

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”



FACULDADE DE ENGENHARIA

CÂMPUS DE ILHA SOLTEIRA

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

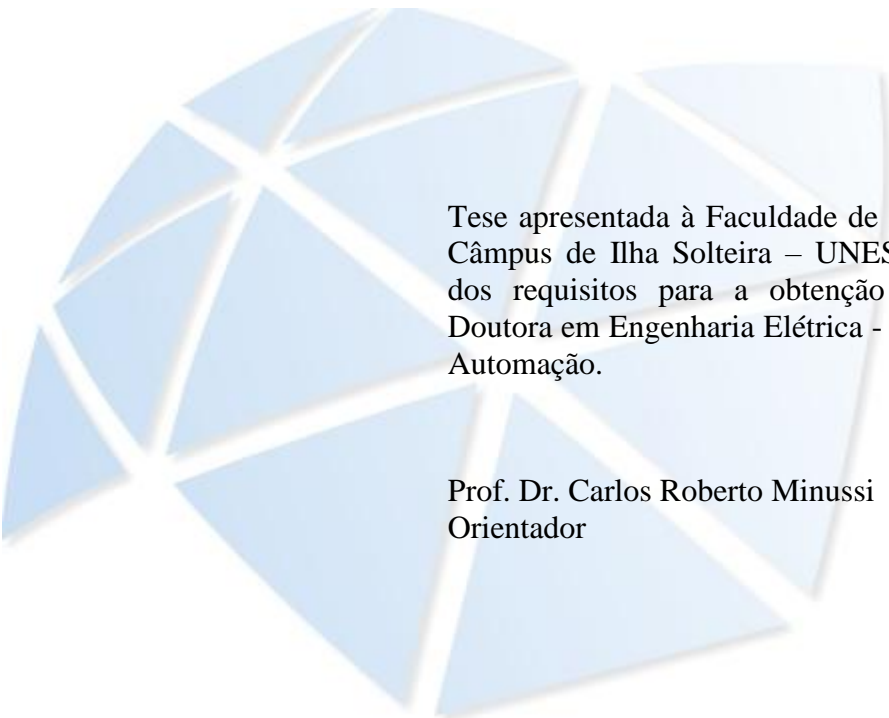
MARCELA ALEXANDRA DA SILVA

**PREVISÃO DE CARGA DE *SMART GRIDS*: NOVA FORMULAÇÃO POR REDES
NEURAIS DA TEORIA DA RESSONÂNCIA ADAPTATIVA**

Ilha Solteira
2020

MARCELA ALEXANDRA DA SILVA

**PREVISÃO DE CARGA DE *SMART GRIDS*: NOVA FORMULAÇÃO POR REDES
NEURAIIS DA TEORIA DA RESSONÂNCIA ADAPTATIVA**



Tese apresentada à Faculdade de Engenharia do
Câmpus de Ilha Solteira – UNESP como parte
dos requisitos para a obtenção do título de
Doutora em Engenharia Elétrica - Especialidade:
Automação.

Prof. Dr. Carlos Roberto Minussi
Orientador

Ilha Solteira
2020

FICHA CATALOGRÁFICA

Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

S586p Silva, Marcela Alexandra da.
Previsão de carga de *Smart Grids*: nova formulação por redes neurais da Teoria da Ressonância Adaptativa / Marcela Alexandra da Silva. -- Ilha Solteira: [s.n.], 2020
70 f. : il.

Tese (doutorado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2020

Orientador: Carlos Roberto Minussi
Inclui bibliografia

1. Sistemas elétricos de potência. 2. Previsão de carga global. 3. Redes Neurais Artificiais. 4. Teoria da Ressonância Adaptativa. 5. Treinamento continuado.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Ilha Solteira

CERTIFICADO DE APROVAÇÃO

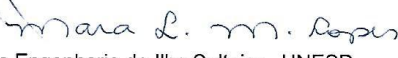
TÍTULO DA TESE: Previsão de Carga de *Smart Grids*: Nova Formulação Por Redes Neurais da Teoria da Ressonância Adaptativa

AUTORA: MARCELA ALEXANDRA DA SILVA

ORIENTADOR: CARLOS ROBERTO MINUSSI

Aprovada como parte das exigências para obtenção do Título de Doutora em ENGENHARIA ELÉTRICA, área: Automação pela Comissão Examinadora:

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

Profª. Drª. MARA LUCIA MARTINS LOPES 
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira - UNESP

Prof. Dr. DILSON AMANCIO ALVES 
Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira - UNESP

Prof. Dr. MAURO DE SOUZA TONELLI NETO 
MacPhail School of Energy - Southern Alberta Institute of Technology - Calgary, AB, Canadá

Prof. Dr. JOSÉ GUILHERME MAGALINI SANTOS DECANINI 
Departamento de Eletrotécnica / Instituto Federal de Educação, Ciência e Tecnologia de São Paulo - IFSP

Ilha Solteira, 11 de março de 2020

DEDICATÓRIA

*Ao meu pai
Wilson
(in memoriam)*

AGRADECIMENTOS

Agradeço a todas as pessoas que me ajudaram a vencer esta etapa e que de alguma forma contribuíram para a realização deste trabalho.

Aos meus pais Wilson (*in memoriam*) e Alzira e minha irmã Daniela, por sempre acreditarem em mim e no meu potencial.

Ao meu esposo Willian, por todo amor, pelo cuidado e carinho, pelo conhecimento compartilhado, por afagar meu coração nos momentos difíceis e por me incentivar a crescer e a ser alguém melhor.

Ao meu orientador Prof. Minussi, por todo conhecimento, pela orientação, por todo apoio e incentivo no decorrer deste trabalho, pela contribuição pessoal e profissional e, sobretudo, por sua dedicação, generosidade e humanidade. Aos professores do grupo de pesquisa, em especial, a Prof. Mara por todo apoio, incentivo, auxílio e amizade.

Aos amigos do laboratório, em especial, a Prof. Thays pelo auxílio, paciência, atenção e companheirismo durante o desenvolvimento deste trabalho.

A Deus, sobretudo, pela vida.

Especiais agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Brasil pelo financiamento desta pesquisa de doutorado – Código de Financiamento 001.

RESUMO

Esta pesquisa de doutorado tem por objetivo o desenvolvimento de um sistema previsor de carga de sistemas de energia elétrica baseado no uso de redes neurais artificiais. A previsão de carga tem sido um assunto abordado há muito tempo. Neste sentido, por certo, pode-se afirmar que a literatura técnica atingiu um alto grau de evolução nesta área do conhecimento. Contudo, observa-se que há, ainda, um longo caminho para explorá-la tendo vista às necessidades de atendimento a um setor que evolui constantemente, em especial, aos atuais e aos futuros sistemas elétricos chamados *smart grids*. Para a realização do planejamento e da operação destes sistemas, várias ações devem ser realizadas: a inserção de modernas tecnologias, uso intenso de dispositivos e de equipamentos eletrônicos, processamento de sinais, inteligência computacional, sistemas de comunicação eficientes, análise e monitoramento das grandezas elétricas, entre outras funções especializadas. Estas ações fazem-se necessárias para tornar exequível o conhecimento do estado do sistema, em regime permanente e, principalmente, em razão da existência de distúrbios elétricos que não são controláveis (manobras acidentais de equipamentos elétricos, acidentes, descargas atmosféricas etc.). Para realizar estas ações, é impreterível o desenvolvimento de metodologias em caráter antecipatório. O principal e indispensável agente antecipatório refere-se à estimação (previsão) de carga futura. A partir do conhecimento de carga futura, pode ser desenvolvida uma série de estudos elétricos (apuração das informações sobre a quantidade de geração necessária para atender a demanda, considerando-se um horizonte predefinido, fluxo de potência, análise de estabilidade, entre outros estudos). Considerando-se que, também, observa-se, na atualidade, um grande desenvolvimento de técnicas para síntese e para a análise de sistemas, em termos de eficiência, precisão e rapidez na obtenção de respostas, é plausível imaginar que o tempo correspondente à execução destas rotinas, ainda em caráter antecipatório, pode se aproximar bastante do tempo real. Assim sendo, nesta pesquisa, dedicar-se-ão os esforços visando contribuir em três ações principais: (1) melhor conhecimento e interpretação das informações contidas na base de dados históricos; (2) explorar a máxima capacidade de realização da rede neural e (3) combinar estes dois objetivos com vistas a proporcionar uma previsão considerando-se um tempo mais próximo do tempo real. Neste sentido, deve-se ressaltar que o tempo futuro, com a máxima aproximação do tempo presente, é bastante crucial, pois à medida que se aproxima do tempo real, perde-se cada vez mais a influência das informações sobre os dados históricos. Neste caso, devem-se contemplar, primordialmente, as medidas advindas do sistema de aquisição de dados. Para a realização deste sistema previsor, usar-se-á uma rede neural da família ART (*Adaptive Resonance Theory*), mais especificamente a arquitetura supervisionada ARTMAP-Fuzzy. Estas redes neurais têm-se mostrado bastante eficientes, rápidas e versáteis à inclusão de inovações em decorrência de serem arquiteturas estáveis e plásticas, as quais são propriedades diferenciais em relação a outras arquiteturas neurais disponíveis na literatura. A plasticidade constitui-se numa propriedade que possibilita a inclusão das medidas fornecidas pelo sistema de aquisição de dados, cuja idealização pode ser implantada por meio da proposição de uma rotina de treinamento continuado. Visando testar a metodologia proposta, apresenta-se uma aplicação tomando-se como exemplo, uma base de dados disponibilizada por uma empresa internacional do setor elétrico.

Palavras-chave: Sistemas elétricos de potência. Previsão de carga global. Redes Neurais Artificiais. Teoria da Ressonância Adaptativa. Treinamento continuado.

ABSTRACT

This research aims to develop a system for load forecasting of electrical power systems based on the use of artificial neural networks. Load forecasting has been a subject studied for a long time. In this sense, of course, it can be said that technical literature has reached a high degree of evolution in this area of knowledge. However, it is observed that there is still a long way to explore it in view of the needs of serving a sector that is constantly evolving, especially to current and future electrical systems called smart grids. To carry out the planning and operation of these systems, several actions must be carried out: the insertion of modern technologies, intense use of electronic devices and equipment, signal processing, computational intelligence, efficient communication systems, analysis and monitoring of electrical quantities, among other specialized functions. These actions are necessary to make the knowledge of the state of the system feasible, in a permanent regime and, mainly, due to the existence of electrical disturbances that are not controllable (accidental maneuvers of electrical equipment, accidents, lightning, etc.). In order to carry out these actions, it is essential to develop methodologies in anticipation. The main and indispensable anticipatory agent refers to the estimation (forecast) of the future load. From the knowledge of the future load, a series of electrical studies can be developed (verification of the information on the amount of generation needed to meet the demand, considering a predefined horizon, power flow, stability analysis, among other studies). Considering that, also, that a great development of techniques for synthesis and systems analysis is observed, in terms of efficiency, precision and speed in obtaining answers, it is plausible to imagine that the time corresponding to the execution of these routines, still in anticipation, can be very close to real time. Therefore, in this research, efforts will be dedicated to contribute to three main actions: (1) better knowledge and interpretation of the information contained in the historical database; (2) explore the maximum realization capacity of the neural network and (3) combine these two objectives in order to provide a forecast considering a time closer to real time. In this sense, it must be emphasized that the future time, with the maximum approximation of the present time, is quite crucial, because as it approaches real time, the influence of information on historical data is increasingly lost. In this case, the measures arising from the data acquisition system must be considered. For the realization of this predictive system, a neural network of the ART family (Adaptive Resonance Theory) will be used, more specifically the supervised architecture Fuzzy-ARTMAP. These neural networks have been shown to be quite efficient, fast and versatile for the inclusion of innovations due to being stable and plastic architectures, which are differential properties in relation to other neural architectures available in the literature. Plasticity is a property that makes it possible to include the measures provided by the data acquisition system, whose idealization can be implemented through the proposal of a continuous training routine. In order to test the proposed methodology, an application is presented taking, for example, a database made available by an international company of the electricity sector.

Keywords: Electric power systems. Global load forecasting. Artificial neural networks, Adaptive Resonance Theory. Continuous learning.

NOTAÇÃO, LISTA DE ABREVIATURAS E SÍMBOLOS

NOTAÇÃO

Variáveis e parâmetros são representados no estilo itálico. Se forem matrizes, ou vetores, o estilo escolhido é o itálico e negrito. Os vetores são apresentados por linhas e não por colunas, como frequentemente usado na notação matemática. Esta notação atende a sugestão dos Autores das redes neurais da família ART (*Adaptive Resonance Theory*) (CARPENTER & GROSSBERG, 1992; GROSSBERG, 2013). Constitui-se numa representação que foi proposta visando tornar os cálculos mais fáceis de compreensão no contexto da rede neural ARTMAP-*Fuzzy*, a qual será o instrumento básico para o desenvolvimento do sistema previsor de carga, que é o objeto desta pesquisa de doutorado.

LISTA DE ABREVIATURAS E SÍMBOLOS

- α : parâmetro de escolha das redes neurais ART-*Fuzzy* e ARTMAP-*Fuzzy*;
- β : taxa de treinamento das redes neurais ART-*Fuzzy* e ARTMAP-*Fuzzy*;
- ρ : parâmetro de vigilância das redes neurais ART-*Fuzzy* e ARTMAP-*Fuzzy*;
- $\rho_{baseline}$: parâmetro de vigilância inicial da rede neural ARTa-*Fuzzy*;
- ε : um número positivo e pequeno;
- M : dimensão do vetor de entrada da rede neural ARTMAP-*Fuzzy*;
- N : dimensão do vetor de saída da rede neural ARTMAP-*Fuzzy*;
- L : quantidade de vetores de entrada das redes neurais ART-*Fuzzy* e ARTMAP-*Fuzzy*;
- $\|I\|_1$: norma-1 vetorial;
- a : vetor de entrada da rede neural ARTMAP-*Fuzzy*;
= $[L T E]$;
- b : vetor de saída da rede neural ARTMAP-*Fuzzy*;
= H ;

- I^a : vetor de entrada da rede neural ARTMAP-*Fuzzy* normalizado e apresentado na codificação complementar;
- I^b : vetor de saída da rede neural ARTMAP-*Fuzzy* normalizado e apresentado na codificação complementar;
- y^a : vetor que compõe os índices dos neurônios ativados da rede neural ARTa-*Fuzzy*;
- \bar{x} : vetor x normalizado aplicando-se a norma-1;
- T_ξ : função de escolha do ξ -ésimo neurônio da rede neural ARTa-*Fuzzy*, ou da rede neural ARTb-*Fuzzy*;
- J : índice do neurônio vencedor da rede neural ARTa-*Fuzzy*;
- K : índice do neurônio vencedor da rede neural ARTb-*Fuzzy*;
- DL : duração do dia com luz solar (número de horas);
- W^a : matriz de pesos do módulo ARTa-*Fuzzy* (entrada) da rede neural ARTMAP-*Fuzzy*;
- W^a_J : J -ésimo vetor da matriz W^a ;
- W^b : matriz de pesos do módulo ARTb (saída) da rede neural ARTMAP-*Fuzzy*;
- W^b_K : K -ésimo vetor da matriz W^b ;
- W^{ab} : matriz de pesos do módulo inter-ART da rede neural ARTMAP-*Fuzzy*;
- W^{ab}_{JK} : J - K elemento da matriz W^{ab} ;
- “*novo*” : refere-se ao vetor atualizado;
- “*velho*” : refere-se ao vetor inicial (anterior à atualização);
- $MAPE$: *Mean Absolute Percent Error*;
- h : tempo na forma discreta (*e.g.*, 1, 5, 10, 15 etc. minutos);
- NA : número de horas (ou frações de hora) da amostragem da série de tempo;
- $y_{atual}(h)$: valor de referência (para fins de comparação) relativo ao instante discreto h ;
- $y_{previsto}(h)$: valor previsto pela rede neural ARTMAP-*Fuzzy* referente ao instante discreto h .
- IS : índice de similaridade;
- Jac : índice de Jaccard;

- sim* : índice de similaridade proposto;
- $A = \|U\|_1$;
- T* : temperatura horária (°C);
- Tmax* : temperatura (°C) máxima do dia;
- Tmin* : temperatura (°C) mínima do dia;
- γ : parâmetro de ajuste para a estimação da temperatura horária;
- L* : vetor cujos componentes são os valores de cargas correspondentes à hora corrente (*h*) que compõem a janela de entrada da rede neural ARTMAP-*Fuzzy*;
- La(h)* : valor de carga correspondente ao tempo discreto *h* do mesmo dia da semana anterior;
- T* : vetor de temperatura (°C);
- E* : vetor contendo as informações exógenas (dia da semana, estação do ano, dia nublado, dia festivo etc.);
- H* : carga correspondente ao horário (*h+1*) correspondente a uma hora à frente: $H \in \mathfrak{R}^N (N = 1)$;
- d* : subvetor referente ao dia da semana na forma binária com 3 *bits* que são suficientes para a representação semanal de domingo a sábado;
- m* : subvetor relativo ao mês do ano com 4 *bits*, sendo suficientes para representar os 12 meses do ano;
- tem* : subvetor visando abrigar as informações sobre a situação atmosférica do dia;
- ad* : subvetor destinado a receber, na forma da codificação binária, outras informações relevantes, cujo número de *bits* será definido em função do número de informações de interesse do usuário;
- NA* : número de horas (ou fração de hora) da amostragem da série de tempo;
- DN* : detecção de novidade;
- ρ_{Acq} : parâmetro de vigilância limitante para fins de descarte de padrões referentes à fase de treinamento continuado;
- I_{Acq}^a : vetor da base de dados advindo do sistema de aquisição;

$\rho_{\alpha_{CL}}$: parâmetro de vigilância da rede neural ARTa-*Fuzzy* referente à fase do treinamento continuado;

Φ : número de ouro (áureo);

= 1,61803399, tal que $\Phi \times \Phi^{-1} = 1$; $\Phi^{-1} = 0,61803399$.

LISTA DE FIGURAS

Figura 1.	Ilustração sobre a formação de classes.....	39
Figura 2.	Estágios temporais da previsão de carga.....	49
Figura 3.	Fluxograma referente à execução da fase de treinamento, treinamento continuado e a previsão de carga global.....	53
Figura 4.	Curvas de carga resultantes da previsão via rede neural ARTMAP- <i>Fuzzy</i> com treinamento realizado por um período de 9 meses SEM, ou COM a atuação do treinamento continuado: resultado relativo a 96 horas (4 dias) à frente.....	61
Figura 5.	Curvas de carga resultantes da previsão via rede neural ARTMAP- <i>Fuzzy</i> com treinamento realizado por um período de 1 mês SEM, ou COM a atuação do treinamento continuado: resultado relativo a 96 horas (4 dias) à frente.....	62

LISTA DE TABELAS

Tabela 1.	Desempenho da rede neural ARTMAP- <i>Fuzzy</i> para distintos períodos de treinamento SEM o treinamento continuado.....	59
Tabela 2.	Resultados da previsão de carga de até 4 dias (96 horas) à frente para um período de treinamento de 9 meses SEM, ou COM a execução do treinamento continuado (TC).....	60
Tabela 3.	Resultados da previsão de carga de até 4 dias (96 horas) à frente para um período de treinamento de 1 mês SEM, ou COM a execução do treinamento continuado (TC).....	60

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO.....	15
1.1. INTRODUÇÃO E JUSTIFICATIVA	15
1.2. ORGANIZAÇÃO DO TEXTO	17
CAPÍTULO 2 – PESQUISAS CORRELATAS	19
CAPÍTULO 3 – PREVISÃO DE CARGA ELÉTRICA	21
3.1. INTRODUÇÃO	21
3.2. FORMULAÇÃO DA PREVISÃO DE CARGA POR REDES NEURAS	23
CAPÍTULO 4 – REDES NEURAS DA FAMÍLIA ART	26
4.1. INTRODUÇÃO	26
4.2. REDE NEURAL ARTMAP- <i>FUZZY</i>	28
4.3. ALGORITMO PARA A EMULAÇÃO DA REDE NEURAL ART- <i>FUZZY</i>	30
4.4. ALGORITMO PARA A EMULAÇÃO DA REDE NEURAL ARTMAP- <i>FUZZY</i>	34
4.5. MELHORIAS DO DESEMPENHO DE REDE NEURAL ARTMAP- <i>FUZZY</i>	34
CAPÍTULO 5 – PREVISÃO DE CARGA GLOBAL	41
5.1. INTRODUÇÃO	41
5.2. COMPOSIÇÃO DOS VETORES DE ENTRADA E DE SAÍDA DA REDE NEURAL ARTMAP- <i>FUZZY</i>	41
5.3. FORMULAÇÃO DA PREVISÃO DE CARGA BASEADA NO	46

TREINAMENTO CONTINUADO (<i>CONTINUOUS LEARNING</i>)	
5.4. FLUXOGRAMA REFERENTE À PREVISÃO DE CARGA	52
5.5. INFERÊNCIA DA QUALIDADE DE PREVISÃO VIA REDE NEURAL ARTMAP- <i>FUZZY</i>	55
CAPÍTULO 6 – RESULTADOS E DISCUSSÕES	57
6.1. INTRODUÇÃO	57
6.2. RESULTADOS E DISCUSSÕES	57
CAPÍTULO 7 – CONCLUSÃO E SUGESTÕES PARA TRABALHOS FUTUROS	64
7.1. CONCLUSÃO	64
7.2. SUGESTÕES PARA TRABALHOS FUTUROS	65
REFERÊNCIAS.....	67

CAPÍTULO 1 – INTRODUÇÃO

1.1. INTRODUÇÃO E JUSTIFICATIVA

Para que um sistema elétrico de potência possa atender plenamente a demanda (carga) de energia elétrica aos seus consumidores, com continuidade e qualidade (tensões, frequência e forma de ondas adequadas), deve-se realizar, permanentemente, uma série de estudos, análises e a adoção de ações antecipatórias contra perturbações, as quais são consideradas de ocorrência endêmica (estarão presentes, independentemente das ações humanas). Para que isto aconteça, deverá ser prevista, com a máxima precisão possível, a carga elétrica demandada. A previsão de carga pode, de forma simplista, ser agrupada em duas categorias: global e multinodal. Esta classificação atende utilidades diferenciadas. A previsão global (GROSS; GALIANA, 1987; O'DONOVAN, 1983) compreende o somatório de todo o consumo atrelado no sistema elétrico. O conhecimento da previsão global se faz necessário com vistas a fornecer informação aos agentes (planejamento e operação) do sistema para proporcionar os estudos e o planejamento para cumprir o atendimento da demanda presente e futura baseada nos recursos próprios de geração e/ou compra de energia nos mercados pertinentes. Com relação à previsão de carga multinodal, há a necessidade de saber o consumo em todas as barras do sistema, *sine qua non*, não se pode realizar as análises pretendidas (despacho da geração, cálculo do fluxo de potência, análise de contingências estática e dinâmica entre outras rotinas importantes). Para que estas previsões sejam exequíveis, necessita-se que a sua execução produza resultados com precisão compatível com as exigências do setor elétrico e com a máxima rapidez de processamento. Em Altran (2010), Nose-Filho *et al.* (2011), Abreu *et al.* (2018) e Amorim *et al.* (2020) foram propostas metodologias para a realização da previsão multinodal a partir da previsão global. Trata-se de concepções visando à repartição de carga global nos vários barramentos do sistema. Uma vez finalizada a previsão global, são previstas as curvas de carga contemplando as diversas barras de interesse.

Assim sendo, nesta pesquisa de doutorado, pretende-se desenvolver um sistema predictor de carga global com dois objetivos principais: (1) disponibilizar um sistema predictor eficiente em termos de precisão, rapidez e confiabilidade e (2) ser compatível visando à

realização da previsão multinodal (não abordada nesta pesquisa). Procurar-se-á explorar vários pontos não devidamente tratados na previsão de carga, tais como: estudo da base de dados históricos, melhor definição da composição dos dados de entrada, tamanho e intervalo de tempo a ser considerado para fins da realização da fase de treinamento a partir da base de dados, método para a realização da previsão, pertinência do uso do treinamento de forma incremental, entre outros elementos que possam influenciar na qualidade da previsão de carga.

A previsão de carga global tem sido realizada, há bastante tempo, baseada no uso de métodos clássicos (O'DONOVAN, 1983), como o algoritmo ARIMA de Box-Jenkins (BOX & JENKINS, 1976), séries temporais (NGUYEN; HANSEN, 2017), filtro de Kalman, regressão linear múltipla ou simples, alisamento exponencial, estimação de estado, métodos probabilísticos (HAUPT *et al.*, 2019) entre outros. Porém, mais recentemente, as redes neurais artificiais (HAYKIN, 2008) (denominada nesta pesquisa, por simplicidade de notação, apenas como “redes neurais”), ou algumas estruturas híbridas que envolvem redes neurais, lógica *fuzzy* (ZADEH, 1965) algoritmos genéticos (GOLDBERG, 1989), métodos heurísticos de programação matemática etc. têm tido seu uso intensificado para a resolução de problemas da previsão de cargas, em especial quando se trata de carga global.

Nesta pesquisa, propõe-se o desenvolvimento de uma metodologia para a previsão global baseada em uma rede neural da família ART (*Adaptive Resonance Theory*) (CARPENTER; GROSSBERG, 1992), mais especificamente, uma rede neural ARTMAP-*Fuzzy*. O emprego da teoria dos conjuntos *fuzzy* (ZADEH, 1965) confere, às redes neurais da família ART, a capacidade de processarem informações analógicas, binárias e, também, combinações dessas informações (analógicas e binárias). Considerando a facilidade de trabalhar com estes tipos de informações, podem-se empregar, também, outros tipos de informações mais genéricas com o propósito de aperfeiçoar o desempenho da rede neural, por exemplo, a inserção do conceito da fusão de dados (BARNETT, 2008). Esta rede neural foi proposta, em 1992, por Carpenter e Grossberg (1992) e tem sido intensamente empregada para a resolução de um grande número de problemas em diversas áreas do conhecimento humano. É uma rede neural moderna, cuja grande atratividade advém, como consequência, da versatilidade e da habilidade destacadas em razão dos atributos plenos de estabilidade e de plasticidade. A estabilidade refere-se à propriedade de garantir sempre a obtenção de uma solução e a plasticidade confere à rede neural a capacidade de agregar novos conhecimentos sem destruir o conhecimento previamente armazenado nas matrizes de pesos (treinamento

incremental). Com estas características, podem-se propor soluções confiáveis e de alta velocidade de processamento. Deste modo, o sistema neural proposto nesta pesquisa (previsão de carga global) pode ser implementado, de forma segura, com tempo de processamento da fase de treinamento bastante rápido, além de prover confiabilidade. A precisão, evidentemente, será uma consequência. As redes neurais da família ART são bastante competitivas em comparação ao uso de outras redes neurais disponíveis na literatura, *e.g.*, rede neural *feedforward* com treinamento via técnica retropropagação (*backpropagation*) (WERBOS, 1974), ou Levenberg-Marquardt (MARQUARDT, 1963; SAINI; SONI, 2002), entre outras redes neurais que, via de regra, não são incrementais e apresentam problemas de convergência (a convergência não é garantida).

1.2 ORGANIZAÇÃO DO TEXTO

Este texto está assim organizado. No Capítulo 1 são enfocados os principais conceitos, particularidades, breve histórico e a justificativa desta pesquisa no âmbito da previsão de carga em sistemas de energia elétrica. Compreende-se de um preâmbulo envolvendo dois principais elementos tratados nesta pesquisa de doutorado que são: (1) o problema da previsão de carga e (2) o sistema predictor proposto. No Capítulo 2, tratar-se-á da análise sobre as principais perspectivas no trato sobre a previsão de carga em sistemas elétricos de potência. Apresentam-se o problema da previsão de carga no Capítulo 3 e a formulação da previsão de carga por redes neurais dando destaque à justificativa da rede neural empregada (ARTMAP-*Fuzzy*). No Capítulo 4, abordam-se as redes neurais da família ART (*Adaptive Resonance Theory*), mais especificamente os algoritmos das redes neurais ART-*Fuzzy* e ARTMAP-*Fuzzy*. São apresentadas, também, algumas sugestões importantes visando à melhoria do desempenho da rede neural ARTMAP-*Fuzzy*. No Capítulo 5, apresenta-se a previsão de carga global, formulada via a rede neural ARTMAP-*Fuzzy*, com o principal enfoque ao projeto das entradas e das saídas para fins da realização das fases de treinamento, validação e do diagnóstico (previsão), bem como a proposição sobre a execução do treinamento continuado. Inclui-se, ainda, a forma tal como será avaliada a precisão dos resultados. No Capítulo 6, são apresentados os resultados da aplicação da previsão de carga considerando-se uma base de dados conhecida na literatura especializada. Nos Capítulos 7 e 8 são apresentadas a conclusão e as sugestões para futuras pesquisas, assim como a relação de referências científicas de

balizamento desta pesquisa de doutorado, respectivamente.

CAPÍTULO 2 – PESQUISAS CORRELATAS

A previsão de carga em sistemas elétricos de potência compreende uma vasta oferta de propostas na literatura especializada. Diante desta realidade, obviamente, há muita dificuldade de se estabelecer a análise crítica. Menor, ainda, se for pensado em algo como o estado da arte. Em vez disto, nesta seção, serão discutidas as principais formulações em termos de metodologias empregadas. Neste ambiente, observa-se, em grande parte, até 30 anos passados, a tendência recaía no uso de metodologias clássicas muito bem exploradas no contexto de Box e Jenkins (BOX & JENKINS, 1976). A partir daquela data, começaram a ser investigadas as aplicações baseadas nas técnicas de inteligência artificial (BENDER, 1996), em especial, as redes neurais como as redes de Hopfield (HOPFIELD, 1982; BIANCHI *et al.*, 2017), Kohonen (KOHONEN, 2000; MARTÍN-MERINO; ROMÁN, 2006), *feedforward* com treinamento realizado via uso do algoritmo retropropagação (*backpropagation*) de Werbos (WERBOS, 1974). A grande maioria das metodologias para a previsão de carga segue uma forte tendência do uso do algoritmo retropropagação para o seu treinamento. Há tantas outras formulações similares à técnica de treinamento retropropagação. Elas são formulações alternativas que usam concepções semelhantes, também, baseadas no uso do cálculo de derivada, *e.g.*, o método de Levenberg e Marquardt (MARQUARDT, 1963; SAINI & SONI, 2002). Essas redes neurais, frequentemente, têm se caracterizado como técnicas que apresentam dificuldades de convergência que podem consumir muito tempo de processamento, inclusive não convergindo por causa do fenômeno de paralisia e a convergência com predileção para pontos mínimos locais. Este fenômeno pode ocorrer em especial quando se trabalha com bases de dados extensas e, também, quando a dimensão da entrada é expressiva. Toda esta modalidade de metodologias, igualmente, não é dotada de capacidade de realização do treinamento de forma incremental o que, desta forma, oferece recursos limitados para a obtenção de soluções para uma boa quantidade de problemas do conhecimento humano. Redes neurais incrementais, grosso modo, emulam a capacidade de aprender continuamente como é observado na intelectualidade humana.

Visando aumentar o desempenho destas propostas, foi sugerido o desenvolvimento de metodologias híbridas usando-se, por exemplo, as redes neurais combinadas com métodos baseados na lógica *fuzzy* (ZADEH, 1965), *soft computing* (KHAN;

ABRAHAN, 2004), transformada *wavelet* (DAUBECHIES, 1992), algoritmos genéticos (GOLDBERG, 1989), métodos de otimização, como enxame de partículas (*particle swarm*) (EBERHART; KENNEDY, 1995), sistemas imunológicos artificiais (DASGUPTA, 1992), entre outras técnicas. Mais recentemente, surgiram novas metodologias como o *deep learning* (GOODFELLOW *et al.*, 2016; HE, 2017), entre outras formulações que prometem revolucionar no contexto das aplicações, principalmente na área de processamento de sinais, de imagens etc.

Outras categorias de redes neurais compreendem a capacidade de aprender de modo incremental. Nesta categoria, podem-se destacar as redes neurais da família ART (ou ART descendente) (CARPENTER; GROSSBERG, 1992) que – aliada à capacidade de receber inovações importantes – possibilitam a realização de processos complexos como é o caso da previsão de carga (objeto desta pesquisa de doutorado). Algumas formulações foram abordadas para a previsão de carga global (LOPES, 2005; LOPES, *et al.*, 2005), assim como para a previsão de carga multinodal (ALTRAN 2010; NOSE-FILHO *et al.*, 2011, ABREU *et al.*, 2018; AMORIM *et al.*, 2020). Diferentemente, nesta pesquisa, usar-se-á uma rede neural ARTMAP-Fuzzy com pretensões de oferecer melhorias, novos resultados, novas interpretações sobre a possibilidade de novas realizações e sugerir inovações, como serão apresentados nos capítulos subsequentes.

CAPÍTULO 3 - PREVISÃO DE CARGA ELÉTRICA

3.1. INTRODUÇÃO

Os sistemas elétricos de potência podem ser considerados como um dos maiores e mais complexos sistemas concebidos pela humanidade. Esses sistemas contemplam interligações entre unidades geradoras, sistemas de transmissão e os sistemas de distribuição envolvendo países (caso do sistema europeu), grandes países como o Brasil. Fazer funcionar esses sistemas, constitui-se um grande desafio em razão da existência de alto grau de não linearidades nos modelos dos seus componentes, principalmente em relação às unidades geradoras síncronas. Os modelos matemáticos que descrevem o comportamento dinâmico desses sistemas são representados por conjuntos de milhares de equações algébricas e diferenciais não lineares altamente interligadas entre si. A dificuldade é aumentada tendo em vista as incertezas decorrentes de cargas elétricas. Os consumidores, que estão conectados ao sistema elétrico, consomem energia elétrica de forma autônoma. Caberá – aos administradores e operadores do sistema – buscar o adequado planejamento e a operação visando acolher qualquer que seja o montante do consumo. Para operar o sistema elétrico adequadamente, devem-se realizar complexos estudos com o propósito de garantir o funcionamento frente a um grande número de falhas que estão sujeitas a ocorrer em função de manobras, mau funcionamento de equipamentos elétricos, ações intempéries (vendavais, descargas atmosféricas etc.) entre outros agentes.

Para enfrentar todas essas adversidades, os sistemas elétricos de potência têm recebido constantes modernizações chegando à atual conjuntura com a implantação dos sistemas chamados *smart grids* (BUCHHOLZ; TYCZYNSKI, 2020). É uma inovação em que são combinadas várias tecnologias como o uso de dispositivos elétricos e eletrônicos de alta velocidade, aquisição e processamento de sinais, medições, inteligência artificial, bem como outras tecnologias de ponta.

Deste modo, percebe-se que conhecer o consumo de energia que o sistema necessita atender é uma tarefa essencial, viz., o consumo de energia elétrica é causa e não consequência. Sem o seu conhecimento, não há a possibilidade de planejar e operar o sistema.

Conhecer o consumo deve ser entendido como uma ação de previsão. Este fato decorre em razão de ser necessária a realização de uma série de estudos (fluxo de potência, estabilidade estática, transitória e dinâmica etc.) com vistas a superar as possíveis contingências que são de natureza involuntária quase impossível de evitá-las. O que se pode fazer é contrapor, ou, pelo menos, mitigar os seus efeitos. Para o funcionamento do sistema elétrico, tais estudos devem ser realizados de modo antecipatório. A análise em tempo real não se constitui em uma ação eficiente, pois ao concluí-la os possíveis eventos já produziram as suas consequências. O efeito de um blecaute somente poderá ser eliminado, ou minorado se medidas forem implementadas com antecipação (ROSEN, 1985). A partir do momento que se constata um problema importante de instabilidade do sistema podendo evoluir para um blecaute, nada se pode fazer para a sua restauração antes de serem consumados os seus efeitos. Desta forma, a palavra-chave, neste contexto, é o fator antecipatório. A antecipação é um fator crucial altamente dependente do conhecimento da demanda elétrica, o qual é nominado na literatura especializada como previsão de carga (*load forecasting*, no idioma inglês) (O'DONOVAN, 1983).

Em se tratando da previsão de carga, pode-se dividi-la em dois conceitos básicos: (1) a previsão global (O'DONOVAN, 1983); e (2) a previsão multinodal (2). A previsão global pode ser classificada por horizontes de previsão: curtos, médios, ou longos prazos. Os horizontes de curto ou de curtíssimo prazo estão associados ao ambiente da operação, enquanto que a previsão de médio prazo e de longo prazo está voltada às necessidades oriundas à conjuntura do planejamento elétrico. A previsão global corresponde às informações que envolvem os estudos referentes à provisão de geração visando suprir a demanda total do sistema. Deve-se considerar a capacidade instalada da empresa geradora, assim como de outros fornecedores por meio da compra de energia. No caso da previsão multinodal (ALTRAN, 2010), o foco é a previsão (estimativa) de carga em cada barra do sistema, que é fundamental para a determinação do estado do sistema e todas as demais conjunções associadas. A obtenção do estado do sistema constitui-se nas condições imprescindíveis para o estudo da operação do sistema, *e.g.*, a análise de contingências sendo os requisitos básicos para garantir minimamente o funcionamento do sistema com continuidade e qualidade (grandezas elétricas sendo atendidas de acordo com os padrões exigidos pelo setor elétrico). Por conseguinte, estes dois enfoques são de natureza e de aplicações distintas. Contudo, de acordo com alguns resultados apurados recentemente (ABREU *et al.*, 2018; AMORIM *et al.* 2020), a formulação da previsão multinodal pode ser

obtida a partir do conhecimento final da previsão global. Para se conseguir uma previsão multinodal adequada, faz-se necessário primar, também, na obtenção da previsão global.

Neste sentido, esta pesquisa tem como objeto o desenvolvimento de um sistema previsor global, que servirá como um aporte para o planejamento de médio prazo, de longo prazo e a operação, assim como para a obtenção de um plano de previsão multinodal. Busca-se, nesta pesquisa, propor uma metodologia rápida, confiável e precisa que atenda os requisitos exigidos pelo setor elétrico. Essa metodologia será desenvolvida mediante o uso das redes neurais, em consonância com a exposição apresentada no Capítulo 3, e baseada na realização de várias experiências – considerando uma base de dados históricos do consumo – que deverão ser realizadas com vistas à obtenção do máximo desempenho, em especial, a precisão.

3.2. FORMULAÇÃO DA PREVISÃO DE CARGA POR REDES NEURAIAS

As redes neurais artificiais (rotuladas, nesta pesquisa, por simplicidade de notação, por apenas “rede neurais”) têm sido usadas intensivamente nos últimos 30 anos para a resolução de um grande número de problemas em variadas áreas do conhecimento humano (medicina, engenharias, sistemas agrícolas, biologia, economia, entre outras). Assim sendo, deliberadamente, dispensa-se, nesta pesquisa, a apresentação de definições etc. da teoria sobre redes neurais. Havendo a necessidade, serão descritas e justificadas as particularidades não usuais na literatura. Esta opção está sendo tomada em razão de tornar o texto mais “enxuto”, no entender da autora, não produzindo perdas importantes na inteligibilidade desta proposta de pesquisa.

As metodologias para a realização da previsão de carga global são classificadas, essencialmente, como clássicas e baseadas na inteligência artificial. Na classificação clássica há um grande conjunto de técnicas empregadas, em destaque, as séries temporais (ULLAH, 2014), ARIMA (BOX; JENKINS, 1976), alisamento exponencial entre outras. Mais recentemente tem sido explorada a aplicação de técnicas inteligentes, *e.g.*, sistemas imunológicos (DASGUPTA, 1997), Redes Neurais Artificiais (RNA) (Haykin, 2008), lógicas *fuzzy* (ZADEH, 1965), *rough sets* (PAWLAK, 1982), sendo usadas individualmente ou composições destas técnicas. Isto tem sido feito com o propósito de obter a máxima eficiência (redução do tempo de processamento, precisão e confiabilidade das respostas).

Nesta pesquisa, será proposta uma metodologia baseada nas RNA. As aplicações iniciais, para a previsão de carga via uso de RNA, tem sido intensificada usando-se a arquitetura *feedforward* com treinamento realizado via algoritmo retropropagação (*backpropagation*) (WERBOS, 1974) e alternativas similares baseadas no cálculo de derivadas, como, por exemplo, o algoritmo Levenberg-Marquardt (MARQUARDT, 1963; SAINI & SONI, 2002). Sabe-se que, na literatura especializada, os algoritmos que empregam o cálculo de derivadas apresentam grande dificuldade de convergência por causa de problemas como paralisia e convergência, frequentemente, na direção de pontos mínimos locais necessitando, habitualmente, de um elevado tempo de processamento (convergência lenta ou inexistência de convergência) e baixa confiabilidade. Estes problemas, frequentemente, tem tornado a aplicação destes algoritmos bastante limitada.

Visando buscar novas configurações e recursos surgiram algumas metodologias alternativas de RNA com propostas de serem mais competitivas no contexto da previsão de carga. Neste cenário, há várias metodologias para se testar a eficiência usando-se as RNA em problemas de previsão de carga de sistemas elétricos de potência ainda não suficientemente explorados. Por exemplo, a computação em reservatório (*reservoir computing*) (MAASS, *et al.*, 2002), e RNA da família ART (*Adaptive Resonance Theory*) (CARPENTER; GROSSBERG, 1992). Nesta família de RNA há a disposição de um conjunto de alternativas para a resolução de diversos tipos de problemas existentes em várias áreas do conhecimento humano (medicina, engenharias, economia etc.). Dentre essas áreas de conhecimento, destaca-se a previsão de carga (O'DONOVAN, 1983).

A previsão de carga pode ser idealizada via a extração do conhecimento de bases de dados históricos. As informações passadas podem ser usadas visando conhecer (estimar) o comportamento de carga no futuro. O futuro, nesta pesquisa, entende-se como o tempo transcorrido, em média, nas 24 horas à frente da curva de carga. Este período pode ser discriminado em intervalos de 5, 10, 30 minutos ou 1 hora, dependendo do tipo de problema a ser abordado na operação do sistema, ou no planejamento da expansão.

Para realizar a extração do conhecimento, adota-se uma parte da base de dados destinada à fase de treinamento e a outra parte para testar a qualidade da previsão de carga tratada, aqui, como previsão, análise, ou diagnóstico. A qualidade do treinamento é dependente de vários fatores que serão discutidos à frente. Porém, os principais parâmetros são a quantidade de informações reservadas à fase de treinamento: dados exógenos,

composição da janela entre outros fatores. A janela corresponde a um conjunto de dados passados de carga, incluindo o valor de carga do tempo corrente que é a entrada do sistema previsor. A carga no instante posterior representa a saída do referido sistema previsor. Assim, a extração do conhecimento da base de dados históricos é realizada baseada nos estímulos de entrada versus saída procedendo-se o deslocamento sequencial da janela desde o tempo inicial até o tempo final da base de dados. Uma vez realizada esta etapa, na sequência, pode-se testar a qualidade das respostas mediante o uso de uma métrica de inferência, por exemplo, o MAPE (*Mean Absolute Percentage Error*) (PARK *et al.*, 1991). Esta métrica será apresentada e discutida adiante. Se o MAPE atender os requisitos exigidos pelo setor elétrico (*e.g.*, igual ou inferior a 5% de acordo com a legislação brasileira), entende-se que o sistema previsor estará apto a realizar a previsão em um horizonte adequado, por exemplo, 24 horas à frente.

A qualidade deste sistema previsor é dependente de uma boa escolha dos componentes da janela e, também, da rede neural usada. Assim sendo, nesta pesquisa dar-se-á especial atenção ao problema associado à escolha da composição da janela, do tamanho da base de dados históricos mais adequados para se atingir os objetivos de cada tipo de aplicação nos sistemas elétricos de potência e, não menos importante em relação à escolha e definição da arquitetura da rede neural usada. Também, para se maximizar o desempenho do sistema previsão, deverá haver a preocupação de propor formas eficientes para a preparação da base de dados (pré-processamento) e o pós-processamento da previsão bruta gerada pelo sistema previsor. O pré-processamento constitui-se de um procedimento de inferência sobre a qualidade das informações disponibilizadas. Pode haver a falta de informações, as quais devem ser repostas de alguma forma e, ainda, podem ocorrer informações imprecisas que necessitam ser corrigidas. O pós-processamento constitui-se no uso de um algoritmo inteligente visando melhorar a previsão.

CAPÍTULO 4 - REDES NEURAI DA FAMÍLIA ART

4.1. INTRODUÇÃO

As redes neurais artificiais, nos últimos 30 anos, têm sido um instrumento e alternativas importantes para a resolução de um grande número de problemas encontrados em várias áreas do conhecimento humano (engenharias, sistemas agrícolas, saúde, economia, robótica etc.). Várias propostas de métodos têm sido apresentadas na literatura, em destaque, as redes neurais de Hopfield (HOPFIELD, 1982), *feedforward* com treinamento realizado via uso do algoritmo retropropagação (WERBOS, 1974), rede de Kohonen (KOHONEN, 2000), redes neurais da família ART (CARPENTER; GROSSBERG, 1992). Há outras propostas, porém, com menos intensidade de uso.

Neste contexto, várias aplicações praticadas no Laboratório SINTEL, vinculado ao Departamento de Engenharia Elétrica da UNESP do Câmpus de Ilha Solteira, foram implementadas baseadas na arquitetura ART descendente (ART-*Fuzzy*, ARTMAP-*Fuzzy*, ARTMAP-Euclidiana) (CARPENTER; GROSSBERG, 1992) em razão do alto desempenho que tem sido observado, tais como:

- (1) Trata-se de concepções diferenciais, em relação à literatura, que são dotadas da plasticidade. O conhecimento é extraído de forma incremental, não havendo a necessidade de ser reinicializado o treinamento toda a vez que novos padrões (informações) são disponibilizados. Esta propriedade será bastante útil no desenvolvimento de um sistema previsor, nesta pesquisa, naturalmente incremental, cujo resultado esperado será a obtenção de previsões, a princípio, que serão melhoradas à medida que se avançará no tempo. Guardadas as devidas proporções, é algo que se aproxima da natureza humana: aprender com o passar do tempo;
- (2) Em razão de ser uma concepção de extração do conhecimento incremental, não se constitui num processo iterativo que, associado a formas simples de ajuste dos pesos sinápticos, a fase de treinamento é executada com extrema rapidez, sendo, portanto, compatível para as aplicações em tempo real;
- (3) Como consequência destas propriedades, as redes ART-descendentes são rigorosamente

estáveis, viz., é garantida a obtenção de uma solução;

- (4) Por se tratar de uma família de redes neurais, cada componente dessa família possui características particulares para resolver problemas específicos. Também há a possibilidade de agregar inovações com largo espectro de alternativas visando maximizar o desempenho para cada tipo de aplicação. No caso que se pretende abordar nesta pesquisa de doutorado, *i.e.*, a previsão de carga, objetiva-se implementar inovações adequadas e pertinentes neste sentido.

As redes neurais ART-Descendente, evidentemente, apresentam limitação. A limitação principal observada é relativa a sua operacionalidade ser baseada no conceito de geração de classes que podem comprometer a precisão dos resultados. Contudo, pode-se superar este impasse; lança-se a possibilidade de agregar as inovações como argumentado anteriormente.

Assim sendo, nesta pesquisa, usar-se-á a arquitetura ARTMAP-*Fuzzy* (CARPENTER-GROSSBERG, 1992). A designação confere à rede neural ART (exclusiva para a resolução de problemas no modo não supervisionado, cujo treinamento é realizado mediante estímulos apenas de entrada à rede neural) a capacidade de resolução na forma supervisionada (estímulos de entrada e de saída). O sufixo “MAP” de (*mapping*, do idioma inglês) significa mapeamento. É um recurso baseado na formação de classes visando gerar o mapeamento entre a entrada e a saída. Esta é a concepção do problema da previsão de carga, *i.e.*, considerando-se um conjunto de informações passadas e corrente, pode-se estimar o valor de carga à frente de forma recursiva. Fazendo-se o tempo discreto variar de um ponto inicial até um ponto final, gera-se uma curva de demanda, por exemplo, 24 horas à frente. A qualidade desta curva depende de escolhas “otimizadas” dos componentes da janela (informação sobre a carga, condições climáticas, informações do tempo, bem como outras informações relevantes). A inclusão dos conceitos de conjuntos *fuzzy* (ZADEH, 1965) objetiva-se à possibilidade de se trabalhar com informações binárias e analógicas juntamente. Isto é, em regra, outro diferencial em relação à literatura que, frequentemente, é observada a dificuldade de manipular dados binários e analógicos simultaneamente. O fato de combinar dados binários e analógicos representa um grande avanço para a abordagem de problemas complexos envolvendo, por exemplo, a linguagem, imagens, bem como outros tipos de dados permitindo o desenvolvimento de sistemas de solução no contexto da teoria da fusão de dados (BARNETT, 2008).

Na sequência será apresentado o algoritmo conceitual básico da rede neural ARTMAP-*Fuzzy*, a qual mais adiante será discutida as possíveis inovações que podem ser agregadas no sistema previsor.

4.2. REDE NEURAL ARTMAP-FUZZY

A rede neural ARTMAP-*Fuzzy*, a partir da concepção ARTMAP, foi desenvolvida agregando-se os recursos da teoria de conjuntos *fuzzy*. Substituindo a álgebra referente à teoria de conjuntos pela teoria de conjuntos *fuzzy* de Zadeh (ZADEH, 1965) resultou a rede neural ARTMAP-*Fuzzy*. Com esta estrutura pode-se trabalhar com problemas que envolvem dados binários e analógicos de modo supervisionado. Isto, evidentemente, aumenta drasticamente a capacidade de realização de problemas complexos, que é o caso da previsão de carga (objeto de estudos nesta pesquisa).

Esta rede neural é composta por três módulos principais: dois módulos ART-*Fuzzy* (por simplicidade de notação é designado apenas por ARTa, ARTb e um módulo inter-ART). ART é uma rede neural não supervisionada, cujos estímulos recebidos são procedentes apenas de entrada. Na rede neural ARTMAP-*Fuzzy* os módulos ARTa-*Fuzzy* e ARTb-*Fuzzy* são empregados para receberem estímulos de entrada e de saída, respectivamente. Com isto, confere à rede neural ARTMAP a capacidade de realização de funções complexas, ou mapeamento entre a entrada e a saída. O módulo Inter-ART é usado visando a realização do “casamento” entre a entrada e a saída compondo um processo matemático similar a $y = f(x)$, sendo x e y dois vetores correspondentes à entrada e à saída. A similaridade advém em razão deste processo ser realizado de modo discreto via formação de classes. A função vetorial f desconhecida pode ser linear, não linear, ou outro tipo de comportamento, a qual é realizada pela rede neural ARTMAP-*Fuzzy*, via processo de treinamento, correspondente à extração do conhecimento baseado em estímulo entrada-saída. No caso da previsão de carga, a extração do conhecimento é determinada a partir de uma base de dados históricos, cuja função é oferecer exemplos. Busca-se, a partir de exemplos de como ocorreu o consumo passado, projetar um possível comportamento de carga no futuro. Evidentemente, este processo necessita ser cuidadosamente projetado e estudado para que os resultados esperados sejam os mais precisos possíveis. Este será o objeto de atuação nesta pesquisa de doutorado.

As informações a serem manipuladas na rede ARTMAP-*Fuzzy* devem ser convertidas resultando um espectro compreendido entre 0 e 1. Portanto, inicialmente, há a necessidade de realizar uma operação de normalização dos dados de entrada e de saída da rede neural, cuja resposta (análise) final deverá ser convertida para as informações reais do problema que está sendo abordado.

Na arquitetura ARTMAP-*Fuzzy*, como toda rede neural, a extração do conhecimento é realizada baseada na adaptação (ajuste) sistemática de pesos. Assim sendo, a rede neural ARTMAP-*Fuzzy* possui três matrizes de pesos nominados W_a , W_b e W_{ab} , para receberem a adaptação referentes aos módulos ARTa, ARTb e Inter-ART, respectivamente, como será ilustrado no algoritmo (Seção 4.4).

Para compreensão plena, consideram-se dois vetores a e b nominados, respectivamente, vetores de entrada e de saída da rede neural ARTMAP-*Fuzzy*. O vetor de entrada a corresponde às informações (as quais serão abordadas e discutidas mais à frente) é correspondente a um conjunto de dados referentes às cargas passadas, temperatura, estações do ano, dias da semana, dados horários, assim como toda a informação que pode contribuir à obtenção de uma solução com a máxima precisão possível. O vetor b , via de regra, corresponde ao valor de carga normalizada no tempo discreto ligeiramente à frente do tempo corrente. O tempo discreto pode ser tratado como intervalo de 5, 10, 15, 30 minutos ou até 1 hora. Isto dependerá do grau de exigência do interessado. Percorrendo todo o intervalo dos dados históricos estabelecidos (compreendendo a fase de treinamento), ao final pode-se gerar a curva de carga, por exemplo, correspondente a 24 horas à frente.

O funcionamento da rede neural ARTMAP-*Fuzzy*, fundamentalmente, consiste nas seguintes fases: (1) fase de treinamento e (2) fase de análise. Estas duas fases ocorrem de modo *off-line* (treinamento) e *on-line* (fase de análise). O treinamento é realizado de modo *off-line*, porém, o tempo transcorrido para a sua realização é, efetivamente, reduzidíssimo, se comparado às principais metodologias da literatura em razão de serem realizadas operações simples de atualização de pesos sem a necessidade de processo repetitivo. Os processos repetitivos (recursivos), além de elevar drasticamente o tempo de processamento, podem incorrer em situações de não convergência.

4.3. ALGORITMO PARA A EMULAÇÃO DA REDE NEURAL ART-FUZZY

Conforme descrito na Seção 4.2, a rede neural ARTMAP-*Fuzzy* é constituída por duas redes neurais ART-*Fuzzy* (redes não supervisionadas) interconectadas por um módulo inter-ART que possui a função de realizar o “casamento” entre a entrada e a saída correspondendo a uma idealização de uma rede neural supervisionada. Por conseguinte, na sequência, será apresentado o algoritmo conceitual da rede neural ART-*Fuzzy*. Esse algoritmo será usado com o propósito de, finalmente, apresentar a concepção da rede neural ARTMAP-*Fuzzy* (Seção 4.4).

A seguir, apresenta-se o algoritmo básico a ser empregado nesta pesquisa. Mais à frente serão discutidas e implementadas algumas melhorias (inovações) com vistas à maximização do seu desempenho.

Nota-se que, quando se trabalha com redes neurais da família ART é conveniente, por sugestão dos autores Carpenter & Grossberg (1992), representar um vetor por linha e não por coluna como habitualmente adotado na literatura, que será a convenção adotada doravante. O algoritmo apresentado adiante é uma versão do apresentado na publicação Carpenter & Grossberg (1992), assim como de outras publicações importantes sobre as redes neurais da família ART (GROSSBERG, 2013).

Passo 1. *Leitura dos parâmetros da rede neural ART-Fuzzy:*

1. $\alpha > 0$ (parâmetro de escolha);
2. $\beta \in [0,1]$ (taxa de treinamento);
3. $\rho \in [0,1]$ (parâmetro de vigilância);
4. M (dimensão do vetor de entrada);
5. L (quantidade de vetores de entrada).

Passo 2. *Leitura do vetor de entrada \mathbf{a} (estímulo de entrada)*

$$\mathbf{a} = [a_1 \ a_2 \ \dots \ a_M]. \quad (1)$$

Passo 3. *Inicialização da Matriz de Pesos da Rede Neural ART-Fuzzy:*

Tendo em vista que no Passo 5, para a realização da rede neural ART-Fuzzy será necessária a duplicação da dimensão do vetor \mathbf{a} , a matriz de pesos da rede neural ART-Fuzzy será assim definida:

$\mathbf{W}_{(L \times 2M)}$, cujos componentes iniciais são postos iguais a 1:

$w_{ij} = 1$, para qualquer $i = 1, 2, \dots, L$ e $j = 1, 2, \dots, 2M$.

Isto indica que não há nenhuma categoria ativa previamente.

Passo 4. *Processo de Normalização*

Tendo em vista que os componentes do vetor de entrada devem estar compreendido entre 0 e 1, via processo de normalização, deve-se processar o vetor de entrada \mathbf{a} obtendo-se o vetor de entrada auxiliar $\bar{\mathbf{a}}$ da seguinte forma:

$$\bar{\mathbf{a}} = \mathbf{a} / \|\mathbf{a}\|_1 \quad (2)$$

sendo:

$\bar{\mathbf{a}}$: vetor de entrada normalizado;

$\|\mathbf{a}\|_1$: norma₁ do vetor \mathbf{a} ;

$$= \sum_{i=1}^M |a_i|. \quad (3)$$

O uso da norma₁ representa o delineamento geométrico da formação de classes por hiper-retângulos. Se for usada a norma₂ refere-se à distância euclidiana, cuja rede neural ART é designada, na literatura, como sendo a rede neural Euclidiana. Na literatura, há, também, o emprego de outras geometrias, inclusive fracionárias (DAGHER, 2006). Todas essas formulações são propostas com o objetivo de contemplar melhorias do desempenho para aplicações específicas. No caso do problema da previsão de carga, o uso da norma₁ tem proporcionado resultados satisfatórios (atendimento de normas técnicas específicas, *e.g.*, erro absoluto não superior a 5%). Ressalta-se que se os componentes do vetor \mathbf{a} , inicialmente, satisfazem a restrição $a_i \in [0,1]$, não há necessidade de realizar a operação de

normalização, como proposta no Passo 4.

Passo 5. *Processo Designado Como Codificação Complementar*

Esta operação visa tornar todos os vetores entre os L exemplares da fase de treinamento do mesmo tamanho. Portanto, tem-se o vetor \mathbf{I} assim definido:

$$\mathbf{I} = [\bar{\mathbf{a}} \ \bar{\mathbf{a}}^c] \quad (4)$$

sendo:

$$\bar{\mathbf{a}}^c = [\bar{a}_1^c \ \bar{a}_2^c \ \dots \ \bar{a}_M^c] \quad (5)$$

$$\bar{a}_i^c = 1 - \bar{a}_i \quad (6)$$

Deste modo a norma₁ do vetor \mathbf{I} corresponde a:

$$\begin{aligned} \|\mathbf{I}\|_1 &= \sum_{i=1}^M [\bar{a}_i + (1 - \bar{a}_i)] \\ &= M \text{ (todos os vetores terão o mesmo comprimento } M\text{).} \end{aligned} \quad (7)$$

Passo 6. *Escolha da Categoria*

Considerando-se o vetor de entrada \mathbf{I} na rede neural *ART-Fuzzy*, para cada nó j a função de escolha T_j é determinada por:

$$T_j = \|\mathbf{I} \wedge \mathbf{W}_j\|_1 / (\alpha + \|\mathbf{W}_j\|_1) \quad (8)$$

sendo:

T_j : função de escolha ou índice de similaridade entre os vetores \mathbf{I} e \mathbf{W}_j ;

\wedge : operador *AND* nebuloso definido por:

$$(\mathbf{I} \wedge \mathbf{W})_i \therefore \min(I_i, W_i), i = 1, 2, \dots, L; \quad (9)$$

α : parâmetro de escolha ($\alpha > 0$);

A categoria é escolhida como sendo o nó v_j ativo:

$$v_j = \arg\{\text{máx } T_j\}, j = 1, 2, \dots, 2M. \quad (10)$$

Considerando-se a equação (10), a categoria escolhida será aquela que possuir menor índice.

Passo 7. Ressonância ou Reset

A ressonância ocorrerá se o critério de vigilância (equação (11)) for satisfeito:

$$\| \mathbf{I} \wedge \mathbf{W}_{vj} \|_1 / \| \mathbf{I} \|_1 \geq \rho \quad (11)$$

Caso o critério, definido pela equação (11), não for satisfeito, o dispositivo *reset* deverá ser acionado. No *reset*, o nó v_j é excluído do processo de busca (conforme estabelecido na equação (10)), arbitrando-se $T_{vj} = 0$. Então, escolhe-se uma nova categoria usando-se a equação (10) no processo de ressonância. Este procedimento deverá ser repetido até que se encontre uma categoria que satisfaça a equação (11).

Passo 8. Atualização dos Pesos (Treinamento)

Uma vez satisfeito o critério de vigilância (11), o índice do neurônio vencedor será designado J . A atualização dos pesos da rede neural ART-Fuzzy é realizada da seguinte forma:

$$\mathbf{W}_J^{novo} = \beta [\mathbf{I} \wedge \mathbf{W}_J^{velho}] + [1 - \beta] \mathbf{W}_J^{velho} \quad (12)$$

sendo:

J : índice referente à categoria ativa;

\mathbf{W}_J^{novo} : vetor peso atualizado;

\mathbf{W}_J^{velho} : vetor peso referente à atualização anterior.

Observa-se que se $\beta = 1$, tem-se o treinamento nominado rápido e simples (*naïve*). Para um valor de $\beta < 1$ o treinamento, também, será rápido, contudo o significado fica por conta de uma realização mais sofisticada.

4.4. ALGORITMO PARA A EMULAÇÃO DA REDE NEURAL ARTMAP-*FUZZY*

De acordo com informações anteriores, a rede neural ARTMAP-*Fuzzy* (CARPENTER & GROSSBERG, 1992) é uma estrutura baseada no treinamento supervisionado. Assim sendo, devem-se empregar os estímulos de entrada-saída. A entrada e a saída correspondem entradas das redes neurais ARTa-*Fuzzy* e ARTb-*Fuzzy*, respectivamente. Porquanto, para a realização do treinamento, a partir da base de dados da fase de treinamento, tem-se L padrões, sendo que cada padrão é definido como sendo o par (\mathbf{a}, \mathbf{b}) . Por medida de simplicidade de notação, deliberadamente, suprime-se a designação “*Fuzzy*”, passando-se apenas o registro como ARTa e ARTb, respectivamente. Os vetores \mathbf{a} e \mathbf{b} são usados como entradas, respectivamente, das redes neurais ARTa e ARTb. Assim, tem-se:

$$\dim(\mathbf{a}) = M$$

$$\dim(\mathbf{b}) = N.$$

As redes neurais ARTa e ARTb são treinadas, usando-se como entradas, respectivamente, os vetores \mathbf{a} e \mathbf{b} empregando-se igualmente o algoritmo ilustrado na Seção 4.3. O treinamento supervisionado da rede neural ARTMAP-*Fuzzy* é realizado via supervisão do módulo inter-ART. Para esta idealização, tem-se três matrizes de pesos ($\mathbf{W}^a_{(L \times 2M)}$, $\mathbf{W}^b_{(L \times 2N)}$ e $\mathbf{W}^{ab}_{(L \times L)}$), as quais estão associadas aos módulos ARTa, ARTb e inter-ART. Estas matrizes são inicializadas da seguinte forma:

$$W_{p,q}^a = 1, \text{ para } p = 1, 2, \dots, L; q = 1, 2, \dots, 2M$$

$$W_{p,q}^b = 1, \text{ para } p = 1, 2, \dots, L; q = 1, 2, \dots, 2N$$

$$W_{p,q}^{ab} = 1, \text{ para } p = 1, 2, \dots, L; q = 1, 2, \dots, L$$

Isto significa que, inicialmente, todos os neurônios estão inativos. Estes neurônios serão ativados à medida que forem indicados os neurônios vencedores das redes neurais

ARTa (neurônio J) e ARTb (neurônio K) ajustando-se as matrizes de pesos \mathbf{W}^a e \mathbf{W}^b . Consequentemente, deve-se ajustar a matriz \mathbf{W}^{ab} tomando os indicativos dos neurônios vencedores J e K , como instruídos na sequência.

A atualização de pesos da rede neural ARTb é realizada de forma independente, não há necessidade da verificação do neurônio vencedor J da rede neural ARTa. Assim sendo, procede-se da seguinte forma. No treinamento da rede neural ARTMAP-*Fuzzy*, o neurônio vencedor associado à rede neural ARTa é estabelecido, se, e somente se, o neurônio candidato atender dois critérios de vigilância simultaneamente:

$$\|\mathbf{I}^a \wedge \mathbf{W}_J^a\|_1 / \|\mathbf{I}\|_1 \geq \rho_a \quad \text{e} \quad (13)$$

$$\|\mathbf{y}_K^b \wedge \mathbf{W}_J^{ab}\|_1 / \|\mathbf{y}_K^b\|_1 \geq \rho_{ab} \quad (14)$$

sendo:

$\rho_{ab} \in [0,1]$: parâmetro de vigilância do módulo inter-ART;

\mathbf{y}_K^b : vetor que compõe os índices dos neurônios ativados da rede neural ARTb;
 $= [0 \ 0 \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0 \ 0 \ 0 \ \dots \ 0]$.

Havendo o cumprimento das restrições estabelecidas pelos critérios de vigilância (13) e (14), deve-se proceder a atualização dos pesos como segue:

1) Rede Neural ARTa:

$$\mathbf{W}_J^{a\,novo} = \beta [\mathbf{I}^a \wedge \mathbf{W}_J^{a\,velho}] + [1 - \beta] \mathbf{W}_J^{a\,velho} \quad (15)$$

2) Rede Neural ARTb:

$$\mathbf{W}_K^{b\,novo} = \beta [\mathbf{I}^b \wedge \mathbf{W}_K^{b\,velho}] + [1 - \beta] \mathbf{W}_K^{b\,velho} \quad (16)$$

3) Módulo inter-ART:

$$W_{JK}^{ab\,novo} = 1 \quad (17)$$

$$W_{ij}^{ab\,novo} = 0, \text{ para } i = J \text{ e } j \neq K. \quad (18)$$

Os valores W_{JK}^{ab} e W_{ij}^{ab} são componentes da matriz \mathbf{W}^{ab} .

4.5. MELHORIAS DO DESEMPENHO DA REDE NEURAL ARTMAP-FUZZY

A seguir, destacam-se as principais melhorias considerando-se a rede neural ARTMAP-Fuzzy:

- (1) Uma das possíveis melhorias que podem ser implementadas na rede neural ARTMAP-Fuzzy consiste no emprego de inferência da maximização da função de escolha, tanto da rede neural ARTa, assim como da ARTb. Por exemplo, pode-se usar o índice de Jaccard (TANIMOTO, 1958), ajustado ao trato com conjuntos *fuzzy*, considerando-se os vetores U e V :

$$Jac(U, V) = \|U \wedge V\|_1 / \|U \vee V\|_1 \quad (19)$$

sendo:

Jac : índice de Jaccard.

Há outros índices de similaridade na literatura que podem ser testados (GOODALL, 1966). Propõe-se, nesta pesquisa, o uso do seguinte índice de similaridade:

$$sim(U, V) = A / (A + \|U - V\|_1) \quad (20)$$

sendo:

sim : índice de similaridade;

A : $\|U\|_1$.

Por exemplo: considerando-se os vetores:

$$U = [0,50 \ 0,40 \ 0,70 \ 0,50 \ 0,60 \ 0,30] \text{ e}$$

$$V = [0,10 \ 0,75 \ 0,60 \ 0,90 \ 0,25 \ 0,40]; \quad \|U\|_1 = \|V\|_1 = 3.$$

As equações (8), (19) e (20) representam métricas de inferência de similaridade entre

dois vetores. Assim sendo, aplicando-se os índices de similaridade, usando-se as equações sobreditas ((8), (19) e (20)), obtêm-se os valores correspondentes a 0,7167; 0,5584 e 0,6383, respectivamente. As equações destes índices podem ser experimentadas com vistas à determinação do desempenho de cada índice e apurar o melhor resultado obtido e, assim, sugerir-se um bom candidato para o índice de medida de similaridade.

Estes índices, como qualquer outro, a ser usado como métrica possuem uma extensão entre 0 e 1. Um valor igual a 0 ou 1, representam a ausência ou a completude da similaridade entre dois vetores, respectivamente. Os valores intermediários indicam o grau de similaridade.

As referidas funções de similaridade podem, também, ser ajustadas para uso no teste de vigilância, em razão de ser, igualmente, uma inferência de similaridade;

- (2) Também, é viável modificar a forma de atualização dos pesos das matrizes \mathbf{W}^a e \mathbf{W}^b . Neste caso, há um forte indício que se pode melhorar bastante a precisão via um “ajuste ótimo” do parâmetro $\beta < 1$, buscando-se realizar um treinamento sofisticado. Deve-se ressaltar que quase na totalidade das aplicações ilustradas na literatura é adotado $\beta = 1$.
- (3) Alternativamente à melhoria, consiste a adoção do parâmetro de vigilância inicial (designado $\rho_{baseline}$) do módulo ARTa igual a:

$$\begin{aligned} \rho^a &\rightarrow \rho_{baseline} \\ &= \varepsilon + \|\mathbf{I}^a \wedge \mathbf{W}_j^a\|_1 / \|\mathbf{I}^a\|_1 \end{aligned} \quad (21)$$

para os vetores \mathbf{I}^a e \mathbf{W}_j^a ,

sendo:

ε : um número pequeno e positivo.

Se forem atendidas as relações (13) e (14) simultaneamente, procede-se a atualização dos pesos dos módulos ARTa, ARTb e de inter-ART usando-se as equações (15)-(18). Os índices temporários v_j e v_k passam a ser considerados neurônios vencedores de ARTa e de ARTb, assumindo como índices J e K , respectivamente. Se (13) for satisfeita e (14) não, o parâmetro ρ^a deve ser ajustado usando a equação (21). Na sequência, busca-se um novo neurônio candidato usando-se a métrica da função de escolha T_j

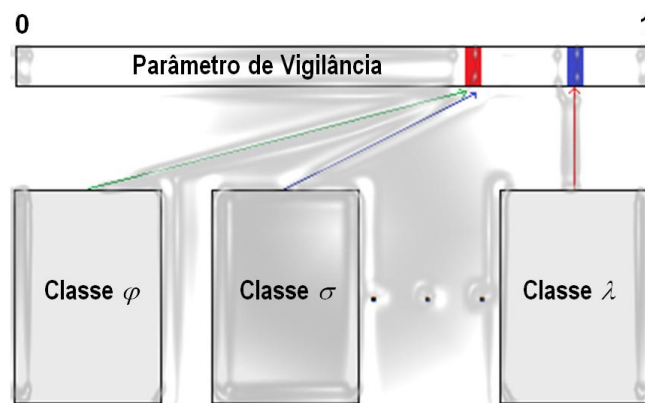
(retorno ao Passo 6, equação (8)) e fazendo-se $T_{vj} = 0$, sendo v_j o índice referente ao neurônio candidato temporário da rede neural ARTa.

Esta opção é introduzida entre a execução das equações (13) e (14) (execução de testes). O parâmetro de vigilância ρa não será um valor fixo, mas sim variável para cada categoria estabelecida pela rede neural ARTMAP-*Fuzzy*. Este detalhe tem sido um ponto importante, pois o tamanho de $\rho a \in [0, 1]$ é ajustado de forma autônoma, via implementação da estrutura computacional da rede neural ARTMAP-*Fuzzy*, mais especificamente, via atuação combinada entre os módulos ARTa e Inter-ART. Evita-se, desta forma, fixar arbitrariamente, por exemplo, um valor $\rho a = 0,7$ o que implica que a semelhança entre os padrões que fazem parte de determinada classe possui medida igual ou superior a 0,7. Isto implica que, objetos com certo grau de semelhança poderão ser confundidos num processo de reconhecimento. No caso da determinação autônoma de ρa , esta possibilidade de ocorrência, por certo, será atenuada. Deve-se ressaltar que nos primeiros anos após a proposição da rede neural ARTMAP-*Fuzzy*, o parâmetro de vigilância ρa era fixado como um valor constante, razão pela qual, frequentemente, havia muitas críticas sobre o desempenho desta rede neural. A qualidade dos resultados era fortemente afetada pelo valor arbitrado de ρa , levando, por vezes, a uma interpretação desfavorável desta arquitetura neural (GEORGIOPOULOS, *et al.*, 1996). Considerando-se os vários aperfeiçoamentos praticados nos últimos anos, esta rede neural, em termos de precisão, passou a ser bastante competitiva. Associada a outras propriedades (rapidez de processamento, estabilidade e plasticidade), pode-se afirmar que se trata de uma excelente alternativa de redes neurais recomendada para várias aplicações em um grande número de áreas do conhecimento humano.

Uma arbitragem interessante – pelo menos se recomenda para ser testada – é usar o valor do parâmetro de vigilância na rede neural ARTa fixo e igual à proporção áurea da distância entre 0 e 1, *i.e.*, $\rho_{baseline} \times \Phi = 1$, sendo, portanto, $\rho_{baseline} = 0,61803399$ (61,8%), tal que $\Phi = 1,61803399$ é o número de ouro, ou áureo (número de Fibonacci) (VOROB'EV, 1961). Um valor próximo (caracterizado como “detector próprio”) de 0,61803399 tem sido usado em aplicações dos algoritmos baseados nos sistemas imunológicos artificiais (SIA's) (DASGUPTA, 1997) proporcionando ótimos resultados. Dasgupta (DASGUPTA, 1997) tem demonstrado a equivalência (estudo comparativo) entre os SIA's e as redes neurais ART-descendentes. Funcionou nos

SIA's. Por essa razão, há bastante indício que funcionará, também, com relação à rede neural *ARTMAP-Fuzzy*, contudo, há a necessidade de se propor um índice de similaridade devidamente ajustado para este desígnio. A proporção áurea tem sido observada em inúmeros casos na natureza, funcionando deste modo, como um importante instrumento para a sua construção, inclusive servindo como uma inferência de beleza estética de rostos humanos, observando-se as relações entre certas áreas e as distâncias relativas à face. Deve-se destacar que uma classe não está associada exclusivamente a um determinado valor do parâmetro de vigilância, mas sim dos valores de similaridade combinados com o parâmetro de vigilância (Figura 1). Daí justifica-se a pertinência desta sugestão ($\rho a = \Phi^{-1}$), ainda que seja um valor relativamente pequeno, se comparado ao uso convencional. Como mencionado previamente, esta concepção, a princípio, deverá ser combinada com um reparo (ajuste) adequado da função para a mensuração da similaridade. Entende-se que esta “nova função” de similaridade pode ser gerada por meio, também, de redes neurais simples usando-se, por exemplo, a rede neural *feedforward* com o treinamento realizado via uso do algoritmo retropropagação (WERBOS, 1974).

Figura 1. Ilustração sobre a formação de classes.



Fonte: Elaborada pela Autora.

- (4) Desenvolver sistemas híbridos combinando-se as rede neural ARTMAP-*Fuzzy* e outras técnicas da inteligência artificial (BENDER, 1996), como a teoria de evidência de Dempster-Shafer (SHAFER, 1976, ZHU *et al.*, 2017) etc. ou métodos clássicos como a lei de Benford (BENFORD, 1938; GÜNNEL & TÖDTER, 2009; MORZY *et al.*, 2016). Compreende o desenvolvimento de uma rede neural, nesta pesquisa, à arquitetura ARTMAP-*Fuzzy*, combinada, por exemplo, com a teoria de evidência de Dempster-Shafer como forma de incluir inovação no contexto da previsão de carga. Também, pode-se sugerir o uso de teoria sobre a Lei Benford (BENFORD, 1938; GÜNNEL; TÖDTER, 2009; CZARNECKI, 2015) para a análise das bases de dados. Isto representa a busca da melhoria na previsão de carga.
- (5) Outro procedimento que se pode melhorar o diagnóstico (análise, previsão etc.) corresponde à implantação do critério de votação (CARPENTER; GROSSBERG, 1992). É um processo em que se usa o treinamento pluralizado de modo aleatório, ou pseudoaleatório (diferentes ordenações a partir de diferentes sementes) dos pares de vetores de dados (entrada, saída) a serem apresentados à rede neural, ao invés de uma única vez, por exemplo, cinco vezes. Uma vez realizada a execução das redes neurais (das cinco vezes), o diagnóstico “correto” é atribuído considerando-se a maior frequência de classes apuradas. Este procedimento é bastante interessante nos casos da maioria das aplicações conhecidas. Entretanto, no caso da previsão de carga, em razão da janela usada ser deslocada sequencialmente (do início ao fim da base de dados históricos), há a necessidade de uma melhor investigação sobre o uso deste procedimento. Da mesma forma, a maior frequência como critério para a melhor escolha da solução pode ser substituída pela média aritmética entre as referidas soluções obtidas nas diversas ordenações aleatórias, tendo em vista que as saídas, na formulação apresentada nesta pesquisa, são valores analógicos. Esta média pode produzir resultados melhores, se comparada ao procedimento usual considerando-se apenas uma ordenação.

CAPÍTULO 5 - PREVISÃO DE CARGA GLOBAL

5.1. INTRODUÇÃO

A previsão de carga total do sistema é formulada, como mencionado anteriormente, via emprego da rede neural ARTMAP-*Fuzzy*. Assim, tem-se uma rede neural supervisionada, em que são considerados os estímulos vetoriais de entrada ($\mathbf{a} \in \mathbb{R}^M$) e de saída ($\mathbf{b} \in \mathbb{R}^N$). O funcionamento de uma rede neural é desdobrado em duas partes principais: (1) treinamento e (2) o diagnóstico. Pode-se, ainda, incluir uma fase intermediária correspondente à validação ou inferência sobre a qualidade das soluções, baseando-se numa métrica adequada. Ela é composta por três módulos principais: ARTa, ARTb e inter-ART. Os dois módulos ART são redes neurais usadas para processarem os dados de entrada e de saída, respectivamente. Durante a fase de treinamento, o módulo Inter-ART é responsável pelo “enlace/casamento” entre a entrada e a saída, *i.e.*, garantir a construção de um mapeamento (formação de classes funcionais). Esta ideia é compatível com o problema a ser abordado nesta pesquisa. A fase de treinamento é realizada procedendo-se o treinamento de modo igual nos módulos ARTa e ARTb e finalizado o armazenamento do conhecimento, para cada par de vetores padrões (\mathbf{a} , \mathbf{b}), mediante a confirmação via teste realizado pelo módulo inter-ART.

5.2. COMPOSIÇÃO DOS VETORES DE ENTRADA E DE SAÍDA DA REDE-NEURAL ARTMAP-FUZZY

Após o uso de toda a base de dados destinada à fase de treinamento, encerra-se o processo de aprendizado da rede neural. A forma adotada neste trabalho, em relação aos vetores padrões (\mathbf{a} , \mathbf{b}), segue a proposta referente ao deslocamento de janela (PARK *et al.*, 1991). Cada janela é assim definida (ABREU *et al.*, 2018):

$$(\mathbf{a}(h), \mathbf{b}(h)) \tag{22}$$

$$\mathbf{a}(h) = [\mathbf{L T E}] \text{ (vetor de entrada)} \quad (23)$$

$$\mathbf{b}(h) = [\mathbf{H}] \text{ (vetor de saída)} \quad (24)$$

sendo:

h : indicador do tempo na forma discreta.

O treinamento consiste na apresentação de todos os pares de vetores padrões (bem como os devidos ajustes das matrizes de pesos caracterizando a extração do conhecimento) desde o início até o término da base de dados referente à fase de treinamento. Este processo é realizado de modo direto (sem interações). Em razão disto, é que a fase de treinamento da rede neural ARTMAP-*Fuzzy* é efetivamente rápida e não oferece problemas de convergência (propriedade de estabilidade). Os subvetores $\mathbf{L T E H}$ são assim definidos:

\mathbf{L} : vetor cujos componentes são os valores de cargas correspondentes à hora corrente (h) e demais valores consonantes a η horas em atraso, cujo atraso é sob a livre escolha do usuário. Neste vetor podem ser incluídos valores, por exemplo, correspondentes ao mesmo horário da semana, mês anterior etc. como forma de aprimorar a capacidade da extração do conhecimento contido na base de dados da fase de treinamento;

\mathbf{T} : vetor de temperatura;

\mathbf{E} : vetor contendo as informações exógenas (dia da semana, estação do ano, dia nublado, dia festivo etc.);

\mathbf{H} : carga correspondente à carga ($h+1$), ou seja, uma hora à frente. Observe que $\mathbf{H} \in \mathfrak{R}^N$ ($N = 1$).

A discretização depende do tipo de amostragem disponível na base de dados. As fases de treinamento, teste e de análise (previsão) são realizadas de acordo com algoritmo apresentado nas Seções 4.3 e 4.4. Observa-se que se trata de um funcionamento bastante simples. É um processo direto (não iterativo), estável, e rápido (demandando apenas uma época). Exemplo: se o tempo discreto for considerado 10 minutos e a contagem h a partir da 24 horas, tem-se a seguinte sequência de tempo: $h = 0$ (0 minuto), $h = 1$ (10 minutos); $h = 2$ (20 minutos); $h = 3$ (30 minutos) e assim por diante até completar o horizonte alvo (por exemplo, 24 horas à frente).

Um importante procedimento que se pode empregar no vetor de entrada \mathbf{a} , refere-se ao uso de algum método visando identificar a correlação (de Pearson, Kendall, Spearman etc.) (CHEN & POPOVICH, 2002) entre os seus componentes. Porém, considerando-se que a dimensão do vetor \mathbf{a} é bastante reduzida (na maioria dos casos é inferior a 20), o estudo sobre a correlação pode ser dispensado. A redução de componentes pode ser adotada baseada na experiência com o trato com a previsão de carga. Por exemplo, a temperatura horária pode ser dispensada em razão do uso da informação horária que, grosso modo, estão frequentemente correlacionadas. Destaca-se que as redes neurais não necessitam de uma análise prévia de correlação dos componentes do vetor de entrada, ou da modelagem de carga. Se houver correlação, a rede neural é capacitada para se ajustar no caso de ocorrência deste detalhe. Também não há a necessidade de definir o conjunto de componentes de entradas que são mais ou menos importantes. Na rede neural, por exemplo, uma *feedforward* com treinamento realizado via algoritmo retropropagação (WERBOS, 1974), após a conclusão da fase de treinamento, pode-se obter os coeficientes de sensibilidade de cada elemento (LOTUFO *et al.*, 2007) em relação ao resultado final (resposta de saída da rede neural). Estes coeficientes são gerados a partir das matrizes de pesos da rede neural. Por conseguinte, podem-se excluir os componentes com baixos valores (em módulo) dos referidos coeficientes. Isto representa um diferencial em relação às metodologias clássicas de previsão.

Muitas vezes, nas bases de dados disponíveis, não constam os dados de temperatura. Quando constam, comumente, são disponíveis os dados de temperaturas máxima (T_{max}) e mínima (T_{min}). Havendo estes valores, pode-se, perfeitamente, introduzi-los no vetor \mathbf{T} . Também, se conhecidos os valores de T_{max} e T_{min} , a temperatura correspondente ao instante t , na forma discreta, pode ser estimada, por exemplo, usando-se as seguintes expressões (LINVILL, 1990):

$$T(t) = (T_{max} - T_{min}) \times \text{sen}[(\pi t)/(DL + \gamma)] + T_{min} \quad (25)$$

sendo:

DL : duração do dia com luz solar (número de horas);

t : tempo após o nascer do sol;

γ : parâmetro constante que deve ser ajustado em função da localização. Na referência Linvill (1990) este parâmetro é fixado igual a 4 unidades.

A duração do dia (DL), no estado de São Paulo, no ano de 2019, variou entre 10h40min (mês de Junho) e 13h35min (Dezembro). A DL diária pode bem ser armazenada em tabela para ser usada a qualquer hora.

Usando-se os parâmetros T_{max} e T_{min} , a previsão tem sido satisfatória. Contudo, havendo a necessidade do conhecimento da temperatura horária, a sugestão é usar a expressão (25) para estimá-la. Ressalta-se que há, na literatura técnico-científica, outras propostas para a estimação da temperatura horária, caso o usuário assim desejar. Se for usada a temperatura horária, pode-se dispensar a consideração dos parâmetros T_{max} e T_{min} . Isto representa uma economia do número de componentes do vetor de entrada \mathbf{a} , mais especificamente, no vetor \mathbf{T} .

O vetor \mathbf{L} deve ser cuidadosamente elaborado (composição da janela). Diferentes janelas podem produzir resultados com maior, ou menor precisão. Inicialmente, pode-se compor este vetor considerando-se a seguinte estratégia:

$$\mathbf{L} = [La(h) \ L(h-3) \ L(h-2) \ L(h-1) \ L(h)] \quad (26)$$

sendo:

h : indicador do tempo na forma discreta;

$L(h)$: valor de carga correspondente ao tempo discreto h ;

$(h-\Omega)$: horário discreto anterior a h , para $\Omega = 1, 2, 3, \dots$;

$La(h)$: valor de carga correspondente ao tempo discreto h do mesmo dia da semana anterior.

Com relação às informações que compõem o vetor \mathbf{E} , podem-se sofisticar tais informações, considerando-se os mais variados casos visando melhor representar as condições do dia (dia da semana, mês do ano, caracterização das condições do tempo (chuvoso, nublado etc.)), assim como outras informações que o usuário entender como relevantes. Como estas informações são representações binarizadas, deve-se usar o tamanho do subvetor de acordo com a necessidade, por exemplo:

$$E = [d \ m \ tem \ ad] \quad (27)$$

sendo:

- d*** : subvetor referente ao dia da semana na forma binária com 3 *bits* que são suficientes para a representação semanal de domingo a sábado;
- m*** : subvetor relativo ao mês do ano com 4 *bits*, sendo suficientes para representar os 12 meses do ano. Esta informação pode ser substituída pela representação por estações do ano, sendo necessários apenas 3 *bits*;
- tem*** : subvetor visando abrigar as informações sobre à situação atmosférica do dia, sendo suficiente apenas 2 *bits* (até 3 modalidades);
- ad*** : subvetor destinado a receber, na forma da codificação binária, outras informações relevantes, cujo número de *bits* será definido em função do número de informações de interesse do usuário.

Portanto, o vetor ***E***, considerando-se o seguinte cenário: quinta-feira do mês de abril e com chuva intensa, é apresentado da seguinte forma:

$$E = [1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1]. \quad (28)$$

Nota-se que o vetor ***a*** é composto por dois grupos de informações: analógicas (vetores ***L*** e ***T***) e binárias (vetor ***E***). Em consonância com informações apresentadas anteriormente, a rede neural ARTMAP-*Fuzzy* é dotada da propriedade de trabalhar com grandezas analógicas e binárias simultaneamente sem obstáculos. Assim sendo, a parte binária não necessita de um processo de normalização. Apenas os vetores ***L*** e ***T*** necessitam de cálculos de normalização usando-se a norma-1 (hiper-retângulos). Contudo, deve-se empregar a normalização dos vetores ***L*** e ***T*** separadamente visando-se evitar a “mistura” de informações. A normalização de cargas e das temperaturas deve ser aplicada a toda a base de dados previamente. Estas informações, quando conduzidas à fase de treinamento, cumprem os requisitos de seus componentes devem pertencer ao intervalo [0, 1]. Na fase de teste e de previsão, os resultados finais devem ser convertidos aos valores reais correspondentes à multiplicação dos valores previstos pela rede neural pelo valor basilar (usado na

normalização).

Ressalta-se que, usando a codificação binária, pode-se sofisticar ao máximo (não há restrição neste sentido) as informações, obviamente com vistas à melhoria da qualidade da previsão. Outro detalhe importante sobre a operacionalidade da rede neural ARTMAP-*Fuzzy* refere-se ao uso de dados analógicos e binários conjuntamente. Como esta rede neural possui a propriedade de manipular símbolos, pode-se realizar esta “mistura” de tipos de dados (analógicos e binários) procedendo-se a normalização de forma individualizada dos vários tipos de dados analógicos, por exemplo, os dados de cargas e os dados de temperatura. Considerando-se estas virtudes, a formulação da previsão de carga baseada na rede neural ARTMAP-*Fuzzy* pode receber inovações sem restrições.

A previsão de carga, alternativamente, pode ser realizada mediante o emprego do modelo proposto por Lopes (2005) usando-se uma arquitetura neural ART-ARTMAP-*Fuzzy*. Os dados analógicos da entrada são inseridos como entrada da rede neural ART-*Fuzzy* (treinamento não supervisionado) que são transformados em um conjunto de dados binários via um bloco designado “código ativo”. Estes dados binários são juntados aos demais dados binários para compor o vetor de entrada binária à rede neural ARTMAP-*Fuzzy*. Assim, a rede neural ARTMAP-*Fuzzy* terá um processamento inteiramente binário. Com isto, agrega-se mais capacidade de realização da previsão de carga. Esta abordagem não será usada nesta pesquisa, mas sim, manter-se-á o processamento considerando-se o vetor de entrada composto por informações analógicas e binárias, com vista à realização de vários experimentos (estudar e analisar as implicações físicas do problema, deixando de lado o enfoque sob o contexto da manipulação de símbolos da representação binária apenas) que podem produzir melhorias da qualidade das previsões. Assim se espera.

5.3. FORMULAÇÃO DA PREVISÃO DE CARGA BASEADA NO TREINAMENTO CONTINUADO (*CONTINUOUS LEARNING*)

O treinamento continuado não deve ser entendido apenas como uma continuidade da fase de treinamento clássica. A sua prática requer critérios e astúcia para que se obtenha a máxima eficiência sem grandes esforços computacionais, pois esta etapa deve ser realizada

concomitantemente com a previsão de carga. Isto, em tese, representa um acréscimo computacional, ainda que pequeno, durante a execução do objeto (previsão). Trata-se de uma forma sutil de seletividade da informação a ser aproveitada. Como compensação, espera-se obter a vantagem de usufruir uma forma incremental do conhecimento que, a princípio, o sistema previsor se tornará mais “inteligente” com o passar do tempo, por um tempo razoável. Este tempo poderá ser calibrado mediante a realização de experimentos. Esta limitação (do tempo de validade) poderá ocorrer em razão das informações que “ensinam” e, por vezes, são acompanhadas por imperfeições que desvirtuam a instrução pretendida. Diante deste escopo, pode-se perceber que o treinamento continuado constitui-se num universo aberto para ser explorado.

Esta capacitação (exercício do treinamento continuado) do sistema previsor é constituída por meio do conceito “Detecção de Novidade” (DN) (BISHOP, 1994) ao ser agregado à rede neural ARTMAP-*Fuzzy*. De uma forma simples, as informações disponíveis são analisadas onde são classificadas como: (1) inclusão de dados e (2) rejeição de dados. Após este processamento, as informações rotuladas no estado (1) são aproveitadas, ou seja, incluídas nas matrizes de pesos da rede neural produzindo, por pressuposto, um incremento do conhecimento. No caso da situação (2), os dados correspondentes podem ser descartados, ou não, mediante um novo teste de consistência e de relevância da informação que está sendo processada.

A formulação mais convencional para a realização da previsão de carga refere-se ao uso de uma base de dados históricos. Engendram-se mecanismos (nesta pesquisa, baseia-se no emprego das redes neurais) levando-se em consideração que o pretérito torna-se um excelente instrumento para emular o comportamento de carga no futuro. Contudo, à medida que se tenha a necessidade do conhecimento de carga o mais próximo possível do tempo real (BAKKER, 2011), esta estratégia deverá ser mais bem elaborada e cuidadosa. Portanto, devem-se contemplar, também, as informações procedentes e disponíveis a partir de sistemas de aquisição de dados (THOMAS; McDONALD, 2017). Ressalta-se que as análises – a serem realizadas com a máxima proximidade do tempo real – devem ser estabelecidas considerando-se o tempo total do processamento envolvido (cálculo do estado do sistema, análise de segurança contra faltas (defeitos) e, ainda, o tempo despendido para a implementação da fase de treinamento). É importante destacar que a disponibilidade dos resultados da análise não será cabalmente útil se o sobredito tempo de processamento for superior à antecipação pretendida, sob a pena de se ter os resultados sem oferecer a possibilidade de agir de forma

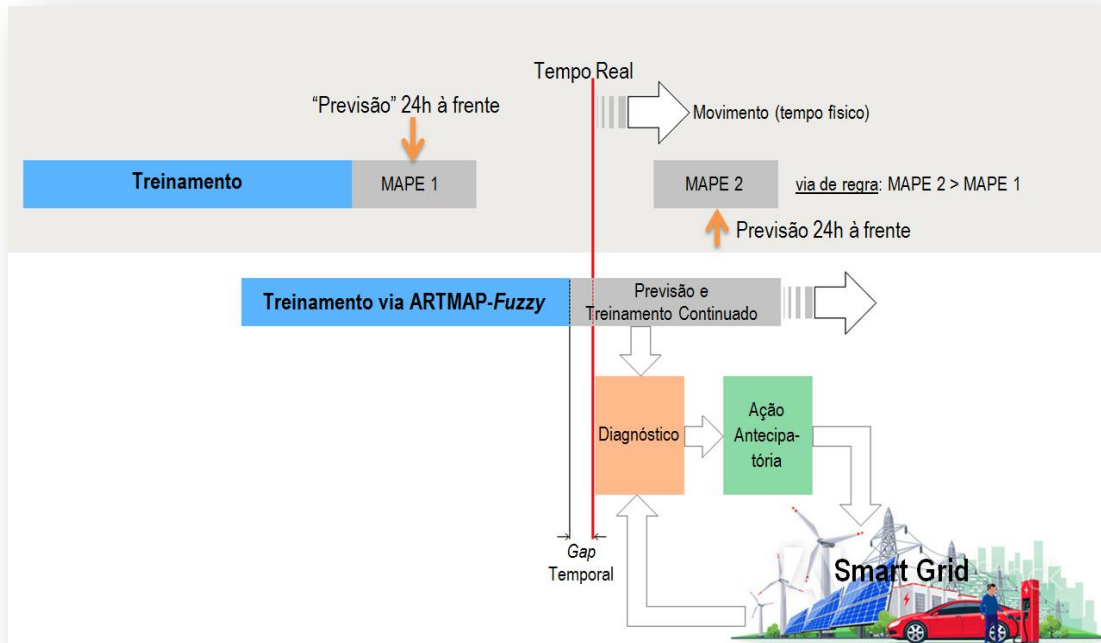
antecipatória, caso ocorra alguma contingência ou perturbação considerada na análise. Se assim for produzido, trata-se de um cálculo que permeia a nulidade.

Na Figura 2, ilustram-se, simbolicamente, as fases necessárias na busca de garantir que a operação ocorra de forma contínua e segura. Neste contexto, o tempo final admissível – a ser considerado na fase de treinamento da rede neural – mais próximo da operação, em tempo real, deverá ser arbitrado levando-se em conta o tempo despendido para a execução das análises e da tomada das ações preventivas. Deve-se destacar que, à medida que a modernidade e a inovação se fazem presentes, as técnicas de análise e as ações preventivas têm-se tornado cada vez mais eficientes e com tempo de execução, também, cada vez mais reduzido. Isto representa que se pode dar maior importância ao uso de informações procedentes do sistema de aquisição de dados (THOMAS; McDONALD, 2017) que, por certo, é a parte mais valiosa das informações disponíveis. Esta idealização pode ser perfeitamente implementada mediante a combinação da fase de treinamento (usando-se uma base de dados históricos) e o treinamento continuado tomando-se os dados advindos e armazenados, mais recentemente, a partir de sistemas de mensuração de carga.

As etapas ilustradas na Figura 2 devem ser entendidas como elaborações referindo-se a um tempo futuro, cuja antecipação deverá ser definida em função do tempo demandado para a realização da previsão, análise de contingências, assim como do cálculo e da implantação das ações antecipatórias.

Por conseguinte, uma forma mais eficiente corresponde à adoção do treinamento continuado (MARCHIORI *et al.* 2011). Trata-se da inclusão do conhecimento advindo de novas informações (procedentes do sistema de aquisição de dados, assim como de qualquer outra fonte) sem haver a necessidade de reiniciar a fase de treinamento. Isto é feito de modo incremental, obviamente, adotando-se formas adequadas para a sua implantação (MARCHIORI *et al.*, 2011). As redes neurais da família ART encontram-se entre as raras classes de redes neurais que são providas da característica de plasticidade. Esta propriedade viabiliza a idealização da inserção de novos conhecimentos sem destruir a cognição previamente adquirida. Grosso modo, compreende uma forma semelhante o que ocorre com os humanos, os quais aprendem continuamente, pelo menos em tese. É neste sentido que se buscará agregar o treinamento continuado ao estimador de carga. Ressalta-se que este procedimento, de tempo em tempo, deverá ser reinicializado quando forem identificados erros na previsão que não atendam as exigências mínimas do setor de eletricidade.

Figura 2. Estágios temporais da previsão de carga.



Fonte: Elaborada pela Autora.

Então, para a implantação do treinamento contínuo, inicialmente, deve-se definir o “melhor” tamanho da base de dados históricos a ser usada. Esta quantidade de informação deve ser bem ajustada (resultados do sistema predictor que atendam à qualidade pretendida), a qual será buscada via a realização de variados experimentos que serão apresentados e discutidos no Capítulo 6. O treinamento contínuo será procedido tomando-se as informações de cargas obtidas a partir do sistema de medição que são armazenadas numa base de dados específica para esta finalidade. As informações provenientes da base de dados históricos e as informações disponíveis a partir do sistema de aquisição de dados devem ser compatíveis em relação à mesma discretização temporal, ou seja, por exemplo, de 5 em 5 minutos.

Para fins da realização da extração do conhecimento, nesta pesquisa, consideram-se duas fases principais: (1) treinamento basilar tomando como objeto as informações originárias da base de dados históricos. Trata-se da forma convencional do processo provisional abordado na literatura; e (2) o treinamento contínuo. O treinamento contínuo compreenderá no uso de informações procedentes de bases de dados recentes que são construídas a partir do sistema de aquisição de dados com certo atraso (a ser devidamente ajustado), tendo em vista que o conceito “previsão” está situado à frente do tempo real.

Porquanto, o procedimento para execução da fase de treinamento da rede neural ARTMAP-*Fuzzy* segue em consonância como apresentado nas Seções 5.3 e 5.4. Uma vez concluída esta fase, a rede passará para a fase da previsão. Durante a previsão, executa-se o treinamento continuado usando-se as informações de carga fornecidas pelo sistema de aquisição de dados mais recentes, os quais convêm serem aproveitados, porém, algumas informações podem ser descartadas. Este aproveitamento é gerenciado por uma rotina específica que irá “julgar” a conveniência e a importância sobre a janela de carga candidata, viz., a confirmação desta janela e a consequente adaptação de pesos nas matrizes W^a , W^b e W^{ab} da rede neural ARTMAP-*Fuzzy*. Esta etapa de realização do treinamento continuado deve ser feita com bastante critério para se evitar o armazenamento de informação nas matrizes de pesos, desnecessariamente, que possam produzir o fenômeno *overtraining* na rede neural. Como o treinamento continuado é realizado simultaneamente à últimação da previsão de carga, qualquer que seja a janela candidata, pode ser fornecida a partir da base de dados históricos, como reforço, até mesmo por alguma janela projetada por algum processo de identificação, ou substituição de medidas visando melhorar a previsão final. É importante destacar que a execução do treinamento continuado é realizada com baixíssimo tempo de processamento, não prejudicando o rito da efetivação da previsão de carga.

Nesta pesquisa, a continuação do aprendizado (treinamento continuado) é estabelecida a partir da conclusão da fase de treinamento. À medida que se dispõe de uma informação considerada relevante (por exemplo, disponibilizada pelo sistema de aquisição de dados) faz-se a atualização da rede neural (matrizes W^a , W^b e W^{ab}) usando-se os parâmetros de vigilância ρ_b e ρ_{ab} iguais aos usados na fase de treinamento. Isto porque estes parâmetros são arbitrados como valores constantes e devidamente testados a partir de experimentos. Contudo, em relação ao parâmetro ρ_a , adota-se este parâmetro fixo (constante) arbitrado como sendo $\rho_{a_{CL}}$. Este parâmetro de vigilância pode ser arbitrado de várias formas:

Alternativa 1. Proposta idealizada por meio de experimentos. Busca-se a “melhor” sintonia deste parâmetro extraído via repetição de experimentos a serem realizados considerando-se as várias bases de dados disponíveis na literatura técnico-científica.

Alternativa 2. É extraído diretamente da fase de treinamento. Observa-se que cada janela corresponde a um determinado horário discreto diário e a um valor do parâmetro de vigilância ρ_a particular. Cada classe estará associada a um

determinado valor de ρ_a e a uma característica de similaridade. Pode-se, por exemplo, arbitrar $\rho_{a_{CL}}$ igual à média aritmética (ou outra métrica de interesse do usuário) do parâmetro de vigilância ρ_a apurado durante a fase de treinamento. Deve-se prever a possibilidade de não haver neurônios candidatos nos módulos ARTa, e/ou ARTb. Para resolver este impasse, basta adicionar um neurônio para cada necessidade para receber a nova informação. Pode ser necessário criar-se um novo neurônio J na contabilidade da matriz \mathbf{W}^a , enquanto que para o neurônio K não, pois ele encontra correspondência (está inserido) na matriz \mathbf{W}^b , ou vice-versa, ou ainda, havendo a necessidade de ser adicionado um neurônio em cada módulo quando se trata de novas classes em ARTa e ARTb, evidentemente, adicionando-se um neurônio no módulo inter-ART.

Alternativa 3. Usa-se o mesmo esquema adotado na fase de treinamento. No entanto, deve-se buscar uma forma de excluir alguns pares (\mathbf{a}, \mathbf{b}) que não terão a oportunidade de produzir contribuição significativa à melhoria da previsão de carga para evitar a ocorrência do fenômeno *overtraining* (HAYKIN, 2008), assim como evitar o uso de recursos desnecessariamente. Considera-se um padrão adequado para ser descartado na fase do treinamento continuado, como sugestão, o seguinte critério:

$$IS(\mathbf{I}_{Acq}^a, \mathbf{W}_K^a) \geq \rho_{Acq}. \quad (29)$$

sendo:

IS : índice de similaridade;

ρ_{Acq} : parâmetro de vigilância limitante para descarte de padrões referentes à fase de treinamento contínuo;

\mathbf{I}_{Acq}^a : vetor da base de dados advindo do sistema de aquisição devidamente processado usando-se as equações (2) e (4) referente à rede neural ARTa (entrada da rede neural ARTMAP-*Fuzzy*);

\mathbf{W}_K^a : K -ésimo vetor (linha) da matriz \mathbf{W}^a .

Na qualidade do índice de similaridade IS , pode-se usar a função de escolha T_j ,

Jac, ou *sim*, ou, ainda, outro índice de interesse do usuário.

Entende-se que é razoável arbitrar $\rho_{Acq} = 0,9$ (ou outro valor ajustado para esta finalidade). Um padrão (a, b) que possua um valor *IS* superior a 90% deve ser descartado para os propósitos do ajuste dos pesos (W^a , W^b e W^{ab}) da rede neural ARTMAP-*Fuzzy*. Para a execução do treinamento continuado, preferencialmente, consideram-se as janelas decorrentes do sistema de aquisição de dados. Pode-se, também, usar dados artificializados em substituição de informações que, por ventura, estejam sob suspensão (erro de medição, não aprovação em teste de validação, ou que não seguem, por exemplo, a lei de Benford) (BENFORD, 1938). No critério (29), pode-se usar, também, os índices de similaridade T_j , ou *Jac* visando estabelecer qual é a melhor proposta. A lei de Benford tem sido bastante usada no contexto de testes de integralidade dos números contidos em bases de dados.

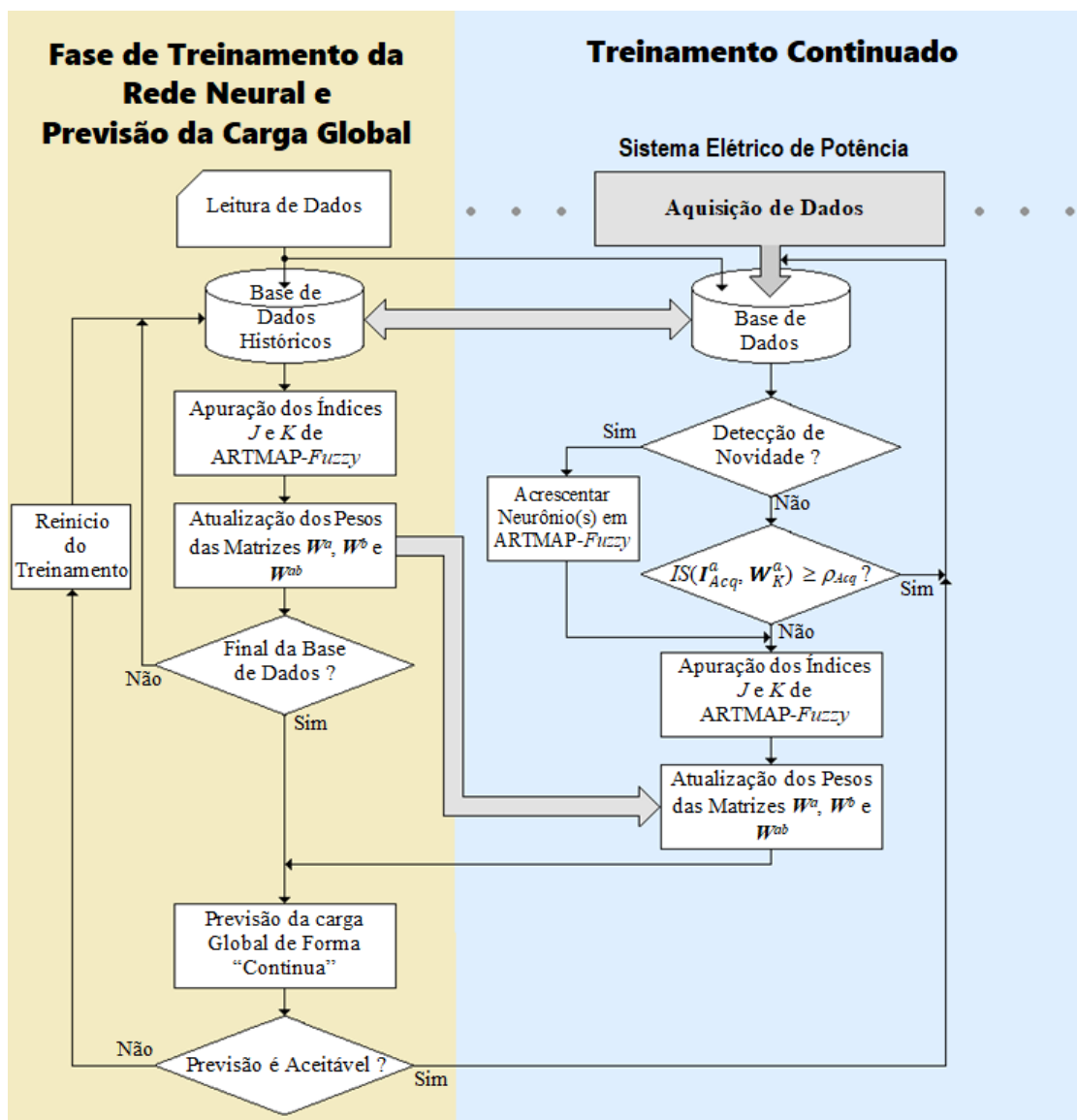
Na maioria das experiências reportadas (formulação convencional) na literatura especializada, é reservado, inicialmente, um número fixo de neurônios nos módulos ARTa, inter-ART e ARTb igual ao número de pares de padrões (a, b) usados na fase de treinamento perfazendo L neurônios. Porém, esta estratégia pode ser substituída por outro procedimento, iniciando-se somente com um neurônio em cada módulo. Um novo neurônio deverá ser criado à medida que for necessário criá-lo. A execução dos passos referentes ao processamento da função escolha será drasticamente reduzida, conseqüentemente, encolhendo o tempo de processamento da fase de treinamento (MORENO, 2008). A execução do treinamento continuado é uma tarefa com baixo custo computacional, além de ser altamente vantajosa para os objetivos da realização da previsão de carga.

5.4. FLUXOGRAMA REFERENTE À PREVISÃO DE CARGA

As informações destinadas à construção do sistema previsor compreendem a arquitetura neural, base de dados históricos e dados originários do sistema de aquisição de dados. A base de dados históricos é usada para fins da realização do treinamento convencional da rede neural, enquanto que as informações advindas do sistema de aquisição são armazenadas em uma unidade específica e usadas exclusivamente para o treinamento continuado. Na Figura 3, apresenta-se o fluxograma relativo à execução das fases do

treinamento convencional, treinamento continuado e da previsão de carga global. No módulo “Leitura de Dados”, além de todos os dados necessários à execução da fase de treinamento, treinamento continuado e a previsão, deve ser fornecida a informação do tempo de retardo para fins da atuação do treinamento continuado.

Figura 3. Fluxograma referente à execução da fase de treinamento, treinamento continuado e da previsão de carga global.



Fonte: Elaborada pela Autora.

Na Figura 3, destaca-se o dispositivo “Detecção de Novidade (DN)”, o qual pode ser entendido, grosso modo, como sendo um mecanismo que visa diferenciar o “díspar (novidade)” do “comum”. O conceito “comum” refere-se às informações já armazenadas nas

matrizes sinápticas (W^a , W^b e W^{ab}) da rede neural ARTMAP-Fuzzy. Como forma de melhor compreender a atuação da DN, pode-se usar um exemplo bastante simples: se desejarmos desenvolver um sistema de segurança domiciliar inteligente para atuar na qualidade de alarme contra a invasão domiciliar, ao invés de treinar o sistema inteligente (por redes neurais, sistemas *fuzzy*, etc.) acompanhando o *modus operandi* do invasor, pode-se optar por um treinamento mais descomplicado e eficiente envolvendo o reconhecimento dos moradores da residência, evidentemente, incluindo-se os animais que habitam no local. Devem-se incluir, também, outros animais, como o gato doméstico, ainda que não habite na residência-alvo, visto que possui a habilidade de se introduzir em locais sem ser convidado, devendo-se considerá-lo no treinamento para evitar a atuação do alarme desnecessariamente. A evolução do *modus operandi* do invasor constitui-se num problema bastante complexo (é um ato contínuo e que requer grande astúcia para a identificação das “novas estratégias” que, decerto, irão acontecer com grande frequência). Este treinamento mais simples pode ser rotulado como sendo a detecção de novidade. Na previsão de carga, a DN pode ser facilmente executada, por meio do descarte de padrões (vetores (a, b) de entrada e de saída da rede neural), durante o treinamento continuado, com certo grau de similaridade que pode ser mais bem desenvolvido via ajuste do valor limítrofe, *e.g.*, de um parâmetro análogo ao parâmetro de vigilância (ρ) e, também, no desenvolvimento de uma função de inferência de similaridade que melhor assimile a referida afinidade, como sugerido anteriormente. Isto pode ser determinado por redes neurais, ou pelo desenvolvimento de uma função, a partir da teoria matemática, harmonizada ao problema-alvo desta pesquisa (previsão de carga).

A inclusão da DN pode ser realizada mediante um grande conjunto de alternativas, as quais se encontram abertas para receber inovações. Para completar o treinamento continuado, baseado no descarte, ou não descarte de padrões, nesta pesquisa, emprega-se, também, outro bloco decisório (vide Figura 3) acolhido via o índice *IS* (equação (29)). Este índice é igualmente uma métrica de similaridade que visa ampliar a capacidade do gerenciamento das informações (padrões) a serem aproveitadas na execução do treinamento continuado com a missão de reforçar o ensinamento. Este bloco decisório, a princípio, possui importância relevante para os casos em que o treinamento continuado é incorporado considerando-se longas durações. A previsão de carga deve ser tratada com excursões de tempo correspondentes a alguns dias à frente (algo não superior a 10, ou 15 dias). Esta abordagem requer uma maior investigação para obter uma conclusão final.

5.5. INFERÊNCIA DA QUALIDADE DE PREVISÃO VIA REDE NEURAL ARTMAP-FUZZY

Para se inferir a qualidade da previsão, usa-se o conceito *MAPE*, o qual é consolidado na literatura especializada (PARK *et al.*, 1991):

$$MAPE = \frac{1}{NA} \sum_{h=1}^{NA} \left| \frac{y_{atual}(h) - y_{fprevisto}(h)}{y_{atual}(h)} \right| \times 100 \quad (30)$$

where:

MAPE : *Mean Absolute Percent Error*;

NA : número da amostragem da série de tempo discreto;

$y_{atual}(h)$: valor de referência (para fins de comparação) relativa ao instante discreto h ;

$y_{previsto}(h)$: valor previsto referente ao instante discreto h .

Considerando esta estrutura, pode-se elaborar um sistema previsor de carga global que seja bastante preciso (atendido um erro máximo permissível em consonância com as exigências do setor elétrico: não superior a 5%) adotando-se vários arranjos com vista a se obter a máxima previsão. Estes arranjos consistem em propostas com diferentes composições da janela em relação à entrada da rede neural ARTa, isto porque a saída corresponde, exclusivamente, à carga referente a uma hora adiante. Deste modo, a melhoria da previsão, basicamente, depende de uma boa escolha da composição do vetor \mathbf{a} . A melhoria da janela refere-se ao aumento dos componentes do histórico de carga em vários horários, dias, meses em atraso à hora corrente. Também, propondo melhor qualidade (mais fidedigna) das informações exógenas.

Ressalta-se que com o uso rede neural ARTMAP-Fuzzy não se identifica nenhum problema em relação ao aumento da dimensão do vetor de entrada \mathbf{a} e, muito menos ainda, em relação ao tamanho da base de dados históricos usada. O aumento da quantidade de pares (\mathbf{a} , \mathbf{b}) representa somente em se dispor um maior contingente de unidades de armazenamento. Versa-se, por conseguinte, de uma rede neural que não impõe limite a sua realização. Assim, pode-se ter liberdade para promover inovações, não se restringindo à formulação basilar que é

frequentemente usada na literatura, isto é, a adoção vetores de entrada bastante reduzida e base histórica pequena para garantir minimamente a convergência durante a execução da fase de treinamento, principalmente, quando se emprega uma rede neural baseada em processos derivativos, por exemplo, o uso da rede neural *feedforward* com treinamento realizado via o algoritmo retropropagação (WERBOS, 1974), Levenberg-Marquardt (MARQUARDT, 1963; SAINI; SONI, 2002) entre outras.

CAPÍTULO 6 - RESULTADOS E DISCUSSÕES

6.1. INTRODUÇÃO

Neste capítulo serão apresentados e discutidos os resultados da previsão de carga obtidos usando a rede neural *ARTMAP-Fuzzy*, contemplando as várias inovações apresentadas nos Capítulos 4 e 5. Serão apresentados experimentos, em especial, visando proporcionar a previsão de carga em tempo mais próximo possível do tempo real. Inclui-se, para fins de treinamento de rede neural, uma base de dados históricos, devidamente dimensionada, com vistas a combiná-la com o uso do treinamento continuado, viz., aproveitando-se as informações colhidas do sistema de aquisição de dados do sistema.

6.2. RESULTADOS E DISCUSSÕES

A base de dados históricos usados nestes experimentos encontra-se disponibilizada na referência CDS (2010), ou seja, de uma companhia do setor elétrico da Nova Zelândia. Compreende os dados referentes às curvas de carga de 9 subestações compondo a carga total. Foram usados para fins do treinamento, treinamento continuado e a previsão, um PC e *software* MATLAB e as suas *toolboxes*. Esta base de dados compreende um período de alguns anos, no entanto, da base utilizou-se o período de 08-Jan-2008 a 07-Jan-2009 (1 ano) por entendimento (conclusão extraída dos experimentos realizados) de ser um intervalo de tempo suficiente para os propósitos desta pesquisa. São fornecidos as curvas de carga de 9 subestações (Cambridge, Hamilton 11 kV, Hamilton 33 kV, Hinuera, Kinleith, Kopu, Te Awamutu, Waihou e Waikino). A carga total é considerada como sendo o somatório de cargas individuais das 9 subestações. Essa carga, representada por uma série temporal, para fins de previsão de carga usando a rede neural *ARTMAP-Fuzzy* sem e com o treinamento continuado, é normalizada ($0 \leq \text{carga normalizada} \leq 1$), cuja base é arbitrada tomando-se o valor máximo de carga no período considerado acrescido em 20%. Nestes experimentos não foram usados os dados de temperatura, em razão de não serem disponibilizados os seus valores. Contudo, podem-se incluí-los sem nenhuma restrição. Basta,

apenas, disponibilizarem-se os valores máximos e mínimos diários da temperatura. A temperatura horária (ou com discriminação mais fragmentada, ou maior resolução) pode ser introduzida usando-se a equação (25).

Para os propósitos previsionais, considera-se o tamanho da base de dados em 6 sistemas: (a) 12 meses; (b) 9 meses; (c) 6 meses; (d) 3 meses; (e) 2 meses e (f) 1 mês. Este experimento foi realizado visando-se observar o comportamento da previsão de carga em função da magnitude de período de treinamento. Para fins do treinamento, foram arbitrados os seguintes valores:

- Parâmetro de vigilância do módulo ARTb : $\rho_b = 1$
- Parâmetro de vigilância do módulo inter-ART : $\rho_{ab} = 0,99$
- Parâmetro de escolha : $\alpha = 0,70$
- Taxa de treinamento : $\beta = 1$

Os valores de α e de β são os mesmos para serem usados nos módulos ARTa e ARTb. Em destaque, usou-se como valor do parâmetro de vigilância do módulo ARTa: $\rho_a = \Phi^{-1} = 0,61803399$. Esta arbitragem foi devidamente argumentada precedentemente.

Para a composição dos dados de entrada de cada janela, foram considerados os seguintes componentes: (1) carga: carga corrente e dois valores referentes às cargas em discretização anteriores [$L(h)$, $L(h-1)$ e $L(h-2)$]. As demais informações (dia, hora e mês) são postas em codificação binária. A saída corresponde ao valor de carga no instante: $L(h+1)$. Na Tabela 1, apresentam-se os resultados referentes ao desempenho do sistema predictor considerando-se os diversos períodos de treinamento da rede neural ARTMAP-*Fuzzy*. Os resultados são apresentados tomando-se o arbítrio dos parâmetros anteriormente relacionados. O período de previsão adotado é de 24 horas à frente sem inclusão do treinamento continuado.

Tabela 1. Desempenho da rede neural ARTMAP-*Fuzzy* para distintos períodos de treinamento SEM o treinamento continuado.

Período de Treinamento (Número de Meses)	MAPE Resultante da Previsão (24 horas adiante)	Número de Neurônios Ativos de ARTa (N_a)	Número de Neurônios Ativos de ARTb (N_b)
12	2,32	13851	97
9	2,79	10412	93
6	3,02	6772	94
3	2,77	3448	82
2	2,96	2063	70
1	2,96	1222	73

Fonte: Elaborada pela Autora.

A partir dos resultados relacionados na Tabela 1, observa-se que o valor do MAPE não se altera significativamente em relação ao tamanho do período considerado no treinamento. Neste sentido, sugere-se um período de treinamento mais adequado de 9 meses em razão de não ser muito extenso e cobrir a maioria dos meses do ano considerando-se a agregação de meses similares, por exemplo, os meses de janeiro e de fevereiro são de comportamento de cargas mensais similares. Assim como observa-se em relação aos meses de junho e julho e assim por diante. Evidentemente, esta é uma sugestão para a previsão de carga de médio, ou de longo prazo. Para horizontes de curto prazo, ou de curtíssimo prazo, o período considerado da base deverá ser bastante estreitado, algo como 30 dias, ou bem menos. Não se pode destacar o uso de bases de dados formadas exclusivamente com informações adquiridas do sistema de aquisição do sistema, notadamente, adotando-se certo atraso. Isto porque o treinamento, via de regra, é realizado usando-se informações passadas visando estimar o consumo de energia no futuro.

Durante a fase de previsão foram observados discrepância da previsão, ou seja, o erro absoluto máximo, em todos os experimentos, em torno de 10%. Este valor representa a possível ocorrência de *spike*, ou, até mesmo, um defeito da previsão. Contudo, este valor pode ser consertado mediante a verificação dissonante entre dois pontos previstos consecutivos. Havendo esta situação, o ponto subsequente (identificado como dissonante) pode ser perfeitamente “corrigido” usando-se algum procedimento heurístico plausível. A ocorrência de *spikes* tem sido observada frequentemente em previsões de carga em relatos da literatura especializada. Nesta pesquisa, este fenômeno tem sido observado somente nos casos sem a

execução do treinamento continuado (vide Figuras 4 e 5).

São apresentados, na Tabela 2, os resultados da previsão de carga para um horizonte de 24, 48, 72 e 96 horas à frente procedentes do treinamento da rede neural ARTMAP-*Fuzzy* por um período de 9 meses com discretização de 30 minutos, correspondendo a 48 amostragens diárias, SEM ou COM o treinamento continuado.

Tabela 2. Resultados da previsão de carga de até 4 dias (96 horas) à frente para um período de treinamento de 9 meses SEM, ou COM a execução do treinamento continuado (TC).

Tempo de Previsão (horas)	MAPE Resultante	
	SEM o TC	COM o TC
24	2,79	1,53
48	3,33	1,29
72	3,13	1,11
96	2,83	1,08

Fonte: Elaborada pela Autora.

Na Tabela 3 são apresentados os resultados da previsão de carga também para um horizonte de 24, 48, 72 e 96 horas à frente, contudo para um período de treinamento de 1 mês e a mesma discretização da previsão anterior.

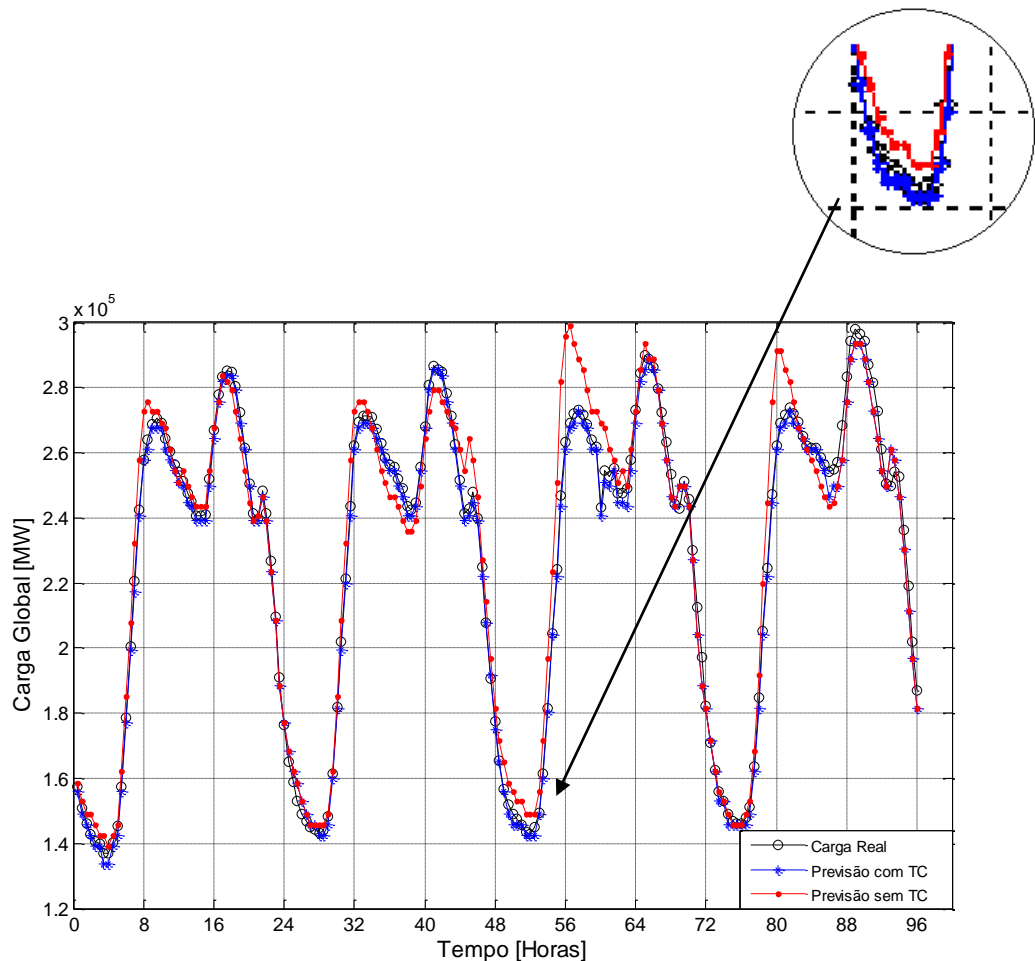
Tabela 3. Resultados da previsão de carga de até 4 dias (96 horas) à frente para um período de treinamento de 1 mês SEM, ou COM a execução do treinamento continuado (TC).

Tempo de Previsão (horas)	MAPE Resultante	
	SEM o TC	COM o TC
24	2,96	1,1
48	3,48	0,83
72	3,29	2,22
96	2,89	1,87

Fonte: Elaborada pela Autora.

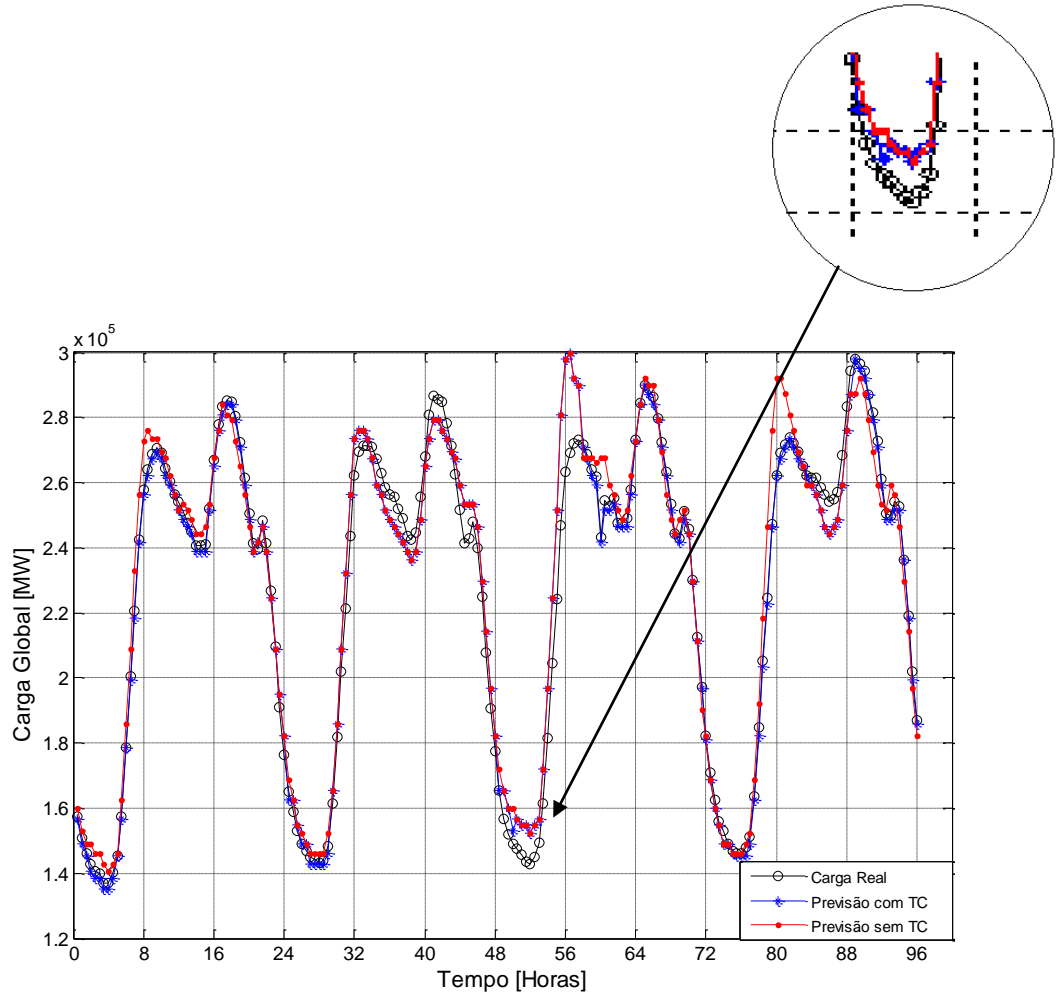
Na Figura 4 ilustram-se as previsões (curvas de carga) para um período de 96 horas, considerando-se um período de treinamento da rede neural ARTMAP-*Fuzzy* de 9 meses SEM, COM a execução do treinamento continuado, assim como da curva real para fins comparativos. Na Figura 5 são apresentadas as previsões de carga considerando-se o mesmo experimento, contudo para um período de treinamento de 1 mês. Ressalta-se que o fracionamento de tempo é de meia hora, perfazendo 48 pontos previstos diariamente.

Figura 4. Curvas de carga resultantes da previsão via rede neural ARTMAP-*Fuzzy* com treinamento realizado por um período de 9 meses SEM, ou COM a atuação do treinamento continuado: resultado relativo a 96 horas (4 dias) à frente.



Fonte: Elaboração pela própria Autora.

Figura 5. Curvas de carga resultantes da previsão via rede neural ARTMAP-Fuzzy com treinamento realizado por um período de 1 mês SEM, ou COM a atuação do treinamento continuado: resultado relativo a 96 horas (4 dias) à frente.



Fonte: Elaboração pela própria Autora.

A partir destes resultados, observa-se que com a aplicação do treinamento continuado houve um incremento de qualidade do parâmetro MAPE para todos os horizontes considerados (previsão entre 24 e 96 horas à frente), se comparados aos resultados onde não foi atuado o dispositivo referente ao treinamento continuado. Nota-se que o MAPE dos casos sem atuação do treinamento continuado recebeu acréscimo à medida que o horizonte da previsão aumentou. Diferentemente, com o treinamento continuado, o MAPE se manteve em torno de 2%, sendo que a maioria fica em torno de 1% para todos os períodos de previsão.

No contexto da previsão de carga, valores acima da curva real, são menos críticos

se comparados aos valores abaixo da curva real quando se trata da análise de segurança, pois se está vislumbrando as situações mais severas, atuando-se de um modo mais conservador. A análise de segurança é uma rotina que visa garantir a integridade do sistema elétrico.

Destas experiências, é importante destacar que, com a implantação de um sistema previsor que contemple o treinamento continuado, por certo, os resultados da previsão serão bastante melhores, em comparação aos modos habituais (sem atuação do treinamento continuado) em dois pontos principais: (1) redução do tempo de treinamento e (2) erro (MAPE) da previsão é muitíssimo baixo ($\leq 2\%$). A redução do tempo de treinamento, combinado com o desenvolvimento de métodos modernos de análise de sistemas elétricos que conseguem executar tarefas de forma rápidas, pode proporcionar aplicações excepcionalmente próximas do tempo real. Trata-se, por conseguinte, de uma ferramenta bastante adequada para emprego nos modernos sistemas *smart grids*.

CAPÍTULO 7 – CONCLUSÃO E SUGESTÕES PARA TRABALHOS FUTUROS

7.1. CONCLUSÃO

Neste texto foram apresentados, formulados e discutidos os detalhes sobre um sistema previsor inteligente de carga global de sistemas elétricos de potência. Para tanto, o objeto-alvo desta pesquisa constitui-se no uso de uma estrutura baseada em uma rede neural ART (*Adaptive resonance Theory*) descendente para a extração do conhecimento na modalidade supervisionada. Trata-se de um conjunto de redes neurais de funcionalidade simbólica que, a princípio, são passíveis de receberem inovações de acordo com a habilidade dos seus desenvolvedores. Esta habilidade advém de algumas propriedades importantes como: estabilidade plena, plasticidade, cálculos matemáticos simples, fácil adaptabilidade à resolução de problemas diversificados, abordagem com dados analógicos e binários simultaneamente, assim como baixo processamento computacional. Isto possibilita a ampliação de capacidade de realização. Com estas propriedades, foi apresentado um sistema previsor de carga que contempla a extração do conhecimento a partir de informações disponíveis de modo contínuo (treinamento contínuo). Esta possibilidade representa uma inovação – no contexto da previsão de carga de sistemas elétricos – em especial, possibilitando realizar a previsão de carga com grande proximidade do tempo real. Compreende, portanto, um ferramental compatível com as aplicações exigidas por sistemas modernos, tal como os *smart grids*.

O sistema proposto destina-se ao trato no âmbito de previsão de carga global. Esta modalidade é afeita como um instrumento para ser empregado no planejamento da expansão, assim como do planejamento da operação. Neste caso, esta previsão visa oferecer as informações indispensáveis para proverem os recursos energéticos (geração própria e/ou compra nos mercados especializados) com vistas ao atendimento pleno da demanda (carga). A previsão de carga pode ser tratada, também, mediante a estimativa de cargas nos pontos de interesse no sistema, por exemplo, nos diversos barramentos (alimentadores, transformadores etc.), os quais são fundamentais para se conhecer o estado do sistema, *sine qua non*, não se

pode garantir o atendimento com segurança e qualidade aos seus usuários, por meio de variados estudos como: fluxo de potência, análise de estabilidade estática e transitória, entre outros. Numa abordagem recente sobre a previsão de carga multinodal (ALTRAN, 2010; NOSE-FILHO, 2001; AMORIM *et al.*, 2020) foi proposto um sistema neural com o propósito de realizar a divisão de carga total do sistema nos referidos barramentos. Daí pode-se destacar a importância de um sistema de previsão de carga total permeando o atendimento de carga do sistema e as divisões (curvas de cargas) individuais. Grosso modo, pode-se afirmar que a previsão multinodal é um subproduto da previsão global. A exigência de qualidade dos resultados (precisão, rapidez e confiabilidade) da previsão multinodal passa, igualmente, pela exigência de qualidade da previsão global, ainda que sejam instrumentos de aplicações distintas, como argumentado anteriormente.

Visando testar a metodologia proposta, foram realizados vários experimentos considerando-se uma base de dados disponibilizada por uma empresa do setor elétrico da Nova Zelândia (CDS, 2010). Esta base de dados foi adotada em razão, evidentemente, pela disponibilidade sem grandes complicações, em especial, pois são oferecidas, também, as curvas de cargas históricas e demais dados importantes por barramento. Com estas informações, pode-se realizar um amplo conjunto de experimentos sobre a previsão global. Os resultados obtidos podem ser considerados satisfatórios (MAPE inferior a 5% que atende as exigências do setor elétrico), com a vantagem associada à confiabilidade, rapidez e ampla flexibilidade para receber aperfeiçoamentos. A inclusão do mecanismo do treinamento continuado tem proporcionado um MAPE médio em torno de 2%.

7.2. SUGESTÕES PARA TRABALHOS FUTUROS

Nesta pesquisa foi apresentada uma nova formulação de resolução do problema referente à previsão de carga global de sistemas elétricos contemplando o treinamento continuado (*continuous learning*). Contudo, trata-se de um projeto aberto à inovações. Além disto, sugerem alguns temas para trabalhos futuros, quais sejam:

- investigar os pontos destacados na Seção 4.5., que não foram plenamente explorados nesta pesquisa de doutorado;
- aperfeiçoar o mecanismo referente ao treinamento continuado como o propósito de

- trabalhar com a previsão de carga muito próxima do tempo real;
- investigar novas concepções de redes neurais da família ART com vistas à obtenção de melhorias, em especial, em relação à precisão;
 - uma vez atingido um desempenho altamente satisfatório, agregar, ao sistema previsor global, a previsão multinodal contemplando, por exemplo, o modelo desenvolvido na referência Amorim *et al.* (2020);
 - explorar a possibilidade de desenvolver metodologias híbridas de redes neurais ART-descendentes, ou outras arquiteturas, com o emprego das técnicas baseadas na fusão de dados (BARNETT, 2008), a qual se constitui uma forma eficiente de trabalhar com informações das mais variadas formas (analógicas, binárias, bem como outros tipos de informação) visando atingir a máxima maturidade do sistema previsor;
 - a melhoria do desempenho da rede neural ARTMAP-*Fuzzy* que está intimamente associada a uma boa escolha de parâmetros (de vigilância, de escolha e taxa de treinamento) via emprego de redes neurais, por exemplo, a rede neural *feedforward*-retropropagação;
 - desenvolvimento de novas formas para a realização da detecção de novidade.

REFERÊNCIAS

- ABREU, T.; AMORIM, A. J.; SANTOS-JUNIOR, C. R.; LOTUFO, A. D. P.; MINUSSI, C. R. Multinodal load forecasting for distribution systems using a ARTMAP-Fuzzy neural network. **Applied Soft Computing**, [s. l.], v. 71, p. 307-316, 2018.
- ALTRAN, A. B. **Sistema inteligente para previsão de carga multinodal em sistemas elétricos de potência**. 2010. 87 f. Tese (Doutorado em Engenharia Elétrica) - Departamento de Engenharia Elétrica, Universidade Estadual Paulista, Ilha Solteira, 2010.
- AMORIM, A. J.; ABREU, T.; TONELLI-NETO, M.S. and MINUSSI, C.R. A new formulation of multinodal short-term load forecasting based on adaptive resonance theory with reverse training. **Electric Power Systems Research**, [s. l.], v. 179, 2020.
- BAKKER, V. **TRIANA: A control strategy for smart grids: forecasting, planning & real-time control**. 2011. Thesis (Master) - Centre for Telematics and Information Technology, University of Twente, Enschede, Netherlands, 2011.
- BARNETT, J. A. Computational methods for a mathematical theory of evidence. In: YAGER, R. R.; LIU, L. (ed.). **Classic works of the Dempster-Shafer theory of belief functions**. Berlin: Springer, 2008. p. 197–216.
- BENDER, E. A. Mathematical methods in artificial intelligence. **IEEE Computer Society Press**, Los Alamitos, California, 1996.
- BENFORD, F. The law of anomalous numbers. **Proceedings of the American Philosophical Society**, [s. l.], v. 78, n. 4, p. 551–572, 1938.
- BIANCHI, F. M.; MAIORINO, M.; KAMPFMEYER, M. C.; RIZZI, A. and JENSSEN, R. **Recurrent neural networks for short-term load forecasting: An overview and comparative analysis**. [S. l.]: Springer, 2017.
- BISHOP, C. M. Novelty detection and neural network validation. **IEE Proceedings Vision, Image and Signal Processing**, [s. l.], v. 141, Issue: 4, p. 217-222, 1994.
- BOX, G. E. and JENKINS, G. M. **Time series analysis: Forecasting and control**. San Francisco: Holden-Day, 1976.
- BUCHHOLZ, B. M. and STYCZYNSKI, Z. **Smart grids: fundamentals and technologies in electric power systems of the future**. 2. ed. [S. l.]: Springer, 2020.
- CARPENTER, G. A. and GROSSBERG, S. A self-organizing neural network for supervised learning, recognition and prediction. **IEEE Communications Magazine**, [s. l.], v. 30, n. 9, p. 38-49, 1992.
- CDS- Centralized Dataset. **Electricity commission for the load dataset**. New Zealand, 2010.
- CHEN, P. Y., and POPOVICH, P. M. **Correlation: parametric and nonparametric measures**. Thousand Oaks, CA: Sage Publications, 2002.

- CZARNECKI, W. M. Weighted Tanimoto extreme learning machine with case study in drug discovery. **IEEE Computational Intelligence Magazine**, [s. l.], v. 10, n. 3, p. 19-29, 2015.
- DAGHER, I. L-p Fuzzy ARTMAP neural network architecture. **Soft Computing**, [s. l.], v. 10, p. 649-656, 2006.
- DASGUPTA, D. Artificial neural networks and artificial immune systems: similarities and differences. **IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation**, Orlando, 6 p., 1997.
- DASGUPTA, D. **Artificial immune systems and their applications**. 1. ed. [S. l.], Springer, 1998.
- DAUBECHIES, I. **Ten lectures on wavelets**. Philadelphia: Society for Industrial and Applied Mathematics, 1992.
- EBERHART, R. and KENNEDY, J. A new optimizer using particle swarm theory. **Proceedings of the Sixth International Symposium on Micro Machine and Human Science-MHS'95**, [s. l.], p. 39- 43, 1995.
- GEORGIOPOULOS, M.; FERNLUND, H.; BEBIS, G. and HEILEMAN, G. L. Order of search in *fuzzy* ART and *fuzzy* ARTMAP: effect of the choice parameter. **Neural Networks**, [s. l.], v. 9, n. 9, p. 1541-1559, 1996.
- GOLDBERG, D. E. **Genetic algorithms in search, optimization, and machine learning**. 1. ed. [S. l.], Addison-Wesley Professional, 1989.
- GOODFELLOW, I.; BENGIO, Y. and COURVILLE, A. **Deep learning**: adaptive computation and machine learning series. [S. l.], The MIT Press, 2016.
- GOODALL, D. W. A New similarity index based on probability. **Biometrics**, [s. l.], v. 22, n. 4, p. 882-907, 1966.
- GROSS, G. and GALIANA, F. D. Short term load forecasting. **Proceedings of the IEEE**, [s. l.], v. 75, n. 12, p. 1558-1573, 1987.
- GROSSBERG, S. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. **Neural Networks**, [s. l.], v. 37, p.1-47, 2013.
- GÜNNEL, S. and TÖDTER, K-H. Does benford's law hold in economic research and forecasting? **Empirica**, [s. l.], v. 36, p. 273-292, 2009.
- HAUPT, S. E.; CASADO, M. G.; DAVIDSON, M.; DOBSCHINSKI, J.; DU, P. LANG, M.; MILLER, T; MÖHRLLEN, CI. MOTLEY, A.; PESTANA and ZACK, J. The use of probabilistic forecasts: Applying them in theory and practice. **IEEE Power & Energy Magazine**, [s. l.], v. 17, n. 6, p. 46-57, 2019.
- HAYKIN, S. **Neural networks and learning machines**. 3. ed. New Jersey, USA: Prentice-Hall, 2008.

HE, W. Load forecasting via deep neural networks. **Procedia Computer Science**, [s. l.], v. 122, p. 308-314, 2017.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the National Academy of Sciences of the United States of America (PNAS)**, [s. l.], v. 79, n. 8, p. 2554-2558, 1982.

KHAN, M.R. & ABRAHAN, A. Short term load forecasting models in Czech Republic using soft computing paradigms. **Computer Science**, Published in ArXiv 2004, 8 p., 2004.

KOHONEN, T. **Self-organizing maps**. 3. ed. [S. l.], Springer, 2000.

LINVILL, D. E. Calculating chilling hours and chill units from daily maximum and minimum temperature observations. **HortScience**, [s. l.], v. 25, n.1, p. 14-16, 1990.

LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistema de energia elétrica**. 2005. 149 f. Tese (Doutorado em Engenharia Elétrica) – Departamento de Engenharia Elétrica, Universidade Estadual Paulista, Ilha Solteira, 2005.

LOPES, M. L. M.; MINUSSI, C. R.; LOTUFO, A. D. P. Electric load forecasting using a fuzzy ART&ARTMAP neural network. **Applied Soft Computing**, Amsterdam, v. 5, n. 2, p. 235–244, 2005.

LOTUFO, A. D. P.; LOPES, M. L. M. and MINUSSI, C. R. Sensitivity analysis by neural networks applied to power systems transient stability. **Electric Power Systems Research**, [s. l.], v. 77, n. 7, p. 730-738, 2007.

MAASS, W.; NAHTSCHLÄGER, T. and MARKRAM, H. Real-time computing without stable states: A new framework for neural computation based on perturbations. **Neural Computation**, [s. l.], v. 14, n. 11, p. 2531–2560, 2002.

MARCHIORI, S. C.; da SILVEIRA, M. C. G.; LOFUFO, A. D. P.; MINUSSI, C. R. and LOPES, M. L. M. Neural network based on adaptive resonance theory with continuous training for multi-configuration transient stability analysis of electric power systems. **Applied Soft Computing**, [s. l.], v. 11, n.1, p. 706-715, 2011.

MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. **SIAM Journal on Applied Mathematics**, [s. l.], v. 11, n. 2, p. 431-441, 1963.

MARTÍN-MERINO, M. and ROMÁN, J. Electricity load forecasting using self organizing maps. In: International Conference on Artificial Neural Networks – ICANN-2006. **Lecture Notes in Computer Science**, [s. l.], v. 4132, p. 709-716, 2006.

MORENO, A. L. **Análise da estabilidade transitória via rede neural Art-Artmap fuzzy Euclidiana modificada com treinamento continuado**. 2010. 143 f. Tese (Doutorado em Engenharia Elétrica) – Departamento de Engenharia Elétrica, Universidade Estadual Paulista, Ilha Solteira, 2010.

MORZY, M.; KAJDANOVICZ, T. and SZYMANSKI, B. K. Benford's distribution in complex networks. **Scientific Report**, [s. l.], v. 6, p. 1-7, 2016.

- NGUYEN, H. and HANSEN, C. K. Short-term electricity load forecasting with time series analysis. **IEEE International Conference on Prognostics and Health Management**, Dallas, TX, p. 214-221, 2017.
- NOSE-FILHO, K. LOTUFO; A. D. and MINUSSI; C. R. Short-term multinodal load forecasting using a modified general regression neural network. **IEEE Transactions on Power Delivery**. New York, v. 26, n. 4, p. 2862-2869, 2011.
- O'DONOVAN, T. M. **Short term forecasting**: an introduction to the Box-Jenkins approach. New York: John Wiley & Sons, 1983.
- PARK, D. C.; EL-SHARKAWI, M. A.; MARKS II, R. J.; ATLAS, E.; DAMBORG, M. J. Electric load forecasting using an artificial neural networks. **IEEE Transactions on Power Systems**, New York, v. 6, n. 2, p. 442-449, 1991.
- PAWLAK, Z. Rough sets. **International Journal of Computer and Information Sciences**, Ontario, v. 11, p. 341-356, 1982.
- ROSEN, R. **Anticipatory systems**: philosophical, mathematical & methodological foundations. Oxford: Pergamon Press, 1985.
- SAINI, L. M.; SONI, M. K. Artificial neural network based peak load forecasting using Levenberg-Marquardt and Quasi-Newton methods. **IEEE Proceedings on Generation, Transmission & Distribution**, New York, v. 149, n. 5, p. 578- 584, 2002.
- SHAFER, G. **A Mathematical theory of evidence**. [S. l.], Princeton University Press, 1976.
- TANIMOTO, T. T. **An elementary mathematical theory of classification and prediction**. New York: International Business Machines Corporation, 1958.
- THOMAS, M. S. and McDONALD, J. D. **Power system SCADA and smart grids**. 1. ed. Boca Raton, FL: CRC Press, 2015.
- ULLAH, M. I. **Time Series Analysis**: Basic Statistics and Data Analysis. [S. l.], WEN Themes, 2014.
- VOROB'EV, N. N. **Fibonacci numbers**. New York: Pergamon Press, 1961.
- WERBOS, P. J. **Beyond regression**: new tools for prediction and analysis in the behavioral sciences. 1974. Thesis (Master) - Harvard University, Cambridge, 1974.
- ZADEH, L. Fuzzy sets. **Information and Control**, New York, v. 8, p. 338-353, 1965.
- ZHU, W.; XU, Y.; WU, Y. and SUN, Y. Determination of evidence correction factors based on the neural network. **Expert Systems**, [s. l.], v. 34, p. 1-9, 2017.