de processos. (SILVA; SPATTI; FLAUZINO, 2010).

3.5 ESTRUTURA DAS REDES NEURAIIS ARTIFICIAIS

Uma rede neural artificial possui uma arquitetura que define como os neurônios estão organizados em relação aos outros. Essa organização é essencial para o direcionamento das conexões sinápticas dos neurônios. O processo de treinamento de uma arquitetura precisa ter um conjunto ordenado de passos com o objetivo de ajustar os pesos e os limiares de seus neurônios. Esse processo de ajuste é conhecido como algoritmo de aprendizagem (ou treinamento), cujo intuito é sintonizar a rede neural para que a resposta de saída da mesma esteja próxima dos valores reais.

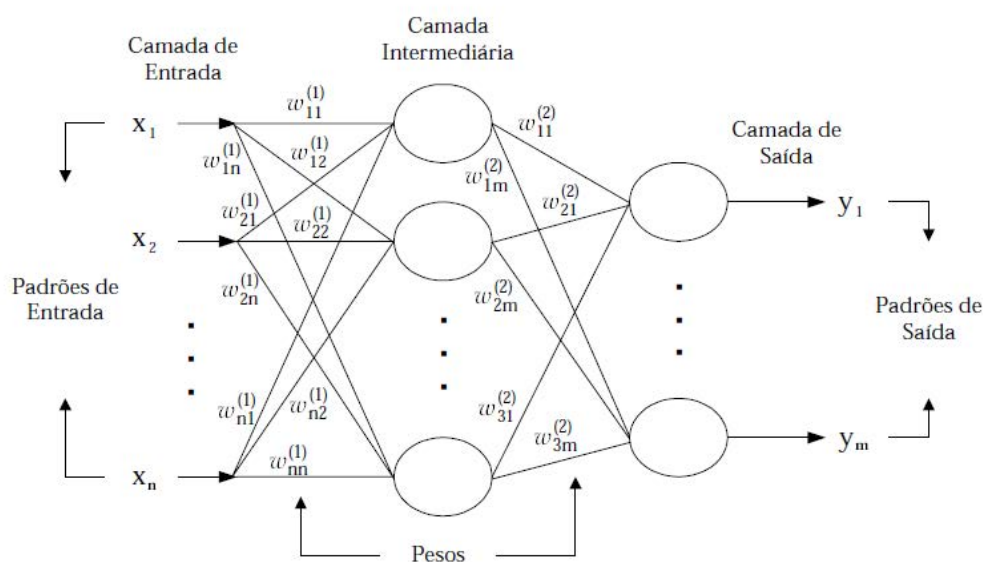
A rede neural artificial é dividida em camadas (SILVA; SPATTI; FLAUZINO, 2010):

Camada de entrada: recebe as informações, características ou medições do meio externo. São geralmente normalizadas em relação à faixa de variações dinâmicas produzidas pela função de ativação.

Camada intermediária, escondida, oculta ou invisível: os neurônios dessa camada são responsáveis pela extração de características associadas ao processo.

Camada de saída: Os neurônios presentes nessa camada são responsáveis pela produção e apresentação dos resultados finais da rede.

Figura 12 - Disposição das Redes Neurais Artificiais.



Fonte: Lopes (2005).

A arquitetura da rede pode ser dividida em: (1) redes *feedforward* de camada simples, (2) arquitetura *feedforward* de camadas múltiplas, (3) arquitetura recorrente ou realimentada e (4) arquitetura em estrutura reticulada. Nessa divisão são consideradas as disposições dos neurônios, a forma como que são interligados e a constituição de suas camadas.

Arquitetura *feedforward* de camada simples: A característica dessa arquitetura é que se considera apenas a camada de entrada e a camada de saída. As informações possuem um único sentido, ou seja, da camada de entrada para camada de saída. Esse tipo de arquitetura de rede é geralmente aplicada em problemas de classificações de padrões e filtragem linear (SILVA; SPATTI; FLAUZINO, 2010). As redes que possuem esse tipo de arquitetura são as *Perceptron* e a rede neural ADALINE (WIDROW; LEHR, 1990).

Arquitetura *feedforward* de camadas múltiplas: Esse tipo rede possui camadas múltiplas que são constituídas de camada de entrada, intermediária e camada de saída. É uma arquitetura que soluciona diversos tipos de problemas. Por exemplo, aproximação de funções, classificação de padrões, identificação de sistemas, otimização, robótica, no controle de processos, etc. As principais redes com essa arquitetura é a *Perceptron* multicamadas e as redes de base radial.

Arquitetura recorrente ou realimentada: Possui uma realimentação, ou seja, a saída dos neurônios são realimentadas como sinais de entrada para outros neurônios. Essa característica de realimentação é utilizada em sistemas variantes em relação a tempo, como previsão de séries temporais, otimização e identificação de sistemas, controle de processo, etc. Por meio desse processo de realimentação, as redes produzem as saídas atuais levando em conta os valores das saídas anteriores.

Arquitetura em estrutura reticulada: A localização dos pesos está diretamente relacionada com o processo de ajuste de seus pesos e limiares (SILVA; SPATTI; FLAUZINO, 2010). Possui uma vasta área de aplicação, tais como problema de agrupamento, reconhecimento de padrões, otimização de sistemas, grafos, etc. A rede de Kohonen possui esse tipo de arquitetura, cujo treinamento é realizado por meio de um processo competitivo.

3.6 TREINAMENTO DA REDE NEURAL

As redes neurais possuem uma característica peculiar que é a capacidade de aprender a partir de padrões de amostra que modelam o comportamento do sistema, e depois de ter aprendido será capaz de generalizar soluções, ou seja, a rede produz uma saída próxima da desejada a partir dos sinais de entrada. O treinamento de uma rede pode ser dividido em:

3.6.1 Treinamento supervisionado

Nesse processo de treinamento considera-se cada amostra dos sinais de entrada e as respectivas saídas desejadas. Esse treinamento necessita de um tutor no processo de aprendizagem, que avalie e informe à rede sobre o seu desempenho. Os pesos sinápticos e limiares são continuamente ajustados mediante a comparação executada pelo próprio algoritmo de aprendizagem, que supervisiona a diferença entre as respostas e a usa para o procedimento de ajuste. Podem-se citar algumas redes que utilizam esse tipo de treinamento: a rede MADALINE (WIDROW; LEHR, 1990), que utiliza o algoritmo de aprendizagem

Backpropagation, a rede Grossberg (rede neural ARTMAP), a rede ARTMAP *Fuzzy* (CARPENTE; GROSSBERG, 1992), entre outras (KROSE; SMAGT, 1996; WASSERMAN, 1989).

3.6.2 Treinamento não supervisionado

Ao contrário do treinamento supervisionado, durante o treinamento não-supervisionado inexistem saídas desejadas. Não há necessidade de tutor, ou seja, é uma rede extremamente autônoma. A rede se auto-organiza identificando subconjuntos e extraindo particularidades da amostra de entrada. Os pesos sinápticos e limiares dos neurônios da rede são ajustados pelo algoritmo de aprendizagem de forma a refletir esta representação dentro da própria rede (SILVA, 2010). As redes de Hopfield (HOPFIELD, 1982), Kohonen (KOHONEN, 1987) e ART (convencional e nebulosa) são exemplos de redes que possuem treinamento não supervisionado (KROSE; SMAGT, 1996; WASSERMAN, 1989).

3.6.3 Treinamento com reforço

Esse método é considerado uma variação do treinamento supervisionado no qual se avalia constantemente a defasagem de valor entre a resposta produzida pela rede e a saída desejada. Os ajustes dos parâmetros internos dos neurônios baseiam-se em qualquer informação quantitativa ou qualitativa advindas da interação com o sistema, as quais são utilizadas para medir o desempenho do aprendizado (SILVA; SPATTI; FLAUZINO, 2010). As redes *Counterpropagation* são um exemplo de redes que possuem esse tipo de treinamento (STORB; WAZLAWICK, 1999).

3.7 TIPOS DE REDES NEURAIAS ARTIFICIAIS

Existem vários modelos de redes neurais sendo utilizados nos mais variados tipos de problemas. Dentre os mais conhecidos destacam-se (STORB; WAZLAWICK, 1999):

- **Rede *Perceptron*:** foi idealizada por Rosenblatt em 1958. É aplicada em problemas de reconhecimento e classificação de padrões e na resolução de problema lógicos que envolvem os operadores AND e OR.
- **Rede *ADALINE*:** foi idealizada por Widrow e Hoff (WIDROW; LEHR, 1990). Sua principal aplicação era em sistema de chaveamento de circuitos telefônicos. Uma rede

neural multicamada, apresenta um alto grau de tolerância a falhas (WIDROW; LEHR, 1990).

- **Redes *Perceptron* Multicamadas (PCM):** são caracterizadas pela presença de pelo menos uma camada intermediária de neurônio, entre a camada de entrada e a camada de saída. Possuem diversas áreas de aplicação nas mais variadas áreas do conhecimento, tais como: aproximação universal de funções; reconhecimento de padrões; identificação e controle de processos, previsão de séries temporais, otimização de sistemas (SILVA; SPATTI; FLAUZINO, 2010).
- **Redes de Base Radial (RBF):** podem ser empregadas em quase todos os tipos de problemas tratados pela rede neural PMC. Diferentemente das redes PCM, a estrutura de uma rede RBF é composta por uma camada intermediária, na qual a função de ativação é do tipo gaussiana.
- **Redes auto-organizáveis de Kohonen:** é uma rede competitiva com a habilidade de realizar mapeamentos entre dados de entrada e de saída (KOHONEN, 1987).
- **Redes ART (*Adaptive Resonance Theory*):** foi desenvolvida em 1975 por Grossberg e tem por base características da teoria da ressonância adaptativa. Ela é adaptada para reconhecimento de sinais e processamento de imagens (GROSSBERG, 1976).

Na busca de solucionar problemas reais que são difíceis, as redes neurais são fundamentais para encontrar a solução. Essas possuem uma estrutura que contempla a capacidade de generalização e consegue responder, de maneira adequada, a novas situações baseada em situações passadas.

As redes possuem características tais quais (HAYKIN, 1994; STORB; WAZLAWICK, 1999):

1. *Aprendizado e generalização:* possui a capacidade de obter resultados por meio da aprendizagem e armazenamento de conhecimento;
2. *Processamento paralelo:* permite que as tarefas complexas sejam executadas com maior rapidez;
3. *Não-linearidade:* elas suprem uma alternativa analítica que as técnicas convencionais estão limitadas;

4. *Robustez*: a rede é apta a processar dados com ruídos ou dados incompletos de forma eficaz;
5. *Flexibilidade*: pelo fato de sua ampla aplicabilidade em qualquer área de atuação.

3.8 COMENTÁRIOS

Neste capítulo foi apresentado um breve histórico sobre o surgimento do estudo de redes neurais e suas principais áreas de aplicação. As redes neurais foram fundamentadas nas estruturas e no comportamento do cérebro humano com base em um neurônio biológico, e a partir do funcionamento deste foi desenvolvido o neurônio artificial. Utilizando o modelo artificial mostrou-se toda estrutura de uma rede tais como suas camadas, tipos de realimentação, tipos de treinamento, tipos de redes existentes, tipos de arquitetura, vantagens da utilização da rede e, principalmente, a motivação de se estudar como consequência da sua aplicabilidade em diversas áreas da engenharia e da ciência.

4 ALGORITMO RETROPROPAGAÇÃO

O algoritmo Retropropagação RP é uma técnica que realiza a adaptação dos pesos da rede em direção oposta à direção do gradiente descendente, cujo objetivo consiste, em minimizar o erro quadrático da rede neural. O somatório do erro quadrático instantâneo de cada neurônio da última camada é dado pela equação (26) (WIDROW; LEHR, 1990):

$$\varepsilon^2 = \sum_{i=1}^{ns} \varepsilon_i^2 \quad (26)$$

sendo:

$$\varepsilon_i : d_i - y_i \quad (27)$$

d_i : saída desejada do i -ésimo elemento da última camada da rede;

y_i : saída do i -ésimo elemento da última camada da rede;

ns : número de neurônios da última camada da rede.

O método utilizado foi o do gradiente descendente, o qual é um algoritmo que é sensível a mínimos locais. Na equação (28) é adotada a direção contrária do gradiente do erro quadrático para minimizar a função objetivo do erro quadrático.

O ajuste foi obtido usando-se a equação (28) (SIMPSON 1989; WIDROW; LEHR, 1990):

$$\mathbf{V}_i (h + 1) = \mathbf{V}_i (h) + \eta \delta_i (h) \quad (28)$$

sendo:

$$\delta_i (h) = - \eta \left[\delta_i (h) \right];$$

η : parâmetro de controle da estabilidade ou taxa de treinamento;

h : representa o índice de iteração;

$\delta_i (h)$: gradiente do erro quadrático com relação aos pesos do neurônio i avaliado em h ;

\mathbf{V}_i : vetor contendo os pesos do neurônio i ;

$$= [w_{0i} \ w_{1i} \ w_{2i} \ \dots \ w_{ni}]^T .$$

O gradiente $\nabla_i(h)$ pode ser expresso por (KROSE; SMAGT, 1996):

$$\nabla_i(h) = \frac{\partial \varepsilon_i^2}{\partial v_i(h)} = 2\varepsilon_i \frac{\partial \varepsilon_i}{\partial v_i(h)} \quad (29)$$

A função de ativação adotada foi a função sigmoide. Esta função possui uma característica que é o equilíbrio gracioso entre o comportamento linear e não-linear (HAYKIN, 1994). Na literatura há dois tipos de funções sigmoidais (WERBOS, 1974):

$$y_i = \frac{1 - e^{-(s_i - \rho)}}{1 + e^{-(s_i - \rho)}} \quad (30)$$

ou:

$$y_i = \frac{1}{1 + e^{-(s_i - \rho)}} \quad (31)$$

sendo:

λ : é o parâmetro de declividade da função;

ρ : constante que determina a translação da curva y_i .

Os valores de saída da função y_i varia entre -1 e 1 para a função sigmoide dada pela equação (30), e 0 e 1 para função sigmoide dada pela equação (31). No caso de problemas de previsão de carga (valores sempre positivos), emprega-se somente a função dada pela equação (31).

Diferenciando a função sigmoide dada pelas equações, (30) e (31), com relação ao vetor V , obtém-se:

$$\frac{\partial \varepsilon_i}{\partial v_i} = -\frac{\partial y_i}{\partial v_i} = -\left(\frac{\partial y_i}{\partial s_i} \frac{\partial s_i}{\partial v_i}\right) \quad (32)$$

ou seja:

$$\frac{\partial \varepsilon_i}{\partial v_i} = -sgm'(s_i) \frac{\partial s_i}{\partial v_i} \quad (33)$$

sendo:

$$sgm'(s_i) = \frac{\partial y_i}{\partial s_i}$$

Derivando a equação (18), com relação aos pesos, tem-se:

$$\frac{\partial s_i}{\partial v_i} = X_i \quad (34)$$

sendo:

$$\begin{aligned} X_i &\triangleq i\text{-ésimo vetor padrão.} \\ &= [x_{0i} \ x_{1i} \ x_{2i} \ \dots \ x_{ni}]^T \end{aligned}$$

Substituindo a equação (34) na equação (33), obtém-se:

$$\frac{\partial \varepsilon_i}{\partial v_i} = -sgm'(s_i)X_i \quad (35)$$

Substituindo a equação (35) na equação (29), resulta em:

$$\nabla_i(h) = -2\varepsilon_i sgm'(s_i)X_i \quad (36)$$

O gradiente obtido na equação (36) é aplicado no método gradiente descendente da equação (28), obtém-se o seguinte algoritmo (VILLALBA; BEL, 2000):

$$V_i(h + 1) = V_i(h) + 2\gamma\beta_iX_i \quad (37)$$

Se o elemento i encontra-se na última camada, então:

$$\beta_i = sgm'(s_i)\varepsilon_i \quad (38)$$

Se o elemento i encontra-se nas demais camadas, tem-se:

$$\beta_i = sgm'(s_i) \sum_{k \in R(j)} w_{ij} \beta_k \quad (39)$$

sendo:

$sgm'(s_i)$: derivada da função sigmoide dada pelas equações (30) e (31), respectivamente, com relação a s_i .

$R(j)$: conjunto dos índices dos elementos que se encontram na fileira seguinte à fileira do elemento i e que estão interligados ao elemento i .

A derivada parcial da função sigmoide, dada pela equação (30), pode ser expressa por:

$$\begin{aligned} \frac{\partial y_i}{\partial s_i} &= \frac{\partial}{\partial s_i} \left\{ \frac{1 - e^{-\lambda(s_i - \rho)}}{1 + e^{-\lambda(s_i - \rho)}} \right\} \\ &= \frac{2\lambda e^{-\lambda(s_i - \rho)}}{(1 + e^{-\lambda(s_i - \rho)})^2} \\ &= \frac{1}{2} \lambda \left\{ \frac{(1 + e^{-\lambda(s_i - \rho)})^2}{(1 + e^{-\lambda(s_i - \rho)})^2} - \frac{(1 - e^{-\lambda(s_i - \rho)})^2}{(1 + e^{-\lambda(s_i - \rho)})^2} \right\} \\ &= 0.5\lambda(1 - y_i^2) \end{aligned} \quad (40)$$

A função sigmoide, dada pela equação (31), possui a sua derivada expressa como:

$$\begin{aligned} \frac{\partial y_i}{\partial s_i} &= \frac{\partial}{\partial s_i} \left\{ \frac{1}{1 + e^{-\lambda(s_i - \rho)}} \right\} \\ \frac{\partial y_i}{\partial s_i} &= \lambda y(1 - y) \end{aligned} \quad (41)$$

Logo o algoritmo de retropropagação que descreve segue os seguintes passos:

- 1- Um padrão de entrada X é apresentado à rede, o qual fornece uma saída Y ;
- 2- Usando-se a diferença entre a saída desejada e a saída obtida da rede, obtém-se o erro para cada saída.
- 3- Há a determinação do erro retropropagado pela rede associado à derivada parcial do erro quadrático de cada elemento relacionado ao peso.
- 4- O próximo passo consiste em ajustar os pesos de cada elemento.
- 5- Finalmente, um novo padrão é fornecido à rede e repete-se o processo até que o erro obtido pela rede seja menor que uma tolerância preestabelecida, isto é, que haja convergência da rede.

Os pesos iniciais são fixados aleatórios pertencentes a um intervalo entre 0 e 1, visto que com a iniciação nula quase sempre proporciona uma baixa convergência (VILLALBA; BEL, 2000).

Há, basicamente, duas formas de treinar o algoritmo RP (MINUSSI; SILVEIRA, 1995):

- **Convergência por padrão:** Consiste em ajustar os pesos da rede (considerando-se todas as camadas), fazendo-se com que haja convergência para cada padrão, até que se complete o conjunto de padrões de entrada. Só há convergência total dos padrões quando o erro quadrático é menor que a tolerância preestabelecida para os padrões.
- **Uma iteração por padrão:** Este procedimento é idêntico ao primeiro, porém, fazendo-se somente uma iteração (ajuste de pesos) por padrão (LOPES, 2005).

4.1 PROBLEMAS GERADOS PELO ALGORITMO RETROPROGAÇÃO

Por mais que o algoritmo RP apresente um bom desempenho por sua capacidade de precisão, existem alguns problemas que comprometem o seu desempenho. Tais problemas são: lentidão no processo de convergência e não convergência, que podem ser causados por:

1. **Paralisia da rede:** Durante o treinamento da rede, os pesos podem ser ajustados assumindo valores altíssimos, e isso pode causar essa paralisia da rede. Quando a entrada total de uma camada intermediária ou de uma camada de saída alcançar valores muito grandes (positivo ou negativo), em decorrência da função de ativação sigmoideal, a unidade terá uma ativação muito próxima de zero ou muito próxima de um. Com isso, os ajustes de pesos estarão próximos de zero, e o processo de treinamento pode chegar a uma completa paralisia (KROSE; SMAGT, 1996).
2. **Mínimos locais:** O confinamento em um mínimo local ocorre quando existe um mínimo próximo, muito profundo que se dá em consequência do gradiente descendente. Existem métodos probabilísticos que podem ajudar evitar esta “armadilha”, mas como consequência tende a tornar o processo de convergência lento. O aumento do número de camada é outra possibilidade que contribuirá para o aumento da dimensão do espaço do erro (KROSE; SMAGT, 1996).

3. **Convergência lenta:** A oscilação do gradiente sobre a superfície do erro, o fato de que os pesos são gerados aleatoriamente e os mínimos locais são os principais fatores que causam a lentidão no processo de convergência da rede.

4.2 MELHORIAS NO ALGORITMO RETROPROPAGAÇÃO

4.2.1 Adaptação dos parâmetros da rede

Na busca por contornar os problemas gerados pelo algoritmo RP, propõe-se melhorias nos parâmetros da rede. Parâmetros tais como o da inclinação e o da translação da função sigmoide que são denotados como, λ, ρ , respectivamente, cujo objetivo é que essa adaptação da translação amplie o espaço de busca da solução do problema por meio do deslocamento da função e da inclinação sigmoide. Os ajustes desses parâmetros são feitos usando-se o método do gradiente descendente, baseado no algoritmo RP sem o termo momento. O ajuste da translação e da inclinação da função sigmoide é semelhante ao ajuste dos pesos, considerando o i -ésimo neurônio empregando-se o método do gradiente descendente.

a) Adaptação do parâmetro de inclinação da função sigmoide

O ajuste do parâmetro de inclinação da função sigmoide pode ser dado pela equação (42) (STAMATIS, 1999):

$$\lambda_i(h + 1) = \lambda_i(h) + \theta_i^\lambda(h) \quad (42)$$

em que:

$$\theta_i^\lambda(h) = -\gamma_\lambda [\nabla_i^\lambda(h)]$$

$\nabla_i^\lambda(h)$: gradiente do erro quadrático com relação a inclinação do neurônio i ;

λ_i : vetor contendo a inclinação do neurônio i .

$$= [\lambda_{0i} \lambda_{1i} \lambda_{2i} \dots \lambda_{ni}]^T.$$

O vetor gradiente ∇_i^λ é representado por:

$$\nabla_i^\lambda(h) = \frac{\partial \varepsilon_i^2}{\partial \lambda_i(h)} = 2\varepsilon_i \frac{\partial \varepsilon_i}{\partial \lambda_i(h)} \quad (43)$$

Novamente equacionando a equação (27), porém com relação ao vetor λ_i , obtém-se a equação (44):

$$\frac{\partial \varepsilon_i}{\partial \lambda_i} = -\frac{\partial y_i}{\partial \lambda_i} = -\sigma_i^\lambda \quad (44)$$

em que:

σ_i^λ $\underline{\Delta}$ derivada da função sigmoide com relação a λ_i .

$$\beta \lambda_i(h+1) = \lambda_i(h) + 2\gamma \lambda \beta_i^\lambda \quad (45)$$

Se o elemento i encontrar-se na última camada, então:

$$\beta_i^\lambda = \sigma_i^\lambda \varepsilon_i \quad (46)$$

Se o elemento i encontrar-se nas demais camadas, tem-se:

$$\beta_i^\lambda = \sigma_i^\lambda \sum \omega_{ik} \beta_k^\lambda \quad (47)$$

A derivada da função sigmoide da equação (31) em relação λ_i , σ_i^λ pode ser descrita pela equação (48):

$$\frac{\partial y_i}{\partial \lambda_i} = \frac{\partial}{\partial \lambda_i} \left\{ \frac{1}{1+e^{-\lambda(s_i-\rho)}} \right\} = (s_i - \rho)y(1-y^2) \quad (48)$$

sendo:

λ : inclinação da função sigmoide;

ρ : translação da função sigmoide;

s_i : saída do neurônio.

b) Adaptação do parâmetro de translação da função sigmoide

De modo similar a adaptação da inclinação, ajusta-se o parâmetro da translação da função sigmoide, usando-se a equação (49) (STAMATIS, 1999):

$$\rho_i(h+1) = \rho_i(h) + \theta_i^\rho(h) \quad (49)$$

em que:

$$\theta_i^\rho(h) = -\gamma_\rho [\nabla_i^\rho(h)]$$

$\nabla_i^\rho(h)$: gradiente do erro quadrático com relação à translação do neurônio i ;

ρ_i : vetor contendo a translação do neurônio i .

$$= [\rho_{0i} \rho_{1i} \rho_{2i} \dots \rho_{ni}]^T$$

O vetor gradiente ∇_i^ρ é representado por:

$$\nabla_i^\rho(h) = \frac{\partial \varepsilon_i^2}{\partial \rho_i(h)} = 2\varepsilon_i \frac{\partial \varepsilon_i}{\partial \rho_i(h)} \quad (50)$$

Novamente, diferenciando-se a equação (27), porém com relação ao vetor ρ_i , obtém-se a equação (51):

$$\frac{\partial \varepsilon_i}{\partial \rho_i} = -\frac{\partial y_i}{\partial \rho_i} = -\sigma_i^\rho \quad (51)$$

em que:

$\sigma_i^\rho \triangleq$ derivada da função sigmoide com relação a ρ_i .

Logo, a regra que define a adaptação da função sigmoide é dada pela equação (52):

$$\rho_i(h+1) = \rho_i(h) + 2\gamma \rho \beta_i^\rho \quad (52)$$

Se o elemento i encontrar-se na última camada, então:

$$\beta_i^\rho = \sigma_i^\rho \varepsilon_i \quad (53)$$

Se o elemento i encontrar-se nas demais camadas, tem-se:

$$\beta_i^\rho = \sigma_i^\rho \sum \omega_{ik} \beta_k^\rho \quad (54)$$

A derivada da função sigmoide da equação (31) em relação ρ_i , σ_i^ρ pode ser descrita pela equação (55):

$$\frac{\partial y_i}{\partial \rho_i} = \frac{\partial}{\partial \rho_i} \left\{ \frac{1}{1+e^{-\lambda(s_i-\rho)}} \right\} = -\lambda y(1-y) \quad (55)$$

sendo:

λ : inclinação da função sigmoide;

ρ : translação da função sigmoide;

s_i : saída do neurônio.

c) Adaptação da Taxa de Treinamento Via Controlador Fuzzy

O objetivo da adaptação da taxa de treinamento via controlador *fuzzy* é de se obter uma estrutura de controle que faça com que o erro global tenda a zero em um número mínimo de iterações, se comparado ao algoritmo retropropagação convencional. Para isso, necessita-se determinar o estado do sistema definido como sendo o erro global (eg) e a variação do erro global (Δeg).

Processo de Controle

A ideia central por trás do controlador *fuzzy* do algoritmo Retropropagação é a implementação da heurística nas regras dos termos *fuzzy if-then*. Para o ajuste dessa taxa faz-se o seguinte procedimento. Quando o treinamento via o algoritmo RP, é realizado, gera-se um erro global (εg) e uma variação do erro global ($\Delta \varepsilon g$). Esses parâmetros serão os componentes do estado do sistema e $\Delta \gamma$ a ação de controle que deve ser exercida sobre o sistema, fazendo com que o erro tenda a zero em um número menor de iterações, se comparado ao algoritmo retropropagação convencional (WERBOS, 1974).

O erro global é definido como sendo:

$$\varepsilon g = \sum_{j=1}^{np} \sum_{i=1}^{ns} \varepsilon_i^2 \quad (56)$$

εg : erro global da rede neural;

np : número de vetores padrão da rede;

ns : número de neurônios da última camada da rede.

O estado do sistema e ação de controle são definidos por (LOPES; MINUSSI, 2000):

$$E^k = [g^k \quad \dot{g}^k]^T \quad (57)$$

$$u^k = \Delta \gamma^k \quad (58)$$

em que:

k : iteração corrente;

$\Delta \gamma^k$: controle adaptativo.

Utiliza-se uma função exponencial decrescente aplicada à resposta do controlador *fuzzy* para obter o controle adaptativo. Dessa maneira o controle adaptativo é definido como (LOPES, 2000):

$$\Delta\gamma^k = \exp(-\alpha_p k)\Delta\psi^k \quad (59)$$

em que:

α_p : um número positivo arbitrário;

$\Delta\psi^k$: acréscimo procedente do controlador nebuloso no instante k .

O parâmetro α exerce uma função de decaimento exponencial imposto pelo controlador cujo objetivo é de evitar que rede neural entre em saturação.

Essa ação de controle é feita usando-se a lógica *fuzzy*. Determina-se um intervalo entre 0 e 1, onde tem-se as variáveis como, por exemplo, PMP (Positivo Muito Pequeno), PMG (Positivo Muito Grande), PP (Positivo Pequeno), ME (Médio), PG (Positivo Grande), PMG (Positivo Muito Grande). O erro global e sua variação serão componentes do estado do sistema, e podem pertencer a esses intervalos predeterminados. Com esses valores, aplica-se a base de regras produzindo os valores das variáveis linguísticas. Posteriormente, é executada a rotina de defuzificação empregando-se o método centroide. Neste caso, é devolvido um novo valor para a taxa de treinamento a cada iteração e assim por diante até a convergência da rede.

O processo, utilizando-se o controlador, é muito simples e de fácil implementação. Além disso, o sistema de controle requer um esforço reduzido por ter como entrada somente duas variáveis e uma de saída. Essa proposta da utilização do controlador *fuzzy*, no algoritmo, foi proposta por Arabshahi et al. (1996), que utiliza as variáveis do erro global (εg) e a variação do erro global ($\Delta\varepsilon g$) para efetuar o controlador *fuzzy* (MINUSSI; LOTUFO, 2011).

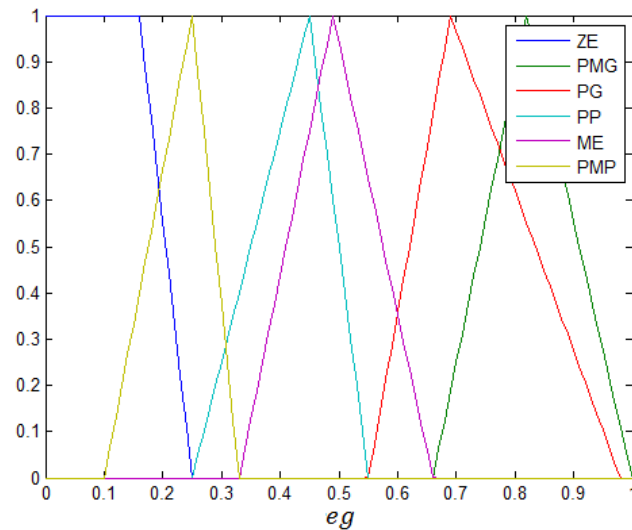
Para a codificação das regras, forma-se uma tabela de decisões, onde cada entrada representa o valor da variável *fuzzy* $\Delta\gamma$ (variação da taxa de treinamento) a partir dos valores do erro global (εg) e da variação do erro global ($\Delta\varepsilon g$). Existe um limite de variação de $\Delta\gamma$, que se chama parâmetro lu . Este parâmetro deve ser ajustado de acordo com cada problema e em função da inclinação (λ) da função de ativação.

A variável εg está compreendida entre 0 e 1. Se a heurística da adaptação estiver sintonizada, a convergência do processo será exponencialmente decrescente (LOPES;

MINUSSI; LOTUFO, 2000) A variável Δeg deverá variar entre -1 e 1 se o processo for exponencialmente decrescente os valores de Δeg deverá ser sempre negativos. Neste caso, embora a escala de Δeg esteja entre -1 e 1 deve se empregar no conjunto de regras que proporcione um ajuste fino entre -1 e 0. No outro intervalo (0, 1] o ajuste poderá ser mais relaxado.

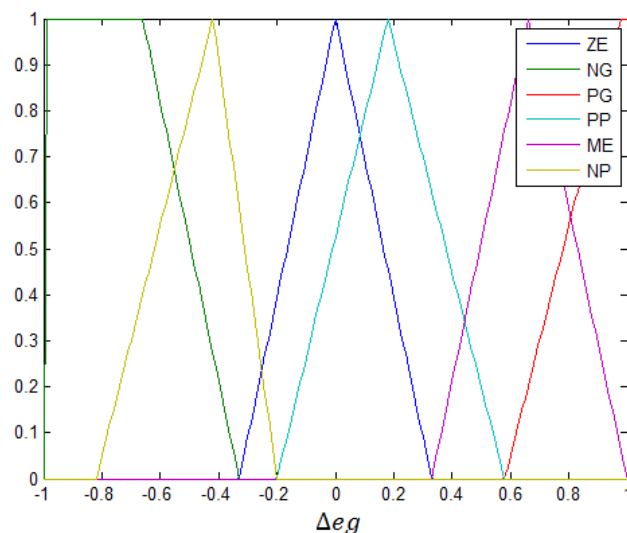
As funções de pertinência, para variáveis envolvidas no treinamento da rede neural com o algoritmo RP, via controlador *fuzzy*, são ilustradas nas Figuras 13, 14 e 15, respectivamente.

Figura 13 - Funções de pertinência para o erro global eg

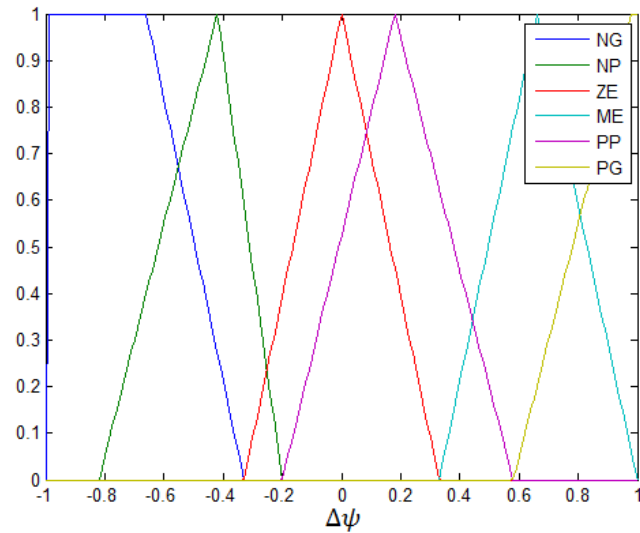


Fonte Dados da pesquisa da autora.

Figura 14 - Funções de pertinência para a variação do erro global Δeg



Fonte Dados da pesquisa da autora.

Figura 15 - Funções de pertinência para a ação de controle $\Delta\psi$ 

Fonte Dados da pesquisa da autora.

sendo:

NG : Negativo Grande;

NP : Negativo Pequeno;

ZE : Zero;

PMG : Positivo Muito Grande;

PG : Positivo Grande;

PP : Positivo Pequeno;

ME : Médio;

PMP : Positivo Muito Pequeno.

Tabela 1 - Regras do controlador *fuzzy*.

<i>Variação do Erro</i>	<i>Erro</i>					
	ZE	PMP	PP	ME	PG	PMG
NGv	NG	NG	NG	-	-	-
NPv	NP	NP	NP	-	-	-
ZEv	ZE	ZE	ZE	-	-	ZE
PPv	PP	ZE	PP	-	-	PP
MEv	ME	ME	ME	-	-	-
PGv	PG	PG	PG	-	-	-

Fonte Dados da pesquisa da autora.

Na tabela 1 é apresentado o conjunto de regras *fuzzy*, totalizando 20 regras. O número de regras poderá ser aumentado visando o melhor desempenho no treinamento da rede.

Pode-se notar que algumas regras não possuem uma ação de controle devido ao fato de assumir valores muito altos e, conseqüentemente, para o problema em questão não satisfaria o objetivo de acelerar o processo de convergência da rede.

4.3 ANÁLISE DE PRECISÃO

Para realizar a análise de precisão, foi definido o erro percentual absoluto médio MAPE (SRINIVASAN et al., 1998) e o erro máximo da previsão diária, comparando os valores reais da carga com os valores estimados pela rede neural da seguinte forma:

$$MAPE = \frac{1}{N} \{ \sum_{i=1}^N (|L(h) - \underline{L}(h)|) / L(h) \} \times 100 \quad (60)$$

$$Erro\ Máximo(\%) = \max\{|L(h) - \underline{L}(h)| / L(h)\} \times 100 \quad (61)$$

sendo:

$L(h)$: valor da carga real referente à hora h ;

$\underline{L}(h)$: valor da carga estimada pela rede referente à hora h ;

N : número de horas total.

4.4 COMENTÁRIOS

Neste capítulo foram apresentados o equacionamento do algoritmo RP convencional e os dois principais tipos de treinamento envolvendo o algoritmo: (1) convergência por padrão e (2) iteração por padrão. Também, foram discutidos os problemas gerados pelo algoritmo tais como: a paralisia da rede, os mínimos locais e a lentidão no processo de convergência. A partir disso, foram propostos melhorias no algoritmo, tais como: a adaptações da inclinação e da translação da função de ativação (função sigmoide), assim como a adaptação da taxa de treinamento realizada via controlador nebuloso, fazendo com que a rede escape dos mínimos locais, não entre em paralisia e minimize o tempo de convergência.

Capítulo 5

5 METODOLOGIA DESENVOLVIDA PARA PREVISÃO DE CARGA ELÉTRICA

O objetivo do trabalho em questão é realizar a previsão multimodal de carga que pode ser composta por subestações, barramento, transformadores etc.

Será apresentado o esquema proposto nesta pesquisa visando à previsão multimodal da carga elétrica. A previsão é realizada por meio de um sistema inteligente considerando vários pontos da rede elétrica. A pesquisa foi desenvolvida para previsão do sistema de distribuição de energia elétrica em consequência da grande importância de conhecer o comportamento da demanda em diversos barramentos do sistema.

A previsão é realizada baseada em uma rede neural composta por módulos. Cada módulo foi implementado utilizando o algoritmo de Retropropagação. Porém, algumas alterações foram realizadas no algoritmo a fim de surgir melhorias nos resultados. Tais alterações como, por exemplo, a adaptação da taxa de treinamento via controlador *fuzzy* e ajuste da inclinação e da translação usando o método do gradiente descendente proporcionaram uma melhoria significativa se comparado ao algoritmo convencional.

5.1 METODOLOGIA DESENVOLVIDA

A previsão de carga é uma atividade de grande importância para o planejamento e a operação das concessionárias fornecedoras de energia elétrica. Por meio de uma previsão confiável, a concessionária poderá planejar e gerar energia elétrica de qualidade e sem desperdício seja hídrica, térmica ou até mesmo eólica e solar. Estudos visam proporcionar ferramentas, aos engenheiros, e informações vitais para que o sistema elétrico possa ser operado de forma segura e eficaz. Na maioria das metodologias para previsão de carga, via redes neurais, emprega-se o algoritmo de Retropropagação, que é um algoritmo que adapta os pesos da rede baseado no método do gradiente descendente (WIDROW; LEHR, 1990).

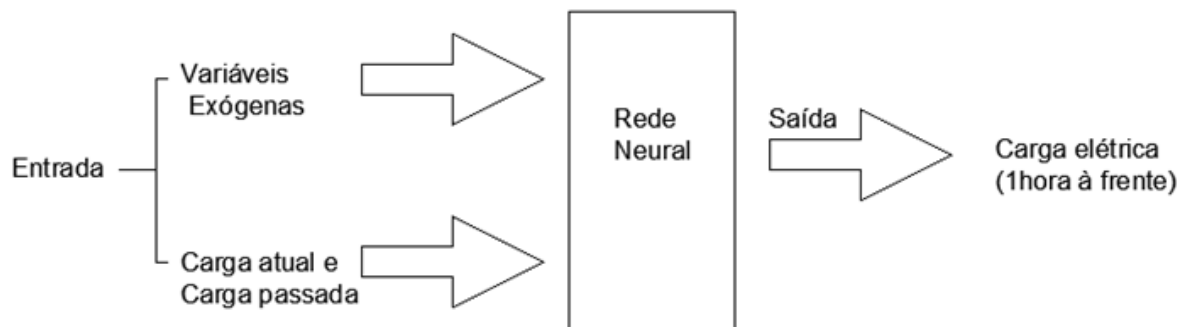
Observa-se através da literatura especializada, que esta arquitetura é bastante utilizada, porém, é uma arquitetura que apresenta determinados problemas que interferem na eficiência e rendimento da rede neural. Com a finalidade de contornar os problemas de

convergência, saturação e mínimos locais foram realizadas alterações na estrutura da rede neural visando a sua melhoria.

Portanto, como contribuição tem-se a utilização das melhorias na estrutura da rede neural: controlador *fuzzy* para adaptar a taxa de treinamento da rede e adaptação dos parâmetros de inclinação e translação da função de ativação da rede neural utilizados na previsão de carga elétrica multimodal a curto prazo.

Foram realizados dois procedimentos de previsão visando a previsão de carga global e a previsão de carga modular. Na Figura 16 descreve-se o procedimento de como foi realizado a inserção dos dados de entrada para o sistema previsor global

Figura 16- Modelo de previsão da carga global via redes neurais.



Fonte: Altran (2010).

Há dois conjuntos de dados de entrada: dados referentes à série histórica da carga elétrica e os dados exógenos. Para os dados referentes à série histórica da carga elétrica é utilizado um procedimento usual na área de previsão de carga que é o movimento de uma janela (PARK; SHARKAWI; MARK II, 1991), cujo deslocamento é arbitrário. Já os dados exógenos são informações importantes para o processo de previsão que pode ser, por exemplo, estação, mês do ano, dias da semana, dados climáticos, etc. Nesta pesquisa foram utilizados como dados exógenos os dias da semana (domingo a segunda) e a hora do dia (24 horas), todos binarizados, já o valor da carga foi normalizada. Para normalizar o valor da carga utilizou o valor da carga na hora h dividido pelo valor da carga máxima acrescido de 30% seu valor.

O problema da previsão global, considerando o i -ésimo módulo, pode ser formulado da seguinte forma:

$$X(h) = [t^T L(h-3)L(h-2)L(h-1)L(h)]^T, \quad X \in R^m \quad (62)$$

$$Y(h) = [L(h+1)], \quad Y \in R^1 \quad (63)$$

sendo:

m : dimensão do vetor X ;

$L(h-p)$: valor da carga p horas anterior à hora corrente h ;

$L(h+1)$: valor da carga elétrica da hora subsequente à hora corrente h ;

t : vetor referente ao tempo dos dados históricos (dia da semana, hora, etc.) representados entre $[0,1]$;

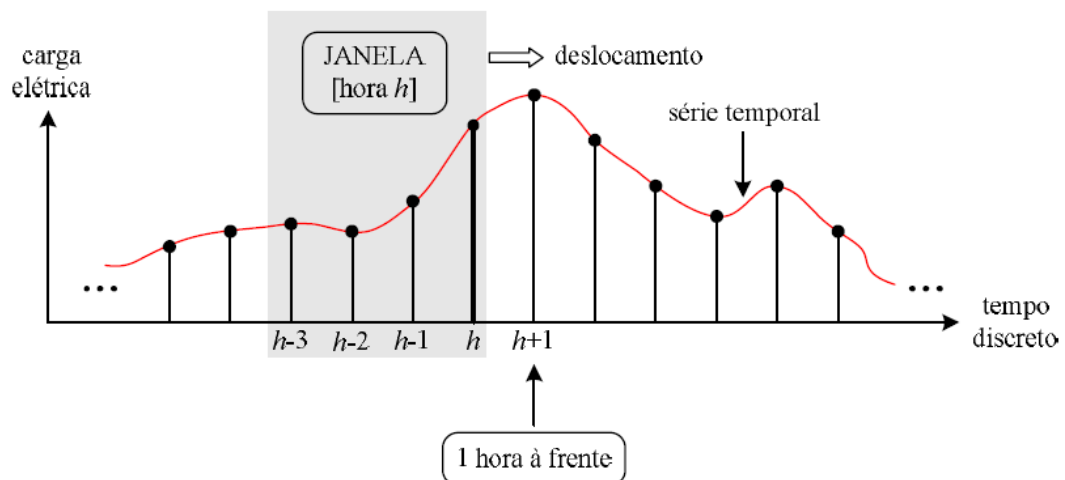
$X(h)$: vetor de entrada da rede neural correspondente à hora h ;

$Y(h)$: estímulo aplicado na saída da rede neural correspondente à hora $h+1$.

Para obter os resultados do sistema predictor global foi utilizado o movimento de janela, isto é, cada janela possui dados das cargas elétricas relativas às 3 horas passadas e a carga atual. As cargas elétricas, $L(h-3), \dots, L(h-1)$, representam o laço de realimentação da saída, constituindo uma rede neural recorrente.

Este processo consiste no movimento da “janela” (DJUKANOVIC et al., 1995) , de tamanho m (m componentes), representado na figura 17.

Figura 17 - Estratégia de previsão via deslocamento de janela.



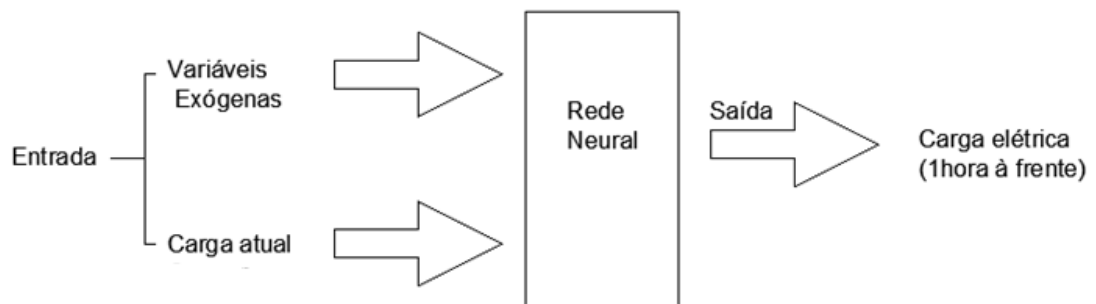
Fonte: Altran (2010).

É importante enfatizar que para ambas as previsões, global e modular, foram utilizadas a rede neural com algoritmo retropropagação convencional, rede neural com algoritmo retropropagação com momento adaptando a taxa de treinamento através do controlador *fuzzy* e a rede neural com algoritmo retropropagação sem momento com a adaptação dos parâmetros de inclinação e translação da função de ativação da rede neural (função sigmoide).

Para o sistema predictor modular não foi utilizada a carga das horas anteriores, foi utilizada somente a carga atual.

A figura 18 ilustra o processo esquemático utilizado na previsão de carga elétrica modular. É possível notar que uma das diferenças existentes com relação ao sistema predictor global é que a entrada de dados para os sistemas modulares é composta por dados exógenos (dia da semana e a hora do dia) e o valor da carga elétrica atual.

Figura 18- Modelo de previsão de carga modular via redes neurais.



Fonte: Altran (2010).

O problema de previsão modular, considerando o i -ésimo módulo, pode ser formulado da seguinte forma:

$$X(h) = [t^T L(h)]^T, \quad X \in R^m \quad (64)$$

$$Y(h) = [L(h + 1)], \quad Y \in R^1 \quad (65)$$

sendo:

m : dimensão do vetor X ;

$L(h + 1)$: valor da carga elétrica da hora subsequente à hora corrente h ;

t : vetor referente ao tempo dos dados históricos (dia da semana, hora, etc.) representados entre $[0,1]$;

$X(h)$: vetor de entrada da rede neural correspondente à hora h ;

$Y(h)$: estímulo aplicado na saída da rede neural correspondente à hora $h + 1$.

Outras estratégias de entrada de dados podem ser utilizadas de acordo com o objetivo do projeto.

O sistema previsor global é composto por uma rede neural responsável pela previsão da carga global. O sistema previsor local é composto por vários módulos, sendo cada módulo representado por uma rede neural responsável pela previsão local.

5.2 COMENTÁRIOS

Neste capítulo foi apresentada a metodologia com vista à realização da previsão de carga: o sistema previsor global e o sistema previsor local. O previsor local é composto por três módulos cada um correspondendo a um barramento. Mostrou-se como foi feita a inserção dos dados de entrada, via movimento de uma janela para o sistema global. Foi apresentada a rotina de treinamento envolvendo esse tipo de estrutura, os tipos de dados os exógenos (dados climáticos, meses, estações do ano, dias da semana) e as cargas elétricas, onde os dados exógenos influenciam significativamente na previsão.

6 RESULTADOS OBTIDOS

Neste capítulo são apresentados os resultados obtidos considerando a rede neural proposta. São apresentados os parâmetros utilizados pela rede durante simulações e os resultados gráficos e numéricos das simulações. Realiza-se também a análise de todos os resultados numéricos obtidos do algoritmo RP convencional comparando com as melhorias que se obteve ao inserir o controlador *fuzzy* e com a implementação das adaptações da inclinação e da translação. Todos os testes do algoritmo foram realizados usando o *software* MATLAB R2013b.

6.1 PARÂMETROS UTILIZADOS NA REDE

Cada aplicação contém uma tabela de parâmetros individual, pois a cada proposta de aplicação mudam-se os parâmetros, que são parâmetros importantes como o número de camada, o número de neurônio por camada, tolerância, taxa de treinamento e a taxa de aprendizagem da inclinação e da translação.

Nenhum método determinístico foi realizado para obter os parâmetros da rede. Tais parâmetros foram obtidos por meio de tentativa e erro.

Observa-se que número de neurônio de entrada do sistema previsor global (12 neurônios) é maior do que a do sistema previsor local (9 neurônios). Esta diferença se dá pelo fato de ter se considerado a técnica de movimento em janela somente no sistema previsor global, e não no sistema previsor local. A entrada da rede para o sistema previsor global é definida como sendo os valores da carga extraídos dos dados históricos em três instantes anteriores, o valor da carga atual e os dados exógenos, isto é, a hora do dia e o dia da semana e a saída da rede neural tanto para o sistema previsor global quanto para o sistema previsor local fornece o valor da carga no instante $(h+1)$.

6.2 RESULTADOS

A previsão abordada, nesta pesquisa, é a previsão de carga de curto prazo. Os dados históricos apresentados foram fornecidos por uma companhia do setor elétrico da Nova

Zelândia (*Electricity Comission* da Nova Zelândia), utilizados para realizar o treinamento da arquitetura da rede neural proposta. Os dados utilizados no presente trabalho possuem os mesmos horizontes de previsão utilizados na previsão modular na tese de Altran (ALTRAN, 2010), visando comparar a eficiência da rede neural. Foram considerados os dados das cargas horárias de 15 dias, compreendidos entre os dias 15 e 29 do mês de outubro do ano de 2007, num total de 360 valores de entrada. Foi realizada uma análise gráfica e numérica dos resultados para o previsor global e para cada módulo do previsor local.

Os resultados foram obtidos via MATLAB R2013b. Primeiramente, apresentam-se os resultados numéricos e gráficos do algoritmo RP convencional, que será utilizado para comparar com a metodologia proposta. Em seguida, os resultados numéricos e gráficos do treinamento da rede com o algoritmo RP, via controlador *fuzzy*, utilizado para adaptar a taxa de aprendizagem dos pesos durante o processo de treinamento. Por fim, é realizada uma comparação separadamente do desempenho dos sistemas no que tange MAPE, erro máximo, número de épocas e o tempo de processamento.

6.2.1 Análise numérica e gráfica para os sistemas previsores com o algoritmo RP convencional

A tabela 2 mostra os parâmetros utilizados pela rede durante o treinamento do algoritmo RP convencional. Nesse treinamento, arbitram-se a taxa de aprendizagem, a tolerância e a quantidade de neurônio na camada intermediária da rede neural, os pesos iniciais são gerados aleatoriamente.

Tabela 2- Parâmetros utilizados no treinamento com o algoritmo RP convencional.

Parâmetros	Global	Módulo 1	Módulo 2	Módulo 3
Número de camada	3	3	3	3
Quantidade de neurônio (por camada)	12-20-1	9-20-1	9-20-1	9-20-1
Tolerância	0,0002	0,0002	0,0002	0,0002
Taxa de Treinamento	0,26	0,25	0,2	0,24
Taxa de aprendizagem da inclinação	-	-	-	-
Taxa de aprendizagem da Translação				

Fonte: Dados de simulações

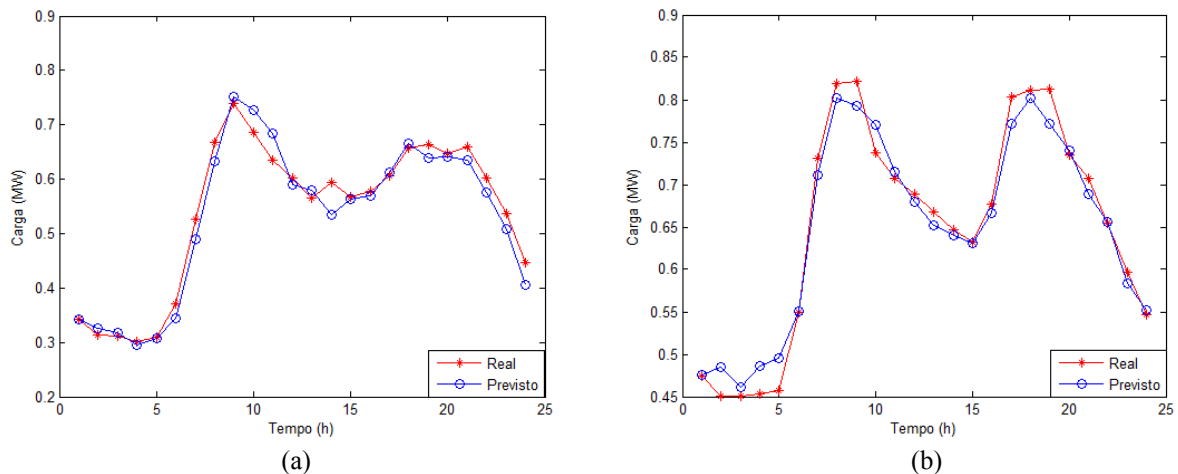
Na tabela 3 é realizada uma análise numérica para os sistemas previsores, global e os modulares, utilizando o algoritmo retropropagação convencional, apresentando os resultados para a análise de precisão MAPE e Erro Máximo obtido no treinamento da rede, tempo de execução e o número de épocas. Na figura 19 são ilustradas as curvas da previsão realizada, sendo a curva na cor vermelha a carga real e a curva na cor azul a carga prevista pela rede via algoritmo RP convencional.

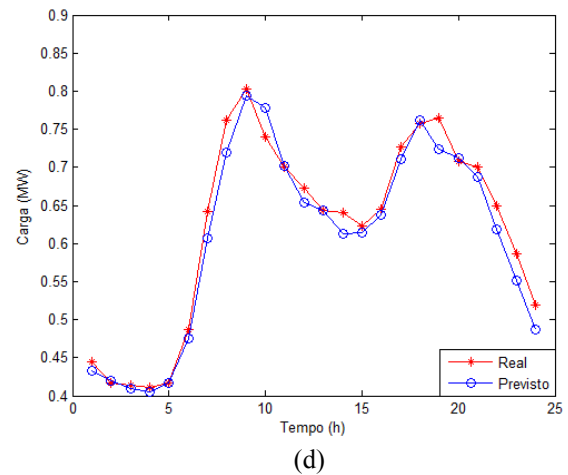
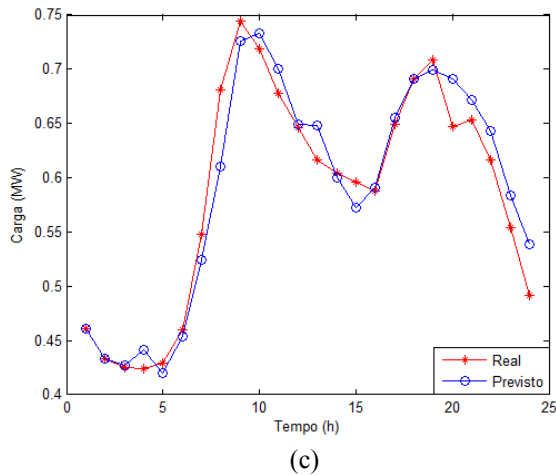
Tabela 3 - Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e para o sistema predictor global treinado com o algoritmo RP convencional.

Sistema Predictor	MAPE	Erro Máximo	Tempo	Época
Global	2,62	6,19	53,94	3001
Módulo 1	3,71	10,25	20,88	1088
Módulo 2	2,64	8,56	5,83	211
Módulo 3	3,02	10,40	6,76	371

Fonte: Dados do próprio autor

Figura 19 - Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema predictor Global (d) treinados com o algoritmo RP convencional





Fonte: Dados do próprio autor

6.2.2 Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação via controlador fuzzy

A tabela 4 mostra os parâmetros utilizados pela rede durante o treinamento do algoritmo RP via controlador *fuzzy*. Nesse treinamento define-se uma tolerância, a quantidade de neurônio na camada intermediária, os pesos iniciais são gerados aleatoriamente.

A partir de uma taxa de treinamento inicial e através do controlador *fuzzy* utilizado no processo de treinamento da rede neural os pesos são ajustados a cada iteração (época). O uso do controlador *fuzzy* faz com que a taxa de treinamento a cada iteração seja modificada de tal forma que o erro seja minimizado e o processo de convergência da rede neural seja acelerado.

Tabela 4- Parâmetros utilizados no treinamento com o algoritmo RP via controlador *fuzzy*.

Parâmetros	Global	Módulo 1	Módulo 2	Módulo 3
Número de camada	3	3	3	3
Quantidade de neurônio (por camada)	12-22-1	9-20-1	9-20-1	9-20-1
Tolerância	0,0002	0,0002	0,0002	0,0002
Taxa de treinamento (inicial)	0,26	0,21	0,19	0,20
Taxa de aprendizagem da inclinação	-	-	-	-
Taxa de aprendizagem da translação	-	-	-	-

Fonte: Dados do próprio autor

Realizam-se a análise numérica e gráfica, considerando-se os sistemas previsores, global e os modulares, utilizando o algoritmo RP via controlador *fuzzy*. Na tabela 5 apresentam-se os resultados de análise de precisão MAPE e Erro Máximo (obtido pelo treinamento da rede), tempo de execução e o número de épocas.

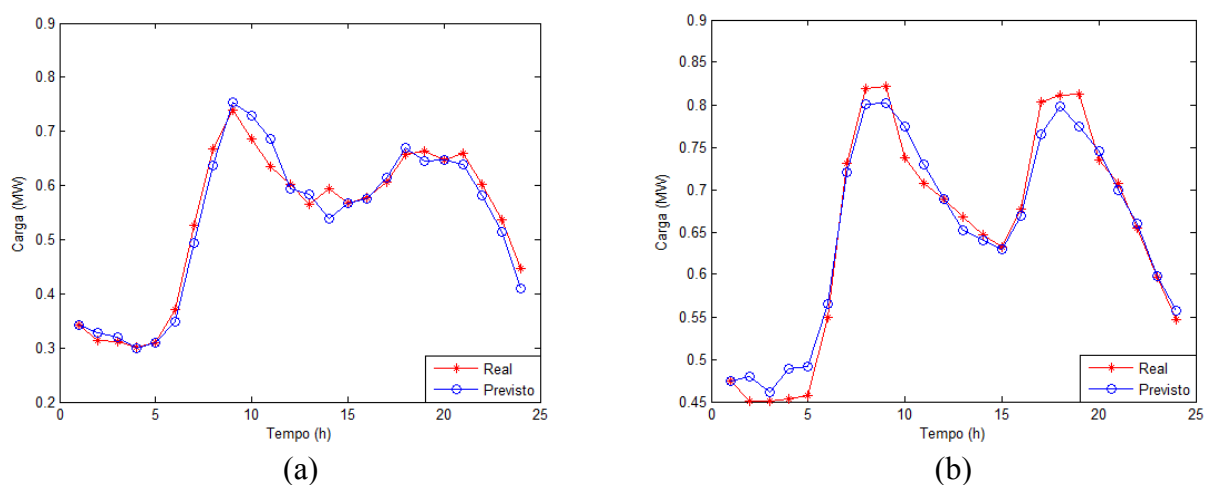
Tabela 5 - Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema predictor global treinado com o algoritmo RP via controlador *fuzzy*.

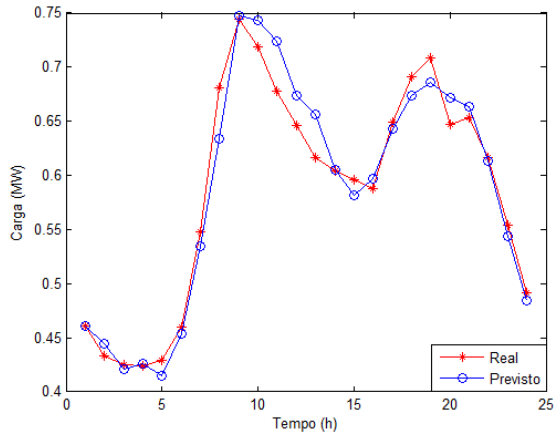
Sistemas Previsores	MAPE	Erro Máximo	Tempo	Épocas
Global	1,98	5,96	3,07	73
Módulo 1	3,41	9,52	16,85	1144
Módulo 2	2,61	8,01	2,43	52
Módulo 3	2,45	6,86	2,97	83

Fonte: Dados do próprio autor

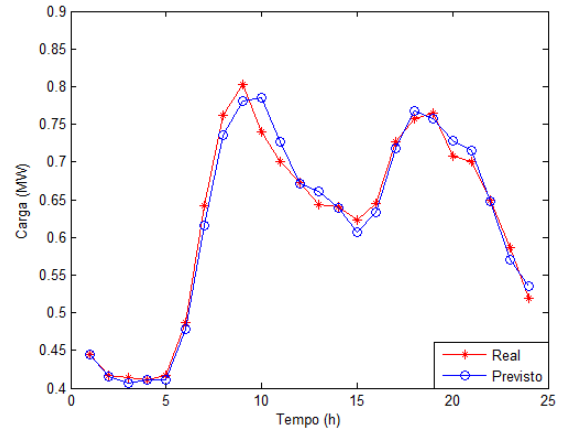
A figura 20 mostra as curvas de carga global e modular (módulos 1, 2 e 3, respectivamente) para a rede neural convencional e a rede neural com adaptação da taxa de treinamento via controlador *fuzzy*.

Figura 20 - Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema predictor Global (d) treinados com o algoritmo RP via controlador *fuzzy*.





(c)



(d)

Fonte: Dados do próprio autor

6.2.3 Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação com a adaptação da inclinação e da translação da função sigmoide

A tabela 6 mostra os parâmetros utilizados pela rede durante o treinamento do algoritmo RP com a adaptação da inclinação e da translação da função sigmoide. Nesse treinamento é inserido na rede uma tolerância, a quantidade de neurônio na camada intermediária, os pesos são gerados aleatoriamente de acordo com a adaptação que ele sofre durante o processo de treinamento e uma taxa de aprendizagem e os parâmetros iniciais da inclinação e da translação da função sigmoide, durante o processo de treinamento da rede os valores da inclinação e da translação são adaptados através do gradiente descendente.

Tabela 6- Parâmetros utilizados no treinamento com o algoritmo RP com a adaptação da inclinação e da translação da função sigmoide.

Parâmetros	Global	Módulo 1	Módulo 2	Módulo 3
Número de camada	3	3	3	3
Quantidade de neurônio (por camada)	12-22-1	9-20-1	9-20-1	9-20-1
Tolerância	0,0002	0,0002	0,0002	0,0002
Taxa de treinamento	0,18	0,16	0,20	0,33
Taxa de aprendizagem da inclinação (inicial)	1,50	0,0001	0,25	0,27
Taxa de aprendizagem da translação (inicial)	1,41	0,10	1,16	1,12

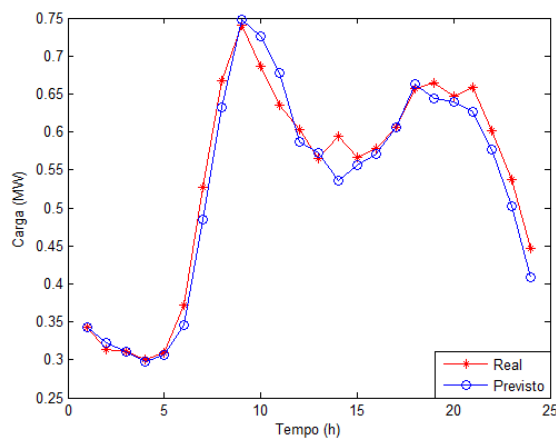
Fonte: Dados do próprio autor

Tabela 7 - Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema predictor global treinado com o algoritmo RP com as adaptações da inclinação e da translação da função sigmoide.

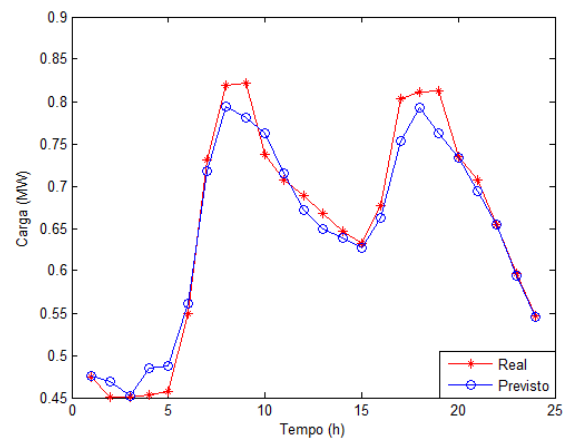
Sistemas	MAPE	Erro Máximo	Tempo	Épocas
Previsores				
Global	2,47	6,59	27,72	779
Módulo 1	3,52	9,96	24,37	1357
Módulo 2	2,54	6,98	5,48	186
Módulo 3	2,22	7,60	6,08	236

Fonte: Dados do próprio autor

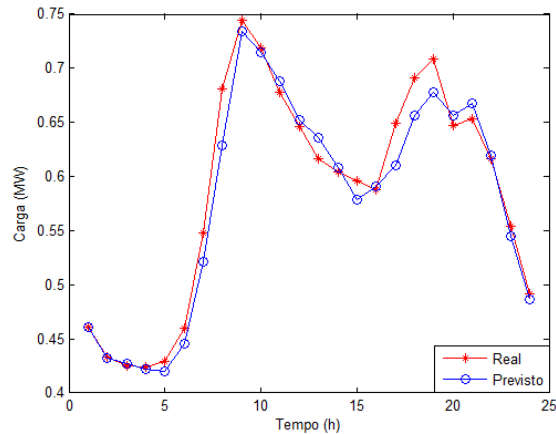
Figura 21 - Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema predictor Global (d) treinados com o algoritmo RP com as adaptações da inclinação e da translação da função sigmoide.



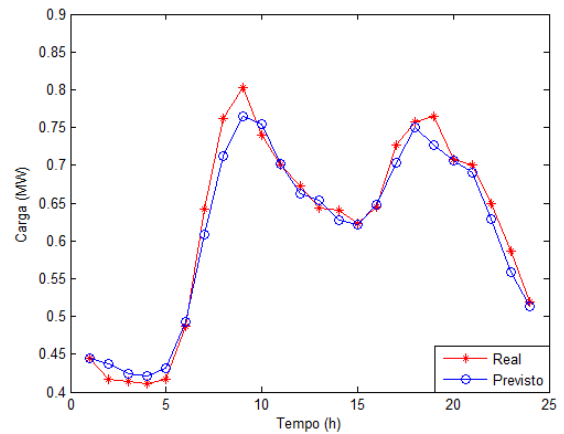
(a)



(b)



(c)



(d)

Fonte: Dados do próprio autor

6.2.4 Análise numérica e gráfica para os sistemas previsores com o algoritmo retropropagação via controlador fuzzy e adaptação da inclinação e da translação da função sigmoide

A tabela 8 mostra os parâmetros utilizados pela rede durante o treinamento do algoritmo RP com a adaptação da inclinação e da translação da função sigmoide e o controlador *fuzzy*. Nesse treinamento é inserido na rede uma tolerância, a quantidade de neurônio na camada intermediária, os pesos são gerados aleatoriamente de acordo com a adaptação que ele sofre durante o processo de treinamento, uma taxa de aprendizagem inicial que muda a cada iteração de acordo com as regras do controlador *fuzzy* e os parâmetros iniciais da inclinação e da translação da função sigmoide, durante o processo de treinamento da rede os valores da inclinação e da translação são adaptados através do gradiente descendente.

Tabela 8- Parâmetros utilizados no treinamento com o algoritmo RP com a adaptação da inclinação e da translação da função sigmoide e o controlador *fuzzy*.

Parâmetros	Global	Módulo 1	Módulo 2	Módulo 3
Número de camada	3	3	3	3
Quantidade de neurônio (por camada)	12-22-1	9-20-1	9-20-1	9-20-1
Tolerância	0,0002	0,0002	0,0002	0,0002
Taxa de treinamento (inicial)	0,21	0,21	0,19	0,33
Taxa de aprendizagem da inclinação (inicial)	1,50	0.000001	0,23	0,27
Taxa de aprendizagem da translação (inicial)	1,41	0.10	1.17	1,12

Fonte: Dados do próprio autor

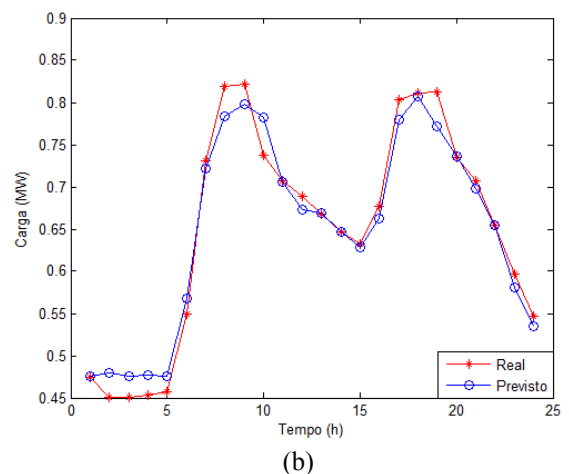
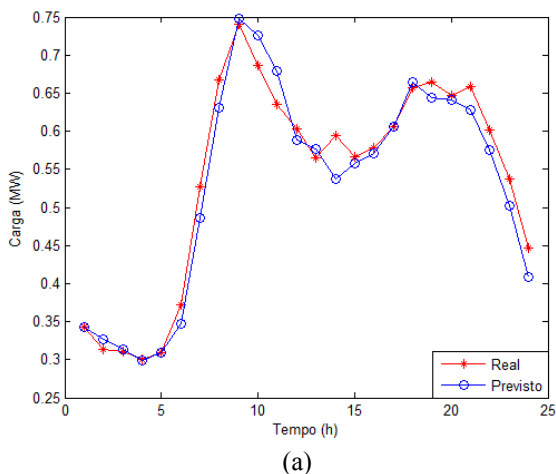
A tabela 9 ilustra a análise numérica para os sistemas previsor global e os modulares, utilizando-se o algoritmo RP, via controlador *fuzzy* e adaptação da inclinação e da translação da função de ativação (sigmoide), na qual são apresentados os resultados de análise de precisão MAPE e Erro Máximo (obtido pelo treinamento da rede), tempo de execução e o número de épocas. Na figura 22 são ilustradas as curvas da previsão realizada, sendo a curva na cor vermelha a carga real e a curva na cor azul a carga prevista pela rede com algoritmo RP via controlador *fuzzy*, e adaptação da inclinação e da translação da função de ativação.

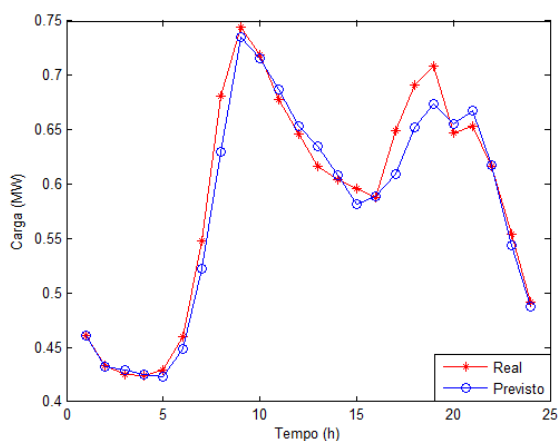
Tabela 9 - Resultados numéricos obtidos para os sistemas previsores locais contendo 3 módulos e o sistema previsor global treinado com o algoritmo RP via controlador *fuzzy* e as adaptações da inclinação e da translação da função sigmoide.

Sistema	MAPE	Erro Máximo	Tempo	Épocas
Previsor				
Global	1,93	4,31	2,77	52
Módulo 1	3,51	9,70	17,96	997
Módulo 2	2,49	6,36	3,77	121
Módulo 3	2,13	7,50	5,34	209

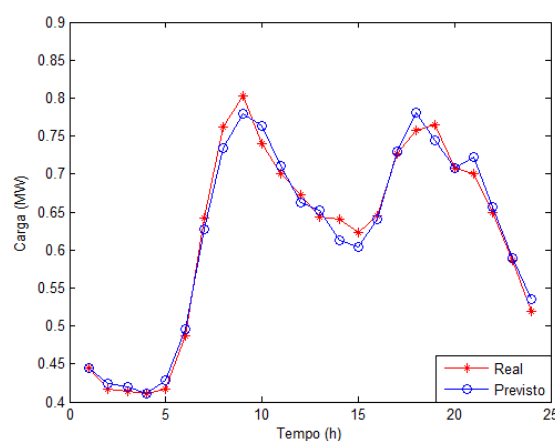
Fonte: Dados do próprio autor

Figura 22 - Análise gráfica realizada para cada um dos sistemas previsores locais sendo: Módulo 1 (a), Módulo 2 (b), Módulo 3 (c) e o sistema previsor Global (d) treinados com o algoritmo RP via controlador *fuzzy* e as adaptações da inclinação e da translação da função de ativação.





(c)



(d)

Fonte: Dados do próprio autor

6.3 COMPARANDO RESULTADOS

O algoritmo RP convencional por si só realiza uma previsão considerada significativa, porém existem meios de aperfeiçoar essa previsão, por esse motivo foi proposta nesta pesquisa a realização de melhorias no algoritmo RP convencional a fim de obter melhores resultados, tanto na precisão da previsão quanto no tempo de processamento durante o processo de convergência, e um dos parâmetros para podermos comparar se a metodologia proposta foi eficaz é através do Erro médio percentual conhecido como MAPE. A tabela 10 mostra a comparação dos valores de MAPE do algoritmo Retropropagação convencional, do algoritmo RP com o controlador *fuzzy*, do algoritmo RP com a adaptação da inclinação e da translação e a junção do controlador *fuzzy* com as adaptações da inclinação e da translação da função sigmoide.

Tabela 10 - Comparando os valores de MAPE do algoritmo RP convencional com o algoritmo com o controlador *fuzzy* e o algoritmo com o controlador *fuzzy* e as adaptações da inclinação e da translação da função de ativação.

Sistema Previsor	Convencional	Controlador <i>Fuzzy</i>	Adaptações	Controlador e as Adaptações
Global	2,62	1,98	2,47	1,93
Módulo 1	3,71	3,41	3,52	3,51
Módulo 2	2,64	2,61	2,54	2,49
Módulo 3	3,02	2,45	2,22	2,13

Fonte: Dados do próprio autor

Comparando o MAPE do sistema predictor global com o treinamento via algoritmo RP convencional com cada uma das propostas podemos perceber uma melhora muito significativa com destaque para o valor obtido com a junção do controlador e as adaptações da inclinação e da translação da função sigmoideal.

Já o sistema predictor composto pelo módulo 1 houve uma melhora significativa utilizando o controlador *fuzzy*, porém pouco significativa se comparada aos outros sistemas predictores, entretanto com o ajuste dos parâmetros pode-se aperfeiçoar a previsão. No sistema predictor composto pelo módulo 2 também houve uma melhoria significativa destacando como o melhor resultado a junção do controlador *fuzzy* que ajusta a taxa de treinamento e as adaptações da inclinação e da translação da função sigmoide através do gradiente descendente. O sistema predictor composto pelo módulo 3 obteve uma melhoria muito significativa obtendo-se o melhor resultado com a junção das duas metodologias.

Além da análise de precisão, usando o conceito da MAPE, tem-se como parâmetro de comparação o erro máximo que será apresentado na tabela 11, no qual cada metodologia proposta é comparada com o algoritmo retropropagação convencional. Pode-se observar que os melhores resultados em termos de erro máximo foram obtidos para a junção das duas metodologias.

Tabela 11- Comparando os valores de Erro Máximo do algoritmo RP convencional com o algoritmo com o controlador *fuzzy* e o algoritmo com o controlador *fuzzy* e as adaptações da inclinação e da translação da função de ativação.

Sistema Predictor	Convencional	Controlador <i>Fuzzy</i>	Adaptações	Controlador e as Adaptações
Global	6,19	5,96	6,59	4,31
Módulo 1	10,25	9,52	9,96	9,70
Módulo 2	8,56	8,01	6,98	6,36
Módulo 3	10,40	6,86	7,60	7,50

Fonte: Dados do próprio autor

Na Tabela 12, realaciona-se a quantidade de épocas, obtida durante o processo de treinamento do algoritmo retropropagação convencional, do algoritmo com o controlador *fuzzy* para ajustar a taxa de aprendizagem durante o processo de treinamento, do algoritmo retropropagação com adaptações da inclinação e da translação por meio do gradiente

descendente e a junção das duas metodologias. Observa-se que o uso do controlador *fuzzy* tem por objetivo acelerar o processo de convergência o que é visto através do número de épocas, tabela 12, e conseqüentemente no tempo de processamento, tabela 13.

Tabela 12 - Comparando os valores do número de épocas do algoritmo RP convencional com o algoritmo com o controlador *fuzzy* e o algoritmo com o controlador *fuzzy* e as adaptações da inclinação e da translação da função de ativação.

Sistema Previsor	Convencional	Controlador <i>Fuzzy</i>	Adaptações	Controlador e as Adaptações
Global	3001	73	779	52
Módulo 1	1088	1144	1357	997
Módulo 2	211	52	186	121
Módulo 3	371	83	236	209

Fonte: Dados do próprio autor

Na tabela 13, são relacionados os valores do tempo de processamento (segundo) do algoritmo retropropagação convencional, do algoritmo com o controlador *fuzzy* para ajustar a taxa de aprendizagem durante o processo de treinamento, do algoritmo retropropagação com adaptações da inclinação e da translação por meio do gradiente descendente e a junção das duas metodologias.

Tabela 13 - Comparando os valores de tempo de processamento (s) do algoritmo RP convencional com o algoritmo com o controlador *fuzzy* e o algoritmo com o controlador *fuzzy* e as adaptações da inclinação e da translação da função de ativação.

Sistema Previsor	Convencional	Controlador <i>Fuzzy</i>	Adaptações	Controlador e as Adaptações
Global	53,94	3,07	27,72	2,77
Módulo 1	20,88	16,85	24,37	17,96
Módulo 2	5,83	2,43	5,48	3,77
Módulo 3	6,76	2,97	6,08	5,34

Fonte: Dados do próprio autor

Comparando o tempo de processamento do sistema predictor global com o treinamento via algoritmo RP convencional com cada uma das propostas podemos perceber uma melhora muito significativa, e se destaca o valor obtido com a junção do controlador e as adaptações da inclinação e da translação da função sigmoidal.

Já no sistema predictor relativo ao módulo 1 houve uma melhora significativa no tempo de processamento utilizando o controlador *fuzzy*, porém pouco significativa se comparada aos outros sistema predictores, entretanto com uma sintonia no ajuste dos parâmetros pode-se reduzir o tempo de processamento. No sistema predictor composto pelo módulo 2 também houve uma melhoria significativa no tempo de processamento, destacando no que tange o melhor resultado o controlador *fuzzy*. O sistema predictor composto pelo módulo 3 obteve uma melhoria muito significativa no tempo de processamento utilizando o controlador *fuzzy*.

6.4 COMPARANDO METODOLOGIAS

A previsão realizada no presente trabalho utilizou dados de um sistema elétrico que foi usado na tese de Altran (2010) no qual sua pesquisa realizou uma previsão (global e modular), porém com metodologias diferentes. No trabalho de Altran (2010) foram realizadas previsões utilizando o algoritmo RP convencional com seis funções de ativação diferentes, comparando qual das funções forneceria uma melhor previsão. Com o intuito de verificar a eficiência da metodologia proposta comparar-se-á com as previsões realizadas por Altran (2010).

As seis funções de ativação utilizadas na pesquisa de Altran (2010) foram: Gaussiana, multiquadrática recíproca, secante hiperbólica, multiquadrática recíproca inversa, secante hiperbólica recíproca e a sigmoidal.

a) Comparando resultados obtidos na presente pesquisa com os melhores resultados obtidos por Altran (2010).

A tabela 14 mostra os melhores resultados de MAPE e erro máximo e tempo de processamento (TP) obtido por Altran (2010) e os resultados obtidos no presente trabalho.

Tabela 14- Comparação entre metodologias de Altran (2010) e a presente pesquisa.

Sistema Previsor	MAPE		Erro Máximo	
	T	D	T	D
Global	1,43	1,93	4,51	9,52
Módulo 1	3,27	3,41	8,69	6,36
Módulo 2	4,35	2,49	9,10	7,50
Módulo 3	2,93	2,13	5,50	4,31

Fonte: Dados do próprio autor

sendo:

TP : Tempo de processamento;

T : tese de doutorado;

D : dissertação de mestrado.

Pode se observar que houve uma melhora muito significativas na previsão dos sistemas modulares 2 e 3, já a previsão do módulo 1 e o sistema previsor global não melhorou a previsão, porém sendo possível melhorias mediante a alterações do parâmetros. É importante mencionar que os melhores resultados obtidos na tese de Altran (2010) utilizavam outras funções de ativação e não a função sigmoide.

b) Comparando resultados obtidos na presente pesquisa com os resultados obtidos por Altran (2010) utilizando a função sigmoideal.

Em sua tese Altran (2010) utilizou a função sigmoideal que é a mesma utilizada nesta pesquisa, e com intuito de comparar a eficiência da utilização dessa função sem as adaptações e com as adaptações os resultados serão ilustrados na tabela 15.

Tabela 15- Comparando os resultados da tese de Altran (2010) utilizando a função sigmoide e o resultado que foi obtido nesse presente trabalho.

Sistema Previsor	MAPE		Erro Máximo	
	T	D	T	D
Global	1,91	1,93	5,86	9,52
Módulo 1	5,56	3,41	18,34	6,36
Módulo 2	5,49	2,49	12,29	7,50
Módulo 3	3,81	2,13	13,73	4,31

Fonte: Dados do próprio autor.

No sistema previsor global os valores de MAPE ficaram bem próximos não havendo uma melhora significativa, e obteve-se um erro máximo bem mais alto que o esperado, porém o tempo de processamento foi reduzido drasticamente em comparação a metodologia proposta por ALTRAN, entretanto com um ajuste sincronizado dos parâmetros pode-se obter resultados melhores.

No sistema previsor modular 1 houve uma redução muito expressiva com mais de 2% no valor do MAPE, o erro máximo obteve valor significativo de redução de quase 12% em sua precisão, mas o valor do tempo de processamento foi acima do esperado.

O sistema previsor módulo 2 obteve uma melhora em seu valor de MAPE de 3% e a redução de seu erro máximo foi de 4,79%.

Já o sistema previsor modular 3 obteve uma melhora expressiva com relação a seu valor de MAPE (redução de 1,68%) e uma diminuição de 9,42% no valor de erro máximo.

No geral tanto a redução do MAPE e do erro máximo foram muito expressivas na utilização da metodologia. Entretanto, o tempo de processamento não foi no geral satisfatório, porém, esse pode ser influenciado por outros fatores, tais como o processador de computador, a linguagem utilizada, o algoritmo desenvolvido.

6.5 COMENTÁRIOS

Neste capítulo foram apresentados os resultados numéricos e gráficos para a previsão de carga de curto prazo (24 horas). Foram considerados dois sistemas previsores, sendo o sistema previsor global e o sistema previsor local constituído de três módulos. O sistema previsor local realiza a previsão por barramento considerando a participação do sistema global. Foram realizados inúmeros testes e percebeu-se que por menor que seja a variação dos parâmetros havia uma mudança muito drástica nos resultados. A teoria apresentada proporcionou uma melhoria significativa da precisão entre o algoritmo RP convencional e o algoritmo com adaptações dos parâmetros da taxa de treinamento via controlador *fuzzy*, adaptação da taxa de inclinação e da translação usando o gradiente descendente. Obteve-se um erro percentual absoluto médio MAPE abaixo de 4%, que é um valor significativo se

comparado aos outros métodos estatísticos encontrados na literatura quando se trata de uma previsão multimodal. Além de obter os resultados de precisão, MAPE e Erro máximo, melhores do que o algoritmo RP convencional houve uma melhoria no tempo de processamento do algoritmo e no número de épocas, fazendo com que o grau de confiabilidade na metodologia seja significativo.

7 CONSIDERAÇÕES FINAIS

A capacidade de aprendizagem, generalização e flexibilidade que as redes neurais possuem fez com que diversas áreas da ciência se interessassem pelo seu estudo.

A rede neural possui vários algoritmos sendo que cada um resolve um problema específico. O mais utilizado para realizar previsões é o algoritmo retropropagação que possui uma capacidade de precisão muito expressiva com relação a outros métodos de previsões encontrados na literatura, mas como todo método o algoritmo possui suas desvantagens que são: a paralisia da rede, o mínimo local e a lentidão no processo de convergência.

Com o intuito de propor melhorias que diminuam essas desvantagens, pesquisadores buscam através de algumas alterações aprimorar a rede. Duas dessas alterações foi proposta neste trabalho, primeiro foi o ajuste da taxa de treinamento através do controlador *fuzzy*, cujo principal objetivo dessa proposta é diminuir o tempo de convergência da rede durante o processo de treinamento evitando, assim que os pesos assumam valores altos e faça com que a rede entre em completa paralisia.

A segunda proposta foi a adaptação dos parâmetros da inclinação e da translação da função sigmoideal através do método do gradiente descendente, na qual esses parâmetros são adaptados de modo similar ao ajuste dos pesos. O objetivo principal dessa proposta consiste em aumentar o espaço de busca de uma solução e reduzir a probabilidade de convergência para pontos de mínimos locais. Uma terceira proposta de melhoria no algoritmo foi a junção dessas duas metodologias.

No intuito de testar essas três metodologias realizou-se a previsão de carga global e multimodal (em vários pontos da rede elétrica: subestação, barramentos, etc.) usando uma rede neural *feedforward* com treinamento realizado via emprego do algoritmo retropropagação. Estes três recursos visam aumentar a capacidade de convergência associada à fase de treinamento da rede neural.

Os resultados obtidos foram satisfatórios no requisito precisão e rapidez. O sistema previsor global e o local (módulos 2 e 3) apresentaram uma melhora significativa com relação ao tempo de processamento e o número de iterações no processo de convergência, tanto no

procedimento da implementação do controlador *fuzzy* como nas adaptações das taxas de inclinação e da translação da função de ativação e na junção do controlador *fuzzy* com as adaptações das taxas de inclinação e de translação da função de ativação.

O módulo 1 apresentou uma melhora significativa no valor de MAPE, porém não houve uma diminuição significativa no número de iterações. Contudo, sendo passivo de aperfeiçoamento mediante uma melhor sintonia dos parâmetros envolvidos. Pode-se perceber que por mais preciso que o algoritmo RP convencional seja sempre se pode melhorar por meio de processos de adaptações.

Assim, sugerem-se metodologias que podem tornar o procedimento ainda mais eficaz, tais como: o estudo e implementação de recursos no algoritmo visando uma melhoria na representação e manipulação das informações do módulo do código ativo da rede neural, o ajuste na inclinação das funções de pertinência, a adaptação da inclinação e da translação via controlador *fuzzy*. Obteve-se um erro percentual absoluto médio *MAPE* abaixo de 4%, o que demonstra sua eficiência se comparado aos outros métodos estatísticos encontrados na literatura quando se trata de uma previsão multimodal.

7.1 SUGESTÕES PARA TRABALHOS FUTUROS

Algumas possibilidades de melhoria no algoritmo foram propostas no trabalho, propostas estas que se mostraram efetivas, mas ainda existem alguns recursos que podem ser explorados, tais como:

- 1 - a troca da função de ativação;
- 2 - a sincronização dos parâmetros através de uma regra;
- 3 - a inserção do fator de participação;
- 4 - ajuste nas regras do controlador *fuzzy*.

REFERÊNCIAS

- AL-KANDARI, A. M.; SOLIMAN, S. A.; and EL-HAWARY, M. E. *Fuzzy* short-term electric load forecasting. **Electrical Power & Energy Systems**, Guildford, v. 26, n. 2, p. 111-122, 2004.
- ALMEIDA, C.; FISHWICH, P. A.; TANG, Z. Time series forecasting using neural network vs. box-jenkins methodology. **Simulation Councils**, p. 303- 310, November 1991.
- ALTRAN, A. B. **Sistema inteligente para previsão de cargas elétricas multinodal em sistemas elétrico de potencia**. 2010. 87 f. Tese (Doutorado) - Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2010.
- ARABSHAHI, P.; CHOI, J. J.; MARKS II, R. J.; CAUDELL, T. P. *Fuzzy* parameter adaptation in optimization. **IEEE Computational Science & Engineering**, Melville, v. 3, p. 57-65.1996.
- AZIZ, S. A.; PARTHIBAN, J. *Fuzzy* logic introduction. [S.l.: s.n], 1996. Disponível em:< http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/sbaa/report.html>. Acesso em: 3 fev 2014.
- BEZDEK, J. C. *Fuzzy* models – what are they, and why?. **IEEE Transactions on Fuzzy Systems**, Florida, v. 1, p. 1-6, 1993.
- BOX, G. E.; JENKINS, G. M. **Times series analysis: forecasting and control**. San Francisco: Holden-Day, 1976. p.575.
- BUNN, D. W.; FARMER, E. D. **Comparative models for electrical load forecasting**. Belfast: Wiley, 1985.
- CARPENTER, G. A. and GROSSBERG, S. A self-organizing neural network for supervised learning, recognition and prediction. **IEEE Communications Magazine**, Piscataway , v. 30, n. 9, p. 38-49, 1992.
- DJUKANOVIC, M.; RUZIC, S.; BABIC, B.; SOBAJIC, D. J.; PAO, Y. H. A neural-net based short term load forecasting using moving window procedure. **Electrical Power & Energy Systems**, Guildford, v. 17, n. 6, p. 391-397, 1995.
- GROSSBERG, S. Adaptive pattern classification and universal recoding I: parallel development and coding of neural feature detectors. **Biological Cybernetics**, Boston, v. 23, n. 3, p. 121-134, 1976.
- GROSS, G.; GALIANA, F. D. Short term load forecasting. **Proceedings of the IEEE**, San Francisco, v. 75, n. 12, p. 1558-1573, 1987.

KARTALOPOULOS, S. V. **Understanding neural networks and Fuzzy logic**. New York: IEEE Press, 1996.

KOHONEN, T. **Self-organizing and associative memory**. 2nd ed. Berlin: Springer-Verlag, 1987.

KROSE, B.; SMAGT, P. V. **An introduction to neural networks**. 8. ed. Amsterdam: University of Amsterdam, 1996. Disponível em: <<http://www.infor.uva.es/~teodoro/neuro-intro.pdf>> . Acesso em: 11 jun 2014.

HAYKIN, S. **Neural networks: a comprehensive foundation**. Upper Saddle River: Prentice-Hall, 1994. p.768 .

HIPPERT, H. S.; PEDREIRA, C. E.; SOUZA, R. C. Neural networks for short term load forecasting: are view and evaluation. **IEEE Transactions on Power Systems**, Juiz de Fora , v. 16, n. 1, p. 44-55, 2001.

HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. **Journal of Physiology**, Cambridge , v.117, p.500-544, 1952.

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

LOPES, M. L. M.; MINUSSI, C. R. Treinamento de redes neurais via back-propagation com controlador nebuloso. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 13., 2000, Florianópolis. **Congresso...** Florianópolis: [s.n.], 2000. p. 1616-1621.

LOPES, M. L. M.; MINUSSI, C. R.; LOTUFO, A. P. A fast electric load forecasting using neural networks. In: MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS, 43., 2000, Lansing. **Symposium...** Lansing: [s.n.], 2000. p. 646-649.

LOPES, M. L. M. **Desenvolvimento de um sistema previsor de cargas elétricas via redes neurais**. 2000. Dissertação (Mestrado)- Faculdade de Engenharia, Universidade Estadual Paulista Júlio de Mesquita Filhos- UNESP, Ilha Solteira, 2000.

LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica**. 2005. Tese (Doutorado)- , Faculdade de Engenharia, Universidade Estadual Paulista Júlio de Mesquita Filhos- UNESP, Ilha Solteira 2005.

MCCULLOCH, W.; PITTS, W. A. A logical calculus of the ideas imminent in nervous activity. **Bull Math Biophys**, Chicago, v. 5, p. 115-133, 1943.

MENDES FILHO, E. F.; CARVALHO, A. C. P. L. Evolutionary design of MLP neural network architectures – In; SIMPÓSIO BRASILEIRO DE REDES NEURAIIS- SBRN, 4., 1997, Campos do Jordão. **Simpósio...** Campos do Jordão: [s.n.], 1997. p. 58-65.

METAXIOTIS, K.; KAGIANNAS, A.; ASKOUNIS, D.; PSARRAS, J. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. **Energy Conversion and Management**, Athens, v. 44, n. 9, p. 1525-1534, 2003.

MINUSSI, C. R.; LOTUFO, A. D. P. **Lógica nebulosa**. Ilha Solteira: Programa de Pós-Graduação em Engenharia Elétrica, 2011. Apostila.

MINUSSI, C. R.; SILVEIRA, M. C. G. Electric power systems transiente stability by neural networks. In: MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS,38., 1995, Rio de Janeiro. **Symposium...** Rio de Janeiro: [s.n.], 1995. v. 2, p. 1305-1308.

MOGHARAM, I.; RAHMAN, S. Analysis and evaluation of five short-term load forecasting techniques. **IEEE Transactions on Power Systems**, Blacksburg, v. 4, n. 4, p. 1484-1491.

MURTO, P. **Neural network models for short-term load forecasting**. 1998. Thesis (Masters) - Department of Engineering Physics and Mathematics, Helsinki University of Technology, Helsinki, 1998.

O'DONOVAN, T. M. **Short term forecasting: an introduction to the box-jenkins approach**. New York: John Wiley & Sons, , 1983.

PARK, D. C.; EL-SHARKAWI, M. A.; MARK II, R. J. Electric load forecasting an artificial neural network. **IEEE Transactions on Power Systems**, New York, v. 6, n. 2, p. 442-449, 1991.

ROSENBLATT, F. **Principles of neurodynamics**. New York: Spartan Books, 1962.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by Back-propagation errors. **Nature**, California , n. 323, p. 533-536, 1986.

SANDRI, S.; CORREA, C. **Lógica nebulosa**. São José dos Campos: INPE, 1999. p. c073-c090. Disponível em: < http://www.gta.ufrj.br/ensino/cpe717-2011/curso_ERN99_fuzzy.pdf >. Acesso em: 12 dez 2014.

SFETSOS, A. Short-term load forecasting with a hybrid clustering algorithm. **IEEE Proceedings Generation, Transmission, and Distribution**, Greece, v. 150, n. 3, p. 257-262, 2003.

SIMPSON, P. K. **Artificial neural systems: foundations, paradigms, applications, and implementations**. New York: Pergamon Press, 1989.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

SRINIVASAN, D.; TAN, S. S.; CHANG, C. S.; CHAN, E. K. Practical implementation of a hybrid *Fuzzy* neural network for one-day-ahead load forecasting. **IEE Proceedings Generation, Transmission and Distribution**, Singapore, v. 145, n. 6, p. 687 – 692, november 1998.

STAMATIS, N.; PARTHIMOS, D.; GRIFFITH, T. M. Forecasting chaotic cardiovascular times series with an adaptive slope multilayer perceptron neural network. **IEEE Transactions on Biomedical Engineering**, Cardiff, v. 46, n. 12, p. 1441-1453, 1999.

STERGIOU, C. Neural networks, the human brain and learning. **Surprise**, Seattle, v. 2, 1996. Disponivel em: <http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol2/cs11/article2.html>. Acesso em: 3 fev 2014

STORB, B. H.; WAZLAWICK, R. S. **Algumas pinceladas sobre redes neurais artificiais**. Santa Catarina: Departamento de Informática e Estatística – Universidade Federal de Santa Catarina - UFSC, 1999.

SWARUP, K. S.; SATISH, B. Integrated ANN approach to forecast load. **IEEE Computer Applications in Power**, Madras, v. 15, n. 1, p. 46-51, 2002.

VILLALBA, S. A.; BEL, C. A. Hybrid demand model for load estimation and short term load forecasting in distribution electric systems. **IEEE Transactions on Power Delivery**, Valencia, v. 15, n. 2, p. 764-769, 2000.

VON NEUMANN, J. **The computer and the brain**. New Haven: Yale University Press, 1958.

WASSERMAN, P. D. **Neural computing: theory and practice**. New York: Van Nostrand Reinhold, 1989.

WERBOS, P. J. **Beyond regression: new tools for prediction and analysis in the behavioral sciences**. 1974. Thesis (PhD.)- Harvard University, 1974.

WIDROW, B.; LEHR, M. A. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. **Proceedings of the IEEE**, v. 78, n. 9, p. 1415-1442, Sep. 1990.

YALCINOZ, T.; EMINOGLU, U. Short term and medium term power distribution load forecasting by neural networks. **Energy Conversion and Management**, London, p. 1-13, 2004.

ZADEH, L. Fuzzy sets. **Information and Control**, California, v. 8, p. 338-353, 1965.

APÊNDICE A - LÓGICA FUZZY

Neste apêndice serão abordados os conceitos de Lógica *Fuzzy* ou Lógica nebulosa, os quais foram empregados com o objetivo de desenvolver um mecanismo de adaptação da taxa de treinamento do algoritmo retropropagação para a previsão de cargas elétricas. Serão apresentadas a definição de conjuntos nebulosos e suas operações, as principais aplicações, as funções de pertinência, o conceito fundamental de lógica nebulosa, assim como o controle *fuzzy*. Além da aplicação em previsão de carga o estudo de lógica *fuzzy* tem várias outras áreas de interesses pela sua capacidade de trabalhar com valores incertos, com ruídos, com a não-linearidade de sistemas e, principalmente, com o reconhecimento de padrões que existem em diversas áreas da ciência. Têm-se alguns exemplos encontrados na literatura, tais como: (BEZDEK, 2005): **Química:** analítica, computacional, industrial, cromatografia, engenharia de alimentos e ciência cervejeira. **Engenharia Elétrica:** imagem e processamento de sinais, redes neurais, sistemas de controle, informática, automação, robótica, sensoriamento remoto e controle, engenharia óptica, visão computacional, computação paralela, *networking*, instrumentação e medição, dielétricos, reconhecimento de voz, circuitos de estado sólido e sistemas de energia elétrica. **Geologia / Geografia:** fotogrametria, pesquisa geofísica, geoquímica, biogeografia e a arqueologia. **Medicina:** ressonância magnética, diagnóstico médico, tomografia, neurologia, farmacologia, física médica, nutrição, ciências dietéticas, anestesia, ultramicroscopia, biomedicina, ciência da proteína, de neuroimagem e interação medicamentosa. **Física:** astronomia, óptica aplicada e física terra. **Ciências do Ambiente:** ciências do solo, floresta e poluição do ar, meteorologia e recursos hídricos.

Em 1965, o professor Zadeh (ZADEH, 1995), da Universidade da Califórnia, em Berkeley, desenvolveu o conceito de conjunto *fuzzy*, que nada mais é do que uma modelagem matemática para resolver problemas de dados imprecisos, vagos e incertos. O conjunto convencional $[0,1]$ é uma generalização do conjunto *fuzzy*. Logo, pode-se dizer que a lógica nebulosa admite valores intermediários de verdade. O grau de pertinência do conjunto nebuloso está compreendido entre 0 e 1. Lógica *fuzzy* é uma das tecnologias mais bem sucedidas no desenvolvimento de controle de processos sofisticados. As operações com os conjuntos nebulosos são muitos similares aos dos conjuntos convencionais, ou seja, são empregados os conceitos de operações como a união, a interseção, a negação, entre outras.

TEORIA DOS CONJUNTOS FUZZY

Um conjunto *fuzzy* A do universo de discurso Ω é definido por uma função de pertinência $\mu_A: \Omega \rightarrow [0, 1]$. Essa função associa cada elemento x de Ω o grau de $\mu_A(x)$ com o qual x pertence a A [x]. A função de pertinência $\mu_A(x)$ indica o grau de compatibilidade entre x e o conceito expresso por A : (SANDRI; CORREA, 1999).

$\mu_A(x) = 1$, indica que x é completamente compatível com A .

$\mu_A(x) = 0$, indica que x é completamente incompatível com A .

$0 < \mu_A(x) < 1$, indica que x é parcialmente compatível com A .

OPERAÇÕES EM CONJUNTOS FUZZY

Similarmente às operações dos conjuntos convencionais, nos conjuntos nebulosos existe a necessidade de proceder as operações de interseção, união, negação entre outras.

Sejam A, B, C e D conjuntos *fuzzy* definidos em Ω (ZADEH, 1995)

1-Interseção de dois conjuntos. A função de pertinência da interseção de dois conjuntos ($C = A \cap B$) é definida por:

$$\mu_C(x) = \min\{\mu_A(x), \mu_B(x)\}, x \in \Omega \quad (26)$$

2-União de dois conjuntos. A função de pertinência da união de dois conjuntos ($D = A \cup B$) é definida por:

$$\mu_D(x) = \max\{\mu_A(x), \mu_B(x)\}, x \in \Omega \quad (27)$$

3-Complemento de um conjunto nebuloso. A função de pertinência do complemento de um conjunto nebuloso A , A' é dado por:

$$\mu_{A'}(x) = 1 - \mu_A(x), x \in \Omega \quad (28)$$

FUNÇÕES DE PERTINÊNCIA

A magnitude da participação de cada entrada possui uma representação gráfica designada de função de pertinência, onde é associado um peso a cada entrada processada, definindo uma superposição funcional e determinando uma resposta de saída. Após as funções serem concluídas, pesadas e combinadas elas são denebulizadas em uma saída real. (LOPES, 2005). As funções de pertinências nebulosas podem ter várias formas, tais como: (1) trapezoidal, (2) gaussiana e (3) triangular. A função triangular foi a utilizada na presente pesquisa.

O grau de pertinência é definido como a projeção vertical do parâmetro de entrada do eixo horizontal no limite mais alto da função de pertinência, o qual tem valores definidos entre $[0,1]$ (BEZDEK, 1993).

A função de pertinência é representada exclusivamente para todo $x \in U$ pela fórmula:

$$A(x) = \sup \{ \mu_A(x) : \mu_A(x) \in [0,1], x \in A_\alpha \} \quad (29)$$

no qual “*sup*” denota supremo (ou máximo quando existir).

O emprego da função de pertinência traz benefícios ao controlador *fuzzy* pelo fato de que caso queira a otimização do controlador, basta modificar a inclinação, aumentando ou diminuindo da função de pertinência.

Outra forma de aperfeiçoar o controlador é por meio de dispositivos de modificação chamados de Hedges, ou seja, ao invés de modificar o conjunto de regras, o *hedge* modifica somente uma regra, enquanto que a modificação da inclinação da função de pertinência afeta todas as regras (MINUSSI; LOTUFO, 2011). Os hedges são genericamente da forma:

$$mA$$

sendo:

m : o agente modificador;

A : o conjunto nebuloso.

ESTRUTURA DO CONTROLADOR *FUZZY*

O controle *fuzzy*, ao contrário do controle convencional em que o algoritmo de controle é descrito por equações algébricas ou diferenciais, é realizado por meio de um modelo matemático utilizando-se de regras lógicas no algoritmo de controle com a intenção de descrever uma experiência humana para controlar o processo. Emprega-se o controlador *fuzzy* por causa de sua grande capacidade de adaptar e sua versatilidade e a possibilidade de trabalhar com não-linearidades, sendo capazes de superar perturbações e ruídos. A estrutura básica de um controlador é a interface de “fuzificação”, inferência e a “defuzificação”.

Interface de “Fuzificação”- Faz a identificação dos padrões de entrada transformando esses padrões em variáveis linguísticas.

Inferência - Manipula a base regras utilizando as declarações *If-Then*, de maneira a caracterizar a estratégia de controle e sua metas.

Defuzificação – Converte as variáveis linguísticas em variáveis reais, as quais se constituem na ação de controle. Os métodos de defuzificação mais utilizados são (SANDRI; CORREA, 1999, ARABSHAHI et. al, 1996):

- **Critério max:** encontra o valor de saída por meio do ponto em que o grau de pertinência da distribuição da ação de controle atinge o primeiro valor máximo.
- **Média do máximo:** encontra o ponto médio entre os valores que possui maior grau de pertinência inferido pelas regras.
- **Centro de área (centroide):** o valor da saída é o centro de gravidade da função de distribuição de possibilidades da ação de controle (SANDRI; CORREA, 1999).

O método de defuzificação utilizado foi o centro de área dado pela equação 30 (ARABSHAHI et. al, 1996).

Para conjuntos nebulosos discretos tem-se:

$$u = \frac{\sum_{i=1}^n \mu_i}{\sum_{i=1}^n \mu_i} \quad (30)$$

sendo:

i : valor do conjunto que tem um valor de pertinência μ_i ;

n : número de regras nebulosas.

Para conjuntos nebulosos contínuos o cálculo do centroide é a divisão do momento dos conjuntos nebulosos pela sua área:

Considere $f(x)$ um conjunto nebuloso, então (AZIZ; PARTHIBAN, 1996; KARTALOPOULOS, 1996):

$$u = \frac{\int x f(x) dx}{\int f(x) dx} \quad (31)$$

Os limites da integração correspondem ao universo inteiro de X .

COMENTÁRIOS

Neste Apêndice foi apresentada a teoria sobre os conjuntos *fuzzy* abordando suas operações, suas funções de pertinência e sua aplicabilidade em diversas áreas da ciência. Foi apresentada, também, a estrutura do controlador *fuzzy*, na qual foi abordada a interface de fuzificação, inferência e defuzificação (os quais são os principais conceitos para se obter um controle), bem como o funcionamento do processo de controle.

APÊNDICES B- DADOS HISTÓRICOS DE CARGA

Esse apêndice apresenta o histórico de carga elétricas fornecido por uma companhia do setor elétrico da Nova Zelândia (Electricity Commission da Nova Zelândia). Estes dados foram utilizados para treinar a rede neural proposta.

Foram considerados o sistema previsor local constituído por três módulos, sendo que cada módulo representa uma subestação, e o sistema previsor global que é a soma dos três módulos (três subestações).

No que tange os dados de carga considerou-se um intervalo de 15 dias, no período do dia 15 de outubro a 29 de outubro do ano de 2007, sendo a carga utilizada nas 24 horas de cada dia, totalizando assim uma entrada de 360 dados para cada um dos sistemas.

O apêndice B1 corresponde ao sistema global, B2 ao sistema local módulo 1, B3 sistema local módulo 2 e B4 ao sistema local módulo 3.

APÊNDICE B1 – Sistema Previsor Global

Dados de Potência Ativa em MW

Hora	Dias															
	15/10	16/10	17/10	18/10	19/10	20/10	21/10	22/10	23/10	24/10	25/10	26/10	27/10	28/10	29/10	30/10
01	60360	65244	64702	66986	69042	71376	70824	65020	59900	66864	64584	61982	65936	62554	60918	64748
02	57670	62166	60912	63794	66154	66668	65184	61234	57088	64542	62114	59182	60876	59296	57144	61464
03	56378	60580	59970	61100	65020	64224	63746	59756	56288	62324	60878	58494	59186	58336	56724	59592
04	56128	59432	58472	59788	63284	63438	61334	59120	56946	60948	58610	58242	59010	57946	56364	58464
05	57920	60802	59874	61176	61638	64866	61972	58868	58952	62278	59434	59030	59312	57958	57188	60300
06	68182	70868	69584	71482	69992	73344	69608	66312	69490	72698	69614	69662	66534	65282	66768	70778
07	88288	91250	91262	91782	92054	88362	84096	81556	90462	91750	90384	88764	82234	79020	88052	91248
08	105480	104978	105748	105188	107450	101420	96036	90614	105872	105142	105708	104236	91870	87480	104484	104822
09	112560	111078	109412	110518	114292	112488	107728	101554	112056	109718	109932	108716	96750	93712	110148	109114
10	107506	103638	104920	106782	110442	110238	106358	107004	107186	100422	103578	107034	95080	94126	101554	105014
11	98882	97036	100570	102006	102772	104854	105158	105812	98940	93608	97596	99042	90318	90144	96036	97812
12	95002	92644	94440	94814	95724	95656	96604	99256	94574	92244	94006	95314	85792	78182	92292	93940
13	91474	90100	91058	95206	91824	90496	90866	95916	91602	91946	90950	94612	82754	73100	88256	90600
14	88226	86200	87522	91986	87752	87528	84860	89892	88906	88484	87906	90430	77504	68886	87838	88368
15	87450	83194	86302	87664	86354	85418	80878	85488	86758	86072	85172	88058	75858	66420	85414	86056
16	91508	87906	90806	92774	90016	89380	83476	87458	89350	89532	88990	90668	80106	70596	88386	91224
17	104312	100566	100960	104728	100942	100110	94418	97122	99186	98474	98110	99144	90866	81498	99620	99378
18	109462	106872	106220	110756	105538	111234	100724	102306	106276	104928	102426	102592	96458	99574	103892	102942
19	107912	106114	105850	113356	104544	111140	102706	99416	106318	102392	100828	100390	93590	96672	104932	101194
20	106062	103456	102246	111182	102040	108042	102548	93814	99672	100640	97664	94674	91162	91944	97114	93542
21	102632	100066	98196	107230	107992	107196	104094	95828	98536	98686	97892	93714	90534	93244	96154	92840
22	92022	89372	90498	97388	104066	104160	95998	88082	89620	87894	86860	87742	83756	85842	89086	86328
23	81156	82238	81482	86506	93210	94458	84450	75662	81396	80140	76770	79122	76022	74844	80404	78650
24	70868	71882	73050	76142	80724	82296	73302	66374	72672	70606	67818	71472	68660	66052	71128	71192

APÊNDICE B2 – Sistema Previsor Local – Módulo 1

Dados de Potência Ativa em MW

Hora	Dias															
	15/10	16/10	17/10	18/10	19/10	20/10	21/10	22/10	23/10	24/10	25/10	26/10	27/10	28/10	29/10	30/10
1	14674	16800	17110	17820	18316	20144	20596	17922	15296	17564	17492	16274	16608	15642	15390	17074
2	13802	15474	15726	16454	16896	17712	17982	16358	14148	16466	16226	15570	14770	14400	14102	15844
3	13408	15354	15438	15814	16238	16472	17134	15558	13778	15796	15570	14400	13826	13832	13958	15008
4	13072	13898	13826	14602	14936	15992	15844	15296	13492	14484	14078	14578	13742	13158	13504	13874
5	13414	14172	14148	14518	15414	16046	16034	14960	13658	14818	14268	17290	13814	13600	13898	14340
6	16776	17026	17182	17314	17934	18088	17648	16502	16560	17660	17564	23344	15892	15534	16668	17014
7	23606	23952	23726	24084	23844	22756	22088	20788	23774	24012	24920	29936	20652	19440	23630	24012
8	30604	30450	30378	30820	31654	28752	27810	24742	30830	30676	31428	32970	24544	23082	29984	30784
9	34188	33936	32730	33244	35166	34988	35334	31572	34080	33040	33650	31822	27966	26760	33208	33292
10	31776	31428	31512	31752	33304	35990	36050	35882	32134	30850	31764	29316	28174	28468	30796	30700
11	29840	29496	29338	30128	30736	33364	36324	35918	29566	28980	30462	28456	25954	27250	28510	28610
12	29004	27428	27954	27882	28408	29722	32264	32360	27668	28312	29232	28910	24378	18544	27058	27214
13	27392	26114	26748	28790	27620	28408	29948	30736	27250	27596	26844	26902	23194	15486	25380	24932
14	26390	26096	26544	28838	26592	27810	27130	28384	26078	27226	26568	26234	21814	14328	26704	25482
15	26032	25374	26044	26330	25924	26856	25160	26592	25374	25876	25232	25542	20764	13684	25420	24014
16	26748	25732	26736	27356	26904	27572	25148	26640	25912	26318	25828	26306	21482	14424	25922	25254
17	29052	27644	28336	29458	29100	30354	27602	28694	27166	28324	27058	27680	24514	16442	27200	26820
18	32552	30462	31106	32862	30952	35084	30982	30080	30152	31452	29494	27732	26880	31142	29480	28826
19	33328	31346	31058	34582	31692	36934	33352	29876	31130	30498	29542	27596	27482	29386	29826	29398
20	31380	31728	31608	34092	32694	37400	34700	27966	30736	32014	29268	28420	27578	27692	29002	28608
21	31632	30628	30592	33996	36528	37340	36372	29292	30986	31034	30080	26228	27674	28444	29564	29338
22	27178	26520	26586	30186	34570	37064	32910	26402	27190	26808	26772	22640	24854	25578	27010	26556
23	23870	24442	24186	26342	30150	34842	27572	21684	24216	24060	24348	19046	21822	21516	24108	23642
24	19786	20024	20682	22042	24466	25636	22186	17660	20252	20228	20550	16274	18448	17516	20072	20108

APÊNDICE B3 – Sistema Previsor Local – Módulo 2

Dados de Potência Ativa em MW

Hora	Dias															
	15/10	16/10	17/10	18/10	19/10	20/10	21/10	22/10	23/10	24/10	25/10	26/10	27/10	28/10	29/10	30/10
1	30452	32502	31516	33166	33952	32614	29966	29904	28720	33196	31618	28174	32736	30406	29472	32076
2	29370	31440	30076	31962	33178	31638	29292	28590	28118	32880	30922	27636	30556	29538	27952	30940
3	28530	30414	29824	30492	33172	30884	29514	28332	27994	31618	30570	28168	30222	29594	27982	30334
4	28674	30658	29938	30592	32770	30998	28508	28218	29090	31632	29852	29200	30330	29832	28096	30284
5	29980	31650	30778	31654	33314	32340	28936	28454	30700	32446	30620	29716	30370	29402	28354	31366
6	35724	37596	36364	37958	38786	37654	34146	33324	37036	38858	36186	36402	34572	33596	34112	37908
7	45874	47996	48134	48554	48714	45450	42302	42168	47468	48794	46492	46506	43452	41784	45364	48714
8	51316	50998	51714	50932	51338	48010	45230	44998	50764	50914	51054	51054	46886	44948	50806	50794
9	50768	50910	50788	51132	50930	48950	45214	45150	50888	50832	50932	50244	45968	45150	51018	50674
10	49588	47224	50516	50264	50012	45888	42396	43344	48938	47694	47554	50696	43688	42518	45754	50494
11	44612	44390	48350	48536	46870	42796	40876	41486	44962	42988	44710	46786	42002	40404	43918	47266
12	42562	42792	44406	44040	43506	39648	38618	39980	43258	41164	42848	43852	39880	38352	42744	45430
13	40980	41342	42096	42720	41150	37714	36688	39722	41364	41372	41912	42906	38546	37132	41436	43932
14	39642	38904	39736	41020	39148	35928	34982	37836	40576	39284	40080	41802	36258	35174	40116	42354
15	39400	37614	39126	39980	38208	35526	34250	36472	39850	38862	39514	40528	36064	33906	39252	42362
16	42622	41070	42440	43616	40726	38840	36774	38862	41314	41536	41770	43376	38876	36932	42018	44922
17	50494	49312	49340	50838	46952	44980	43646	45460	48134	47180	48164	49534	45008	44046	49842	49196
18	50996	51050	50702	51234	48394	47350	44498	47804	50928	49416	49146	50298	46714	46198	50358	50258
19	51032	49418	50064	50968	46184	45588	42656	45108	48912	46980	46816	47776	43186	44298	50458	47446
20	49362	47632	46894	50078	42476	42042	40740	42190	44046	43696	44204	44204	40324	41550	45602	42620
21	46320	45294	43860	47044	43150	40970	39256	41894	42602	42312	43036	41522	39084	40856	43860	40562
22	42200	40762	41622	43258	41680	38094	36630	38682	40418	38796	37856	38660	36372	37978	40618	37806
23	37434	37878	37548	39132	38532	34630	33690	33906	37394	36372	32542	35932	34042	33582	37018	35432
24	33784	34386	34666	35940	35268	31732	31466	31252	34948	33030	29796	33980	32176	31180	33948	34048

APÊNDICE B4 – Sistema Previsor Local – Módulo 3

Dados de Potência Ativa em MW

Hora	Dias															
	15/10	16/10	17/10	18/10	19/10	20/10	21/10	22/10	23/10	24/10	25/10	26/10	27/10	28/10	29/10	30/10
1	30452	32502	31516	33166	33952	32614	29966	29904	28720	33196	31618	28174	32736	30406	29472	32076
2	29370	31440	30076	31962	33178	31638	29292	28590	28118	32880	30922	27636	30556	29538	27952	30940
3	28530	30414	29824	30492	33172	30884	29514	28332	27994	31618	30570	28168	30222	29594	27982	30334
4	28674	30658	29938	30592	32770	30998	28508	28218	29090	31632	29852	29200	30330	29832	28096	30284
5	29980	31650	30778	31654	33314	32340	28936	28454	30700	32446	30620	29716	30370	29402	28354	31366
6	35724	37596	36364	37958	38786	37654	34146	33324	37036	38858	36186	36402	34572	33596	34112	37908
7	45874	47996	48134	48554	48714	45450	42302	42168	47468	48794	46492	46506	43452	41784	45364	48714
8	51316	50998	51714	50932	51338	48010	45230	44998	50764	50914	51054	51054	46886	44948	50806	50794
9	50768	50910	50788	51132	50930	48950	45214	45150	50888	50832	50932	50244	45968	45150	51018	50674
10	49588	47224	50516	50264	50012	45888	42396	43344	48938	47694	47554	50696	43688	42518	45754	50494
11	44612	44390	48350	48536	46870	42796	40876	41486	44962	42988	44710	46786	42002	40404	43918	47266
12	42562	42792	44406	44040	43506	39648	38618	39980	43258	41164	42848	43852	39880	38352	42744	45430
13	40980	41342	42096	42720	41150	37714	36688	39722	41364	41372	41912	42906	38546	37132	41436	43932
14	39642	38904	39736	41020	39148	35928	34982	37836	40576	39284	40080	41802	36258	35174	40116	42354
15	39400	37614	39126	39980	38208	35526	34250	36472	39850	38862	39514	40528	36064	33906	39252	42362
16	42622	41070	42440	43616	40726	38840	36774	38862	41314	41536	41770	43376	38876	36932	42018	44922
17	50494	49312	49340	50838	46952	44980	43646	45460	48134	47180	48164	49534	45008	44046	49842	49196
18	50996	51050	50702	51234	48394	47350	44498	47804	50928	49416	49146	50298	46714	46198	50358	50258
19	51032	49418	50064	50968	46184	45588	42656	45108	48912	46980	46816	47776	43186	44298	50458	47446
20	49362	47632	46894	50078	42476	42042	40740	42190	44046	43696	44204	44204	40324	41550	45602	42620
21	46320	45294	43860	47044	43150	40970	39256	41894	42602	42312	43036	41522	39084	40856	43860	40562
22	42200	40762	41622	43258	41680	38094	36630	38682	40418	38796	37856	38660	36372	37978	40618	37806
23	37434	37878	37548	39132	38532	34630	33690	33906	37394	36372	32542	35932	34042	33582	37018	35432
24	33784	34386	34666	35940	35268	31732	31466	31252	34948	33030	29796	33980	32176	31180	33948	34048

APÊNDICE C - TRABALHO SUBMETIDO

AMORIM, A. J; LOPES, M. L; MINUSSI, C. R. Desenvolvimento de uma rede neuro-*fuzzy* para a previsão da carga. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 12, 2015, Natal. **Simpósio...** Natal: [s.n], 2015. P. 25-28.