

CÍCERO MARCELO DE OLIVEIRA

**PREVISÃO DE CARGAS ELÉTRICAS ATRAVÉS DE UMA
REDE NEURAL HÍBRIDA BACK-ART FUZZY**

Ilha Solteira – SP

2012

CÍCERO MARCELO DE OLIVEIRA

Previsão de Cargas Elétricas Através de uma Rede Neural Híbrida Back-ART Fuzzy

Orientadora: Prof. Dra. Anna Diva Plasencia Lotufo

Co-orientadora: Dra. Maria do Carmo Gomes da Silveira

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

Ilha Solteira – SP

2012

FICHA CATALOGRÁFICA

Elaborada pela Seção Técnica de Aquisição e Tratamento da Informação
Serviço Técnico de Biblioteca e Documentação da UNESP - Ilha Solteira.

- | | |
|------|---|
| O48e | <p>Oliveira, Cícero Marcelo de.
Previsão de cargas elétricas através de uma Rede Neural Híbrida Back-ART Fuzzy/
Cícero Marcelo de Oliveira. – Ilha Solteira : [s.n.], 2012
74 f. : il.</p> <p>Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2012</p> <p>Orientadora: Anna Diva Plasencia Lotufo
Co-orientadora: Maria do Carmo Gomes da Silveira
Inclui bibliografia</p> <p>1. Previsão de cargas elétricas. 2. Redes neurais artificiais. 3. Art FUZZY.
4. Perceptron multicamadas. 5. Híbridismo.</p> |
|------|---|



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

CERTIFICADO DE APROVAÇÃO

TÍTULO: PREVISÃO DE CARGAS ELÉTRICAS ATRAVÉS DE UMA REDE NEURAL HÍBRIDA
BACK-ART FUZZY

AUTOR: CICERO MARCELO DE OLIVEIRA

ORIENTADORA: Profa. Dra. ANNA DIVA PLASENCIA LOTUFO

CO-ORIENTADORA: Profa. Dra. MARIA DO CARMO G DA SILVEIRA

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


Profa. Dra. ANNA DIVA PLASENCIA LOTUFO

Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira


Profa. Dra. LILIAN YULI ISODA

Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira


Prof. Dr. GELSON DA CRUZ JUNIOR

Escola de Engenharia Elétrica e de Computação / Universidade Federal de Goiás

Data da realização: 19 de outubro de 2012.

Dedicatória

Agradecimentos

Agradeço a Deus pela oportunidade, força e capacidade para cumprir as etapas deste percurso. À minha mãe, Lourdes, pelas preces nos momentos de dificuldade. À minha irmã, Marli, meu cunhado, João, minha sobrinha, Carolina e meu padraсто, Galdino, pelo apoio dispensado.

Aos colegas de laboratório Marleide, Thays, Junior, Ana Cláudia e Juliana, pelos conselhos e torcida pela obtenção deste título.

Aos meus amigos Marco, Mariana, Virgilian, Virlaine, Rodrigo, Fausto, Anna Helena, Mariane, Helvia, Rogério, Aline, Bruno e a todos os demais, que tiveram sabedoria e paciência para passar comigo as fases difíceis.

A minha orientadora, Prof^a. Anna Diva, por ter acreditado em minha capacidade de desenvolver o trabalho proposto.

A minha co-orientadora, Maria do Carmo Gomes da Silveira, por ter aceitado estar nessa condição quando da ausência da Prof^a Anna Diva.

A Prof^a. Mara Lúcia Martins Lopes, do Departamento de Matemática, pelo empenho e dedicação dispensados no decorrer do projeto, o que a fez se tornar muito mais do que uma colaboradora, mas uma amiga para o resto de minha vida!

Resumo

Nos dias atuais, a previsão de cargas elétricas tem se mostrado cada vez mais importante às empresas distribuidoras de tal serviço, especialmente para que sejam possibilitados o planejamento, análise e operação dos sistemas elétricos, restando clara a necessidade de se antever o comportamento da carga, tornando possível o fornecimento eficiente de energia aos consumidores, visando que isso ocorra de forma econômica e contínua, valendo ressaltar ainda que, a tais empresas resta o interesse na lucratividade do setor para que se mantenham sólidas no mercado. Em um primeiro momento, a solução para tais problemas foi a utilização de técnicas matemáticas e estatísticas, podendo citar como exemplo, as séries numéricas, com resultados satisfatórios, mas de difícil modelagem. A inteligência artificial tem se mostrado uma técnica que supera os resultados anteriormente obtidos e, como prova de tal afirmação, a presente dissertação apresenta uma metodologia baseada em redes neurais, possibilitando a obtenção de resultados bastante satisfatórios, demonstrando ser um modelo robusto, com baixo custo computacional, rápido e eficiente. O objetivo deste trabalho é a utilização do hibridismo de redes neurais, sendo a primeira delas, uma rede ART Fuzzy e a segunda, a Perceptron multicamadas, via algoritmo backpropagation, aproveitando as melhores características de cada uma delas para a obtenção de resultado viável quando de sua utilização. A metodologia utilizada apresenta níveis de erro aceitáveis comparado a outros métodos que se encontram na literatura ou, até mesmo, em uso pelas empresas distribuidoras do setor elétrico brasileiro. No intuito de se obter a previsão de cargas citada acima, foi utilizado um banco de dados histórico de uma empresa distribuidora nacional, valendo-se de técnicas como o janelamento, entre outras que serão devidamente descritas no decorrer do texto.

Palavras-chave: Previsão de cargas elétricas. Redes neurais artificiais. ART Fuzzy. Perceptron multicamadas. Hibridismo.

Abstract

Electrical load forecasting is very important to the electrical companies to advice in planning, operation and analysis of electrical power systems. Knowing the load in advance leads to provide energy efficiently, continuously and economically. Firstly the solution for such problems was executed with mathematical and statistical techniques, e.g. time series analysis with good results but with difficult modeling. The use of Artificial Intelligence overcomes these problems and the proposal of this work is to present a robust hybrid method using neural networks providing good results with low computational costs.

The hybrid model proposed in this work uses a Fuzzy ART neural network and a MLP (multi layerPerceptron) by backpropagation training emphasizing the best characteristics of each one. The errors obtained are compatible with those found in the literature. Results are shown for data from a Brazilian electrical company.

Keywords: Electrical load forecasting.; Artificial neural networks. ART fuzzy.Hybridmethods.Backpropagation.MultilayerPerceptron.

Lista de Figuras

Figura 1 -Filtro linear, com Entrada a_t , saída Z_t , e Função de Transferência $\Psi(B)$	19
Figura 2 -Modelo de Série Temporal Estocástica	22
Figura 3 - Rede Neural Perceptron Multicamadas	29
Figura 4 - Estrutura das Redes Feedforward	30
Figura 5 - Modelo de um Neurônio Artificial	30
Figura 6 - Adaptação dos Pesos do Algoritmo Retropropagação	34
Figura 7 - Rede Neural ART Fuzzy	35
Figura 8 - Algoritmo da Rede Neural ART Fuzzy	36
Figura 9 - Fluxograma da Rede Neural ART Fuzzy	39
Figura 10 - Rede Neural Back-ART Fuzzy	41
Figura 11 - ART Fuzzy- Processo de Classificação de Dados Analógicos	42
Figura 12 - Resultado da Rede Híbrida Back-ART Fuzzy(744 vetores padrão)	51
Figura 13 - Resultado da Rede Híbrida Back-ART Fuzzy(1.464 vetores padrão)	52
Figura 14 - Resultado da Rede Híbrida Back-ART Fuzzy(2.208 vetores padrão)	53
Figura 15 - Neurônio Biológico	61
Figura 16 - Modelo do Neurônio de McCulloch-Pitts	62
Figura 17 - Função Relé	63
Figura 18 - Função Lógica Threshold	63
Figura 19 - Função Sigmóide (1)	64
Figura 20 - Função Sigmóide (2)	64
Figura 21 - Rede Neural de Hopfield	65
Figura 22 - Rede Neural Não-recorrente	65
Figura 23 - Treinamento Supervisionado	66
Figura 24 - Treinamento Não-supervisionado	66
Figura 25 - Função de Pertinência do Tipo Linear por Partes	70
Figura 26 - Função de Pertinência do Tipo Gaussiana	70
Figura 27 —VariáveisLingüísticas de Temperatura	71

Lista de Tabelas

Tabela 1 - Cargas Elétricas Reais	46
Tabela 2 - Períodos de Testes	50
Tabela 3 - Parâmetros da Rede Neural Back-ART Fuzzy - Aplicação 1	50
Tabela 4 - MAPE, Erro Máximo e Tempo de Processamento - Aplicação 1	51
Tabela 5 - Parâmetros da Rede Neural Back-ART Fuzzy - Aplicação 2	51
Tabela 6 - MAPE, Erro Máximo e Tempo de Processamento - Aplicação 2	52
Tabela 7 - Parâmetros da Rede Neural Back-ART Fuzzy - Aplicação 3	53
Tabela 8 - MAPE, Erro Máximo e Tempo de Processamento - Aplicação 3	54
Tabela 9 - Comparação de MAPE e Erro Máximo	54

Lista de Abreviaturas e Siglas

ANEEL	Agência Nacional de Energia Elétrica
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto RegressiveMovingAverage
ART	AdaptiveResonanceTheory
BACK	BackPropagation
CMO	Custo Marginal de Operações
IA	Inteligência Artificial
MA	MovingAverage
MAPE	Erro Percentual Médio Absoluto
MIT	Instituto de Tecnologia de Massachusetts
MLP	MultilayerPerceptron
RNA	Rede Neural Artificial

Lista de Símbolos

$yc(t)$	Carga no Instante t
a_0, a_1, \dots, a_n	Coefficientes de Regressão
$a(t)$	Sinal de Ruído Branco com Variável Aleatória de Média Zero e Variância Constante
$xe_1(t), \dots, xe_n(t)$	Variáveis Exógenas
ϕ_p	Constantes da Função Auto-Regressiva
p	Número de Instantes Anteriores Usados na Regressão
q	Ordem do Processo
b	Componente Constante
ε	Componente de Erro
s	Valores Previstos das Observações
xo	Valores Observados
A	Constante de Alisamento
t	Instante em que a Série Temporal é Calculada
ψ_0	Parâmetro que Determina o Nível da Série
Z_t	Processo Linear (discreto)
$\pi(B)$	Operador
ε_i	$d_i - y_i$
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
$\theta_i(h)$	$\gamma(\nabla_i(h))$
γ	Taxa de Treinamento (Parâmetro de Controle de Estabilidade)
h	Índice de Iteração

$\nabla_i(h)$	Gradiente do Erro Quadrático Relacionado aos Pesos do Neurônio i
V_i	$[w_{0i}w_{1i}w_{2i} \dots w_{ni}]^T$
σ_i	Δ Derivada da Função de Ativação Referente a Saída s_i
Q_i	Índices da Camada Seguinte, do i -ésimo Elemento, Interligados ao i -ésimo elemento
ρ	Parâmetro de Vigilância
α	Parâmetro de Escolha
β	Parâmetro de Treinamento
\wedge	Operador Fuzzy AND
$L(h)$	Valor da Carga Real Referente a Hora h
$\underline{L}(h)$	Valor da Carga Estimada pelo Modelo Híbrido Referente a Hora h
N	Número Total de Horas
λ	Inclinação da Curva
U	Universo de Discurso
CAT^{bin}	Vetor Binarizado da Categoria Obtida pelo Módulo ART Fuzzy
t^{bin}	Vetor Contendo Dados de Tempo (Dia da Semana, Hora e Feriado) Binarizado
$L(h+1)$	Carga Elétrica Referente a Hora $h+1$
X_{MLP}	Padrão de Entrada para a Rede Neural Perceptron Multicamadas
Y_{MLP}	Padrão de Saída para a Rede Neural Perceptron Multicamadas

Sumário

1 INTRODUÇÃO	14
1.1 REDES NEURAIS ARTIFICIAIS APLICADAS À PREVISÃO DE CARGAS ELÉTRICAS	14
1.2 OBJETIVO	15
1.3 DESCRIÇÃO DO TRABALHO	15
2 ESTADO DA ARTE	16
2.1INTRODUÇÃO	16
2.2 MOTIVAÇÃO PARA O ESTUDO DE CARGAS ELÉTRICAS	16
2.3 VIABILIDADE DA PREVISÃO DE CARGAS ELÉTRICAS FUTURAS	17
2.3.1 Séries Temporais	18
2.3.2 Séries Temporais Estacionárias	18
2.3.3 Séries Temporais Não-Estacionárias	20
2.4 METODOLOGIAS DE PREVISÃO DE CARGAS ELÉTRICAS	21
2.4.1 Métodos Estatísticos de Previsão	21
2.4.1.1 Regressão Linear Múltipla	22
2.4.1.2 Alisamento Exponencial	24
2.4.2 Métodos de Inteligência Artificial	25
2.4.2.1 Redes Neurais Artificiais	26
2.4.2.2 Lógica Fuzzy	26
2.4.2.3 Algoritmos Genéticos	27
2.5 MÉTODOS HÍBRIDOS	27
3 REDES NEURAIS	29
3.1 REDE NEURAL PERCEPTRON MULTICAMADAS, VIA ALGORITMO BACKPROPAGATION	29
3.2 REDE NEURAL ART FUZZY	34
4 REDE NEURAL HÍBRIDA	40
4.1 INTRODUÇÃO	40

4.2 REDE NEURAL BACK-ART FUZZY	40
4.3 COMENTÁRIOS	47
5 APLICAÇÃO, TESTES E RESULTADOS	49
6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	55
6.1 CONCLUSÕES	55
6.2 SUGESTÕES PARA TRABALHOS FUTUROS	56
REFERÊNCIAS	57
APÊNDICE A	61
APÊNDICE B - LÓGICA FUZZY	67

1INTRODUÇÃO

1.1 REDES NEURAIS ARTIFICIAIS APLICADAS À PREVISÃO DE CARGAS ELÉTRICAS

A previsão de demanda de energia elétrica tem se mostrado cada vez mais importante às empresas distribuidoras, especialmente pelo fato de que auxilia no planejamento relativo à própria demanda, análise, controle, necessidade de eventual expansão, fluxo de potência, etc. Para se atingir referida previsão de cargas, diversos métodos tem sido utilizados, merecendo enfoque a utilização da inteligência computacional, através do uso de Redes Neurais Artificiais, levando-se em conta períodos anteriores para obtenção das futuras cargas elétricas.

Redes Neurais Artificiais (RNAs) têm se mostrado um dos métodos mais eficientes para a obtenção da previsão precisa das cargas elétricas (HAYKIN, 1999), valendo ressaltar que, quando utilizado o hibridismo, através do uso de mais um tipo de rede neural, torna-se ainda melhor o resultado quanto à previsão anteriormente citada (ABRAHAM; NETH, 2001).

Diversos métodos relacionados à previsão de cargas elétrica se encontram na literatura especializada e, no presente caso, a proposta se dá pela utilização de uma Rede Neural ART Fuzzy, aliada a uma segunda rede, Perceptron Multicamadas, via algoritmo backpropagation, possibilitando a obtenção da previsão de cargas futuras de maneira bastante satisfatória.

A utilização de tais redes neurais se justifica, uma vez que, considerando cargas reais de período anterior, é possível realizar a classificação destas pela Rede Neural ART Fuzzy. Obtida a referida classificação, tal dado é aliado ao dia e hora respectivos, bem como ao fato de ser ou não feriado, dados estes utilizados como entrada da rede neural Perceptron Multicamadas (MultiLayerPerceptron - MLP), a qual prevê a primeira hora posterior ao período apresentado. Obtida a hora posterior, é possível retornar à RNA ART Fuzzy, a fim de obter nova classificação, passando as redes a trabalhar de forma ininterrupta e simultânea, com a primeira realizando a classificação de dados de entrada e a segunda, realizando a previsão da hora seguinte.

1.2 OBJETIVO

O presente trabalho tem como objetivo estudar a aplicação de Redes Neurais Artificiais na previsão de cargas elétricas, por meio de estudos relacionados às próprias RNAs, análise de dados relativos às cargas elétricas de períodos anteriores, viabilidade da utilização do hibridismo das RNAs, visando obtenção de resultados precisos e eficazes quanto à previsão, demonstrando que o modelo proposto é capaz de alcançar robustez, baixo custo computacional na intenção de atingir resultados satisfatórios e erros aceitáveis em relação a outros modelos já vistos no mercado.

1.3 DESCRIÇÃO DO TRABALHO

O presente trabalho se encontra disposto da seguinte forma: o capítulo 1, representado pela presente introdução; o capítulo 2 traz os conceitos relacionados a Cargas Elétricas; o capítulo 3 diz respeito às RNAs, especialmente as utilizadas neste estudo (ARTFuzzy e MLP, via algoritmo backpropagation); o capítulo 4 trata da RNA híbrida ora proposta, demonstrando a metodologia abordada; o capítulo 5 apresenta a aplicação, testes e resultados do presente modelo e o capítulo 6, traz a conclusão do estudo, bem como sugestões para trabalhos futuros.

2 ESTADO DA ARTE

2.1 INTRODUÇÃO

Neste capítulo é apresentado o estado da arte no que tange aos problemas de previsão de cargas elétricas, conforme disposto na literatura especializada. Para melhor entendimento do tema, são trazidas ao presente trabalho, noções relativas ao assunto em tela, visando o conhecimento de previsão de cargas elétricas pelos métodos anteriormente estudados em trabalhos já descritos, por meio de artigos e/ou na bibliografia disponível.

2.2 MOTIVAÇÃO PARA O ESTUDO DE CARGAS ELÉTRICAS

Segundo Schmidt (2004), o setor elétrico é de fundamental importância ao desenvolvimento econômico do país e a privatização deste se demonstrou uma das grandes precursoras no processo de reforma do Estado. Assim, resta clara a necessidade de expansão do setor por meio de investimentos na área de pesquisa e desenvolvimento, a fim de evitar crises futuras em um setor tão primordial em todos os segmentos, seja público ou privado.

Há muitos anos busca-se atingir a eficiência no setor de energia elétrica, valendo-se o governo de uma política competitiva entre empresas fornecedoras do mercado. Porém, a grande dificuldade reside na capacidade de investimento dos países em desenvolvimento, bem como em relação ao fornecimento de energia a curto prazo (VIEIRA FILHO et al., 2000).

Urge a necessidade de que se obtenha qualidade e eficiência no atendimento, motivo pelo qual foram desvinculadas da comercialização de energia elétrica, a transmissão, operação e administração da mesma, dando margem ao surgimento da “commodity” energia elétrica, a qual se diferencia das demais, em vista de impossibilidade de armazenamento em grande quantidade do produto (QUEIROZ, 2007).

Os preços da energia elétrica estão intimamente ligados ao investimento de geração da mesma, o que pode representar risco iminente em relação ao retorno esperado, bem como à carteira de contratos (QUEIROZ, 2007).

Devido a tais fatos, torna-se importante o estudo de métodos de previsão de cargas, uma vez que esta se torna importante nas áreas anteriormente citadas. E tais métodos podem ser dispostos em categoria, quais sejam: métodos de simulação, previsão de séries temporais, análise de equilíbrio, modelos econométricos e sistemas inteligentes (LI et al., 2005).

A previsão de cargas elétricas em relação ao intervalo de tempo é descrita da seguinte forma (AL-KANDARI et al., 2004; YALCINOZ; EMINOGLU, 2005):

- Curtíssimo prazo: toma por base intervalos de menos de uma hora, com instantaneidade do perfil da carga, servindo de base para decisões relacionadas ao gerenciamento de demanda de carga, maior disponibilidade de potência, servindo como parâmetro de realocação de manutenções de unidades geradoras, otimização do uso de potência decorrente de cargas diversas, etc.;
- Curto prazo: se fundamenta em intervalos semanais, no intuito de obter uma maior vantagem econômica dos geradores, tornando menor o curso de geração de energia, possibilitando, ainda, a otimização quanto à utilização de geradores térmicos;
- Médio prazo: neste caso, o intervalo considerado é mensal, com o objetivo de melhor aproveitamento de máquinas geradoras ou, até mesmo, avaliação da instabilidade da rede de energia, para fins de segurança quanto ao fornecimento;
- Longo prazo: o intervalo, aqui, é anual, sendo realizado para avaliação relativa à capacidade de fornecimento ao mercado consumidor. No Brasil, tendo em vista o setor ser basicamente hidrelétrico, o que possibilita transferência de carga entre empresas, são utilizados intervalos temporais de 10, 15 e até 20 anos para que ocorra a expansão do sistema, considerando-se o período compreendido entre a decisão e o real aumento da capacidade de fornecimento.

2.3 VIABILIDADE DA PREVISÃO DE CARGAS ELÉTRICAS FUTURAS

A previsão de cargas elétricas futuras se demonstra viável, tendo em vista a competitividade de um mercado em franca e contínua expansão e, visando principalmente, minimizar a perda às empresas fornecedoras de energia (HIPPERT et al., 2001).

Desta forma, torna-se necessária a utilização de um sistema com capacidade de alcançar o mínimo possível de erro, através de técnica precisa, possibilitando a maximização dos lucros e diminuindo a possibilidade de perdas, através de estratégias desenvolvidas pelo próprio sistema (LOPES, 2005).

2.3.1 Séries Temporais

A definição de séries temporais está ligada a observações de uma determinada variável em relação a períodos subsequentes de tempo, visando analisar e modelar a dependência da referida vizinhança, o que se torna a característica mais importante desse tipo de série, possível de ser utilizado em diversos segmentos do mercado, inclusive no segmento de energia elétrica. Segundo Moretin (1981), a análise de séries temporais tem como objetivos:

- 1) Descrição: propriedades da série;
- 2) Explicação: descreve o comportamento da série no período observado;
- 3) Controle de Processos: como em caso de controle estatístico de qualidade; e
- 4) Previsão: obtenção de valores futuros tem por base valores anteriormente coletados.

No Brasil, com a Lei de Concessões de 1995, ocorreu a reestruturação do setor elétrico, quando foi criada a ANEEL (Agência Nacional de Energia Elétrica), com a finalidade de regular a tarifação dos serviços prestados, bem como as condições de contratação.

Havia a necessidade de prever a quantidade de carga necessária para um determinado período de tempo, bem como o valor a ser estipulado em relação a seu fornecimento, motivo pelo qual passaram a ser usados modelos matemáticos capazes de atingir o custo marginal de operações (CMO). Para tanto, foram desenvolvidos diversos modelos para que se chegasse à previsão das séries temporais, tendo merecido atenção em um primeiro momento, os modelos estatísticos lineares para tratamento de dados estacionários e Gaussianos (BOX; JENKINS; REINSEL, 1994).

2.3.2 Séries Temporais Estacionárias

Os modelos estacionários são viáveis para a realização da previsão de séries também estacionárias, ou seja, que ocorrem em um determinado intervalo de tempo, em torno de uma média constante. Sendo assim, torna-se necessária a geração da série temporal por meio de filtro linear, demonstrado na Figura 2 ou, até mesmo, de um sistema linear, tomando como entradas o ruído branco (SILVA, 2012), o qual será explicado a seguir.

Figura 1 - Filtro Linear, com Entrada a_t , Saída Z_t , e Função de Transferência $\psi(B)$



Fonte: Silva, (2012)

Segundo Moretin (1985), Z_t é uma série de saída, ou seja, o valor no instante “ t ” é uma soma dos choques aleatórios da entrada do sistema, advinda desde o passado remoto até o instante “ t ”, sendo a primeira diferença o ruído branco.

Moretin (1985) dispõe que, formalmente:

$$Z_t = \mu + a_t + \varphi_1 a_{t-1} + \varphi_2 a_{t-2} + \dots = \mu - \varphi(B)a_t, \quad 13$$

sendo que:

$$\varphi(B) = 1 + \varphi_1 B + \varphi_2 B^2 + \dots \quad 14$$

definida como função de transferência do filtro, sendo ψ_0 parâmetro que determina o nível da série.

Deve-se considerar ainda que Z_t é um processo linear (discreto).

Chamando $\tilde{Z}_t = Z_t - \mu$, temos:

$$\tilde{Z}_t = \psi(B)a_t. \quad 15$$

Considerando a soma ponderada de valores passados $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$ somado ao ruído a_t , obtém-se, de forma alternativa, \tilde{Z}_t , da seguinte forma:

$$\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t. \quad 16$$

Vale ainda que:

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \tilde{Z}_t = a_t \quad 17$$

ou:

$$\pi(B)\tilde{Z}_t = a_t \quad 18$$

sendo $\pi(B)$ o operador.

Temos ainda:

$$\pi(B)\varphi(B)a_t = a_t, \quad 19$$

em que:

$$\pi(B) = \varphi^{-1}(B). \quad 20$$

Por meio de tal relação, torna-se possível obter pesos π_j em função dos pesos φ_j , bem como o oposto disso.

2.3.3 Séries Temporais Não-Estacionárias

Modelos não-estacionários são desenvolvidos para séries que se desenvolvem no tempo, considerando médias inconstantes, o que acontece com maior frequência na prática. Tal fato ocorre, por exemplo, nas séries econômicas que, quando aplicadas, as diferenças, tornam-se estacionárias (MORETIN, 1985).

A título de exemplo, vale salientar que Z_t é não estacionária, porém

$$W_t = Z_t - Z_{t-1} = (1 - B)Z_t = \Delta Z_t \quad 21$$

é estacionária.

Dentre os métodos descritos na literatura, para a obtenção da estacionariedade, pode ser citado, como exemplo, o ARIMA (p,d,q) . Muitos métodos são utilizados para aplicação de processos não-lineares, mas muitos não o são (AMARAL, 2003).

Uma série temporal é definida por y_t , onde $t = 1, 2, \dots, n$, com condições iniciais de $y_0, y_{-1}, \dots, y_{1-p}$ disponível sempre que necessário. As informações em $t - 1$ são representadas por Ω_{-1} , com informações suficientes para se obter previsão de séries futuras, descrita por $y_t, y_{t+1}, y_{t+2}, \dots$. No caso de Ω_{-1} não conter nenhuma informação para a previsão de um modelo linear para y_t , a série temporal correspondente passa a ser chamada de ruído branco,

descrita como ε_t . É necessário, ainda, que haja constância da média da série, bem como da variância, sendo que a primeira deve ser igual a zero. Ou seja, a série temporal, que também é um ruído branco, é definida por Amaral (2003):

$$E[\varepsilon_t] = 0 \quad 22$$

$$E[\varepsilon_t^2] = \sigma^2 \quad 23$$

$$E[\varepsilon_t \varepsilon_s] = 0 \quad \forall s \neq t \quad 24$$

No caso em tela, verifica-se que as covariâncias de ε_t são iguais a zero, o que corresponde a dizer que as informações de Ω_{-1} são insuficientes para prever ε_t por meio de modelos lineares.

2.4 METODOLOGIAS DE PREVISÃO DE CARGAS ELÉTRICAS

2.4.1 Métodos Estatísticos de Previsão

Os métodos estatísticos de previsão de carga podem ser subdivididos em métodos baseados em séries temporais e métodos regressivos, sendo que, os primeiros levam em consideração alguns efeitos diversos, como temperatura, comportamento de carga e efeitos aleatórios, tendo como exemplo: Auto-regressivo Média Móvel (ARMA). Já os modelos regressivos tomam por base a combinação linear de funções, que podem ser exponenciais, senóides, entre outras, podendo ser citadas como exemplo: alisamento exponencial, decomposição espectral e regressão linear.

Dentre os modelos apresentados, os métodos estatísticos apresentam vantagem, tendo em vista serem séries matemáticas já estudadas mais aprofundadamente. Como desvantagem, pode ser trazida a dependência de modelo matemático complexo e a dificuldade de adaptação a rápidas mudanças, como de condições, como o caso de feriados e fatores climáticos, valendo ressaltar que estas técnicas têm sido substituídas por outras, como inteligência artificial ou esta combinada com referidos métodos (RAN-CHANG; DONG-XIAO; ZHENG-YUAN, 2004; ELDESOUKY; ELKATEB, 2000).

2.4.1.1 Regressão Linear Múltipla

A modelagem da carga pode ser obtida através de uma função linear de múltiplas variáveis, da seguinte forma (MOGHRAN; RAHMAN, 1989):

$$y(t) = a_0x_1(t) + \dots + a_nx_n(t) + a(t)$$

1

em que:

$y(t)$ = carga no instante t ;

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

$a(t)$ = sinal de ruído branco com variável aleatória de média zero e variância constante;

$xe_1(t), \dots, xe_n(t)$ = variáveis exógenas;

Variáveis exógenas são capazes de influenciar a carga do sistema como, por exemplo, a temperatura.

O método dos mínimos quadrados e de análise estatística possibilita obter os coeficientes de regressão, definindo a importância de cada variável na previsão de cargas. Porém, é um método pouco utilizado, tendo em vista que outros métodos apresentam erros mais aceitáveis que este (MOGHRAN; RAHMAN, 1989).

2.4.1.2 Séries Temporais Estocásticas

As séries temporais estocásticas consistem em um método bastante útil na previsão de cargas, por meio de modelagem desta. Em cada instante, definido como “ t ”, a carga é obtida como resposta de um filtro linear a um sinal de ruído branco, com média zero e variância constante, valendo ressaltar que, o modelo da previsão é fornecido pelo filtro (GUIRELLI, 2006).



Fonte: Guirelli, (2006)

Em relação ao filtro apresentado na Figura 1, vale trazer alguns modelos encontrados na literatura, conforme segue (GUIRELLI, 2006):

- **Processo Auto-Regressivo (AR):**

O valor de carga $y(t)$ é uma função linear de cargas anteriores $y(t - 1)$, além de ser considerado, ainda, um ruído aleatório $a(t)$, sendo (BOX, 1976):

$$y(t) = \phi_1 y(t - 1) + \phi_2 y(t - 2) + \dots + \phi_p y(t - p) + a(t) \quad 2$$

sendo:

ϕ_p = constantes da função auto-regressiva;

p = número de instantes anteriores usados na regressão.

As cargas passadas até o instante $t - p$ definem a carga no instante t sendo, portanto, o processo auto-regressivo, de ordem p . Já o operador atraso B , este pode ser definido por:

$$y(t - 1) = B y(t) \quad 3$$

seguido por:

$$y(t - m) = B^m y(t) \quad 4$$

sendo a carga da equação:

$$\phi(B) y(t) = a(t) \quad 5$$

em que:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad 6$$

- **Processo da Média Móvel (MA):**

No caso do processo MA (moving average), tem-se como valor de carga, uma função linear que considera valores atuais e passados de uma série de ruído branco $a(t)$, $a(t - 1)$, ...,

que se definem por erros de previsão ou resíduos observados na amostra. Sendo assim (BOX, 1976):

$$y_c(t) = a(t) - \theta_1 a(t-1) - \theta_2 a(t-2) - \dots - \theta_q a(t-q) \quad 7$$

sendo que:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B - \dots - \theta_q B^q \quad 8$$

em que q é a ordem do processo.

- **Processo Auto-Regressivo Integrado da Média Móvel (ARIMA):**

É referencial teórico que séries temporais estejam ligadas a processos estacionários, ou seja, sem alterações sistemáticas de média ou tendências. Para que seja possível trabalhar com processos não-estacionários, é necessário transformá-lo em estacionário, ou seja, remover sazonalidade e tendências, obtendo-se o modelo (MORETIN, 1985):

$$\phi(B) \nabla^d y(t) = \theta(B) a(t) \quad 9$$

sendo:

$$y_c(t) = \phi_1 y_c(t-1) + \dots + \phi_p y_c(t-p) + a(t) - \theta_1 a(t-1) - \dots - \theta_q a(t-q) \quad 10$$

ou seja, o processo acima é de ordem p, d, q .

2.4.1.2 Alisamento Exponencial

O método de alisamento exponencial toma por base que, cada elemento de uma série, se vale de uma constante e uma componente de erro, ou seja:

$$x_o(t) = b + \varepsilon_t \quad 11$$

sendo:

b =componente constante;

ε =componente de erro.

Portanto, verifica-se que b é componente constante de cada segmento, mas também pode variar no tempo. Isso significa que o valor atual de b é resultado das anteriores, ou seja, quanto mais antiga a medição, menor é o peso exponencial desta, como descrito na equação 12:

$$s(t) = \alpha xo(t) + (1 - \alpha) s(t - 1) \quad 12$$

no qual:

s = valores previstos das observações;

xo = valores observados;

α = constante de alisamento;

t = instante em que a série temporal é calculada.

Outros métodos também devem ser considerados, como alguns exemplos encontrados na literatura especializada, quais sejam: Método de Holt-Winters; Previsão Bayesiana; Filtro de Kalman; Método de Box-Jenkins e Decomposição Espectral, merecendo ser considerado que muitos métodos estatísticos ainda se encontram sendo utilizados e ainda aperfeiçoados.

2.4.2 Métodos de Inteligência Artificial

Modelos foram desenvolvidos com base em Inteligência Artificial (IA), possibilitando aos pesquisadores elaborar programas de computadores, conhecidos como sistemas especialistas, a fim de descrever o comportamento de um determinado sistema considerando seu histórico, o que é conhecido por treinamento, interpolando sistemas que, até então, eram independentes de modelos matemáticos complexos ou profundo conhecimento do sistema. Isto somente foi possível por meio de redes neurais artificiais (RNAs) e lógica fuzzy (GUIRELLI, 2006).

Os sistemas especialistas nada mais são do que programas que contam com uma grande base de dados a respeito de determinado assunto, valendo-se de um raciocínio complexo baseado no conhecimento de um especialista humano. Para fins de previsão de carga, torna-se necessário transformar o conhecimento do especialista em regras matemáticas

que, normalmente se encontra aliada a outras técnicas, como, por exemplo: lógica fuzzy e redes neurais, o que possibilita obtenção de melhores resultados, do que quando utilizados separadamente (GUIRELLI, 2006).

2.4.2.1 Redes Neurais Artificiais

Segundo Haykin (1999) “uma rede neural é um processador maciçamente paralelo distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Ela se assemelha ao cérebro em dois aspectos:

1. O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
2. Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimentos adquirido”.

Podem ser citadas como algumas particularidades das redes neurais: não-linearidade, mapeamento de entrada e saída, adaptabilidade, resposta a evidências, informação contextual, tolerância a falhas e uniformidade de análise e projeto. Em função destas características, torna-se possível que uma rede neural se adapte a uma resposta previamente determinada, realize a classificação a padrões, se adapte a modificações com o decorrer do tempo, obtenha robustez computacional, podendo manter ainda o dilema elasticidade/plasticidade, tudo em decorrência da metodologia a ser utilizada, por meio de análise da melhor estrutura neural adequada a cada situação específica (HAYKIN, 1999).

2.4.2.2 Lógica Fuzzy

O conceito de lógica fuzzy surgiu na década de 60 (ZADEH, 1965), mas já vinha sendo estudado desde a década de 20, pelo matemático polonês Jan Lukasiewicz, que desenvolveu a lógica dos conceitos vagos, considerando o grau de pertinência 0, $\frac{1}{2}$, 1, expandindo posteriormente para um conjunto infinito compreendido no intervalo $[0, 1]$.

No conceito fuzzy, um elemento possui grau de pertinência parcial a um dado conjunto, dando margem à imprecisão, tendo em vista um ambiente de incertezas, como ocorre no caso da habilidade humana (ZADEH, 1988), possibilitando sua utilização em diversas áreas, ainda que distintas.

Na teoria clássica dos conjuntos, A pode somente pertencer ou não a um universo U , enquanto na teoria fuzzy, pode assumir um número infinito de valores compreendidos no intervalo $[0, 1]$. Deve ser considerado ainda que, na teoria fuzzy, são consideradas variáveis lingüísticas, como termos primários (alto, baixo, pequeno, médio, grande), modificadores (muito, pouco, levemente, extremamente), etc. (ISODA, 2009).

Quanto à função de pertinência anteriormente citada, vale conceituá-la como uma representação gráfica da magnitude de participação de cada entrada, sendo exemplos: gaussiana, exponencial, triangular, trapezoidal, entre outras. Ao considerar mais de um conjunto fuzzy, como por exemplo: A e B , vários são os operadores que merecem ser considerados, como *AND*, *OR* e *NOT* (KARTALOPOULOS, 1996).

Sendo assim, verifica-se que a teoria dos conjuntos fuzzy se encontra baseada no raciocínio aproximado e não exato, motivo pelo qual acaba por ser utilizada em diversas aplicações de modo bastante eficiente.

2.4.2.3 Algoritmos Genéticos

Os algoritmos genéticos podem ser considerados uma ferramenta viável para a solução de problemas de solução combinatória. Tais algoritmos têm fundamento na teoria de que a evolução natural possibilita a obtenção de soluções eficientes, por meio de mecanismos simples de sobrevivência do indivíduo mais adequado, o qual transfere suas melhores características aos da geração seguinte, seja por passagem ou por mutação (GUIRELLI, 2006).

Apesar de não serem utilizados diretamente para a previsão de cargas elétricas, podem servir de auxílio a alguns métodos que se propõe a atingir tal objetivo, podendo citar como exemplo, as RNAs.

2.5 MÉTODOS HÍBRIDOS

Os métodos híbridos representam uma técnica já bastante difundida na literatura especializada e que demonstra a viabilidade de sua utilização, visando especialmente extrair as melhores características de modelos distintos, em favor da obtenção dos melhores resultados.

No presente trabalho, optou-se por uma rede ART Fuzzy, que combina técnicas próprias das redes neurais, aliadas às características da lógica fuzzy, possibilitando o raciocínio aproximado e não exato, o que traz grandes benefícios a este tipo de rede, como por exemplo, no que diz respeito ao dilema estabilidade/ plasticidade. Tal rede foi utilizada para a classificação de padrões de cargas a ela apresentadas.

Outra rede que compõe o modelo híbrido é a Perceptron Multicamadas, via algoritmo backpropagation, responsável pela previsão de cargas, valendo citar como suas principais características a facilidade de solução de problemas complexos.

Para fins de previsão de cargas elétricas, os modelos híbridos têm sido bastante difundidos, com resultados satisfatórios em relação a outros já descritos na literatura especializada. O presente trabalho demonstra de forma clara a viabilidade de se combinar métodos distintos, no intuito de extrair as melhores características de cada modelo, visando a precisa previsão de cargas de curto prazo.

3 REDES NEURAIS

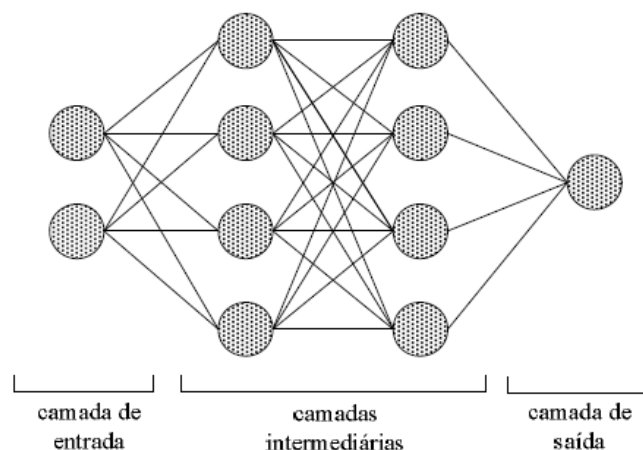
3.1 REDE NEURAL PERCEPTRON MULTICAMADAS, VIA ALGORITMO BACKPROPAGATION

Na literatura especializada, as primeiras redes feedforward que se têm notícias são a Perceptron (ROSENBLATT, 1958) e Adaline (WIDROW, 1987), sendo que, a MLP tem sido utilizada de maneira bastante eficiente ao longo dos anos, principalmente para problemas de reconhecimento de padrões, controle e processamento de sinais.

A MLP possibilita a solução de problemas complexos, especialmente em função da utilização do algoritmo backpropagation, merecendo ser considerado, ainda, o treinamento supervisionado (WERBOS, 1974). O algoritmo backpropagation ou retropropagação tem se demonstrado um método eficiente para o treinamento da rede Perceptron Multicamadas (HAYKIN, 1994).

Quanto à sua estrutura, redes MLP são compostas por camadas sucessivas, com a primeira delas denominada camada de entrada ou input layer, uma ou mais camadas intermediárias ou hidden layers e a camada de saída, também denominada output layer (HAYKIN, 1994), conforme Figura 3.

Figura 3 - Rede Neural Perceptron Multicamadas

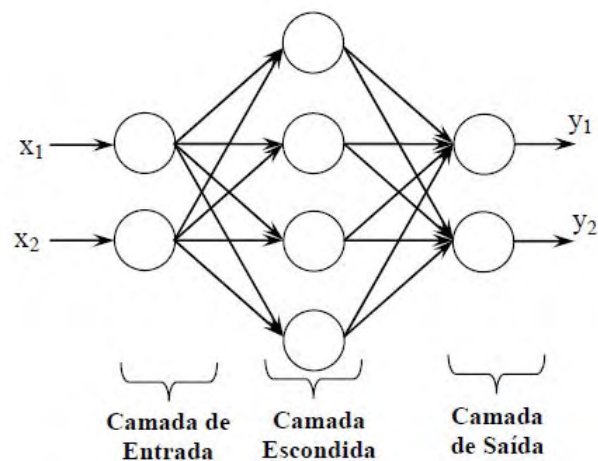


Fonte: Altran, (2010)

Na rede neural MLP, os sinais de entrada se propagam sequencialmente rumo à camada de saída, passando por todos os neurônios das camadas estruturais da rede, ou seja, da esquerda para a direita, de acordo com a Figura 3. As camadas intermediárias têm a função de

extrair características, com pesos que são uma codificação das características dos sinais de entrada, permitindo à rede criar uma representação mais particular, em um formato mais rico e complexo. Sendo corretas as conexões da entrada até um conjunto suficiente de unidades das camadas intermediárias, torna-se possível o correto mapeamento da camada de entrada rumo à saída (HAYKIN, 1994).

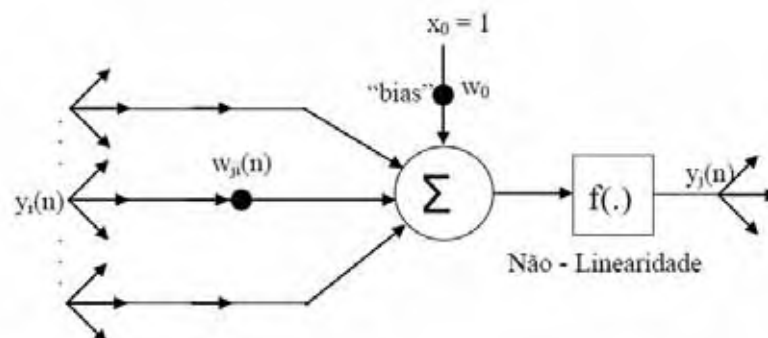
Figura 4 - Estrutura das Redes Feedforward



Fonte: Pereira, (2009)

A Figura 5 ilustra o modelo de um neurônio artificial e nela os sinais de entrada são ponderados por pesos sinápticos e somados no corpo celular do neurônio. A ativação do neurônio se dá por meio de uma função não-linear $f(\cdot)$, transformando o sinal de entrada em estado de ativação. Na Figura 4, é possível visualizar que o neurônio j de uma determinada camada é alimentado por sinais recebidos de neurônios da camada anterior (HAYKIN, 1994).

Figura 5 - Modelo de um Neurônio Artificial



Fonte: Pereira, (2009)

Segundo Cybenko (1989), a rede MLP necessita de, no máximo, duas camadas intermediárias, possuindo unidades suficientes por camada, o que já torna possível obter qualquer tipo de mapeamento, tendo chegado à conclusão que, uma única camada intermediária é suficiente para aproximação de qualquer função contínua.

Três características devem ser consideradas quanto ao projeto de uma rede MLP, quais sejam: a) determinar o número de camadas intermediárias; b) definir o número de neurônios de cada camada intermediária; e c) especificar os pesos sinápticos dos neurônios interconectados de diferentes camadas (CYBENKO, 1989).

Deve-se considerar que, quanto às duas primeiras características, não há uma regra a ser seguida, tendo em vista que a(s) camada(s) intermediária(s) tem o objetivo de servir como influência na relação entrada-saída, o que servirá também influência no comportamento dos dados de entrada, por meio dos quais a rede adquire o conhecimento (CYBENKO, 1989).

No que tange à quantidade de neurônios da camada de saída, só é possível chegar a um número viável se considerarmos a dimensão da resposta desejada, valendo salientar que, o número de neurônios da camada de entrada será definido em vista da dimensão dos dados a serem observados pela rede (HAYKIN, 1994).

A determinação quanto aos pesos sinápticos está intimamente ligada ao algoritmo de treinamento que será utilizado, sendo este o responsável pela adaptação de pesos, sempre considerando os pares de treinamento, representados pelo conjunto entrada/saída e treinamento supervisionado.

Quando o treinamento da rede MLP se dá via algoritmo backpropagation, ocorre da seguinte forma (WERBOS, 1974):

- Na primeira fase, apresenta-se o padrão às unidades da camada de entrada, que se incumbe de propagá-lo à(s) camada(s) intermediária(s). São então calculadas as respostas ainda na(s) camada(s) intermediária(s), produzindo-as na camada de saída, onde também é possível obter o erro;
- Obtido o erro, este é propagado às camadas anteriores (o que justifica o termo “retropropagação”), até atingir a camada de entrada, adaptando-se os pesos das conexões intermediárias, modificando-os.

Quando treinada a rede MLP e minimizado o nível de erro, esta serve como ferramenta para classificação de novos dados. Porém, em vista do fato de já ter ocorrido o treinamento da rede, esta passa a ser ativada somente no modo feedforward (progressivo), ou seja, os sinais apresentados à camada de entrada passam pela(s) camada(s) intermediária(s),

com os resultados apresentados à camada de saída, sem a retropropagação do erro, uma vez que não há mais a necessidade de adaptação de pesos.

O algoritmo backpropagation foi criado por Werbos (1974) e pode ser melhor entendido pelos seguintes passos:

- 1- Apresenta-se um padrão X à rede neural, produzindo uma saída Y;
- 2- É obtida a diferença entre o valor desejado e a saída, ou seja, é efetuado o cálculo de erro de cada saída;
- 3- O erro é retropropagado pela rede, estando associado à derivada parcial do erro quadrático de cada elemento associado aos pesos;
- 4- Os pesos de cada elemento são ajustados;
- 5- Apresenta-se um padrão desconhecido à rede, repetindo o processo até sua convergência, ou seja, $|\text{erro}| \leq \text{tolerância}$ arbitrada.

O algoritmo de retropropagação é baseado no método do gradiente descendente (WASSERMAN, 1989), com adaptação de pesos baseada na minimização dos erros quadráticos. Em primeiro momento, os pesos recebem valores aleatórios do intervalo $\{0,1\}$ e a soma dos erros quadráticos de cada neurônio da camada de saída definido por (WIDROW; LEHR, 1990):

$$\varepsilon^2 = \sum_{i=1}^{ns} \varepsilon_i^2 \quad 25$$

em que:

$$\varepsilon_i = d_i - y_i$$

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.

Tomando por base o i -ésimo neurônio da rede, valendo-se do método do gradiente descendente, obtém-se o ajuste dos pesos, conforme descreve Widrow; Lehr (1990):

$$V_i(h+1) = V_i(h) - \theta_i(h) \quad 26$$

em que:

$$\theta_i(h) = \gamma(\nabla_i(h);$$

γ = taxa de treinamento (parâmetro de controle de estabilidade);

h = índice de iteração;

$\nabla_i(h)$ = gradiente do erro quadrático relacionado aos pesos do neurônio i ;

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

A Equação 26, utilizada para minimização da função objetivo do erro quadrático corresponde exatamente à direção oposta do gradiente. O comprimento do vetor $\theta_i(h)$ é definido pelo parâmetro γ , utilizado para controle de estabilidade do processo iterativo (LOPES; LOTUFO, 2003).

A adaptação dos pesos é definida por (WIDROW; LEHR, 1990):

$$V_i(h+1) = V_i(h) + 2\gamma\beta_i X_i. \quad 27$$

Se acaso, o i -ésimo elemento estiver na última coluna, têm-se:

$$\beta_i = \sigma_i \varepsilon_i. \quad 28$$

Caso o i -ésimo elemento esteja em outra coluna, então:

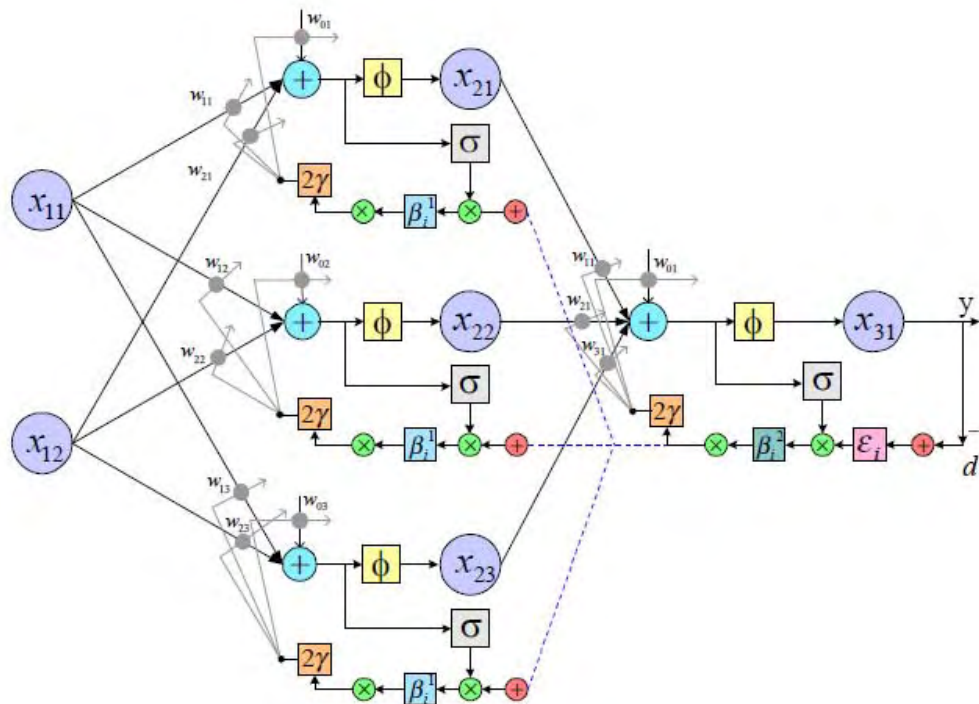
$$\beta_i = \sigma_i \sum_{j \in Q_i} w_{ij} \beta_j \quad 29$$

sendo:

$\sigma_i = \Delta$ derivada da função de ativação referente a s_i ;

Q_i = índices da camada seguinte, do i -ésimo elemento, interligados ao i -ésimo elemento.

Figura 6 - Adaptação dos pesos do algoritmo Retropropagação



Fonte: Altran, (2010)

A adaptação de pesos ocorre conforme demonstra a Figura 6.

3.2 REDE NEURAL ART FUZZY

A teoria da ressonância adaptativa (adaptive resonance theory – ART), proposta originalmente por Grossberg (1976a, 1976b), foi desenvolvida a partir da observação de alguns fenômenos biológicos envolvendo a visão, a fala, o desenvolvimento cortical e as interações cognitivo-emocionais (SILVA, 2010).

Redes neurais artificiais têm como uma de suas principais características a capacidade de generalização, ou seja, produzir respostas a padrões de entrada de forma similar e não idênticos aos padrões apresentados à rede durante o seu treinamento. Valendo-se do treinamento adequado de uma rede multicamadas, torna-se possível que ela forneça como resposta a classificação correta correspondente ao padrão de entrada, e que continue a responder de maneira satisfatória a quaisquer padrões de entrada, dentro do domínio, apresentados futuramente (WERBOS, 1974). Porém, é comum que o conhecimento a respeito de determinado domínio do mundo real se modifique ao longo do tempo. O mesmo acontece com os padrões de entrada de uma RNA, que podem modificar com o tempo para acomodar

as mudanças ocorridas. Assim, para um número de aplicações, o desempenho da rede pode decair gradativamente à medida que o tempo passa, uma vez que os pesos definidos na fase de treinamento não refletem as novas situações. Para adaptar novos padrões de entrada indefinidamente, um algoritmo de aprendizado de RNA deve ser plástico.

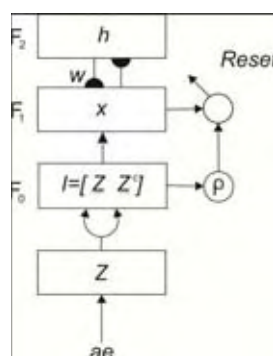
O dilema estabilidade/plasticidade (CARPENTER; GROSSBERG, 1987) está ligado aos atributos das redes neurais da família ART, ou seja, diz respeito à habilidade de aprender novos padrões, preservando os conhecimentos adquiridos anteriormente, por meio de apresentação de outras amostras. Sendo assim, o sistema deve ser flexível (adaptativo), incorporando mudanças do ambiente, devendo também ser estável, no intuito de manter o conhecimento adquirido com o passar do tempo (SILVA, 2010).

A plasticidade está ligada à capacidade de uma RNA se adaptar ou agrupar padrões indefinidamente, enquanto a estabilidade diz respeito à preservação de conhecimentos adquiridos anteriormente, sem que haja sobreposição aos conhecimentos anteriores. Para esse tipo de arquitetura, utiliza-se um algoritmo baseado no critério racional de alocação de padrões, por meio dos vetores de entrada em direção a geração de agrupamentos, que se baseiam na similaridade dos padrões apresentados (CARPENTER; GROSSBERG, 1987).

Uma diversidade de redes neurais da família ART é baseada na neurofísica, valendo citar que a implementação destas RNAs ocorre como soluções analíticas ou aproximação para equações diferenciais.

Já as Redes Neurais ART Fuzzy utilizam a teoria dos conjuntos fuzzy (ZADEH, 1965), com execução de treinamento em tempo reduzido, mantendo o dilema de estabilidade/plasticidade, ou seja, tem capacidade de aprendizado com ajuste de pesos, preservando conhecimento anteriormente adquirido. O reconhecimento de padrão ocorre em tempo real, respondendo a uma sequência arbitrária de entradas e possibilidade de treinamento rápido ou lento, de modo não-supervisionado, com três níveis de atividade: F_0 , F_1 e F_2 .

Figura 7 - Rede Neural ART Fuzzy



Fonte: Silveira, (2003)

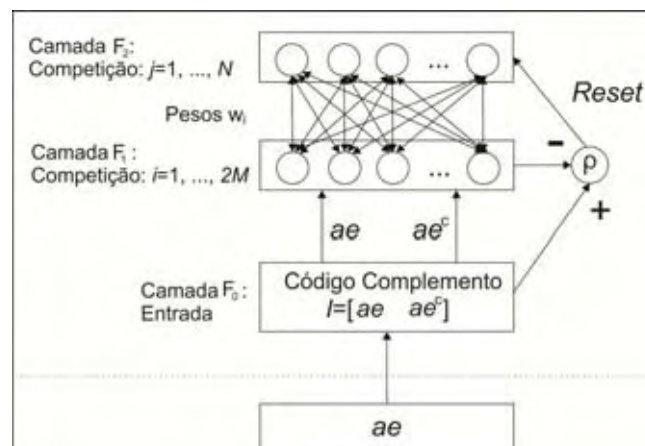
A estrutura da rede ART Fuzzy, disposta na Figura 7, consiste das seguintes camadas (CARPENTER; GROSSBERG, 1987):

- F_0 : onde ocorre o pré-processamento de neurônios, modificando o vetor de entrada ae ;
- F_1 : recebe entradas de F_0 e F_2 , consistindo em $2N$ nós;
- F_2 : camada composta pela quantidade suficiente de nós para que sejam acomodados padrões de entrada em categorias, realizando o treinamento.

Vale trazer ainda, a título de conhecimento, o algoritmo ART Fuzzy:

Redes Neurais ART Fuzzy são baseadas no operador de interseção, representado pelo operador \wedge utilizado na teoria dos conjuntos nebulosos (ZADEH, 1965), possibilitando o processamento de dados binários ou analógicos.

Figura 8 - Algoritmo da Rede Neural ART Fuzzy



Fonte: Lopes, (2005)

A Figura 8 demonstra o funcionamento do algoritmo ART Fuzzy e que se encontra descrito a seguir:

Passo 1. Valores iniciais dos pesos e inicialização de parâmetros da rede:

O processo é iniciado, atribuindo-se o valor 1 a todos os pesos, ou seja, os nós estão descomprometidos ou desativados (CARPENTER & GROSSBERG, 1992), conforme abaixo:

$$w_{j1}(0) = \dots = w_{jm}(0) = 1$$

$$\rho \in [0, 1], \alpha > 0, \text{ e } \beta \in [0, 1];$$

sendo que:

ρ = parâmetro de vigilância : $\rho \in [0,1]$;

α = parâmetro de escolha : $\alpha > 0$;

β = parâmetro de treinamento : $\beta \in [0,1]$.

Passo 2. Normalização de vetor de entrada (CARPENTER; GROSSBERG, 1992):

Esta fase tem como função evitar a proliferação de categorias, sendo que, a normalização é realizada da seguinte forma:

$$I = (ae) / |ae| \quad 30$$

sendo:

ae = vetor de entrada;

$I = [I_1, I_2, \dots, I_m]$ (normalizado);

$|\cdot|$ = função norma.

Define-se a norma por:

$$|ae| \equiv \sum_{i=1}^M |ae_i| \quad 31$$

Passo 3. Codificação complementar:

Nesta fase, preserva-se a amplitude de informação dos vetores de entrada (CARPENTER & GROSSBERG, 1992). Para se obter o complemento do vetor de entrada, temos:

$$ae_i^c \equiv 1 - ae_i; \quad 32$$

fazendo do vetor de entrada:

$$I = [ae \ ae^c] \equiv [ae_1 \dots ae_M \ ae_1^c \dots ae_M^c]. \quad 33$$

Ou seja:

$$|I| = |[aeae^c]|; \quad 34$$

$$= \sum_{i=1}^M ae_i + (M - \sum_{i=1}^M ae_i);$$

$$= M.$$

Passo 4. Escolha de categoria:

O cálculo da função de escolha para cada uma das entradas é feito (CARPENTER; GROSSBERG, 1992), utilizando-se:

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|}. \quad 35$$

sendo:

\wedge =operador fuzzy AND, que se define por (KARTALOPOULOS, 1996):

$$(p \wedge q)_i \equiv \min(p_i, q_i);$$

O índice J é o responsável pela indexação da categoria escolhida, sendo:

$$T_J = \max\{T_j: j = 1, \dots, N\}.$$

No sistema de escolha, a atividade em F_I , vetor x , é dado por:

$$x = I, \text{ se } F_2 \text{ estiver inativo; ou} \\ x = I \wedge w_J, \text{ se o } J\text{-ésimo nó de } F_2 \text{ é escolhido.}$$

Passo 5. Ressonância:

Ocorre ressonância quando a função de combinação $|I \wedge w_j| / |I|$ da categoria escolhida satisfaz o critério de vigilância descrito pela Equação 36 (CARPENTER et al., 1992; GEORGIOPOULOS et al., 1996):

$$\frac{|I \wedge w_j|}{|I|} \geq \rho \quad 36$$

Em caso de não ocorrência do que descreve a equação acima, o valor da função de escolha T_j é fixado em 0 evitando que o processo de busca persista na mesma categoria. Sendo assim, a busca continua até que a categoria escolhida J satisfaça a equação.

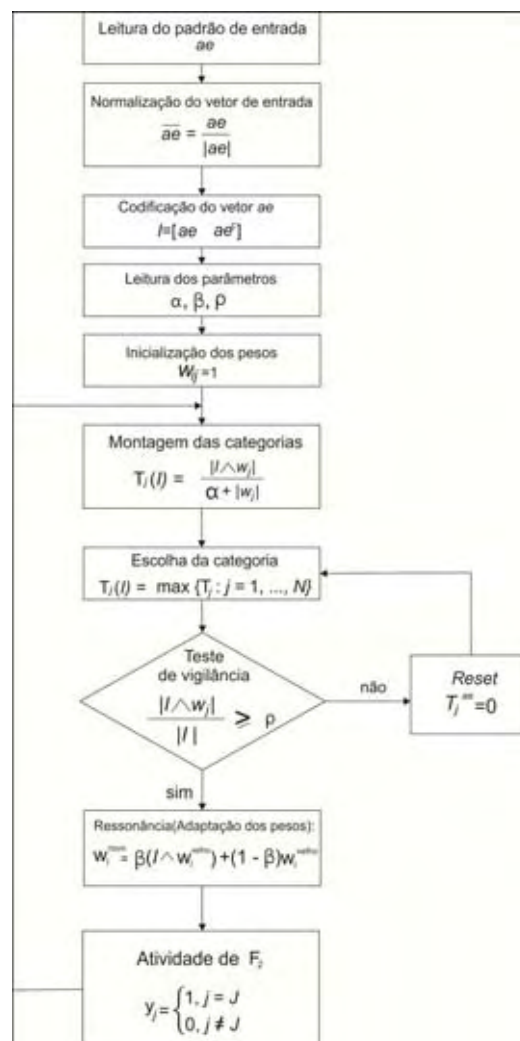
Passo 6. Aprendizagem:

Finalizada a busca, ocorre a adaptação de pesos do vetor w_j , conforme a equação (CARPENTER et al., 1992; GEORGIPOULOS et al., 1996):

$$w_j^{novo} = \beta(I \wedge w_j^{velho}) + (1 - \beta)w_j^{velho} \quad 37$$

Se $\beta = 1$, ocorre o treinamento rápido e o sistema se adapta com mais rapidez. Quando $\beta < 1$, a adaptação de pesos ocorre lentamente, tendo muitos ciclos de aprendizagem.

Figura 9 -Fluxograma da Rede Neural ART Fuzzy



Fonte: Moreno, (2010)

4 REDE NEURAL HÍBRIDA

4.1 INTRODUÇÃO

Neste capítulo será apresentada a metodologia híbrida proposta para a aplicação em previsão de cargas elétricas a curto prazo. O propósito da utilização de uma rede neural artificial híbrida está em obter previsão de cargas elétrica mais precisa e baseado neste fundamento fez-se um modelo híbrido composto de uma rede neural ART Fuzzy, responsável pela classificação de cargas elétricas através de “categorias”, e uma rede neural artificial, MLP via algoritmo backpropagation, a qual realiza a previsão das cargas utilizando como entrada de dados para a rede neural a classificação proveniente da rede neural ART Fuzzy.

Verifica-se, portanto, o uso de uma rede neural baseada na teoria da ressonância adaptativa (CARPENTER; GROSSBERG, 1991): a rede neural ARTFuzzy, a qual mantém o dilema estabilidade/plasticidade, com possibilidade de treinamento rápido ou lento, de maneira não-supervisionada e capacidade de generalização, produzindo respostas similares, mas não idênticas aos padrões apresentados no treinamento. A estabilidade se relaciona à capacidade de rede neural de se adaptar a novos padrões, enquanto a plasticidade garante o treinamento continuado, sem necessidade de novo treinamento. Tais características proporcionam flexibilidade do sistema podendo, portanto, ser considerado um sistema adaptativo.

Uma segunda rede neural que compõe o modelo é a rede neural Perceptron Multicamadas, via algoritmo backpropagation (WERBOS, 1974), bastante popular nos últimos anos. Esta rede neural utiliza o treinamento supervisionado e este possibilita a atualização de pesos da rede neural baseado no método do gradiente descendente (WASSERMAN, 1989) e que tem o objetivo de minimizar o erro quadrático da rede.

4.2 REDE NEURAL BACK-ART FUZZY

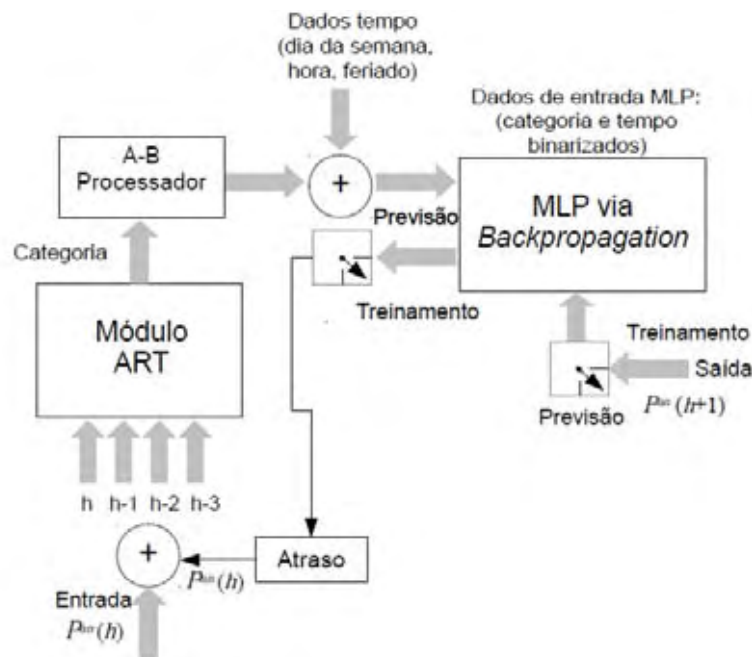
A rede neural Back-ART Fuzzy é uma arquitetura híbrida que favorece a utilização de dados de entrada na forma binária e dados de saída analógicos. Trata-se de uma rede neural híbrida composta por duas arquiteturas de redes neurais artificiais:

1. ART Fuzzy (CARPENTER; GROSSBERG, 1991);
2. MLP via algoritmo backpropagation (WERBOS, 1974).

A formulação das variáveis aplicadas ao problema de previsão de carga é feita através de variáveis híbridas, ou seja, dados binários (utilizados para expressar os dias, meses, anos, etc.) e dados analógicos (que são representados por valores de carga elétrica, temperatura, etc.). O motivo de utilizar a rede neural ART Fuzzy parte do princípio em classificar em “categorias” ou “clusters” os dados analógicos, para desta forma, o conjunto de dados de entrada utilizados no processamento da rede neural MLP seja constituído apenas de dados binários (provenientes das categorias binarizadas e dados binários referentes ao tempo (dia, mês, etc.)).

Na Figura 10, apresenta-se a rede neural Back-ART Fuzzy ajustada para resolução do problema de previsão de carga elétrica. Pode-se notar facilmente, as duas arquiteturas de redes neurais artificiais utilizadas: ART Fuzzy e MLP. À esquerda e à direita encontram-se, respectivamente, as duas redes neurais ART Fuzzy e MLP via backpropagation.

Figura 10 - Rede Neural Back-ART Fuzzy



Fonte: do próprio autor

A rede neural híbrida é constituída de duas fases:

- (1) treinamento;
- (2) previsão.

Para realização da fase de treinamento, na rede neural híbrida, as chaves deverão indicar a modo “treinamento”. A entrada da rede neural híbrida, Back-ART Fuzzy, é composta pelo vetor $P^{an}(h)$ (dados analógicos), pelo escalar $P^{an}(h+1)$, correspondente ao

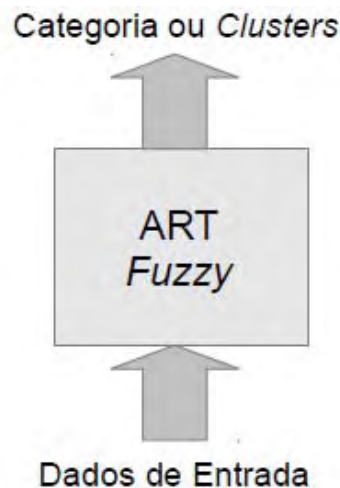
dado de saída da rede neural sofrendo um atraso de uma hora, e pelo vetor correspondente aos dados de tempo binarizados (dia da semana - domingo a sábado, hora e feriado). A saída da rede neural Back-ART Fuzzy é representada pela carga elétrica referente a 1 hora à frente. Para a efetuar a previsão, as chaves devem ser colocadas no modo “previsão”. Desta forma, a entrada escalar $P^{an}(h)$ é gerada pela rede neural híbrida correspondente a carga com um atraso de 1 hora em relação a carga prevista $P^{an}(h + 1)$.

A entrada de dados da rede neural híbrida, Back-ART Fuzzy, é formada pela junção de dois subconjuntos de dados.

1. “Categorias” ou “clusters” binarizados obtidos pela rede neural ART Fuzzy;
2. Dados de tempo (dia da semana – domingo a sábado, hora e feriado).

A obtenção do primeiro conjunto de dados a ser agregado a entrada de dados da rede neural híbrida é dado através da rede neural ART Fuzzy. Os dados de entrada para a rede neural ART Fuzzy é um conjunto que contém apenas dados analógicos. A arquitetura ART Fuzzy tende a agrupar os dados de entrada analógicos, através de similaridades tratadas dentro da arquitetura da rede neural ART Fuzzy ilustrada na Figura 11.

Figura 11 -ART Fuzzy- Processo de Classificação de Dados Analógicos



Fonte: do próprio autor

Os dados de entrada da rede neural ART Fuzzy é dada por:

$$P^{an}(h) = [L(h - q) \dots L(h - 2) L(h - 1) L(h)]$$

sendo:

$L(h - q)$ = valor da carga, q horas antes da hora atual h ;

O conjunto de entrada de dados para a rede neural ART Fuzzy é formado de dados de carga elétrica horária obtidas através de um janelamento, ou seja, carga atual, $L(h)$, uma - $L(h-1)$, duas - $L(h-2)$ e três cargas anteriores - $L(h-3)$. Este janelamento representa o laço de realimentação da saída com um atraso, respectivamente, de 3, 2 e 1 hora da hora corrente h .

A estratégia utilizada neste trabalho é o janelamento realizado levando em consideração a carga elétrica corrente e três valores de carga elétricas anteriores. Contudo, deve-se considerar que existem inúmeras estratégias propostas na literatura. O que distingue uma das demais, refere-se aos dados que são usados em cada janela.

O vetor $P^{an}(h)$ constitui a entrada da rede neural ART Fuzzy e a saída da rede neural ART Fuzzy é dada pela classificação de similaridade do conjunto de dados através de “categorias” ou “clusters” e estes são convertidos via processador códigos ativo-(A)/binário-(B) em informação binária.

A formação do segundo conjunto de dados se dá através da informação de dados binários correspondente a dados que se referem ao tempo, neste caso, o dia da semana – domingo a sábado, a hora correspondente a cada carga elétrica e a informação se este dia da semana é um feriado. Este conjunto de dados possui um total de informação de 9 bits.

A partir dos dois conjuntos de dados tem-se o conjunto de padrões de entrada de dados para a rede neural Perceptron Multicamadas via backpropagation. Portanto, a entrada da rede neural MLP via backpropagation é composta apenas por dados binários formadas pelo vetor binário da saída do processador A/B, acrescido do vetor binário que representa os dados do tempo.

A rede neural MLP tem a incumbência de realizar todo o processamento e consequentemente a previsão de cargas elétricas. O conjunto de dados de entradas da MLP será composta da categoria proveniente do módulo ART Fuzzy dos dados de tempo relacionados a dia da semana (domingo a segunda), hora e feriado. Os dados de saída da rede neural MLP é formada por valores de cargas elétricas horárias relativas à hora subsequente aos valores de carga horária, ou seja, $h+1$. Portanto, os padrões de entrada e saída de dados da rede neural MLP, respectivamente, é dado por:

$$X_{MLP} = [CAT^{bin} t^{bin}]^T \quad 38$$

$$Y_{MLP} = [L(h+1)]^T \quad 39$$

sendo:

CAT^{bin} = vetor binarizado da categoria obtida pelo módulo ART Fuzzy;

t^{bin} = vetor contendo dados de tempo (dia da semana, hora e feriado) binarizado;

$L(h+1)$ = Carga elétrica referente a hora $h+1$.

Considerando uma formação de até 480 categorias distintas, tem-se que o conjunto de dados obtidos pela rede neural ART Fuzzy é decodificado em 9 bits. O conjunto de dados com informação de tempo – dia da semana, hora e feriado também possui uma configuração binária de 9 bits (3 bits correspondente ao dia da semana, 5 bits referentes a hora e 1 bit para indicação do feriado). Portanto, A dimensão do vetor de entradas do MLP via Backpropagation tem um total de informação de 18 bits.

Assim, a previsão de carga elétrica corresponde aos valores da saída da rede neural híbrida Back-ART Fuzzy para o intervalo de tempo considerado (por exemplo, a previsão diária, 24 horas à frente).

Em resumo, a previsão de cargas elétricas futuras é realizada por meio da utilização de duas redes neurais distintas, sendo a primeira, uma rede ART Fuzzy e, em segundo momento, por uma rede MLP via algoritmo backpropagation.

Utilizou-se valores de cargas elétricas reais fornecidas por uma empresa do setor elétrico brasileiro, valendo-se de intervalos temporais pré-definidos, de acordo com o interesse de previsão proposto, quais sejam: a) cargas horárias de trinta e um dias, referentes ao período compreendido entre 1º e 31 de julho de 1998, ou seja, 744 cargas; e b) cargas horárias de sessenta e um dias, referentes ao período compreendido entre 1º de junho e 31 de julho de 1998, representando 1464 cargas.

Em qualquer dos períodos acima, foram acrescentadas três cargas das últimas horas imediatamente anteriores à primeira, uma vez que, por meio de janelamento, foi possível considerar como entrada da primeira rede neural (ART Fuzzy): a carga relativa à primeira hora do período, definida como “ h ”; uma hora anterior à primeira, definida como “ $h - 1$ ”; duas horas anteriores à primeira, definida como “ $h - 2$ ” e, por último, três horas anteriores à primeira, descrita como “ $h - 3$ ”. Sendo assim, o primeiro banco de dados, que era formado por uma matriz coluna de dimensão 744x1, passou a um novo formato: 747x1, tendo em vista

a inclusão de carga das três horas anteriores à primeira e, por meio do janelamento, formou-se uma nova matriz; agora de dimensão 744x4.

Obtido o janelamento e formada a nova matriz (744x4), tais dados foram utilizados como entrada da primeira rede neural (ART Fuzzy), a qual é responsável pela classificação destes, conforme descrito passo a passo no item 4.2. Ou seja, a rede ART Fuzzy, neste primeiro momento, ainda se apresenta com exclusividade no método proposto, o que é alterado a partir da classificação, o que ocorre linha a linha, tomando por base a matriz de dimensão 744x4, considerando então a hora “ h ”, hora “ $h - 1$ ”, hora “ $h - 2$ ” e hora “ $h - 3$ ”, ou seja, a primeira das 744 cargas e suas três anteriores, a segunda das 744 cargas e suas três anteriores e assim sucessivamente, obtendo-se a classe a que pertence cada uma das cargas. Esta classe (categoria) é binarizada para que seja utilizada como parte das entradas da segunda rede neural (MLP via backpropagation).

Verifica-se que, a partir da geração da carga referente à hora $h+1$, ambas redes neurais passam a ter atividade simultânea, uma vez que tal carga, obtida pela rede MLP (745ª carga), aliada à três cargas anteriores, servem de entrada para a rede ART Fuzzy que, mais uma vez, classifica o conjunto de dados, novamente unidos ao dia, hora e o fato de ser ou não feriado, formando uma nova entrada para a segunda rede neural (MLP), que prevê a segunda carga elétrica posterior ao período considerado e assim sucessivamente, demonstrando que as redes passam a atuar, como já dito anteriormente, de forma simultânea e também, ininterrupta.

O mesmo ocorre em relação ao período de dois meses, que conta com 1464 cargas do banco de dados histórico da mesma companhia do setor de energia elétrica brasileira, sendo estes acrescidos das três horas anteriores para efeito de janelamento, vindo posteriormente a ser realizada a classificação pela rede neural ART Fuzzy e, em seguida, adicionados dia e hora respectivos, além do fato de ser ou não feriado, dados estes que são utilizados como entrada da rede MLP via algoritmo backpropagation para previsão da primeira hora posterior às apresentadas à tal rede. Assim, passa-se à fase de atividade simultânea de ambas redes neurais, como descrito anteriormente.

A Tabela 1 ilustra os valores de cargas elétricas reais, a fim de que seja visualizado o formato do banco de dados original.

TABELA 1 - CARGAS ELÉTRICAS REAIS

DIA	TEMPO (HORA)	CARGA (MVA)
01/07/1998	01	2515,5
01/07/1998	02	2345,4
01/07/1998	03	2350,6
01/07/1998	04	2318,9
01/07/1998	05	2334,9
01/07/1998	06	2429,0
01/07/1998	07	2616,4
01/07/1998	08	2778,1
01/07/1998	09	2969,1
01/07/1998	10	3090,0
01/07/1998	11	3201,8
01/07/1998	12	3261,9
01/07/1998	13	3240,2
01/07/1998	14	3241,3
01/07/1998	15	3251,9
01/07/1998	16	3237,9
01/07/1998	17	3301,4
01/07/1998	18	3450,8
01/07/1998	19	3706,5
01/07/1998	20	3518,5
01/07/1998	21	3444,3
01/07/1998	22	3413,9
01/07/1998	23	3191,3
01/07/1998	24	2899,1

Fonte: do próprio autor

O sistema proposto se vale de redes neurais distintas, caracterizando assim, hibridismo do método, dando a ele particularidades bastante vantajosas se comparado a outros já vistos.

Vale salientar que os métodos híbridos, combinando duas ou mais técnicas e valendo-se de vantagens particulares de cada uma, torna possível obter mais exatidão na previsão de cargas, merecendo ser citados, a título de exemplo, o neural genético, o neuro-fuzzy, etc.

Prova do acima alegado se encontra no presente estudo, no qual foi utilizado o hibridismo de redes neurais e obtendo, em função disso, bons resultados, o que será apresentado nos tópicos subsequentes.

4.3 COMENTÁRIOS

Neste capítulo demonstra-se a viabilidade de utilização do hibridismo proposto para a previsão de cargas elétrica a curto prazo. Por meio desse método é possível conseguir vantagem quanto ao custo computacional, bem como robustez do sistema em relação a outros já descritos na literatura especializada.

Este trabalho propõe o desenvolvimento do sistema para previsão de carga elétrica por meio de um modelo neural híbrido, Back-ART Fuzzy, baseado numa composição de duas arquiteturas de redes neurais artificiais: ART Fuzzy e Perceptron Multicamadas via algoritmo backpropagation. O hibridismo se torna necessário de forma a utilizar as características e vantagens de ambas arquiteturas de redes neurais, sendo que, o Perceptron Multicamadas via backpropagation é considerado na literatura um benchmark em termos de precisão e as redes neurais baseadas na teoria da ressonância adaptativa (ART - Adaptive Resonance Theory) possui uma das principais características necessárias para o funcionamento da rede (dilema da plasticidade e estabilidade).

Através dos resultados obtidos pela rede neural híbrida e que serão apresentados no próximo capítulo (capítulo 6) pode-se observar a viabilidade da utilização da metodologia proposta neste trabalho e que representa grande vantagem quando comparado a outros já vistos anteriormente, tendo apresentado resultados eficientes, tanto em relação ao MAPE, como ao erro máximo, estando em conformidade com limites aceitáveis para previsão de cargas elétricas de curto prazo sendo que, segundo Ghiassi (2006), os resultados de MAPE variam tipicamente de 3 a 12%.

O método em análise demonstrou viabilidade de sua utilização por empresas do setor elétrico, restando cabalmente provada sua eficiência através dos resultados obtidos ao final dos testes.

5 APLICAÇÃO, TESTES E RESULTADOS

O presente trabalho tem como objetivo demonstrar a viabilidade da aplicação da rede neural Back-ART Fuzzy na previsão de cargas elétricas a curto-prazo e, para tanto, foram feitas duas modalidades de aplicação utilizando dados de carga elétrica horária, com base em dados históricos de uma empresa do setor elétrico brasileiro, respectivamente em relação ao mês de julho de 1998 (744 vetores padrão) e junho e julho do mesmo ano (1464 vetores padrão), na intenção de prever as cargas elétricas para o dia 1º de agosto de 1998.

A utilização do hibridismo, como ocorre no presente trabalho, combinando uma rede neural ART Fuzzy e a PerceptronMulticamadas, via algoritmo backpropagation, se mostra uma técnica bastante eficiente na previsão das referidas cargas elétricas, com resultados satisfatórios e viáveis quando da utilização do método.

Com o intuito de verificar a precisão da previsão de cargas elétricas para metodologia proposta, foi utilizado o cálculo do erro médio percentual absoluto (MAPE), bem como do erro máximo da previsão, o que ocorre da seguinte forma:

$$MAPE = \frac{1}{N} \sum_{h=1}^N \left\{ \frac{|L(h) - \underline{L}(h)|}{L(h)} \right\} 100\% \quad 40$$

$$Erro\ máximo\ (\%) = \max \left\{ \frac{|L(h) - \underline{L}(h)|}{L(h)} \right\} 100\% \quad 41$$

em que:

$L(h)$ =valor da carga real referente a hora h ;

$\underline{L}(h)$ =valor da carga estimada pelo modelo híbrido referente a hora h ;

N = número total de horas.

A Tabela 2 abaixo descreve os períodos adotados para efeito de testes do modelo Back-ART Fuzzy:

TABELA 2 - PERÍODOS DE TESTES

Aplicação	Período	Vetores padrão	Data prevista
1	01/07/98a 31/07/98 31 dias	744	01/08/98
2	01/06/98 a 31/07/98 61 dias	1464	01/08/98
3	01/05/98 a 31/07/98 92 dias	2208	01/08/98

Fonte: do próprio autor

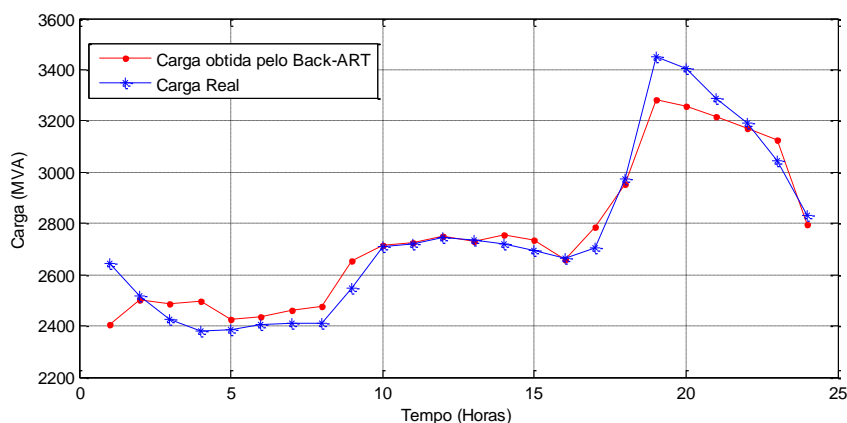
Cumpra, então, passar à análise dos resultados obtidos a partir dos dados acima descritos, sendo que, na Tabela 3 têm-se os parâmetros utilizados para a aplicação 1, os quais possibilitaram o seguinte resultado da rede Back-ART Fuzzy:

**TABELA 3 -PARÂMETROS DA REDE NEURAL BACK-ART FUZZY
APLICAÇÃO 1 - 744 VETORES PADRÃO**

PARÂMETROS	VALORES
Número de vetores padrão	744
Taxa de treinamento - ART Fuzzy	1,0
Parâmetro de escolha	0,1
Parâmetro de vigilância - Treinamento	0,98
Parâmetro de vigilância - Diagnóstico	0,5
Número de Camadas	3
Quantidade de neurônios (por camada)	18-35-1
Tolerância	0,08
Taxa de Treinamento	2,0
Constante Momento	0,9
Lambda	1,0

Fonte: do próprio autor

A Figura 12 mostra o resultado obtido pela rede neural híbrida Back-ART Fuzzy, para previsão de cargas elétricas de curto prazo, utilizando-se as cargas do mês de julho de 1998 para o treinamento da mesma.

Figura 12 - Resultado da Rede Híbrida Back-ART Fuzzy(744 vetores padrão)

Fonte: do próprio autor

Valendo-se dos parâmetros acima, foi possível a obtenção dos seguintes resultados de MAPE e Erro Máximo quanto aos 744 vetores padrão de treinamento (julho de 1998), para fins de previsão de carga de curto prazo, conforme Tabela 4.

**TABELA 4 - MAPE E ERRO MÁXIMO OBTIDOS
APLICAÇÃO 1**

MAPE (%)	2,176844
Erro Máximo (%)	9,021447
Tempo de processamento (centésimos de segundo)	14

Fonte: autor da dissertação

A seguir, se encontram os resultados da aplicação 2, em que foram considerados para treinamento da rede híbrida, 61 dias de carga reais, representando 1464 cargas e, na Tabela 5, têm-se os parâmetros utilizados para a aplicação 2, os quais possibilitaram o seguinte resultado da rede Back-ART Fuzzy.

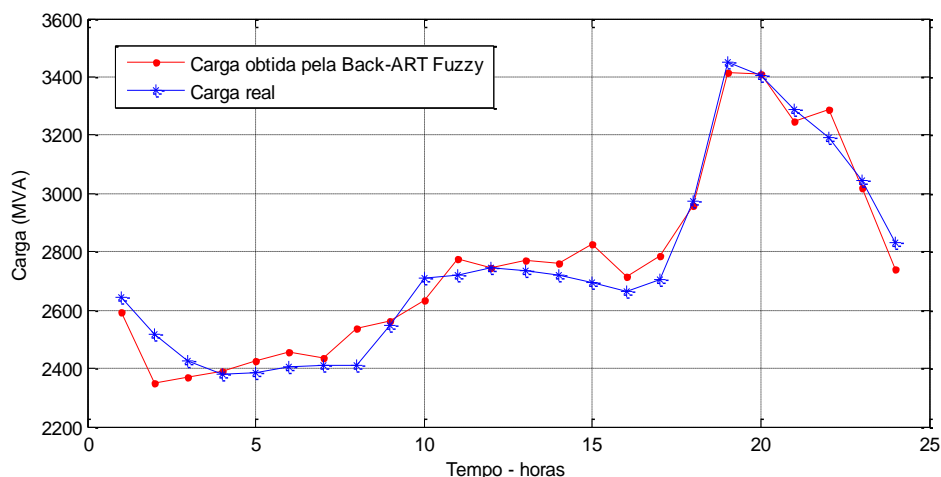
**TABELA 5 - PARÂMETROS DA REDE NEURAL BACK-ART FUZZY
APLICAÇÃO 2 - 1464 VETORES PADRÃO**

PARÂMETROS	VALORES
Número de vetores padrão	1464
Taxa de treinamento - ART Fuzzy	0,1
Parâmetro de escolha	1,0
Parâmetro de vigilância - Treinamento	0,98
Parâmetro de vigilância - Diagnóstico	0,6
Número de Camadas	3
Quantidade de neurônios (por camada)	18-34-1
Tolerância	0,06
Taxa de Treinamento	4,0
Constante Momento	0,8
Lambda	0,64

Fonte: do próprio autor

A Figura 13 expressa o resultado obtido pela rede neural híbrida Back-ART Fuzzy, para previsão de cargas elétricas de curto prazo, utilizando-se as cargas dos meses de junho e julho de 1998 para o treinamento.

Figura 13 - Resultado da Rede Híbrida Back-ART Fuzzy(1464 vetores padrão)



Fonte: do próprio autor

Quando utilizados os parâmetros descritos na Tabela 5, obteve-se os seguintes resultados de MAPE e Erro Máximo quanto aos 1464 vetores padrão de treinamento (junho e julho de 1998), para fins de previsão de carga de curto prazo, descritos na Tabela 6.

**TABELA 6 - MAPE E ERRO MÁXIMO OBTIDOS
APLICAÇÃO 2**

MAPE (%)	2,068939
Erro Máximo (%)	6,743864
Tempo de Processamento (centésimos de segundo)	95

Fonte: do próprio autor

Ainda a título de análise, encontram-se abaixo os resultados obtidos com base na aplicação 3, em que foram considerados para treinamento da rede híbrida, 92 dias de carga reais, representando 2208 cargas e, na Tabela 7, têm-se os parâmetros utilizados para referida aplicação, os quais possibilitaram o seguinte resultado da rede Back-ART Fuzzy.

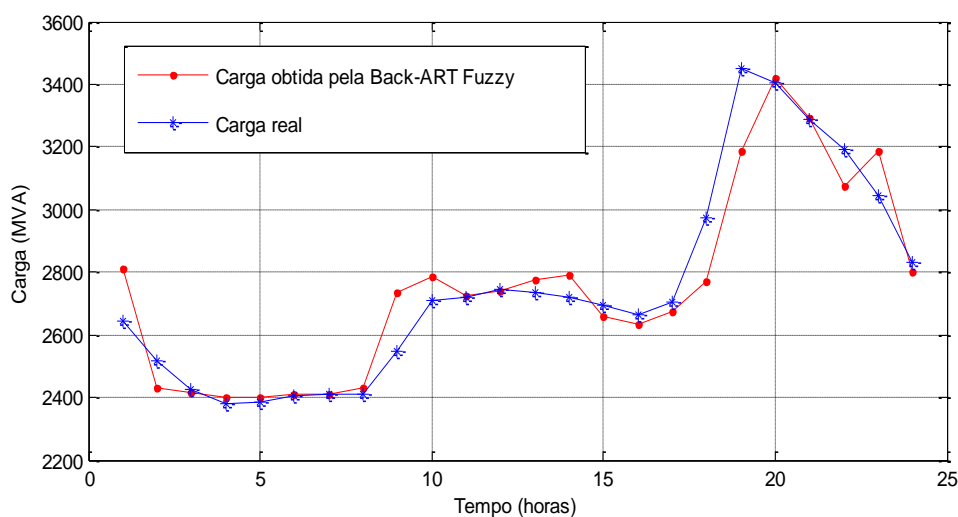
**TABELA 7 - PARÂMETROS DA REDE NEURAL BACK-ART FUZZY
APLICAÇÃO 3 - 2208 VETORES PADRÃO**

PARÂMETROS	VALORES
Número de vetores padrão	2208
Taxa de treinamento - ART Fuzzy	0,1
Parâmetro de escolha	1,0
Parâmetro de vigilância - Treinamento	0,99
Parâmetro de vigilância - Diagnóstico	0,7
Número de Camadas	3
Quantidade de neurônios (por camada)	18-39-1
Tolerância	0,06
Taxa de Treinamento	2,5
Constante Momento	0,7
Lambda	0,8

Fonte: do próprio autor

Na Figura 14 expressa o resultado obtido pela rede neural híbrida Back-ART Fuzzy, para previsão de cargas elétricas de curto prazo, utilizando-se as cargas dos meses de maio, junho e julho de 1998 para o treinamento.

Figura 14 - Resultado da Rede Híbrida Back-ART Fuzzy(2208 vetores padrão)



Fonte: do próprio autor

Utilizando-se dos parâmetros descritos na Tabela 7, obteve-se os seguintes resultados de MAPE e Erro Máximo quanto aos 2208 vetores padrão de treinamento (maio, junho e julho de 1998), para fins de previsão de carga de curto prazo, descritos na Tabela 8.

TABELA 8 - MAPE E ERRO MÁXIMO OBTIDOS

APLICAÇÃO 3

MAPE (%)	2,320467
Erro Máximo (%)	7,635786
Tempo de Processamento (segundos)	20,68

Fonte: do próprio autor

Para análise dos resultados aqui obtidos em relação ao MAPE e erro máximo, para previsão de carga de curto prazo, comparado com uma arquitetura anteriormente apresentada no trabalho de (NOSE FILHO et al, 2009), o que se encontra expresso na Tabela 9 abaixo:

TABELA 9 - COMPARAÇÃO DE MAPE E ERRO MÁXIMO

Teste	Modelos	MAPE (%)	Erro máximo (%)
Aplicação 1	Back-ART (744 vetores padrão)	2,176844	9,021447
	RNA feedforwardcom algoritmo gradiente descendente com momento	1,998	4,845
Aplicação 2	Back-ART (1464 vetores padrão)	2,068939	6,743864
	RNA feedforwardcom algoritmo gradiente descendente com momento	2,093	5,022
Aplicação 3	Back-ART (2208 vetores padrão)	2,320	7,635
	RNA feedforwardcom algoritmo gradiente descendente com momento	2,313	7,088

Fonte: do próprio autor

6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

6.1 CONCLUSÕES

A utilização de redes neurais, aliadas ao hibridismo, tem possibilitado grandes avanços na previsão mais precisa de energia elétrica, seja de curto, médio ou longo prazo de forma satisfatória e com baixo custo computacional. O presente trabalho teve o intuito de utilizar as técnicas do hibridismo, fazendo com que o método proposto aproveitasse as melhores características dos dois modelos de RNAs, quais sejam: ARTFuzzy e Perceptron Multicamadas, via algoritmo backpropagation.

Por meio da rede ART Fuzzy, foi possível obter a classificação das cargas de um banco de dados histórico de uma companhia do setor elétrico nacional, o que foi possibilitado com o auxílio de janelamento do vetor padrão, fazendo com que a primeira hora (hora " h ") fosse aliada às três anteriores (" $h - 1$ ", " $h - 2$ " e " $h - 3$ "), para fins de classificação, independente da quantidade de dados do vetor padrão. Obtido o número da classe, a qual pertencia a carga, atrelando a ela o dia da semana, hora respectiva e o fato de ser ou não feriado, serviu de entrada para a RNA Perceptron Multicamadas, via algoritmo backpropagation, sendo esta a responsável pela previsão da primeira carga imediatamente posterior à última do vetor padrão, retornando-a à primeira das redes para nova classificação e passando, assim, a ocorrer a atividade simultânea das redes neurais.

No intuito de obter a previsão de carga de curto-prazo, mais precisamente, 24 horas seguintes às contidas no vetor padrão, foram realizados dois testes, respectivamente, em relação a 744 cargas, relativas ao período de 1 a 31 de julho de 1998, bem como 1464 cargas, referentes ao período compreendido entre 1º de junho a 31 de julho do mesmo ano. Tanto no caso da primeira aplicação (744 cargas), como no da segunda (1464 cargas) e terceira aplicações (2208 cargas) foram obtidos MAPE pouco superior a 2% e erros máximos de 9,021447%, 6,743864% e 7,635786%, demonstrando a viabilidade da aplicação e da metodologia proposta, por meio de um programa robusto, com baixo custo computacional e alto desempenho..

No que diz à previsão de carga de curto prazo, o trabalho demonstrou que a metodologia proposta se mostra bastante eficiente, com capacidade de apresentar resultados

extremamente satisfatórios, obtendo-se resultados que representam grande vantagem em relação a outros modelos já descritos na literatura especializada.

6.2 SUGESTÕES PARA TRABALHOS FUTUROS

A título de sugestão para trabalhos futuros, podem ser citadas:

- * aperfeiçoar o método híbrido proposto, considerando outros fatores de influência quanto à previsão de cargas elétricas;
- * desenvolvimento de método similar para previsão de cargas de médio e longo prazo;
- * comparar o modelo proposto a outro com a utilização de algoritmo distinto ao ora apresentado;
- * realizar outros tipos de previsão, como preços, ventos, afluência, utilizando o algoritmo descrito na presente dissertação;
- * melhoria da estratégia para a constituição de dados de entrada.

REFERÊNCIAS

ABE, S. **Neural networks and fuzzy systems: theory and applications**. Boston: Kluwer Academic, 1997. 258 p.

ABRAHAM, A; NATH, B.A Neuro-fuzzy approach for modeling electricity demand in victoria., **Applied Soft Computing Journal**, Amsterdam, v. 1, n. 1, p. 127-138, 2001.

AL-KANDARI, A. M.; SOLIMAN, S. A.; EL-HAWARY, M. E. Fuzzy short-term electric load forecasting. **Electrical Power & Energy Systems**, Guildford, v. 26, n.2, p. 111-122, 2004.

ALTRAN, A. B. **Sistema inteligente para previsão de carga multinodal em sistemas elétricos de potência**. 2010. 86 f. Tese (Doutorado) – Faculdade de engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2010.

AMARAL, L. F. M. **Modelos lineares e não-lineares na modelagem do preço Spot de energia elétrica do Brasil**. 2003. 77 f. Dissertação (Mestrado) - Faculdade de Engenharia, Pontifícia Universidade Católica, Rio de Janeiro.

BOX, G.E.; JENKINS, G.M. **Times series analysis: forecasting and control**. San Francisco: Holden - Day, 1976. 575 p.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. **Times series analysis: forecast and control**. 3. ed. Englewood Cliffs: Prentice Hall, 1994. 598 p.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2.^a ed. Rio de Janeiro: LTC, 2000. 250 p.

CAMPOS, J. R. **Desenvolvimento de um sistema dinâmico para predição de cargas elétricas por redes neurais através do paradigma de programação orientada a objeto sob a linguagem java**. 2010. 66f. Dissertação (Mestrado)- Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2010.

CARPENTER, G. A.; GROSSBERG, S. A. Massively parallel architecture for a self-organizing neural pattern recognition machine. **Computer Vision, Graphics and Image Processing**, San Diego, v. 37, n. 1, p. 54-115, 1987a.

CARPENTER, G. A.; GROSSBERG, S. ART 2: self-organizing of stable category recognition codes for analog input patterns. **Applied Optics**, New York, v. 26, n. 23, p. 4919-4930, 1987b.

CARPENTER, G. A.; GROSSBERG, S. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system., **Neural Network**, New York, v. 4, n. 6, p. 759-771, 1991.

CARPENTER, G. A. et al. A neuralnetwork architecture for incremental supervised learning of analog multidimensional maps. **IEEE Transactions on Neural Networks**, Piscataway, v. 3, n.5, p. 698-713, 1992.

CYBENKO, G. Approximation by superpositions of sigmoidal function. **Mathematics of Control Signals, and Systems**, New York, v. 2, n. 4, p. 303-314, 1989.

EL DESOUKY, A. A.; ELKATEB, M. M. Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA: generation, transmission and distribution., **IEE Proceedings-Generation Transmission and Distribution**, Hertford, v. 147, n. 4, p. 213-217, 2000.

GEORGIOPOULOS, M. et al. Order of search in fuzzy ART and Fuzzy ARTMAP: effect of the choice parameter. **Neural Networks**, New York, v. 9, n. 9, p. 1541-1559, 1996.

GOMIDE, F. A. C.; GUDWIN, R. R.; TANSCHKEIT, R. Conceitos fundamentais da teoria de conjuntos fuzzy, lógica fuzzy e aplicações. **IFSA WORLD CONGRESS, 6.**, São Paulo. **Proceedings...**, São Paulo: IFSA, Brasil, 1995. p. 1-38. Tutorials.

GUIASSI, M.; ZIMBRA, D. K.; SAIDANE, H. Medium term system load forecasting with a dynamic artificial neural network model. **Electric Power Systems Research**, Lausanne, v. 76, n. 5, p. 302-316, 2006.

GUIRELLI, C. R. **Previsão da carga de curto prazo de áreas elétricas através de técnicas de inteligência artificial**. 2006. 115 f. Tese (Doutorado) - Escola Politécnica, Universidade de São Paulo. USP, 2006.

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

HAYKIN, S. **Neural networks: a comprehensive foundation**. 2. ed. New Jersey: Prentice-Hall, 1999. 842 p.

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

ISODA, L. Y. **Análise da estabilidade estática de tensão de sistemas elétricos de potência usando uma rede neural baseada na teoria da ressonância adaptativa**. 2009. 113 f. Tese (Doutorado) - Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2009.

QUEIROZ, A. R. **Simulação de custos marginais em mercado de energia elétrica utilizando redes neurais**. 2007. 133 f. Dissertação (Mestrado) - Universidade Federal de Itajubá, Itajubá, 2007.

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

LIMA, C. M.; LABIDI, S. **Introdução à inteligência artificial**. São Luiz: FAPEMA, 1999. 166 p.

LOPES, M. L. M.; MINUSSI, C. R.; LOTUFO, A. D. P. Electrical load forecasting formulation by a fast neural network. **Engineering Intelligent System**, Leicester, v. 11, n. 1, p. 51-57, 2003.

LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica**. 2005. 169 f. Tese (Doutorado) – Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2005.

MENDES FILHO, E. F.; CARVALHO, A. C. P. L. **Tutorial introdutório sobre redes neurais artificiais**. São Paulo: Departamento de Ciências de Computação e Estatística – USP, 1997. Apostila.

MINUSSI, C. R.; LOTUFO, A. D. P. **Redes neurais** - introdução e principais conceitos. Ilha Solteira: Departamento de Engenharia Elétrica - UNESP, 2008. Apostila.

MINUSSI, C. R. **Lógica nebulosa** – fuzzy logic. Ilha Solteira: Departamento de Engenharia Elétrica – UNESP, 2009. Apostila.

MOGHAN, I; RAHMAN, S. Analysis and evaluation of five short-term load forecasting techniques. **IEEE Transactions on Power Systems**, Piscataway, v. 4, n. 4, p. 1484-1491, 1989.

MORENO, A. L. **Análise da estabilidade transitória via rede neural art-artmap fuzzy euclidiana modificada com treinamento continuado**. 2010. 143 f. Tese (Doutorado) - Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho", – UNESP, Ilha Solteira, 2010.

MORETTIN, P. A.; TOLOI, C. M. C. **Modelos para previsão de séries temporais**. Rio de Janeiro: IMPA, 1981. v. 1-2, p. 356 e 267.

MORETTIN, P. A.; TOLOI, C. M. C. **Previsão de séries temporais**. São Paulo: Atual, 1985.

NOSE-FILHO, K. et al. Previsão de cargas elétricas utilizando diferentes arquiteturas de redes neurais artificiais com o auxílio das toolboxes do MATLAB. In: LATIN-AMERICAN CONGRESS ON ELECTRICITY GENERATION AND TRANSMISSION - CLAGTEE, 8., 2009, Ubatuba. **Proceedings...** São Paulo: UNESP/UNMP/UCA, 2009. p. 85.

PEREIRA, A. L. S. **Utilização de redes neurais artificiais no ajuste de controladores suplementares e dispositivos FACTS STATCOM para melhoria da estabilidade a pequenas perturbações do sistema elétrico de potência**. 2009. 110 f. Tese (Doutorado) – Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2009.

RAN-CHANG, L.; DONG-XIAO, N.; ZHENG-YUAN, J. A study of short-term load forecasting based on ARIMA-ANN; machine learning and cybernetics. **INTERNATIONAL**

CONFERENCE ON MACHINE LEARNING AND CYBEMETICS,3, 2004, Shanghai.**Proceedings...** New York: IEEE, 2004. v. 5, p. 3183-3187.

ROSENBLANT, F. The Perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, Washington, v. 65, n. 6, p. 386-408, 1958.

SCHMIDT, C. A. J.;LIMA, M. A. M.A demanda por energia elétrica no Brasil. **Revista Brasil Econômico**, Campinas, v. 58, n. 1, p. 68-98, 2004.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais: para engenharia e ciências aplicadas.**São Paulo: Artliber, 2010. 399 p.

SILVA, T. A. A. **Previsão de cargas elétricas através de um modelo híbrido de regressão com redes neurais.** 2012. 63 f. Dissertação (Mestrado) – Faculdade de engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho" - UNESP, Ilha Solteira, 2012.

SILVEIRA, M. C. G. **Análise de estabilidade transitória de sistemas elétricos por redes neurais ARTMAP nebulosas modulares.** 2003. 92 f. Tese (Doutorado) - Faculdade de Engenharia,Universidade Estadual Paulista "Julio de Mesquita Filho"– UNESP, Ilha Solteira, 2003.

VIEIRA FILHO, X.et al. **Energy trading: the south American experience.**SIMPÓSIO DE ESPECIALISTAS EM PLANEJAMENTO DA OPERAÇÃO E EXPANSÃO ELÉTRICA - **SEPOPE**,7., 2000, Curitiba. **Simpósio...** Curitiba: [S.n.], 2000.[S.p.].

WASSERMAN, P. D. **Neural computing:theory and practice.** New York: Van Nostrand Reinhold, 1989. 240 p.

WERBOS, P. J. **Beyond regression: new tools for prediction and analysis in the behavioral sciences.** 1974. 102 f. Dissertação (Mestrado)- Harvard University, Harvard, 1974.

WIDROW, B. The original adaptive neural net broom-balancer. In: IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS- ISCS,1987, Philadelphia. **Proceedings...**Philadelphia: IEEE, 1987. p. 351.357.

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

YALCINOZ, T.; EMINOGLU, U. Short term and medium term power distribution load forecasting by neural networks. **Energy Conversion and Management**, Oxford, v. 46, n.9-10, p. 1393-1405, 2005.

ZADEH, L. A. Fuzzy sets. **Information and Control**, New York, v. 8, n. 3, p. 338-353, 1965.

ZADEH, L. A. Fuzzylogic. **Computer**, Los Alamitos, v. 92, n.4, p. 83-93, 1988.

Apêndice A

Conceitos Básicos de Redes Neurais Artificiais

A.1 Neurônio Biológico

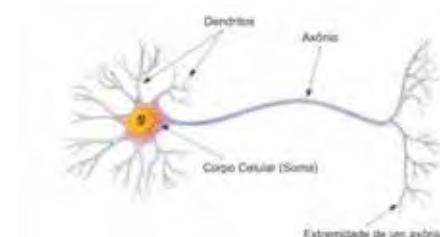
Para que haja melhor entendimento em relação às redes neurais artificiais, torna-se importante o estudo de alguns conceitos básicos em relação ao neurônio biológico.

O cérebro humano é um processador complexo composto por aproximadamente 10 bilhões de neurônios com capacidade de controlar, por exemplo, o reconhecimento de padrões, percepção e controle motor. Esses neurônios são ligados entre si por meio de sinapses transmissoras de estímulos, formando uma gigantesca rede neural que estende o resultado por todo o corpo humano. O desenvolvimento do cérebro humano é adquirido, principalmente, nos dois primeiros anos de vida, porém, continua nos anos subseqüentes por meio de processamento e armazenamento de informações obtidas do meio externo (BRAGA_{et al.}, 2000). Conforme demonstrado na Figura 14, o neurônio biológico conta com os seguintes componentes:

- Dendritos: têm a função de receber estímulos de outros neurônios;
- Corpo celular (também conhecido como soma): responsável por coletar e combinar informações recebidas de outros neurônios; e
- Axônio: constituído de uma fibra tubular, que é responsável por transmitir estímulos para outras células.

Como dito anteriormente, a conexão de um axônio aos dendritos de outros neurônios são feitas pelas sinapses, que possibilitam a propagação de pulsos nervosos, excitando ou inibindo os pulsos, conforme pode ser observado na Figura 14.

Figura 15 - Neurônio Biológico

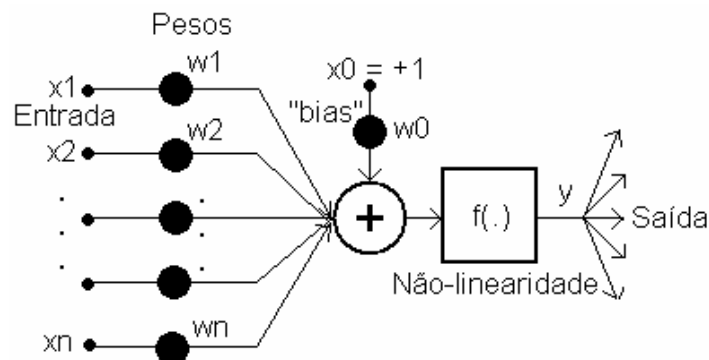


Fonte: Campos, (2010)

A.20 Neurônio Artificial

Neurônios artificiais foram desenvolvidos tomando por base o funcionamento do neurônio biológico e, em 1943, o neurofisiologista Warren McCulloch, do MIT, e o matemático, Walter Pitts, da universidade de Illinois, apresentaram a primeira proposta de neurônio artificial, como demonstrado na Figura 15 (MENDES FILHO; CARVALHO, 1997).

Figura 16 - Modelo do Neurônio de McCulloch-Pitts



Fonte: Minussi; Lotufo, (2008)

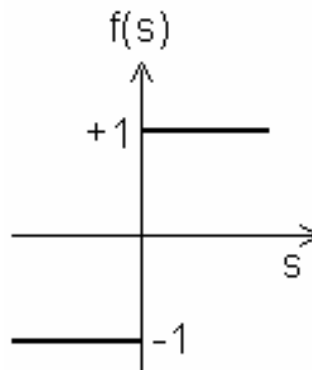
O neurônio artificial é composto de um conjunto de entradas X , representadas por x_1 , x_2 , x_3 , ..., x_n , representando os sinais dentro das sinapses de um neurônio biológico (MENDES FILHO; CARVALHO, 1997). Cada um desses sinais é multiplicado por um peso W , representado por w_1 , w_2 , w_3 , ..., w_n , que será somado no corpo celular, chamado soma e que corresponde ao corpo celular do neurônio biológico. O neurônio de McCulloch-Pitts pode, ainda, conter o peso Bias w_0 , que é alimentado por uma constante $x_0 = +1$, que controla o nível de saída do neurônio.

Existe, ainda, uma função não linear (FN) denominada função de ativação, sendo que, a cada entrada (X), existe um peso sináptico (W) correspondente. Caso a soma ponderada das conexões de X e W for superior ao valor do peso Bias (w_0), é enviado um pulso de saída, que provoca ativação do neurônio. Caso contrário, o neurônio não será ativado. Ao peso Bias é associado um valor de entrada, $x_0=1$, que controla o nível de saída do neurônio (WIDROW; LEHR, 1990; HAYKIN, 1999).

As funções de ativação $f(.)$ são as responsáveis pela ativação do neurônio artificial, valendo trazer alguns exemplos de funções de ativação, nas Figuras 16 a 19 (MINUSSI, 2008):

1- Relé:

Figura 17 - Função Relé



Fonte: Minussi; Lotufo, (20080)

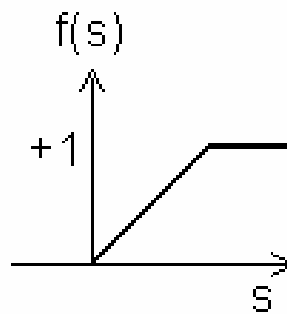
$$f(s) = +1, \text{ se } s \geq 0$$

ou

$$f(s) = -1, \text{ se } s \leq 0$$

2- Lógicathreshold

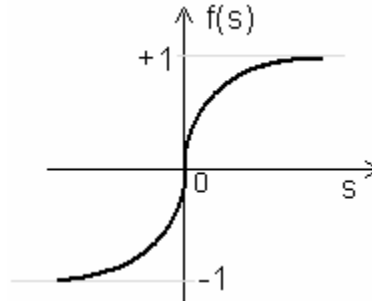
Figura 18 - Função Lógica Threshold



Fonte: Minussi; Lotufo, (2008)

3. Função Sigmóide (1)

Figura 19 - Função Sigmóide (1)



Fonte: Minussi; Lotufo, (2008)

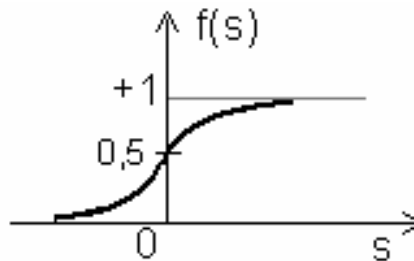
$$f(s) = (1 - e^{-\lambda s}) / (1 + e^{-\lambda s})$$

42

Sendo que, λ = inclinação da curva.

4. Função Sigmóide (2)

Figura 20 - Função Sigmóide (2)



Fonte: Minussi; Lotufo, (2008)

$$f(s) = 1 / (1 + e^{-\lambda s})$$

43

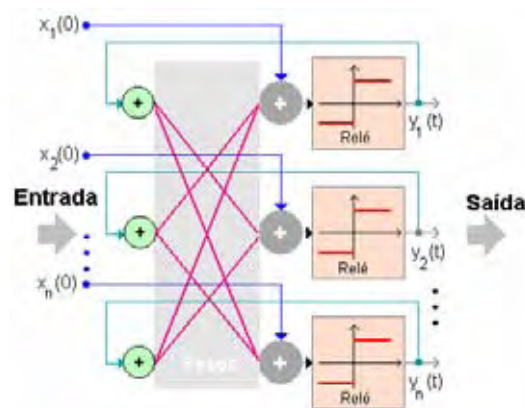
Importante salientar que um neurônio pode contar com várias entradas, porém, com apenas uma saída, a qual, por meio de ramificações, pode servir de entrada de vários outros neurônios. Em vista das várias conexões posteriores, cada uma delas possui um peso que determina sua contribuição da decisão de disparo (MINUSSI, 2008).

A.3 Arquitetura e Treinamento das Redes Neurais Artificiais

Quanto à sua estrutura, as redes neurais artificiais podem ser classificadas em dois tipos: recorrentes (feedforward) e redes não-recorrentes.

As redes neurais recorrentes (Figura 20) possuem laços de realimentação, ou seja, a conexão de saída de uma camada se liga à entrada da mesma camada ou de camadas anteriores, sendo que, neste tipo de arquitetura, as redes possuem memória de longo prazo nos neurônios internos, tendo como exemplo, a rede neural de Hopfield (MINUSSI, 2008).

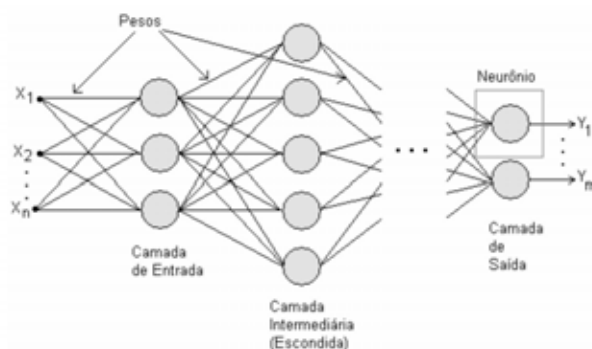
Figura 21 - Rede Neural de Hopfield



Fonte: Minussi, (2008)

Redes neurais não-recorrentes, Figura 21, são dispostas em camadas que recebem sinais exclusivamente de camadas anteriores e, portanto, sem laços de realimentação. Em função disso, não possuem memória de longo prazo e sua saída é determinada somente pela entrada e valores dos pesos (MINUSSI, 2008).

Figura 22 - Rede Neural Não-recorrente

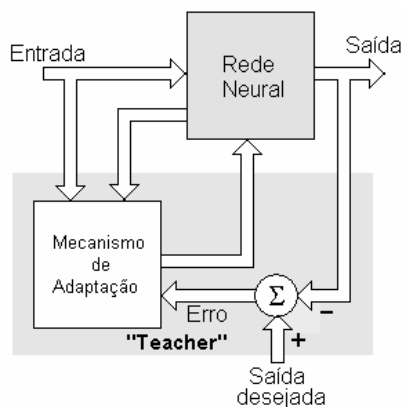


Fonte: Minussi; Lotufo, (2008)

Quanto ao treinamento, podem ser utilizados alguns algoritmos específicos às redes neurais, a saber:

- Treinamento supervisionado: consiste em um método de aprendizagem que considera um vetor de entrada e um vetor alvo, que representa a saída desejada, conhecidos como par treinado, ou seja, possui um tutor ou “professor”, como mostra a Figura 22. Um dos algoritmos mais conhecidos deste tipo de rede é o retropropagação (backpropagation), criado por Werbos, em 1974 (LIMA; LABIDI, 1999).

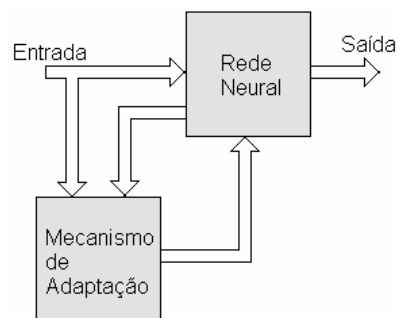
Figura 23 - Treinamento Supervisionado



Fonte: Minussi; Lotufo, (2008)

- Treinamento não-supervisionado: consiste em um método em que não existe saída desejada, treinando-se a rede por meio de excitações ou padrões de entrada organizando, arbitrariamente, padrões em categorias, como ilustra a Figura 23. Caso o padrão de entrada não corresponda a uma classe já existente, gera-se uma nova classe. É normalmente utilizado para redes auto-organizáveis (LIMA; LABIDI, 1999).

Figura 24 - Treinamento Não-supervisionado



Fonte: Minussi; Lotufo, (2008)

Apêndice B - Lógica Fuzzy

A lógica fuzzy é baseada na Teoria dos Conjuntos Fuzzy (ZADEH, 1965), servindo de base para o raciocínio aproximado e não exato. Levando-se em consideração a teoria clássica dos conjuntos, um elemento pertence ou não a um determinado conjunto, ou seja, o valor verdade pode assumir somente dois valores: verdadeiro (1) ou falso (0). Assim, para se verificar se um determinado elemento “x” pertence a um conjunto A, utiliza-se a função característica:

$$\begin{aligned} \mu_A(x): U \rightarrow \{0,1\} \quad \mu_A(x) &= 1 \text{ se } x \in A \\ \mu_A(x) &= 0 \text{ se } x \notin A \end{aligned}$$

Na teoria dos conjuntos convencionais, se um elemento x é dado, este é determinado unicamente se está ou não contido no conjunto A. Ou seja, x está incluído em A ou não. Mas essa definição é muitas vezes inadequada para expressar coisas reais (ABE, 1997).

Zadeh propôs uma caracterização mais ampla, com um número infinito de valores compreendidos no intervalo $[0, 1]$, sendo que, um conjunto fuzzy A no universo U se define pela função de pertinência $\mu_A(x): U \rightarrow [0,1]$, representado por um conjunto de pares ordenados $A = \{\mu_A(x) / x\}, x \rightarrow U$.

A função de pertinência $\mu_A(x)$ define o grau de compatibilidade entre x e o conceito descrito por A (GOMIDE; GUDWIN; TANSCHKEIT, 1995), da seguinte forma:

- $\mu_A(x) = 1$ significa que x é totalmente compatível com A;
- $\mu_A(x) = 0$ significa que x é totalmente incompatível com A;
- $0 < \mu_A(x) < 1$ significa que x é parcialmente compatível com A, com grau $\mu_A(x)$.

O conjunto A da teoria de conjuntos clássica pode ser considerado um conjunto nebuloso, definido por conjunto “crisp”, caso em que, a pertinência é do tipo “tudo ou nada”, “sim ou não”, diferentemente da pertinência gradual dos conjuntos nebulosos, que dá margem a muitas outras opções de resultados.

Considerando as operações de dois conjuntos fuzzy A e B, no universo U , e $\forall x \in U$, temos (GOMIDE; GUDWIN; TANSCHKEIT, 1995):

1- Conjunto vazio:

$$A = \emptyset \text{ se e somente se } \mu_A(x) = 0$$

2- Complemento A':

$$\mu_{A'}(x) = 1 - \mu_A(x)$$

3- Conjuntos iguais:

$$A = B \text{ se e somente se } \mu_A(x) = \mu_B(x)$$

4- A subconjunto de B:

$$A \subset B \text{ se e somente se } \mu_A(x) \leq \mu_B(x)$$

Em operações com conjuntos fuzzy, valem os conceitos de norma triangular, conhecida como norma-t, e co-norma triangular, definida como norma-s. Norma triangular é uma função $\tau : [0,1] \times [0,1] \rightarrow [0,1]$, $\forall x, y, z, w \in [0,1]$, de forma que (GOMIDE; GUDWIN; TANSCHKEIT, 1995):

$$(i) x \tau w \leq y \tau z, \text{ se } x \leq y, w \leq z$$

$$(ii) x \tau y = y \tau x$$

$$(iii) (x \tau y) \tau z = x \tau (y \tau z)$$

$$(iv) x \tau 0 = 0; x \tau 1 = x$$

A co-norma triangular ocorre de tal forma que $\delta : [0,1] \times [0,1] \rightarrow [0,1]$, o que satisfaz as propriedades i, ii e iii supra e

$$(iv) x \delta 0 = x; x \delta 1 = 1$$

Como exemplos da norma-t podem ser citados o mínimo (*min* ou \wedge), bem como o produto algébrico (\cdot). Já os exemplos de co-norma-t, valem ser citados o máximo (*max* ou \vee) e a soma limitada (\oplus) (GOMIDE; GUDWIN; TANSCHKEIT, 1995).

Normas e co-normas triangulares servem para representar a união e a intersecção dos conjuntos fuzzy, sendo que, utilizando os operadores *min* e *max* como norma-t e co-norma-t, obtém-se:

1- União de A e B:

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

2- Intersecção de A e B:

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

Por meio da utilização de união e intersecção, as quais são implementadas por operadores *max* e *min* e complemento, torna-se possível verificar que as seguintes propriedades algébricas de conjuntos ordinários são válidas, também, para os conjuntos fuzzy (GOMIDE; GUDWIN; TANSCHKEIT, 1995), quais sejam:

Involução: $(A')' = A$

Idempotência: $A \cap A = A$ e $A \cup A = A$

Comutatividade: $A \cap B = B \cap A$ e $A \cup B = B \cup A$

Associatividade: $(A \cap B) \cap C = A \cap (B \cap C)$ e $(A \cup B) \cup C = A \cup (B \cup C)$

Distributividade: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Lei Transitiva: se $A \subset B$ e $B \subset C$ então $A \subset C$

Quanto às funções de pertinência dos conjuntos vazio e universo, estas são definidas como 0 e 1, respectivamente, $\forall u \in U$ e, portanto, também são válidas as propriedades a seguir:

$$A \cap \emptyset = \emptyset \text{ e } A \cap U = A$$

$$A \cup \emptyset = A \text{ e } A \cup U = U$$

Propriedades de conjuntos clássicos que não se aplicam a conjuntos fuzzy são:

$$A \cap A' \neq \emptyset \text{ e } A \cup A' \neq U$$

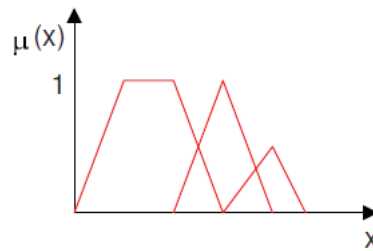
Pelas definições acima, percebe-se que, em algumas ocasiões, pretende-se trabalhar com conjuntos em que as fronteiras não se encontram bem definidas, tornando impossível trabalhar com conjuntos clássicos, em vista da transição entre classes ocorrer de forma suave. Portanto, para se obter os conjuntos fuzzy, bem como suas operações, é necessário generalizar a função característica da lógica clássica para o intervalo $[0,1]$, conforme segue:

$$\mu_A: U \rightarrow [0, 1],$$

considerando valores contínuos e não apenas “pertence ou não pertence”. Ou seja, um elemento x pertencerá a um subconjunto A com grau de pertinência que é um valor do intervalo $[0,1]$ (MORENO, 2010).

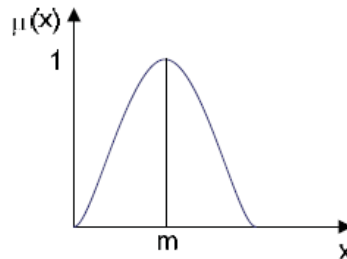
Em caso de conjuntos fuzzy contínuos, estes podem ser representados pela própria função de pertinência. Dentre as funções de pertinência mais utilizadas, temos a linear por partes, representada pela Figura 24, a quadrática e a Gaussiana, conforme Figura 25 (MORENO, 2010), conforme os exemplos:

Figura 25 - Função de Pertinência do Tipo Linear por Partes



Fonte: Moreno, (2010)

Figura 26 - Função de Pertinência do Tipo Gaussiana



Fonte: Moreno, (2010)

Quanto à relação fuzzy, deve ser considerado que, se A_1, A_2, \dots, A_n são conjuntos fuzzy em U_1, U_2, \dots, U_n , são expressos da seguinte forma (GOMIDE; GUDWIN; TANSCHKEIT, 1995):

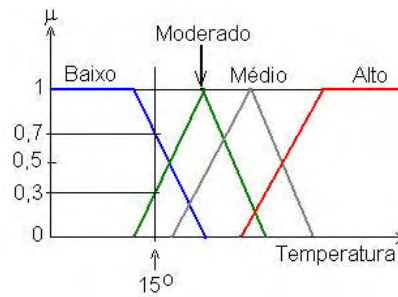
$$R = \{ \mu_R(x_1, \dots, x_n) / (x_1, \dots, x_n) \} \text{ tal que } (x_1, \dots, x_n) \in U_1 \times U_2 \times \dots \times U_n$$

Sejam R e P relações fuzzy, respectivamente em $U \times V$ e $V \times W$, então, a composição de R e P é denotada por $R \circ P$, conforme segue (GOMIDE; GUDWIN; TANSCHKEIT, 1995):

$$R \circ P = [(x,z), \sup_y (\mu_R(x,y) \tau \mu_P(y,z))] \quad x \in U, y \in V, z \in W$$

Variável linguística é o nome dado aos valores dos conjuntos fuzzy e, como exemplo, pode ser citada a temperatura de um determinado processo, com valores como *baixa*, *moderada*, *média*, *alta*, *etc.*, sendo que, tais valores são descritos por meio de conjuntos fuzzy (MINUSSI, 2009), conforme representado na Figura 26.

Figura 27 - Variáveis Linguísticas de Temperatura



Fonte: Minussi, (2009)

Observa-se que:

$$\mu(15^\circ)_{\text{baixo}} = 0,7$$

$$\mu(15^\circ)_{\text{moderado}} = 0,3$$

$$\mu(15^\circ)_{\text{médio}} = 0$$

$$\mu(15^\circ)_{\text{alto}} = 0$$

Uma variável linguística pode ser construída a partir de (GOMIDE; GUDWIN; TANSCHKEIT, 1995):

- Termos primários: como por exemplo, *alto*, *baixo*, *pequeno*, *médio*, *grande* ou *zero*;
- Conectivos lógicos: negação (NÃO), conectivos E e OU, conectivos mascarados (MAS, PORÉM);

- Modificadores: *muito*, *pouco*, *levemente* e *extremamente*;

- Delimitadores: parênteses.

Termos primários são associados a conjuntos fuzzy e são definidos por meio de:

- Funções de pertinência analítica;
- Funções de pertinência lineares por partes (formas triangulares ou trapezoidais);
- Funções de pertinência discretizadas.

A negação NÃO, bem como os conectivos E e OU se definem, respectivamente, em termos de operações de complementação, intersecção e união, enquanto os conectivos mascarados correspondem a E e OU. Já os delimitadores visam a não ocorrência de ambigüidades (GOMIDE; GUDWIN; TANSCHKEIT, 1995).

As variáveis lingüísticas têm como função fornecer uma maneira sistemática de fenômenos complexos ou mal definidos (GOMIDE; GUDWIN; TANSCHKEIT, 1995).

Quanto às proposições, pode-se dizer que, utilizando uma frase do tipo Π é A, sendo Π uma variável lingüística e A um conjunto fuzzy do universo de discurso U de Π , define-se como proposição fuzzy. Diferentes operadores com combinações fuzzy, como por exemplo, os conectivos lógicos E e OU, operador de implicação SE ... ENTÃO, com proposições fuzzy descritas em termos de relações fuzzy e valores determinados de inúmeras maneiras (GOMIDE; GUDWIN; TANSCHKEIT, 1995).

A título de exemplo, pode-se considerar as variáveis lingüísticas x e y , o universo de discurso X e Y , sendo respectivamente, conjuntos fuzzy A e B , os quais são definidos em X e Y , como proposições fuzzy (x é A) e (y é B). Conectando tais proposições por meio do operador OU, obtém-se o seguinte: (x é A) OU (y é B), podendo ser escrito como $((x,y) \text{ é } R_{A \text{ OU } B})$, sendo $R_{A \text{ OU } B}$ uma relação fuzzy em $X \times Y$, definida por uma função do tipo f_{OU} (normalmente uma co-norma-t):

$$R_{A \text{ OU } B} = \{ \mu_R(x,y) / (x,y) \} \text{ tal que } \mu_R(x,y) = f_{OU}(\mu_A(x), \mu_B(y))$$

Se as proposições forem conectadas por meio do operador E, tem-se $((x,y) \text{ é } R_{A \text{ E } B})$, onde:

$$R_{A \text{ E } B} = \{ \mu_R(x,y) / (x,y) \} \text{ tal que } \mu_R(x,y) = f_E(\mu_A(x), \mu_B(y))$$

sendo que, f_E é normalmente uma norma-t.

Existe ainda a declaração condicional fuzzy SE... ENTÃO, que descreve dependência do valor de uma variável lingüística em relação ao valor de outra. Em diversas ocasiões, as declarações condicionais são definidas como regras (lingüísticas) e são expressas por frases da forma SE (x é A) ENTÃO (Y é B), denominada implicação e que se reduz à proposição fuzzy binária $((x,y) \text{ é } R_{A \rightarrow B})$ (GOMIDE; GUDWIN; TANSCHKEIT, 1995), onde:

$$R_{A \rightarrow B} = \{ \mu_R(x,y) / (x,y) \} \text{ tal que } \mu_R(x,y) = f_{\rightarrow}(\mu_A(x), \mu_B(y))$$

sendo f_{\rightarrow} : uma função de implicação.

Se a declaração condicional tiver mais de uma variável *antecedente* (x é A), estas são combinadas pelo conectivo E:

$$\text{SE } ((x_1 \text{ é } A_1) \text{ E } (x_2 \text{ é } A_2) \text{ E } \dots \text{ E } (x_m \text{ é } A_m)) \text{ ENTÃO } (y \text{ é } B),$$

que pode ser reduzido a: $((x_1, x_2, \dots, x_m, y) \text{ é } R)$, tendo como função de pertinência:

$$\mu_R(x_1, x_2, \dots, x_m, y) = f \rightarrow (f_E(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_m}(x_m)), \mu_B(y))$$

Caso haja diversas declarações, estas podem ser combinadas por meio do conectivo OU, da seguinte forma:

$$\text{SE } (x \text{ é } A^1) \text{ ENTÃO } (y \text{ é } B^1)$$

OU

$$\text{SE } (x \text{ é } A^2) \text{ ENTÃO } (y \text{ é } B^2)$$

OU

.

.

$$\text{SE } (x \text{ é } A^n) \text{ ENTÃO } (y \text{ é } B^n)$$

ou seja:

$$((x,y) \text{ é } R^N): ((x,y) \text{ é } R^1) \text{ OU } ((x,y) \text{ é } R^2) \text{ OU } \dots \text{ OU } ((x,y) \text{ é } R^n).$$

Assim, a função de pertinência é:

$$\mu_R^N(x,y) = f_{OU}(\mu_R^1(x,y), \mu_R^2(x,y), \dots, \mu_R^n(x,y))$$

ou seja:

$$\begin{aligned} \mu_R^N(x,y) = f_{OU} [& f \rightarrow (\mu_A^1(x), (\mu_B^1(y)), \\ & f \rightarrow (\mu_A^2(x), (\mu_B^2(y)), \dots, \\ & f \rightarrow (\mu_A^n(x), (\mu_B^n(y)))] \end{aligned}$$

Verifica-se, portanto, que as regras não definidas pelas implicações lógicas SE... ENTÃO, que representa uma relação $R_{A \rightarrow B}$ entre antecedente(s) e conseqüente(s). A função de pertinência μ_R se define pelo operador de implicação $f \rightarrow$, relacionada a um ramo da matemática, conhecido como lógica proposicional.

Em modelos fuzzy, a inferência é definida pelas famílias de regras. Diferentes modelos são resultados de interpretações distintas de conhecimentos contidos nas regras, bem como dos mecanismos de raciocínio adotados (MORENO, 2010).

A inferência se difere pelo tipo de operador utilizado, bem como pelos tipos de antecedentes e conseqüentes, podendo, até mesmo, conter entrada e saída únicas, conhecido como sistema “SINGLE IMPUT/SINGLE OUTPUT” (SISO). Quando ocorrem múltiplas entradas e múltiplas saídas (MIMO), obtém-se a dinâmica do sistema por meio de regras iterativas e a iteração do passo de saída do passo anterior atua como entrada do passo posterior.

Dentre os diversos modelos de inferência, deve-se optar pelo que melhor se adapta ao sistema específico. Um dos mais utilizados, em função de sua simplicidade e facilidade de adaptação é o *Método Mamdani* (MORENO, 2010).

Quando a intenção é obter um resultado clássico, a partir de uma saída fuzzy, utiliza-se métodos de defuzzificação (MORENO, 2010), podendo citar como exemplos mais utilizados:

- Média dos máximos (MM): é um método descontínuo, obtido por meio da média dos valores de saída, com maiores graus de possibilidade, tendo como maior desvantagem a não utilização total do conjunto fuzzy de saída;
- Centro de área (CA): trata-se de um método contínuo, sendo a técnica mais utilizada de defuzzificação, em vista de utilizar totalmente a saída para se obter o valor clássico, tendo como principal desvantagem o custo computacional;
- Método das alturas (MA): é um método contínuo, que utiliza somente os máximos das funções de pertinência. Trata-se de um método simplificado de obtenção do valor clássico e tem como vantagem a redução do custo computacional.