

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE ENGENHARIA
CAMPUS DE ILHA SOLTEIRA**

MARCOS RICARDO MÜLLER

**ANÁLISE DE DESEMPENHO DA REDE NEURAL ARTIFICIAL ARTMAP
FUZZY APLICADA PARA PREVISÃO MULTI-STEP DE CARGAS
ELÉTRICAS EM DIFERENTES NÍVEIS DE AGREGAÇÃO**

Ilha Solteira

2018

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MARCOS RICARDO MÜLLER

**ANÁLISE DE DESEMPENHO DA REDE NEURAL ARTIFICIAL ARTMAP
FUZZY APLICADA PARA PREVISÃO MULTI-STEP DE CARGAS
ELÉTRICAS EM DIFERENTES NÍVEIS DE AGREGAÇÃO**

Tese apresentada à Faculdade de Engenharia de
Ilha Solteira – UNESP como parte dos requisitos
para obtenção do título de Doutor em Engenharia
Elétrica.

Área de Conhecimento: Automação

Anna Diva Plasencia Lotufo

Orientadora

Ilha Solteira

2018

FICHA CATALOGRÁFICA

Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

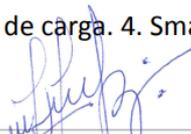
Müller, Marcos Ricardo.

M958a Análise de desempenho da rede neural artificial ARTMAP fuzzy aplicada para previsão multi-step de cargas elétricas em diferentes níveis de agregação / Marcos Ricardo Müller. -- Ilha Solteira: [s.n.], 2018
97 f. : il.

Tese (doutorado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2018

Orientador: Anna Diva Plasencia Lotufo
Inclui bibliografia

1. Redes neurais artificiais. 2. ARTMAP fuzzy. 3. Previsão de carga. 4. Smart Grid.



João José Barbosa
Serviço Técnico de Biblioteca e Documentação
Diretor Técnico
CRB 8-5642

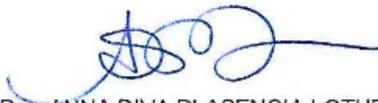
CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: Análise de Desempenho da Rede Neural Artificial ARTMAP Fuzzy Aplicada para Previsão Multi-Step de Cargas Elétricas em Diferentes Níveis de Agregação

AUTOR: MARCOS RICARDO MÜLLER

ORIENTADORA: ANNA DIVA PLASENCIA LOTUFO

Aprovado como parte das exigências para obtenção do Título de Doutor em ENGENHARIA ELÉTRICA, área: AUTOMAÇÃO pela Comissão Examinadora:



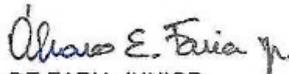
Profa. Dra. ANNA DIVA PLASENCIA LOTUFO
Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira



Prof. Dr. EDGAR MANUEL CARREÑO FRANCO
Departamento de Engenharia Elétrica / Universidade Estadual do Oeste do Paraná



Profa. Dra. MARA LUCIA MARTINS LOPES
Departamento de Matemática / UNESP/ Câmpus de Ilha Solteira



Prof. Dr. ALVARO EDUARDO DE FARIA JUNIOR
School Of Mathematics and Statistics / School Of Mathematics and Statistics Faculty of Science Technology Engineering and Mathematics



Profa. Dra. HUEI DIANA LEE
Centro de Ciências Exatas e Tecnológicas / Universidade Estadual do Oeste do Paraná

Ilha Solteira, 26 de fevereiro de 2018

DEDICATÓRIA

Dedico este trabalho a minha Mãe Marinês e a meu Pai João (*in memoriam*), com todo o meu amor e gratidão.

AGRADECIMENTOS

Agradeço ao Senhor do universo, meus mentores espirituais e a todos que amo e que já transcenderam ao plano espiritual.

Agradeço ao meu irmão Bruno e a minha família por todo apoio, amor, aprendizado e momentos inesquecíveis compartilhados. Em especial agradeço a minha amada Mãe, por ser fortaleza, amparo, carinho e exemplo.

Thai e Pri, obrigado pelo amor que trouxeram para minha vida.

Agradeço aos colegas, amigos e professores por terem sido parte importante desta jornada. Agradeço especialmente a Marleidinha, Thais, Tânia, Monara, Gionei e Junior pelos conhecimentos compartilhados, horas de estudo, conversas, *coffee breaks* e amizade.

Em especial agradeço a minha orientadora Profa. Anna Diva, pela oportunidade, confiança, auxílio e dedicação durante estes anos. Ao Prof. Edgar por ter me recebido em Foz, pelas orientações oferecidas e por ter proporcionado condições para conclusão deste trabalho. A Profa. Mara, por todas as orientações, por estar sempre disposta a ajudar, pelos conselhos e incentivos. A Profa. Huei, pelos livros emprestados e conhecimentos compartilhados.

Agradeço a CAPES pela concessão da bolsa de estudos.

Agradeço ainda ao SINTEL, PPGEE e FEIS por proporcionarem a estrutura para a concretização deste trabalho.

“Debulhar o trigo
Recolher cada bago do trigo
Forjar no trigo o milagre do pão
E se fartar de pão

Decepar a cana
Recolher a garapa da cana
Roubar da cana a doçura do mel
Se lambuzar de mel

Afagar a terra
Conhecer os desejos da terra
Cio da terra propícia estação
E fecundar o chão”

O Cio da Terra, Milton Nascimento &
Chico Buarque

RESUMO

A maior inserção de tecnologias da informação nas redes de distribuição de energia elétrica vem permitindo que maiores volumes de dados de consumo sejam capturados em níveis cada vez mais detalhados, menos agregados e com maiores resoluções. Com a evolução dos mercados de energia elétrica, esses tipos de dados alcançam maior importância, uma vez que a comercialização de energia também passa a considerar estes níveis de consumo. Diversas técnicas têm sido aplicadas para previsão de cargas elétricas, como modelos estatísticos, de inteligência computacional e híbridos. Na literatura especializada é possível encontrar trabalhos que aplicam a rede neural artificial ARTMAP *Fuzzy* para tarefas de previsão de cargas elétricas, no entanto, a técnica ainda é pouco explorada em cenários de consumo menos agregados, e com maiores níveis de detalhe. Neste trabalho a rede ARTMAP *Fuzzy* é aplicada em tarefas de previsão *multi-step* de cargas elétricas reais com distintos níveis de agregação. Considerando o impacto do ruído sobre os previsores, sobretudo na capacidade de generalização das redes neurais artificiais, a técnica *singular spectrum analysis* é aplicada na tarefa de remoção de ruído. Os resultados de previsão permitiram analisar desempenho da rede ARTMAP *Fuzzy*, que foi comparada com outros dois previsores utilizados como *benchmark*, a saber, *seasonal autoregressive integrated moving average* e a rede neural *multiLayer perceptron*. A remoção de ruído permitiu melhora nos níveis de generalização da rede, impactando positivamente sua capacidade preditiva.

Palavras-chave: Redes neurais artificiais. ARTMAP *Fuzzy*. Previsão *Multi-Step*. *Smart Grid*. SSA.

ABSTRACT

The increased insertion of information technologies in electricity distribution networks has allowed larger volumes of consumption data to be captured at increasingly detailed, less aggregated and higher resolution levels. With the evolution of electric energy markets, these types of data become more important, since the commercialization of energy also begins to consider these levels of consumption. Several techniques have been applied to predict electrical loads, such as statistical, computational intelligence and hybrids models. In the specialized literature it is possible to find works that apply the artificial neural network ARTMAP Fuzzy for tasks of prediction of electric charges, however, the technique is still little explored in less aggregated consumption scenarios, and with greater levels of detail. In this work the ARTMAP Fuzzy network is applied in multi-step forecasting tasks of real electric loads with different levels of aggregation. Considering the impact of noise on predictors, especially in the generalization capacity of artificial neural networks, the singular spectrum analysis technique is applied in the noise removal task. The prediction results allowed to analyze the performance of the ARTMAP Fuzzy network, which was compared with other two predictors used as benchmark, namely seasonal autoregressive integrated moving average and the multiLayer perceptron neural network. The noise removal allowed an improvement in the levels of network generalization, positively impacting its predictive capacity.

Keywords: Artificial neural networks. Fuzzy ARTMAP. Multi-step forecasting. Smart grid. SSA.

LISTA DE FIGURAS

Figura 1 - Curva de carga de uma <i>microgrid</i> - Janeiro de 2017	30
Figura 2 - Arquitetura da RNA <i>MultiLayer Perceptron</i>	39
Figura 3 - Família de redes ART	45
Figura 4 - Arquitetura da RNA ART <i>Fuzzy</i>	46
Figura 5 - Processos realizados na camada de Entrada - ART <i>Fuzzy</i>	47
Figura 6 - Processos realizados na camada de comparação - ART <i>Fuzzy</i>	48
Figura 7 - Processos realizados na camada de reconhecimento - ART <i>Fuzzy</i>	49
Figura 8 - Fluxograma completo da RNA ART <i>Fuzzy</i>	50
Figura 9 - Arquitetura da RNA ARTMAP <i>Fuzzy</i>	52
Figura 10 - Fluxograma completo da rede ARTMAP <i>Fuzzy</i>	55
Figura 11 - Curva de intercâmbio de energia elétrica – Fevereiro	58
Figura 12 - Curvas de carga das subestações Kopu, Waikino, Waihou e Combinadas .	61
Figura 13 - Curva de carga com 48 e 96 dados	69
Figura 14 - Curva de carga semanal com 96 dados	69
Figura 15 - Curva de carga incluindo dias atípicos influenciados por feriado nacional	70
Figura 16 - Previsão de 07 dias a frente com dados originais - Curva de carga com 96 dados.....	72
Figura 17 - ARTMAP <i>Fuzzy</i> com previsão utilizando dados originais e suavizados por SSA.....	74
Figura 18 - Dados históricos dos 05 prédios	76
Figura 19 - Curvas de carga originais e submetidas à remoção de ruído por SSA	77
Figura 20 – Representação de um neurônio biológico	95
Figura 21 - Modelo de neurônio artificial de McCulloch e Pitts.....	96

LISTA DE TABELAS

Tabela 1 - Parâmetros adotados para a rede MLP	43
Tabela 2 - Parâmetros Testados para a RNA ARTMAP <i>Fuzzy</i>	59
Tabela 3 - Resultados de MAPE para horizonte preditivo de 24 horas - Intercâmbio de EE	59
Tabela 4 - Resultados de MAPE para horizonte preditivo de 72 horas - Intercâmbio de EE	59
Tabela 5 - Divisão dos dados em subconjuntos – Subestações de EE	62
Tabela 6 - Parâmetros testados para a RNA ARTMAP <i>Fuzzy</i>	62
Tabela 7 - Resultados de previsão – Subestações de EE agregadas	63
Tabela 8 - Parâmetros testados para a RNA ARTMAP <i>Fuzzy</i>	65
Tabela 9 - Resultados de previsão com os subconjuntos de dados da subestação Kopu	66
Tabela 10 - Resultados de previsão com os subconjuntos de dados da subestação Waikino	66
Tabela 11 - Resultados de previsão com os subconjuntos de dados da subestação Waihou	67
Tabela 12 - Resultados de previsão da rede ARTMAP <i>Fuzzy</i> com os subconjuntos de dados submetidos a remoção de ruído - Kopu, Waikino e Waihou	67
Tabela 13 - Média do MAPE para as subestações Waihou, Waikino, Kopu e Agregadas	68
Tabela 14 - Parâmetros testados para a RNA ARTMAP <i>Fuzzy</i> - Microrrede.....	71
Tabela 15 - Resultados de previsão com CC de 48 dados - Microrrede	71
Tabela 16 - Resultados de previsão com CC de 96 dados - Microrrede	71
Tabela 17 - Parâmetros testados para a RNA ARTMAP <i>Fuzzy</i> - Conjunto de prédios .	73
Tabela 18 - Resultados de previsão de um dia a frente - Conjunto de prédios	74
Tabela 19 - Parâmetros testados para a RNA ARTMAP <i>Fuzzy</i> – Nível predial	78
Tabela 20 - Resultados de previsão de um dia à frente - Nível predial.....	78

LISTA DE SIGLAS E ABREVIATURAS

AI	Inteligência Artificial
AIC	Critério de Informação Akaike
ANEEL	Agência Nacional de Energia Elétrica
ARIMA	Autoregressive Integrated Moving Average
ART	<i>Adaptive Resonance Theory</i>
BIC	Critério de Informação Bayesiano
CC	Curva de Carga
DLVQ	<i>Dynamic Learning Vector Quantization</i>
EE	Energia Elétrica
NaN	<i>Not a Number</i>
RBF	<i>Radial Basis Function</i>
RNA	Rede Neural Artificial
SSA	<i>Singular Spectrum Analysis</i>
ST	Série Temporal
STP	<i>Shielded Twisted Pair</i>
SVD	<i>Singular Value Decomposition</i>
wcor	<i>Weighed-correlations</i>
Weka	<i>Waikato Environment for Knowledge Analysis</i>

LISTA DE SÍMBOLOS

F_0	Camada de entrada
F_1	Camada de comparação
F_2	Camada de reconhecimento
a	Vetor de entrada
I	Vetor de atividade da camada F_0
x	Vetor de atividade da camada F_1
y	Vetor de atividade da camada F_2
ρ	Parâmetro de vigilância
β	Taxa de treinamento
α	Parâmetro de escolha
M	Número de neurônios da camada F_0
N	Número de neurônios da camada F_2
T_j	Função de escolha
J	Categoria ativa da camada F_2
w_{ij}	Matriz de pesos
F_0^a	Camada de entrada do módulo ART_a
F_1^a	Camada de comparação do módulo ART_a
F_2^a	Camada de reconhecimento do módulo ART_a
F_0^b	Camada de entrada do módulo ART_b
F_1^b	Camada de comparação do módulo ART_b
F_2^b	Camada de reconhecimento do módulo ART_b
F^{ab}	Camada do módulo <i>Inter</i> – ART
b	Vetor de entrada do módulo ART_b
I^a	Vetor de atividade da camada F_0 do módulo ART_a
I^b	Vetor de atividade da camada F_0 do módulo ART_b
x^a	Vetor de atividade da camada F_1 do módulo ART_a
x^b	Vetor de atividade da camada F_1 do módulo ART_b
x^{ab}	Vetor de atividade do módulo <i>Inter</i> – ART
y^b	Vetor de atividade da camada F_2 do módulo ART_b
ρ_a	Parâmetro de vigilância do módulo ART_a
ρ_b	Parâmetro de vigilância do módulo ART_b

ρ_{ab}	Parâmetro de vigilância do módulo <i>Inter – ART</i>
ε	Valor de incremento do parâmetro ρ_a
K	Categoria ativa da camada F_2 do módulo ART_b
w_j^a	Vetor de pesos do módulo ART_a
w_k^b	Vetor de pesos do módulo ART_b
w_{jk}^{ab}	Matriz de pesos do modulo <i>Inter – ART</i>

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	Objetivos e Contribuições.....	20
1.2	Organização do Texto.....	21
2	REVISÃO BIBLIOGRÁFICA.....	23
3	DEMANDA POR ENERGIA ELÉTRICA.....	26
3.1	Previsão da Demanda de Energia Elétrica.....	26
3.2	Níveis de Agregação.....	27
3.2.1	<i>Impacto na Previsão.....</i>	28
3.3	Fatores que Influenciam a Demanda por Energia Elétrica.....	29
3.4	Horizontes de Previsão em Sistemas de Potência.....	31
3.5	Pré-Processamento dos Dados de Consumo de Energia Elétrica.....	32
3.5.1	<i>Deteção de Inconsistências.....</i>	34
3.5.2	<i>Tratamento de Inconsistências.....</i>	35
3.5.3	<i>Remoção de Ruídos utilizando SSA.....</i>	35
4	REDES NEURAIS ARTIFICIAIS.....	38
4.1	Introdução a Redes Neurais Artificiais.....	38
4.2	Rede Neural Artificial <i>MultiLayer Perceptron</i>	38
4.2.1	<i>Treinamento por Backpropagation.....</i>	39
4.2.2	<i>Implementação Computacional e Treinamento da RNA MLP.....</i>	41
4.3	Redes Neurais Artificiais da Família ART.....	43
4.3.1	<i>Dilema da Estabilidade-Plasticidade.....</i>	43
4.3.2	<i>Teoria da Ressonância Adaptativa.....</i>	44
4.3.3	<i>Redes ART e ART Fuzzy.....</i>	45
4.3.4	<i>Rede ARTMAP e ARTMAP Fuzzy.....</i>	51
5	CENÁRIOS DE ESTUDO E ENSAIOS DE PREVISÃO.....	57
5.1	Intercâmbio de Energia Elétrica.....	57
5.1.1	<i>Cenário Gerador dos Dados.....</i>	57
5.1.2	<i>Ensaio de Previsão.....</i>	58
5.1.3	<i>Considerações.....</i>	60
5.2	Conjunto de Subestações de Energia Elétrica.....	60
5.2.1	<i>Cenário Gerador dos Dados.....</i>	60

5.2.2	<i>Ensaaios de Previsão</i>	61
5.2.3	<i>Considerações</i>	64
5.3	Subestação de Energia Elétrica	64
5.3.1	<i>Cenário Gerador dos Dados</i>	64
5.3.2	<i>Ensaaios de Previsão</i>	64
5.3.3	<i>Considerações</i>	67
5.4	Microrrede	68
5.4.1	<i>Cenário Gerador dos Dados</i>	68
5.4.2	<i>Ensaaios de Previsão</i>	70
5.4.3	<i>Considerações</i>	71
5.5	Conjunto de Prédios	73
5.5.1	<i>Cenário Gerador dos Dados</i>	73
5.5.2	<i>Ensaaios de Previsão</i>	73
5.5.3	<i>Considerações</i>	74
5.6	Nível Predial	75
5.6.1	<i>Cenário Gerador dos Dados</i>	75
5.6.2	<i>Ensaaios de Previsão</i>	75
5.6.3	<i>Considerações</i>	79
6	CONCLUSÕES	80
7	TRABALHOS FUTUROS	83
	REFERÊNCIAS	85
A.1	Modelos Box & Jenkins	92
A.1.2	Implementação Computacional do SARIMA	93
B.1	Redes Neurais Artificiais	95
B.1.1	Modelo Biológico	95
B.1.2	Modelo de McCulloch e Pitts	96

1 INTRODUÇÃO

Este trabalho apresenta uma análise quantitativa e qualitativa do desempenho da rede neural ARTMAP *Fuzzy* aplicada na previsão *multistep* de cargas elétricas em diferentes cenários de consumo, onde são observados distintos níveis de agregação. Tendo em vista a aleatoriedade presente nas séries de demanda por energia elétrica – EE, especialmente nos níveis menos agregados, a técnica *Singular Spectrum Analysis - SSA* é aplicada para remoção de ruído, tendo como objetivo permitir melhores resultados preditivos. Como resultado da aplicação das técnicas descritas, pretende-se expandir a utilização da rede ARTMAP *Fuzzy* em cenários menos agregados, permitindo propor melhorias e modificações alinhadas a estes níveis de consumo. Para poder comparar resultados de previsão, outros dois previsores foram utilizados como *benchmark*, um estatístico, a saber, SARIMA, e outra arquitetura de rede neural artificial, a *MultiLayer Perceptron - MLP*.

Uma série temporal pode ser definida informalmente como um conjunto de medidas de uma variável aleatória de interesse, realizadas de maneira ordenada no tempo. Assim sendo, uma curva de carga - CC pode ser entendida como uma série temporal, de modo que seu comportamento futuro pode ser estimado utilizando técnicas de análise e previsão de séries temporais (MORETIN; TOLOI, 2006).

Com a modernização das redes de distribuição de energia elétrica, através da inserção de tecnologias da informação, conceito conhecido como redes inteligentes (*smart grids*) somado a popularização dos medidores inteligentes de energia (*smart meters*), a captação de dados foi intensificada, permitindo que as séries de consumo de energia pudessem ser menos agregadas, mais extensas, com maior nível de detalhe, e como resultado, com maiores níveis de ruído. Atualmente estão disponíveis dados de consumo de EE de cenários antes ignorados, como consumidores individuais ou pequenos aglomerados (AGARWAL; WENG; GUPTA, 2011). Essas séries apresentam maior aleatoriedade, uma vez que decisões individuais (imprevisíveis) passaram a ter efeito significativo no formato (*shape*) das curvas de carga.

Características não estacionárias e não lineares presentes nas séries de cargas, que são afetadas por efeitos sazonais, condições climáticas, questões econômicas e efeitos aleatórios que interferem na sua previsibilidade, acabam por caracterizar a tarefa de previsão de cargas elétricas como um desafio importante e de interesse por parte dos agentes envolvidos (ZHENG et al., 2017). O planejamento energético se dá considerando a capacidade de se prever adequadamente a demanda por energia elétrica. Uma vez que os previsores sejam robustos e apresentem acurácia, tarefas como o planejamento da expansão, operação econômica, análise de segurança e controle, são impactadas de maneira positiva. Porém, uma projeção inadequada da demanda é capaz de gerar grandes perdas financeiras para os envolvidos. Os comercializadores de energia elétrica perdem ao realizar compras emergenciais de EE a preços elevados, com o desgaste ou perda de ativos, além de estarem sujeitos a pesadas sanções financeiras. Os

consumidores de energia perdem de acordo com a classe de consumo a que pertencem, podendo ocorrer a interrupção de processos produtivos, de serviços, de transporte, comunicação e dados, entre outros.

Os métodos de previsão normalmente encontram-se divididos na literatura especializada em dois grandes grupos: estatísticos e de inteligência computacional. Entre os métodos estatísticos, podem-se destacar os de Box e Jenkins (1970), tais como *Autoregressive Integrated Moving Average* - ARIMA, *seasonal* ARIMA (SARIMA) e ARIMA com variáveis exógenas (ARIMAX). O segundo grupo é representado pelas Redes Neurais Artificiais - RNA, pelos métodos que utilizam Lógica *Fuzzy*, entre outros. Cada método pode ser avaliado por suas vantagens e desvantagens, porém, tal análise é incompleta ou inviável se forem desprezadas as características da série temporal em estudo, assim como, as necessidades do agente ou sistema que dependa desses dados; é necessário considerar a acurácia esperada, custo computacional e tempo de processamento admissíveis, disponibilidade de pessoal, recursos entre outros.

Entre os métodos baseados em inteligência computacional é possível destacar a RNA ARTMAP *Fuzzy*. Esta rede é aplicada a uma gama de problemas de previsão nas mais diversas áreas de estudo. Entre suas vantagens, quando comparada a outros modelos de RNA, destacam-se a aprendizagem rápida, facilidade de parametrização e de codificação (CARPENTER, et al., 1992c). Estas características somadas a sua aprendizagem incremental e estável, se devem a *Adaptive Resonance Theory* - ART, que designa a família de redes que integra. A ART foi concebida com objetivo de superar o dilema da estabilidade-plasticidade, ou seja, permitir que a rede seja capaz de aprender a medida que novos dados são apresentados, e que o conhecimento adquirido a partir dos dados passados seja mantido (CARPENTER; GROSSBERG; ROSEN, 1991; CARPENTER, et al., 1992c; BENITES; SAPOZHNIKOVA, 2017).

Ao realizar a previsão de cargas elétricas em níveis menos agregados, as principais características a serem consideradas incluem a faixa de resolução e o volume de dados, presença de dados atípicos, maior aleatoriedade da série e maiores níveis de ruído.

O ruído se apresenta como um desafio a parte para as tarefas de modelagem e previsão de séries temporais. Para os modelos baseados em inteligência artificial, ele é capaz, de influenciar o processo de aprendizagem, diminuindo a capacidade de generalização, causando *overfitting*. Nos modelos estatísticos, o ruído pode afetar a otimização dos coeficientes e na escolha dos modelos. De tal modo, a remoção prévia do ruído das séries temporais pode trazer vantagens para a tarefa de previsão, se traduzindo em maior acurácia.

Nos cenários de estudo apresentados neste trabalho, que consideram cargas elétricas em diferentes níveis de agregação, o algoritmo SSA foi empregado para tarefa de remoção de ruído. O SSA apresenta algumas vantagens, como a necessidade de escolha de apenas dois parâmetros, a saber, tamanho da janela e os autovetores a serem

utilizados na reconstrução da série. Para a escolha do tamanho da janela não existe uma maneira analítica definida, porém pequenas variações deste parâmetro apresentam pouco impacto no resultado; por outro lado, a escolha dos autovetores pode ser automatizada, utilizando seleção automática baseada em clusterização hierárquica, por exemplo. Diferente dos métodos de filtragem tradicionais, o SSA é adequado tanto para sinais estacionários quanto não estacionários. A codificação da técnica não apresenta grande complexidade, e seu custo computacional é considerado baixo (ALONSO; DEL CASTILLO; PINTADO, 2005; GOLYADINA; NEKRUTKIN; ZHIGLIAVSKY, 2001; ELSNER; TSONIS, 1996).

A remoção prévia de ruído para posterior modelagem não é comum na literatura especializada, principalmente em se tratando de previsão de cargas elétricas. Porém diversos trabalhos, como Nose-Filho, Lotufo e Minussi (2011); Li, Cui e Guo(2014); Niu, et al. (2016); Zhang, Wang e Zhang (2017) tem demonstrado vantagens em se aplicar remoção de ruído, resultando em maior acurácia, melhorando resultados de previsão. Porém, estes trabalhos consideram curvas de carga de menor resolução e com maiores níveis de agregação, de maneira que o presente trabalho se diferencia por considerar dados mais desagregados, com maiores resoluções, e por empregar uma RNA da família ART.

Além do ruído, outro aspecto a ser considerado ao trabalhar com séries de carga, diz respeito as curvas de carga atípicas, que se distinguem significativamente das demais, e, além de não poderem ser adequadamente previstas, prejudicam a previsão de dias típicos (TAYLOR; ARORA, 2018). Estas curvas de carga são provenientes da estrutura da demanda analisada, e podem descrever eventos raros, não sendo determinadas por erros de medição ou quaisquer outros eventos que acarretem a inserção de *outliers* nas séries históricas. As curvas de carga podem ser influenciadas em diferentes níveis a partir da ocorrência de eventos dos mais diversos tipos, como feriados, finais de semana, férias, pontos facultativos, mudanças bruscas de temperatura, ocorrência de precipitação entre outros. Os previsores tradicionalmente empregados não são capazes de incorporar diretamente conhecimento prévio ao modelo, isto é, precisam que determinados tipos de dias sejam assinalados como pertencentes a determinadas classes, para que possam lidar de maneira satisfatória com dias atípicos. Por este motivo, normalmente os dados são submetidos a pré-processamento ou adaptações.

O tratamento dos dias tidos como atípicos não é comum na literatura especializada, de modo que tais dados frequentemente são substituídos na série original por outros tidos como típicos, ou por previsões dentro da amostra para aquele dia ou tipo de dia, transformando o *shape* atípico da curva de carga, em um tido como típico. Um dos maiores problemas nesta abordagem está em desprezar que tais informações supostamente poderiam ajudar a explicar o comportamento futuro da carga (HIPPERT; BUNN; SOUZA, 2005; TAYLOR, 2010).

Nem sempre os eventos que causam dias atípicos são conhecidos ou antecipados, sendo assim, é desejável que o previsor seja capaz de lidar com tais informações. Para investigar o comportamento da RNA ARTMAP *Fuzzy* neste tipo de cenário, um conjunto de dados reais com dias atípicos foi considerado para tarefa de previsão *multistep* de cargas elétricas. Para comparação de resultados, além da RNA ARTMAP *Fuzzy*, utilizou-se o modelo SARIMA e a RNA MLP.

Apesar dos avanços na integração das tecnologias da informação nas redes elétricas, a obtenção de dados reais para estudo ainda é uma tarefa difícil. Neste sentido, o presente trabalho se destaca por lidar com cenários reais com distintos níveis de agregação de carga, submetendo os previsores a desafios não simulados, oriundos de cenários influenciados por todas as volatilidades inerentes ao mundo real. Apresenta ainda a codificação da rede ARTMAP *Fuzzy*, sua aplicação juntamente com a rede MLP e o SARIMA na previsão de cargas elétricas, além da aplicação da técnica SSA na remoção de ruído de séries históricas.

Este trabalho pretende desenvolver uma análise sobre o comportamento da RNA ARTMAP *Fuzzy* na previsão *multistep* de cargas elétricas para diferentes horizontes preditivos, considerando cenários com distintos níveis de agregação e o impacto da remoção de ruído das séries históricas por SSA.

A principal contribuição deste trabalho é a análise de uma rede neural consolidada na previsão de cargas elétricas com altos níveis de agregação, então analisada considerando dados reais de diferentes níveis de agregação, com destaque para os níveis menos agregados e para remoção de ruído via SSA.

1.1 Objetivos e Contribuições

Partindo da hipótese que a RNA ARTMAP *Fuzzy* deva apresentar melhores resultados preditivos considerando dados de consumo de energia elétrica em níveis mais agregados quando comparado com níveis menos agregados, e, considerando ainda que a remoção de ruído por SSA represente maior impacto sobre dados de consumo menos agregados, visto seus níveis mais significantes de ruído, são apresentados os objetivos e contribuições esperadas deste trabalho.

O objetivo principal deste trabalho é analisar o comportamento da RNA ARTMAP *Fuzzy* na previsão *multistep* de cargas elétricas para diferentes horizontes preditivos, considerando cenários com distintos níveis de agregação, onde o impacto da remoção de ruído das séries históricas utilizando SSA também é objeto de investigação.

Objetivos específicos:

- Realizar pré-processamento dos dados, de modo a prepará-los para etapas de treinamento e previsão;

- Codificar a RNA ARTMAP *Fuzzy*;
- Realizar previsão de cargas elétricas em distintos níveis de agregação;
- Comparar os resultados de previsão com outras técnicas estabelecidas (RNA MLP e SARIMA);
- Codificar a técnica SSA e aplicá-la na tarefa de remoção de ruído;
- Realizar previsão com as séries históricas suavizadas por SSA.

Contribuições:

- Investigar o impacto de séries reais de consumo EE com distintos níveis de agregação sobre uma rede ARTMAP *Fuzzy*;
- Investigar o impacto da remoção de ruído sobre a previsão nos distintos níveis de agregação;
- Compilar os resultados de previsão obtidos nos distintos níveis de agregação, de modo a propor novas abordagens.

De modo sintético, este trabalho visa aplicar uma ferramenta de previsão estabelecida para cenários de consumo de energia elétrica altamente agregados, em distintos cenários de consumo, com menores níveis de agregação. Além de analisar deficiências e potencialidades apresentadas pela ferramenta, assim como o impacto da remoção de ruído dos dados históricos sobre sua capacidade preditiva.

1.2 Organização do Texto

Este trabalho está organizado em nove capítulos, descritos brevemente na sequência:

- 2) Revisão Bibliográfica: este contém a revisão bibliográfica a respeito da previsão de cargas elétricas. São apresentados trabalhos que consideram dados de consumo de energia elétrica em diferentes níveis de agregação. O principal enfoque é a utilização de redes neurais artificiais e sistemas híbridos aplicados para previsão de curto prazo de cargas elétricas. A técnica SSA também é referenciada.
- 3) Demanda por Energia Elétrica: conceitua uma curva de carga; aborda os desafios para previsão de cargas elétricas nos diferentes níveis de agregação; apresenta os horizontes preditivos normalmente utilizados, além de abordar o pré-processamento dos dados.

- 4) Redes Neurais Artificiais: introduz brevemente as redes neurais artificiais, apresenta a rede MultiLayer Perceptron, a teoria da ressonância adaptativa, assim como as redes ART, ART *Fuzzy*, ARTMAP e ARTMAP *Fuzzy*.
- 5) Cenários de Estudo e Ensaio de Previsão: neste capítulo são apresentados os cenários que fornecem os dados de consumo de energia elétrica utilizados nesta pesquisa, além de apresentar os métodos utilizados e os resultados de previsão para cada nível de agregação de cargas considerado.
- 6) Conclusões: apresenta as conclusões considerando os estudos realizados e os resultados apresentados no capítulo anterior.
- 7) Trabalhos Futuros: capítulo dedicado a sugestão de trabalhos futuros, aplicação e expansão da pesquisa.
- 8) Referências: lista as referências bibliográficas utilizadas neste trabalho.

2 REVISÃO BIBLIOGRÁFICA

O principal objetivo deste capítulo é apresentar trabalhos que abordam conteúdos correlatos, destacar seus resultados, e contextualizar com a pesquisa realizada. Para tanto, são apresentados preponderantemente trabalhos dedicados a tarefa de previsão de cargas elétricas. De maneira breve são citados trabalhos que empregam técnicas analíticas, e, com maior ênfase são apresentados trabalhos dedicados a previsão de cargas elétricas de curto prazo que utilizam técnicas de inteligência artificial - AI, sobretudo redes neurais artificiais. Também são apresentados trabalhos que aplicam técnicas para remoção de ruído das séries históricas utilizadas nas tarefas de previsão.

É importante destacar que são esperados resultados preditivos menos assertivos nos níveis mais desagregados de consumo de EE, de acordo com o trabalho de (SEVLIAN; RAJAGOPAL, 2017), que salienta que o erro preditivo para consumidores individuais pode facilmente chegar a 30%, enquanto em nível de subestação estes erros são significativamente inferiores, podendo ficar abaixo de 2%.

Entre os trabalhos pertencentes ao grupo de técnicas analíticas, existem os que empregam métodos baseados em regressão linear, como Dudek (2016) que aplica regressão linear e padrões de ciclos diários de demanda na previsão de cargas elétricas, de modo que o uso de tais padrões simplificam a série, permitindo maior acurácia preditiva ao modelo; Yildiz, Bilbao e Sproul (2017) apresentam uma revisão de modelos de regressão aplicados a previsão de cargas elétricas de curto prazo em prédios universitários. Os trabalhos (DAUT et al., 2017; SADA EI et al., 2017) empregam método de Box & Jenkins, enquanto o primeiro considera um conjunto de dados de carga predial, o segundo explora 12 conjuntos de dados com maiores níveis de agregação. Pertencentes a mesma categoria, porém utilizando métodos de regressão não-paramétricos, têm-se (CLEMENTS; HURN; LI, 2016; ANEIROS; VILAR; RANA, 2016) com resultados promissores.

O trabalho de (FAN; CHEN; LEE, 2009) utiliza uma estratégia baseada em coeficientes adaptativos para considerar informações meteorológicas de diversas fontes durante a tarefa de previsão, na sequência apresenta uma RNA aplicada na previsão de cargas de curto prazo em um sistema de potência americano, alcançando bons resultados de previsão.

Em (RAZA; KHOSRAVI, 2015) os autores se dedicaram a revisar vários trabalhos sobre aplicação de redes neurais artificiais na previsão de cargas no curto prazo em *smart grids* e em nível predial, apresentando como conclusão que a melhor combinação para os conjuntos de entrada para as redes neurais envolve dados de carga, variáveis climáticas (incidência ou não de precipitação) e tipos de dia.

O trabalho de (BUI TRAGO; ASFOUR, 2017) apresenta um modelo baseado em RNA para previsão de carga horária de um determinado dia, utilizando para tal tarefa dados de microrredes. Os autores utilizaram como entrada dados de carga do dia

anterior, o dia da semana, o mês e a carga total. Também apresentaram o *Principal Component Analysis* aplicado para detecção de *outliers*, demonstrando ganhos preditivos.

Em (PANDIAN et al., 2006) é apresentado um modelo baseado em lógica *fuzzy*, tendo como objetivo investigar o impacto causado pelas variáveis tempo, temperatura e carga do dia anterior na margem de erro da previsão de carga horária. A base de regras *fuzzy* foi projetada considerando o método de Mamdani (MAMDANI; ASSILIAN, 1975).

O trabalho de (KHOSRAVI; NAHAVANDI; CREIGHTON, 2011) apresenta um modelo de previsão de carga de curto prazo cujo objetivo é diminuir os níveis de incerteza em relação a carga e otimizar o esquema despacho ótimo de EE.

Diversos trabalhos empregam redes neurais da família ART na previsão de cargas elétricas, entre eles, pode-se destacar (LOPES, 2005) que introduz a rede ART&ARTMAP *Fuzzy*, uma arquitetura de rede baseada na teoria da ressonância adaptativa, concebida para o problema de previsão de cargas elétricas a curto prazo, onde foram obtidos bons resultados, sendo possível destacar sua rápida convergência e acurácia preditiva; no trabalho de (OLIVEIRA; LOTUFO; LOPES, 2013) é apresentada uma abordagem híbrida para previsão de cargas elétricas a curto prazo, sendo baseada em duas arquiteturas de redes neurais, a saber, ART *Fuzzy* e MLP; em (ABREU et al., 2017) é apresentada uma rede neural ARTMAP *Fuzzy* com um treinamento continuado, sendo que tal técnica pode ser concebida graças as características de estabilidade e plasticidade da ART, e, diferentemente do treinamento incremental, permite a inclusão de novos padrões à memória da rede continuamente de forma permanente, sendo especialmente útil para cenários altamente mutáveis, dispensando a necessidade de reiniciar o treinamento para que a rede permaneça atualizada; em (GOMES; LOPES; SANTOS JUNIOR; LOTUFO, 2017) uma rede ARTMAP Euclidiana é aplicada para previsão global de cargas elétricas a curto prazo, tendo como uma de suas principais vantagens não necessitar de normalização e complementação dos dados de entrada, permitindo redução das matrizes consideradas, e, conseqüentemente, reduzindo custos computacionais.

O trabalho de (FARIA; SOUZA, 1995) trata da combinação de previsores, levando em conta metodologias Bayesianas, *quasi*-Bayes, e modelo clássico de combinação ótima, aplicadas à combinação linear de previsões de preço do petróleo. O trabalho apresenta descrição metodológica e teórica, além de comparar desempenhos de previsões e combinações.

A prévia remoção de ruído das séries históricas utilizadas para as tarefas de previsão de curvas de carga não é muito recorrente na literatura especializada, porém, ainda assim é possível citar trabalhos que indicam a existência de vantagens em se aplicar este tipo de tratamento prévio aos dados históricos. Por exemplo, o trabalho de Zhang, Wang e Zhang (2017) apresenta bons resultados utilizando *Support Vector*

Regression, otimizado pelo algoritmo *Cuckoo Search*, assim como redução de ruído com SSA. Niu, et al. (2016), realizam previsão de cargas elétricas de curto prazo através de redes neurais MLP não lineares, otimizadas por meta heurísticas, com redução de ruído, também por SSA. Nose-Filho, Lotufo e Minussi (2011) mostram ganho de acurácia preditiva ao aplicar, aos dados, um filtro de médias móveis com corretor de defasagem, sendo utilizado para previsão de carga de subestação com uma rede GRNN. Já Lisi, Nicolis e Sandri (1995) utilizam o filtro SSA para remoção de ruído de séries temporais a serem preditas por redes neurais *feedforward*, conseguindo, assim, melhores resultados em comparação a predição dos dados brutos.

3 DEMANDA POR ENERGIA ELÉTRICA

Este capítulo contextualiza a previsão de séries temporais considerando dados históricos de consumo de EE. Conceitos referentes à demanda, curvas de carga, níveis de agregação, horizontes preditivos, fatores que influenciam a demanda, além do pré-processamento dos dados, são apresentados a seguir.

Uma série temporal representa um conjunto de medidas de uma variável aleatória de interesse tomadas de maneira ordenada no tempo. Portanto, uma curva de carga pode ser entendida como uma série temporal. E como tal, apresenta dependência serial, isto é, dependência entre instantes de tempo.

A curva de carga de um consumidor específico, ou de um aglomerado de consumidores, apresenta informações relativas ao comportamento da carga e à sua solicitação ao sistema que a supre, ou seja, descrevendo a demanda (KAGAN; OLIVEIRA; ROBBA, 2010).

É possível observar que as curvas de carga são tipicamente moldadas pela composição das mais diversas cargas elétricas associadas e distribuídas no tempo. Tipicamente uma curva de carga diária é formada por 24, 48 ou 96 medições equiespaçadas, por 1h, 30min ou 15min, respectivamente.

As características de consumo apresentadas pelas curvas de cargas diárias podem apresentar alterações bastante significativas, causadas por diversos fatores, onde destacam-se as mudanças organizacionais, estruturais, de rotina (influenciadas ou não por fatores externos, como meteorológicos e econômicos, que por sua vez, podem ou não ser cíclicos), troca e a inserção de equipamentos elétricos, entre outros.

Tendo em vista que determinadas mudanças em uma curva de carga diária típica ocorre em considerável parte por alterações do modo de uso, substituição ou inserção de novos equipamentos, e que isto normalmente ocorre dentro de intervalos significativos de tempo, há de se considerar que o conhecimento adquirido sobre o ‘comportamento’ das cargas não deve ser necessariamente descartado, mas tido como mutável com o decorrer do tempo e com a evolução do cenário responsável pela demanda.

3.1 Previsão da Demanda de Energia Elétrica

Os sistemas previsores de carga, sobretudo de curto prazo, utilizam informações referentes ao consumo de energia elétrica medidas no passado recente, além da possibilidade de considerar informações correlatas e/ou exógenas, como tipo de dia e dados climáticos.

O previsor deve ser capaz de relacionar informações previamente medidas, caracterizadas como conjuntos de entrada, com um conjunto de estimativas de carga futura, ou seja, saídas.

É necessário definir quais informações serão consideradas pelo previsor, sendo que tal aspecto está relacionado a disponibilidade de dados, correlação entre as séries de dados, horizontes de previsão entre outros. O horizonte de previsão pode ser do tipo *one-step*, prevendo apenas a carga para a próxima janela de tempo da série temporal. Outros problemas podem exigir que o previsor seja capaz de prever uma sequência de valores futuros utilizando apenas os valores observados no passado, esta tarefa é conhecida como *multistep-ahead time series prediction* (GERSHENFELD; WEIGEND, 1993). Uma abordagem típica para este problema é utilizar um previsor ou construir um modelo a partir dos dados históricos, e realizar previsão *step-wise* (passo a passo), considerando o valor previsto no passo corrente para previsão do próximo passo, esta abordagem é conhecida como *multi-stage prediction*, técnica utilizada nos previsores deste trabalho.

O horizonte de dados históricos a ser utilizado na etapa de previsão também precisa ser definido, alguns trabalhos utilizam semanas ou poucos meses, enquanto outros utilizam dados de mais de um ano, tal decisão é capaz de grande influência no desempenho do previsor. Uma vez que não existe consenso e um método estabelecido suficientemente abrangente que indique qual horizonte de dados históricos deve ser utilizado em cada caso, esta tarefa cabe ao especialista, que na prática costuma se basear em conhecimento empírico, testes com subconjuntos dos dados, estudos correlatos, tentativa e erro entre outros.

Após definido o horizonte de dados históricos e o de previsão, é necessário definir os parâmetros do modelo, número de entradas entre outros. Esta etapa também depende dos critérios utilizados pelo especialista, haja vista a inviabilidade de uma busca exaustiva na maioria dos casos; de tal modo, somente um grupo restrito de possibilidades é considerado.

3.2 Níveis de Agregação

Em muitos casos a energia elétrica é gerada em locais geograficamente distantes de seus consumidores, transportada através dos sistemas de transmissão em grandes blocos com elevados níveis de tensão. Próximo a seus centros de consumo, é entregue para os sistemas de distribuição, que operam com blocos menores e com níveis de tensão mais baixos, de modo a atender os consumidores finais.

O consumo de energia elétrica normalmente é apresentado com grande nível de agregação, isto é, fornece informações referentes a grandes aglomerados de consumidores ou de grandes consumidores de EE.

A agregação das cargas elétricas normalmente ocorre em nível de alimentadores, transformadores ou subestações de energia. Esta agregação caracteriza-se pela integralização das demandas individuais a partir de determinado ponto de medição.

Consumidores do ramo industrial podem apresentar grande complexidade, em parte por suas etapas produtivas e de trabalho, emprego de diferentes tecnologias, condições do mercado entre outras, características que por vezes podem ser modeladas adequadamente. Já a volatilidade com que se configura uma curva de carga residencial representa outro tipo de desafio, tendo em vista a impossibilidade de modelar de maneira satisfatória diversos fatores responsáveis por essas flutuações, originários da vontade humana (MOCANU, 2016).

Antes do advento das redes inteligentes, que se caracterizam pela inserção de tecnologias da informação nas redes de distribuição, existiam alternativas menos integradas e mais elementares capazes de captar dados de consumo de energia elétrica em diferentes resoluções e extensões, porém, tais aplicações eram pouco comuns, visto sua baixa importância ao se considerar o mercado de energia elétrica, onde a comercialização ocorria em níveis mais agregados; este cenário muda com as *smart grids* e a possibilidade de comercialização em níveis menos agregados. Atualmente estão disponíveis dados de consumo de EE de grandes períodos de tempo, com resoluções maiores e em níveis menos agregados, como consumidores individuais ou pequenos grupos (AGARWAL; WENG; GUPTA, 2011).

Algumas características presentes nestes novos cenários, como a disponibilidade de maiores horizontes de dados históricos, somado a sua alta resolução, implicam diretamente na etapa de pré-processamento dos dados, resultando por vezes em maiores custos computacionais.

3.2.1 Impacto na Previsão

A partir de uma abordagem *bottom-up*¹ em um cenário de *smart grid*, as séries de consumo de EE tornam mais extensas, mais desagregadas, com maior nível de detalhe, e de modo consequente, apresentam maiores níveis de ruído (CHIN, 2016). Desta forma, as características descritas devem ser consideradas para tarefa de previsão de cargas elétricas em níveis menos agregados, além de outras, como a maior aleatoriedade da série e a presença de dados atípicos.

Modelos baseados em inteligência artificial, como as redes neurais artificiais são afetadas de diferentes formas pela presença de dados ruidosos, dependendo de sua

¹ Usualmente os previsores são aplicados em dados de demanda agregados, em uma abordagem *top-down*, porém nos últimos anos graças à popularização dos *smart meters*, dados de consumidores individuais passaram a estar cada vez mais disponíveis, permitindo que a previsão fosse realizada para cada consumidor individualmente, e então, a demanda agregada obtida a partir da soma das previsões de demanda individuais, este tipo de abordagem é dito *bottom-up* (DANGERFIELD; MORRIS, 1992).

arquitetura. Porém, de maneira geral o processo de aprendizagem é um dos mais afetados, fazendo com que a rede não seja capaz de lidar adequadamente com dados de entrada distintos daqueles considerados durante o treinamento da rede. No caso dos modelos estatísticos, a presença de ruído é capaz de influenciar no processo de otimização dos coeficientes e na escolha dos modelos, afetando diretamente a capacidade do modelo extrapolar de maneira adequada o comportamento futuro da série. Neste sentido, a inclusão de uma etapa de pré-processamento dos dados é necessária, permitindo que sejam adotadas estratégias para a remoção prévia dos dados ruidosos, de maneira a mitigar seu impacto sobre os mais distintos previsores.

A maior aleatoriedade dos níveis menos agregados, assim como a impossibilidade de modelar adequadamente diversos fatores que podem ser responsáveis pelas flutuações no consumo de EE, fazem com que as tarefas de previsão nestes níveis se configurem como um desafio, exigindo previsores robustos, atenção no pré-processamento dos dados, e escolha adequada dos horizontes dos conjuntos de dados utilizados nas etapas de treinamento.

A previsão de cargas elétricas em maiores níveis de agregação, como cidades ou grandes regiões metropolitanas, comporta-se como um problema menos complexo que aquele de realizar previsão em níveis bastante menos agregados, como residencial ou pequenos aglomerados, onde as cargas são mais estocásticas e voláteis (YU; MIROWSKI; HO, 2017). Os erros obtidos em previsões de cargas elétricas altamente agregadas costumam ser baixos, facilmente ficando entre 1 e 2% em nível de subestação, por exemplo. Porém, quando são considerados níveis altamente desagregados, estes erros podem ser consideravelmente maiores, chegando a casa de 30% (SEVLIAN; RAJAGOPAL, 2017).

O desenvolvimento de uma abordagem de previsão robusta exige análise dos dados, e que inconsistências sejam identificadas e tratadas adequadamente, de modo que seu impacto sobre o sistema predictor seja mitigado, permitindo ao modelo capturar o comportamento presente na série de dados.

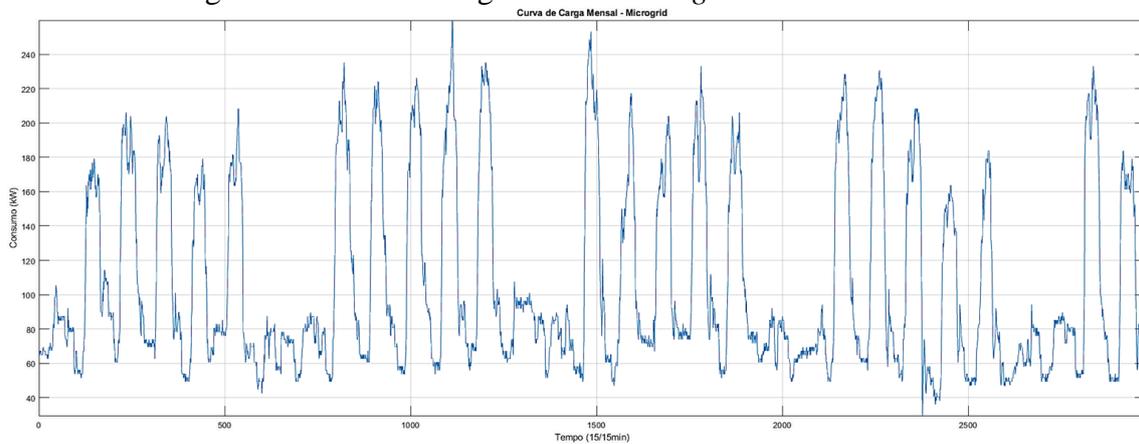
3.3 Fatores que Influenciam a Demanda por Energia Elétrica

É fato que a curva de carga apresenta todas as informações relativas ao comportamento da carga e à sua solicitação ao sistema (KAGAN; OLIVEIRA; ROBBA, 2010). Entretanto, a curva de carga não torna explícito em si todos os mecanismos responsáveis por moldar seu formato (*shape*).

Ao observar curvas de carga diária dos dias compreendidos em uma semana, é possível observar diferenciações significativas; ao considerar horizontes de tempos maiores, como um mês ou um ano, estas variações tendem a ser ainda mais evidentes, influenciadas pelas estações do ano, períodos de férias entre outros. A Figura 1

apresenta a demanda de uma *microgrid* para o mês de janeiro de 2017, onde os dias úteis são facilmente diferenciados dos finais de semana, uma vez que estes últimos registram menor demanda por EE.

Figura 1 - Curva de carga de uma *microgrid* - Janeiro de 2017



Fonte: Dados da pesquisa

Diversos fatores são capazes de exercer influência em maior ou menor grau na forma que os usuários consomem energia elétrica; tendo em vista níveis menos agregados, como residencial, as incertezas envolvidas tendem a ser maiores, onde dados referentes a diversos fatores não estão disponíveis e, portanto, não podem ser considerados em um sistema de previsão, de modo que fatores alheios a vontade do consumidor são considerados na tentativa de *explicar* o comportamento da série, dentre eles, pode-se destacar as condições climáticas, hora, horário de verão, temperatura e incidência de luz solar. Feriados, finais de semana, férias, dia anterior e posterior a férias e feriados também são capazes de exercer grande impacto sobre o formato das curvas de cargas (KHATOON et al., 2014). Alguns destes fatores são descritos a seguir (KADOWAKI et al., 2004):

- Condições climáticas: capaz de apresentar elevada influência sobre determinadas curvas de carga, devido, por exemplo, a seu impacto sobre a carga de climatização e iluminação, que pode ser mais requerida em dias muito quentes, muito frios, e/ou nublados;
- Tipo de dia: os dias normalmente são divididos entre dias úteis, feriados e finais de semana. É possível indicar cada dia da série como pertencente a um dos sete dias da semana, ou ainda, que determinados dias da semana pertençam a um mesmo grupo (i.e., segundas, terças e quartas-feiras podem apresentar comportamento similar, então são consideradas como pertencentes ao mesmo tipo, excluindo-se, portanto, os outros dias da semana desta categoria). De acordo com o cenário em estudo, é possível que outros tipos de dia possam ser considerados, como pré-feriado, pós-feriado, feriado nacional, feriado regional entre outros;

- Horário de verão: quando presente, sua influência sobre a curva de carga normalmente se dá como um deslocamento das cargas; também pode ser observado algum impacto sobre a demanda máxima;
- Horário do pôr do sol: no inverno do sudeste brasileiro, o pôr do sol se inicia próximo às 17h30min, e gradualmente se desloca até próximo às 19h no verão. Sendo assim, é possível conceber que tal fenômeno seja capaz de causar considerável impacto sobre as cargas de iluminação, causando deslocamento da rampa de consumo da esquerda para direita;
- Temperatura: sua influência sobre a carga está em grande parte relacionada ao tipo de consumidor em questão (por exemplo: residencial, comercial, industrial ou especial). A quantidade de energia elétrica necessária para manter a temperatura de um equipamento de solda na indústria automobilística pode ser maior no inverno que no verão, por exemplo, assim como, sistemas de climatização e de aquecimento (elétrico) de água tendem a ser mais ou menos exigidos de acordo com a variação da temperatura. A temperatura apresenta significativa variação durante o decorrer do dia, e sua variação pode se tornar ainda mais evidente ao ser observada durante o passar das estações do ano, com a variação da umidade do ar, entre outros.

Nem sempre estão disponíveis dados exógenos como temperatura, índice de precipitação, umidade relativa do ar, entre outros. Porém, informações como hora, tipo de dia, horário de verão, estação do ano, entre outras, independem de medidores, de modo que mais comumente são consideradas ao se ponderar fatores que exercem influência sobre as curvas de carga.

Desde 1985 o horário de verão é adotado anualmente no Brasil. Porém, devido a mudanças nos hábitos de consumo dos brasileiros, que em considerável parcela passaram a utilizar aparelhos de ar condicionado durante os horários mais quentes do dia, compreendido entre o final da manhã e o início da tarde, contribuiu para que o horário de ponta fosse transferido do período entre 17h e 20h para o período entre 14h e 15h. De tal modo, o horário de verão parece não mais se justificar sob a perspectiva do setor elétrico, conforme fala do diretor-geral da Agência Nacional de Energia Elétrica - ANEEL amplamente divulgada pela mídia (VILLAS BÔAS, 2017; WARTH, 2017).

3.4 Horizontes de Previsão em Sistemas de Potência

Inicialmente os horizontes preditivos em sistemas de potência eram divididos em três grupos principais, sendo: longo prazo, médio prazo e curto prazo. O surgimento de medidores eletrônicos, *smart meters*, o conceito de *smart-grids*, e o aumento da complexidade operacional dos sistemas de potência, contribuíram para que novos horizontes preditivos se popularizassem. Os tópicos a seguir apresentam característica

dos horizontes preditivos citados, destacando que tais informações são sensíveis para determinados mercados e países (HONG; FAN, 2016):

- **Previsão de Longo Prazo:** este é o mais longínquo horizonte preditivo considerado, normalmente abrange de alguns anos a décadas. Este tipo de previsão apresenta menos detalhe que as previsões de curto e médio prazo, uma vez que seu objetivo também é distinto, servindo como referência para orientação de investimentos e decisões comerciais. No contexto da distribuição, esse tipo de previsão está relacionado a decisões de compra e a venda de energia elétrica, que podem ser tomadas com vários anos de antecedência. Entre os principais desafios neste tipo de previsão está a necessidade de considerar e estimar diversos fatores, como a disponibilidade de geração, preços do carvão e do petróleo, taxas de câmbio, de juros entre outros, considerando horizontes bastante distantes.
- **Previsão de Médio Prazo:** compreende horizontes preditivos de semanas a 3 anos. Seu objetivo está em servir de base para verificação do volume hídrico necessário nas hidrelétricas, de combustível nas termoelétricas, para operações de manutenção e para o planejamento de intercâmbio energético. Sendo assim, se destina a orientar tarefas que visam otimizar recursos disponíveis.
- **Previsão de Curto Prazo:** neste modelo de previsão, o horizonte de tempo envolvido é bastante curto, compreendendo horizontes de uma hora a uma semana. Rahman e Hazin (1993) argumentam que esse tipo de previsão é fundamental para orientar o planejamento da operação, transferência de energia e gerenciamento de demanda, sendo fundamental para o planejamento da operação do sistema. Uma previsão de curto prazo adequada pode permitir otimizar recursos, e, até mesmo reduzir custos de geração.
- **Previsão de Curtíssimo Prazo:** designado na literatura como *very short-term electric load forecasting*, corresponde aos menores horizontes preditivos geralmente considerados, indo de poucos minutos até uma hora. Esta previsão é utilizada na etapa de operação em tempo real dos recursos energéticos despacháveis.

3.5 Pré-Processamento dos Dados de Consumo de Energia Elétrica

Uma vez que por definição uma curva de carga é uma série temporal, isto significa que ela também pode ser analisada a partir de sua decomposição em componentes de tendência, ciclo e sazonalidade, onde (MORETTIN; TOLÓI, 2006):

- **Tendência:** representa o comportamento de longo prazo observado em uma série temporal, como aquela que representa um conjunto de curvas de carga diária, indicando se elas permanecem em um mesmo patamar, se crescem, decrescem e em que velocidade estes efeitos se apresentam.
- **Ciclo:** ou componente cíclica, descreve variações suaves de subida e queda na série, por vezes, relacionadas a questões de ordem econômica e meteorológica. Devido sua natureza, movimentos cíclicos tendem a ser irregulares.
- **Sazonalidade:** está relacionada às oscilações de subida e queda observadas em intervalos regulares em uma série temporal, como as oscilações observadas na demanda por EE, que são influenciadas entre outros, pela jornada de trabalho, processos de produção e descanso.

Assim posto, é importante que os dados sejam observados ao longo do tempo, e que o mecanismo predictor seja capaz de detectar as variações descritas por cada um dos componentes, de maneira a projetar sua influência para determinados horizontes futuros.

Ao analisar a série temporal de acordo com seus componentes, o especialista terá acesso a indicadores que farão com que determinados métodos de previsão sejam preferíveis em detrimento de outros, uma vez que os primeiros são capazes de lidar de maneira mais adequada com determinadas características verificadas.

De acordo com Pyle (1999), o pré-processamento dos dados descreve um conjunto de verificações e tarefas que permitam preparar os dados históricos, de modo que possam ser adequadamente analisados. Nesta etapa, normalmente os dados são analisados quanto a existência de valores faltantes, *outliers* e amostragem irregular:

- **Valores Faltantes:** a ausência ou supressão de dados pode ser decorrente de erros no preenchimento de formulários, inconsistências entre registros, mau funcionamento nos equipamentos de medição, transporte e armazenamento de dados;
- **Outliers:** também são conhecidos como valores discrepantes ou aberrantes, e representam informações que se distanciam do esperado. Tais informações devem ser tratadas com cuidado, de modo a determinar se essas informações se configuram como raridades ou erros;
- **Amostragem Irregular:** indica séries de dados cujos valores não foram tomados em intervalos de tempo equiespaçados. Uma vez que determinados métodos lidam exclusivamente com séries cujos valores foram tomados em intervalos igualmente espaçados, é necessário que sejam realizados ajustes, de modo que a série passe a refletir, da maneira mais adequada possível, os valores que seriam verificados caso a amostragem fosse regular.

Uma inconsistência se caracteriza como uma observação que se distancia das demais, ou ainda, uma observação ou subconjunto delas que parece ser atípica em relação ao restante do conjunto de dados (HAWKINS, 1980; BARNETT; LEWIS, 1994).

A tarefa de detectar inconsistências não é trivial, uma vez que dados de tal natureza normalmente são originários de eventos aleatórios, não sendo possível mapeá-los por intermédio de uma equação determinística.

Por outro lado, se os limiares tangíveis de determinados conjuntos de dados são conhecidos, o processo de identificar e assinalar valores que ultrapassam essas fronteiras como *outliers* é facilitado.

A literatura apresenta diversas técnicas para detecção de inconsistências, tais abordagens vão de observação visual da série, métodos estatísticos, técnicas de clusterização até métodos computacionais avançados. Tendo em vista que um único método pode não ser capaz de detectar todas as inconsistências de uma série temporal, uma abordagem alternativa é utilizar mais que uma técnica com este objetivo. Na continuação são apresentadas técnicas utilizadas neste trabalho para a etapa de pré-processamento dos dados.

3.5.1 Detecção de Inconsistências

As séries temporais foram analisadas previamente por meio do Weka (*Waikato Environment for Knowledge Analysis*), um *software* livre de código aberto para mineração de dados, que pode ser utilizado para classificação, mineração de regras de associação, clusterização, previsão de séries temporais, além de permitir visualizar e analisar séries históricas (WITTEN; FRANK, 2005).

Posteriormente desenvolveu-se um algoritmo capaz de efetuar varredura nos dados em busca de elementos inválidos para os previsores, como elementos negativos ou aberrantes, além daqueles identificados no *software* Matlab® como tipo NaN (*not a number*, que indica um dado do tipo indefinido, irrepresentável ou inexistente).

O Critério de Chauvenet, também conhecido como Teste de Chauvenet, foi empregado na etapa de pré-processamento dos dados. A técnica permite identificar se um determinado dado é discrepante considerando os demais valores da amostra. Baseia-se em encontrar o maior desvio da amostra de n dados, para tanto, calcula-se o desvio d_i de cada dado em relação à média $d_i = x_i - \bar{x}$. O critério de eliminação é dado por (CHAUVENET, 1891):

$$\frac{d_i}{\sigma} > \frac{d_{max}}{\sigma} \quad (3.1)$$

onde d_{max} representa o máximo desvio admissível, e σ representa o desvio padrão da amostra. O algoritmo gera um alerta sempre que um dado é tido como discrepante.

Quando determinados dados indicados pelo Critério de Chauvenet são eliminados, suas posições originais são armazenadas, para que posteriormente sejam inseridas informações nestas posições, de acordo com a técnica utilizada.

3.5.2 Tratamento de Inconsistências

Por vezes dados faltantes e *outliers* são tratados de maneira similar, com a inserção ou substituição por valores obtidos a partir da utilização de determinadas técnicas. Entre as técnicas mais utilizadas pode-se citar: média simples ou ponderada dos valores vizinhos mais próximos, e substituição por previsão dentro da amostra.

Para este trabalho, uma vez que um dado d_i de uma determinada curva de carga c seja identificado como inconsistente ou *NaN*, este valor é substituído utilizando um operador estatístico baseado em medidas de posição, neste caso, média simples:

$$d_i = \frac{d_{i-1} + d_{i+1}}{2}.$$

É importante observar que a utilização de técnicas que permitam introduzir informações na série, via de regra, não asseguram que o dado gerado seja igual ou muito próximo ao real. Introduzir ou substituir dados pode significar incorporar ruído à série, de maneira que, por vezes, pode ser preferível não preencher valores faltantes ou substituir *outliers*, escolhendo remover os segundos, e lidar com séries sem estes dados.

3.5.3 Remoção de Ruídos utilizando SSA

O *Singular Spectrum Analysis* se apresenta como uma importante técnica para análise de séries temporais. O SSA incorpora elementos da análise clássica de séries temporais, estatística multivariada, geometria multivariada, sistemas dinâmicos e processamento de sinais (GOLYANDINA; NEKRUTKIN; ZHIGLJAVSKY, 2001).

O objetivo do SSA é realizar a decomposição de uma série original em um pequeno número de componentes independentes e interpretáveis, como componentes de tendência, oscilatórios e ruído branco.

A técnica SSA pode ser empregada para a resolução de uma gama de problemas, dentre os quais destacam-se: encontrar tendências de diferentes resoluções, detecção de mudanças de regime, extrair componentes sazonais e suavização de séries temporais. O SSA apresenta algumas vantagens ao ser aplicado na realização das tarefas descritas, dentre as quais destaca-se a necessidade de poucos dados para a obtenção de resultados adequados, além da abordagem depender de poucos parâmetros, exigindo menos conhecimento prévio sobre as séries temporais (GOLYANDINA; NEKRUTKIN; ZHIGLJAVSKY, 2001; ELSNER; TSONIS, 1996).

A abordagem clássica SSA consiste em dois estágios principais, o primeiro engloba a decomposição da série temporal - ST, e consiste em decompor a série em uma matriz Hankel, e, decompor a matriz Hankel por meio de *Singular Value Decomposition* - SVD; o segundo estágio garante a reconstrução da ST, para tanto, engloba reconstruir

as subséries por meio da média diagonal, além do agrupamento das séries, de modo a identificar componentes de tendência, oscilatórios e ruído.

Os trabalhos de Li, Cui e Guo (2014), Wu e Chau (2011), Marques et al. (2006) e Golyandina, Nekrutkin e Zhigljavsky (2001) descrevem o algoritmo SSA, apresentado na sequência:

Considerando a primeira parte do estágio de decomposição, seja $y_t (t = 1, 2, 3, \dots, N)$ uma série temporal unidimensional de tamanho N , que deseja-se submeter ao SSA, y_t é transformada em uma sequência de vetores L – *dimensionais* (em que L é chamado tamanho da janela), que são utilizados para construir uma matriz trajetória X , que apresenta o seguinte formato:

$$X = [x_1, \dots, x_k] = \begin{pmatrix} y_1 & y_2 & \dots & y_K \\ y_2 & y_3 & \dots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \dots & y_N \end{pmatrix} \quad (3.2)$$

Esta matriz apresenta L linhas e K colunas, sendo $K = N - L + 1$. Não existe uma regra restrita para a determinação de L , porém, ele deve pertencer ao intervalo $2 \leq L \leq N - 1$. Maneiras de determinar seu valor ótimo ainda são tópicos de pesquisas.

A decomposição da matriz por SVD consiste em transformar a matriz trajetória em uma soma de matrizes de rank 1. Sendo $S = XX^T$, $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_L$ seus respectivos autovalores em ordem decrescente, assim como $U_1, U_2, U_3, \dots, U_L$ os autovetores associados a $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_L$, além de $V_i = X^T U_i / \sqrt{\lambda_i}$, a matriz trajetória X pode ser escrita como $X = X_1 + X_2 + X_3 + \dots + X_L$, onde $X_i = \lambda_i U_i V_i^T$. A coleção (λ_i, U_i, V_i) é designada como i -ésima autotripla na matriz trajetória X .

O agrupamento tem como objetivo a junção das autotriplas, de forma que os grupos remanescentes sejam interpretáveis como tendência, um ou mais componentes oscilatórios com períodos variáveis, e ruído. Para tanto, nesta subetapa as L autotriplas são separadas em m conjuntos *disjuntos*. Seja $I = \{i_1, i_2, i_3, \dots, i_p\}$ os índices das autotriplas agrupadas em um dos m subconjuntos desejados, determina-se a matriz resultante X_I como a soma das X_i matrizes correspondentes, ou seja, $X_I = X_{i_1} + X_{i_2} + X_{i_3} + \dots + X_{i_p}$.

A subetapa média diagonal é responsável por transformar cada uma das m matrizes X_I , em uma subsérie temporal univariada de tamanho N , conhecida como componente SSA. Esta etapa consiste em calcular a média dos elementos de X_I , denotados $x_{I(i,j)}$, onde $i + j = k$, sendo k uma constante qualquer. Desta forma, sendo $L^* = \min(L, K)$ e $K^* = \max(L, K)$, onde $y_{I_t} (t = 1, 2, 3, \dots, N)$ denota uma componente SSA, cada elemento de y_{I_t} é dado por:

$$y_{I_t} \begin{cases} \frac{\sum_{i=1}^t x_{I(i,t-i+1)}}{t}, \forall 1 \leq t \leq L^* \\ \frac{\sum_{j=1}^{L^*} x_{I(i,t-i+1)}}{L^*}, \forall L^* \leq t \leq K^* \\ \frac{\sum_{j=t-K^*}^{N-K^*-1} x_{I(i,t-i+1)}}{N-K^*-1}, \forall K^* \leq t \leq N \end{cases} \quad (3.3)$$

3.5.3.1 Codificação e Parametrização do SSA

A codificação do SSA foi realizada no *software* R, tendo como entradas a série histórica e o tamanho da janela SSA. Como saída são produzidas três séries, uma produto da reconstrução dos dados considerando apenas tendência, outra produto da reconstrução dos dados considerando apenas harmônicos, e por fim, uma terceira corresponde ao produto da reconstrução dos dados considerando apenas ruído.

A técnica SSA foi empregada com objetivo de remover ruído da série original, para tanto, o parâmetro tamanho da janela SSA recebeu valor correspondente ao dobro da periodicidade dos dados, uma vez que (GOLYADINA; NEKRUTKIN; ZHIGLIAVSKY, 2001) argumenta que utilizar múltiplos da periodicidade pode permitir uma melhor separabilidade dos dados. Ainda que esta não seja uma questão principal para o problema de suavização da série, permite identificar quando uma janela é suficientemente grande para permitir o efeito desejado.

As componentes SSA foram clusterizadas hierarquicamente em três grupos, sendo utilizado como critério de distância entre as componentes a matriz *weighted-correlations* (*wcor*). Os grupos resultantes foram interpretados como componente oscilatório de baixa frequência de variação, componente oscilatório de alta frequência de variação, e ruído, de modo que este último não foi utilizado para reconstrução da série, resultando em uma série menos ruidosa, que pode ser dita: suavizada.

4 REDES NEURAIIS ARTIFICIAIS

Neste capítulo é apresentada uma introdução sobre as redes neurais artificiais. A família de redes da teoria da ressonância adaptativa é introduzida, e com maior destaque é apresentada a RNA ARTMAP *Fuzzy*, empregada para previsão *multistep* de cargas elétricas em distintos níveis de agregação. A rede MLP também é destacada, sendo igualmente utilizada para previsão de cargas elétricas.

4.1 Introdução a Redes Neurais Artificiais

As Redes Neurais Artificiais - RNA tiveram seu início em 1943 no trabalho de McCulloch e Pitts (MCCULLOCH; PITTS, 1943). Os autores buscaram modelar matematicamente o comportamento fundamental de um neurônio biológico conhecido à época; seus resultados foram considerados embaixadores para as futuras implementações computacionais (EBERHART; DOBBINS, 1990).

Uma RNA compreende uma técnica bio-inspirada, uma vez que busca assimilar mecanismos da estrutura neural de organismos inteligentes, capazes de adquirir conhecimentos e efetuar assimilações ao serem expostos a determinadas experiências.

Um dos aspectos mais importantes de uma RNA diz respeito a sua capacidade de generalização. Esta característica denota a habilidade da rede em apresentar respostas coerentes para dados de entrada diversos daqueles utilizados durante a etapa de treinamento.

Técnicas de inteligência artificial, especialmente redes neurais artificiais, já se mostraram capazes de aprender relações não lineares complexas; uma vez que tais relações são difíceis de modelar, estas técnicas ganharam especial destaque (LUTHULI; FOLLY, 2016).

As redes neurais são amplamente utilizadas na área de previsão de cargas elétricas (RAZA; KHOSRAVI, 2015; SCHELLONG; HENTGES, 2007; HIPPERT; BUNN; SOUZA, 2005). Na sequência são apresentadas com mais detalhes a RNA MLP, a ART e algumas das principais redes desta família, com especial enfoque são apresentadas as redes ART, ART *Fuzzy*, ARTMAP e ARTMAP *Fuzzy*.

4.2 Rede Neural Artificial *Multilayer Perceptron*

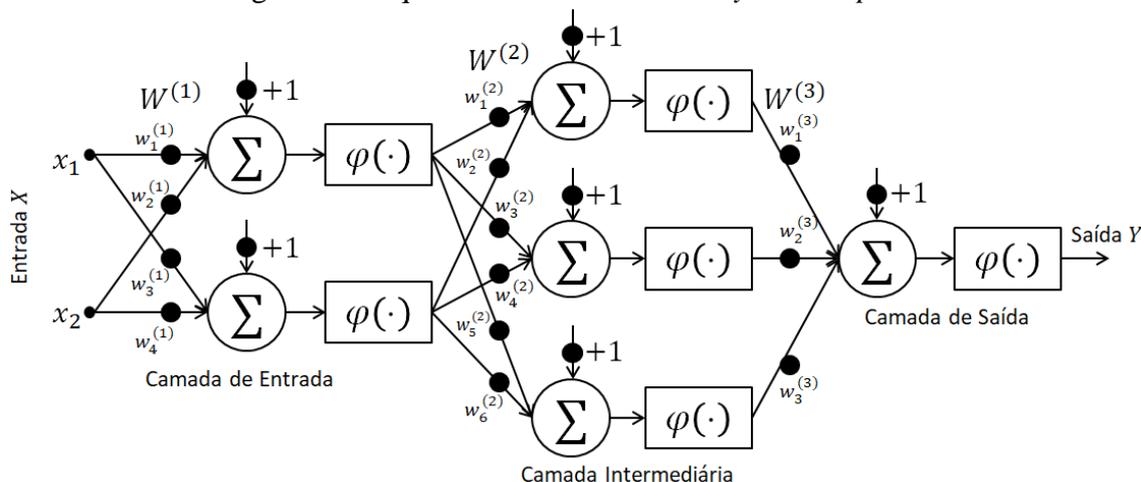
Neste trabalho a RNA MLP é empregada como *benchmark*, de modo que as previsões realizadas a partir da rede ARTMAP *Fuzzy*, objeto de estudo, pudessem ser comparadas com aquelas resultantes de uma arquitetura distinta de rede neural, porém amplamente utilizada.

Durante o processo de treinamento de redes neurais artificiais de camada simples, os padrões de entrada são diretamente mapeados para conjuntos de padrões de saída, de modo que não existe qualquer representação interna na rede.

Este tipo de estrutura simples pode permitir que padrões de entrada similares sejam associados a padrões de saída similares, de modo que a rede seja incapaz de aprender mapeamentos mais complexos. Sendo assim, padrões de entrada com estruturas similares que levem a saídas distintas, como o caso da função XOR (ou exclusivo), não são passíveis de serem solucionadas por redes Perceptron de camada única. Porém, Minsky e Papert (1969) demonstraram que esse tipo de problema poderia ser solucionado utilizando uma rede com camada intermediária.

A rede neural *MultiLayer Perceptron* é uma das arquiteturas de redes neurais artificiais mais utilizadas, podendo ser aplicada em uma vasta gama de problemas (HAYKIN, 2009). Sua arquitetura é baseada em neurônios do tipo McCulloch & Pitts dispostos em camadas. Estas camadas são designadas como camada de entrada, camada(s) intermediária(s) e camada de saída. A Figura 2 apresenta a arquitetura de uma rede MLP com uma camada intermediária (HAYKIN, 2009).

Figura 2 - Arquitetura da RNA *MultiLayer Perceptron*



Fonte: Adaptado de (HAYKIN, 2009)

A rede apresentada na Figura 2 é do tipo *feedforward*, portanto, o sinal de entrada X é combinado com os pesos sinápticos $W^{(1)}$ e então propagados através da rede. A saída de cada neurônio da camada de entrada e da(s) camada(s) intermediária(s) compõe a entrada das camadas subsequentes, até que seja gerada uma saída Y .

4.2.1 Treinamento por Backpropagation

Para que uma rede neural seja capaz de fornecer respostas adequadas para padrões de entradas diversos, é necessário que ela esteja devidamente treinada. Neste sentido, é apresentado o algoritmo de treinamento supervisionado *backpropagation*, que utiliza a técnica de gradiente descendente no ajuste dos pesos sinápticos, de modo que o erro produzido pela rede fique abaixo de um limite pré-estabelecido (HAYKIN, 2009).

Na sequência são apresentados os passos *forward* e *backward* que descrevem a RNA MLP com algoritmo de treinamento *backpropagation*, descritos em (HAYKIN, 2009; WIDROW; LEHR, 1990):

4.2.1.1 *Passo forward*

1) Um conjunto de entradas X é apresentado, e o sinal resultante propagado pela rede. Sendo a saída linear do neurônio representada por:

$$s = \sum_{i=1}^m w_i x_i + b \quad (4.1)$$

em que:

w_i : representa o i -ésimo peso sináptico;

x_i : representa a i -ésima entrada;

b : peso bias;

m : número de entradas.

A saída não linear de cada neurônio é dada por:

$$y_j(n) = \varphi_j(s_j(n)) \quad (4.2)$$

em que:

$y_j(n)$: corresponde a saída do j -ésimo neurônio na iteração n ;

φ_j : função não linear.

2) O erro da RNA na camada de saída para a n -ésima iteração é representado por:

$$e_j(n) = d_j(n) - y_j(n) \quad (4.3)$$

em que:

d_j : representa a j -ésima saída desejada;

y_j : representa a j -ésima saída calculada pela RNA.

3) Realiza-se o cálculo do gradiente local $\delta_j(n)$ referente a última camada, considerando o erro gerado na camada de saída e a derivada do erro dada pela equação (4.4):

$$\delta_j(n) = e_j(n) \varphi'_j(s_j(n)) \quad (4.4)$$

sendo que a função φ escolhida deve ser diferenciável.

4.2.1.2 *Passo backward*

4) O gradiente local $\delta_i(n)$ de cada neurônio das camadas anteriores é calculado de acordo com:

$$\delta_j(n) = \varphi'_j(s_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (4.5)$$

em que:

j : é o índice do neurônio da camada atual;

k : é o índice do neurônio da camada posterior.

5) Uma vez calculado cada gradiente local, os pesos sinápticos são ajustados de acordo com:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (4.6)$$

em que:

η : taxa de aprendizagem da rede.

A adaptação dos pesos se dá de acordo com:

$$w_{ji}^{k+1}(n) = w_{ji}^k(n) + \Delta w_{ji}(n) \quad (4.7)$$

O erro instantâneo $\varepsilon(n)$ é medido a cada novo padrão que é apresentado a rede, de acordo com:

$$\varepsilon(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (4.8)$$

em que:

C : conjunto de neurônios da camada de saída.

O algoritmo segue até que todos os padrões sejam submetidos a rede, sendo cada um destes ciclos é dito uma época. A convergência ocorre quando o erro médio dos erros instantâneos atende a tolerância estabelecida, de acordo com:

$$\varepsilon_{m\u00e9dio} = \frac{1}{N} \sum_{n=1}^N \varepsilon(n) \quad (4.9)$$

em que:

N : número total de padrões submetidos a rede.

4.2.2 *Implementação Computacional e Treinamento da RNA MLP*

A rede MLP utilizada neste trabalho é parte do pacote *Stuttgart Neural Network Simulator* - RSNNS disponível para o *software* R. Este pacote apresenta várias

implementações padrão de redes neurais, como *Dynamic Learning Vector Quantization* - DLVQ, *Radial Basis Function* - RBF, e, inclui ainda algumas redes da família ART, a saber, ART1, ART2 e ARTMAP (BERGMEIR; BENÍTEZ, 2017). O treinamento da rede pode ser realizado a partir de um conjunto de funções de treinamento disponíveis, porém a mais utilizada e aqui considerada é a *Backpropagation* convencional (*Std_Backpropagation*).

A função de treinamento *Std_Backpropagation* possui dois parâmetros, sendo, taxa de aprendizado e a diferença máxima de saída. A taxa de aprendizado é estabelecida geralmente entre 0,1 e 1, sendo responsável por definir o tamanho do passo do gradiente descendente. O parâmetro diferença máxima determina até qual valor para a diferença entre a saída e o valor alvo é tratado como zero, não sendo retropropagado; tal parâmetro busca prevenir *overtraining*.

Os argumentos que podem ser estabelecidos para a função MLP são apresentados na sequência (BERGMEIR; BENÍTEZ, 2017):

- *x*: define uma matriz com entradas para o treinamento da rede;
- *y*: contém as saídas desejadas;
- *size*: número de unidades nas camadas escondidas;
- *maxit*: define o número máximo de iterações para a etapa de aprendizado;
- *initFunc*: define a função de inicialização a ser utilizada;
- *initFuncParams*: contém os parâmetros da função de inicialização;
- *LearnFunc*: determina a função de treinamento (por padrão é *Std_Backpropagation*);
- *LearnFuncParams*: corresponde aos parâmetros para a função de treinamento;
- *updateFunc*: determina a função de atualização a ser utilizada;
- *updateFuncParams*: contém os parâmetros para a função de atualização;
- *hiddenActFunc*: determina a função de ativação para todas as unidades das camadas escondidas;
- *shufflePatterns*: indica se os padrões devem ser embaralhados;
- *linOut*: indica se a função de ativação das unidades da camada de saída são lineares ou logísticas;

- *outputActFunc*: determina a função de ativação para todas as unidades da camada de saída;
- *inputTest*: representa uma matriz de entradas para testar a rede;
- *targetTest*: determinas as saídas esperadas para o teste;
- *pruneFunc*: indica a função de poda a ser utilizada;
- *PruneFuncParams*: contém os parâmetros para a função de poda.

A Tabela 1 apresenta os valores utilizados pela MLP em todos os cenários de previsão:

Tabela 1 - Parâmetros adotados para a rede MLP

Função	Valor
maxit	100
initFunc	Randomize_Weights
initFuncParams	c(-0.3, 0.3)
learnFunc	Std_Backpropagation
learnFuncParams	c(0.2, 0)
updateFunc	Topological_Order
updateFuncParams	0.0
hiddenActFunc	Act_TanH
shufflePatterns	FALSE
linOut	TRUE
normalization.method	premnmx
monte.carlo	FALSE

Com base em experiências prévias foram testadas 09 configurações para o número unidades na camada de entrada, sendo definidos de acordo com o número de elementos correspondente a uma curva de carga diária (48 ou 96 elementos), com o dobro e com o triplo deste valor, foram testados também valores próximos a estes, decrescidos e acrescidos de 02; para a camada intermediária foram considerados valores correspondentes a metade daqueles utilizados na camada de entrada.

4.3 Redes Neurais Artificiais da Família ART

Na sequência são apresentadas considerações a respeito da Teoria da Ressonância Adaptativa. A ART representa uma evolução no contexto de redes neurais artificiais, originando uma família de redes adequadas para uma considerável gama de aplicações.

4.3.1 Dilema da Estabilidade-Plasticidade

O dilema da estabilidade/plasticidade está relacionado a capacidade de um sistema em efetuar assimilações, isto é, aprender novas informações sem

corromper/perder conhecimento prévio, mantendo a estabilidade do sistema. A teoria da ressonância adaptativa é proposta para enfrentar tal dilema (GROSSBERG; 1982a; GROSSBERG, 1982b).

A estabilidade diz respeito a capacidade do sistema em se manter estável em relação a informações já aprendidas, tal característica é possível nas redes da família ART devido a garantia de agrupamento de todos os elementos nas categorias criadas pelo modelo, uma vez que o processo de adaptação dos pesos da rede tem por característica o decrescimento, ou seja, a medida que os pesos são adaptados, os valores tendem ao valor mínimo (entre a entrada da rede e o peso a ser atualizado) até a estabilização (LOPES, 2005).

A plasticidade está relacionada a capacidade da rede aprender novos padrões sem perder aquilo que foi assimilado anteriormente, tal característica é garantida por meio da ativação de novas categorias.

4.3.2 *Teoria da Ressonância Adaptativa*

A teoria da ressonância adaptativa pretende explicar a maneira que o cérebro é capaz de processar e categorizar informações ao passo que novas informações lhe são apresentadas.

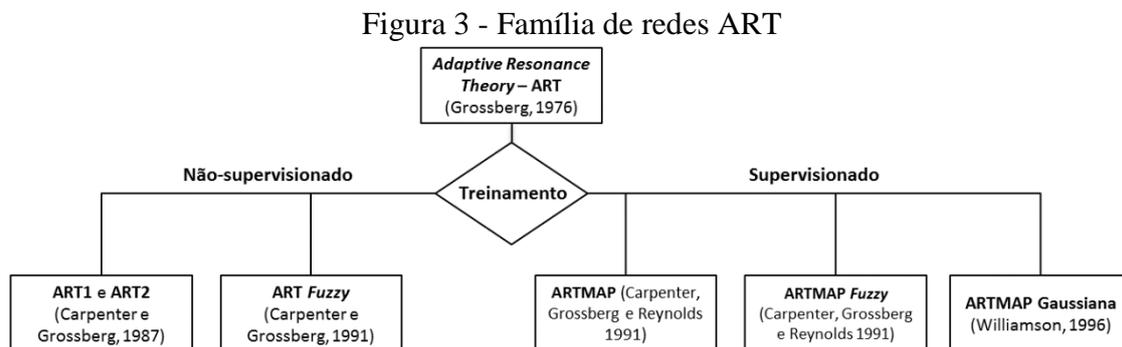
Em uma rede da família ART, a aprendizagem ocorre devido as informações que são representadas pelas saídas dos neurônios apresentarem efeito de ressonar (ou de oscilar) entre as camadas de unidades de processamento da rede, apresentando um estado de ativação de equilíbrio (AMORIM, 2006).

A partir da teoria da ressonância adaptativa apresentada por Grossberg, surgiram três classes de arquitetura ART, sendo: ART1, ART2 e ART *Fuzzy*. Utilizando paradigmas de aprendizado tanto supervisionado como não supervisionado, foram apresentadas variações destas redes, compreendendo importantes adaptações, necessárias para lidar com problemas específicos. Dentre os modelos mais utilizados, pode-se destacar:

- ART1: modelo precursor dentro da teoria, apresenta treinamento não-supervisionado. Caracteriza-se por reconhecer padrões de entrada binários arbitrariamente apresentados; (CARPENTER; GROSSBERG, 1987a);
- ART2: também utiliza treinamento não-supervisionado; difere da ART1 ao se apresentar como um modelo dedicado ao processamento de dados tanto binários como analógicos (contínuos) (CARPENTER; GROSSBERG, 1987b);
- ART3: treinamento do tipo não-supervisionado; utiliza o conceito de transmissores químicos empregados para busca hierárquica (CARPENTER; GROSSBERG, 1990);

- ART *Fuzzy*: com treinamento não-supervisionado, caracteriza-se por incorporar em sua arquitetura cálculos baseados na lógica *fuzzy* (CARPENTER; GROSSBERG; ROSEN, 1991);
- ARTMAP: utiliza treinamento supervisionado; é composta por dois módulos ART interconectados por um módulo *Inter-ART*. É capaz de processar padrões de entrada tanto binários quanto analógicos (CARPENTER; GROSSBERG; REYNOLDS, 1991);
- ARTMAP *Fuzzy*: arquitetura e treinamento similar a rede ARTMAP convencional, se destaca por apresentar todos os cálculos fundamentados na lógica *fuzzy* (CARPENTER et al., 1992c);
- ARTMAP Gaussiana: utiliza treinamento supervisionado e é apresentada como uma síntese de um classificador gaussiano e de uma rede neural da família ART (WILLIAMSON, 1996).

A Figura 3 apresenta uma estrutura em árvore, que permite observar desde o surgimento da teoria da ressonância adaptativa até algumas das principais implementações de redes ART.



Fonte: Próprio autor

4.3.3 Redes ART e ART *Fuzzy*

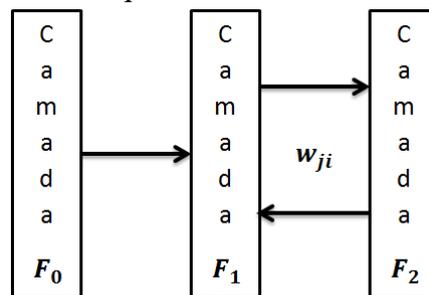
A diferença fundamental entre a rede ART e a ART *Fuzzy* se baseia no operador de interseção, que, na primeira é definido por \cap , já na segunda é definido por \wedge (operador AND da teoria dos conjuntos nebulosos (ZADEH, 1965)). No caso da rede nebulosa, essa diferença permite o processamento de dados analógicos e digitais (LOPES, 2005).

A RNA ART *Fuzzy*, incluindo as demais redes da família que utilizam aprendizado não-supervisionado, baseiam-se em um sistema de aprendizagem competitivo capaz de auto-organizar categorias de forma estável em resposta a sequências arbitrárias de padrões de entrada (analógicos ou binários). Os padrões são agrupados em categorias (clusters), ilustradas através de hiper-retângulos. A formação

dos clusters se dá a partir de regras abstratas sobre a distribuição dos dados, permitindo certos níveis de generalização (AMORIM, 2006; CARPENTER et al., 1992).

A arquitetura da rede ART *Fuzzy* é composta pelas camadas F_0 (entrada), F_1 (comparação) e F_2 (reconhecimento), que em conjunto permitem comparar a similaridade entre um novo padrão apresentado a rede com outro já representado por determinado neurônio. As camadas conforme apresentadas na Figura 4, exibem conexões, sendo a interconexão entre as camadas F_1 e F_2 formadas por pesos tipo *feedforward* (w_i) e *feedback* (w_j), encarregados do armazenamento das informações.

Figura 4 - Arquitetura da RNA ART *Fuzzy*



Fonte: Próprio autor

Os tópicos a seguir apresentam com maior detalhamento as funções específicas de cada uma das camadas da rede ART, que permitem realizar o processo de classificação de padrões. Os passos apresentados são descritos em (CARPENTER et al., 1992).

4.3.3.1 Camada F_0

A camada de entrada é a responsável pelo tratamento do vetor de entrada M -dimensional a , realizando a normalização e codificação em complemento, formando o vetor de atividade I (CARPENTER et al., 1992).

$$I = \frac{a}{|a|} \quad (4.10)$$

em que:

a : vetor de entrada;

I : $[I_1, I_2, I_3 \dots I_M]$ (normalizado);

$|\cdot|$: função norma, definida por:

$$|p| = \sum_{i=1}^M |p_i| \quad (4.11)$$

A codificação em complemento do padrão de entrada a é realizada de tal forma que produz um vetor de entrada I do tipo $2M$ -dimensional, ou seja:

$$I = (\mathbf{a}, \mathbf{a}^c) = (a_1, a_2, a_3, \dots, a_M, a_1^c, a_2^c, a_3^c, \dots, a_M^c) \quad (4.12)$$

em que:

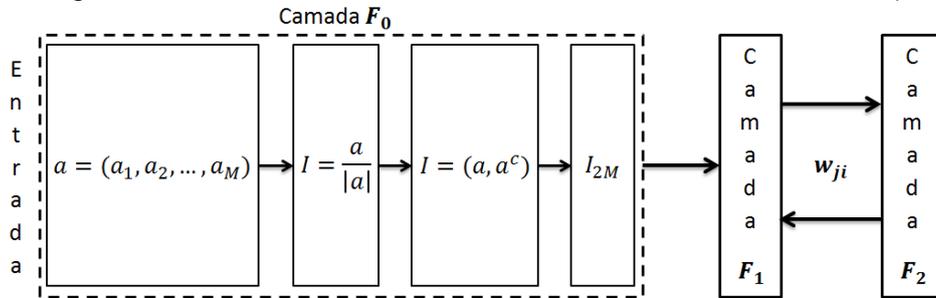
$$a_i^c = 1 - a_i \quad (4.13)$$

A codificação em complemento tem por objetivo de preservar a amplitude da informação:

$$|I| = |\mathbf{a}, \mathbf{a}^c| = \sum_{i=1}^M a_i + \left(M - \sum_{i=1}^M a_i \right) = M \quad (4.14)$$

A saída da camada de entrada é dada por seu vetor de atividade I que segue em direção a camada F_1 . A Figura 5 descreve os passos realizados na camada F_0 .

Figura 5 - Processos realizados na camada de Entrada - ART *Fuzzy*



Fonte: Próprio autor

4.3.3.2 Camada F_1

A camada de comparação recebe sinais tanto da camada F_0 quanto da camada F_2 . O vetor de ativação (ou ativação) de F_2 (que representa categoria ativa) é denotado por $y = [y_1, y_2, y_3, \dots, y_N]$, sendo N o número de categorias criadas em F_2 (CARPENTER et al., 1992).

O vetor de ativação da camada de comparação (F_1) é dado por $x = x_1, x_2, x_3, \dots, x_{2M}$, sendo resultante da operação \wedge (AND *fuzzy*) entre o vetor de atividade I (da camada de entrada) e o vetor de pesos w_j relativo a categoria ativa J de F_2 , obedecendo a equação (4.15):

$$x = \begin{cases} I & \text{se } F_2 \text{ estiver inativa} \\ I \wedge w_j & \text{se existir categoria ativa em } F_2 \end{cases} \quad (4.15)$$

em que:

I : vetor de ativação de F_0 ;

w_j : vetor de pesos entre F_1 e F_2 ;

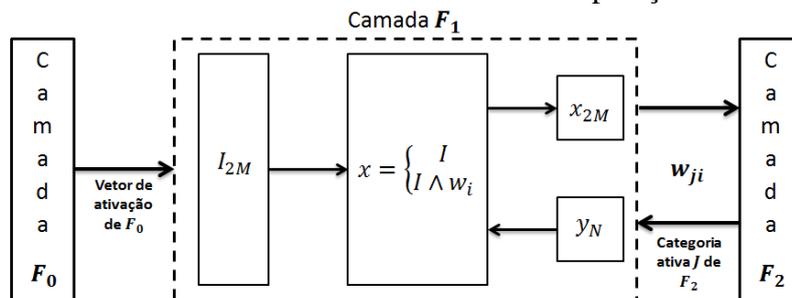
\wedge : operador AND *fuzzy*, definido por:

$$(p \wedge q)_i = \min(p_i, q_i) \quad (4.16)$$

para quaisquer vetores M -dimensionais p e q .

Os pesos w_{ji} entre as camadas F_1 e F_2 representam a memória da rede. Em w_{ji} , i (de dimensão $2M$) representa os elementos da categoria armazenada, e j (de dimensão N) representa as categorias, a dimensão N aumenta sempre que uma nova categoria é armazenada em w pela RNA ART *Fuzzy*. A Figura 6 apresenta os procedimentos realizados em F_1 .

Figura 6 - Processos realizados na camada de comparação - ART *Fuzzy*



Fonte: Próprio autor

4.3.3.3 Camada F_2

A camada F_2 contém N unidades de processamento capazes de classificar os padrões de treinamento em categorias de reconhecimento aprendidas pela rede (CARPENTER et al., 1992).

A seleção da categoria ativa J ocorre de tal forma que, o valor da função de escolha T_j é atribuído a cada unidade de processamento em F_2 , de maneira que a unidade J com maior valor atribuído de acordo com (4.17) é a vencedora (em caso de empate, é selecionada a categoria J com $T_j = \max_j \{T_j\}$ e menor índice J).

O neurônio selecionado em F_2 envia para F_1 seu protótipo por meio do vetor de ativação $y = y_1, y_2, y_3, \dots, y_N$ (de F_2) e o vetor de pesos w_j referente ao neurônio selecionado.

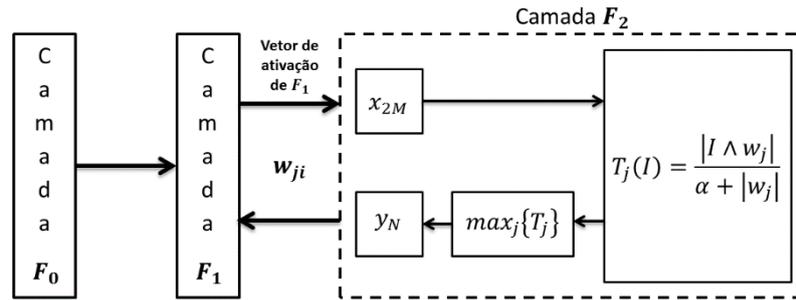
A função de escolha é definida por:

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|} \quad (4.17)$$

em que:

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

A Figura 7 descreve os procedimentos realizados na camada F_2 .

Figura 7 - Processos realizados na camada de reconhecimento - ART *Fuzzy*

Fonte: Próprio autor

4.3.3.4 Ressonância

A similaridade verificada entre o padrão de entrada e a categoria ativa J (em F_2) determinam se a rede ativa o estado de ressonância ou de reset. Se a função de combinação ($\frac{|I \wedge w_j|}{I}$) entre o padrão de entrada e a categoria ativa J atender ao critério de vigilância apresentado em (5.18), a rede entra em estado de ressonância, caso contrário ocorre o reset. No reset o neurônio rejeitado é desabilitado (fixando $T_j = 0$) até o final do processo corrente, evitando assim que uma seleção persista na mesma categoria durante o processo de busca. Este processo de busca continua desabilitando neurônios de F_2 , até que seja encontrado um neurônio que atenda o critério de vigilância. Na impossibilidade de encontrar um neurônio que atenda ao critério, uma nova categoria é criada em F_2 associada ao padrão de entrada corrente (LOPES, 2005; CARPENTER et al., 1992).

O parâmetro de vigilância ρ determina o grau de semelhança exigido no teste de similaridade, se o valor da comparação for maior ou igual ao de ρ , o padrão de entrada é atribuído a categoria ativa, e por consequência, no processo de aprendizagem, caso contrário, ocorre reset. O teste de similaridade é definido por:

$$\frac{|I \wedge w_j|}{|I|} \geq \rho \quad (4.18)$$

em que:

I : vetor de ativação da camada F_0 ;

w_j : vetor de pesos da categoria ativa;

ρ : parâmetro de vigilância ($\rho \in [0,1]$).

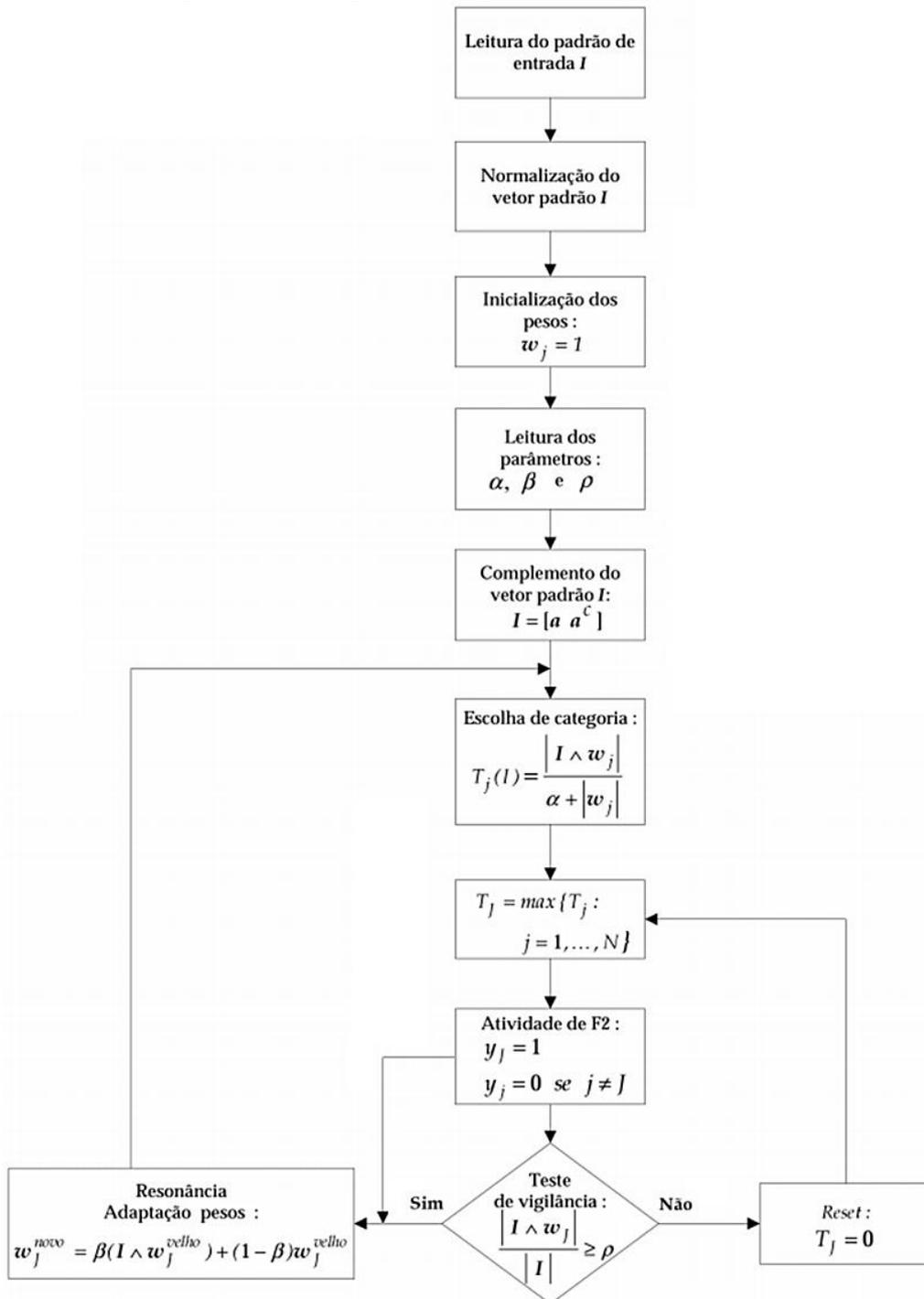
4.3.3.5 Aprendizagem

Quando o processo de busca de categoria e o teste de vigilância são finalizados, o vetor de pesos w_j é adaptado de acordo com a equação (LOPES, 2005; CARPENTER et al., 1992):

$$w_j^{novo} = \beta(I \wedge w_j^{velho}) + (1 - \beta)w_j^{velho} \quad (4.19)$$

em que $\beta \in [0,1]$ é a velocidade da aprendizagem; quando $\beta = 1$, tem-se o treinamento rápido, permitindo que o sistema se adapte com maior rapidez ($w_j^{novo} = I \wedge w_j^{velho}$). O treinamento lento ocorre quando $\beta < 1$, neste caso os pesos se adaptam lentamente sobre muitos ciclos (4.19).

Figura 8 - Fluxograma completo da RNA ART Fuzzy



Fonte: (LOPES, 2005)

4.3.3.6 Fluxograma

Os processos apontados no tópico anterior (4.3.3) e seus subtópicos, podem ser vistos sintetizados no fluxograma apresentado na Figura 8.

4.3.3.7 Parâmetros Importantes

A escolha adequada dos valores para os parâmetros da rede é de grande importância, uma vez que estes são capazes de impactar diretamente no desempenho da rede, indo do desenvolvimento da aprendizagem até os custos computacionais envolvidos. Considerando uma rede ART, citam-se os parâmetros (WEENINK, 1997; CARPENTER; GROSSBERG, 2003):

- α : parâmetro de escolha, assume valores de acordo com $\alpha > 0$. Quando uma entrada de dados pertence a mais de uma categoria (ilustradas através de hiper-retângulos), o parâmetro de escolha faz com que o menor dos hiper-retângulos seja o escolhido;
- β : representa a taxa de treinamento, assumindo valores de acordo com $0 \leq \beta \leq 1$. Um valor baixo para β implica em um treinamento lento, por outro lado, se $\beta = 1$, indica um treinamento rápido;
- ρ : representa o parâmetro de vigilância, sendo que $\rho \in (0,1]$. Implica diretamente na quantidade de clusters que serão formados, uma vez que, define o grau máximo de variação entre os componentes que podem pertencer a uma mesma categoria.

4.3.4 Rede ARTMAP e ARTMAP Fuzzy

O treinamento da RNA ARTMAP *Fuzzy* é supervisionado e auto-organizável, e seus cálculos são fundamentados na lógica nebulosa (CARPENTER et al., 1992c). A rede destina-se à aproximação de funções lineares multidimensionais.

Sua arquitetura é formada por dois módulos *ART Fuzzy* (ART_a e ART_b), de tal forma que, o módulo ART_a é encarregado pelo processamento do vetor de entrada, e o módulo ART_b é responsável pelo processamento do vetor da saída desejada.

O módulo de memória associativa *Inter – ART* recebe entradas do módulo ART_a (referente à conexão associativa $J \rightarrow K$, onde: J indica a categoria ativa em ART_a , e K a ativa em ART_b) e a categoria ativa do módulo ART_b . Este módulo conta com um mecanismo auto-regulador, chamado *Match-Traking*, que busca verificar se há casamentos entre as categorias de ART_a com as de ART_b (CARPENTER; GROSSBERG, 1992a). A arquitetura da rede ARTMAP *Fuzzy* pode ser vista na Figura 9.

em que $|a| = \sum_i |a_i|$;

- Complemento do padrão de entrada (A) e saída (B):

$$\bar{a}_i^c = 1 - a_i \text{ e } \bar{b}_i^c = 1 - b_i; \quad (4.21)$$

- Vetor de entrada (A) e saída (B) normalizados e complementados:

$$I^e = [a \ \bar{a}] \text{ e } I^s = [b \ \bar{b}]; \quad (4.22)$$

- Inicialização das matrizes de pesos com valor igual a 1:

$$w_j^a = 1; w_k^b = 1; w_j^{ab} = 1;$$

indicando que não existem categorias ativas;

- Leitura dos parâmetros da rede:

$$\alpha, \beta, \rho_a, \rho_b, \rho_{ab}, \varepsilon;$$

Módulo ART_b Fuzzy:

- Cálculo da função T_k^b :

$$T_k^b(I^s) = \frac{|I^s \wedge w_k^b|}{\alpha + |w_k^b|}; \quad (4.23)$$

- Escolha da categoria (K) para o módulo ART_b Fuzzy:

$$T_K^b = \max\{T_k^b : k = 1, \dots, N_b\}; \quad (4.24)$$

- Verifica se atende o critério de vigilância do módulo ART_b Fuzzy:

$$|x^b| = \frac{|I^s \wedge w_K^b|}{|I^s|} \geq \rho_b; \quad (4.25)$$

- Se não atender: Reset: $T_K^b = 0$;

- Se atender:

- Ressonância, adaptação de pesos do módulo ART_b Fuzzy:

$$w_K^{novo} = \beta(I^s \wedge w_K^{velho}) + (1 - \beta)w_K^{velho}; \quad (4.26)$$

e

- Cálculo do vetor de atividade em F_2 :

$$y_K^b = [y_1^b, y_2^b, \dots, y_N^b],$$

sendo: (4.27)

$$y_k^b = \begin{cases} 1, & \text{se } k = K \\ 0, & \text{se } k \neq K \end{cases}$$

- Segue para o módulo de Match Tracking:

Módulo ART_a Fuzzy:

- Cálculo da função T_j^a :

$$T_j^a(I^e) = \frac{|I^e \wedge w_j^a|}{\alpha + |w_j^a|}; \quad (4.28)$$

- Escolha de categoria (J) para o módulo ART_a Fuzzy:

$$T_J^a = \max\{T_j^a : j = 1, \dots, N_a\}; \quad (4.29)$$

- Verificação do critério de vigilância do módulo ART_a Fuzzy:

$$|x^a| = \frac{|I^e \wedge w_J^a|}{|I^e|} \geq \rho_a \quad (4.30)$$

- Se atendido: Cálculo do vetor de atividade em F_2 : $y_j^a = [y_1^a, y_2^a, \dots, y_N^a]$, sendo:

$$y_j^a = \begin{cases} 1, & \text{se } j = J \\ 0, & \text{se } j \neq J \end{cases} \quad (4.31)$$

- Segue para o módulo de Match Tracking:

Módulo Match Tracking:

- Verifica se atende o critério de vigilância do módulo *Inter – ART*:

$$|x^{ab}| = \frac{|y^b \wedge w_j^{ab}|}{|y^b|} \geq \rho_{ab} \quad (4.32)$$

- Se não atendido: incrementa o parâmetro de vigilância:

$$\rho_a = \frac{|I^e \wedge w_j^a|}{|I^e|} + \varepsilon \quad (4.33)$$

- Reset: $T_j^a=0$;
- Se Match Tracking atendido:
 - Ressonância, adaptação de pesos do módulo ART_a Fuzzy:

$$w_j^{novo} = \beta(I^e \wedge w_j^{velho}) + (1 - \beta)w_j^{velho} \quad (4.34)$$

- Atualização dos pesos do módulo *Inter – ART*:

$$W_{JK}^{ab} = [y_1^{ab}, y_2^{ab}, \dots, y_N^{ab}], \quad (4.35)$$

sendo:

$$y_{jk}^{ab} = \begin{cases} 1, & \text{se } j = J; k = K \\ 0, & \text{se } j = J; k \neq K \end{cases}$$

- Verifica se todos os pares de treinamento foram processados;
- Se não: $\rho_a = \rho_a^{original}$, continua em ART_a Fuzzy e ART_b Fuzzy;
- Se sim: fim do treinamento.

sendo:

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

β : taxa de treinamento $[0, 1]$;

ρ_a : parâmetro de vigilância do módulo ART_a $(0, 1]$;

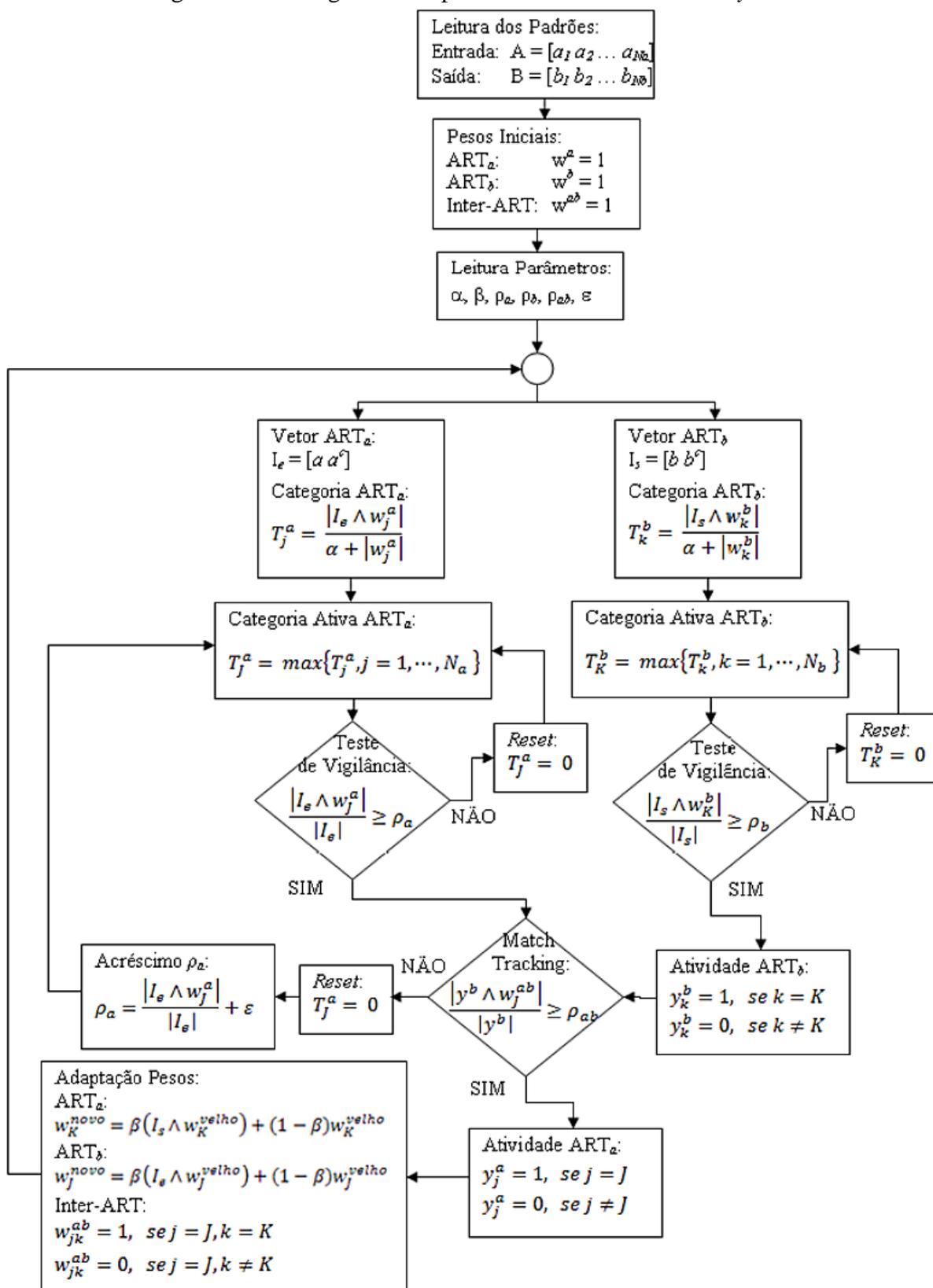
ρ_b : parâmetro de vigilância do módulo ART_b $(0, 1]$;

ρ_{ab} : parâmetro de vigilância do módulo ART_{ab} $(0, 1]$;

ε : incremento do parâmetro ρ_a ;

n : número de padrões de entrada.

A RNA ARTMAP Fuzzy também pode ser visualizada através de seu fluxograma completo, apresentado na Figura 10.

Figura 10 - Fluxograma completo da rede ARTMAP *Fuzzy*

Fonte: Baseado em (LOPES, 2005)

4.3.4.2 Codificação da Rede ARTMAP Fuzzy

A rede ARTMAP Fuzzy utilizada nas simulações apresentadas neste trabalho foi codificada utilizando o ambiente interativo para computação numérica MATLAB® versão R2015a (64-bit), nome que também designa a linguagem utilizada. De maneira a obter maior controle sobre os processos realizados pela rede, funções prontas não verificáveis foram substituídas por codificação equivalente.

Os testes foram realizados em um mesmo computador, configurado com Windows 8 Professional 64-bit, processador Intel® Core™ i7-3770 8M Cache 3.90GHz e 8GB de memória RAM.

4.3.4.3 Vetores de Entrada e Saída

Os vetores de entrada da RNA são formados por um dígito binário que indica se o dia é útil ou final de semana/feriado, em seguida um vetor de dígitos binários representa a posição na curva de carga para o valor da saída esperada, considerando o treinamento. Isto é, para uma curva de carga diária formada por 24 dados, a posição na curva de carga é representada por um valor entre 00001 e 11000.

Após o dígito que indica se o dia é útil ou não, e do vetor que aponta a posição na curva de carga, aparecem as cargas históricas consideradas. O número de cargas anteriores a saída esperada é definida durante os testes. Deste modo, um vetor de entrada (ve) é formado por:

$$ve = [td, pc, L_{(t-q)}, \dots, L_{(t-2)}, L_{(t-1)}]; \quad (4.36)$$

em que:

td : indica se é dia útil ou final semana/feriado;

pc : corresponde a um vetor binário que indica a posição na curva de carga, ou seja, posição temporal;

L_t : representa a carga histórica no instante t ;

q : indica o número de cargas históricas consideradas em ve .

O padrão de saída por sua vez é formado pelos valores de carga relativos aos instantes futuros $t + n$.

5 CENÁRIOS DE ESTUDO E ENSAIOS DE PREVISÃO

Nesta seção são apresentados os cenários de estudos e os ensaios de previsão realizados, considerando a ordem de agregação das cargas elétricas, de modo que os níveis mais agregados são apresentados primeiro.

É abordada a previsão de cargas elétricas de intercâmbio de energia, de um conjunto de subestações, de uma subestação, de uma *microgrid*, de um conjunto de prédios, e, por fim, em nível predial.

Considerando todos os cenários apresentados, realizou-se previsão de carga com os dados originais, e com aqueles submetidos a remoção de ruído por SSA. Considerando que o nível de ruído tende a ser mais evidente quanto menor o nível de agregação das cargas, é esperado que a remoção de ruído permita melhora nos resultados preditivos nos níveis menos agregados, e que seu impacto seja menor nos níveis mais agregados.

Os parâmetros considerados para a rede ARTMAP *Fuzzy* são destacados em cada cenário de estudo. O modelo estatístico SARIMA e a RNA MLP são aplicados como *benchmark*, e suas configurações seguem o apresentado no APÊNDICE A e nas seções 4.2 e 4.2.2. A remoção de ruído por SSA deu-se conforme apresentado nas seções 3.5.3 e 3.5.3.1.

Os dias considerados alvo de previsão foram removidos dos conjuntos históricos utilizados para os testes e previsão. De modo que foi adotado como valor alvo dentro dos conjuntos de teste, o dia (ou conjunto de dias) equivalente imediatamente anterior aos valores removidos, sendo necessário apenas que este pertença a mesma classe de dia daquele removido, isto é, dia útil ou final de semana/feriado, podendo, caso contrário, ser removido uma quantidade maior de dias até que a condição seja atendida.

5.1 Intercâmbio de Energia Elétrica

5.1.1 Cenário Gerador dos Dados

Neste cenário de estudos são considerados dados reais de intercâmbio de grandes blocos de energia elétrica. Devido a dificuldade em se obter uma autorização clara para a divulgação de informações complementares, os dados dos agentes envolvidos serão preservados.

O intercâmbio de energia se dá respeitando o excedente de geração para um determinado mercado consumidor, que comercializa a parcela não utilizada. A geração se dá a partir de hidrelétrica, e entre as diversas regras estabelecidas, há uma que delimita o uso de volume hídrico máximo em um intervalo de 24 horas, esta regra tem por objetivo evitar impactos ambientais e navegacionais. Portanto, há necessidade de se

prever o intercâmbio de energia elétrica para um horizonte futuro de 24 horas, de modo que tais dados sejam utilizados para guiar tarefas que permitam melhor aproveitamento do sistema e de recursos, respeitando resoluções vigentes.

Os dados de intercâmbio de energia foram submetidos aos três previsores, a saber, rede ARTMAP *Fuzzy*, RNA MLP *Feedforward* e SARIMA. Foram realizadas previsões *multistep* de carga horária para dois horizontes considerados de importância para os agentes envolvidos, sendo estes, 24 e 72 horas.

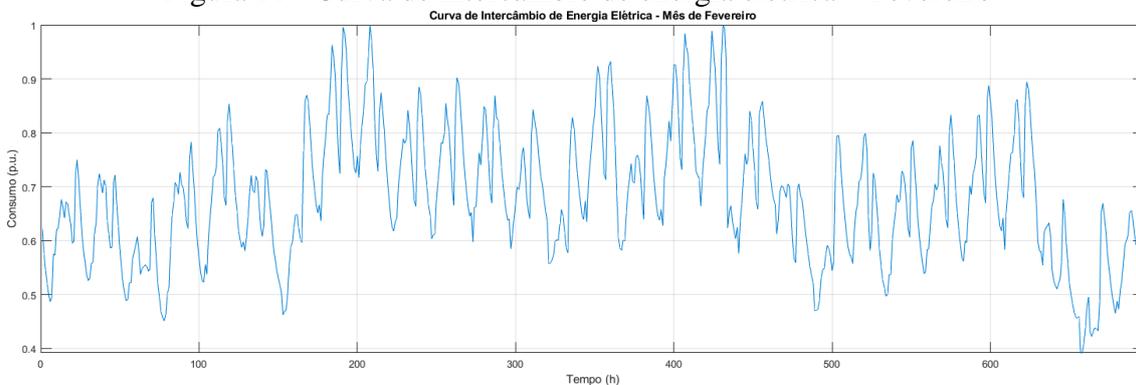
5.1.2 Ensaios de Previsão

5.1.2.1 Introdução

Foi considerada uma série histórica formada pelos 06 primeiros meses do ano de 2016 (janeiro a junho), composta pelo intercâmbio médio horário entre os agentes envolvidos. Os dias objeto de previsão foram removidos das séries históricas, no caso, dia 30 de junho para o horizonte de previsão de 24 horas, e os dias 28, 29 e 30 para a previsão de três curvas de carga diária.

Uma vez que os dados de intercâmbio dependem de diversos fatores, seu comportamento pode ser mais complexo de prever que dados de consumo em nível de subestação de energia. A Figura 11 apresenta a variação do intercâmbio para a o mês de fevereiro, que compõem o conjunto de dados adotado neste trabalho.

Figura 11 - Curva de intercâmbio de energia elétrica – Fevereiro



Fonte: Dados da pesquisa

5.1.2.2 Metodologia

Para previsão *multistep* de 24 horas a frente foram considerados 181 dias históricos, equivalentes a 4.344 dados de carga. Já para a previsão de 72 horas, foram considerados 4.296 dados, equivalente a 179 dias.

Uma *grid search* foi conduzida sobre um subconjunto de teste de modo a determinar a combinação de parâmetros a ser utilizada na etapa de previsão pela rede ARTMAP *Fuzzy*. Os parâmetros α , ρ_a , ρ_b , β , ρ_{ab} e ε foram definidos e variados de acordo com o apresentado na Tabela 2.

Tabela 2 - Parâmetros Testados para a RNA ARTMAP *Fuzzy*

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,003	1,323	0,66
ρ_a	0,93	0,97	0,02
ρ_b	0,995	0,999	0,002
β	1	1	-
ρ_{ab}	1	1	-
ε	0,001	-	-

A dimensão dos vetores de entrada foi estabelecida como 24 para a previsão de um dia, e, 72 para a previsão de três dias a frente.

5.1.2.3 Resultados

A Tabela 3 apresenta os resultados de MAPE para os três métodos de previsão testados, incluindo aquele obtido com a RNA ARTMAP *Fuzzy* com a remoção de ruído por SSA. Com o horizonte preditivo de 24 horas o método estatístico apresentou seu melhor desempenho, com um MAPE de 3,68. A ARTMAP *Fuzzy* obteve um desempenho superior que a rede MLP, com um MAPE de 9,91 para o conjunto de dados originais, e um pouco superior, 9,81 com os dados filtrados por SSA.

Tabela 3 - Resultados de MAPE para horizonte preditivo de 24 horas - Intercâmbio de EE

Previsor	MAPE
MLP Feedforward	14,38
SARIMA	3,68
ARTMAP Fuzzy	9,91
ARTMAP Fuzzy+ SSA	9,81

Os resultados de MAPE apresentados na Tabela 4 referem-se a previsão de carga de 03 dias futuros. Neste horizonte, como esperado, o desempenho de todos os previsores foi afetado negativamente, uma vez que se tratam de previsões *multistep*, o previsor não pode ser ajustado com a correção do erro a cada novo dado previsto, uma vez que o método não toma conhecimento do erro cometido no instante anterior. Neste cenário o desempenho marginal foi registrado pela rede MLP, com um MAPE de 23,35; a RNA ARTMAP *Fuzzy* obteve um desempenho superior que o SARIMA tanto com os dados originais, quanto com os dados submetidos a remoção de ruído, com MAPEs de 6,85 e 6,04, respectivamente, enquanto o SARIMA obteve um MAPE de 12,90.

Tabela 4 - Resultados de MAPE para horizonte preditivo de 72 horas - Intercâmbio de EE

Método	MAPE
MLP Feedforward	23,35
SARIMA	12,90
ARTMAP Fuzzy	6,85
ARTMAP Fuzzy + SSA	6,04

5.1.3 Considerações

Apesar da série de dados retratar o intercâmbio de grandes blocos de energia elétrica, as volatilidades envolvidas, como a disponibilidade mutável de energia para intercâmbio por parte do gerador (uma vez que seu consumo se apresenta em considerável parte atrelado as mudanças de temperatura), a dependência de indicadores hídricos, a demanda variável do agente que adquire o excedente, entre outras, acabam por impor maiores dificuldades aos previsores utilizados.

Foi possível verificar que ao treinar a rede com os dados submetidos a remoção de ruído com auxílio do SSA, houve pequena melhora preditiva para ambos os horizontes considerados, estes resultados são coerentes ao verificado em outros trabalhos da literatura especializada.

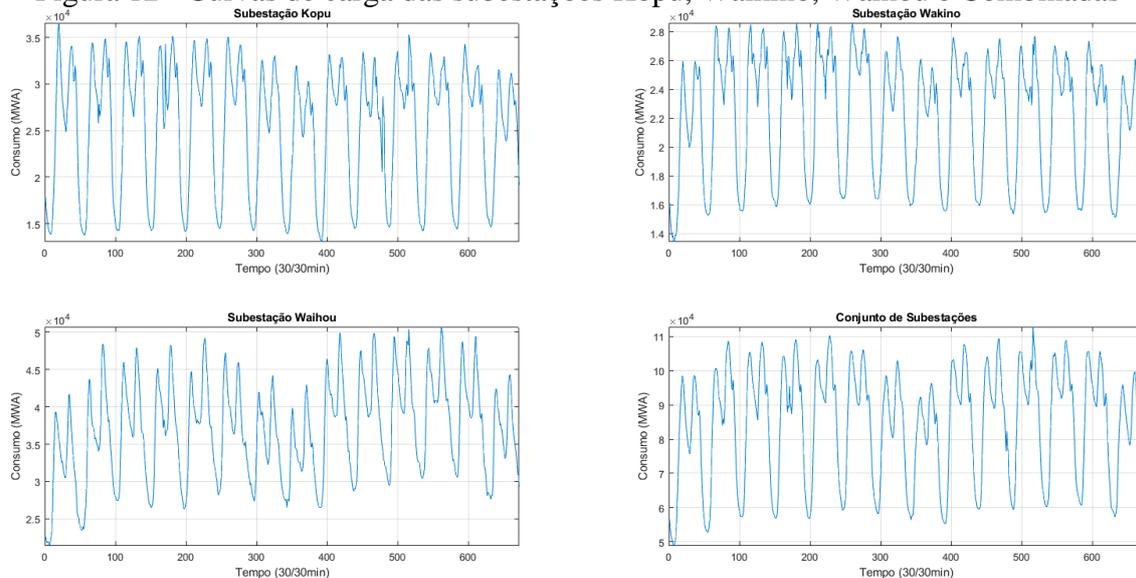
5.2 Conjunto de Subestações de Energia Elétrica

5.2.1 Cenário Gerador dos Dados

São utilizados dados reais de consumo de EE em nível de subestação, cedidos pela *New Zealand Electricity Commission*, extraídos do *Centralized Dataset - CDS* (2010). Neste conjunto de dados, são fornecidas informações de consumo de energia ativa medidos de meia em meia hora, sendo assim, uma curva de carga diária é formada por 48 medições distintas. Além das informações de carga, também são fornecidas informações referentes ao dia do mês, mês e ano. Os dados representam a agregação das cargas de três subestações, Kopu, Waikino e Waihou, que atendem uma região de Waikato, Nova Zelândia.

Foram utilizados dados de consumo de EE referentes aos meses de janeiro, fevereiro e março de 2009, totalizando 4320 padrões. A agregação de curvas de carga tende a formar uma curva global mais suavizada, por vezes, permitindo obter melhores resultados de previsão. A Figura 12 apresenta curvas de carga de cada uma das subestações, além daquela resultante da agregação das três, os dados apresentados são referentes aos dias 4 a 17 de janeiro de 2009, representando duas semanas completas (domingo a sábado).

Figura 12 - Curvas de carga das subestações Kopu, Waikino, Waihou e Combinadas



Fonte: Dados da pesquisa

Ao observar a escala lateral dos gráficos da Figura 12, é possível verificar que as subestações atendem a demandas bastante distintas. Os períodos que correspondem a um dia são facilmente identificados, assim como os períodos que correspondem a uma semana. Os domingos e sábados se destacam pela queda na demanda, principalmente para as subestações Waikino, Waihou e Conjunto de Subestações.

5.2.2 Ensaio de Previsão

5.2.2.1 Introdução

Neste cenário também foram considerados os três previsores, aplicados na previsão *multistep* de cargas elétricas.

Durante previsões realizadas nos conjuntos de teste, verificou-se que a RNA ARTMAP *Fuzzy* e o modelo SARIMA apresentavam comportamentos singulares, isto é, a rede se apresentava pessimista, com saídas em considerável parte abaixo dos valores reais, por outro lado, o modelo SARIMA apresentava comportamento otimista, de modo que suas previsões por diversas vezes ultrapassavam os dados reais, portanto, foi traçada uma estratégia para combinar ambas as previsões. O comportamento da rede pode ser explicado em parte pela dificuldade que arquiteturas de redes neurais artificiais apresentam em ultrapassar máximos históricos. O SARIMA por sua vez, pode ter identificado uma tendência que se mostrou menos acentuada para o período previsto.

A estratégia utilizada neste cenário ocorre devido a singularidade do comportamento observado por parte do SARIMA e da rede ARTMAP *Fuzzy*, não sendo expandido para séries históricas de outros cenários de demanda investigados neste trabalho.

Excepcionalmente para o cenário que considera os dados históricos das subestações de Kopu, Waikino, Waihou e Agregadas (seções 5.2 e 5.3), os valores de p, q e P, Q do SARIMA foram limitados a 06.

5.2.2.2 Metodologia

Foram considerados três horizontes históricos distintos para embasar as tarefas de previsão, conforme apresentados na Tabela 5.

Tabela 5 - Divisão dos dados em subconjuntos – Subestações de EE

Horizonte	Período	Nº de Padrões
1	01/03/2009 a 30/03/2009 (30 dias)	1440
2	01/02/2009 a 30/03/2009 (58 dias)	2784
3	01/01/2009 a 30/03/2009 (89 dias)	4272

A curva de carga do dia 31 de março de 2009 não aparece na Tabela 5, uma vez que ela é objeto de previsão, não fazendo parte dos subconjuntos utilizados para modelagem/treinamento e previsão.

A divisão dos dados em três subconjuntos históricos tem por objetivo permitir investigar o impacto da extensão das séries históricas sobre a capacidade preditiva das técnicas empregadas.

Os parâmetros da rede foram ajustados a partir de uma *grid search*, onde variações para α , ρ_a e ρ_b foram testadas. Os parâmetros ε , β e ρ_{ab} foram mantidos fixos, de acordo com a Tabela 6.

Tabela 6 - Parâmetros testados para a RNA ARTMAP *Fuzzy*

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,663	1,323	0,33
ρ_a	0,93	0,97	0,01
ρ_b	0,995	0,999	0,001
β	1	1	-
ρ_{ab}	1	1	-
ε	0,001	-	-

A dimensão dos vetores de entrada foi determinada a partir de testes na etapa de validação, que considerou horizontes entre 48 e 144 dados, variando de 12 em 12. As configurações dos modelos SARIMA são apresentadas juntamente com os resultados de previsão na sequência.

As previsões obtidas pela rede neural ARTMAP *Fuzzy* e pelo modelo SARIMA foram combinadas, de modo que ambas as previsões tivessem o mesmo peso sobre a curva de carga resultante. Portanto, adotou-se a média aritmética simples para tal combinação, de modo que as previsões S_j e S_k foram atribuídas a S_i de acordo com $S_i = w.S_j + (1 - w).S_k$, onde $w = 0,5$.

A definição de $w = 0,5$ se deu com base nos comportamentos da rede ARTMAP *Fuzzy* e do modelo SARIMA nos conjuntos de teste. É de conhecimento que o comportamento de cada previsor dependente das características das séries históricas utilizadas, por este motivo, é descrita uma abordagem genérica que permite definir variação do peso w para cada instante da série prevista, considerando para tanto, o desempenho do previsor no conjunto de validação. Deste modo, as previsões são combinadas de acordo com $S_i = w_j.S_j + w_k.S_k$, onde w_j é definido de acordo com o desempenho do primeiro previsor no conjunto de teste, e, w_k é definido de acordo com o desempenho do segundo previsor, também no conjunto de teste. Apesar desta abordagem não garantir que ocorra melhora preditiva, é esperado que os previsores se saiam melhor em determinadas porções da curva de carga, e tal característica seja mantida com dados históricos do mesmo tipo. É necessário que o previsor com melhor desempenho no conjunto de validação receba um peso w maior, de modo que $w_j + w_k = 1$ seja respeitado.

5.2.2.3 Resultados de Previsão

A Tabela 7 apresenta os valores de MAPE obtidos a partir da previsão com os subconjuntos das subestações agregadas. A combinação das previsões da RNA ARTMAP *Fuzzy* com as do modelo SARIMA permitiram obter os menores valores de MAPE com os dois primeiros subconjuntos de dados. O SARIMA obteve o melhor resultado geral considerando o subconjunto mais extenso de dados. A rede MLP apresentou os piores resultados de MAPE, com destaque para subconjunto 1 e 2, onde obteve valores de MAPE próximos a 5. Sem considerar a combinação das previsões, o modelo SARIMA obteve os melhores resultados preditivos com o menor e o maior subconjunto de dados históricos, enquanto a ARTMAP *Fuzzy* se mostrou superior com o conjunto intermediário.

Tabela 7 - Resultados de previsão – Subestações de EE agregadas

Horizonte Histórico	Previsor	MAPE
1) 01/03/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,4911
	ARTMAP <i>Fuzzy</i> + SSA	2,4694
	SARIMA(1,0,4)x(6,1,6)48	1,3522
	SARIMA + ARTMAP <i>Fuzzy</i>	1,2621
	MLP <i>Feedforward</i>	4,8005
2) 01/02/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	1,9105
	ARTMAP <i>Fuzzy</i> + SSA	1,8803
	SARIMA(1,1,2)x(5,1,6)48	2,3457
	SARIMA + ARTMAP <i>Fuzzy</i>	1,8435
	MLP <i>Feedforward</i>	2,6748
3) 01/01/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,2651
	ARTMAP <i>Fuzzy</i> + SSA	2,3217
	SARIMA(2,0,1)x(5,1,6)48	1,6882
	SARIMA + ARTMAP <i>Fuzzy</i>	1,5359
	MLP <i>Feedforward</i>	4,6314

5.2.3 Considerações

Neste cenário a rede MLP obteve alguns de seus melhores resultados, com valores de MAPE entre 2,6748 e 4,8005, porém mesmo assim seu desempenho foi marginal quando comparado com os outros previsores. O SARIMA obteve resultados de MAPE entre 1,3522 e 2,3457, já os resultados da rede ARTMAP *Fuzzy* ficaram entre 1,9105 e 2,4911. Os resultados superiores obtidos a partir da combinação das previsões da RNA ARTMAP *Fuzzy* e do SARIMA variaram com MAPEs entre 1,2621 e 1,8435.

Considerando a extensão dos conjuntos históricos utilizados para a previsão, os resultados para o SARIMA foram ambíguos, uma vez que seus menores valores de MAPE foram obtidos com o maior e o menor conjunto; os melhores desempenhos das redes ARTMAP *Fuzzy* e MLP foram obtidos com o subconjunto intermediário, equivalente a dois meses de dados históricos.

A utilização da técnica SSA para remoção de ruído das séries históricas submetidas para o treinamento da rede ARTMAP *Fuzzy*, permitiram obter previsões com MAPEs moderadamente menores para os subconjuntos 1 e 2 ao serem comparados aos resultados observados com os dados originais.

5.3 Subestação de Energia Elétrica

5.3.1 Cenário Gerador dos Dados

Para este cenário também são utilizados dados reais de consumo de EE cedidos pela *New Zealand Electricity Commission*, os dados descrevem curvas de carga das subestações Kopu, Waikino e Waihou. As medições das grandezas elétricas também são realizadas em intervalos de 30min, de modo que uma curva de carga diária é composta por 48 medições.

Para o treinamento da rede foram utilizados dados das subestações Kopu, Waikino e Waihou, separadamente. São considerados 4320 padrões para cada uma das subestações, correspondentes aos meses de janeiro, fevereiro e março de 2009.

Cada uma das subestações é responsável por atender distintas regiões, com diferentes índices demográficos, e, consumidores de diversas classes, sendo assim, cada qual impõem determinado nível de complexidade para a tarefa de previsão.

5.3.2 Ensaio de Previsão

5.3.2.1 Introdução

Os três previsores foram utilizados na tarefa de previsão *multistep* de cargas elétricas em nível de subestação.

Considerando o comportamento otimista do modelo SARIMA e pessimista da rede ARTMAP *Fuzzy* no cenário formado pela agregação das subestações, agora individualmente estudadas, foram realizados testes para verificar se tal comportamento se mantinha para as subestações separadamente, esta hipótese foi confirmada, e com base no comportamento peculiar observado, a estratégia de combinar as saídas do modelo estatístico com a saída da rede da teoria da ressonância adaptativa foi novamente adotada. A composição da previsão foi realizada de modo a permitir equidade entre a influência de cada previsor.

Tendo em vista que os conjuntos de dados apresentam a mesma dimensão que o formado pela agregação das 03 subestações consideradas no cenário anterior (seção 5.2), os presentes dados foram submetidos a mesma divisão apresentada na Tabela 5. As 48 cargas que formam a curva de carga diária de 31 de março de 2009 é objeto de previsão.

5.3.2.2 Metodologia

Os três horizontes históricos de cada uma das subestações foram submetidos aos três previsores, sendo assim, cada previsor foi aplicado na realização de três previsões *multistep* de 01 dia a frente para cada subestação.

A escolha dos parâmetros da rede se deu através de uma *grid search*, onde foram testados os parâmetros apresentados na Tabela 8. As configurações dos modelos SARIMA são apresentadas juntamente com os resultados de previsão na sequência.

Tabela 8 - Parâmetros testados para a RNA ARTMAP *Fuzzy*

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,663	1,323	0,33
ρ_a	0,93	0,97	0,01
ρ_b	0,995	0,999	0,001
β	1	1	-
ρ_{ab}	1	1	-
ϵ	0,001	-	-

5.3.2.3 Resultados de Previsão

A Tabela 9 apresenta os resultados de previsão para os 48 dados de carga referentes ao dia 31 de março de 2009, obtidos a partir dos três subconjuntos de dados da subestação Kopu. Os melhores resultados de MAPE foram observados ao combinar as previsões da rede ARTMAP *Fuzzy* com o modelo SARIMA. Considerando os resultados de previsão obtidos com os subconjuntos 1 e 2, é possível constatar que isoladamente a ARTMAP *Fuzzy* e o SARIMA apresentaram resultados próximos, porém com pequena vantagem para a técnica estatística. O resultado marginal da rede da família ART foi observado com o maior subconjunto de dados, que considera 4272 padrões para o treinamento da rede.

Tabela 9 - Resultados de previsão com os subconjuntos de dados da subestação Kopu

Horizonte Histórico	Previsor	MAPE
1) 01/03/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	3,4704
	SARIMA(4,1,5)x(6,1,6)48	3,1695
	SARIMA + ARTMAP <i>Fuzzy</i>	2,1184
	MLP <i>Feedforward</i>	4,5821
2) 01/02/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,5056
	SARIMA(1,0,1)x(5,1,6)48	2,9572
	SARIMA + ARTMAP <i>Fuzzy</i>	1,9507
	MLP <i>Feedforward</i>	5,1179
3) 01/01/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	4,7041
	SARIMA(1,0,2)x(5,1,6)48	2,6444
	SARIMA + ARTMAP <i>Fuzzy</i>	2,1224
	MLP <i>Feedforward</i>	12,7630

A Tabela 10 apresenta os valores de MAPE para os três subconjuntos de dados históricos da subestação Waikino. Novamente todos os melhores resultados de MAPE foram obtidos a partir da combinação das previsões da RNA ARTMAP *Fuzzy* e do modelo SARIMA. Ao contrário da situação anterior, a rede ARTMAP *Fuzzy* se saiu melhor com os maiores subconjuntos de dados. Isoladamente o SARIMA obteve os menores valores de MAPE.

Tabela 10 - Resultados de previsão com os subconjuntos de dados da subestação Waikino

Horizonte Histórico	Previsor	MAPE
1) 01/03/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	3,9183
	SARIMA(2,1,5)x(6,0,5)48	2,5639
	SARIMA + ARTMAP <i>Fuzzy</i>	2,3731
	MLP <i>Feedforward</i>	3,5714
2) 01/02/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	3,2758
	SARIMA(1,1,6)x(6,1,6)48	1,9463
	SARIMA + ARTMAP <i>Fuzzy</i>	1,2694
	MLP <i>Feedforward</i>	6,7555
3) 01/01/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,0714
	SARIMA(2,0,6)x(5,1,5)48	1,5166
	SARIMA + ARTMAP <i>Fuzzy</i>	1,1947
	MLP <i>Feedforward</i>	6,7871

Na Tabela 11 são apresentados os resultados para a subestação Waihou. A combinação das previsões das redes ARTMAP *Fuzzy* e dos modelos SARIMA permitiram obter os menores resultados de MAPE nos dois menores subconjuntos de dados. O SARIMA obteve o melhor resultado geral com o subconjunto mais extenso. Considerando o subconjunto 02 de dados sem a combinação das técnicas, a RNA ARTMAP *Fuzzy* obteve o melhor resultado de previsão.

Tabela 11 - Resultados de previsão com os subconjuntos de dados da subestação Waihou

Horizonte Histórico	Previsor	MAPE
1) 01/03/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,7169
	SARIMA(6,1,6)x(5,1,6)48	2,0941
	SARIMA + ARTMAP <i>Fuzzy</i>	1,1396
	MLP <i>Feedforward</i>	13,3480
2) 01/02/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	2,9440
	SARIMA(4,1,3)x(4,1,5)48	3,4154
	SARIMA + ARTMAP <i>Fuzzy</i>	1,7204
	MLP <i>Feedforward</i>	4,2950
3) 01/01/2009 a 30/03/2009:	ARTMAP <i>Fuzzy</i>	3,3621
	SARIMA(6,1,3)x(4,1,5)48	1,7498
	SARIMA + ARTMAP <i>Fuzzy</i>	2,3853
	MLP <i>Feedforward</i>	10,8749

A Tabela 12 apresenta o desempenho (MAPE) da rede ARTMAP *Fuzzy* com os dados históricos submetidos a redução de ruído via SSA.

Tabela 12 - Resultados de previsão da rede ARTMAP *Fuzzy* com os subconjuntos de dados submetidos a remoção de ruído - Kopu, Waikino e Waihou

Horizonte Histórico	Kopu	Waikino	Waihou
1) 01/03/2009 a 30/03/2009:	3,4655	3,8337	2,7435
2) 01/02/2009 a 30/03/2009:	2,4874	3,2465	2,8240
3) 01/01/2009 a 30/03/2009:	4,5611	2,1825	3,2809

5.3.3 Considerações

As combinações das previsões somente não representaram o melhor resultado de MAPE em um dos nove cenários. Sem considerar a combinação de previsões, o modelo SARIMA obteve a maior parcela dos melhores resultados, seguido pela rede ARTMAP *Fuzzy*.

Considerando todas as subestações e os subconjuntos de dados históricos, os resultados obtidos com a rede MLP *Feedforward* foram marginais, somente superando um dos previsores no subconjunto de dados 01 da subestação Waikino, onde a RNA ARTMAP *Fuzzy* apresentou seu segundo pior resultado.

A Tabela 13 apresenta as médias dos resultados de MAPE obtidos a partir dos dados das subestações agregadas (seção 6.3.2), também são apresentadas as médias dos resultados de MAPE verificados a partir dos dados de cada uma das subestações individualmente. Para a média foram considerados os valores de MAPE obtidos nos três subconjuntos de cada cenário.

É possível verificar que a agregação dos dados apresenta efeito positivo sobre os previsores, de modo que as melhores médias de MAPE foram verificadas no cenário que considera a agregação das cargas das três subestações. Neste cenário a RNA

ARTMAP *Fuzzy* apresentou alguns de seus melhores resultados; os resultados da MLP se apresentaram mais estáveis; o modelo SARIMA também apresentou bom desempenho.

Tabela 13 - Média do MAPE para as subestações Waihou, Waikino, Kopu e Agregadas

Previsor	Média MAPE			
	Agregadas	Waihou	Waikino	Kopu
Todas as Técnicas	2,4000	4,1704	3,1036	4,0088
ARTMAP <i>Fuzzy</i>	2,2222	3,0076	3,0885	3,5600
SARIMA	1,7953	2,4197	2,0089	2,9237
SARIMA + ARTMAP <i>Fuzzy</i>	1,5471	1,7484	1,6124	2,0638
MLP <i>Feedforward</i>	4,0355	9,5059	5,7046	7,4876

Ao observar o desempenho da RNA ARTMAP *Fuzzy* com os dados submetidos a redução de ruído através da decomposição da série original via SSA, excluindo a componente de ruído, foi possível observar pequena diminuição dos valores de MAPE para a maioria dos ensaios de previsão.

5.4 Microrrede

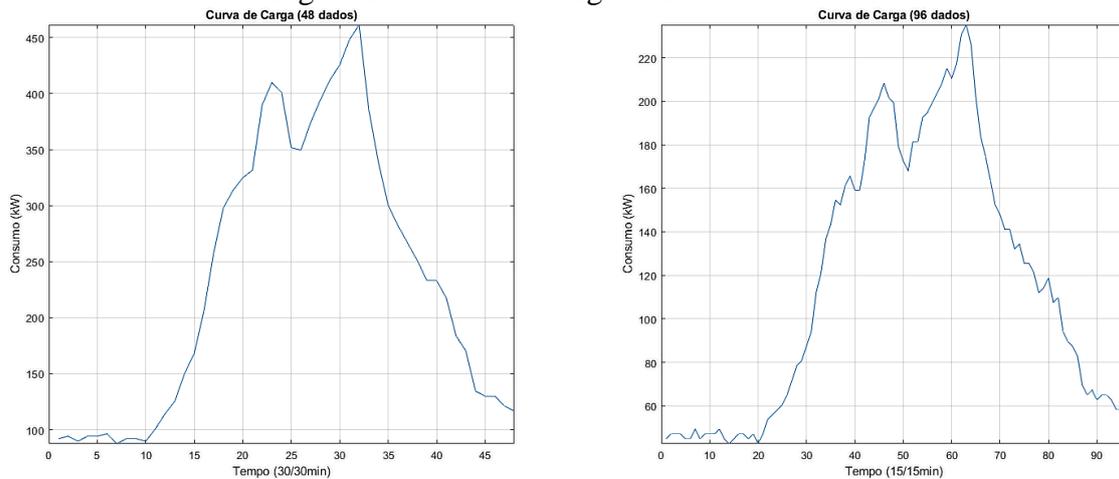
5.4.1 Cenário Gerador dos Dados

Nesta etapa são utilizados dados reais de consumo de energia elétrica de uma microrrede inteligente. As cargas provenientes desta estrutura são bastante distintas, formadas principalmente pela demanda de estabelecimentos comerciais, *startups*, centros de pesquisa, centros de ensino, laboratórios, escritórios administrativos e serviços. A origem dos dados será preservada.

Considerável parcela da carga é formada por equipamentos de condicionamento de ar, devido principalmente as altas temperaturas registradas, com média anual superior a 20°C, e temperaturas máximas médias no verão que ultrapassam 30°C, capazes de gerar sensação térmica próxima a 40°C nos períodos mais quentes.

Neste trabalho são consideradas curvas de carga diária com duas resoluções, a primeira formada por 48 dados de carga, e a segunda por 96 medições distintas do consumo de energia ativa igualmente espaçadas no tempo, com intervalos de 30min e 15min respectivamente. São utilizados dados de consumo de energia referentes ao primeiro trimestre de 2017, especificamente do dia 01 de janeiro a 31 de março de 2017. A Figura 13 apresenta duas curvas de carga, a primeira com 48 dados e a segunda com 96 dados, referentes a última sexta feira de março de 2017; ao verificar as imagens é possível constatar o maior nível de detalhe daquela formada por 96 dados.

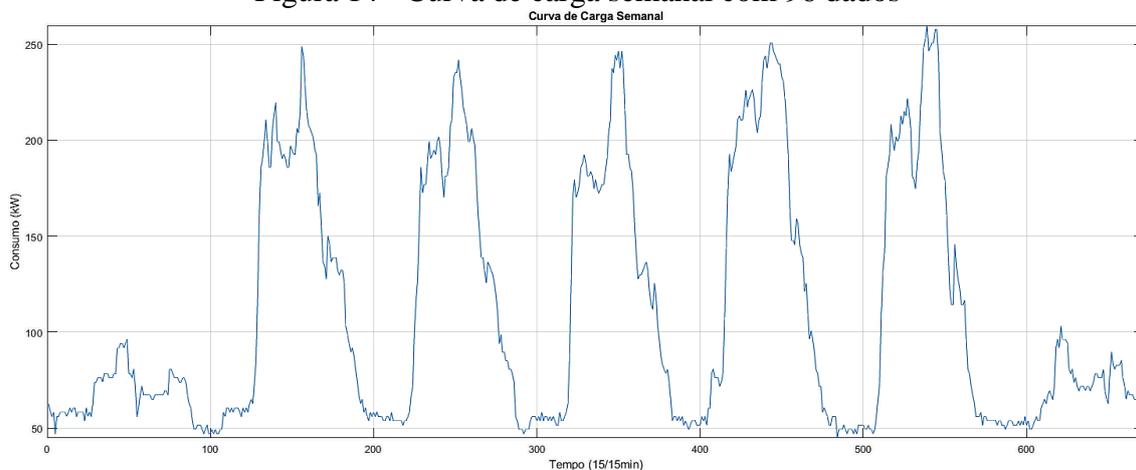
Figura 13 - Curva de carga com 48 e 96 dados



Fonte: Dados da pesquisa

Ao analisar os dados de consumo da microrrede é possível constatar que as cargas dos dias da semana apresentam considerável diferença e variação, de maneira esperada, o consumo de energia é bastante inferior nos finais de semana quando comparado com dias úteis. A Figura 14 apresenta o gráfico da curva de carga semanal referente os dias 19 a 25 de março de 2017 (domingo a sábado).

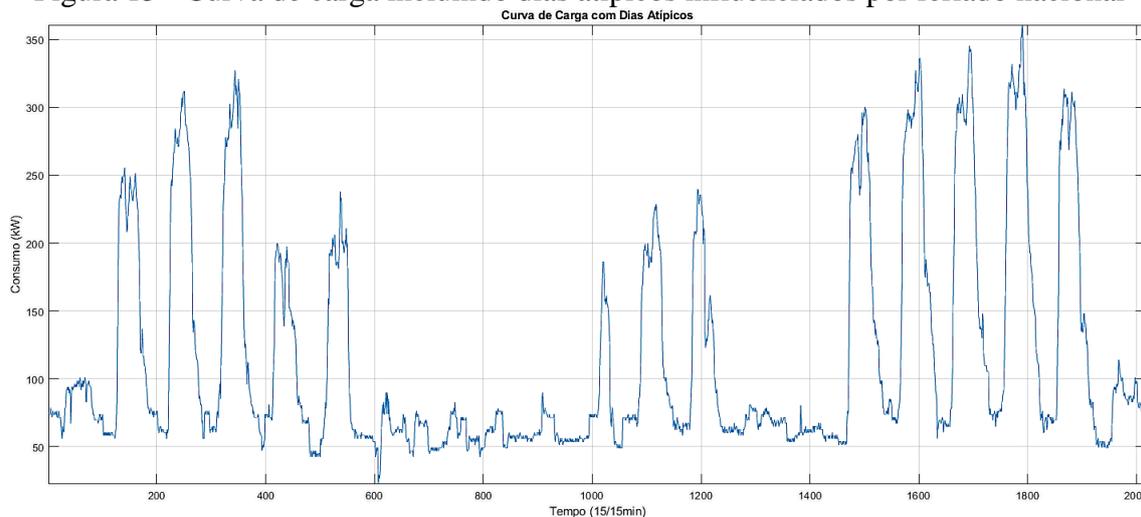
Figura 14 - Curva de carga semanal com 96 dados



Fonte: Dados da pesquisa

Além das diferenciações entre os dias da semana, existem outras ainda mais complexas, que são os dados atípicos, normalmente identificados por sua discrepância com a maior parcela dos dados históricos, tal tipo de fenômeno é causado por mudanças bruscas nas rotinas; este tipo de dados está presente na série histórica utilizada. A Figura 15 apresenta a semana que precede o feriado de Carnaval, a semana que inclui o feriado e a semana imediatamente posterior a ele (19 de fevereiro a 11 de março). Estes tipos de dias não devem ser confundidos com *outliers*, mesmo que alguns possam ser classificados como raridades. A maioria destes dias podem ser facilmente contextualizados, explicados a partir da ocorrência de determinados eventos/fenômenos.

Figura 15 - Curva de carga incluindo dias atípicos influenciados por feriado nacional



Fonte: Dados da pesquisa

Ao analisar a Figura 15 é possível observar a discrepância entre os dias apresentados; e que o efeito não se limita aos dias do feriado nacional, sendo possível verificar a ocorrência de uma influência pré-feriado nos dias anteriores ao Carnaval, e um efeito pós-feriado, afetando os dias imediatamente posteriores, com efeito de redução na demanda por EE. As curvas de carga dos dias do feriado apresentam semelhanças com aquelas verificadas nos finais de semana (sábado e domingo), sobretudo devido a drástica redução de pessoas e atividades dentro da área da microrrede.

5.4.2 Ensaios de Previsão

5.4.2.1 Introdução

Neste cenário foram realizadas previsões *multistep* de cargas elétricas de curto prazo, com horizontes preditivos de 1, 3 e 7 dias. Os ensaios de previsão foram realizados com os dados originais, e com o conjunto submetido a redução de ruído por SSA.

5.4.2.2 Metodologia

Considerando os dados originais, e então, a série reconstruída sem o componente de ruído, os dados foram utilizados para o treinamento da RNA ARTMAP *Fuzzy*. Os parâmetros foram adaptados de acordo com o apresentado na Tabela 14.

A dimensão dos vetores de entrada utilizados para a etapa de previsão foi definida de acordo com os resultados obtidos na etapa de teste, onde foram verificados horizontes de 1 a 7 dias, com variação de 12 dados para a CC formada por 48 medições, e 24 para a CC formada por 96 dados. O critério de seleção utilizado para definir os parâmetros da rede e a dimensão do vetor de entrada foi o menor MAPE no conjunto de testes.

Tabela 14 - Parâmetros testados para a RNA ARTMAP *Fuzzy* - Microrrede

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,003	1,323	0,66
ρ_a	0,93	0,97	0,01
ρ_b	0,995	0,999	0,001
β	1	1	-
ρ_{ab}	1	1	-
ε	0,001	-	-

O SARIMA e a RNA *MultiLayer Perceptron Feedforward* foram utilizados como benchmark, sendo testados com os conjuntos originais e com aqueles suavizados por SSA. Os resultados foram comparados aos obtidos com a rede ARTMAP *Fuzzy* com os mesmos conjuntos de dados.

5.4.2.3 Resultados de Previsão

A Tabela 15 apresenta os valores de MAPE obtidos com cada um dos previsores nos três horizontes preditivos considerando os dados originais de consumo de EE, onde uma curva de carga diária é formada por 48 medições. Também são apresentados os resultados de MAPE obtidos pela rede ARTMAP *Fuzzy* com os dados suavizados via SSA.

Tabela 15 - Resultados de previsão com CC de 48 dados - Microrrede

Previsor	1 dia	3 dias	7 dias
SARIMA	8,3100	19,2500	36,3600
RNA MLP <i>Feedforward</i>	15,5000	37,6000	57,7000
RNA ARTMAP <i>Fuzzy</i>	11,6700	12,8400	45,0400
RNA ARTMAP <i>Fuzzy</i> com SSA	11,4300	12,5600	33,8900

Na Tabela 16 são apresentados valores de MAPE obtidos a partir do conjunto de dados original, onde uma curva de carga diária é formada por 96 dados. Além daqueles obtidos a partir da aplicação do SSA na tarefa de remoção de ruído dos dados, submetidos a previsão via RNA ARTMAP *Fuzzy*.

Tabela 16 - Resultados de previsão com CC de 96 dados - Microrrede

Previsor	1 dia	3 dias	7 dias
SARIMA	10,9000	22,6000	38,6000
RNA MLP <i>Feedforward</i>	14,1000	35,6000	56,0000
RNA ARTMAP <i>Fuzzy</i>	11,4700	14,7300	21,3800
RNA ARTMAP <i>Fuzzy</i> com SSA	11,4400	14,2100	15,1300

5.4.3 Considerações

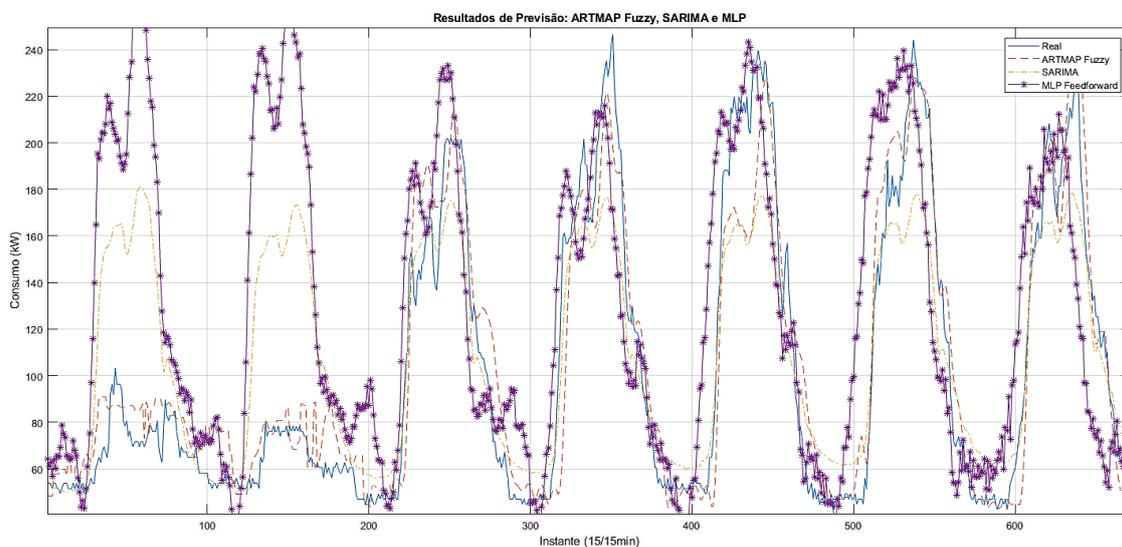
Ao analisar os resultados de previsão, é possível verificar que a rede ARTMAP *Fuzzy* acompanhada da redução de ruído por SSA obteve melhores resultados que os demais métodos utilizados como *benchmark*, a exceção ocorre para o horizonte preditivo de um dia, onde o SARIMA obteve melhores resultados, este fato se deve em

parte a alta correlação do dia previsto com os dias imediatamente anteriores, possibilitando uma boa aproximação pela combinação linear característica deste modelo, situação não verificada com os outros horizontes preditivos.

Devido a natureza dos dados, a modelagem é bastante dificultada, isto ocorre por dois motivos principais, o primeiro é o menor nível de agregação dos dados, e o segundo diz respeito a presença de dias atípicos na série de dados. O menor nível de agregação causa aumento relativo da aleatoriedade da série, significando também maiores níveis de ruído, isto se deve a decisões individuais (imprevisíveis) causarem maior impacto na curva de carga. A presença de dados atípicos é um desafio a parte, devido ao fato dos previsores normalmente utilizados não estarem preparados para incorporar diretamente conhecimento prévio à modelagem, tais como feriados e dias atípicos no geral; nesta situação específica, a rede ARTMAP Fuzzy mostrou superioridade, apresentando ainda um desempenho melhorado pela utilização de SSA na remoção de ruído, que proveu maior capacidade de generalização para a rede.

A Figura 16 permite observar uma das situações em que a RNA ARTMAP Fuzzy foi capaz de representar o formato das curvas reais, conseguindo diferenciar dias úteis de dias não úteis (ou, atípicos), ao contrário dos demais métodos.

Figura 16 - Previsão de 07 dias a frente com dados originais - Curva de carga com 96 dados



Fonte: Dados da pesquisa

O horizonte de 7 dias deve ser analisado com cautela, haja vista presença de diferentes tipos de dia no intervalo a ser previsto, onde o SARIMA e a rede MLP Feedforward não foram capazes de se ajustar adequadamente, e, uma vez diminuída a capacidade preditiva, a remoção de ruído apresenta pouco impacto.

5.5 Conjunto de Prédios

5.5.1 Cenário Gerador dos Dados

O complexo de prédios que forma a *microgrid* que provém os dados de consumo de EE para este estudo, está localizada a 180 metros acima do nível do mar, em uma região de clima subtropical temperado superúmido. A origem dos dados será preservada.

As medições das grandezas elétricas consideradas neste estudo são provenientes da agregação das cargas de 05 prédios. As curvas de carga destes prédios são compostas principalmente pela demanda de estabelecimentos comerciais, *start-ups*, centros de pesquisa, laboratórios, escritórios administrativos, serviços e estabelecimentos alimentícios. Os dados foram aferidos por medidores eletrônicos instalados junto aos transformadores que atendem cada um dos prédios, e então transportados através de uma rede com proteção contra interferências eletromagnéticas (que utiliza cabos *Shielded Twisted Pair - STP*) até o servidor de armazenagem.

Foram considerados dados de carga referente ao período 01 de julho a 31 de julho de 2013, correspondendo a 2976 padrões. As medições das grandezas elétricas foram realizadas em intervalos de 15 min, de modo que uma curva de carga diária é formada por 96 observações.

5.5.2 Ensaios de Previsão

5.5.2.1 Introdução

Os ensaios de previsão foram realizados utilizando os três previsores. O dia alvo de previsão, 31 de julho (quarta feira), foi removido do conjunto utilizado para os testes e previsão. A rede ARTMAP *Fuzzy* foi utilizada tanto com o conjunto original de dados, quanto com aquele submetido a redução de ruído por SSA.

5.5.2.2 Metodologia

A escolha dos parâmetros da rede ARTMAP *Fuzzy* foram realizados a partir de uma *grid search* no subconjunto de teste, onde foram testadas combinações dos parâmetros apresentados na Tabela 17.

Tabela 17 - Parâmetros testados para a RNA ARTMAP *Fuzzy* - Conjunto de prédios

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,003	1,323	0,66
ρ_a	0,93	0,97	0,01
ρ_b	0,995	0,999	0,001
β	1	1	-
ρ_{ab}	1	1	-
ε	0,001	-	-

5.5.2.3 Resultados de Previsão

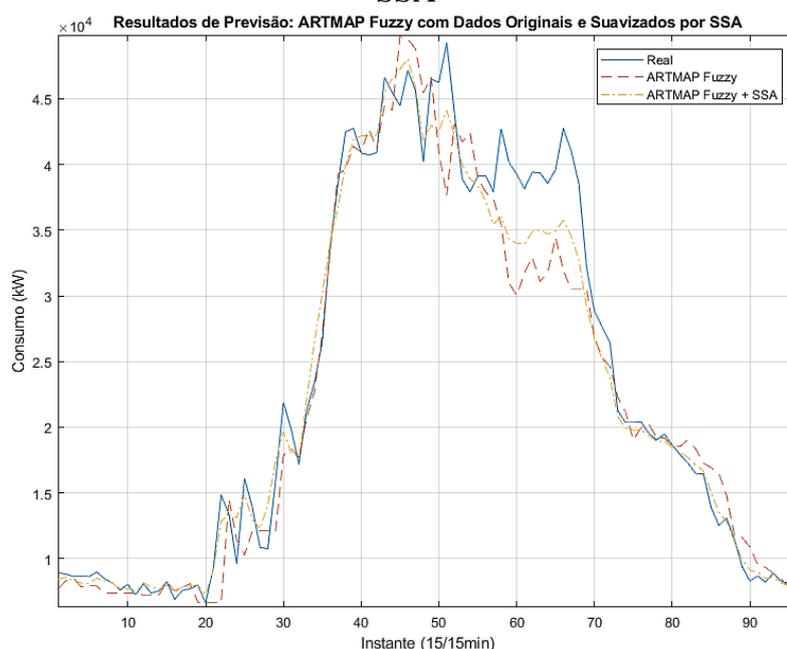
Na Tabela 18 são apresentados os valores de MAPE obtidos por cada um dos previsores utilizando a série agregada de demanda do conjunto de prédios.

Tabela 18 - Resultados de previsão de um dia a frente - Conjunto de prédios

Previsor	MAPE
SARIMA	23,6848
RNA MLP <i>Feedforward</i>	8,9357
RNA ARTMAP <i>Fuzzy</i>	10,1203
RNA ARTMAP <i>Fuzzy</i> + SSA	9,0474

A Figura 17 apresenta as previsões obtidas pela rede ARTMAP *Fuzzy* com o conjunto de dados original e com aquele submetido a redução de ruído via SSA, além da curva de carga real alvo da previsão.

Figura 17 - ARTMAP *Fuzzy* com previsão utilizando dados originais e suavizados por SSA



Fonte: Dados da pesquisa

5.5.3 Considerações

Observando os resultados de previsão é possível constatar o impacto positivo que a redução de ruído por SSA significou para a rede ARTMAP *Fuzzy*, reduzindo em mais de 10% o valor de MAPE quando comparado ao resultado obtido com os dados originais.

O melhor resultado de MAPE foi atingido pela rede MLP, já o desempenho marginal foi obtido pelo modelo SARIMA.

5.6 Nível Predial

5.6.1 Cenário Gerador dos Dados

Nesta etapa foram considerados dados de consumo de energia ativa de prédios individuais inseridos em uma *microgrid*. Os prédios estão localizados em um complexo formado por vários outros prédios que apresentam cargas bastante distintas.

Os dados de consumo de EE foram capturados por medidores eletrônicos instalados junto aos transformadores que atendem exclusivamente a cada um dos prédios considerados. Uma vez que os dados são medidos junto ao transformador, estes são transportados por uma rede de dados dedicada, concebida de maneira a apresentar alta proteção contra interferências eletromagnéticas, aumentando assim a confiança que os dados cheguem íntegros até o servidor de armazenagem.

As diversas cargas dos prédios são compostas pelas demandas de uma agência bancária, uma agência dos Correios, estabelecimentos comerciais, *start-ups*, centros de pesquisa, laboratórios, salas de aula, escritórios administrativos, serviços e estabelecimentos alimentícios.

O prédio 05 abriga um restaurante que serve diariamente em torno de 700 refeições, sendo que este número pode ser bastante maior quando são realizados eventos no complexo de prédios, e expressivamente menor em períodos de férias acadêmica, recessos e feriados. O restaurante apresenta uma área útil climatizada de aproximadamente 1.200m², que no verão exige a utilização de um expressivo número de condicionadores de ar. A utilização de refrigeradores, fornos elétricos e lava-louças também compõem as cargas do restaurante, tido como o principal consumidor de EE no referido prédio.

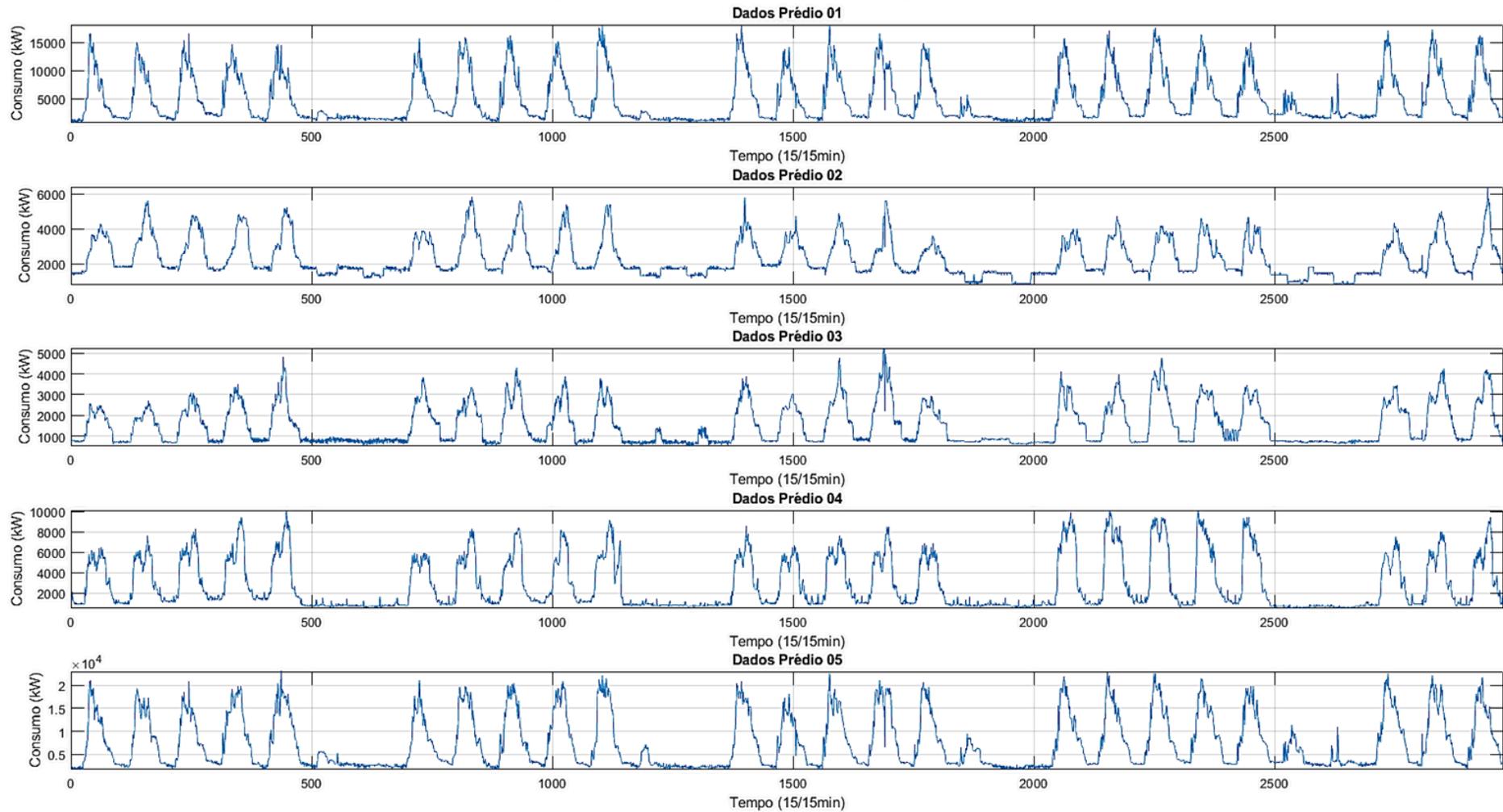
Foram considerados dados históricos de consumo de EE referente ao período compreendido entre 01 de julho a 31 de julho de 2013, onde cada curva de carga diária é formada por 96 medições equiespaçadas do consumo de EE.

5.6.2 Ensaios de Previsão

5.6.2.1 Introdução

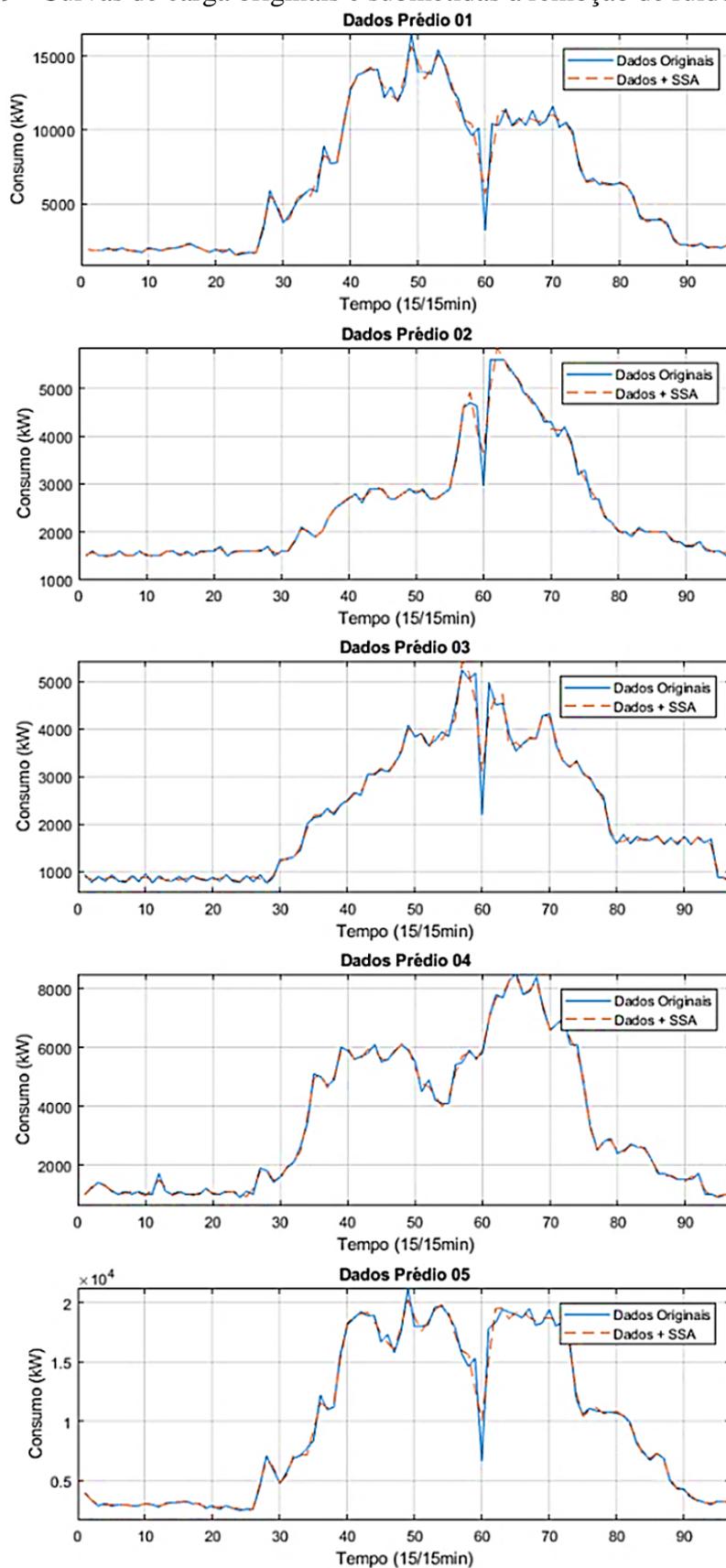
A previsão de cargas em nível predial foi realizada pela rede ARTMAP *Fuzzy*, além das duas outras técnicas utilizadas como *benchmark*, todos os métodos foram aplicados na previsão *multistep* de cargas elétricas de curto prazo, com horizonte preditivo de 96 dados a frente, equivalentes a uma CC diária. Os ensaios de previsão foram realizados com os dados originais, e com aqueles submetidos a redução de ruído por SSA.

Figura 18 - Dados históricos dos 05 prédios



Fonte: Dados da pesquisa

Figura 19 - Curvas de carga originais e submetidas à remoção de ruído por SSA



Fonte: Dados da pesquisa

A Figura 18 apresenta as séries históricas originais dos 05 prédios, onde é possível verificar seus distintos patamares de demanda. A diferenciação entre dia útil e final de semana é bastante evidente, e a ocorrência de eventos esporádicos e mudanças na rotina afetam diretamente as curvas de carga diária, como é possível observar com a alteração da demanda registrada para o segundo sábado e domingo do prédio 03 (13ª e 14ª curvas de carga diária, respectivamente).

A Figura 19 apresenta a terceira quinta-feira de julho de cada um dos 05 prédios, são apresentados os dados originais e aqueles resultantes da remoção de ruído por SSA. É possível verificar que a remoção de ruído por SSA causa pequena suavização nas séries de carga, porém sem alterar significativamente seu formato.

5.6.2.2 Metodologia

Os parâmetros da rede ARTMAP *Fuzzy* foram testados a partir de uma *grid search* no subconjunto de teste, sendo selecionada a combinação de parâmetros apresentados na Tabela 19 que permitiu obter os menores valores de MAPE dentro do referido subconjunto de testes.

Tabela 19 - Parâmetros testados para a RNA ARTMAP *Fuzzy* – Nível predial

Parâmetro	Valor Inicial	Valor Final	Incremento
α	0,003	1,323	0,66
ρ_a	0,93	0,97	0,01
ρ_b	0,995	0,999	0,001
β	1	1	-
ρ_{ab}	1	1	-
ε	0,001	-	-

5.6.2.3 Resultados de Previsão

A Tabela 20 apresenta os resultados de previsão obtidos a partir dos previsores SARIMA, MLP e ARTMAP *Fuzzy* para os dados históricos de cada um dos 05 prédios. São apresentados ainda os resultados obtidos a partir dos dados submetidos à remoção de ruído via SSA.

Tabela 20 - Resultados de previsão de um dia à frente - Nível predial

Previsor	Resultado Previsão (MAPE)				
	Prédio 01	Prédio 02	Prédio 03	Prédio 04	Prédio 05
SARIMA	16,2537	15,9588	29,7547	22,1944	27,5749
SARIMA + SSA	13,4333	15,9991	31,0014	21,4682	27,9278
MLP	13,2649	14,7233	16,6375	15,1227	30,1443
MLP + SSA	14,9661	14,6985	18,8250	15,1108	26,2832
ARTMAP <i>Fuzzy</i>	10,8003	30,8295	20,2417	10,8683	18,1162
ARTMAP <i>Fuzzy</i> + SSA	9,4255	30,6151	21,1053	9,9289	18,0129

5.6.3 Considerações

A utilização da remoção de ruído por SSA permitiu os melhores resultados preditivos em 04 dos 05 cenários, quando submetido à previsão via RNA ARTMAP *Fuzzy*. O SARIMA se saiu melhor com os dados originais, de maneira que a remoção de ruído somente permitiu resultados superiores com os dados dos prédios 01 e 04. Considerando a rede MLP, os dados suavizados permitiram obter 03 dos 05 melhores resultados. Ao contrário do verificado com a rede ARTMAP *Fuzzy*, o MLP apresentou variações mais significativas nos resultados ao considerar o conjunto original e aquele submetido a remoção de ruído, com destaque para os dados dos prédios 03 e 05; o SARIMA apresentou variação considerável do valor de MAPE com os dados do prédio 01.

6 CONCLUSÕES

Os resultados obtidos com a RNA ARTMAP *Fuzzy* na previsão de cargas da *microgrid* permitiram observar a superioridade da rede em diferenciar dias úteis e finais de semana de dias atípicos. Mesmo que esta característica tenha sido observada na maioria dos ensaios, e não em sua totalidade, evidencia-se que a RNA MLP não foi capaz de apresentar resultados próximos durante os experimentos, se comportando de maneira similar ao verificado com a técnica estatística.

A rede ARTMAP *Fuzzy* se mostrou superior a rede MLP, permitindo resultados mais assertivos na maioria dos testes realizados. Apesar da capacidade da rede em modelagem não linear, o SARIMA permitiu bons resultados em considerável parcela dos ensaios de previsão. Porém a RNA ARTMAP *Fuzzy* apresenta pontos desejáveis, como não necessitar ajustar um modelo para toda a série de dados, e a obtenção de resultados aceitáveis a partir de poucos testes em subconjuntos reduzidos.

Nos testes de previsão realizados no cenário de intercâmbio de energia, a rede ARTMAP *Fuzzy* apresentou resultados superiores para o período de 03 dias a frente, enquanto o SARIMA se saiu melhor com a previsão de um dia a frente. Os bons resultados para previsão de um dia se devem ao seu comportamento altamente correlacionado com as cargas dos dias imediatamente anteriores, permitindo boa aproximação pela combinação linear característica do modelo. A remoção de ruído por SSA permitiu pequena melhora de MAPE para a rede ARTMAP *Fuzzy*, provavelmente influenciado pela melhora na capacidade de generalização. Devido a diversos fatores externos, existe considerável flutuação nas curvas diárias de intercâmbio de energia, trazendo maiores desafios aos previsores utilizados, mesmo que tal cenário se destaque pelos dados de grandes blocos de energia.

Ao considerar as curvas de carga que atendem a região de Waikato, formadas a partir da agregação das subestações de Kopu, Waikino e Waihou, foram obtidos os seguintes resultados de MAPE: com o SARIMA, ficaram entre 1,3522 e 2,3457, com a RNA ARTMAP *Fuzzy* variaram entre 1,9105 e 2,4911, ficando entre 2,6748 e 4,8005 com a MLP, sendo os melhores resultados obtidos a partir da combinação das saídas da rede ARTMAP *Fuzzy* e do modelo SARIMA, assim como o melhor resultado geral foi obtido a partir da combinação de ambas as saídas, com um MAPE de 1,2621. Enquanto o SARIMA se saiu melhor com os conjuntos de dados menos e mais extenso, as duas arquiteturas de redes neurais se saíram melhor com o conjunto intermediário de dados. A suavização da série resultante do maior nível de agregação dos dados de consumo permitiu maior generalização por parte da rede ARTMAP *Fuzzy*, impactando diretamente nas previsões. A remoção de ruído da série histórica permitiu diminuição dos valores de MAPE considerando a rede ARTMAP *Fuzzy* nos subconjuntos 1 e 2.

Considerando dados históricos em nível de subestação, a combinação das saídas da RNA ARTMAP *Fuzzy* com a previsão utilizando modelagem SARIMA representou

todos os melhores resultados para as subestações Kopu e Waikino, obtendo também 02 dos 03 melhores resultados verificados na subestação de Waihou. A exceção se deu para o conjunto mais extenso de dados, onde o SARIMA permitiu um MAPE de 1,7498. Sem considerar a combinação de previsões, o SARIMA foi responsável por considerável parcela dos melhores resultados preditivos, seguido pela RNA ARTMAP *Fuzzy*, sendo a maioria dos resultados marginais obtidos pela rede MLP. Ao comparar os resultados de previsão individual das subestações com o conjunto formado pela agregação das três subestações, é possível constatar que os níveis mais agregados tenderam a permitir maior acurácia. A rede ARTMAP *Fuzzy* obteve melhores resultados para as subestações Kopu e Waihou quando foram utilizados os dois menores subconjuntos históricos, formados por 1440 e 2784 dados, respectivamente, e, obteve o melhor resultado para subestação Waikino ao considerar o subconjunto mais extenso de dados, e segundo menor MAPE com o subconjunto intermediário. A RNA MLP também se saiu melhor com os dois menores subconjuntos para as subestações Kopu e Waikino, além de obter seu melhor desempenho para subestação Waihou com o subconjunto intermediário. Os resultados com o SARIMA foram ambíguos, sendo que para subestação Kopu obteve seu melhor desempenho com o conjunto intermediário, e para as subestações Waihou e Waikino obteve melhor desempenho com o menor e o maior subconjunto histórico, respectivamente. Os dados históricos submetidos a remoção de ruído por SSA permitiram obter considerável parcela dos melhores resultados com a RNA ARTMAP *Fuzzy*, quando comparado com as previsões realizadas no conjunto original.

O resultado obtido para o conjunto de prédios foi singular, de modo que o modelo SARIMA foi o responsável pelo pior resultado de MAPE, e a RNA MLP pelo melhor. A rede ARTMAP *Fuzzy* obteve desempenho intermediário, com pequena redução do MAPE utilizando o conjunto de dados submetido a remoção de ruído.

No cenário que considera curvas de carga em nível predial, todos os previsores foram testados com os dados originais e aqueles submetidos a remoção de ruído por SSA. As duas arquiteturas de rede neural obtiveram a maioria dos melhores resultados com os dados suavizados, ao comparar com o desempenho nos dados originais, sendo verificado o inverso com o modelo SARIMA.

A remoção de ruído por SSA se mostrou vantajosa em considerável parcela dos experimentos realizados com a RNA ARTMAP *Fuzzy*. As séries suavizadas permitiram que a rede da família ART alcançasse melhores resultados preditivos em diversos cenários, beneficiada por uma melhor generalização.

O nível de agregação de carga exerce influência sobre os previsores, de modo que alguns dos maiores valores de MAPE foram registrados nos níveis menos agregados de consumo, enquanto alguns dos melhores resultados foram observados justamente nos níveis com maior agregação. Apesar de existir relação entre o nível de agregação dos dados e a acurácia dos previsores, esta não é determinada somente por

este, e sim pela quantidade de agentes capazes de exercer influência sobre as cargas, e quão aleatoriamente o fazem. Exemplo disto são alguns dos resultados obtidos em nível predial, que se aproximam daqueles obtidos a partir do conjunto de prédios. Os resultados observados para o intercâmbio de energia também devem ser destacados, uma vez que, apesar de se tratarem de grandes blocos de energia, os diversos fatores que influenciam esta transação prejudicam diretamente os previsores.

7 TRABALHOS FUTUROS

Considerando que em cenários reais nem sempre é possível que um especialista verifique as séries históricas de carga, assinalado dias típicos e atípicos, da mesma maneira que nem sempre dados exógenos são capazes de apontar determinados dias como atípicos, sendo tais informações frutos de situações isoladas, sendo assim, é importante que o sistema previsor seja capaz de assimilar tais informações, e tratar diretamente essas características, garantindo robustez e automação dos processos preditivos. Portanto, é de interesse que se investigue em quais níveis a RNA ARTMAP *Fuzzy* é capaz de fazer isso, e quais adaptações permitiriam potencializar esta característica.

Uma vez que a RNA ARTMAP *Fuzzy* apresentou considerável variação nos resultados de MAPE quando foram testados diferentes conjuntos históricos, e que em considerável parcela das vezes obteve bons resultados com os menores conjuntos históricos, um passo de interesse é automatizar o processo de escolha do horizonte de dados históricos a ser utilizado para o treinamento da rede, uma vez que quanto menor a série histórica utilizada neste processo, menor será a carga computacional. Uma estratégia possível inicia com o treinamento da rede com distintos subconjuntos históricos, cujo processo de otimização dos parâmetros se dá por *grid search* ainda mais esparsa, sendo eleito como o horizonte mais adequado aquele que permitir menores valores de MAPE nos testes de previsão realizados. O horizonte histórico escolhido deve ser extrapolado para uso no treinamento da rede principal, sendo então realizada uma *grid search* mais abrangente para escolha dos parâmetros, com objetivo de melhorar os resultados de previsão.

Os bons resultados obtidos utilizando a técnica SSA para tarefa de remoção de ruído permitem considerar a codificação da técnica como um módulo da rede ARTMAP *Fuzzy*, criando um sistema híbrido, permitindo testar variações da parametrização do SSA com objetivo de ganho em acurácia preditiva.

Outra possibilidade a ser considerada é a realização de previsões diretamente a partir das subséries somáveis resultantes da decomposição por SSA. De modo que determinados previsores sejam aplicados com determinadas subséries, deste modo, por exemplo, os componentes de tendência poderiam ser previstos utilizando SARIMA, uma vez que as redes neurais podem apresentar maior dificuldade em transpor máximos históricos, enquanto os componentes de sazonalidade poderiam ser previstos diretamente pela rede ARTMAP *Fuzzy*. A partir da soma das saídas de ambos os previsores é formada a previsão para a série temporal em questão.

Expandir os testes com a rede ARTMAP *Fuzzy* com conjuntos de dados que disponham de variáveis exógenas, como temperatura, umidade do ar, entre outras, de modo a investigar o impacto de um maior número de variáveis explicativas das séries

históricas. Destaca-se que não estavam disponíveis variáveis exógenas como temperatura, índice de precipitação e outras específicas para os cenários aqui estudados.

REFERÊNCIAS

ABREU, T.; SANTOS JUNIOR, C. R.; LOPES, M. L. M.; LOTUFO, A. D.; MINUSSI, C. R. Previsão de cargas elétricas utilizando uma rede Neural ARTMAP fuzzy com treinamento continuado. In: BRAZILIAN CONGRESS ON COMPUTATIONAL INTELLIGENCE, 13, 2017, Niterói. Disponível em: <cbic2017.org/papers/cbic-paper-88.pdf>. Acesso em: 20 nov. 2017.

AGARWAL, Y.; WENG, T.; GUPTA, R. K. Understanding the role of buildings in a smart microgrid. In: DESIGN, AUTOMATION & TEST IN EUROPE CONFERENCE & EXHIBITION, 2011, Grenoble. **Proceedings of the...** Grenoble: IEE, 2011. 6 p.

ALONSO, F. J.; DEL CASTILLO, J. M.; PINTADO, P. Application of singular spectrum analysis to the smoothing of raw kinematic signals. **Journal of Biomechanics**, New York, v. 38, n. 5, p. 1085-1092, 2005.

AMORIM, D. G. **Redes ART com categorias internas de geometria irregular**. 2006. 246 f. Tese (Doutorado) - Universidade de Santiago de Compostela, Departamento de Eletrônica e Computação, Santiago de Compostela, 2006.

ANEIROS, G.; VILAR, J. M.; RAÑA, P. Short-term forecast of daily curves of electricity demand and price. **International Journal of Electrical Power & Energy Systems**, Oxford, v. 80, p. 96-108, 2016.

BARNETT, V.; LEWIS, T. **Outliers in statistical data**. 3. ed. New York: John Wiley & Son, 1994.

BENITES, F.; SAPOZHNIKOVA, E. Improving scalability of ART neural networks. **Neurocomputing**, Libertyville, v. 230, p. 219-229, 2017.

BERGMEIR, C.; BENÍTEZ, J. M. **Neural networks using the stuttgart neural network simulator (SNNS)**. v. 0.4-10, p. 1-74. Disponível em: <<https://cran.r-project.org/web/packages/RSNNS/>>. Acesso em: 27 nov. 2017.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. Oakland: Scientific Computing Associates® Corp, 1970.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. 4. ed. Hoboken: John Wiley & Sons, 2008. 746 p.

BUITRAGO, J.; ASFOUR, S. Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs. **Energies**, Paris, v. 10, n. 40, p. 1–24, 2017. Disponível em: <<http://www.mdpi.com/1996-1073/10/1/40>>. Acesso em: 13 abr. 2018.

CARPENTER, G. A.; GROSSBERG, S. A self organizing neural network for supervised learning, recognition, and prediction. **IEEE Communications Magazine**, Piscataway, v. 30, n. 9, p. 38-49, 1992a.

CARPENTER, G. A.; GROSSBERG, S. ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. **Neural Networks**, Kidlington, v. 3, p. 129-152, 1990.

CARPENTER, G. A.; GROSSBERG, S. Adaptive resonance theory. In: _____. **The handbook of brain theory and neural networks**. Cambridge: MIT Press, 2003. p. 87-90.

CARPENTER, G. A.; GROSSBERG, S.; IIZUKA, K. Comparative performance measures of fuzzy ARTMAP, learned vector quantization, and back propagation for handwritten character recognition. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS – IJCNN, 1992, Baltimore. **Proceedings of the...** Baltimore: IEEE, 1992b. v. 1, p. 794-799.

CARPENTER, G. A.; GROSSBERG, S.; MARKUZON, N.; REYNOLDS, J. H.; ROSEN, D. B. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. **Neural Networks, IEEE Transactions on**, Piscataway, v. 3, p. 698-713, 1992c.

CARPENTER, G. A.; GROSSBERG, S.; REYNOLDS, J. H. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. **Neural Networks**, Kidlington, v. 4, n. 5, p. 565–588, 1991.

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

CARVALHO, A. P. L. F. **Redes neurais artificiais**. Universidade de São Paulo - USP, Departamento de Ciências da Computação, 2009. (Notas de aula).

CENTRALIZED DATASET - CDS. **Electricity commission for the load dataset**. New Zealand, 2010.

CHAUVENET, W. **A manual of spherical and practical astronomy**. J. B. Lippincott Company, v. 2, 1891. Disponível em: <
http://www.wilbourhall.org/pdfs/sphericalAstronomy/A_Manual_of_Spherical_and_Practical_AstrVOL_I.pdf>. Acesso em: 13 abr. 2018.

CHIN, H. An analytical evaluation of Top-Down versus Bottom-Up forecast in the electricity demand. IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2016, Nantou. **Proceedings of the...** Nantou: IEEE, 2016. p. 1-2.

CLEMENTS, A. E.; HURN, A. S.; LI, Z. Forecasting day-ahead electricity load using a multiple equation time series approach. **European Journal of Operational Research**, Amsterdam, v. 251, n. 2, p. 522-530, 2016.

DANGERFIELD, B. J.; MORRIS, J. S. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. **International Journal of Forecasting**, Amsterdam, v. 8, n. 2, p. 233–241, 1992.

- DAUT, M. A. M.; HASSAN, M. Y.; ABDULLAH, H.; RAHMAN, H. A.; ABDULLAH, M. P.; HUSSIN, F. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. **Renewable and Sustainable Energy Reviews**, v. 70, p. 1108-1118, 2017.
- DUDEK, G. Pattern-based local linear regression models for short-term load forecasting. **Electric Power Systems Research**, Lausanne, v. 130, p. 139-147, 2016.
- EBERHART, R. C.; DOBBINS, R. W. **Neural network PC tools: a practical guide**. San Diego: Academic Press, 1990.
- ELSNER, J.; TSONIS, A. **Singular spectrum analysis: a new tool in time series analysis**. New York: Plenum Press, 1996.
- FAN, S.; CHEN, L.; LEE, W. Short-Term Load Forecasting Using Comprehensive Combination Based on Multimeteorological Information. **IEEE Transactions on Industry Applications**, Piscataway, v. 45, n. 4, p. 1460-1466, 2009.
- FARIA, A. E.; SOUZA, R. C. A re-evaluation of the quasi-Bayes approach to the linear combination of forecasts. **Journal of Forecasting**, Hoboken, v. 14, n. 6, p. 533-542, 1995.
- GERSHENFELD, N. A.; WEIGEND, A. S. **The future of time series: in time series prediction: forecasting the future and understanding the past**. Boston: Addison-Wesley, 1993. p. 1-70. Disponível em: <<https://sfi-edu.s3.amazonaws.com/sfi-edu/production/uploads/sfi-com/dev/uploads/filer/86/5d/865d74c0-3246-4552-89b8-66f0cc9c69d6/93-08-053.pdf>>. Acesso em: 13 abr. 2018.
- GOLYADINA, N.; NEKRUTKIN, V.; ZHIGLIAVSKY, A. A. **Analysis of time series structure: SSA and related techniques**. Boca Raton: Chapman & Hall/CRC, 2001. 2001.
- GOMES, T. T.; LOPES, M. L. M.; SANTOS JUNIOR, C. R.; ALVES, M. F.; LOTUFO, A. D. P. Rede ARTMAP Euclidiana na solução do problema de previsão de cargas elétricas. In: LATIN-AMERICAN CONGRESS ON ELECTRICITY GENERATION AND TRANSMISSION - CLAGTEE, 12, 2017, Mar del Plata. **Proceedings of the...** Mar del Plata: [s.n.], 2017. p. 1-6, 2017. Disponível em: <www3.fi.mdp.edu.ar/clagtee/2017/articles/01-003.pdf>. Acesso em: 13 abr. 2018.
- GROSSBERG, S. A Theory of human memory: self-organization and performance of sensor-motor codes, maps, and plans. In: _____. **Studies of mind and brain**. Boston: Reidel. 1982a. p. 498-639.
- GROSSBERG, S. Behavioral contrast in short-term memory: serial binary memory models or parallel continuous memory models. In: _____. **Studies of mind and brain**. Boston: Reidel. p. 425-447, 1982b.
- HAMILTON, J. D. **Time series analysis**. Princeton: Princeton University Press, 1994. 799 p.

- HAWKINS, D. **Identification of outliers**. London: Chapman & Hall, 1980.
- HAYKIN, S. *Neural Networks and Learning Machines*. 3. ed. Upper Saddle River: Prentice Hall, 2009.
- HIPPERT, H. S.; BUNN, D. W.; SOUZA, R. C. Neural networks for short-term load forecasting: a review and evaluation. **International Journal of Forecasting**, Amsterdam, v. 21, p. 425–434, 2005.
- HONG, T.; FAN, S. Probabilistic electric load forecasting: A tutorial review. **International Journal of Forecasting**, Amsterdam, v. 32, p. 914-938, 2016.
- HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, Innsbruck, v. 27, n. 2, 22 p., 2008. Disponível em: < <https://robjhyndman.com/publications/automatic-forecasting/>>. Acesso em: 13 abr. 2018.
- KADOWAKI, M.; OHISHI, T.; SOARES FILHO, S.; LIMA, W. S. Modelo de previsão de demanda de carga de curtíssimo prazo para o período da ponta. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL - SBPO, 36, 2004, São João Del Rei. **Anais...** São João Del Rei: SBPO, 2004. p. 2160-2171, 2004. Disponível em: < <http://www.din.uem.br/sbpo/sbpo2004/pdf/arq0029.pdf>>. Acesso em: 13 abr. 2018.
- KAGAN, N.; OLIVEIRA C. C. B; ROBBIA, E. J. **Introdução aos sistemas de distribuição de energia elétrica**. São Paulo: Blucher, 2010.
- KHATOON, S.; NASIRUDDIN, I.; SINGH, A. K.; PRITI. Effects of various factors on electric load forecasting: an overview. In: IEEE POWER INDIA INTERNATIONAL CONFERENCE, PIICON, 6, 2014, Delhi. **Proceedings of the...** Delhi: IEEE, 2014. 5 p.
- KHOSRAVI, A.; NAHAVANDI, S.; CREIGHTON, D. Short term load forecasting using interval type-2 fuzzy logic systems. In: IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, FUZZ-IEEE 2011, Taipei. **Proceedings of the...** Taipei: IEEE, 2011. p. 502–508.
- KRITHARAS, P. **Developing a SARIMAX model for monthly wind speed forecasting in the UK**. 2014. 211 f. Thesis (Doctor) - Loughborough University, Loughborough, 2014. Disponível em: <<https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/16350/2/Thesis-2014-Kritharas.pdf>>. Acesso em: 13 abr. 2018.
- LI, H.; CUI, L.; GUO, S. A hybrid short-term power load forecasting model based on the singular spectrum analysis and autoregressive models. **Advances in Electrical Engineering**, London, v. 2014, p. 1–7, 2014. Disponível em: <<https://www.hindawi.com/journals/aee/2014/424781/>>. Acesso em: 13 abr. 2018.
- LISI, F.; NICOLIS, O.; SANDRI, M. Combining singular-spectrum analysis and neural networks for time series forecasting. **Neural Processing Letters**, v. 2, n. 4, p. 6–10, 1995.

- LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica**. 2005. 149 f. Tese (Doutorado) – Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2005.
- LUTHULI, Q. W.; FOLLY, K. A. Short term load forecasting using artificial intelligence. In: IEEE PES POWER AFRICA CONFERENCE, 2016, Livingstone. **Proceedings of the...** Livingstone: IEEE, 2016. p. 129-133.
- MAMDANI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. **International Journal of Man-Machine Studies**, London, v. 7, n. 1, p. 1-13, 1975.
- MARQUES, C. A. F.; FERREIRA, J. A.; ROCHA, A.; CASTANHEIRA, J. M.; MELO-GONÇALVES, P.; VAZ, N.; DIAS, J. M. Singular spectrum analysis and forecasting of hydrological time series. **Physics and Chemistry of the Earth**, Oxford, v. 31, p. 1172-1179, 2006.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, p. 115-133, 1943.
- MINSKY, M.; PAPERT, S. **Perceptrons: an introduction to computational geometry**. Cambridge: M.I.T. Press, 1969. 258 p.
- MOCANU, E.; NGUYEN, P. H.; GIBESCU, M.; LARSEN, E. M.; PINSON, P. Demand forecasting at low aggregation levels using factored conditional restricted boltzmann machine. In: PROCEEDINGS OF POWER SYSTEMS COMPUTATION CONFERENCE, 19, 2016, Genoa. **Proceedings of the...** Genoa: IEEE, 2016. 7 p.
- MORETTIN, P. A.; TOLÓI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Edgard Blücher, 2006. 538 p. (Coleção Projeto Fisher).
- NIU, M.; SUN, S.; WU, J.; WANG, J. An innovative integrated model using the singular spectrum analysis and nonlinear multi-layer perceptron network optimized by hybrid intelligent algorithm for short-term load forecasting. **Applied Mathematical Modelling**, Guildford, v. 40, n. 5/6, p. 4079–4093, 2016.
- NOSE-FILHO, K.; LOTUFO, A. D. P.; MINUSSI, C. R. Preprocessing data for short-term load forecasting with a general regression neural network and a moving average filter. In: TRONDHEIM POWER TECH 2011 IEEE, 2011, Trondheim. **Proceedings of the...** Trondheim: IEEE, 2011. p. 1-7.
- OLIVEIRA, C. M.; LOTUFO, A. D. P.; LOPES, M. L. M.; SILVEIRA, M. C. G. Previsão de cargas elétricas através de uma rede neural híbrida back-ART fuzzy. In: BRAZILIAN CONGRESS ON COMPUTATIONAL INTELLIGENCE, CBIC, 11, Porto de Galinhas. **Anais...** Porto de Galinhas: [S.n.], 2013. Disponível em: <http://abricom.org.br/wp-content/uploads/2016/03/bricsccicbic2013_submission_35.pdf>. Acesso em: 13 abr. 2018.

PANDIAN, S. C.; DURAISWAMY, K.; RAJAN, C. C. A.; KANAGARAJ, N. Fuzzy approach for short term load forecasting. **Electric Power Systems Research**, Lausanne, v. 76, n. 6, p. 541–548, 2006.

PYLE, D. **Data preparation for data mining**. San Francisco: Morgan Kaufmann Publishers, 1999.

RAHMAN, S.; HAZIM, O. A generalized knowledge-based short-term Load-forecasting technique. **IEEE Transaction Power Systems**, Piscataway, v. 8, n. 2, p. 508-514, 1993.

RAZA, M. Q.; KHOSRAVI, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. **Renewable and Sustainable Energy Reviews**, v. 50, p. 1352–1372, 2015.

RODRIGUES, J. **Modelo mcculloch e pitts**. Espírito Santo: Universidade Federal do Espírito Santo - UFES, Centro de Ciências Agrárias – CCA, Departamento de Computação, Alegre, Espírito Santo. Notas de aula.

ROUGIER, N. Schemat budowy neuronu. **Wikimedia Commons**, 2007. Disponível em: <<https://upload.wikimedia.org/wikipedia/commons/1/19/Neuron-figure-fr.svg>>. Acesso em: 21 dez. 2017.

SADAEI, H. J.; GUIMARAES, F. G.; DA SILVA, C. J.; LEE, M. H.; ESLAMI, T. Short-term load forecasting method based on fuzzy time series, seasonality and long memory process. **International Journal of Approximate Reasoning**, New York, v. 83, p. 196-217, 2017.

SCHELLONG, W.; HENTGES, F. Forecast of the heat demand of a district heating system. In: PROCEEDINGS OF THE CONFERENCE ON POWER AND ENERGY SYSTEMS, Spain. **Proceedings of the...** Spain: Seventh IASTED, v. 7, 2007, p. 383-388.

SEVLIAN, R.; RAJAGOPAL, R. A scaling law for short term load forecasting on varying levels of aggregation. **International Journal of Electrical Power & Energy Systems**, Oxford, v. 98, p. 350-361, 2018.

TAYLOR, J. W. Triple seasonal methods for short-term electricity demand forecasting. **European Journal of Operational Research**, Amsterdam, v. 204, p. 139–152, 2010.

TAYLOR, J. W.; ARORA, S. Rule-Based autoregressive moving average models for forecasting load on special days: a case study for France. **European Journal of Operational Research**, Amsterdam, v. 266, p. 259-268, 2018.

VILLAS BÔAS, B. Horário de verão traz poucos benefícios para setor elétrico, diz Aneel. **Valor Econômico**, São Paulo, 2017. Disponível em: <<http://www.valor.com.br/brasil/5130458/horario-de-verao-traz-poucos-beneficios-para-setor-eletrico-diz-aneel>>. Acesso em: 21 dez. 2017.

WARTH, A. Estudo do governo mostra que horário de verão não reduz consumo de energia. **O Estado de São Paulo**, São Paulo, 2017. Disponível em: <<http://economia.estadao.com.br/noticias/70001872855>>. Acesso em: 21 dez. 2017.

WEENINK, D. **Category ART**: a variation on adaptive resonance theory neural networks. Institute of Phonetic Sciences – University of Amsterdam, IFA Proceedings, n. 21, p. 117-129, 1997. Disponível em: <<https://pdfs.semanticscholar.org/e3bf/a47a3a6e589da8fdb544749acea3c4f10029.pdf>>. Acesso em: 05 jan. 2018.

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

WILLIAMSON, J. Gaussian ARTMAP: a neural network for fast incremental learning of noisy multidimensional maps. **Neural Networks**, Kidlington, v. 9, n. 5, p. 881-897, 1996.

WITTEN, I. H.; FRANK, E. **Data mining**: practical machine learning tools and techniques. San Francisco: Morgan Kaufmann, 2005. 525 p.

WOLF, A. S. **Análise automática de sinais eletrocardiográficos por redes neurais artificiais**. 2004. 124 f. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2004.

WU, C. L.; CHAU, K. W. Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. **Journal of Hydrology**, Amsterdam, v. 399, n. 3/4, p. 394-409, 2011.

YILDIZ, B.; BILBAO, J. I.; SPROUL, A. B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. **Renewable and Sustainable Energy Reviews**, v. 73, p. 1104-1122, 2017.

YU, C-N.; MIROWSKI, P.; HO, T. K. A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids. **IEEE Transactions on Smart Grid**, Piscataway, v. 8, n. 2, p. 738-748, 2017.

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

ZHANG, X.; WANG, J.; ZHANG, K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. **Electric Power Systems Research**, Lausanne, v. 146. p. 270-285, 2017.

ZHENG, J.; XU, C.; ZHANG, Z.; LI, X. Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: 2017 51st ANNUAL CONFERENCE ON INFORMATION SCIENCES AND SYSTEMS - CISS, 51, 2017, Baltimore. **Proceedings of the...** Baltimore: IEEE, 2017. p. 1-6.

APÊNDICE A

A.1 MODELOS BOX & JENKINS

Nos testes de previsão realizados neste trabalho, que consideram cargas elétricas em diferentes níveis de agregação, um dos previsores empregados é um modelo SARIMA, de Box & Jenkins. Portanto o método é brevemente introduzido neste capítulo.

Os modelos de Box e Jenkins (1970) buscam identificar um sistema probabilístico plausível gerador de uma série temporal que apresenta média e covariância constantes (i.e., estacionariedade de segunda ordem) e autocorrelação, considerando apenas informações contidas na própria série.

De acordo com Hamilton (1994), um processo é fracamente estacionário se a média e as autocovariâncias para quaisquer defasagens não dependam do tempo e, dito estacionário se, para uma série temporal $Y_t (t = 1, 2, 3, \dots, T)$ as funções de distribuição conjunta entre Y_t e Y_{t+j} , $\forall j \in \mathbb{Z}$, dependam somente do inteiro j .

Um processo estacionário homogêneo não se enquadra diretamente nas definições de estacionariedade dadas acima, porém, gera uma série estacionária ao passar por um determinado número de diferenciações, este número é representado pelo parâmetro d do modelo ARIMA.

Um modelo ARIMA (p, d, q) é baseado em modelos de filtros lineares estáveis, sendo assim, considera-se que a série é gerada por um filtro linear que recebe como entrada um ruído branco, onde p indica a ordem da parte auto-regressiva do modelo e q representa a ordem da parte médias móveis do modelo, compondo as saídas e entradas do filtro respectivamente.

Para Morettin e Toloí (2006), seja $W_t = \nabla^d Y_t$, uma série estacionária, é possível representar W_t por um modelo ARMA (p, q) , sendo assim, um modelo ARIMA genérico pode ser escrito como:

$$\phi(B)\nabla^d Y_t = \theta(B)a_t \quad (\text{A.1})$$

onde B representa o operador de atraso; ∇ , o operador diferença; $\phi(B)$ o polinômio auto-regressivo de grau p ; $\theta(B)a_t$ o polinômio de médias móveis de grau q . Outra maneira de representar tal modelo é sob a forma de equações de diferenças:

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \dots + \varphi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (\text{A.2})$$

O ajuste de um modelo ARIMA passa pela identificação das ordens (p, d, q) do modelo, e então, pela estimação dos coeficientes dos polinômios $\phi(B)$ e $\theta(B)$.

Dois princípios são considerados ao se utilizar modelagem ARIMA, são eles: parcimônia, ou seja, estabelecer um modelo com o menor número de parâmetros possível; e ciclo iterativo, que consiste na estratégia de seleção de modelos (MORETTIN; TOLOI, 2006).

As ordens (p, d, q) iniciais podem ser determinadas a partir da análise do perfil dos gráficos das funções de autocorrelação - FAC e função de autocorrelação parcial - FACP (HAMILTON, 1994).

Após identificar um modelo plausível, o próximo passo é a identificação de seus parâmetros. Neste ponto utilizam-se métodos iterativos não lineares em associação com o método da máxima verossimilhança (MORETTIN; TOLOI, 2006; BOX; JENKINS; REINSEL, 2008).

Uma vez que um modelo tenha sido estimado, sua validação ocorre a partir da aplicação de testes de diagnósticos, como, por exemplo, teste de Box-Pierce, Ljung-Box, Durbin-Watson, critérios de informação de Akaike (AIC) e Bayesiano (BIC), periodograma acumulado, autocorrelação simples e parcial (HAMILTON, 1994).

Modelos SARIMA são utilizados para séries não estacionárias sazonais. Tido como expansões dos modelos ARIMA convencionais, sua expressão leva em consideração além das partes já descritas nos modelos ARIMA, diferenciações, médias móveis e auto-regressões sazonais; o modelo pode ser representado como (BOX; JENKINS; REINSEL, 2008):

$$\phi(B)\Phi(B^s)\nabla_s^D\nabla^d Y_t = \theta(B)\Theta(B^s)a_t \quad (\text{A.3})$$

onde ∇_s^D é o operador de diferença sazonal de ordem D , $\Phi(B^s)$ e $\Theta(B^s)$ são, respectivamente, os polinômios auto-regressivo e de médias móveis da parte sazonal do modelo, com ordens P e Q respectivamente. De maneira resumida, um modelo com estas características pode ser representado por SARIMA $(p, d, q)(P, D, Q)_s$.

Estes modelos apresentam maior flexibilidade, possibilitando a adição de variáveis exógenas explicativas ao modelo, bastando para isso, adicionar um componente da variável exógena ao modelo já descrito. Na literatura, este modelo sazonal é descrito como SARIMAX $(p, d, q)(P, D, Q)_s$ (KRITHARAS, 2014).

Para este trabalho não estavam disponíveis variáveis exógenas climáticas específicas de cada cenário em estudo. Uma vez que as séries de dados empregadas são não estacionárias e apresentam sazonalidade, o modelo SARIMA foi utilizado em determinados cenários de estudo.

A.1.2 IMPLEMENTAÇÃO COMPUTACIONAL DO SARIMA

Neste trabalho os dados reais de consumo de energia elétrica foram submetidos a biblioteca *forecast* do software R para computação estatística, que implementa entre outros um algoritmo automático (*auto.arima*) de previsão *step-wise* com modelos ARIMA. O algoritmo completo é descrito em (HYNDMAN; KHANDAKAR, 2008).

O algoritmo testa a série para verificar se ela é estacionária, realizando diferenciações se necessário, o mesmo vale para a parte sazonal. São ajustados inicialmente 04 modelos, sendo selecionado como modelo corrente aquele com menor valor AIC; o modelo corrente é submetido a 30 variações, sendo novamente selecionado como modelo corrente aquele que apresentar menor valor AIC, o processo se repete até que não seja mais possível encontrar um modelo próximo ao corrente com um valor AIC menor. Por default p e q não ultrapassam 05, do mesmo modo que P e Q estão limitados a 02. Estes e outros parâmetros podem ser customizados, porém neste trabalho é mantida a configuração padrão.

APÊNDICE B

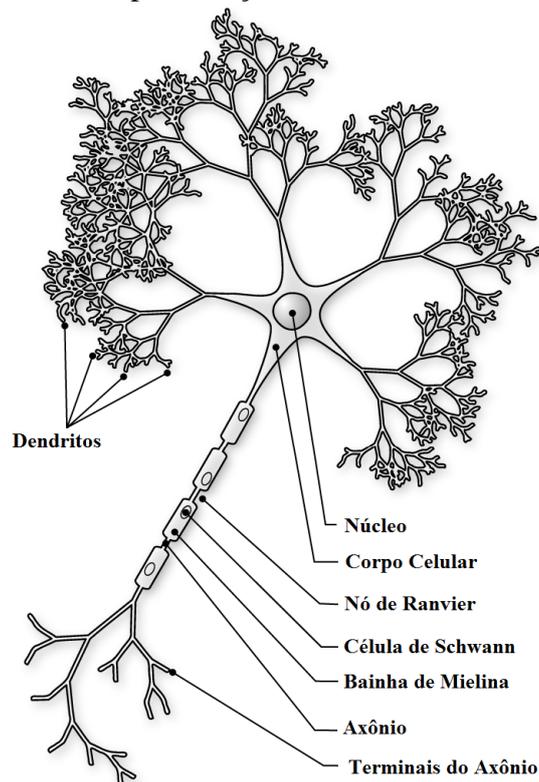
B.1 REDES NEURAIS ARTIFICIAIS

B.1.1 MODELO BIOLÓGICO

O sistema nervoso é formado por um complexo conjunto de células, os neurônios. Essas células são responsáveis pela determinação do modo de funcionamento e comportamento do corpo e da mente.

Os neurônios são formados entre outros pelos dendritos, que podem ser compreendidos como terminais de entrada, pelo corpo celular (também conhecido como soma), e pelo axônio, cujas prolongações podem ser vistas como terminais de saída. A Figura 20 ilustra um neurônio biológico:

Figura 20 – Representação de um neurônio biológico



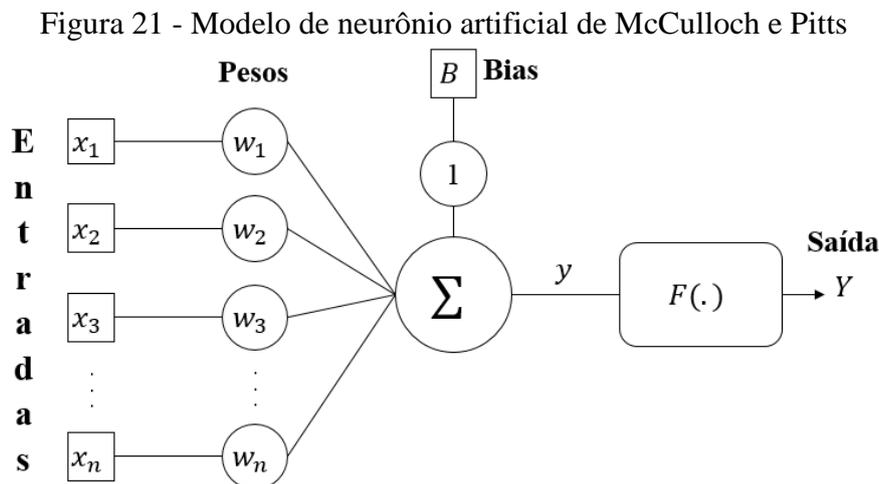
Fonte: Adaptado de (ROUGIER, 2007)

A comunicação entre os neurônios ocorre através das sinapses, que são regiões através das quais impulsos nervosos são transmitidos, configurando a comunicação entre dois neurônios. Os impulsos recebidos por um determinado neurônio num tempo específico são processados e, caso seja atingido um determinado limiar de excitação o neurônio dispara, produzindo uma substância neurotransmissora que flui da soma para o axônio e seus terminais, estando estes conectados aos dendritos de outros neurônios é

possível dizer que uma comunicação foi estabelecida. O neurotransmissor pode aumentar ou diminuir a polaridade da membrana pós-sináptica, de maneira a excitar ou inibir a geração de pulsos no segundo neurônio. Tal processo depende da geometria da sinapse, do tipo de neurotransmissor entre outras características (CARVALHO, 2009).

B.1.2 MODELO DE MCCULLOCH E PITTS

McCulloch e Pitts (1943) propuseram um modelo de neurônio com base e a partir da simplificação daquilo que se conhecia da estrutura e do funcionamento de um neurônio biológico na época. O modelo matemático resultante apresentava n terminais de entrada x_1, x_2, \dots, x_n , um elemento 'bias' de peso unitário, e um terminal de saída Y (que é resultante de uma função linear ou não - $F(y)$), conforme ilustrado pela Figura 21:



Fonte: Adaptado de (HAYKIN, 2009)

Com o objetivo de emular o comportamento das sinapses de um neurônio real, cada terminal de entrada do neurônio possui um peso sináptico associado w_1, w_2, \dots, w_n , podendo assumir valores positivos ou negativos, de tal maneira que as sinapses associadas possam ser dos tipos inibitórias ou excitatórias. O produto de $x_i w_i$ é o efeito da sinapse i no neurônio pós-sináptico.

Os pesos sinápticos determinam em que grau o neurônio artificial deve considerar os sinais de entrada apresentados às conexões (HAYKIN, 2009). Quando a soma dos impulsos recebidos por um neurônio ultrapassa seu limiar de excitação (*threshold*) ele produz uma determinada resposta de saída.

A ativação ou não do neurônio no modelo de McCulloch e Pitts - MCP é realizada por uma função de ativação, cuja entrada é a composição linear dos dados de entrada multiplicados pelos pesos associados somados ao elemento 'bias'. A função de

ativação original descrita no modelo de MCP é dada pela seguinte função limiar (WOLF, 2004):

$$\sum_{i=1}^n x_i w_i + B \geq \theta \quad (\text{B.1})$$

onde n representa o número de entradas do neurônio, w_i é o peso associado a entrada x_i , B é o peso ‘bias’ e θ é o limiar de excitação do neurônio.

O modelo de MCP é bastante simplificado em relação aos sistemas biológicos, exemplo disto é que no modelo apresentado assume-se que os neurônios em cada camada da rede disparam sincronamente e que entradas em um instante de tempo t produzem saídas no tempo $t + 1$ (WOLF, 2004). Entre as características do modelo de MCP, pode-se destacar (RODRIGUES, ____):

- A atividade do neurônio é binária para cada instante, de modo que dispara sempre que a atividade = 1, e não dispara quando a atividade = 0;
- A rede é constituída por linhas sem pesos que ligam os neurônios, estas linhas podem ser excitatórias (positivas) ou inibitórias (negativas);
- O limiar de excitação L dos neurônios é fixo, de modo que ele somente dispara se a entrada total recebida em um determinado instante for maior ou igual a L ;
- De modo a reproduzir o atraso sináptico, um sinal leva uma determinada unidade de tempo para passar de um neurônio para outro.
- O recebimento de uma sinapse inibitória em um determinado instante, independentemente do recebimento de outras excitatórias no mesmo instante, inibe o disparo do neurônio;