



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

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

LEONARDO BRAIN GARCÍA FERNÁNDEZ

**MODELO PREVISOR DIÁRIO PARA CARGAS EM DIFERENTES NÍVEIS DE
DESAGREGAÇÃO BASEADO NAS TÉCNICAS DE *MACHINE LEARNING***

Ilha Solteira
2023

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LEONARDO BRAIN GARCÍA FERNÁNDEZ

**MODELO PREVISOR DIÁRIO PARA CARGAS EM DIFERENTES
NÍVEIS DE DESAGREGAÇÃO BASEADO NAS TÉCNICAS DE
*MACHINE LEARNING***

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

Prof. Dr. Carlos Roberto Minussi
Orientador

FICHA CATALOGRÁFICA
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

G216m García Fernández, Leonardo Brain.
Modelo predictor diário para cargas em diferentes níveis de desagregação baseado nas técnicas de machine learning / Leonardo Brain García Fernández. -- Ilha Solteira: [s.n.], 2023
137 f. : il.

Tese (doutorado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2023

Orientador: Carlos Roberto Minussi

Inclui bibliografia

1. Previsão de cargas elétricas em nível desagregado. 2. Machine learning. 3. Teoria da ressonância adaptativa. 4. Máquina de suporte vetorial. 5. Filtros Wavelet.


Amanda Sertori dos Santos

Bibliotecária - CRB/8-9061
Seção Técnica de Referência, Atendimento ao
Usuário e Documentação
Diretoria Técnica de Biblioteca e Documentação

Impacto potencial desta pesquisa

O modelo previsor flexível proposto pode auxiliar tanto na tomada de decisões estratégicas dos agentes, na operação dos sistemas elétricos pelo coordenador quanto no desenvolvimento de políticas públicas voltadas para o setor energético. Isso pode resultar em benefícios econômicos, como o aumento da eficiência energética, a redução de custos na operação e o estímulo ao crescimento sustentável do Brasil.

Potential impact of this research

The proposed flexible forecasting model can help both in the strategic decision-making of the agents, in the operation of the electrical systems by the coordinator and in the development of public policies aimed at the energy sector. This can result in economic benefits, such as increased energy efficiency, reduced operating costs, and stimulating Brazil's sustainable growth.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Ilha Solteira

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: MODELO PREVISOR DIÁRIO PARA CARGAS EM DIFERENTES NÍVEIS DE DESAGREGAÇÃO BASEADO NAS TÉCNICAS DE *MACHINE LEARNING*

AUTOR: LEONARDO BRAIN GARCÍA FERNÁNDEZ

ORIENTADOR: CARLOS ROBERTO MINUSSI

Aprovado como parte das exigências para obtenção do Título de , área: Automação pela Comissão Examinadora:


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

Profª. Drª. ANNA DIVA PLASENCIA LOTUFO (Participação Presencial)
Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira - UNESP

Profa. Dra. THAYS APARECIDA DE ABREU SANTOS (Participação Presencial)
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira - UNESP

Prof. Dr. JOEL DAVID MELO TRUJILLO (Participação Virtual)
Centro de Engenharia, Modelagem e Ciências Sociais / Universidade Federal do ABC - UFABC

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

Ilha Solteira, 19 de junho de 2023

AGRADECIMENTOS

Gostaria de expressar minha sincera gratidão a todos aqueles que contribuíram para a realização desta tese e tornaram esta jornada possível.

Em primeiro lugar, gostaria de agradecer a Deus pela oportunidade de viver com saúde e felicidade, e por me capacitar a adquirir novos conhecimentos durante todo o processo de desenvolvimento desta tese. Também sou grato por ter sido agraciado com a chance de conhecer a cultura brasileira incrível e diversa.

Gostaria de expressar minha profunda gratidão à minha família, especialmente à minha mãe Ruth Noemí Fernández Flores e ao meu pai Leonardo García Olarte, pelo apoio incondicional que me deram ao longo da minha jornada. Eles sempre acreditaram em mim e me proporcionaram as ferramentas necessárias para alcançar o meu potencial. Além disso, quero agradecer aos meus irmãos Angélica María e Víctor Hugo por me inspirar e motivar a buscar sempre o melhor.

Também gostaria de agradecer ao Professor Doutor Carlos Roberto Minussi por me oferecer a oportunidade e a confiança para desenvolver todo o meu potencial no laboratório SINTEL. Sou grato às professoras Mara Lúcia Martins Lopes e Anna Diva Plasencia Lotufo por suas contribuições valiosas no processo de desenvolvimento desta pesquisa.

Não posso deixar de agradecer aos meus colegas do laboratório de Sistemas Inteligentes, Tânia Gomes, Monara Maeda, Driely Santos, Danieli Vilela e Paula Montenegro, pela amizade, companheirismo e pela troca constante de conhecimentos acadêmicos e humanísticos. Em especial, quero agradecer à Marleide Alves por sua contribuição importante neste trabalho.

Também sou grato aos meus colegas de república Marco Meneses, Gustavo Akita e Raphael Bueno por serem parte da cultura brasileira e por terem me mostrado o que é viver no Brasil. Quero agradecer às senhoras Maria Aparecida e Sandra, responsáveis por manter as repúblicas em que vivi em ordem.

Por fim, gostaria de expressar minha gratidão a todos os funcionários da UNESP pelo esforço incansável em fornecer as melhores condições de ensino para todos os alunos. Sei que sem o apoio de todos vocês, eu não teria chegado tão longe.

“O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001”.

“Deixem que o futuro diga a verdade e avalie cada um de acordo com o seu trabalho e realizações. O presente pertence a eles, mas o futuro, pelo qual eu sempre trabalhei, pertence a mim”.

(Nikola Tesla, 1927)

RESUMO

A previsão de carga elétrica é fundamental para o gerenciamento dos sistemas de energia, especialmente em um contexto de desregulamentação do setor. Gerenciar esses sistemas levando em consideração diferentes perspectivas, como a operação, a manutenção e o planejamento se tornam tarefas desafiadoras devido à evolução tecnológica e às demandas variáveis dos usuários. Esses desafios exigem uma abordagem complexa para garantir o fornecimento adequado de energia. Neste trabalho propõe-se o desenvolvimento de um sistema previsor flexível para cargas elétricas desagregadas no horizonte do dia seguinte via métodos de *Machine Learning* (ML). Por conseguinte, há necessidade da aplicação de um procedimento especializado que produza resultados satisfatórios. Neste sentido são abordadas três metodologias baseadas em dois métodos de ML, os quais se alinham às redes neurais ARTMAP *Fuzzy* modificada (plasticidade, estabilidade e rapidez) e às *Support Vector Machine* (SVM) de regressão. Portanto, a primeira proposta está desenvolvida utilizando o método SVM_ARTMAP *Fuzzy* (*Kernel* Linear, Gaussiano e Otimizado). Para a segunda, utiliza-se um modelo híbrido de treinamento em paralelo baseado no método WPT-SVM_ARTMAP *Fuzzy*. Finalmente a terceira, utiliza-se o método WPT-MMQ_ARTMAP *Fuzzy*. Essas duas últimas propostas são treinadas utilizando dados via filtro *Wavelet*. Para fins de comparação, implementou-se o método ARTMAP *Fuzzy* como o *benchmark* 1, adicionalmente, o método MMQ-ARTMAP *Fuzzy* como o *benchmark* 2. Nas metodologias apresentadas são consideradas o impacto da escolha dos *inputs* nas redes, uma modelagem matemática aplicada à tarefa de identificação de dias atípicos, por último, uma análise de correlação com dados meteorológicos. A aplicação do treinamento em paralelo e a análise dos *inputs* resultaram em uma melhoria significativa na capacidade preditiva da terceira metodologia, ampliando sua capacidade de generalização.

Palavras-chave: previsão de cargas elétricas em nível desagregado; *machine learning*; teoria da ressonância adaptativa; máquina de suporte vetorial; filtros *wavelet*.

ABSTRACT

Electrical load forecasting is fundamental for the management of energy systems, especially in a context of deregulation of the sector. Managing these systems taking into account different perspectives such as operation, maintenance and planning become challenging tasks due to technological evolution and the electrical loads of users. These challenges require a complex approach to ensure adequate energy supply. This work proposes the development of a flexible forecaster system for disaggregated electrical loads on the horizon of the next day via Machine Learning (ML) methods. There is therefore a need for a specialized procedure to produce satisfactory results. In this sense, three methodologies based on two ML methods are approached, which align with the modified Fuzzy ARTMAP neural networks (plasticity, stability and speed) and the regression Support Vector Machine (SVM). Therefore, the first proposal is developed using the Fuzzy SVM_ARTMAP method (Linear, Gaussian and Optimized Kernel). For the second, a hybrid model of parallel training based on the Fuzzy WPT-SVM_ARTMAP method is used. Finally, the third method uses the Fuzzy WPT-MMQ_ARTMAP method. These last two proposals are trained using data via Wavelet filter. For comparison purposes, the Fuzzy ARTMAP method was implemented as benchmark 1, additionally, the Fuzzy MMQ-ARTMAP method as benchmark 2. In the methodologies presented are considered the impact of the choice of inputs in the networks, a mathematical modeling applied to the task of identification of atypical days, finally, a correlation analysis with meteorological data. The application of parallel training and the analysis of inputs resulted in a significant improvement in the predictive capacity of the third methodology, expanding its generalization capacity.

Keywords: electrical load forecasting in disaggregated level; machine learning; adaptive resonance theory; support vector machine; wavelet filters.

LISTA DE FIGURAS

Figura 1 - Esquema do processo de filtragem.	41
Figura 2 - Processo de decomposição.	42
Figura 3 - Processo de reconstrução.	42
Figura 4 - Diferença entre classificação e regressão.	44
Figura 5 - Representação da maldição da dimensionalidade.....	46
Figura 6 - Esquema de validação cruzada de 5 vezes.....	48
Figura 7 - Classificação do SVM.	49
Figura 8 - Hiperplano de separação ótima.....	51
Figura 9 - Hiperplano de separação de margem suave.....	53
Figura 10 – Diagrama esquemático da Regressão SVM usando ϵ - épsilon.....	59
Figura 11 - Arquitetura ARTMAP Fuzzy.....	64
Figura 12 - Fluxograma da RNA ARTMAP Fuzzy Modificada.....	76
Figura 13 – Pseudocódigo do modelo híbrido ARTMAP Fuzzy-MMQ.....	77
Figura 14 - Curvas das cargas em diferentes níveis de desagregação.....	81
Figura 15 - Temperatura Lima, Peru (2017-2019).	82
Figura 16- Umidade Lima, Peru (2017-2019).	82
Figura 17 - Funcionamento do movimento da janela na etapa de previsão das cargas.....	85
Figura 18 - Correlação da Potência de Pico vs. Temperatura média (2017-2019).....	86
Figura 19 - Variações de Energia Sistema Global SEIN.....	87
Figura 20 - Distribuição normal das variações de energia SEIN.	88
Figura 21 - RNA ARTMAP Fuzzy modificada.	89
Figura 22 - Metodologia 1 utilizando SVM_ARTMAP Fuzzy.....	91
Figura 23 - Fluxograma do modelo híbrido SVM_ARTMAP Fuzzy.	93
Figura 24 - Análise Multirresolução utilizando filtro Wavelet db 20 - carga nível SEIN.	95
Figura 25 - Processamento dos dados Wavelet de detalhe utilizando o modelo MMQ_ARTMAP Fuzzy.	96
Figura 26 – Estrutura da metodologia 2 composta pelo modelo WPT-MMQ_ARTMAP Fuzzy.	97
Figura 27 - Processamento dos dados Wavelet de detalhe utilizando o modelo SVM_ARTMAP Fuzzy.	99

Figura 28 - Estrutura da metodologia 3 composta pelo modelo WPT-SVM_ARTMAP Fuzzy.	99
Figura 29 - Validação cruzada para o SEIN.	102
Figura 30 - Validação cruzada para a área norte.	102
Figura 31 - Validação cruzada para a subárea sudoeste.	103
Figura 32 - Validação cruzada para grandes usuários livres.	103
Figura 33 - Validação cruzada para grandes usuários livres área centro.....	104
Figura 34 - Validação cruzada para grandes usuários livres área sul.....	104
Figura 35 - Curvas da previsão da metodologia 1 versus benchmark 1 e 2 (Global SEIN)...	108
Figura 36 - Curvas da previsão das metodologias 2 e 3 versus benchmark 1(Global SEIN).	109
Figura 37 - Curvas da previsão da metodologia 1 versus benchmark 1 e 2 (Área norte SEIN).	111
Figura 38 - Curvas da previsão das metodologias 2 e 3 versus benchmark 1 (Área norte SEIN).	112
Figura 39 - Curvas de carga das metodologias 1, 2 e 3 (Previsão subárea sudoeste).....	114
Figura 40 - Curvas de cargas das metodologias 4 e 5 (Previsão subárea sudoeste).	115
Figura 41 - Curvas de carga das metodologias 1, 2 e 3 (Previsão dos grandes usuários livres SEIN).	117
Figura 42 - Curvas de carga das metodologias 4 e 5 (Previsão dos grandes usuários livres SEIN).	118
Figura 43 - Curvas de carga das metodologias 1, 2 e 3 (Previsão dos grandes usuários livres área centro).	121
Figura 44 - Curvas de carga das metodologias 4 e 5 (Previsão dos grandes usuários livres área centro).	122
Figura 45 - Curvas de carga das metodologias 1, 2 e 3 (Previsão grandes usuários livres área sul).	124
Figura 46 - Curvas de carga das metodologias 4 e 5 (Previsão grandes usuários livres área sul).	125

LISTA DE TABELAS

Tabela 1 - Resultados das metodologias (Previsão Global SEIN).	107
Tabela 2 - Resultados das metodologias (Previsão da área norte SEIN).....	110
Tabela 3 - Resultados das metodologias (Previsão subárea sudoeste).	113
Tabela 4 - Resultados das metodologias (Previsão dos grandes usuários livres SEIN).	116
Tabela 5 - Resultados das metodologias (Previsão dos grandes usuários livres área centro).	120
Tabela 6 - Resultados das metodologias (Previsão dos grandes usuários livres área sul).....	123

LISTA DE QUADROS

Quadro 1 - Metodologias e características	79
Quadro 2 - Informações dos métodos utilizados para as previsões das cargas elétricas	100
Quadro 3 - Configuração dos parâmetros internos da RNA ARTMAP Fuzzy Modificada...	101

LISTA DE ABREVIATURAS

AR	<i>Autoregressive</i>
ARIMA	<i>Autoregressive Integrated Moving Average</i>
ARMA	<i>Autoregressive Moving Average</i>
ART	Teoria da Ressonância Adaptativa (<i>Adaptive Resonance Theory</i>)
CLC	<i>Clustering Method</i>
CLSAF	<i>Convolutional LSTM neural network</i> com características autorregressivas selecionadas
CNN	<i>Convolution Neural Network</i>
COES	Operador do Sistema Interligado Nacional Peruano
EB	<i>Ensemble Bagging</i>
EHM	<i>Ensemble Hidden Markow</i>
EMD	<i>Empirical Model Decomposition</i>
GBR	<i>Gradient Boosting Regressor</i>
GMRAE	<i>Geometric Median Relative Absolute Error</i>
GRU	<i>Gated Recurrent Units</i>
IA	Inteligência Artificial
IEOD	Relatórios de Avaliação da Operação Diária
LSTM	<i>Long Short-Term Memory</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MdRAE	<i>Median Relative Absolute Error</i>
ML	<i>Machine Learning</i>
MLBD_LSTM	<i>Multilayer Bidirectional Long Short-Term Memory</i>
MMQ	Métodos dos Mínimos Quadrados
MSE	<i>Mean Square Error</i>
PCC	<i>Pearson Correlation Coefficient</i>
PLSTM	<i>Parallel Long Short-Term Memory</i>
PLSTM	<i>Parallel Long Short-Term Memory</i>
RMSE	<i>Root Mean Square Error</i>
RNA	Redes Neurais Artificiais
SEIN	Sistema Interligado Nacional
SEP	Sistemas Elétricos de Potência
sMAPE	<i>Symmetric Mean Absolute Percentage Error</i>
SSA	<i>Singular Spectrum Analysis</i>
SVR	<i>Support Vector Regression</i>
TWC	Transformada <i>Wavelet</i> Contínua
TWD	Transformada <i>Wavelet</i> Discreta

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	MOTIVAÇÕES	17
1.2	OBJETIVOS	19
1.2.1	Objetivo geral.....	19
1.2.2	Objetivos específicos.....	19
1.3	REVISÃO DA LITERATURA	19
1.4	ORGANIZAÇÃO DO TEXTO	23
2	PREVISÃO DE CARGAS ELÉTRICAS.....	26
2.1	CLASSIFICAÇÃO DA PREVISÃO	26
2.2	FATORES QUE INFLUENCIAM A PREVISÃO DE CARGA.....	27
2.3	MÉTODOS DE PREVISÃO DE CARGAS ELÉTRICAS	29
2.3.1	Método de média.....	29
2.3.2	Método baseado em modelo matemático.....	30
2.3.3	Método de inteligência artificial.....	32
2.4	MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO DA PREVISÃO.....	33
2.4.1	Medidas dependentes de escala.....	33
2.4.2	Medidas baseadas em erros percentuais.....	34
2.4.3	Medidas baseadas em erros relativos.....	35
3	TRANSFORMADA WAVELET.....	37
3.1	TRANSFORMADA WAVELET CONTÍNUA.....	38
3.2	TRANSFORMADA WAVELET DISCRETA	38
3.3	BANCO DE FILTRO UTILIZANDO A TRANSFORMADA WAVELET.....	40
4	MACHINE LEARNING.....	43
4.1	TIPOS DE ALGORITMOS DE APRENDIZAGEM.....	43
4.1.1	Aprendizagem supervisionada.....	43
4.1.2	Aprendizagem não supervisionada.....	45
4.2	CONCEITOS BÁSICOS DE MACHINE LEARNING.....	45
4.2.1	A maldição da dimensionalidade.....	45
4.2.2	Overfitting	46
4.2.3	Seleção do modelo.....	47

4.2.4	Validação cruzada.....	47
5	MÉTODOS DE <i>MACHINE LEARNING</i> UTILIZADOS.....	49
5.1	MÁQUINA DE VETORES DE SUPORTE.....	49
5.1.1	Hiperplano de separação ótimo	50
5.1.2	Hiperplano de margem suave	53
5.1.3	O método de convolução do produto escalar no espaço característico	56
5.1.4	Regressão utilizando a máquina de vetores de suporte.....	59
5.2	TEORIA DA RESSONÂNCIA ADAPTATIVA	62
5.2.1	Rede neural artificial ARTMAP	62
5.2.2	Rede neural artificial ARTMAP <i>Fuzzy</i>	63
5.2.3	Parâmetros da RNA ARTMAP <i>Fuzzy</i>	69
5.2.4	Algoritmo de treinamento da RNA ARTMAP <i>Fuzzy</i> – Modificada	69
5.2.5	Modelo híbrido da RNA ARTMAP <i>Fuzzy</i> – MMQ.....	76
6	METODOLOGIAS DO MODELO PREVISOR FLEXÍVEL	79
6.1	METODOLOGIA 1	80
6.1.1	Constituição e preparação dos dados.....	80
6.1.2	Características dos dados abordados no estudo.....	85
6.1.3	Metodologia SVM_ARTMAP <i>Fuzzy</i>	89
6.2	METODOLOGIA 2	94
6.2.1	Decomposição da base de dados	94
6.2.2	Processo da metodologia 2 - híbrida	95
6.3	METODOLOGIA 3	97
6.3.1	Processo da metodologia 3 - híbrida	98
7	RESULTADOS	100
7.1	APLICAÇÃO DA VALIDAÇÃO CRUZADA.....	101
7.2	RESULTADOS OBTIDOS DOS DIVERSOS NÍVEIS DE DESAGREGAÇÃO .	105
7.2.1	Previsão de carga global SEIN	106
7.2.2	Previsão de carga da área norte	109
7.2.3	Previsão de carga da subárea sudoeste	112
7.2.4	Previsão de carga dos grandes usuários livres	115
7.2.5	Previsão de carga dos grandes usuários livres da área centro	118
7.2.6	Previsão de carga dos grandes usuários livres da área sul	122

7.3	DISCUSSÕES	125
8	CONCLUSÕES E SUGESTÕES PARA FUTUROS TRABALHOS	128
8.1	CONCLUSÕES	128
8.2	SUGESTÕES PARA TRABALHOS FUTUROS	130

1 INTRODUÇÃO

1.1 MOTIVAÇÕES

A previsão das cargas elétricas no contexto da desregulamentação do setor energético mundial é uma atividade imprescindível inserida ao gerenciamento dos Sistemas Elétricos de Potência (SEP) (KYRIAKIDES; POLYCARPOU, 2007; WOOD *et al.*, 2013). Gerenciar os sistemas desde diversas perspectivas, tanto na operação, na manutenção quanto no planejamento, consequência da evolução tecnológica no SEP, torna-se uma tarefa complexa e desafiadora pelos requerimentos variáveis de cargas dos usuários decorrentes das inserções de fontes renováveis e autossuprimentos (FEINBERG; GENETHLIOU, 2005; KROPOSKI *et al.*, 2017; OGIMOTO; WANI, 2020).

Partindo desde essa perspectiva da operação, o gerenciamento do sistema elétrico no curto prazo é resultado de um despacho econômico a mínimo custo preservando a segurança e garantindo o melhor aproveitamento dos recursos energéticos (WOOD *et al.*, 2013). Como consequência, torna-se imprescindível realizar um correto acompanhamento do equilíbrio entre oferta e demanda nos diferentes barramentos, as quais serão inseridas a um modelo de otimização. Prever essa demanda em curto prazo, dados relevantes para a programação do despacho do dia seguinte, nos diversos barramentos é um processo que evolui desafios significativos, resultado da volatilidade das cargas decorrentes dos autossuprimentos e da presença de fontes renováveis (KROPOSKI *et al.*, 2017; SEVLIAN; RAJAGOPAL, 2018).

Portanto, a programação do mercado do dia seguinte precisa dispor de perfis de cargas que refletem fielmente a área em análise em relação aos diferentes níveis de desagregação para, dessa forma, a otimização da programação possa garantir a segurança no sistema e o melhor aproveitamento dos recursos naturais, evitando a operação de usinas inflexíveis, as quais poderiam inserir altos custos operativos (WOOD *et al.*, 2013).

A complexidade dessas previsões é quantificada pelo nível de desagregação das cargas elétricas. Um sistema sem desagregação de cargas, consegue alcançar um alto rendimento. Entretanto, para diferentes níveis de desagregação, dada as condições atuais que envolvem os sistemas elétricos, os métodos tornam-se complexos na modelagem dos padrões característicos das cargas (SEVLIAN; RAJAGOPAL, 2018). Por esta razão, os padrões de consumo que apresentam incertezas servem como uma fonte de motivação para a exploração de novas ferramentas na previsão visando procurar métodos mais eficientes e robustos.

Uma das ferramentas aplicadas na previsão de cargas, utilizadas na literatura especializada é *Machine Learning* (ML). Esse pertence ao campo da Inteligência Artificial (IA) que abrange um conjunto de métodos, os quais têm a capacidade de detectar e generalizar, de forma automática, padrões nos dados e, conseqüentemente, ser um suporte para prever o futuro ou tomar decisões sob diversas condições de incerteza (MURPHY, 2012). O campo da IA vem sendo utilizado no mundo, além das previsões de carga, para resolver problemas em diversas áreas do conhecimento (LAKEMEYER; NEBEL, 2002). Nas últimas duas décadas, diversos modelos têm sido aplicados aos sistemas energéticos para resolver problemas de previsão de energia, irradiação solar (NOURANI *et al.*, 2022), velocidade do vento, aquecimento urbano, classificação de distúrbios da qualidade de energia, previsão de preços do mercado de eletricidade, dentre outros (MOSAVI *et al.*, 2019).

A motivação desta tese reside, portanto, na necessidade de desenvolver um modelo previsor flexível e eficiente que permita lidar com a complexidade e volatilidade das cargas elétricas em um contexto de transformação estrutural dos SEP, *e.g.*, inserção de fontes renováveis intermitentes no sistema e autossuprimentos. Os métodos tradicionais de previsão muitas vezes falham ou tendem a ter limitações em capturar nuances e padrões não lineares presentes nos dados elétricos, resultando em previsões nada satisfatórias, imprecisas e ineficientes. Por tanto, é imperativo explorar abordagens avançadas baseadas em ML para aprimorar a precisão e a confiabilidade das previsões de cargas elétricas.

Diante dessas motivações, esta tese visa desenvolver um sistema previsor flexível para cargas elétricas desagregadas em diversos níveis, utilizando métodos de ML. Através dessa abordagem inovadora, busca-se melhorar a acurácia das previsões de carga no intuito de otimizar a operação do sistema e impulsionar o desenvolvimento de políticas energéticas sustentáveis.

Nesta pesquisa de doutorado, considera-se como o conceito de desagregação a fato de separar, fragmentar ou dividir cargas elétricas totais do Sistema Interligado Nacional (SEIN) "global" em partes menores para uma análise de demanda por áreas e subáreas "local". Além disso, visando proporcionar maior inteligibilidade ao texto, variáveis e parâmetros são representados no estilo *itálico*. Se forem vetores, ou matrizes, adiciona-se o estilo **negrito**.

1.2 OBJETIVOS

1.2.1 Objetivo geral

- Propor um modelo previsor flexível para a previsão das cargas elétricas do dia seguinte utilizando técnicas de *Machine Learning*, aplicado a diferentes níveis de desagregação, visando obter resultados precisos e eficientes.

1.2.2 Objetivos específicos

- Identificar e analisar as características meteorológicas e geográficas que influenciam na demanda elétrica e utilizá-las como entrada para o modelo previsor;
- Desenvolver e implementar uma metodologia de pré-processamento baseada em modelos matemáticos que permitam a identificação e rotulagem de valores atípicos na base de dados, incluindo dias festivos e anomalias na operação do sistema elétrico;
- Implementar uma sequência de avaliação de métodos de *Machine Learning* e aplicá-los na base de dados composta por diferentes níveis de desagregação para a previsão do dia seguinte;

Comparar e analisar os resultados das previsões dos diferentes modelos de *Machine Learning* avaliados, identificando aquele que oferece a melhor precisão e flexibilidade para os distintos níveis de desagregação das cargas elétricas.

1.3 REVISÃO DA LITERATURA

A necessidade de informações precisas sobre a distribuição de cargas nos diversos níveis de desagregação é crucial para garantir o fornecimento de energia de alta qualidade e atender às demandas do setor elétrico. Nesse sentido, há estudos na literatura especializada que abordam em maior proporção temas relacionados a esses níveis empregando técnicas de ML.

No intuito de fornecer uma compreensão das características de cargas elétricas decorrentes de um nível desagregado, métodos de ML estão sendo potenciados combinando-os com diversas técnicas (modelos híbridos), assim, visando compreender e identificar certos padrões aleatórios nos dados. Portanto, nesta seção serão apresentados trabalhos correlatos à previsão das cargas elétricas utilizando os métodos híbridos de ML aplicados à diversos níveis de cargas.

Sevlian e Rajagopal (2018) propõem uma lei de escalonamento empírico simples e descrevem a qualidade da previsão da carga em vários níveis de agregação. O objetivo deste

trabalho é o de desenvolver uma lei para medir o desempenho em relação aos diferentes níveis de agregação, partindo de dados desagregados, que se ajustem aos dados experimentais. Demonstra-se, experimentalmente, como parte de uma suavização intuitiva, o estabelecimento de um certo limite no qual a agregação não ajudará mais a melhorar o desempenho da previsão. Além disso, também foi proposto um modelo de processo estocástico simples que descreve a curva de erro agregado experimental.

Em Eseye *et al.* (2019) foi proposta uma abordagem de seleção do espaço de recurso, especificamente dos agentes externos que influenciam o comportamento das cargas elétricas a curto prazo. Segundo esta proposta, a seleção é realizada a partir da implementação de um algoritmo genético binário que se utiliza de uma proposta de organização em termos de função *fitness* de regressão gaussiana para escolher a melhor combinação deste recurso. No processo de avaliação de desempenho os autores propuseram utilizar o modelo das redes neurais *feedforward* aplicado a diferentes tipos de consumidores de cargas elétricas na previsão do dia seguinte.

Em Peng *et al.* (2019) salienta-se a importância da previsão das cargas elétricas na operação confiável e econômica dos sistemas elétricos. Estes mesmos autores propõem ainda que se encaminhem novas pesquisas sob uma abordagem de previsão de carga com alto nível de agregação e sugerem, a priori, seu direcionamento com a proposição de se levar em consideração os baixos níveis de agregação e discutem também as características das pequenas e médias empresas e cargas residenciais em diferentes níveis de agregação, quantificando a previsibilidade por meio da entropia aproximada. Ao final são comparadas as técnicas de regressão linear convencional até o aprendizado profundo de última geração, por meio de uma análise detalhada.

Em Kong *et al.* (2019) constata-se que os sistemas energéticos estão em transição para um sistema mais flexível, inteligente e com maior penetração das energias renováveis. E para uma prospecção futura deste cenário estaria tratada a previsão de carga aos clientes individuais quando da operação e planejamento destas redes em consequência do alto grau de complexidade ao qual estão inseridas. Os autores, neste artigo, abordaram o problema da previsão elétrica a curto prazo para cargas residenciais individuais, levando-se em consideração a diferença entre as associações de padrões das cargas – tanto em nível granular, quanto em relação às subestações. No intuito de resolver a localização de padrões de alta volatilidade, propuseram a utilização da estrutura de uma rede neural recorrente *Long Short-Term Memory* (LSTM) que

fosse capaz de gerenciar de forma eficiente inconsistências, melhorando desta forma, a previsão.

Na referência de Yan *et al.* (2019) é proposta uma metodologia híbrida baseada na transformada estacionária *Wavelet* conjuntamente com a rede LSTM para resolver o problema da previsão de carga elétrica a curto prazo das residências individuais. Os dados foram transformados via *Wavelet* em subsinais, para logo ser avaliados individualmente com a rede LSTM e finalmente serem integrados utilizando da transformada inversa. A proposta superou em termos de precisão os métodos do *Support Vector Regression* (SVR), da rede LSTM e do modelo híbrido *Convolution Neural Network-LSTM* (CNN-LSTM).

Fallah *et al.* (2019) destacam o verdadeiro desafio da previsão de carga elétrica como parte da programação dos sistemas de potência para diferentes níveis no setor de energia. Em relação a isto, neste artigo, discute-se quatro metodologias de última geração para a previsão de carga a curto prazo, ou seja – padrão semelhante, seleção de variáveis, previsão hierárquica e seleção de estação meteorológica. Cada um destes métodos propõem uma solução específica para previsão de carga. Eventualmente, ao destacar-se as principais vantagens e desvantagens de cada abordagem, os autores concluíram que o modelo de carga pode se beneficiar da robustez dos métodos individuais em um esquema híbrido. Finalmente, o esboço geral de uma estratégia híbrida é proposto para a avaliação futura.

Na referência de Müller *et al.* (2020) é abordada a previsão de carga elétrica em níveis desagregados, focados em resolver a problemática da qualidade dessas previsões, consequência do alto grau estocástico dos dados da série temporal. Com o intuito de mitigar o problema anteriormente descrito e obter resultados de boa performance na previsão propôs-se uma metodologia para a remoção do ruído por meio de *Singular Spectrum Analysis* (SSA), conjuntamente com a rede neural artificial ARTMAP *Fuzzy*, os quais geraram excelentes resultados conseguindo uma redução de quase 50% em termos do *Mean Absolute Percentage Error* (MAPE).

Na referência Wang *et al.* (2020) se estabelece uma metodologia para a previsão de carga elétrica a curto prazo em nível industrial utilizando-se da *Ensemble Hidden Markow* (EHM), a qual tem por característica aprender as dinâmicas envolvidas nos padrões de consumo dos usuários industriais em séries de tempo multivariadas correlacionadas. As características temporais dinâmicas de múltiplas sequencias foram capturadas por meio da estratégia de log-verossimilhança e para estabelecer uma melhora na capacidade de generalização e estabilidade

adota-se o algoritmo de aprendizagem nomeado *ensemble Bagging*, melhorando significativamente a precisão na previsão.

Em Sajjad *et al.* (2020) foi proposto um modelo híbrido entre a redes CNN e a *Gated Recurrent Units*, aplicado à previsão de cargas em edifícios residenciais. No trabalho destaca-se a necessidade de encontrar uma metodologia robusta e que consiga uma maior capacidade de generalização no domínio de previsão de cargas. Os resultados desse trabalho foram comparados com modelos como regressão linear, SVM, LSTM Multilayer Bidirecional, modelo híbrido CNN-LSTM, autoencoder, e rede neural profunda, demonstrando um alto desempenho do modelo híbrido.

Aguilar Madrid e Nuno (2021) avaliaram o desempenho de um conjunto de métodos de ML, utilizando como dados de entrada a demanda, condições meteorológicas, dias da semana e identificação de dias feriados, aplicados à previsão semanal de carga elétrica agregada do Panamá. Como resultado, verificou-se que o *Gradient Boosting Regressor* (GBR) consiste na melhor alternativa. Os autores expõem que esta previsão, que propicia uma alta precisão, pode reduzir as incertezas do planejamento introduzidas pela intermitência da produção renovável, resultando em um despacho econômico eficiente e provendo uma redução dos custos na operação no SEP.

Na referência Li *et al.* (2021) se desenvolveu o modelo *Convolutional LSTM* neural network com características autorregressivas selecionadas (CLSAF) para melhorar a previsão de carga residencial de curto prazo. O modelo utilizou as estratégias de seleção de características autorregressivas e exógenas, e um estado "padrão" para evitar *overfitting* em momentos de alta volatilidade de carga. Nesse estudo, demonstrou-se que o modelo CLSAF pode prever com até 25% a mais de precisão em comparação com um modelo de persistência, em diferentes níveis de granularidade espacial. Além do modelo, estabeleceu-se uma técnica de triagem para estimar níveis de confiança das previsões de carga, otimizando recursos computacionais e riscos associados a previsões incertas.

Jin *et al.* (2022) propuseram um modelo híbrido que combina SSA e a rede neural *Parallel Long Short-Term Memory* (PLSTM). Os autores identificaram na literatura que os modelos para previsão da demanda não fornecem resultados confiáveis por consequência de flutuações irregulares e mudanças repentinas nas cargas. A remoção de ruído é umas das principais tarefas em nível desagregado de cargas, por isso os autores propõem extrair os recursos importantes dos dados brutos em subsinais, para posteriormente, de forma paralela,

alimentar uma rede LSTM produzindo resultados de previsão em cada subsinal. Esta proposta teve um ótimo desempenho para a previsão a curto prazo.

Em Yang *et al.* (2022), um algoritmo chamado *decompose-cluster-feedback* foi proposto para melhorar o desempenho na previsão das cargas com estrutura hierárquica. Esse algoritmo é uma versão melhorada do método *Clustering Method* (CLC), que utiliza o *Empirical Model Decomposition* (EMD), mecanismo empregado para decompor a carga em subséries com base na amplitude-frequência. Essa abordagem foi aplicada a três cenários reais de previsão de cargas: 12 regiões dos Estados Unidos, 180 subestações de zona de distribuição australiana e o consumo de 370 clientes em um nível alto. Os resultados demonstraram um superior em comparação com outros métodos, no entanto a limitação do trabalho encontra-se em um alto esforço computacional do EMD.

A partir dos trabalhos correlatados, observa-se que um amplo número de métodos tem sido explorado e utilizado na previsão das cargas elétricas em diversos níveis. No entanto é importante ressaltar que há apenas um número limitado de estudos focados na adaptação de modelos de ML flexíveis aplicados aos problemas das previsões em diferentes níveis. Na tese, o objetivo é apresentar uma abordagem versátil capaz de lidar como diversos níveis de desagregação.

1.4 ORGANIZAÇÃO DO TEXTO

O conteúdo desta pesquisa está dividido em oito capítulos. O Capítulo 1 refere-se à introdução da tese e às motivações que levaram ao desenvolvimento do trabalho, seus objetivos e revisão da literatura correlato ao tema.

Capítulo 2. **Previsão de cargas elétricas** – neste capítulo estão apresentados os horizontes de previsão das cargas elétricas, os fatores que influenciam no comportamento dos perfis das curvas de curto prazo, as metodologias utilizadas na literatura para resolver o problema de previsão e, finalmente, as métricas de avaliação dos resultados.

Capítulo 3. **Transformada Wavelet** – neste capítulo, apresenta-se a descrição dos principais conceitos da teoria da transformada *Wavelet*, transformada contínua e discreta para finalmente introduzir os conceitos da análise multirresolução como principal ferramenta na decomposição e análise de dados.

Capítulo 4. **Machine Learning** – neste capítulo estão dispostos os principais conceitos da área de *Machine Learning*, os tipos de algoritmos de aprendizado existentes e difundido na literatura especializada, adicionalmente, os desafios relacionados à maldição de uma alta

dimensionalidade e o problema do *overfitting*, para finalmente mostrar uma forma de selecionar baseado em *Machine Learning* utilizando a validação cruzada.

Capítulo 5. **Métodos de *Machine Learning*** – neste capítulo, realiza-se a descrição dos principais conceitos relativos à dois métodos de *Machine Learning*, *i.e.*, os da teoria da ressonância adaptativa e os das Máquinas de Vetores de Suporte. Mais especificamente, o primeiro focado no algoritmo da RNA ARTMAP *Fuzzy* modificada e o último baseado em SVM – aplicação do modelo de regressão na procura de um hiperplano ótimo de separação dos dados, os quais apresentam-se como uma das principais ferramentas para o desenvolvimento das previsões de cargas em diferentes níveis de desagregação.

Capítulo 6. **Metodologia do Modelo Previsor Flexível** – neste capítulo são descritas as metodologias propostas para solucionar o problema de previsão de cargas com diferentes níveis de desagregação. Será apresentada a constituição e preparação dos dados a serem inseridos no processo de aprendizagem dos modelos, bem como uma decomposição dos dados em aproximação e detalhes, utilizando filtros *Wavelet*.

A primeira metodologia é baseada em uma análise de categorias multidimensional, utilizando o modelo de regressão SVM e a ARTMAP *Fuzzy* modificada. A segunda metodologia é baseada na aplicação do modelo híbrido WPT-MMQ_ARTMAP *Fuzzy*, que realiza um treinamento paralelo dos dados descompostos via *Wavelet* empregando as redes ARTMAP *Fuzzy* modificada e MMQ_ARTMAP *Fuzzy*. A terceira metodologia consiste no desenvolvimento de outro modelo híbrido, denominada WPT-SVM_ARTMAP *Fuzzy*, que realiza um treinamento paralelo dos dados descompostos via *Wavelet* por meio das ARTMAP *Fuzzy* modificada e SVM_ARTMAP *Fuzzy*.

Por fim, serão estabelecidos dois métodos *benchmarks* para comparar o desempenho de cada metodologia: ARTMAP *Fuzzy* modificada e MMQ_ARTMAP *Fuzzy*, respectivamente. Essas metodologias passarão por uma avaliação utilizando uma metodologia estabelecida e relevante para o nível de desagregação definido, a fim de compará-las em termos de diversidade métrica, com o objetivo de atingir o MAPE.

Capítulo 7. **Resultados** – neste capítulo estão dispostos os resultados das previsões de cargas elétricas derivadas das metodologias propostas no sexto capítulo por meio de tabelas e representações gráficas das curvas de cargas. Em uma próxima etapa, efetua-se a análise comparativa empregando as das métricas *Root Mean Square Error* (RMSE), *Pearson Correlation Coefficient* (PCC), *Mean Square Error* (MSE), *Mean Absolute Error* (MAE) e MAPE com o intuito de determinar o desempenho dos sistemas inteligentes previsores de carga.

Capítulo 8. **Conclusões e sugestões para futuros trabalhos** – neste capítulo são apresentadas as conclusões relacionadas à aplicação e resultados dos sistemas previsores flexíveis propostos, e a importância dos mesmos para os diversos níveis de desagregação no que tange a predição de curto prazo e, finalmente, propõe-se sugestões de trabalhos futuros para a expansão da pesquisa.

2 PREVISÃO DE CARGAS ELÉTRICAS

O setor de energia elétrica é fundamental para a economia global, sendo responsável por garantir o fornecimento de energia elétrica aos consumidores finais. Nesse sentido, uma previsão de alta acurácia das cargas elétricas desempenha um papel fundamental na tomada de decisões dos reguladores, operadores e agentes do setor elétrico. Essas decisões são influenciadas por uma série de fatores, *e.g.*, aspectos econômicos, técnicos, governamentais e ambientais. No entanto, para garantir a estabilidade e a eficiência do sistema elétrico, é essencial contar com previsões precisas. A acurácia das previsões permite que sejam tomadas medidas adequadas para lidar com a demanda de energia, otimizar a operação do sistema, gerenciar os recursos de forma eficiente e atender as necessidades dos consumidores finais de forma confiável. Portanto, a qualidade nas previsões das cargas elétricas é de extrema importância para assegurar um funcionamento adequado e eficaz do setor elétrico (ISLAM *et al.*, 2019).

Cada decisão tomada é crucial para o desenvolvimento adequado do setor elétrico, e cada comportamento dos participantes fornece uma visão das escolhas a serem consideradas. Portanto, um comportamento adequado no curto prazo, pode atender aos interesses individuais dos agentes, garantindo objetivos comuns na gestão dos mercados elétricos, como confiabilidade, segurança e qualidade do suprimento de energia (MUNASINGHE, 2013; ISLAM *et al.*, 2019).

2.1 CLASSIFICAÇÃO DA PREVISÃO

A previsão da carga elétrica é categorizada em três faixas principais: curto, médio e longo prazo (FEINBERG; GENETHLIOU, 2005). Cada faixa atende a diferentes objetivos. Por exemplo, previsões de curto prazo são úteis para determinar a disponibilidade das unidades e para despachos econômicos, enquanto as previsões de médio prazo são necessárias para programar manutenções, aquisição de combustível e intercâmbios. Já as previsões de longo prazo são vitais para planejar a expansão do sistema e analisar a viabilidade financeira (ISLAM, 1997; FEINBERG; GENETHLIOU, 2005). Por conta da importância dessa classificação, são definidos os intervalos de duração para cada categoria.

Curto prazo – esta categoria de curto prazo abrange um intervalo de tempo que varia de algumas horas a alguns dias. A partir deste detalhe, as previsões desempenham uma função importante na tomada de decisões, visando uma correta operação do sistema elétrico porque auxilia o operador na estimativa dos fluxos de carga, geração por restrição *Unit Commitment*, reserva girante, intercâmbios econômicos de energia e gerenciamento de cargas (ISLAM, 1997;

FEINBERG; GENETHLIOU, 2005). As implementações oportunas das decisões acima mencionadas garantem continuidade, segurança e melhora da confiabilidade do sistema elétrico.

Médio prazo – a previsão do médio prazo abrange um horizonte geralmente no período de uma semana a um ano, permitindo que as empresas estimem a demanda em um intervalo mais amplo, auxiliando nas negociações de contratos com outras empresas, na previsão da quantidade de combustível a ser comprada, na programação do plano de manutenção do sistema elétrico e nos intercâmbios de blocos de energia entre áreas (ISLAM, 1997; HAHN, 2009).

Longo prazo – esta categoria abrange períodos que, geralmente excedem um ano, sendo que a sua importância está disposta tanto na tomada de decisões quanto no planejamento da expansão da geração e transmissão, estudos de factibilidade para interconexões, compra de combustíveis e planejamento das tarifas elétricas (ISLAM, 1997; FEINBERG; GENETHLIOU, 2005).

2.2 FATORES QUE INFLUENCIAM A PREVISÃO DE CARGA

O sistema elétrico desempenha um papel fundamental ao fornecer energia para atender às diversas demandas de cargas elétricas em diversos níveis de consumo. As variações na magnitude da demanda dos usuários têm um impacto significativo nas características das curvas de carga. Portanto, é importante identificar o intervalo de consumo, tanto para os usuários regulamentados quanto usuários livres, a fim de compreender o grau de relevância dessas variações nas cargas. Isso permite uma melhor compreensão das necessidades específicas dos consumidores e contribui para uma gestão mais eficiente e eficaz do sistema elétrico.

Usuários regulamentados ou demanda vegetativa – são os Usuários sujeitos à regulamentação dos preços de energia ou pela potência que estes consomem. Segundo a regulamentação da lei das concessões elétricas em PERU (1993) estão classificados nesta categoria, os usuários cuja máxima demanda seja menor ou igual a 200 kW.

Usuários Livres – são aqueles que não estão sujeitos à regulamentação dos preços de energia ou pela potência que eles consomem. Segundo PERU (1993) são considerados, nesta condição, usuários cuja demanda máxima anual esteja superior a 2500 kW.

Os usuários que estão dispostos no intervalo entre 200 kW e 2500 kW, em relação ao consumo de carga, têm a opção de escolher entre pertencer ou não a uma das opções já mencionadas, correlato às abordagens para estratégias comerciais dos usuários. Outro fator, que

a posteriori será explanado detalhadamente diz respeito ao horizonte de análise na demanda de cargas a curto, médio e a longo prazo.

Para o curto prazo a carga é afetada por vários fatores, como – condições sociais, meteorológicas e econômicas, clima, temperatura ambiente e padrões de carregamento diário que cumprem um papel importante na previsão das cargas. Além disso, cabe mencionar também que para um curtíssimo prazo as condições sociais e econômicas não têm impacto significativo nas cargas.

Para o médio prazo baseado em sessões de *brainstorming* e consulta com operadores de sistemas elétricos, a temperatura, umidade relativa, velocidade do vento, radiação, precipitações e número de conexão do consumidor são fatores que influenciam diretamente as cargas elétricas, tais variáveis apresentam uma alta correlação linear com a demanda (ISLAM, 1997).

Na medida em que o horizonte de tempo para a previsão aumenta, no caso do longo prazo, as variáveis que afetam a previsão também variam. Os fatores socioeconômicos tendem a caracterizar papel de fundamental importância, pois geram um grau de correlação aceitável nesta previsão, *e.g.*, crescimento demográfico, graus do índice de conforto e produto interno bruto (ISLAM, 1997).

Depois de mencionar os diversos fatores que apresentam um alto índice de correlação com a previsão de carga em diferentes horizontes de tempo, define-se os mais imprescindíveis no curto prazo.

Fatores meteorológicos – são mudanças da natureza que acontecem por si próprias, processos permanentes de movimentos e transformações pelos quais a natureza passa a afetar de maneira direta a vida humana, *e.g.*, temperatura, a velocidade do vento, a cobertura de nuvens e a umidade (FEINBERG; GENETHLIOU, 2005; WERON, 2007).

Fatores sociais – estes fatores levam em consideração o estilo de vida e o crescimento demográfico de uma população.

Fatores econômicos – estes fatores são representados pelas variáveis causais do crescimento da demanda a médio e longo prazo, sendo o indicador mais importante o produto interno bruto.

Fatores associados ao tempo – são os fatores que incluem a época do ano, o dia da semana, e a hora do dia. Desempenham uma função relevante na previsão das cargas e buscam caracterizar o comportamento cíclico e sazonal (MURTO, 1998; FEINBERG; GENETHLIOU, 2005; WERON, 2007).

Fatores esporádicos – são fatores que acontecem ocasionalmente, *e.g.*, eleições presidenciais, eventos esportivos, eventos culturais, ataques terroristas, dentre outros. Podendo colocar como desafio o gerenciamento da operação dos sistemas elétricos consequência das variações esporádicas nas cargas, isto decorrentes de um aumento de carregamentos ou desligamentos em grandes centros de consumo (MURTO, 1998).

2.3 MÉTODOS DE PREVISÃO DE CARGAS ELÉTRICAS

Visando resolver o problema da previsão das cargas elétricas, a literatura evidencia diversos métodos desenvolvidos e utilizados na solução deste problema, *e.g.*, séries temporais, modelos de regressão, modelos econométricos, autorregressivos, modelos inteligentes (SUGANTHI; SAMUEL, 2012). Estas ferramentas, conforme Islam *et al.* (2019) são elencados nos subitens 2.3.1 a 2.3.3.

2.3.1 Método de média

Uma maneira simples de realizar uma previsão de carga é extrapolando os dados das séries temporais usando os métodos da média. A média móvel simples – a previsão está representada pela média representativa dos valores de um determinado número fixo de dados mais recentes, descartando os mais antigos. Os pontos de dados utilizados pela média móvel se deslocam ao longo do tempo.

$$F_{t+1} = \frac{\sum E_t}{n} = \frac{E_t + E_{t-1} + \dots + E_{t-n}}{n}, \quad (1)$$

Em que:

F_{t+1} : previsão da carga para $t+1$;

E_t : valor real para o período atual t ;

n : número de dados utilizados para o cálculo da média móvel.

Suavização exponencial – similar à média móvel ponderada, no entanto com duas exceções. A primeira exceção decorre da atribuição de pesos a cada intervalo de dados. A outra refere-se, portanto, aos valores dos pesos selecionados para serem decrescidos exponencialmente dos dados mais recentes para os dados mais antigos.

$$F_{n+1} = \alpha E_n + (1 - \alpha)F_n, \quad (2)$$

Em que:

F_{n+1} : previsão para a carga no instante $n+1$;

E_n : valor real da carga para o período atual, n ;

F_n : previsão da carga no instante n ;

α : coeficiente de suavização valorado no intervalo $[0,1]$.

2.3.2 Método baseado em modelo matemático

Método de regressão – considera-se umas das técnicas estatísticas mais usadas para temas de previsão de forma geral, neste trabalho a previsão das cargas elétricas. Segundo Fumo e Biswas (2015) este método pode ser classificado como regressão linear, que inclui a monovariável, a regressão de múltiplas variáveis, e a não linear. A previsão de carga elétrica usando o método de regressão gira em torno das duas seguintes etapas.

- a) Construção de um modelo de regressão procurando a relação geral entre demanda e variáveis que afetam o comportamento desta, *e.g.*, clima, tempo, comportamento do consumidor e número de horas de trabalho.
- b) Análise da regressão em um processo iterativo verificando e modificando repetidamente os coeficientes do modelo até atingir uma boa relação entre variáveis dependentes e independentes.

Método autorregressivo (AR) – método de modelagem expresso como uma combinação linear de valores reais anteriores e com um ruído aleatório (ISLAM *et al.*, 2019). O processo abrange uma combinação linear de dados da série atual contra um ou mais valores anteriores da mesma série.

$$Y(t) = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t, \quad (3)$$

$$\delta = \left(1 - \sum_{i=1}^p \phi\right) \mu, \quad (4)$$

sendo:

δ : constante definida pela equação;

μ : média do processo;

p : ordem de regressão automática;

Y_{t-p} : os valores passados da série.

Por meio da combinação de dois métodos, como *Autoregressive* (AR) e *Moving Average* (MA), foi possível obter um novo modelo *Autoregressive Moving Average* (ARMA) representado na forma:

$$Y(t) = \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \beta_j e_{t-j} \quad (5)$$

sendo:

q : a ordem do q-ésimo termo de erro médio móvel;

b_j : j-ésimo coeficiente de MA;

e_t : ruído branco – parâmetros aleatórios não correlacionados na média.

No intuito de modelar séries temporais foi desenvolvido o modelo *Autoregressive Integrated Moving Average* (ARIMA) de Box-Jenkins (BOX; JENKINS,1976) que considera uma função de três variáveis *ARIMA*(p, d, q), sendo ‘ p ’ a ordem da parte AR e ‘ q ’ a ordem do movimento MA, finalmente ‘ d ’ a ordem diferenciadora segundo:

$$\nabla^d Y(t) = \nabla^{d-1} Y_t - \nabla^{d-1} Y_{t-1} \quad (6)$$

sendo:

∇^d : Operador de diferenças.

2.3.3 Método de inteligência artificial

Além dos métodos referidos acima, os métodos de previsão e cargas elétricas baseado em inteligência artificial têm ganhado popularidade em razão da robustez contra as incertezas. Os métodos concebidos via IA são adequados e especialmente úteis em situações nas quais o modelo matemático envolvido nos dados não é suficientemente claro (ISLAM et al., 2019). Dentro deste contexto, destacam-se:

Redes Neurais Artificiais (RNA) – as redes neurais são modelos computacionais baseados no funcionamento dos neurônios biológicos. Consiste em um grupo de neurônios artificiais em camadas, associados essencialmente a funções não lineares, as quais estão interconectadas por meio de pesos. Portanto, ao se ajustar os pesos por meio do “treinamento” a rede pode ser condicionada a fornecer uma relação entrada-saída que corresponda aos dados do treinamento original (HAYKIN, 2008).

A rede é amplamente utilizada na literatura para resolver problemas de dados não lineares e complexos nos quais não é possível a determinação analítica de uma correlação entre estes (ISLAM et al., 2019). Além disso, é possível encontrar uma grande variedade de arquiteturas de redes neurais aplicadas à previsão e cargas elétricas, *e.g.*, *Feedforward*, *Feedback* e *Self Organizing*, na qual cada uma destas está acompanhada de algoritmos de aprendizado supervisionado ou não-supervisionado para ajustar seus parâmetros internos (KARAYIANNIS; VENETSANOPOULOS, 1993).

Lógica *Fuzzy* – conhecida também como lógica nebulosa, visa modelar, por meio de funções de pertinência, os modos imprecisos de raciocínio que desempenham a capacidade humana a respeito da tomada de decisões racionais em um entorno de incertezas e imprecisão (ZADEH, 1988).

Nestes últimos anos houve grandes contribuições da *Lógica Fuzzy* (ZADEH, 1988) aos projetos aplicados aos sistemas de energia elétrica (KOCAARSLAN; ÇAM, 2005; KAVOUSIFARD; SAMET, 2011; SUGANTHI; INIYAN; SAMUEL, 2015), sem, contudo, deixar de lado as previsões de cargas elétricas (ABREU et al., 2020; DIMITROULIS; ALAMANIOTIS, 2022).

As vantagens de utilizar o método a pouco descrito são muitas; entretanto, as principais são a possibilidade de condensar uma grande quantidade de dados em um conjunto menor por meio de regras variáveis, e por último, a capacidade de incorporar intuição e experiência

humana, consequência de projetos de controladores que baseiam-se em heurística (MAMLOOK; BADRAN; ABDULHADI, 2009).

Máquina de vetores de Suporte – é uma técnica clássica de ML também empregada para analisar dados, procurando um hiperplano separador em um espaço amostral, o qual pode separar amostras em diferentes classes. Este hiperplano tem a capacidade de generalização de dados lineares e não lineares de forma sólida e pode prover resultados mais robustos (CORTES; VAPNIK, 1995; ZHOU, 2021).

Ensemble Learning – este método visa integrar fusão, modelagem e mineração de dados em uma estrutura unificada. Como parte do processo característico, o *Ensemble Learning* extrai um conjunto de recursos com uma variedade de transformações, baseadas em múltiplos algoritmos de aprendizagem no sentido de produzir resultados preditivos com um alto grau de precisão e finalmente, funde todos os conhecimentos concernentes ao primeiro caso para desta forma alcançar a descoberta de conhecimento e, portanto, melhorar o desempenho preditivo por meio de valoração adaptativa (WANG; WANG; SRINIVASAN, 2018; DONG; YU; CAO, 2020).

2.4 MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO DA PREVISÃO

As métricas são ferramentas utilizadas em análises de dados, úteis na medição e comparação de erros nas previsões de demanda. No processo de avaliação da precisão diversas métricas são propostas, *e.g.*, RMSE, MAE, MAPE, entre outras. Em Hyndman e Koehler (2006) realiza-se um levantamento crítico das medidas de precisão baseadas nos erros considerando dados dependentes e não dependentes de escala, as quais estão arroladas abaixo de 2.4.1 a 2.4.3.

2.4.1 Medidas dependentes de escala

Existem algumas métricas cuja escala está em função da escala dos dados. Estes são úteis para comparar métodos aplicados a um mesmo conjunto de dados, concretamente na mesma escala, conforme descrita pelas relações (8) (9) e (10):

$$Err_t = L_t - \underline{L_t} \quad (7)$$

sendo:

Err_t : erro no tempo t ;

L_t : valor real no tempo t ;

\underline{L}_t : valor da previsão no tempo t .

Apesar de RMSE ser a métrica constantemente usada, ao invés do MSE, deve-se indicar que ambas são muito sensíveis a *outliers*, motivo pelo qual se levou a desaconselhar seu uso na avaliação de previsões (HYNDMAN; KOEHLER, 2006):

$$MAE = \frac{1}{n} \sum_{t=1}^n |Err_t| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n |Err_t|^2} \quad (9)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n |Err_t|^2 \quad (10)$$

sendo:

n : número de entradas apresentadas na comparação.

2.4.2 Medidas baseadas em erros percentuais

O erro percentual descrito na relação (11) possui a vantagem de ser independente da escala e é frequentemente utilizado para comparar desempenhos de previsões em diferentes conjuntos de dados.

$$pErr_t = \frac{Err_t}{L_t} \times 100 \quad (11)$$

As métricas mais usadas são baseadas nos erros percentuais, no entanto, apresentam a desvantagem de serem infinitos ou indeterminados se houver valores iguais a zero na base de dados ou uma distribuição assimétrica quando o valor for muito próximo a zero (HYNDMAN; KOEHLER, 2006):

$$MAPE = \frac{1}{n} \sum_{t=1}^n |pErr_t| \quad (12)$$

$$sMAPE = \frac{2}{n} \sum_{t=1}^n \frac{|L_t - \underline{L}_t|}{(L_t + \underline{L}_t)} \quad (13)$$

Adicionalmente, cabe aqui referenciar que diversos autores notaram que as médias baseadas em erros percentuais podem ser melhoradas e tornando-as mais estáveis pela utilização de transformações baseadas em similaridade.

2.4.3 Medidas baseadas em erros relativos

O erro relativo baseia-se na divisão do erro de escala pelo erro obtido usando outro método padrão de previsão, conhecido como erro *benchmark*, o qual está descrito na relação expressa em (14).

$$rErr_t = \frac{Err_t}{Err_t^*} \quad (14)$$

sendo:

Err_t^* : valor do erro na previsão obtido pelo método *benchmark*.

A partir da relação (14), define-se os seguintes erros:

$$MdRAE = \frac{1}{n} \sum_{t=1}^n |rErr_t| \quad (15)$$

$$GMRAE = \sqrt[n]{\prod_{t=1}^n rErr_t} \quad (16)$$

Estas métricas apresentam deficiência quando o erro relativo *benchmark* tem seu valor perto do valor zero.

3 TRANSFORMADA WAVELET

Uma *Wavelet* (MORLET, 1982; GROSSMANN; MORLET, 1984; MEYER, 1985) é uma função, particularmente uma forma de onda de duração limitada que possui um valor médio igual a zero. As *Wavelet* podem ser irregulares ou assimétricas, estas características garantem uma melhor adaptação comparadas com as séries de Fourier. Na literatura existem uma ampla quantidade de famílias *Wavelet*, e.g., *Daubechies*, *Symelet*, *Haar* entre outras (CORTÉS; GARAZÓN; CHAVES, 2007; JIMÉNEZ *et al.*, 2008).

A *Wavelet* escolhida para a aplicação da transformada *Wavelet* em um sinal é conhecida como *Wavelet* mãe, esta sofre algumas modificações, e.g., pode ser expandida ou comprimida e se movimenta ao longo do sinal. As modificações são responsáveis pelos parâmetros de escala e deslocamento (GROSSMANN; MORLET, 1984). Define-se a *Wavelet* mãe da seguinte forma:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad a, b \in \mathbb{R}; a \neq 0, \quad (17)$$

sendo:

a : parâmetro de escala;

b : parâmetro de deslocamento.

A transformada *Wavelet* permite executar a análise de resolução múltipla, que infere o sinal com uma resolução adaptativa em diversas frequências (MALLAT, 1989a; MALLAT, 1989b). Esta análise é baseada na decomposição do sinal em diferentes componentes, a qual altera a resolução no tempo e frequência para resolver problemas do processamento de sinais que inclui estruturas com diferentes resoluções de tempo-frequência. Portanto, a transformada *Wavelet* é projetada para produzir uma alta resolução no tempo e uma baixa resolução espectral para sinais de alta frequência e uma alta resolução espectral para sinais de baixa frequência (DEBNATH; SHAH, 2015).

Existem variantes nas transformadas *Wavelet*, como a transformada *Wavelet* contínua, a transformada *Wavelet* discreta, transformada rápida de *Wavelet* e a transformada *Wavelet* de *packet*. Os dois primeiros serão explicitados a seguir.

3.1 TRANSFORMADA WAVELET CONTÍNUA

A transformada *Wavelet* contínua (TWC) em $L^2(\mathbb{R}^2)$ de um sinal de tempo contínuo $x(t) \in L^2(\mathbb{R})$ é definida como a soma da multiplicação deste sinal contínuo e a *Wavelet* mãe em sua forma deslocada e escalonada $\psi_{a,b}(t)$, que possui a forma:

$$TWC(a, b) = \langle x(t), \psi_{ab}(t) \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (18)$$

Em que:

$x(t)$: o sinal contínuo.

A transformada *Wavelet* satisfaz a condição de inversibilidade, e o sinal é recuperado a partir dos coeficientes *Wavelet* conforme (GROSSMANN; MORLET, 1984; DEBNATH; SHAH, 2015):

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{CWT(a, b) \psi_{ab}(t)}{a^2} da db \quad (19)$$

Desde que C_ψ satisfaça a condição de admissibilidade.

Os parâmetros de escala e deslocamento permitem a obtenção dos coeficientes *Wavelet*. Os coeficientes da *Wavelet* mostram qual é a relação existente entre o sinal e a *Wavelet* mãe (CORTÉS; GARAZÓN; CHAVES, 2007). Esta relação permite conhecer os componentes de frequência do sinal. Entretanto, essa transformada não é prática e torna-se redundante, portanto, foram propostos parâmetros discretizados de deslocamento e de escalamento na transformada *Wavelet* discreta (AKANSU; SERDIJN; SELESNICK, 2010).

3.2 TRANSFORMADA WAVELET DISCRETA

A transformada *Wavelet* discreta (TWD), usadas para superar as redundâncias da *Wavelet* contínua, é obtida por meio da discretização dos parâmetros de escala e deslocamento da transformada *Wavelet* contínua (ZHANG; WANG; HAN, 2010; AKANSU; SERDIJN;

SELESNICK, 2010). Foram escolhidos coeficientes representativos para efetuar a discretização, apresentados da seguinte forma:

$$a = 2^{-j} \quad (20)$$

$$b = k2^{-j} \quad (21)$$

a é a escala, b é o deslocamento e, j e k devem ser valores inteiros.

Tendo-se definido os parâmetros de escala e deslocamento com valores discretos, será obtido uma amostragem diádica nos eixos de tempo e de frequência, e, portanto, obteve-se a seguinte *Wavelet* diádica:

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k); j, k \in Z \quad (22)$$

Assim, permite-se definir a transformação *Wavelet* discreta da função de escala complementar $\phi(t)$ como:

$$TWD_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle = 2^{\frac{j}{2}} \int_{-\infty}^{\infty} x(t) \phi(2^j t - k) dt \quad (23)$$

sendo:

$x(t)$: sinal discreto.

Esta *Wavelet* de característica ortonormal e as bases de escala satisfazem conjuntamente sua propriedade de base complementar:

$$\int \psi_{j,k}(t) \phi_{r,s}(t) dt = 0, \quad \forall j, k, r, s \quad (24)$$

A transformada *Wavelet* discreta permite reconstruir o sinal, uma vez que se calcula os coeficientes *Wavelet*. Para realizar esta reconstrução, duas funções são necessárias: a função *Wavelet* $\psi(t)$ e $\phi(t)$ a função de escala. A reconstrução do sinal é obtida empregando a equação (25):

$$x(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t), \quad j, k \in Z \quad (25)$$

$c_{j,k}$ são os coeficientes de aproximação e $d_{j,k}$ os coeficientes de detalhe. Deve-se mencionar que os coeficientes de aproximação estão associados à função escala, enquanto os coeficientes de detalhe estão associados à função *Wavelet*.

Os coeficientes de aproximação e detalhe nos permitem obter informações sobre as características do sinal, e com sua manipulação pode-se obter um novo sinal eliminando componentes indesejáveis do sinal original.

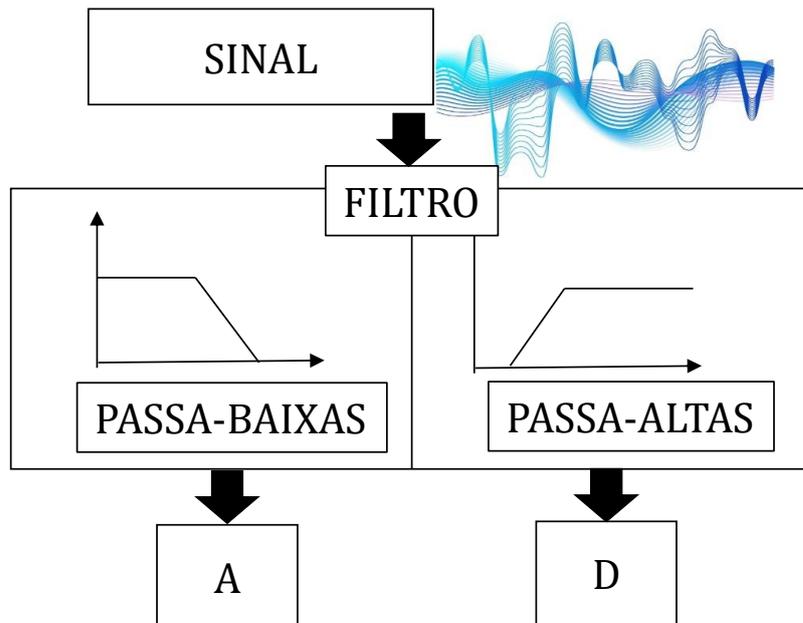
O desenvolvimento do algoritmo para avaliar o TWD levou ao desenvolvimento da análise multirresolução (MALLAT, 1989_a; MALLAT, 1989_b), contemplando a implementação de banco de filtros (VETTERLI; HERLEY, 1992). Estes filtros correspondem a um filtro passa-baixas e passa-altas, sendo que quando o sinal original passa por tais filtros, obtêm-se os coeficientes de saída $c_{j,k}$ e $d_{j,k}$, respectivamente. Uma decomposição de um sinal em diferentes níveis pode ser obtida passando os coeficientes de escala obtidos na filtragem anterior por um par de filtros idênticos, obtendo assim os coeficientes do próximo nível (MERRY; STEINBUCH, 2005).

3.3 BANCO DE FILTRO UTILIZANDO A TRANSFORMADA WAVELET

A análise multirresolução dos sinais, que utiliza a transformada *Wavelet* diádica ortogonal, é um processo de aplicação sucessiva da transformada *Wavelet* em resoluções múltiplas. Este equivale a sequencias de subespaços encaixantes nos filtros dispostos na Figura 1. Neste processo, é realizado uma divisão dos coeficientes, obtendo-se os coeficientes de aproximação (A) e detalhe (D), os quais são resultados dos operadores de amostragem como *downsampling* no processo de decomposição e *upsampling* na reconstrução de sinais (VETTERLI; HERLEY, 1992). Os coeficientes de aproximação representam os valores altos de escala, correspondendo aos componentes de baixa frequência do sinal, portanto, está

associada à função de escala que é determinada com um filtro passa-baixa. De forma contrária, os coeficientes de detalhes são os valores baixos de escala correspondentes aos componentes de alta frequência, e está associada à função *Wavelet* que é determinada como um filtro passa-altas (JIMÉNEZ *et al.*, 2008).

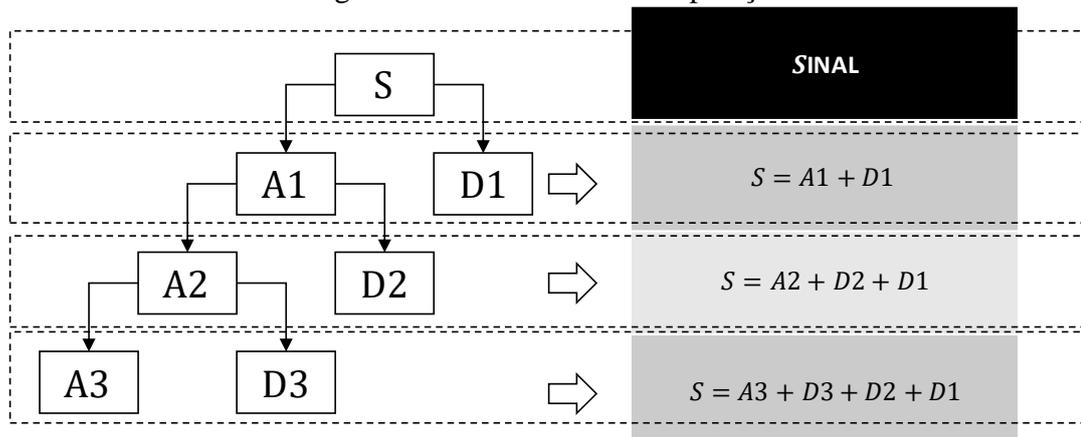
Figura 1 - Esquema do processo de filtragem



Fonte: Elaborado pelo próprio autor.

O processo de filtragem ou decomposição é repetido para descompor N níveis. O número de vezes que o sinal é filtrado é determinado pelo número de níveis de decomposição, em cada nível de decomposição origina-se o dobro de amostras do sinal original, portanto precisa-se descartar a duplicidade de amostras utilizando o operador *downsampling*. Na Figura 2 mostra-se a decomposição em três níveis de resolução.

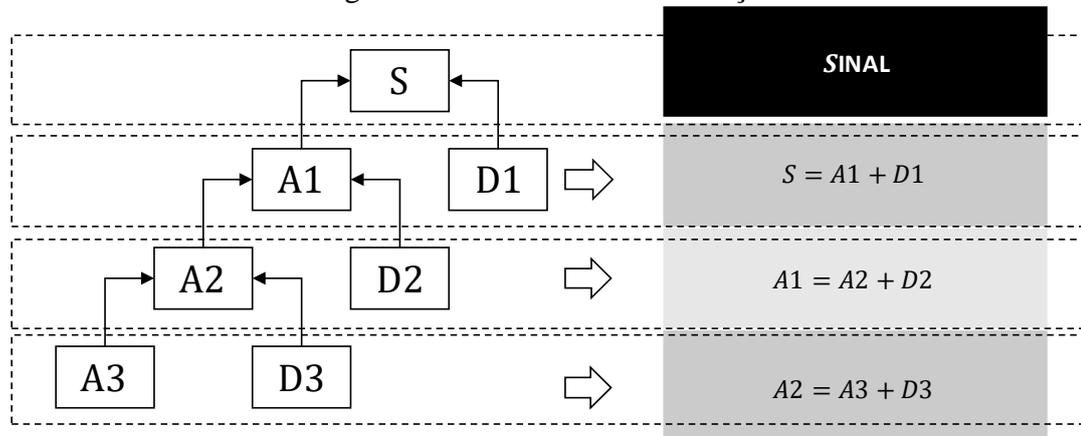
Figura 2 - Processo de decomposição



Fonte: Elaborado pelo próprio autor.

No processo reconstrução Figura 3, o operador *upsampling* é aplicado aos componentes dos canais passa-altos e passa-baixos no intuito de recuperar e/ou reconstruir o sinal original a partir da combinação dos componentes processados.

Figura 3 - Processo de reconstrução



Fonte: Elaborado pelo próprio autor.

4 *MACHINE LEARNING*

Machine Learning é a aplicação de recursos baseados na computação para implementar algoritmos de aprendizado, considera-se um processo computacional complexo de reconhecimento automático de padrões, consistindo em procedimentos inteligentes para tomadas de decisão baseados em um processo de treinamento (DUA; DU, 2011). Particularmente é um programa de computador que consegue melhorar em termos de desempenho as tarefas atribuídas nos modelos por meio da experiência.

A ML tem como função imprescindível automatizar tarefas, construindo modelos analíticos que permitam executá-las em uma ordem cognitiva como detecção de objetos ou tradução à linguagem natural (DUA; DU, 2011; JANIESCH; ZSCHECH; HEINRICH, 2021).

A característica fundamental do ML reside no conjunto de métodos que possuem a capacidade de detectar padrões de forma automática em dados. Após identificados os padrões nos dados, estes são utilizados para prever cenários futuros por meio de uma análise, de forma que se garanta decisões de baixa incerteza.

Desde uma abordagem associativa na criação de conhecimentos, destacam-se algoritmos de treinamento como processo de adaptação de informação (HAYKIN, 2008). Classificam-se estes algoritmos tanto como preditivo ou supervisionado, quanto descritivo ou não supervisionado, os quais são associados aos diversos métodos de ML (MURPHY, 2012).

Portanto, nesta seção apresentar-se-ão, com mais detalhe os algoritmos de aprendizagem utilizados e os principais conceitos envolvidos para uma melhor compreensão das ML.

4.1 TIPOS DE ALGORITMOS DE APRENDIZAGEM

Nesta subseção, apresentam-se os conceitos básicos sobre os tipos de algoritmos de aprendizagem pertencentes aos métodos de ML. Na literatura destaca-se a capacidade da ML para criar associações automáticas por meio de padrões no intuito de generalizar informações. Este processo é conhecido como aprendizagem, a qual pode ser supervisionada ou não.

4.1.1 **Aprendizagem supervisionada**

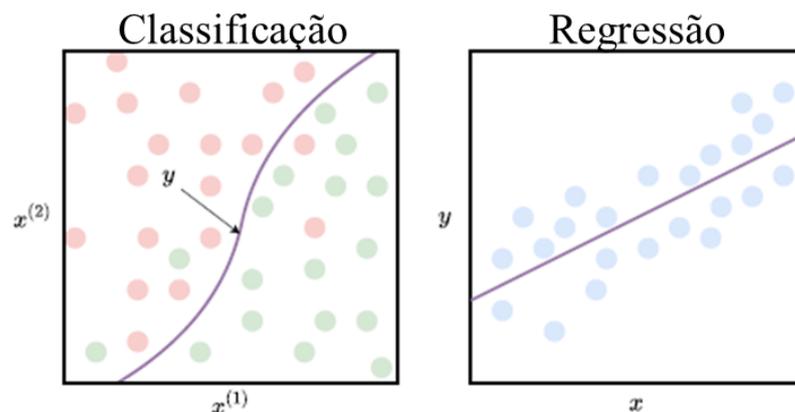
A aprendizagem supervisionada tem como objetivo encontrar um mapeamento ou generalização entre dados relativos às entradas e às saídas de um modelo. Neste processo é utilizado uma base de dados para produzir uma captura das características relevantes entre entradas e saídas de um vetor, permitindo assim deduzir rótulos na amostra (BURKOV, 2020).

Este processo está baseado em um algoritmo alimentado com dados da amostra que são rotulados. O algoritmo usa a amostra rotulada para treinar e obter um modelo. Após o aprendizado por conta do treinamento, o modelo tem a capacidade de rotular informações dos dados que nunca foram utilizados pelo algoritmo, desta forma visa alcançar uma alta precisão em classificação e regressão (DUA; DU, 2011).

O aprendizado supervisionado utiliza a classificação e regressão como meio de prever uma ampla variedade de informações. O problema de prever uma classe ou múltiplas classes é conhecido como classificação, conquanto o processo de prever um número real é denominado regressão (DUA; DU, 2011).

A diferença entre classificação e regressão disposta na Figura 4 reside no fato que o algoritmo de aprendizagem durante a classificação procura uma linha que separa classes entre si. No entanto, na regressão, busca-se encontrar uma linha ou hipersuperfície que acompanhe de perto os exemplos de treinamento. Um exemplo de problema de classificação, dar-se-á quando um médico ingressa as características dos sintomas de um paciente em um software, e seguidamente este fornece um diagnóstico. Por outro lado, um problema de regressão pode ser a previsão do mercado de energia baseado nas características de variáveis exógenas, *e.g.*, dados meteorológicos, dias festivos, condições hidrológicas e programas de manutenção (BURKOV, 2020).

Figura 4 - Diferença entre classificação e regressão



Fonte: Adaptado (BURKOV, 2020).

Em síntese, a aprendizagem supervisionada consiste na procura de associações por meio de um “professor externo” que vincula entradas e saídas em uma base de dados tanto em problemas de classificação, quanto de regressão.

4.1.2 Aprendizagem não supervisionada

A aprendizagem não supervisionada, contrária à aprendizagem supervisionada, não permite a criação de rótulos nos dados. A inexistência de rótulos é consequência da ausência de dados de saída, a qual possui a grande vantagem de criar modelos que permite capturar características dos vetores de entrada e atribuí-las em categorias, situação que não acontece com a aprendizagem supervisionada.

Neste processo de captura das características ou categorização, o modelo retorna um número de identificação pertencente à categoria de cada vetor de entrada. Este processo é útil para localizar grupos de objetos semelhantes em uma grande coleção de objetos. Na categorização é possível identificar um subconjunto bastante representativo de uma grande coleção de exemplos não rotulados.

4.2 CONCEITOS BÁSICOS DE *MACHINE LEARNING*

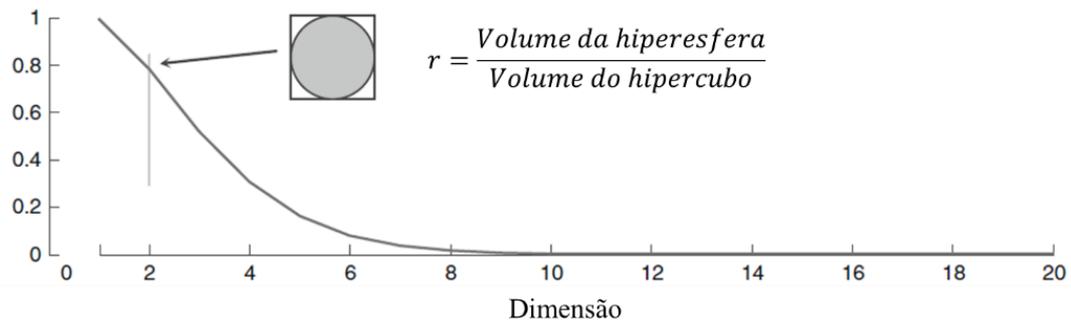
Nesta subseção apresenta-se uma introdução das ideias chave para a compreensão teórica dos métodos de ML a serem utilizados.

4.2.1 A maldição da dimensionalidade

A maldição da dimensionalidade é um problema que ocorre em ML quando é acrescentado nas entradas uma quantidade alta de dimensões, o modelo não consegue funcionar bem com entradas de alta dimensão e, como consequência, apresenta um pobre desempenho (MURPHY, 2012).

Este termo descreve o problema provocado pelo aumento exponencial de dimensões extras adicionadas ao espaço euclidiano da entrada, apresentado na Figura 5. Na medida que se acrescenta dimensões na entrada, a distância entre elas torna-se remota, portanto, o modelo não garante uma ótima generalização no processo de classificação.

Figura 5 - Representação da maldição da dimensionalidade



Fonte: Adaptado (SAMMUT; WEBB, 2017).

Este conceito normalmente está contrário à ideia errada de que um desempenho eficiente do classificador é geralmente atribuído às altas dimensões na entrada. Uma forma de mitigar então a maldição da dimensionalidade seria adicionando um número elevado de amostras, no intuito de minimizar as distâncias entre elas; entretanto, o problema apresentado envolveria um problema complexo de alto esforço computacional no processo de avaliação do desempenho (MURPHY, 2012; SAMMUT; WEBB, 2017).

4.2.2 *Overfitting*

No campo do ML, uma das questões a serem examinadas é a capacidade de generalização de um modelo, isso a partir dos dados disponíveis. Espera-se de um modelo, portanto, a capacidade de realizar previsões precisas e confiáveis em dados nunca antes vistos. Entretanto, às vezes no processo de reconhecimento de informações depara-se com um fenômeno conhecido como *overfitting*, que pode comprometer com a capacidade de generalização de um modelo.

O termo de *overfitting* é utilizado para descrever quando um modelo se ajusta excessivamente ao conjunto de dados de treinamento, este descreve de forma exata recursos que surgem de ruídos ou variações de dados, em vez de aprender os padrões subjacentes que podem ser generalizados para novos dados. Em contrapartida, na avaliação dos dados de teste, mostra-se uma ineficácia nos resultados porque o modelo memorizou padrões no treinamento, em vez de aprender relações mais amplas entre as variáveis (MURPHY, 2012; SAMMUT; WEBB, 2017; YING, 2019).

Existem várias causas potenciais para o *overfitting*. Um modelo pode ser muito complexo em relação à quantidade de dados disponíveis, o que permite que este seja sobreestimado na adaptação aos exemplos de treinamento. Além disso, se não houver cuidado na seleção de recursos relevantes, *e.g.*, presença de *outliers* ou se houver um desequilíbrio na

distribuição dos dados de treinamento, o modelo pode se tornar sensível às variações específicas nos dados de treinamento (YING, 2019).

Afortunadamente, existem estratégias e técnicas para lidar com problemas de *overfitting*. Uma dessas baseia-se na estratégia *Early-Stopping*, essa consiste em parar o treinamento antes do modelo aprender ruídos; além disso, tem-se a *Reduce The Size of Network*, a qual permite reduzir o ruído no conjunto de treinamento. Por outro lado, a estratégia da *Data-Expansion* é proposta para modelos complicados que requerem dados abundantes para ajustar sus parâmetros. E finalmente pode mencionar, a *Regularization*, a qual distingue ruídos com e sem significado, para posteriormente atribuir pesos (YING, 2019; BURKOV, 2020).

4.2.3 Seleção do modelo

Ao lidar com o desafio de selecionar um modelo adequado em ML, é comum buscar aquele que minimize o erro durante os estágios de treinamento, validação e teste. A abordagem óbvia seria escolher o modelo com o menor erro de teste. No entanto, quando o conjunto de dados é usado para o treinamento, não é possível acessar os dados de teste para a seleção do modelo com a complexidade correta. Para contornar essa limitação, uma estratégia comumente empregada é dividir os dados em conjuntos de treinamento e validação, permitindo a seleção do modelo mais adequado (MURPHY, 2012).

Nesse processo, recomenda-se destinar no treinamento em torno de 80% dos dados e 20% para a validação. Essa divisão permite avaliar o desempenho do modelo em um conjunto de dados independentes, além de evitar o *overfitting*. No entanto, quando a base de dados é pequena, essa divisão pode resultar em um conjunto de treinamento com dados insuficientes para o ajuste adequado do modelo, bem como uma estimativa imprecisa do seu desempenho. Para superar essa limitação, a literatura recomenda o uso da validação cruzada (MURPHY, 2012).

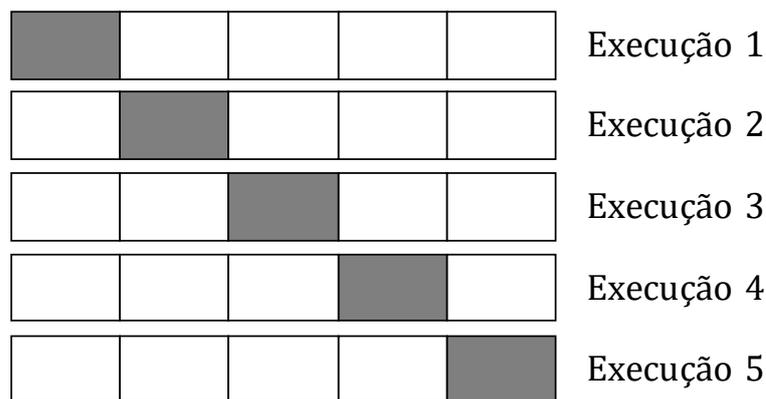
4.2.4 Validação cruzada

A validação cruzada é uma técnica essencial na avaliação de modelos de ML, utilizada para determinar a capacidade de generalização de um modelo em relação aos dados de treinamento. Além de ajudar a evitar o *overfitting*, a validação cruzada é especialmente útil quando se dispõe de um conjunto limitado de dados, pois permite avaliar a capacidade de generalização de um modelo a partir de um número reduzido de amostras (KING; ORHOBOR; TAYLOR, 2021).

Uma das principais vantagens da validação cruzada é fornecer uma estimativa mais precisa do desempenho do modelo em dados de teste não vistos, o que é útil na seleção do modelo mais adequado para um determinado problema. No entanto, a escolha adequada do número de subconjuntos “dobras”, a qual é resultado de dividir o conjunto em “ k ” vezes, pode ser um desafio, pois um número muito baixo pode resultar em uma estimativa imprecisa para o desempenho do modelo, enquanto um número muito alto pode aumentar o esforço computacional do processo (JAMES, *et al.*, 2013).

A validação cruzada consiste em dividir os dados de treinamento em k dobras, para cada dobra $k \in \{1,2,3,\dots,k\}$ treinam-se todas as dobras excetuando a k -ésima, a qual é empregado para o teste, conforme à Figura 6. Em seguida, calcula-se o erro médio de todas as etapas de teste, ou seja, realiza-se o procedimento de treinamento/teste k vezes (MURPHY, 2012).

Figura 6 - Esquema de validação cruzada de 5 vezes



Fonte: (MURPHY, 2012).

Em resumo, a validação cruzada é uma técnica essencial na avaliação de modelos de ML, sendo utilizadas para evitar problemas de *overfitting*, avaliar a capacidade de generalização e selecionar o modelo mais adequado para um determinado problema. Ao permitir uma avaliação mais precisa e robusta do desempenho do modelo, a validação cruzada é uma ferramenta poderosa para aprimorar a confiabilidade e a eficácia dos modelos de ML.

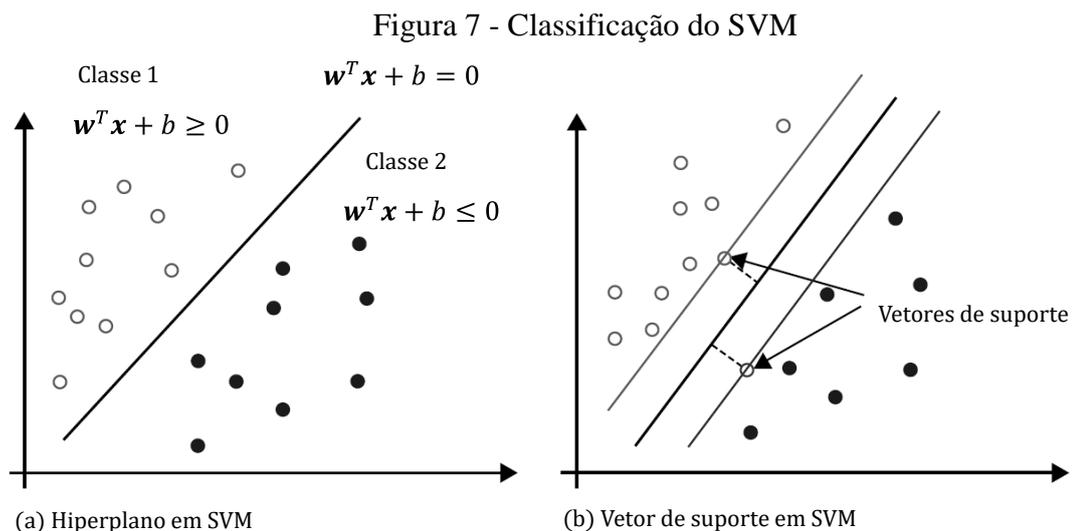
5 MÉTODOS DE MACHINE LEARNING UTILIZADOS

Neste capítulo apresentam-se os conceitos sobre dois métodos amplamente empregados em ML. O primeiro trata-se da *Support Vector Machine* (SVM) que visa classificar dados por meio de hiperplanos de separação, neste trabalho dar-se-á maior relevância ao processo de regressão das SVM. O segundo corresponde à rede neural da *Adaptive Theory Resonance* (ART), mais especificamente, o algoritmo de treinamento modificado (ARTMAP *Fuzzy-Modificada*), considerado o foco principal desta tese.

5.1 MÁQUINA DE VETORES DE SUPORTE

A Máquina de Vetores de Suporte (CORTES; VAPNIK, 1995) é um potente método de aprendizagem supervisionado utilizado no âmbito de *Machine Learning*, aplicado aos problemas de classificação e regressão de dados.

Na classificação mostrada na Figura 7, o marco conceptual deste processo está baseado na construção de um hiperplano que permite maximizar uma margem de separação entre os dados de um espaço de entrada, com ou sem atribuição de erros no seu desenvolvimento, em um cenário de erros busca-se a minimizá-los. Caso os dados sejam não lineares é possível por meio do mapeamento transformá-los a um espaço característico na procura de um hiperplano linear que consiga separar e/ou classificar notavelmente os dados (CORTES; VAPNIK, 1995).



Fonte: Adaptado de (DUA, 2012).

O processo de regressão das Máquinas de Vetores de Suporte consiste em uma ferramenta capaz de resolver estimativas de regressão em funções multidimensionais, neste

processo o objetivo é maximizar a margem no intervalo de confiança minimizando ao mesmo tempo os erros atribuídos fora desse intervalo, no qual o desempenho depende particularmente dos parâmetros de regularização, épsilon e do *Kernel* (VAPNIK; GOLOWICH; SMOLA, 1997; SMOLA; SCHÖLKOPF, 2004).

Neste sentido, para uma melhor compreensão, aprofundam-se os conceitos básicos sobre SVM nas subseções a seguir. Os temas estão desenvolvidos em torno dos conceitos sobre o hiperplano de separação ótimo que considera as margens rígidas no processo de classificação de dados, os hiperplanos de margens suaves, o método de convolução utilizando produto escalar para conseguir separar dados não lineares em um espaço característico e, finalmente, uma visão geral da regressão SVM e seu processo de otimização.

5.1.1 Hiperplano de separação ótimo

O algoritmo para a construção de um hiperplano de separação ótimo foi estabelecido detalhadamente no começo da década de sessenta, e apresentada no livro *Theory of pattern recognition* (VAPNIK; CHERVONENKIS, 1974). Entretanto, no artigo *Support-Vector Networks* (CORTES; VAPNIK, 1995) é fornecido uma versão mais simplificada do algoritmo, a qual será detalhada a seguir.

A proposta deste algoritmo é a construção de um hiperplano de separação ótima, mostrados na

Figura 8, considerando os dados de treinamento e que garanta uma ampla margem de distanciamento, sem erros entre elas (CORTES; VAPNIK, 1995).

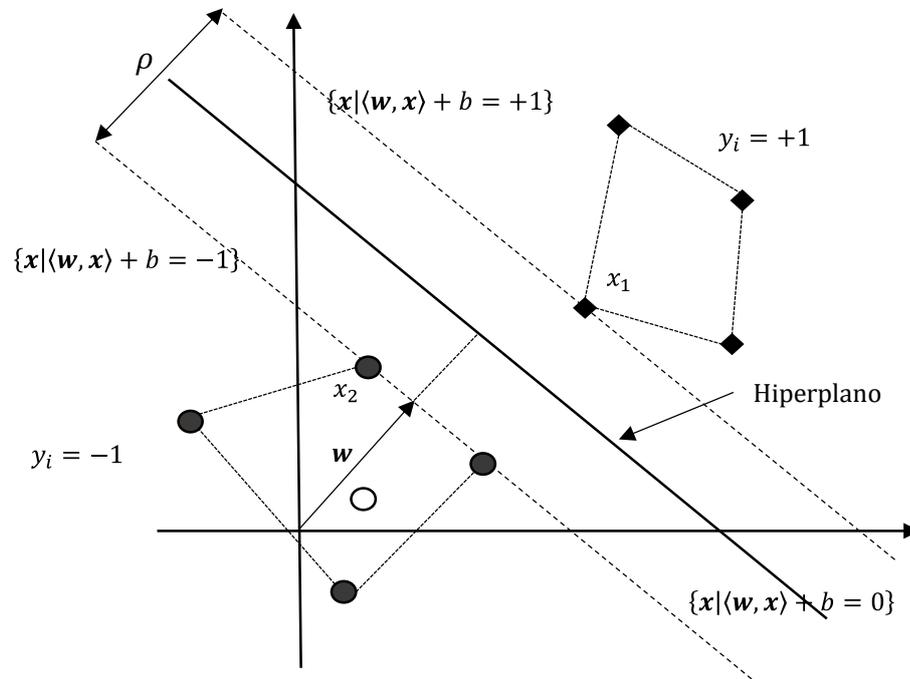
Os dados de treinamento são representados pela relação (26):

$$(y_1, \mathbf{x}_1), \dots, (y_l, \mathbf{x}_l), \quad y_i \in \{-1, 1\} \quad (26)$$

O hiperplano ótimo da separação é representado por meio da relação (27), a qual separa os dados de treinamento com uma margem máxima:

$$\mathbf{w}_0 \cdot \mathbf{x} + b_0 = 0 \quad (27)$$

Figura 8 - Hiperplano de separação ótima



Fonte: Elaborado pelo próprio autor.

Portanto, o objetivo é maximizar a margem de separação ρ entre os dados por meio da seguinte função:

$$\rho(\mathbf{w}, b) = \min_{\{x:y=1\}} \frac{\mathbf{x} \cdot \mathbf{w}}{|\mathbf{w}|} - \max_{\{x:y=-1\}} \frac{\mathbf{x} \cdot \mathbf{w}}{|\mathbf{w}|} \quad (28)$$

Sob as restrições:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \quad i = 1, \dots, l \quad (29)$$

O hiperplano ótimo (\mathbf{w}_0, b_0) , então, é o argumento que maximiza a distância da Equação (30), a qual é representado da seguinte forma.

$$\rho(\mathbf{w}_0, b_0) = \frac{2}{|\mathbf{w}_0|} = \frac{2}{\sqrt{\mathbf{w}_0 \cdot \mathbf{w}_0}} \quad (30)$$

Isto significa que o hiperplano ótimo é o único que minimiza a relação (31) sob a restrição disposta em (28):

$$R_1(\mathbf{w}, b) = \mathbf{w} \cdot \mathbf{w} \quad (31)$$

Neste sentido, a construção de um hiperplano ótimo é sintetizada como um problema de programação quadrática.

A representação compacta obtida para a resolução do problema de otimização quadrática consiste na determinação do vetor \mathbf{w}_0 , o qual determina o hiperplano que pode ser representado pela combinação linear de vetores de treinamento, conforme mostrado em (32).

$$\mathbf{w}_0 = \sum_{i=1}^l y_i \alpha_i^0 \mathbf{x}_i \quad (32)$$

Sendo o multiplicador de Lagrange $\alpha_i^0 \geq 0$. Adicionalmente, deve-se considerar que para encontrar o vetor do parâmetro α_i :

$$\mathbf{\Lambda}_0^T = (\alpha_1^0, \dots, \alpha_l^0), \quad (33)$$

Portanto, deve-se resolver o seguinte problema de programação quadrática (34):

$$W(\mathbf{\Lambda}) = \mathbf{\Lambda}^T \mathbf{1} - \frac{1}{2} \mathbf{\Lambda}^T \mathbf{D} \mathbf{\Lambda}, \quad (34)$$

Em relação à $\mathbf{\Lambda}^T = (\alpha_1, \dots, \alpha_l)$, considerar-se-ão as restrições:

$$\mathbf{\Lambda} \geq \mathbf{0}, \quad (35)$$

$$\Lambda^T Y = 0, \quad (36)$$

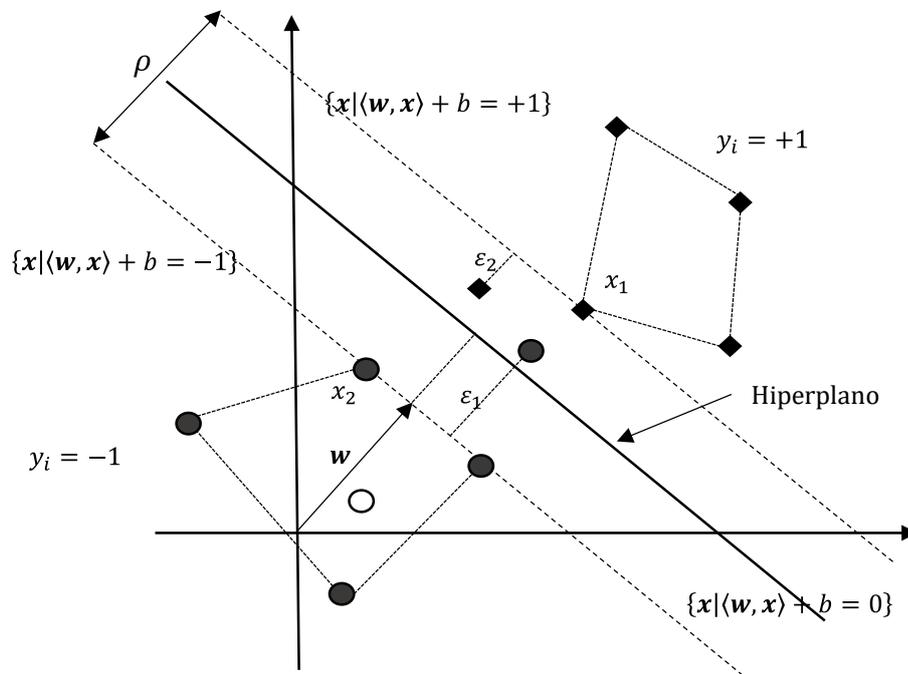
Sendo $\mathbf{1}^T = (1, \dots, 1)$ é a transposta de um vetor unitário l -dimensional, $Y^T = (y_1, \dots, y_l)$ é a transposta de um vetor l -dimensional de rótulos e D é uma matriz simétrica $l \times l$.

$$D_{ij} = y_i y_j x_i \cdot x_j, \quad i, j = 1, \dots, l \quad (37)$$

5.1.2 Hiperplano de margem suave

Ao contrário do algoritmo de separação ótima, o qual não considera erros na construção do hiperplano de separação, o algoritmo de margem suave (CORTES; VAPNIK, 1995) foca em separar os dados de treinamento com um número mínimo de erros, conforme disposta na figura Figura 9.

Figura 9 - Hiperplano de separação de margem suave



Fonte: Elaborado pelo próprio autor.

Deste modo, define-se o termo ε_i sendo o erro variável não negativa no intuito de minimizar a seguinte função:

$$\phi(\varepsilon) = \sum_{i=1}^l \varepsilon_i^\sigma \quad (38)$$

Para um valor pequeno $\sigma > 0$, sob as restrições:

$$\begin{aligned} y_i(\mathbf{w} \cdot \mathbf{x}_i + b) &\geq 1 - \varepsilon_i, & i = 1, \dots, l \\ \varepsilon_i &\geq 0, & i = 1, \dots, l \end{aligned} \quad (39)$$

No processo de otimização da relação (38) é possível encontrar um subconjunto mínimo de erros no treinamento, excluindo estes, pode-se construir um hiperplano de margem ótimo.

$$(y_{i_1}, \mathbf{x}_{i_1}), \dots, (y_{i_k}, \mathbf{x}_{i_k}),$$

Para conseguir construir um hiperplano suave de separação, deve ser minimizado o erro e maximizado a margem segundo a função objetivo apresentado na relação (40):

$$\frac{1}{2} \mathbf{w}^2 + CF \left(\sum_{i=1}^l \varepsilon_i^\sigma \right) \quad (40)$$

Em Cortes e Vapnik (1995) se detalha que para uma função F convexa monotônica com uma constante C suficientemente grande e σ pequeno, é possível determinar um hiperplano que consiga minimizar o número de erros no conjunto de treinamento, e separar os dados restantes dos elementos com uma margem máxima.

No intuito de procurar uma solução única e visando evitar o problema de NP-completo, foi considerado um σ igual a 1, para depois ser aplicado na relação (41). Isto para construir um hiperplano separador que minimize a soma dos desvios dos erros de treinamento e maximize a

margem com vetores classificados corretamente e, assim, garantir uma solução eficiente chamado hiperplano de margem suave (CORTES; VAPNIK, 1995).

$$\frac{1}{2} \mathbf{w}^2 + CF \left(\sum_{i=1}^l \varepsilon_i \right) \quad (41)$$

A função F apresentado na relação (41), pode ser simplificada a uma de segundo grau $F(u) = u^2$ para, desta forma, resolver um problema de programação quadrática. O vetor \mathbf{w} pode ser expresso como uma combinação linear de vetores suporte mostrado a seguir.

$$\mathbf{w}_0 = \sum_{i=1}^l y_i \alpha_i^0 \mathbf{x}_i,$$

Para encontrar o vetor Λ , tem de ser resolvido o problema de programação quadrática dual de maximização:

$$W(\Lambda, \delta) = \Lambda^T \mathbf{1} - \frac{1}{2} \left[\Lambda^T \mathbf{D} \Lambda + \frac{\delta^2}{C} \right], \quad (42)$$

Sob as restrições:

$$\Lambda^T \mathbf{Y} = 0, \quad (43)$$

$$\delta \geq 0, \quad (44)$$

$$\mathbf{0} \leq \Lambda \leq \delta \mathbf{1}, \quad (45)$$

Sendo Λ , \mathbf{Y} e \mathbf{D} os elementos utilizados para a otimização do problema na construção do hiperplano ótimo.

O menor valor admissível na função está representado pelo valor escalar δ da equação 46.

$$\delta = \alpha_{\max} = \max(\alpha_1, \dots, \alpha_l), \quad (46)$$

Substituindo na relação (42) é possível obter um problema de programação convexa, porque a relação (47) deixa de ser quadrática como consequência de α_{\max} . Portanto, para construir um classificador de margem suave, pode-se resolver o problema de programação convexa no espaço l -dimensional dos parâmetros Λ , ou pode-se resolver o problema de programação quadrática no espaço dual $l + 1$ dos parâmetros Λ e δ (CORTES; VAPNIK, 1995).

$$W(\Lambda) = \Lambda^T \mathbf{1} - \frac{1}{2} \left[\Lambda^T \mathbf{D} \Lambda + \frac{\alpha_{\max}^2}{C} \right], \quad (47)$$

5.1.3 O método de convolução do produto escalar no espaço característico

Conforme apresentado nas seções anteriores, os algoritmos descritos constroem hiperplanos de separação no espaço de entrada (espaço original dos dados) tendo em consideração margens rígidas ou suaves. Entretanto, há casos em que certos dados não lineares não podem ser adequadamente classificados por meio de um hiperplano, consequência disso são mapeados e levados a um espaço característico (espaço na qual as variáveis residem em uma alta dimensão) para consequentemente ser separado utilizando uma SVM linear (LORENA; DE CARVALHO, 2007). Portanto, para transformar um hiperplano existente no espaço de entrada e garantir uma alta probabilidade de classificação para dados não lineares, é necessário mapear um vetor de dimensão n para outro de dimensão N , utilizando a função vetorial ϕ . Essa função realiza o mapeamento para um espaço característico em uma dimensão superior, na qual os dados podem ser linearmente separáveis, mesmo que não fossem no espaço de entrada original (CORTES; VAPNIK, 1995):

$$\phi : \mathfrak{R}^n \rightarrow \mathfrak{R}^N$$

Um separador linear N -dimensional é construído a partir dos conjuntos de vetores mostrados na relação (48):

$$\phi(x_i) = \phi_1(x_i), \phi_2(x_i), \dots, \phi_N(x_i) \quad , i = 1, \dots, l. \quad (48)$$

Desta forma, a classificação de um vetor x desconhecido é feita primeiro transformando o vetor de espaço de separação $x \rightarrow \phi(x)$, em seguida, tomando o sinal da função.

$$f(x) = \mathbf{w} \cdot \phi(x) + b \quad (49)$$

Tendo-se em consideração o método do classificador de margem suave, o vetor \mathbf{w} pode ser escrito como uma combinação linear de vetores suporte:

$$\mathbf{w} = \sum_{i=1}^l y_i \alpha_i \phi(\mathbf{x}_i), \quad (50)$$

A linearidade do produto escalar, implica que a função de classificação dependa apenas de produtos escalares. A ideia é construir vetores de suporte considerando formas gerais do produto escalar em um espaço de Hilbert (YOUNG, 1988).

$$f(x) = \mathbf{w} \cdot \phi(x) + b = \sum_{i=1}^l y_i \alpha_i \phi(\mathbf{x}_i) \cdot \phi(x) + b, \quad (51)$$

$$\phi(\mathbf{u}) \cdot \phi(\mathbf{v}) = K(\mathbf{u}, \mathbf{v}), \quad (52)$$

Considerando a teoria de Hilbert (YOUNG, 1988), qualquer função simétrica $K(\mathbf{u}, \mathbf{v})$ pode ser expandida da seguinte forma:

$$K(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{u}) \cdot \phi_i(\mathbf{v}), \quad (53)$$

sendo:

$\lambda_i \in \mathbb{R}$ e ϕ_i autovalores e autofunções.

Uma condição suficiente para garantir que defina um produto escalar em um espaço característico é que todos os autovalores na expressão sejam positivos. Nesse sentido é imprescindível que se garanta o teorema de Mercer. Portanto, a convolução do produto escalar pode estar dada por qualquer função que satisfaça a condição de Mercer (CORTES; VAPNIK, 1995).

A função representada na relação (53) é denominada função *Kernel*. Esta recebe dados do espaço de entradas e leva o produto escalar desses dados no espaço característico. Este pode ser empregado sem conhecer o mapeamento, conferindo-lhe a simplicidade do seu cálculo e capacidade de representar satisfatoriamente os espaços característicos (LORENA; DE CARVALHO, 2007). Os *Kernel* mais utilizados apresentam-se nas relações (54), (55) e (56).

Kernel RBF:

$$K(\mathbf{u}, \mathbf{v}) = \exp\left(-\frac{|\mathbf{u} - \mathbf{v}|}{\sigma}\right), \quad (54)$$

Kernel Polinomial:

$$K(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \cdot \mathbf{v} + 1)^d, \quad (55)$$

Kernel Sigmoidal:

$$k(\mathbf{u}, \mathbf{v}) = \tanh(n\mathbf{u} \cdot \mathbf{v} + p), \quad (56)$$

Usando o produto escalar no *Kernel*, é possível construir diversos ML com arbitrários hiperplanos de decisão.

$$f(x) = \sum_{i=1}^l y_i \alpha_i K(x, x_i), \quad (57)$$

Para encontrar os vetores x_i e pesos α_i , é aplicado o mesmo esquema de solução para o classificador de margem ótimo ou suave. A única diferença é a inclusão da matriz D .

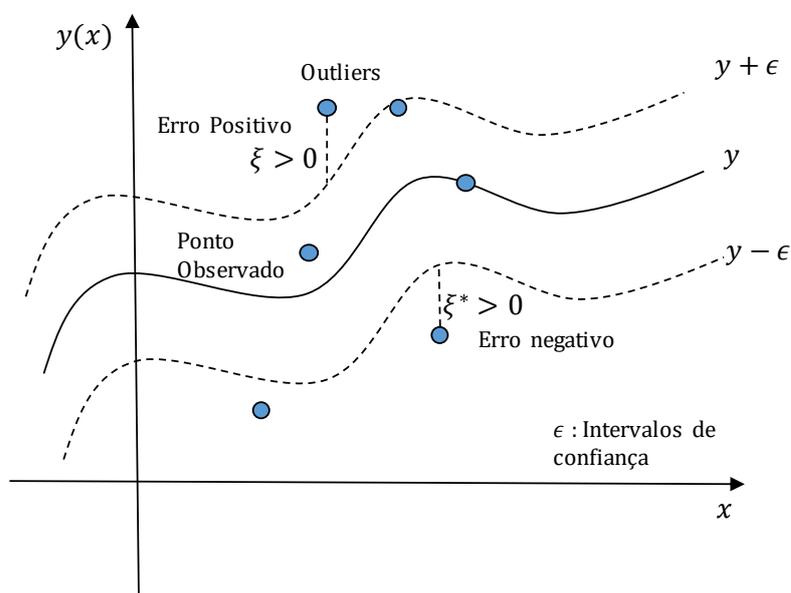
$$D_{ij} = y_i y_j K(x_i, x_j), \quad (58)$$

5.1.4 Regressão utilizando a máquina de vetores de suporte

O processo de regressão das Máquinas de Vetores de Suporte (SVM) é uma técnica amplamente utilizada para resolver problemas de estimativas de funções multidimensionais (VAPNIK; GOLOWICH; SMOLA, 1997). Essa técnica tem sido demonstrada como uma ferramenta universal para resolver problemas de reconhecimento de padrões, incluindo problemas não lineares, com boa capacidade de generalização, como foi evidenciado em estudos como de Cortes e Vapnik (1995). Além disso, em Vapnik *et al.*, (1997) foi demonstrado que os vetores de suporte são uma ferramenta valiosa na estimativa de regressão.

O treinamento na regressão consiste em resolver a função objetivo que garanta a maximização da margem e minimize o erro no processo de treinamento, conforme disposto na Figura 10.

Figura 10 – Diagrama esquemático da Regressão SVM usando ϵ - épsilon



Fonte: Elaborado pelo próprio autor.

Função de perda insensível ao épsilon, apresentada na relação (59), determina que os dados dentro da faixa criada por ϵ não são penalizados. Esta função é imprescindível na formulação da função objetivo do processo de treinamento visando a regressão.

$$L_{\epsilon}(y, \hat{y}) \triangleq \begin{cases} 0 & \text{se } |y - \hat{y}| < \epsilon \\ |y - \hat{y}| - \epsilon & \text{caso contrário} \end{cases} \quad (59)$$

A função objetivo cujas características relevantes são representadas pela forma convexa e a irrestrita, foram apresentadas por meio da relação (60). Esta se torna uma função não diferenciável em consequência do valor absoluto na função L_{ϵ} . Portanto, visando resolver este problema se reformula a equação (60) procurando um problema de otimização restrito (MURPHY, 2012).

$$J = C \sum_{i=1}^N L_{\epsilon}(y_i, \hat{y}_i) + \frac{1}{2} \|\mathbf{w}\|^2 \quad (60)$$

$$\hat{y}_i = f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i + w_0 \quad (61)$$

A constante de regularização (C) na relação (62) determina a troca entre os nivelamentos da função $f(\mathbf{x}_i)$ e o grau de desvios maiores do que ϵ são tolerados (SMOLA; SCHÖLKOPF, 2004).

$$C = \frac{1}{\lambda} \quad (62)$$

A função quadrática apresentada na relação (63) considera a abordagem de um problema de otimização restrito, na qual foram introduzido variáveis de folga para representar dados encontrados fora do tubo, apresentada na Figura 10 (MURPHY, 2012). Desta forma, foi reescrita a função objetivo quadrática que deverá ser minimizada considerando as restrições (64) (65) (66).

$$J = C \sum_{i=1}^N (\xi_i^+ + \xi_i^-) + \frac{1}{2} \|\mathbf{w}\|^2 \quad (63)$$

Sob as restrições:

$$y_i \leq f(\mathbf{x}_i) + \epsilon + \xi_i^+ \quad (64)$$

$$y_i \geq f(\mathbf{x}_i) - \epsilon - \xi_i^- \quad (65)$$

$$\xi_i^+ > 0; \xi_i^- > 0 \quad (66)$$

A solução ótima pertencente ao processo da minimização da função objetivo é apresentada pela seguinte relação (67). Deve-se mencionar também que o rendimento do algoritmo generalizado neste processo depende de um bom ajuste do parâmetro regularização C e do parâmetro épsilon e adicionalmente dos parâmetros relacionados ao *Kernel* (SMOLA; SCHÖLKOPF, 2004).

$$\hat{\mathbf{w}} = \sum_i \alpha_i \mathbf{x}_i \quad (67)$$

sendo:

$$\alpha_i \geq 0 \quad (68)$$

Após treinado o modelo, pode-se utilizar a seguinte equação para realizar previsões.

$$\hat{y} = \hat{w}_0 + \hat{\mathbf{w}}^T \mathbf{x} \quad (69)$$

Substituindo a solução de $\hat{\mathbf{w}}$ obtém-se:

$$\hat{y} = \hat{w}_0 + \sum_i \alpha_i \mathbf{x}_i^T \mathbf{x} \quad (70)$$

No intuito de generalizar a equação é possível apresentá-la em relação ao *Kernel*.

$$\hat{y} = \hat{w}_0 + \sum_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) \quad (71)$$

5.2 TEORIA DA RESSONÂNCIA ADAPTATIVA

A teoria da ressonância adaptativa (*Adaptive Resonance Theory* - ART) foi desenvolvida por Gail Carpenter e Stephen Grossberg com o objetivo de oferecer uma solução para o dilema de plasticidade e estabilidade que os designers de sistemas de aprendizagem enfrentam (CARPENTER; GROSSBERG, 1988). A plasticidade refere-se à capacidade da rede neural de adquirir novos conhecimentos sem perder informações previamente aprendidas, enquanto a estabilidade é a capacidade de agrupar corretamente os elementos em categorias definidas pelo modelo, à medida que os pesos são ajustados, os valores tendem a diminuir até alcançar um estado estável. (MORENO, 2010). As redes ART foram projetadas para codificar reconhecimentos estáveis em tempo real, usando a auto-organização como base, para responder a padrões de entrada arbitrários. Além disso, os princípios utilizados para resolver o problema de plasticidade-estabilidade incluem o aprendizado por correspondência, o aprendizado baseado em erro e o aprendizado rápido e estável (CARPENTER; GROSSBERG, 1988).

Quando se tratar do ambiente das redes neurais da família ART, a notação de vetor será usada por linha e não por coluna, como habitualmente na área da matemática. Esta forma tem sido recomendada pelos autores (CARPENTER; GROSSBERG, 1988) da linhagem das redes neurais ART. Ressalta-se que essa notação tem proporcionado um melhor entendimento das operações matemáticas a serem realizadas.

5.2.1 Rede neural artificial ARTMAP

A arquitetura de rede neural denominada RNA ARTMAP foi proposta por Carpenter, Grossberg e Reynolds (1991a), sendo caracterizada pela sua capacidade de aprendizado supervisionado. Esta rede tem como principal funcionalidade a habilidade de classificar automaticamente vetores ordenados arbitrariamente em categorias de reconhecimento, utilizando o treinamento como base dos sucessos preditivos.

O algoritmo de treinamento do tipo ARTMAP ocorre de forma supervisionada com base na estrutura dos módulos ART_a e ART_b . No decorrer do processo de treinamento, em termos gerais, o módulo ART_a é encarregado da recepção dos vetores de entrada $\mathbf{a}^{(p)}$ enquanto ART_b recebe $\mathbf{b}^{(p)}$ como entrada, na qual $\mathbf{b}^{(p)}$ torna-se a previsão correta para o padrão $\mathbf{a}^{(p)}$. Quando os dois módulos funcionam em conjunto, eles estão conectados a uma rede de aprendizado associativo e a um controle interno que garante o funcionamento autônomo do sistema em tempo real por meio do conceito de inter-ART, um terceiro módulo (CARPENTER; GROSSBERG; REYNOLDS, 1991a).

O inter-ART, como terceiro módulo, é responsável por analisar a correspondência e/ou compatibilidade entre os padrões de entrada e saída esperados, citados acima, operação conhecida como *match tracking*. Este mecanismo de operação tem a função de maximizar generalizações e, ao mesmo tempo, minimizar erros. Ao fazer uma previsão errada, por meio desta conexão associativa instruída, o parâmetro de vigilância que é responsável pela seletividade no módulo ART_a , é progressivamente aumentado até que seja possível atender a uma correção do erro, a fim de satisfazer uma categoria coincidente de previsão no módulo ART_b (CARPENTER; GROSSBERG; REYNOLDS, 1991a).

5.2.2 Rede neural artificial ARTMAP *Fuzzy*

A Rede Neural Artificial ARTMAP *Fuzzy* é uma rede de aprendizagem incremental supervisionada aplicada ao reconhecimento de categorias em mapeamentos multidimensionais, estáveis às sequências arbitrárias de vetores de entrada analógicos ou binários, isto por causa da sua capacidade de representar conjuntos *Fuzzy* (CARPENTER *et al.*, 1992).

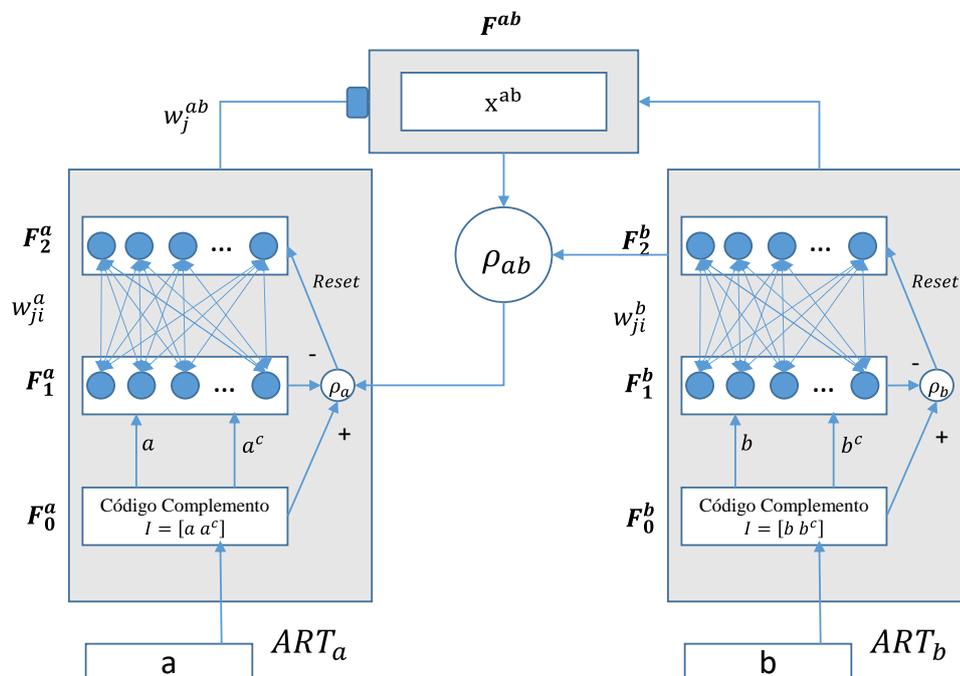
A estrutura da RNA ARTMAP *Fuzzy* está constituída por duas RNA não supervisionadas ART_{Fuzzy} (CARPENTER; GROSSBERG; ROSEN, 1991b). Estas redes não supervisionadas denominadas ART_a_{Fuzzy} e ART_b_{Fuzzy} criam categorias de reconhecimento estáveis em resposta a sequências de padrões arbitrários na entrada, encontram-se vinculadas entre si por meio de um módulo de controle interno denominado inter-ART (CARPENTER *et al.*, 1992).

Os módulos neurais ART_a e ART_b , desempenham a função de receber fluxos de padrões na entrada por meio do vetor \mathbf{a} e vetor \mathbf{b} , respetivamente; cabe mencionar que, \mathbf{b} está estritamente relacionada ao prognóstico de \mathbf{a} . Por outro lado, o módulo inter-ART é responsável pela análise de correspondência entre os dois módulos *Fuzzy*, um correto gerenciamento dos parâmetros de controle interno garante ao conjunto de módulos (modelo) rapidez, eficiência e

precisão, minimizando erros preditivos e maximizando generalizações (CARPENTER *et al.*, 1992).

A arquitetura mostrada na Figura 11 representa a RNA ARTMAP *Fuzzy*, sendo constituída pelos três módulos previamente tratados, dois sistemas ART *Fuzzy* e um sistema inter-ART. Os sistemas *Fuzzy* estão constituídos por três campos, respetivamente: um campo F_0 formado por nós que representam o vetor de entrada, um campo F_1 alimentado de baixo para cima por entradas *bottom-up* do campo F_0 e pelas entradas *top-down* do campo F_2 , que representam a categoria ativa (CARPENTER *et al.*, 1992).

Figura 11 - Arquitetura ARTMAP *Fuzzy*



Fonte: Adaptação (CARPENTER *et al.*, 1992).

Camada de entrada F_0 – esta camada cumpre a função de receber os vetores de entrada, atribui-se o processo de normalização e codificação de complemento desses dados (CARPENTER *et al.*, 1992).

A normalização garante no modelo o controle da proliferação de categorias, evitando-a. Neste caso, nos módulos *Fuzzy*, realiza-se a normalização de cada vetor apresentado pela rede da seguinte forma (CARPENTER *et al.*, 1992).

$$I = \frac{\mathbf{a}}{|\mathbf{a}|} \quad (72)$$

sendo.

\mathbf{a} : vetor de entrada;

$\mathbf{I} = [I_1 I_2 \dots I_M]$ (Normalizado);

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

$$|p| = \sum_{i=1}^M p_i; \quad (73)$$

A codificação de complemento, após realizada a normalização, cumpre a função de preservar a amplitude da informação fornecida (CARPENTER *et al.*, 1992).

$$a_i^c = 1 - a_i \quad (74)$$

É dessa forma que o vetor de entrada de dimensão M para cada módulo *Fuzzy*, após o processamento nesta camada terá em F_1 uma dimensão $2M$, pois:

$$\mathbf{I} = [\mathbf{a} \ \mathbf{a}^c] \quad (75)$$

Camada de comparação F_1 – nesta camada apresenta-se a recepção de informação da camada F_0 , este modo de intercâmbio de informação de baixo para cima se conhece como *bottom-up*, e da mesma forma, F_1 recebe informação fornecida de cima para baixo, ou seja, da camada F_2 , conhecida como *top-down*.

A camada F_1 recolhe informação da categoria ativa em F_2 , e da entrada, do vetor da ativação da camada F_0 , a partir disso obedece à equação (76).

$$\mathbf{x} = \begin{cases} \mathbf{I} & \text{se } F_2 \text{ estiver inativa,} \\ \mathbf{I} \wedge \mathbf{w}_j & \text{se existir categoria ativa em } F_2 \end{cases} \quad (76)$$

sendo:

\mathbf{I} : Vetor de ativação da camada F_0 ;

\mathbf{w}_j : Vetor de pesos entre as camadas F_1 e F_2 ;

\wedge : Operador *Fuzzy* AND, definido por:

$$(\mathbf{p} \wedge \mathbf{q})_i = \min(p_i, q_i). \quad (77)$$

sendo \mathbf{p} e \mathbf{q} vetores M -dimensionais.

A memória neste módulo ART *Fuzzy* está sustentada nos pesos das camadas F_1 e F_2 . Os pesos são representados pela matriz \mathbf{w}_{ji} , sendo i de dimensão $2M$ definida após realizada a codificação, que representa cada elemento da categoria armazenada do processo, e j de dimensão N representa cada categoria. Essa memória irá em aumento a cada nova categoria armazenada na matriz de pesos \mathbf{w} (CARPENTER *et al.*, 1992).

Camada de reconhecimento F_2 – nesta camada N -dimensional (N representa as unidades de *cluster*) ocorre a seleção da categoria ativa J . Esta categoria J é selecionada por meio de um processo atribuído à função de escolha T_j , isto a cada unidade de processamento da camada F_2 . Se o valor máximo de T_j for a ganhadora, então esse índice será a categoria ativa em J , caso exista empate será escolhida o índice mínimo (CARPENTER *et al.*, 1992).

$$T_j = \frac{|\mathbf{I} \wedge \mathbf{w}_j|}{\alpha + |\mathbf{w}_j|}; \quad (78)$$

sendo:

α : parâmetro de escolha ($\alpha \geq 0$).

Ressonância – a RNA entra em ressonância quando a similaridade entre padrões de entrada e a categoria J ativa consiga satisfazer a exigência da relação (79). Caso contrário,

ocorrerá o *reset*, que irá procurar um novo neurônio na camada F_2 . Esse processo de rejeição será realizado até encontrar um neurônio que seja aprovado no teste de similaridade, disposto em (79). No pior dos casos, quando nenhum neurônio for encontrado o vetor passará à condição de classe desconhecida, sendo criada uma nova categoria F_2 (CARPENTER *et al.*, 1992):

$$\frac{|I \wedge w_j|}{|I|} \geq \rho; \quad (79)$$

sendo:

I : Vetor de ativação da camada F_0 ;

w_j : Vetor de pesos da categoria ativa;

ρ : Parâmetro de vigilância.

Módulo Inter-ART – após o processo de ressonância nos módulos *Fuzzy* ocorrido satisfatoriamente, o módulo Inter-ART verifica a correspondência entre as categorias ativas destes, modulo ART_a e ART_b (CARPENTER *et al.*, 1992), resultado o vetor de ativação apresentado em (80).

$$x^{ab} = \begin{cases} \mathbf{y}^b \wedge \mathbf{w}_j^{ab}, & F_2^a \text{ ativo e } F_2^b \text{ ativo;} \\ \mathbf{w}_j^{ab}, & F_2^a \text{ ativo e } F_2^b \text{ inativo;} \\ \mathbf{y}^b, & F_2^a \text{ inativo e } F_2^b \text{ ativo;} \\ 0, & F_2^a \text{ inativo e } F_2^b \text{ inativo;} \end{cases} \quad (80)$$

sendo:

\mathbf{y}^b : Vetor de ativação da camada Fb_2 do módulo ART_b ;

\mathbf{w}_j^{ab} : Vetor de índice J da matriz de pesos do módulo Inter-ART, \mathbf{w}^{ab} .

A matriz de pesos \mathbf{w}^{ab} , de característica binária pertencente ao módulo Inter-ART, mapeia a categoria do modulo ART_a com sua respectiva em ART_b . Conseguindo isto, confirma-se a correspondência entres as categorias por meio do vetor \mathbf{w}_j^{ab} (CARPENTER *et al.*, 1992).

Match Tracking – o objetivo deste mecanismo é forçar a escolha de outra categoria ART_a em resposta à ausência de correspondência entre as categorias ativas ART_a e ART_b .

No começo de cada padrão de entrada em ART_a , o parâmetro de vigilância ρ_a será igual ao valor de vigilância base (parâmetro de vigilância *baseline*), $\overline{\rho_a}$. Para casos em que a relação (81) for satisfeita, o parâmetro de vigilância ρ_a terá um ligeiro incremento ε até que a relação (82) seja satisfeita (CARPENTER *et al.*, 1992):

$$|\mathbf{x}^{ab}| < \rho_{ab} |\mathbf{y}^b|; \quad (81)$$

sendo:

ρ_{ab} : Parâmetro de vigilância do módulo Inter-ART:

$$|\mathbf{I}^a \wedge \mathbf{w}_j^a| < \rho_a |\mathbf{I}^a|; \quad (82)$$

Nesse processo de incremento, o módulo ART_a busca outra categoria para ativar, seguidamente o módulo Inter-ART realiza uma nova verificação procurando uma correspondência entre os padrões de saída e entrada (CARPENTER *et al.*, 1992).

Após alcançado sucesso no processo de ressonância dos três módulos anteriormente explicados, a RNA passará a estabelecer as respectivas adaptações dos pesos, referente aos módulos ART_a e Inter-ART, mostrados nas relações (83) (84) (CARPENTER *et al.*, 1992).

$$\mathbf{w}_j^{novo} = \beta (\mathbf{I} \wedge \mathbf{w}_j^{velho}) + (1 - \beta) \mathbf{w}_j^{velho}; \quad (83)$$

$$\mathbf{w}_{JK}^{novo} = 1; \quad (84)$$

sendo:

β : Taxa de treinamento, com valor pertencente ao intervalo [0,1]

Cabe indicar que a classe ativa de ART_b não será alterada até acontecer uma próxima entrada e a adaptação dos pesos em ART_b é independente do módulo Inter-ART. A adaptação dos pesos é realizada depois de ressonância em ART_b , conforme a relação (85) (CARPENTER *et al.*, 1992).

$$\mathbf{w}_K^{novo} = \beta(\mathbf{I} \wedge \mathbf{w}_K^{velho}) + (1 - \beta)\mathbf{w}_K^{velho}; \quad (85)$$

5.2.3 Parâmetros da RNA ARTMAP Fuzzy

A estrutura da RNA ARTMAP Fuzzy é composta por vários parâmetros que devem ser ajustados previamente. Os parâmetros estão intrinsecamente relacionados com o problema. É desta forma que desempenham um papel relevante em relação à velocidade de aprendizagem, à eficiência e eficácia do sistema (CARPENTER *et al.*, 1992).

Parâmetro de escolha (α) – a função deste parâmetro é controlar a sequência de busca entre os nós da camada F_2 , seu valor deve ser $\alpha > 0$ (CARPENTER *et al.*, 1992).

Taxa de treinamento (β) – o valor da taxa de treinamento deve estar compreendido no intervalo $[0; 1]$, é responsável pela velocidade de adaptação dos pesos na rede (CARPENTER *et al.*, 1992). Quanto mais baixo for o valor atribuído ao parâmetro, mais lento será o processo de aprendizagem.

Parâmetro de vigilância (ρ) – (parâmetro especificado no intervalo $[0; 1]$) é responsável por controlar a ressonância na rede. Verifica a compatibilidade dos padrões de entrada e pesos para que ocorra a ressonância. Está diretamente relacionado ao número de categorias criadas. Quanto maior for o valor atribuído ao parâmetro, maior será o número de categorias criadas, no entanto, menor será a generalização da rede (CARPENTER *et al.*, 1992). A análise da ocorrência de ressonância, também chamada de critério de vigilância, é aplicada no módulo ART_a (ρ_a), no módulo ART_b (ρ_b) e no módulo interART (ρ_{ab}).

5.2.4 Algoritmo de treinamento da RNA ARTMAP Fuzzy – Modificada

O algoritmo de treinamento da RNA ARTMAP Fuzzy Modificada (MORENO, 2010) é um processo cujo objetivo é incluir um comportamento mais racional do funcionamento do neurônio biológico à ARTMAP Fuzzy convencional (CARPENTER *et al.*, 1992). Este procedimento visa aperfeiçoar o treinamento em uma forma mais racional e econômica, isto baseado no processo análogo à maneira como o cérebro reage frente a novos conhecimentos.

Uma forma prática de compreender este algoritmo é da seguinte forma – o cérebro contém uma grande quantidade de neurônios não ativos, as quais são utilizados para a construção de conhecimentos como resultados de um processo de aprendizagem. Deste processo, os neurônios entrarão ao estado ativado à medida que os novos conhecimentos sejam requeridos.

Esse processo começa apenas com um *cluster*, cujos dados constituem os dados da primeira entrada, e no decorrer do processo de treinamento começará a se expandir para assim acolher novos padrões. Desta forma, o cálculo da função de ativação é executado para um número reduzido de *clusters*, e também para o teste de vigilância (MORENO, 2010).

A seguir apresenta-se o algoritmo de treinamento da RNA ART *Fuzzy* Modificada (MORENO, 2010). Adicionalmente, na Figura 12 mostra-se o fluxograma aplicado ao treinamento da RNA ARTMAP *Fuzzy* Modificada, que é constituído por dois módulos ART *Fuzzy* e um Inter-ART.

Passo 1: Leitura dos dados de entrada.

Os dados de entrada são representados pelo seguinte vetor:

$$\mathbf{a} = [a_1 \ a_2 \ \dots \ a_M]; \quad (86)$$

Dimensão M . O vetor de entrada é normalizado para dessa forma evitar a proliferação de categorias:

$$\bar{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} \quad (87)$$

sendo:

$\bar{\mathbf{a}}$: Vetor normalizado.

$$|\mathbf{a}| = \sum_i^M a_i; \quad (88)$$

Passo 2: Codificação do vetor de entrada.

Visando preservar a amplitude da informação realiza-se a codificação:

$$\bar{a}_i^c = 1 - \bar{a}_i ; \quad (89)$$

Em que o vetor \bar{a}_i^c corresponde ao complemento do vetor da entrada normalizado. Desta forma, o vetor codificação terá dimensão $2M$ e será representado por:

$$\mathbf{I} = [\bar{\mathbf{a}} \ \bar{\mathbf{a}}^c]; \quad (90)$$

Observe que:

$$|\mathbf{I}| = \sum_i^M \bar{a}_i + \sum_i^M \bar{a}_i^c = M \quad (91)$$

Todos os vetores de entrada depois de realizada a normalização e codificação terão a mesma magnitude M .

Passo 3: Vetor de atividade.

O vetor de atividade será representado por:

$$\mathbf{y} = [y_1 \ y_2 \ \dots \ y_N]; \quad (92)$$

N é o número de categorias criadas em F2, assim tem-se:

$$y_j = \begin{cases} 1, & \text{se o nó } j \text{ de } F_2 \text{ for ativo} \\ 0, & \text{caso contrário} \end{cases} \quad (93)$$

Passo 4: Parâmetros da rede.

Os parâmetros utilizados no processamento da rede ART *Fuzzy* Modificada são:

1. Parâmetro de escolha: $\alpha > 0$
2. Taxa de Treinamento: $\beta \in [0,1]$
3. Parâmetro de Vigilância: $\rho \in [0,1]$

Passo 5: Inicialização dos pesos.

Comparada à rede ART *Fuzzy* convencional, a proposta do treinamento com a característica modificada começa com somente um *cluster* que contém dados do primeiro vetor de entrada.

$$w_{11}(0) = I_{11}; \quad w_{12}(0) = I_{12}; \quad \dots \quad w_{jM}(0) = I_{1,2M}; \quad (94)$$

Isto significa que existe apenas um *cluster* ativo com os dados do primeiro padrão.

Passo 6: Inicialização do contador de *clusters*:

$$N = 1; \quad (95)$$

Sendo,

N : Indica o número de *clusters* ativos.

Início do Treinamento:

Passo 7: Inicialização do contador.

No intuito de verificar se todos os *clusters* foram testados para acolher a referida entrada e, no caso que nenhum *cluster* existente seja adequado para a entrada atual, cria-se um *cluster* temporário para resolver esse problema, portanto foi imprescindível incluir um contador:

$$Cont = 1; \quad (96)$$

Em cada momento que uma entrada é apresentada à rede, este contador é inicializado.

Passo 8: Cálculo da função de escolha:

Tendo o vetor \mathbf{I} em F_1 , para cada nó em F_2 , a função de escolha é determinada por:

$$T_j = \frac{|\mathbf{I} \wedge \mathbf{w}_j|}{\alpha + |\mathbf{w}_j|}; \quad (97)$$

Busca do *cluster*:

Passo 9: Escolha da categoria.

A categoria é escolhida como sendo o nó J ativo, ou seja:

$$J = \arg \max\{T_j; \quad j = 1, 2, \dots, N\}; \quad (98)$$

Utilizando a equação mostrada é possível encontrar mais de uma categoria ativa, a categoria selecionada será aquela que possuir menor índice.

Passo 10: Teste de Vigilância.

O critério de vigilância é representado por:

$$\frac{|I \wedge w_j|}{|I|} \geq \rho; \quad (99)$$

Se o critério definido nesta equação for satisfeito, dirigir-se para o Passo 13.

Passo 11: Verificação se todos os *clusters* foram testados.

A verificação se todos os *clusters* já foram testados é feita por meio da inequação:

$$N < Cont; \quad (100)$$

Caso o critério definido não seja satisfeito, dirigir-se ao Passo 12. Caso contrário, o contador é atualizado da seguinte forma:

$$Cont^{novo} = Cont^{velho} + 1; \quad (101)$$

E o *reset* ocorre. No *reset*, o nó J de F_2 é excluído do processo de busca por meio de:

$$T_j = 0; \quad (102)$$

Então, retorna-se ao Passo 9.

Passo 12: Criação de *clusters*.

Para o caso na qual a relação (100) não for satisfeita significa que nenhum dos *clusters* já existentes comporta a entrada atual. Deste modo, se faz necessário a criação de um novo *cluster*, assim o contador de *clusters* é atualizado.

$$N^{novo} = N^{velho} + 1 ; \quad (103)$$

e,

$$J = N ; \quad (104)$$

daí

$$\mathbf{w}_J = \mathbf{I} ; \quad (105)$$

Após a criação do *cluster*, dirigir-se ao Passo 14.

Passo 13: Atualização dos pesos (Treinamento)

Nesse passo acontece o treinamento, o vetor de pesos é atualizado:

$$\mathbf{w}_J^{novo} = \beta(\mathbf{I} \wedge \mathbf{w}_J^{velho}) + (1 - \beta)\mathbf{w}_J^{velho} ; \quad (106)$$

sendo.

J : Categoria ativa;

\mathbf{w}_J^{velho} : Vetor peso referente à atualização anterior;

\mathbf{w}_J^{novo} : Vetor peso atualizado.

Passo 14: Vetor de atividade:

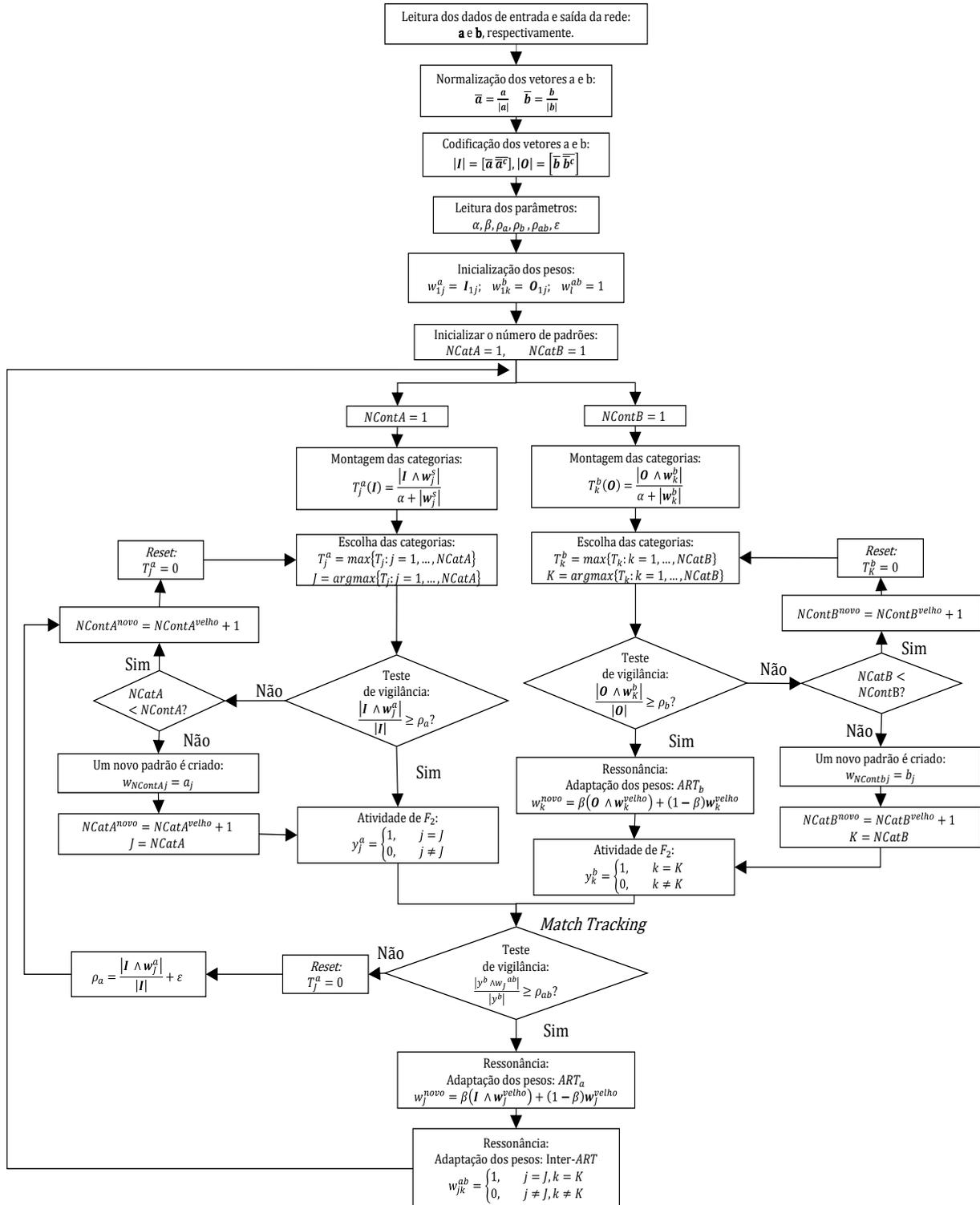
Este vetor de atividade F_2 é simbolizado por:

$$\mathbf{y} = [y_1 \ y_2 \ \dots \ y_N]; \quad (107)$$

Neste caso, N será o número de *clusters* criadas em F_2 . Dessa forma, apresenta-se a seguinte relação.

$$y_j = \begin{cases} 1, & \text{se o nó } j \text{ de } F_2 \text{ for ativo} \\ 0, & \text{caso contrário} \end{cases} \quad (108)$$

Figura 12 - Fluxograma da RNA ARTMAP Fuzzy Modificada



Fonte: Adaptação (MORENO, 2010).

5.2.5 Modelo híbrido da RNA ARTMAP Fuzzy – MMQ

Alves (2019) propôs um modelo híbrido utilizando a RNA ARTMAP Fuzzy - Modificada conjuntamente com a teoria dos mínimos quadrados para melhorar os resultados na

previsão de cargas elétricas residenciais mistas. Esta proposta conserva o treinamento realizado pelo algoritmo da ARTMAP *Fuzzy* Modificada (MORENO, 2010), no entanto, na fase de diagnóstico a rede desempenha o papel de selecionar as categorias para que o MMQ realize um ajuste em termos de regressão às categorias armazenadas, fornecendo dessa forma uma previsão de alta qualidade.

A parte fundamental do trabalho de Alves (2019) desenvolve-se em torno das quatro etapas da fase do funcionamento da arquitetura híbrida proposta, conforme apresentado na Figura 13.

Figura 13 – Pseudocódigo do modelo híbrido ARTMAP *Fuzzy*-MMQ

```

1.  #Primeira Etapa
2.  Treinamento da ARTMAP Fuzzy modificada.
3.
4.  # Segunda Etapa
5.  Para j = 1 até n :
6.    T(j) = soma(mínimo(I, wa(j)) / (alfa + soma (wa(j)))
7.  Fim
8.  i = 1
9.  Para j = 1 até n :
10.   J = arg(max(T(j))
11.   xa = soma(mínimo(I, wa(j)) / soma (I)
12.   Se xa >= ro_a :
13.     V(i) = J
14.     T(j) = 0
15.     i = i + 1
16.   Senão
17.     T(j) = 0
18.   Fim
19.
20. %%Terceira etapa
21. m = i - 1  # O índice i veio do item anterior
22. Para j = 1 até m:
23.   A = wa(v(j))
24.   K = arg(max(wab(v(j))
25.   B = wb(K)
26. Fim
27.
28. %%Quarta etapa
29. coeficientes_regressão = pseudo inversa(A) * B
30. saída_calculada = Ie * coeficientes_regressão

```

Fonte: Adaptação (ALVES, 2019).

Primeira etapa: Os padrões de entrada e saída desejada são apresentados à rede neural ARTMAP *Fuzzy* para treinamento.

Segunda etapa: Efetua-se o diagnóstico da rede. Portanto, é utilizado um vetor de entrada para realizar a previsão. Esta entrada será comparada com a matriz \mathbf{w}^a com a finalidade de encontrar os valores para a escolha de categorias.

De acordo com o algoritmo da RNA ARTMAP *Fuzzy*, o índice pertencente ao maior valor da categoria T_j ganha a oportunidade de ser avaliado se satisfizer o parâmetro de vigilância ρ_a . Os índices que satisfazem este parâmetro, serão armazenados em um vetor V .

Terceira etapa: Uma matriz A , das variáveis independentes, que será composta pelos elementos da matriz de pesos w^a , que estão associados a cada um dos índices armazenados em V . E a matriz B , das variáveis dependentes, que será composta pelas categorias de w^b que estão ligadas às categorias selecionadas em w^a . Assim, a matriz A , terá os elementos mais similares com a entrada para a previsão apresentada, enquanto B é composta pelos elementos com as saídas possíveis.

Quarta etapa: Esta última parte é encarregada da preparação dos dados, na regressão linear e da saída calculada.

As matrizes A e B que foram obtidas anteriormente estão configuradas de acordo com as matrizes de pesos da rede neural. Visando utilizá-las no método dos mínimos quadrados, será necessário adicionar uma coluna unitária à matriz A . Por outro lado, a matriz B , a qual possui duas colunas, sendo que a segunda coluna corresponde à codificação complementar que foi herdada da RNA, torna-se desnecessária para o MMQ, por esta razão esta coluna deve ser retirada de B .

6 METODOLOGIAS DO MODELO PREVISOR FLEXÍVEL

Neste capítulo são apresentadas as metodologias visando a previsão de cargas elétricas em diferentes níveis de desagregação. Estas metodologias foram desenvolvidas com base em técnicas de *Machine Learning*, utilizando como ferramentas a RNA ARTMAP *Fuzzy* modificada, o método dos Mínimos Quadrados e a Máquina de Vetor de Suporte, nono Quadro 1 estão descritas as características dos métodos inteligentes empregados. As métricas de desempenho adotadas para comparação e avaliação incluem RMSE, MSE, MAE e MAPE.

Quadro 1 - Metodologias e características

Metodologia		Características
Metodologia 1	SVM_ARTMAP <i>Fuzzy</i>	Processo simples: (1) Treinamento e teste de dados utilizando a versão híbrida composta pela RNA ARTMAP <i>Fuzzy</i> modificada e SVM (regressão). A metodologia é aplicada aos diversos níveis de desagregação.
Metodologia 2	Híbrido WPT-MMQ_ARTMAP <i>Fuzzy</i>	Processo paralelo: (1) Decomposição de dados com a transformada <i>Wavelet</i> ; (2) Treinamento e teste de dados suavizados com a RNA ARTMAP <i>Fuzzy</i> modificada; (3) Treinamento e teste de <i>outliers</i> com MMQ_ARTMAP <i>Fuzzy</i> . O processamento ocorre de forma paralela e é aplicado aos diversos níveis de desagregação.
Metodologia 3	Híbrido WPT-SVM_ARTMAP <i>Fuzzy</i>	Processo paralelo: (1) Decomposição de dados com a transformada <i>Wavelet</i> ; (2) Treinamento e teste de dados suavizados com a RNA ARTMAP <i>Fuzzy</i> modificada; (3) Treinamento e teste de <i>outliers</i> com SVM_ARTMAP <i>Fuzzy</i> . O processamento ocorre de forma paralela e é aplicado aos diversos níveis de desagregação.

Fonte: Elaborado pelo próprio autor.

Visando aprimorar as curvas de previsão, em diferentes níveis de desagregação, é empregado um procedimento que utiliza a análise multirresolução por meio da transformada *Wavelet*. Essa abordagem é um método poderoso para analisar diferentes níveis de decomposição, permitindo a filtragem de *outliers* e fornecendo dados suavizados com ausência de ruídos, preservando dessa forma as características da das cargas. Os dados suavizados e ruídos são armazenados em vetores para posterior avaliação por meio das metodologias 2 e 3 apresentadas neste trabalho. Nas subseções seguintes, apresenta-se com detalhes as metodologias utilizadas.

6.1 METODOLOGIA 1

Esta metodologia tem como objetivo realizar uma regressão, em um espaço multidimensional, das categorias identificadas no processo de diagnóstico da RNA ARTMAP *Fuzzy* modificada (MORENO, 2010), utilizando a técnica de regressão das Máquinas de Suporte de Vetores (SVM) para obter um modelo de previsão flexível que possa ser aplicado em diferentes níveis de desagregação das cargas elétricas.

Para o desenvolvimento desta metodologia, foram considerados aspectos importantes, como a seleção e preparação adequada dos dados com base em suas características, a estruturação clara do fluxograma e pseudocódigo para a construção do modelo SVM_ARTMAP *Fuzzy*, tudo isso com o propósito de alcançar o objetivo-alvo desta pesquisa.

A seguir, será apresentado o desenvolvimento detalhado desta metodologia.

6.1.1 Constituição e preparação dos dados

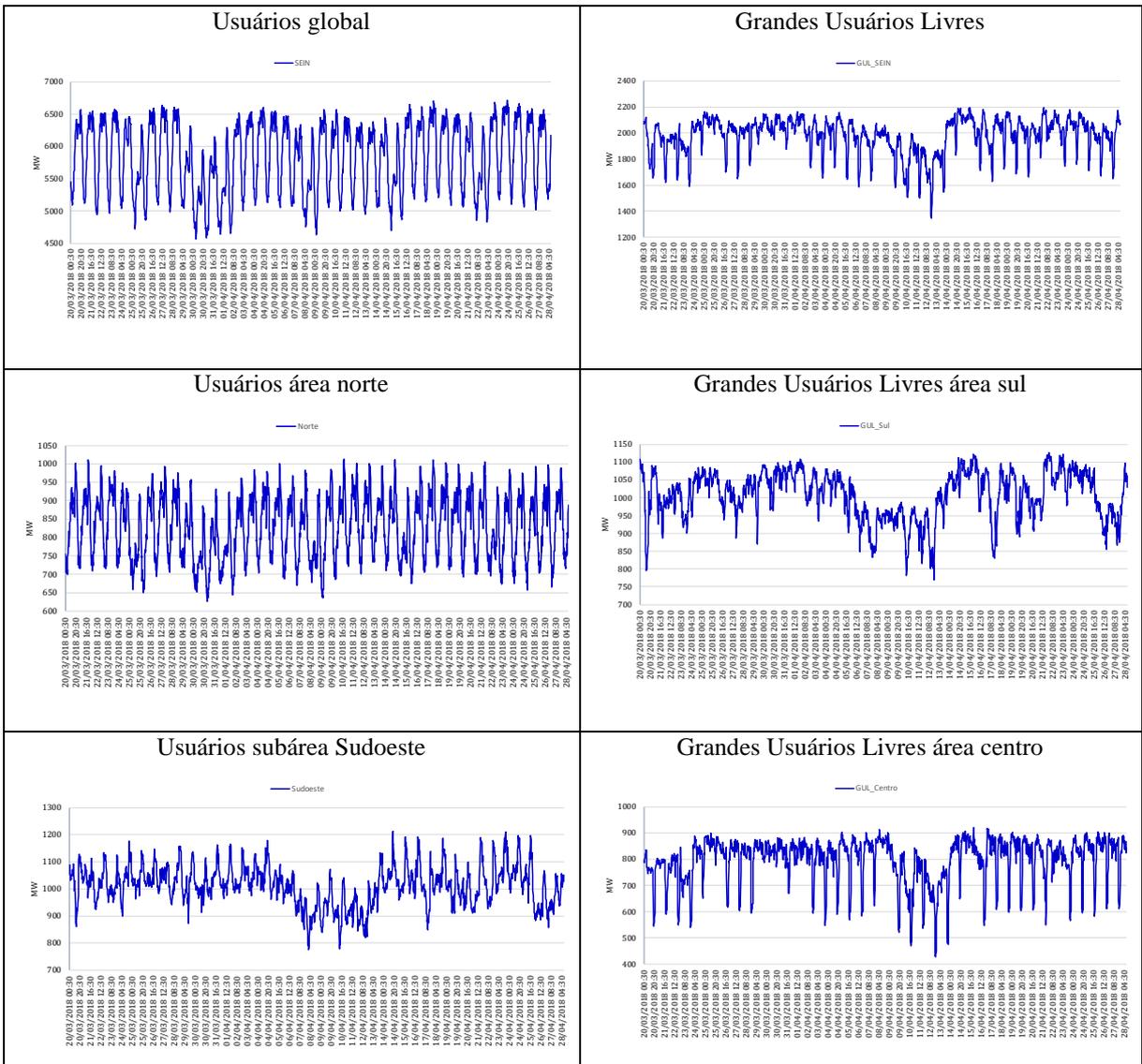
Nesta subseção, será descrita a composição e a estrutura dos dados utilizados para os processos de treinamento e teste empregado nesta primeira metodologia. Todos os dados usados nesta pesquisa, incluindo informações sobre as cargas elétricas e os dados meteorológicos, estão disponíveis publicamente. Esses dados foram coletados cada meia hora no período de 2017 a 2019, e podem ser encontrados no site do Operador do Sistema Interligado Nacional Peruano (COES¹) e no Serviço Nacional de Meteorologia e Hidrologia no Peru (Senamhi²).

A composição dos dados das cargas elétricas é apresentada a seguir. Tais cargas elétricas históricas são obtidas diariamente a partir dos relatórios da operação diário (IEOD) do sistema, publicados pelo COES. Estes dados são compostos pelas cargas globais, área norte, subárea sudoeste, grandes usuários livres, grandes usuários livres da área centro e sul. Na Figura 14 mostram-se os comportamentos das cargas em diferentes níveis de agregação. O nível de agregação global exibe um comportamento previsível durante períodos semanais; no entanto, o comportamento em níveis desagregados para os grandes usuários livres adiciona flutuações e complexidades nos padrões.

¹.Disponível em <<https://www.coes.org.pe/Portal/PostOperacion/Reportes/Ieod>>, último dia acessado 07/02/2023.

².Disponível em <<https://www.senamhi.gob.pe/?p=estaciones>>, último dia acessado 07/02/2023.

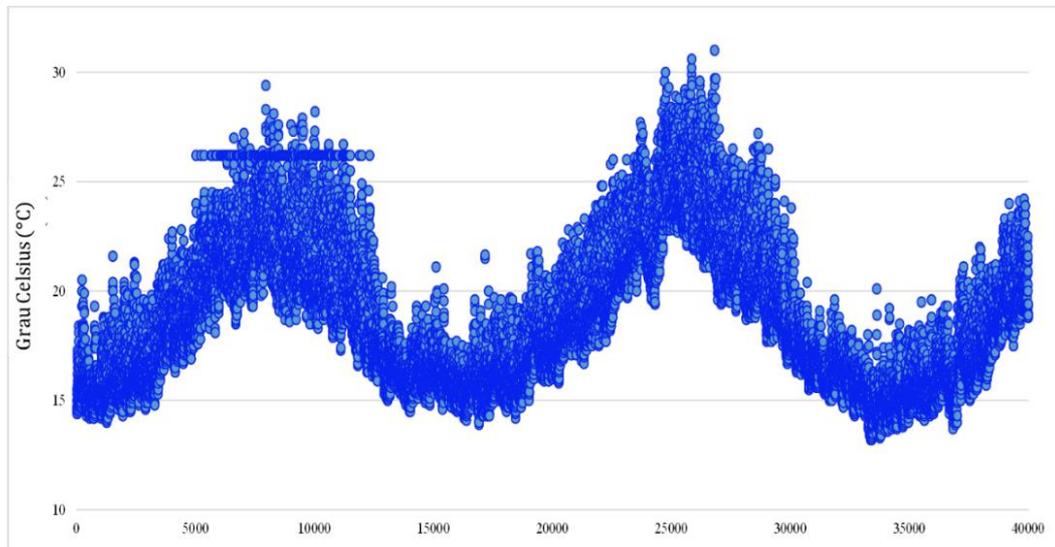
Figura 14 - Curvas das cargas em diferentes níveis de desagregação



Fonte: Elaborado pelo próprio autor.

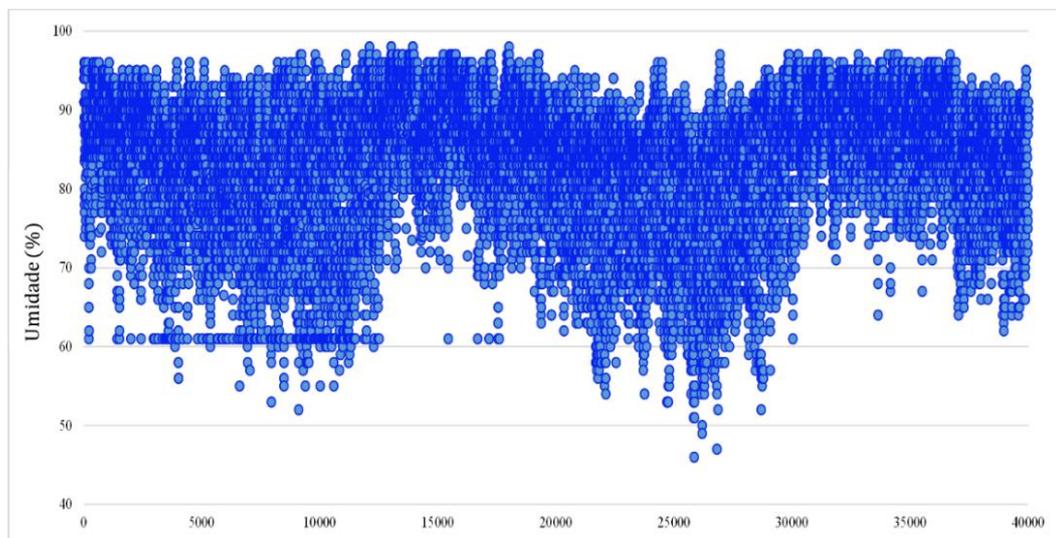
As variáveis meteorológicas, como temperatura, umidade, precipitação e velocidade do vento, da capital do Peru são fornecidas em tempo real pelo Senamhi, uma instituição pública peruana responsável pelo serviço nacional de meteorologia e hidrologia. Nesta pesquisa foram utilizados os dados de temperatura e de umidade como variáveis exógenas. Na Figura 15 ilustra-se a evolução da temperatura de setembro de 2017 até dezembro de 2019, demonstrando um comportamento sazonal com altas temperaturas no verão. Na Figura 16 apresenta-se o comportamento da umidade na cidade metropolitana de Lima no mesmo período, observando-se uma correlação inversa entre a evolução de uma variável em relação à outra.

Figura 15 - Temperatura Lima, Peru (2017-2019)



Fonte: Elaborado pelo próprio autor.

Figura 16- Umidade Lima, Peru (2017-2019)



Fonte: Elaborado pelo próprio autor.

O conjunto de dados das cargas elétricas e meteorológicas encontra-se em formato Excel publicado diariamente em padrões granulares de meia em meia hora. Leva-se em consideração que os dados coletados para este trabalho são obtidos pelas instituições públicas referidas acima e de livre acesso nos respectivos sites.

A composição do vetor de entrada para o sistema predictor inteligente tem como referência o movimento da janela, utilizada para gerenciar os padrões de entrada (PARK *et al.*, 1991). Neste sentido, o vetor de entrada a é definido da seguinte forma:

A entrada para os níveis de carga elétrica global, área norte, e subárea sudoeste será da seguinte forma.

$$\mathbf{a} = [t \quad dS \quad dAt \quad \mathbf{n1} \quad \mathbf{n2} \quad L(t-3) \quad L(t-2) \quad L(t-1) \quad L(t)], \quad \mathbf{a} \in \mathbb{R}^m \quad (109)$$

sendo.

t : codificação em relação ao tempo em intervalos de 30 minutos (vetor de seis bits de 00:30- [000001] às 24:00 - [110000]).

dS : codificação em relação ao dia da semana (vetor de três bits segunda-feira - [001] até Domingo - [111]).

dAt : codificação em relação aos dias normais e atípicos (um bit, dia normal [0] e dia atípico [1]).

$\mathbf{n1}$: vetor que armazena dois componentes normalizados, a temperatura $Temp(t)$, e a umidade $Hum(t)$.

$\mathbf{n2}$: vetor que armazena quatro componentes normalizados das quatro últimas semanas nos mesmos períodos horários, $L(t-1334)$, $L(t-1008)$, $L(t-672)$ e $L(t-336)$.

$L(t-3)$: carga relativa ao instante $t-3$.

$L(t-2)$: carga relativa ao instante $t-2$.

$L(t-1)$: carga relativa ao instante $t-1$.

$L(t)$: carga relativa ao instante t .

m : a dimensão do vetor de entrada.

A entrada para os níveis de carga elétrica dos grandes usuários livres, grandes usuários livres da área norte, e sul será da seguinte forma.

$$\mathbf{a} = [t \quad dS \quad dAt \quad \mathbf{n2} \quad L(t-3) \quad L(t-2) \quad L(t-1) \quad L(t)], \quad \mathbf{a} \in \mathbb{R}^{m_1} \quad (110)$$

sendo:

- t : codificação em relação ao tempo em intervalos de 30 minutos (vetor de seis bits de 00:30- [000001] às 24:00 - [110000])
- dS : codificação em relação ao dia da semana (vetor de três bits segunda-feira - [001] até Domingo - [111])
- dAt : codificação em relação aos dias normais e atípicos (um bit, dia normal [0] e dia atípico [1])
- $n2$: vetor que armazena quatro componentes normalizados das quatro últimas semanas nos mesmos períodos horários, $L(t-1334)$, $L(t-1008)$, $L(t-672)$ e $L(t-336)$.
- $L(t-3)$: carga relativa ao instante $t-3$.
- $L(t-2)$: carga relativa ao instante $t-2$.
- $L(t-1)$: carga relativa ao instante $t-1$.
- $L(t)$: carga relativa ao instante t .
- m_1 : a dimensão do vetor de entrada.

O vetor de saída b corresponde ao valor da carga no instante $t + 1$, *i.e.*, 30 minutos adiante.

$$b = [L(t + 1)], \quad b \in \mathbb{R}^1 \quad (111)$$

Uma das grandes vantagens desta metodologia é a possibilidade de utilizar tanto os valores de carga presentes quanto passados utilizando o método do movimento da janela, além de incluir variáveis exógenas como as meteorológicas, entre outras. Isto se deve ao fato de que a arquitetura ARTMAP *Fuzzy* não impõe limitações quanto à dimensão do vetor de entrada, bem como para o conjunto de padrões utilizados durante a fase de treinamento. Dessa forma, é possível explorar um amplo conjunto de informações para a construção dos modelos, permitindo uma análise mais abrangente e precisa das situações do sistema elétrico. Na Figura 17 mostra-se a estrutura organizada para os dados de carga elétrica global, área norte, e subárea sudoeste.

Figura 17 - Funcionamento do movimento da janela na etapa de previsão das cargas

Entradas para a previsão													Saída prevista
Tempo	Dia da semana	Dia atípico	Temperatura	Umidade	$L(t-1344)$	$L(t-1008)$	$L(t-672)$	$L(t-336)$	$L(t-3)$	$L(t-2)$	$L(t-1)$	$L(t)$	$L(t+1)$
h(1345)	[1 1 1]	0	0.3708	0.7963	L(1)	L(337)	L(673)	L(1009)	L(1342)	L(1343)	L(1344)	L(1345)	L(1346)
h(1346)	[1 1 1]	0	0.3708	0.7963	L(2)	L(338)	L(674)	L(1010)	L(1343)	L(1344)	L(1345)	L(1346)	L(1347)
h(1347)	[1 1 1]	0	0.3652	0.7778	L(3)	L(339)	L(675)	L(1011)	L(1344)	L(1345)	L(1346)	L(1347)	L(1348)
h(1348)	[1 1 1]	0	0.3652	0.7778	L(4)	L(340)	L(676)	L(1012)	L(1345)	L(1346)	L(1347)	L(1348)	L(1349)

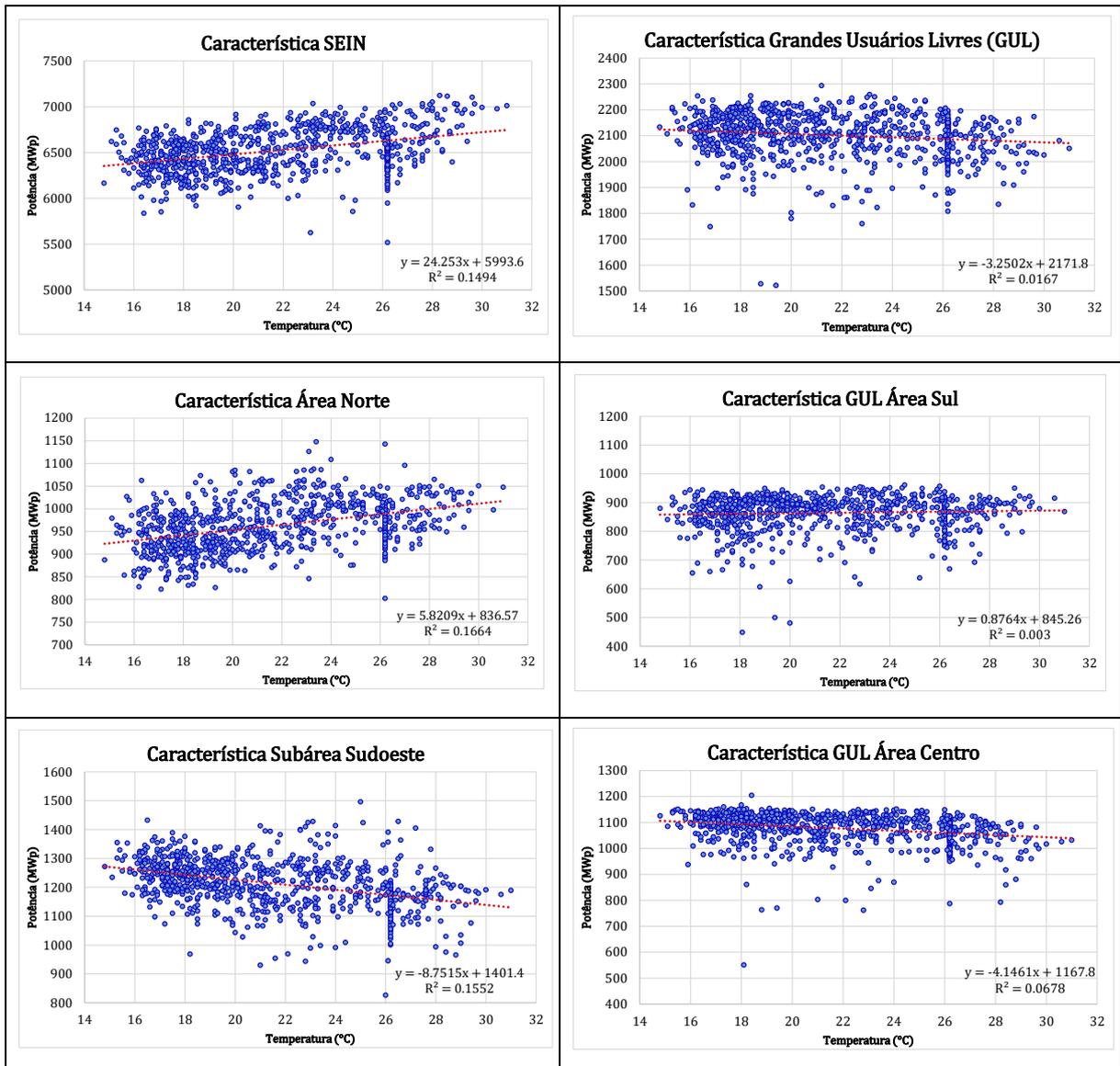
Fonte: Elaborado pelo próprio autor.

6.1.2 Características dos dados abordados no estudo

Nesta subseção é apresentada a descrição das características relevantes dos dados abordados neste trabalho, identificam-se duas particularidades que será explicado a seguir. Na primeira serão abordadas as correlações existentes entre dados cargas elétricas pico com os dados meteorológicos, conforme ilustrado na Figura 18. Já na segunda parte será realizada uma análise de variações em grandes blocos de energia, a fim de identificar dias atípicos.

A primeira característica dos dados abordados está relacionada à composição dos usuários, englobando tanto os usuários regulamentados quanto os livres. Na Figura 18 pode-se observar a correlação entre a potência de pico e a temperatura média, destacando a importância desses dados como *input* no processo de previsão para as metodologias apresentadas. No entanto, a análise dos dados também revelou que as cargas provenientes exclusivamente de usuários livres não apresentam uma correlação com a temperatura ou umidade, uma vez que não são influenciadas pelas condições meteorológicas em seus requerimentos de carga. Como resultado dessa análise, esse tipo de usuários não levaram em consideração os *inputs* meteorológicos nas previsões.

Figura 18 - Correlação da Potência de Pico vs. Temperatura média (2017-2019)



Fonte: Elaborado pelo próprio autor.

A segunda característica relevante direciona-se à operação dos sistemas elétricos de potência, que sempre buscam o equilíbrio entre oferta e demanda. Contudo, ocorrências de contingências e operações atípicas podem ter um impacto negativo e até mesmo modificar completamente o comportamento das curvas de cargas elétricas, resultando em grandes desvios, conforme mostrado na Figura 19. Lidar com essas situações requer uma modelagem matemática baseada em variações de energia diárias, permitindo identificar anomalias nas curvas de carga. Esses dias anormais são posteriormente rotulados e incorporados como *inputs* nas metodologias analisadas, a fim de garantir alta precisão nos resultados.

No intuito de identificar variações significativas em grandes blocos de energia, propõe-se uma modelagem matemática baseada na análise de desvios diários, nesse sentido a energia presente será comparada com a média das energias passadas e futuras, a média é realizada conforme à equação (112). Esse cálculo é aplicado aos seis níveis de desagregação abordados na tese, portanto os desvios de energia serão determinados por meio da relação (113):

$$EM(d) = \frac{1}{2r} \left(\sum_{n=-r, r \neq 0}^r E(d - 7n) \right) \quad (112)$$

$$\Delta Energia(d) = E(d) - EM(d) \quad (113)$$

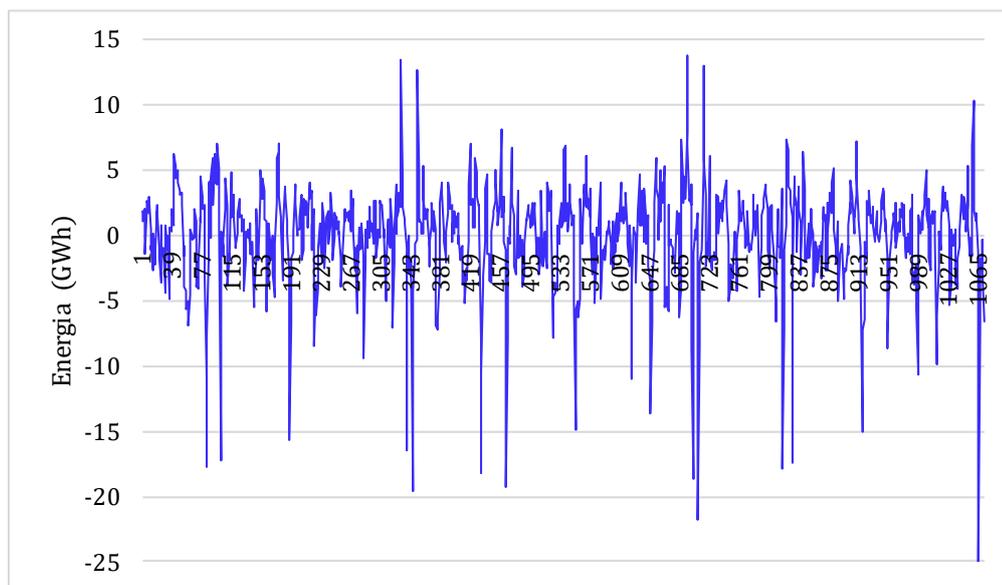
sendo:

$E(d)$: Energia diária, somatório de cargas no dia d .

$EM(d)$: Energia média de referência pertencente ao dia d .

r : Intervalo que abrange à média da energia diária, para esta tese o valor será 2.

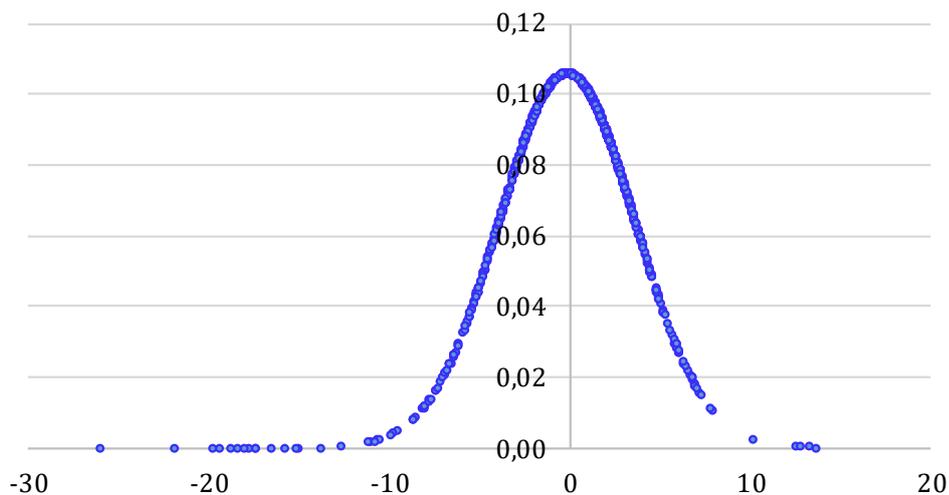
Figura 19 - Variações de Energia Sistema Global SEIN



Fonte: Elaborado pelo próprio autor.

Aplica-se a relação (113) aos seis níveis de desagregação de cargas elétricas apresentadas na subseção 6.6.1, isto no intuito de identificar as variações diárias de energia envolvidas na operação do sistema elétrico Peruano. Após identificado essas variações, é estabelecido uma análise de dados baseado na distribuição normal para cada nível de desagregação exemplificadas na Figura 19. O critério escolhido para a seleção dos dias de maiores erros, erros atípicos, baseia-se em um processo empírico que satisfaz o critério de ativação dos dias festivos no Peru. Além disso, observou-se também nesse processo a ativação de dias na qual aconteceram problemas fortuitos tanto nos desligamentos de linhas de transmissão, quanto problemas de suprimento elétrico. A estatística que permitiu essa validação no processo de ativação foram, como processo de dias típicos, o percentil do intervalo de 5% até 95%, conforme apresentado na Figura 20, enquanto valores fora desse intervalo foram considerados atípicos. Sem contravensão ao estabelecido na literatura sobre o processo de filtro de *outliers* conhecida como regra de Tukey na análise de dados (TUKEY, 1997), a proposta mostrada neste apartado, sobre a análise empírica, resgata efetivamente os fenômenos diários sem perder a natureza dos dados. Portanto, essa abordagem permite identificar e rotular padrões de entrada atípicos, como dias feriados e faltas fortuitas no sistema, atribuindo-lhe um valor 1, e dias típicos, atribuindo-lhe um valor 0, a qual dará um maior entendimento no processo de treinamento nas metodologias propostas.

Figura 20 - Distribuição normal das variações de energia SEIN



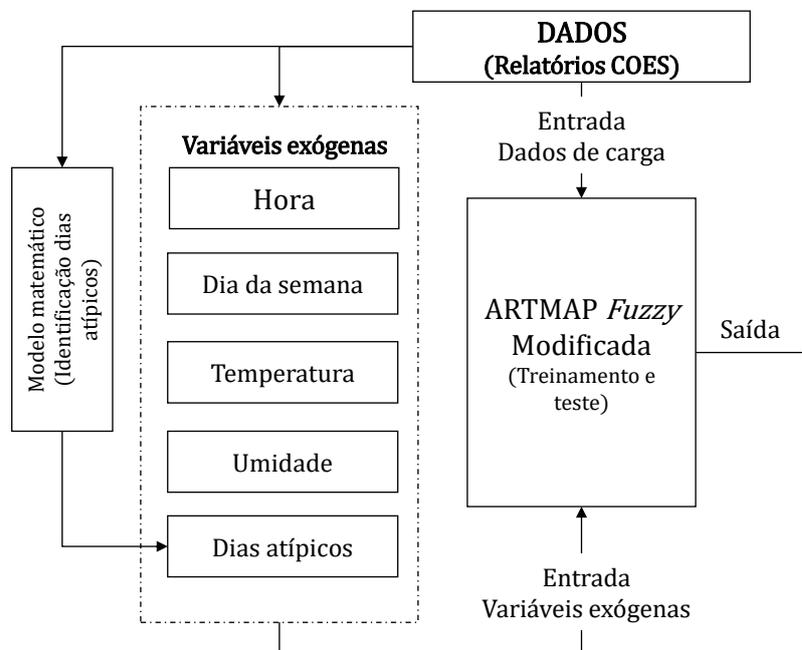
Fonte: Elaborado pelo próprio autor.

6.1.3 Metodologia SVM_ARTMAP Fuzzy

Nesta subseção, apresenta-se o processo detalhado da metodologia SVM_ARTMAP Fuzzy, que consiste em uma combinação entre a RNA ARTMAP Fuzzy Modificada (MORENO, 2010) e o recurso de regressão das Máquinas de Vetor de Suporte (SVM).

Inicialmente, os dados são organizados tendo em consideração a preparação de dados disposta na seção 6.1.1 e incluindo o processo de rotulado relacionados aos dias atípicos mostrados na seção 6.1.2. Essas informações são inseridas na RNA ARTMAP Fuzzy modificada, conforme ilustra-se na Figura 21.

Figura 21 - RNA ARTMAP Fuzzy modificada



Fonte: Elaborado pelo próprio autor.

Na estrutura interna, especificamente na camada de entrada da RNA, os dados são pré-processados e normalizados para remover quaisquer inconsistências e garantir que os valores estejam no mesmo intervalo. Cabe indicar que este processo ajuda evitar a proliferação de categorias que poderia acontecer no sistema (CARPENTER *et al.*, 1992).

Após organizado os dados de entrada e efetuado o passo de pré-processamento no sistema predictor ARTMAP Fuzzy modificado, passa-se à fase de treinamento. Deve-se ressaltar que o módulo ART_a é alimentado pelo vetor entrada \mathbf{a} , e o módulo ART_b pelo vetor saída \mathbf{b} , na qual a saída \mathbf{b} está estritamente relacionada ao prognóstico de \mathbf{a} .

Durante a fase de treinamento, os dados referentes aos padrões de entrada e de saída são apresentados à RNA ARTMAP *Fuzzy*, seguindo o algoritmo proposto por Moreno (2010). O módulo ART_b identifica a categoria ativa em \mathbf{w}^b mais próxima ao padrão de saída esperada, utilizando a função de escolha T_k e teste de similaridade com o parâmetro de vigilância ρ_b . Caso a categoria já existir, os pesos correspondentes em \mathbf{w}^b serão adaptados, caso contrário, uma nova categoria será criada.

No módulo ART_a o mesmo processo ocorre: a procura da categoria em \mathbf{w}_a mais próxima da entrada esperada. Assim que uma categoria é escolhida, o processo de correspondência entre os módulos ART_a e ART_b é realizado por meio do *Match Tracking* no módulo Inter-ART.

No processo de *Match Tracking* no sistema ARTMAP *Fuzzy* modificado, o objetivo é garantir a escolha de uma categoria ART_a diferente, caso não haja correspondência entre as categorias ativas ART_a e ART_b, utilizando o parâmetro de vigilância ρ_{ab} . Se a categoria escolhida em ART_a for confirmada pelo *Match Tracking*, os pesos \mathbf{w}_a e \mathbf{w}_{ab} são adaptados. Caso contrário, o parâmetro ρ_a será incrementado em um *mismatch*, e o processo será repetido procurando por uma nova categoria no módulo ART_a até que a correspondência seja confirmada. Se nenhuma categoria for aceita, uma nova será criada em \mathbf{w}_a e \mathbf{w}_{ab} .

O parâmetro de vigilância desempenha um papel fundamental no processo de treinamento dos módulos ART *Fuzzy* e Inter-ART, permitindo a busca por características de similaridade por meio de hiper-retângulos e atingindo a ressonância. Estabelecer um limite aceitável para o parâmetro de vigilância é crucial para satisfazer o critério de similaridade e, assim, adaptar os pesos. No entanto, ao utilizar parâmetros de vigilância menor do limite estabelecido como máximo, é possível obter resultados ainda melhores.

Esta premissa foi a motivação para a proposta de Alves (2019), que consiste em estabelecer um limite no parâmetro de vigilância durante o processo de diagnóstico, permitindo a identificação dos índices das características de similaridade maiores que esse limite e, em seguida, localizá-los em \mathbf{w}_a e \mathbf{w}_b . Dessa forma, é possível construir uma matriz baseada nessas informações que será utilizada posteriormente em uma regressão utilizando o método dos mínimos quadrados. Com o objetivo de encontrar categorias mais adequadas, permitindo a inclusão de similaridades que excedam o limite estabelecido no parâmetro de vigilância durante o teste, Alves (2019) propôs o método híbrido ARTMAP *Fuzzy*-MMQ, que, nesta pesquisa, será denominado como MMQ_ARTMAP *Fuzzy* por questões de praticidade.

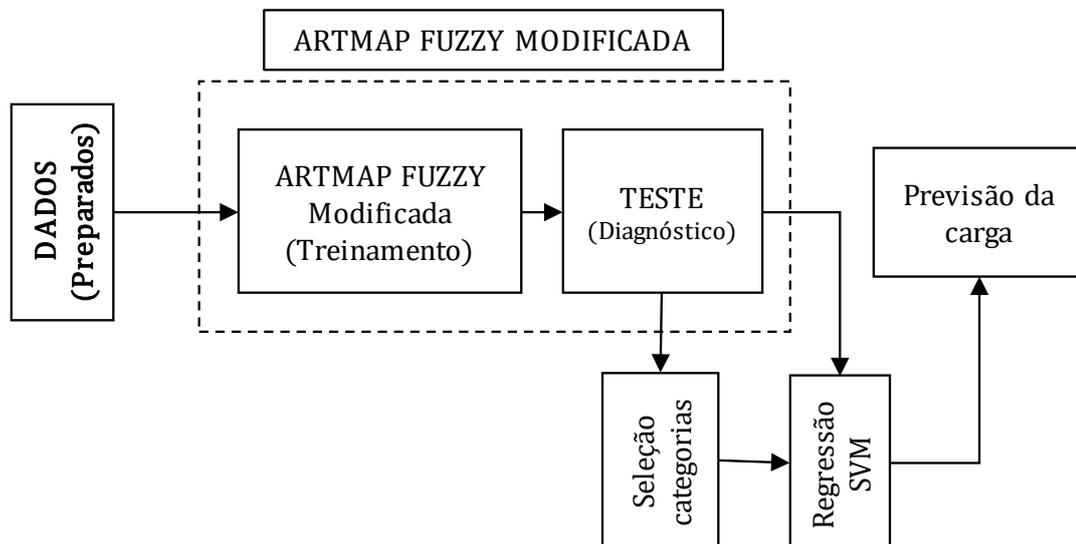
Nessa proposta, utiliza-se a RNA ARTMAP *Fuzzy* modificada para realizar o treinamento e adaptação dos pesos. Na etapa de diagnóstico, estabelece-se um limite no

parâmetro de vigilância ρ_a , com o objetivo de armazenar as categorias de similaridade acima desse limite. Esses dados armazenados são, então, submetidos a uma regressão linear por meio do MMQ. Em seguida, os vetores de entrada são avaliados para prever a respectiva saída ou previsão com base nos parâmetros encontrados na regressão.

Observa-se no trabalho de Alves (2019) que os vetores que são submetidos ao método MMQ possuem características multidimensionais e valores não lineares. Diante disso, propõe-se, nesta pesquisa, o uso da regressão SVM para buscar um hiperplano de separação ótimo entre essas categorias, por meio da adição de dimensões extras na análise por meio do *Kernel* Linear e Gaussiano. Além disso, será utilizada a ferramenta de otimização do toolbox SVR do Matlab 2022 para buscar um melhor hiperplano de separação, ajustando os parâmetros internos e fornecendo resultados mais precisos na regressão, o que contribuirá para atingir os objetivos da tese.

Assim, na proposta, o algoritmo passará a utilizar o recurso de treinamento da RNA ARTMAP *Fuzzy* modificada e a regressão do SVM, que será denominada SVM_ARTMAP *Fuzzy*, conforme ilustrado na Figura 22. O processo desta proposta pode ser descrito em quatro etapas, e está representado em um fluxograma na Figura 23.

Figura 22 - Metodologia 1 utilizando SVM_ARTMAP *Fuzzy*



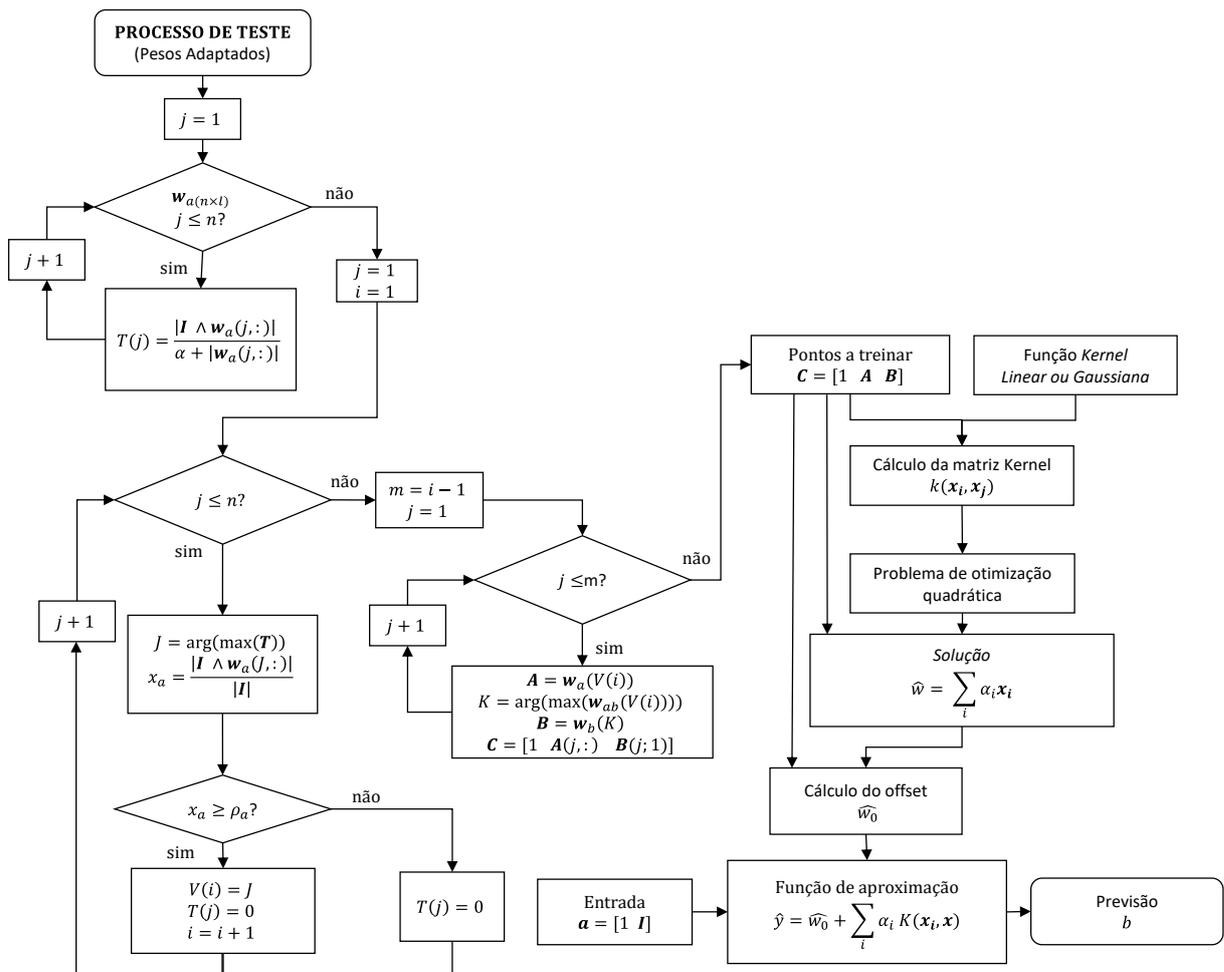
Fonte: Elaborado pelo próprio autor.

- Passo 1: O primeiro passo consiste no treinamento da base de dados utilizando a RNA ARTMAP *Fuzzy* modificada apresentada na Figura 21, conforme estabelecido nas seções 6.1.1 e 6.1.2.
- Passo 2: O segundo passo envolve a adaptação da RNA ARTMAP *Fuzzy* modificada na fase de diagnóstico. A função de escolha avalia a correspondência entre o primeiro vetor de entrada testado com a matriz de pesos adaptados em ART_a , e os valores obtidos são organizados e armazenados. Em um *loop*, cada valor da função escolha T_j é avaliado. A similaridade é usada para escolher o índice com o maior valor, representado pela posição da função de escolha. Se o valor do corte estabelecido pelo operador é atingido, o índice é armazenado em um vetor V , e a função de escolha é apagada para que o próximo *loop* possa identificar outro valor máximo. Esse processo é repetido até que todos os índices sejam identificados.
- Passo 3: O terceiro passo refere-se à construção dos vetores de entrada para o modelo de regressão SVM, que utiliza esses dados para otimizar o melhor hiperplano que se adapte aos dados. Os índices armazenados no vetor V são usados para construir o referido vetor, alocando cada índice na respectiva linha do vetor adaptado em w_a . Em seguida, um vetor temporal K é criado para armazenar a coluna do vetor w_{ab} do neurônio ativado na linha do índice do vetor V . Com isso, a linha K do vetor w_b é identificada e armazenada em B . Para concluir esse passo, A e B são armazenados em um vetor C , que será usado para treinar o modelo de regressão SVM em busca do melhor hiperplano de separação que possa representar melhor os dados.
- Passo 4: No quarto passo é realizado o treinamento do modelo de regressão SVM usando os dados armazenados no vetor C . Para ajustar os parâmetros de controle, como a regularização C e o erro épsilon ξ , é utilizado a toolbox de regressão do Matlab, que ajusta automaticamente esses parâmetros com base nos dados de entrada. Na otimização da regressão SVM, é utilizada a validação cruzada com k igual a 5 para cada iteração das 30 configuradas pela toolbox, buscando os melhores parâmetros internos e atingindo o melhor hiperplano. As restrições da caixa, a escala do *kernel* e épsilon são otimizados nessas 30 iterações. A regressão fornece uma função aproximada dos valores treinados, baseada na busca ótima de hiperplanos de corte em uma dimensão adicional, graças ao *kernel* linear, e Gaussiano e a um ajuste dos parâmetros internos no processo de otimização. Após a obtenção da função

aproximada de regressão, a primeira entrada do vetor de teste, que mantém as características dos vetores de entrada no processo de treinamento, é avaliada na função aproximada de regressão, que fornece a previsão para essa característica de entrada.

Finalmente, ao aplicar o modelo ARTMAP *Fuzzy* modificado (MORENO, 2010) nos diferentes níveis de desagregação, busca-se oportunizar a categorização na fase de diagnóstico da rede, identificando índices que atendam ao grau de similaridade maior que o parâmetro de vigilância estabelecido no pelo operador. Para isso, é realizada uma regressão utilizando o método de Máquina de Suporte Vetorial, aplicando os *kernels* Linear, Gaussiana e otimizando os parâmetros internos. A função de regressão desse vetor de categorias criadas é utilizada para prever os valores futuros das cargas visando uma melhor precisão nos resultados.

Figura 23 - Fluxograma do modelo híbrido SVM_ARTMAP *Fuzzy*



Fonte: Elaborado pelo próprio autor.

6.2 METODOLOGIA 2

Nesta metodologia propõe-se a utilização da decomposição de dados históricos das cargas elétricas por meio de um filtro *Wavelet* da família *Daubechies* db20, realizada em três níveis, para obtenção de dados de aproximação e de detalhe. Para tal fim, serão utilizados dois modelos de aprendizado de máquina: ARTMAP *Fuzzy* modificado (MORENO, 2010) para o processamento dos dados de aproximação e MMQ_ARTMAP *Fuzzy* (ALVES, 2019) para os dados de detalhe.

Os dados coletados são armazenados em vetores que serão as entradas para os modelos estabelecidos. O treinamento é realizado, em paralelo, com a parte de aproximação dos dados sendo treinada pelo ARTMAP *Fuzzy* modificado e a parte dos detalhes sendo treinada pelo MMQ_ARTMAP *Fuzzy*. Para os dados de aproximação, no terceiro nível de decomposição, as entradas são preparadas de acordo com a seção 6.1 na Figura 21. Porém, para os dados de detalhe, resultados da soma aritmética pertencentes ao primeiro, segundo e terceiro nível de decomposição, as entradas são realizadas com valores exógenos reduzidos, utilizando a MMQ_ARTMAP *Fuzzy* como método de previsão, as que serão explicadas a seguir.

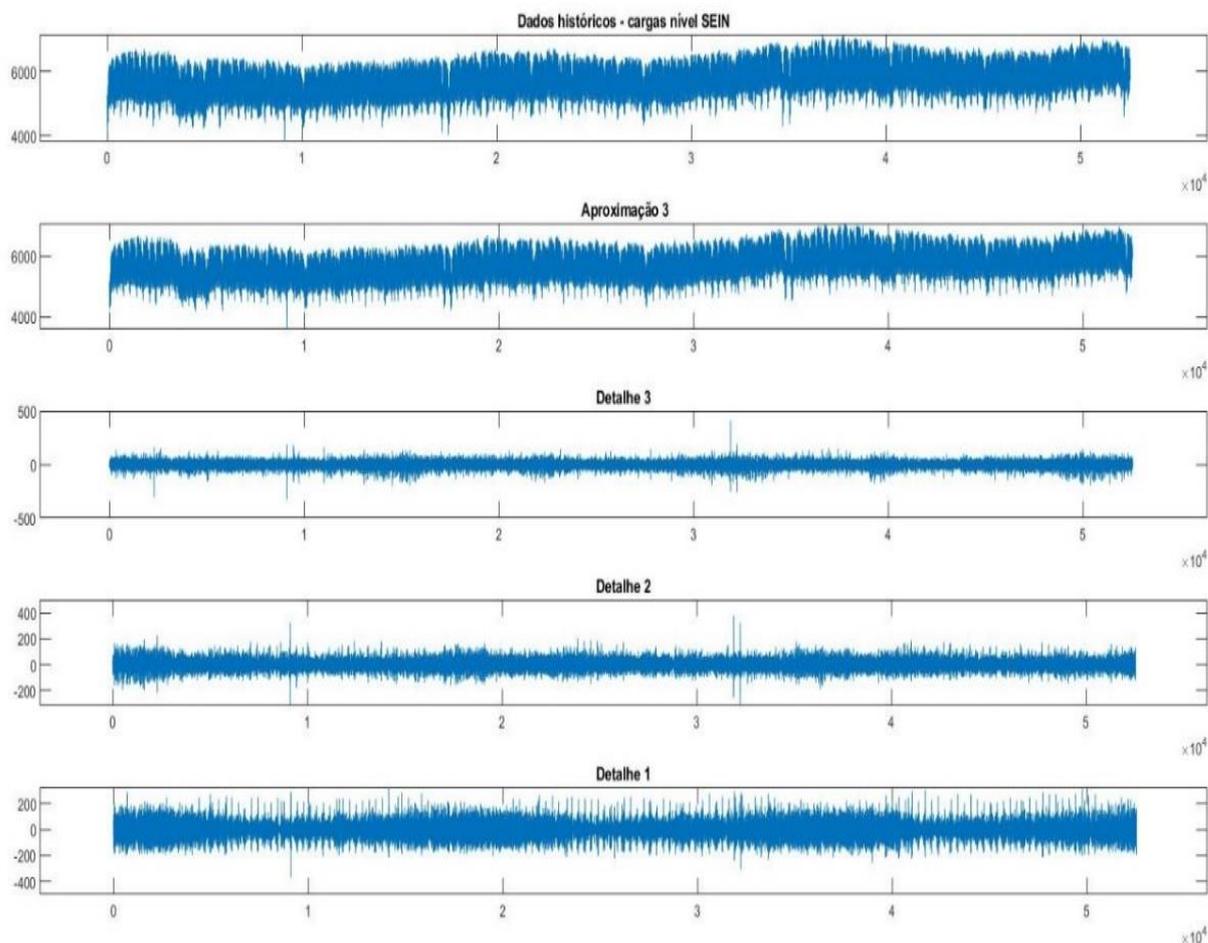
O processo completo desta metodologia será detalhado nas próximas subseções.

6.2.1 Decomposição da base de dados

Para realizar a análise multirresolução dos dados históricos de carga, utilizou-se a transformada *Wavelet* da família *Daubechies* db20 até o terceiro nível de decomposição por meio da toolbox *Wavelet* de Matlab. A escolha da *Daubechies* se deu em virtude de sua propriedade para formar uma base ortonormal, possuir suporte compacto e ter baixo custo computacional. A ortonormalidade garante independência nos diferentes níveis de decomposição, enquanto a forma compacta garante a maior concentração de energia em db^k para $k > 2$, como destacado por SHENG (1996). Além disso, a transformada *Wavelet* permite a análise do sinal via multirresolução, sendo um fator relevante para esta metodologia.

Na decomposição até o terceiro nível, processo da análise multirresolução, foram obtidos coeficientes de aproximação e de detalhe. Após à análise foram reconstruídos utilizando a função "wrcoef" do toolbox *Wavelet*, como mostrado na Figura 24. Esses coeficientes reconstruídos, de aproximação e detalhe na mesma escala, são armazenados em vetores para após serem utilizados como entradas nos modelos propostos nesta metodologia. É importante destacar que esse processo é realizado para os seis níveis de desagregação propostos nesta tese.

Figura 24 - Análise Multirresolução utilizando filtro *Wavelet* db 20 - carga nível SEIN



Fonte: Elaborado pelo próprio autor.

6.2.2 Processo da metodologia 2 - híbrida

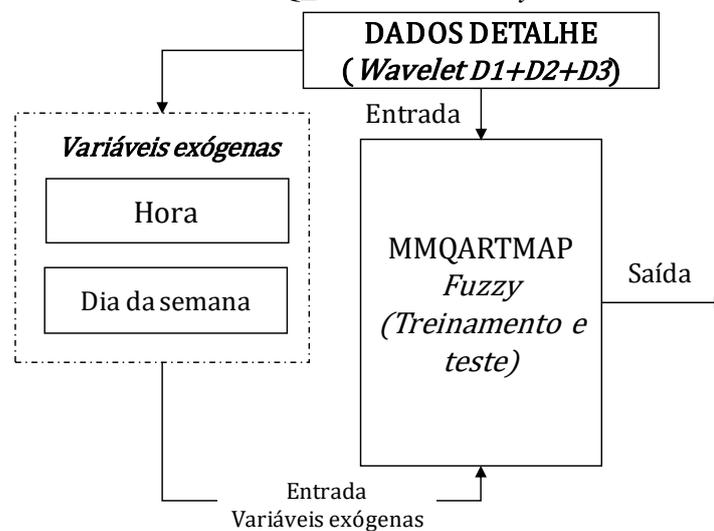
Nesta subseção, será apresentado o processo detalhado da metodologia 2, que consiste em uma abordagem híbrida que combina a RNA ARTMAP *Fuzzy* modificada com MMQ_ARTMAP *Fuzzy* para resolver o problema de previsão de cargas em vários níveis de desagregação. O desenvolvimento dessa metodologia será explicado nos seguintes passos a seguir.

- Passo 1: Na etapa de decomposição dos dados, é utilizado a toolbox *Wavelet* do Matlab para aplicar a análise multirresolução nos diferentes níveis de desagregação da carga, como apresentado na Figura 24.
- Passo 2: No terceiro nível de decomposição, os dados de aproximação, que representam a suavização dos dados sem *outliers*, e os dados de detalhe *outliers* em cada nível são reconstruídos utilizando o toolbox *Wavelet*. Nesse sentido, é possível obter os valores

de aproximação e a soma dos detalhes 1, 2 e 3 completamente reconstruídos no nível original. Em seguida, os valores de aproximação e de detalhe são armazenados em vetores para serem utilizados como entrada nos modelos.

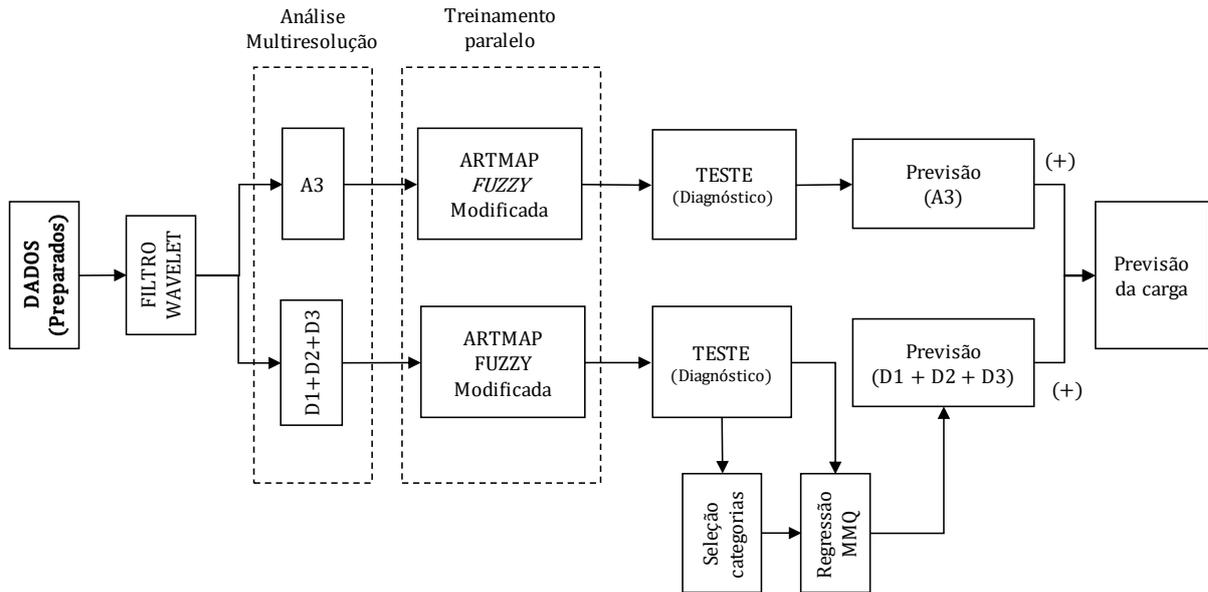
- Passo 3: Realiza-se um treinamento em paralelo, como mostrado na Figura 26, utilizando os vetores obtidos no passo 2. Os dados de aproximação são inseridos no modelo ARTMAP *Fuzzy* modificado, preservando os valores exógenos conforme apresentado na Figura 21. Paralelamente, os dados de detalhe são treinados com o modelo MMQ_ARTMAP *Fuzzy*, seguindo a metodologia proposta por Alves (2019) e apresentada na Figura 25.
- Passo 4: Após o processo de treinamento é realizado o diagnóstico, isso para prever os valores futuros dos modelos em paralelo. A previsão será uma superposição das previsões obtidas pelos modelos no processo de teste em paralelo.

Figura 25 - Processamento dos dados *Wavelet* de detalhe utilizando o modelo MMQ_ARTMAP *Fuzzy*



Fonte: Elaborado pelo próprio autor.

Figura 26 – Estrutura da metodologia 2 composta pelo modelo *WPT-MMQ_ARTMAP Fuzzy*



Fonte: Elaborado pelo próprio autor.

6.3 METODOLOGIA 3

Nesta metodologia propõe-se a utilização da decomposição de dados históricos das cargas elétricas por meio de um filtro *Wavelet* da família *Daubechies db20*, realizada em três níveis de decomposição, para obtenção de dados de aproximação e de detalhe, segundo a subseção 6.2.1. Para tal fim, serão utilizados dois modelos de aprendizado de máquina: o modelo *ARTMAP Fuzzy* modificado (MORENO, 2010) para a parte de aproximação dos dados e o modelo *SVM_ARTMAP Fuzzy*, baseada no processo da primeira metodologia, para os dados de detalhe.

Os dados coletados são armazenados em vetores que serão as entradas para os modelos estabelecidos. O treinamento é realizado, em paralelo, com a parte de aproximação dos dados sendo treinada pela *ARTMAP Fuzzy* modificada e a parte dos detalhes sendo treinada pela *SVM_ARTMAP Fuzzy*. No que tange os dados de aproximação, mais especificamente, A3, as entradas são preparadas de acordo com a seção 6.1 disposta na Figura 21. Já para os dados de detalhe, D1, D2 e D3, as entradas são realizadas de acordo com a Figura 27, utilizando a *SVM_ARTMAP Fuzzy* como método de previsão. O processo completo desta metodologia será detalhado na próxima subseção.

6.3.1 Processo da metodologia 3 - híbrida

Nesta subseção, será apresentado o processo detalhado da metodologia 3, uma abordagem híbrida que combina as técnicas de RNA ARTMAP *Fuzzy* modificada e SVM_ARTMAP *Fuzzy* para resolver o problema de previsão de cargas em diferentes níveis de desagregação. O desenvolvimento desta metodologia será explicado nos passos a seguir.

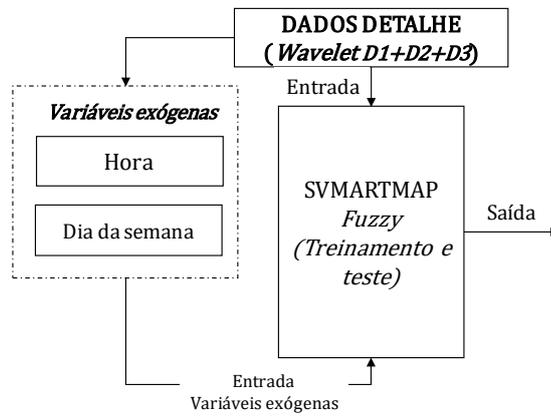
Passo 1: Durante a etapa de decomposição dos dados, utilizamos a toolbox *Wavelet* do Matlab para aplicar a análise multirresolução nos diferentes níveis de desagregação da carga. Essa etapa é ilustrada na Figura 24.

Passo 2: No terceiro nível de decomposição, os dados de aproximação - que representam a suavização dos dados sem *outliers* - e os dados de detalhe em cada nível são reconstruídos utilizando a toolbox *Wavelet*. Dessa forma, é possível obter os valores de aproximação e a soma dos detalhes 1, 2 e 3 completamente reconstruídos no nível original. Em seguida, os valores de aproximação e de detalhe são armazenados em vetores para serem utilizados como entrada nos modelos propostos.

Passo 3: Realiza-se, então, um treinamento em paralelo, conforme apresentado na Figura 28, utilizando os vetores obtidos no passo 2 como entrada. Os dados de aproximação são inseridos no modelo ARTMAP *Fuzzy* modificado, preservando os valores exógenos conforme apresentado na Figura 21. Já os dados de detalhe são treinados com o modelo SVM_ARTMAP *Fuzzy*, seguindo o processo da metodologia 1 e conforme a Figura 27.

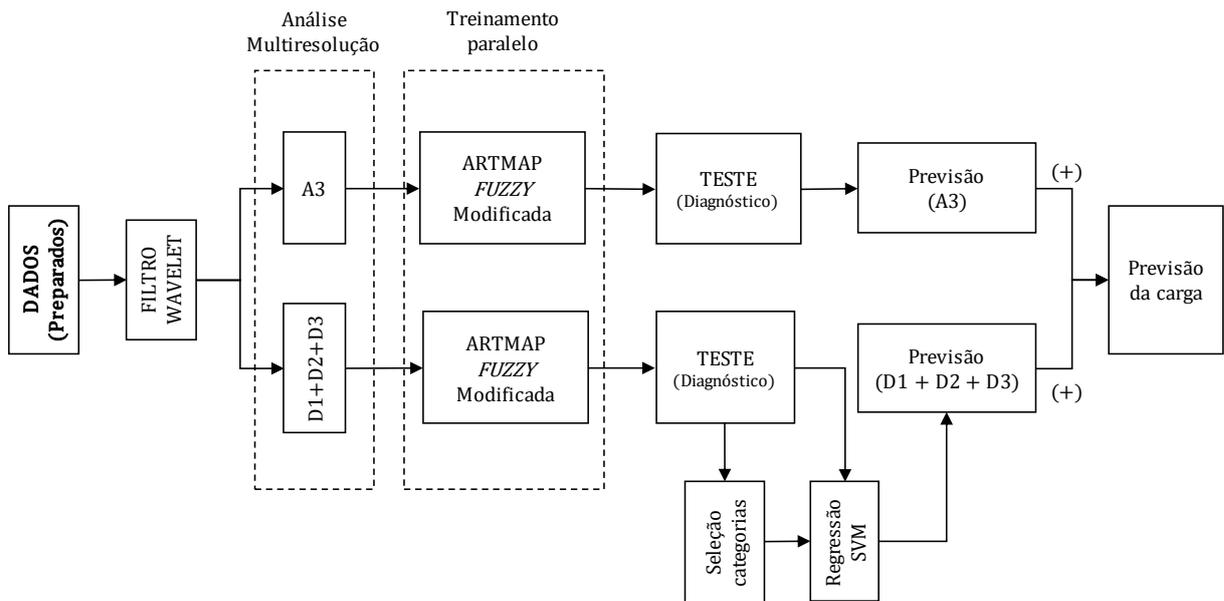
Passo 4: Após o processo de treinamento é realizado o diagnóstico, isso para prever os valores futuros dos modelos em paralelo. A previsão final será uma combinação das previsões obtidas pelos modelos durante o processo de teste em paralelo.

Figura 27 - Processamento dos dados *Wavelet* de detalhe utilizando o modelo SVM_ARTMAP *Fuzzy*



Fonte: Elaborado pelo próprio autor.

Figura 28 - Estrutura da metodologia 3 composta pelo modelo *WPT-SVM_ARTMAP Fuzzy*



Fonte: Elaborado pelo próprio autor.

7 RESULTADOS

Na busca das melhores soluções em relação às métricas estabelecidas nesta pesquisa, adotou-se o MAPE como uma medida para comprar a precisão das previsões. Portanto, para alcançar essa precisão, avaliou-se três metodologias detalhadas no capítulo 6, além de dois métodos como a RNA ARTMAP *Fuzzy* modificada e o modelo híbrido MMQ_ARTMAP *Fuzzy*, que foram escolhidas como *benchmarks* 1 e 2, respectivamente, neste estudo.

Nesse sentido, no intuito de encontrar o modelo predictor flexível e determinar o melhor parâmetro de vigilância que se ajuste às metodologias mencionadas, utilizou-se a validação cruzada com k igual a 5 para a RNA ARTMAP *Fuzzy* modificada, diante dos dados de cargas desagregadas, considerando um conjunto de 4032 dados históricos de cargas, correspondente ao período de 20 maio de 2019 a 11 de agosto de 2019.

Quadro 2 - Informações dos métodos utilizados para as previsões das cargas elétricas

Metodologias	Métodos de ML	Previsão do dia seguinte			Dados Treinados
		Dia	Mês	Ano	
Metodologia 1	SVM_ARTMAP <i>Fuzzy</i> (SVMFAM, Linear/Gaussiana/Otimizada)	12 até 18	Agosto	2019	4032
Metodologia 2	WPT-MMQ_ARTMAP <i>Fuzzy</i> (WPTMMQFAM)				
Metodologia 3	WPT-SVM_ARTMAP <i>Fuzzy</i> (WPTSVMFAM Linear/Gaussiana/Otimizada)				
(MORENO, 2010)	ARTMAP <i>Fuzzy</i> modificada (FAM)				
(ALVES, 2019)	MMQ_ARTMAP <i>Fuzzy</i> (MMQFAM)				

Fonte: Elaborado pelo próprio autor.

Após a validação cruzada, foram selecionados os melhores parâmetros de vigilância e avaliados nas três metodologias propostas e nas duas adicionais mencionadas como *benchmark*, com o intuito de garantir um maior desempenho em termos de precisão. Nesse processo, também, teve-se em consideração os seguintes parâmetros internos para o treinamento da RNA ARTMAP *Fuzzy* modificada, os quais foram aplicadas à metodologias propostas e apresentadas no Quadro 3. Todas as simulações foram realizadas em um computador equipado com

processador Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz 2.00 GHz e 8.00 GB de memória RAM, utilizando o software Matlab (com Toolbox *Wavelet* e SVM).

Quadro 3 - Configuração dos parâmetros internos da RNA ARTMAP *Fuzzy* Modificada

ρ_b	ρ_{ab}	β	α	ε
0,99	0,99	1	0,05	0,001

Fonte: Elaborado pelo próprio autor.

Os dados históricos utilizados, neste estudo, foram obtidos a partir dos IEOB do SEIN, disponíveis de forma *open access* no site do COES. Para a realização das previsões de carga, foram utilizados seis níveis de desagregação: previsão de carga global do SEIN, área norte do SEIN, subárea sudoeste SEIN, grandes usuários livres, grandes usuários livres área centro e grandes usuários livres área sul. Em cada nível, foram aplicados os métodos de *Machine Learning* descritos no Quadro 2.

7.1 APLICAÇÃO DA VALIDAÇÃO CRUZADA

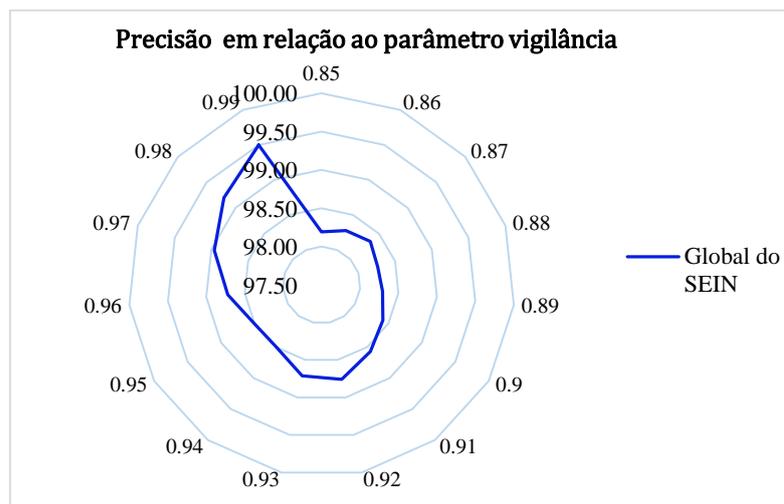
A validação cruzada é uma técnica amplamente utilizada na avaliação de modelos, que consiste em dividir o conjunto de dados em partes para que o modelo seja treinado em uma parte e avaliado em outra, repetindo esse processo várias vezes. Dessa forma, é possível obter uma avaliação mais robusta do desempenho do modelo e evitar problemas de *overfitting*.

Com o objetivo de garantir a capacidade de generalização do modelo ARTMAP *Fuzzy* modificada para determinados parâmetros de vigilância ρ_a , a validação cruzada com k igual 5 foi utilizada em 4032 dados históricos anteriores ao dia 12 de agosto de 2019, aplicados a cada um dos seis níveis de desagregação estabelecidos neste trabalho. Foi estabelecido um intervalo de avaliação do valor de vigilância adequado às características das cargas elétricas e selecionados os melhores parâmetros baseados no desempenho em termos de precisão. Esses parâmetros serão utilizados nas metodologias descritas no Quadro 2, visando encontrar aquela que melhor se adapte aos dados e atenda aos objetivos do estudo.

Nesta subseção, é apresentada uma análise comparativa da validação cruzada para os diversos níveis de desagregação, a fim de identificar os melhores parâmetros de vigilância ρ_a a serem utilizados no processo de treinamento das metodologias propostas. Isso permitirá garantir os melhores resultados na previsão da carga elétrica do dia seguinte, em cada um dos níveis de desagregação considerados.

Na Figura 29 apresenta-se a evolução do desempenho do modelo em função das variações do parâmetro de vigilância ρ_a . É possível observar que, no nível SEIN, a RNA ARTMAP *Fuzzy* apresenta uma boa precisão na validação para os parâmetros de vigilância ρ_a iguais a 0,92, 0,93, 0,97, 0,98 e 0,99.

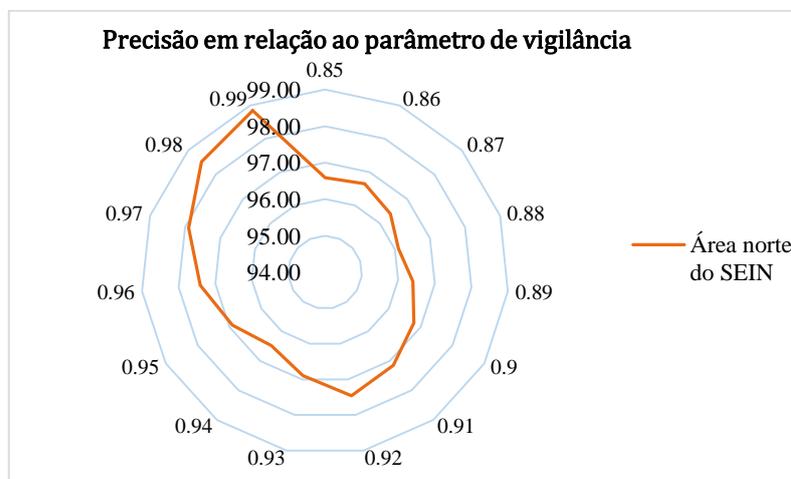
Figura 29 - Validação cruzada para o SEIN



Fonte: Elaborado pelo próprio autor.

A Figura 30 ilustra a evolução do desempenho do modelo em relação às variações do parâmetro de vigilância ρ_a , para o nível de desagregação correspondente à área norte do SEIN. É possível observar que a RNA ARTMAP *Fuzzy* modificada apresenta boa precisão na validação cruzada para os valores de ρ_a iguais a 0,92, 0,96, 0,97, 0,98 e 0,99.

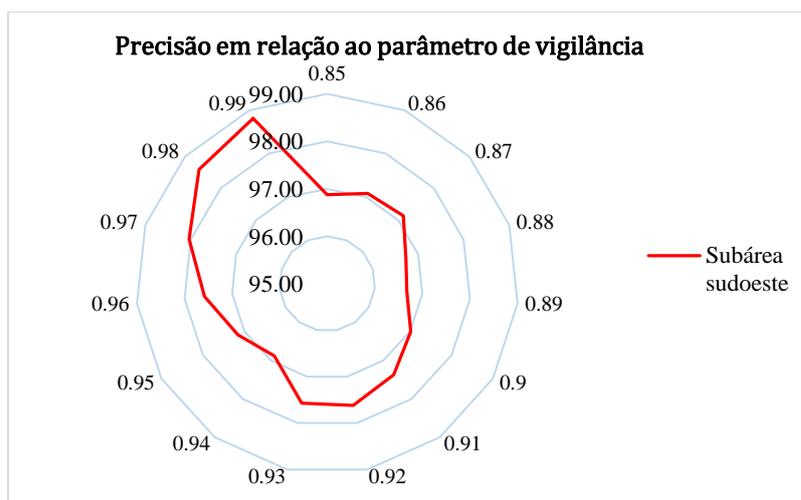
Figura 30 - Validação cruzada para a área norte



Fonte: Elaborado pelo próprio autor.

Durante o processo de validação cruzada para a subárea sudoeste do SEIN, conforme apresentado na Figura 31, foi possível observar que a RNA ARTMAP *Fuzzy* modificada apresentou bons resultados de precisão para os parâmetros de vigilância ρ_a iguais a 0,92, 0,93, 0,97, 0,98 e 0,99.

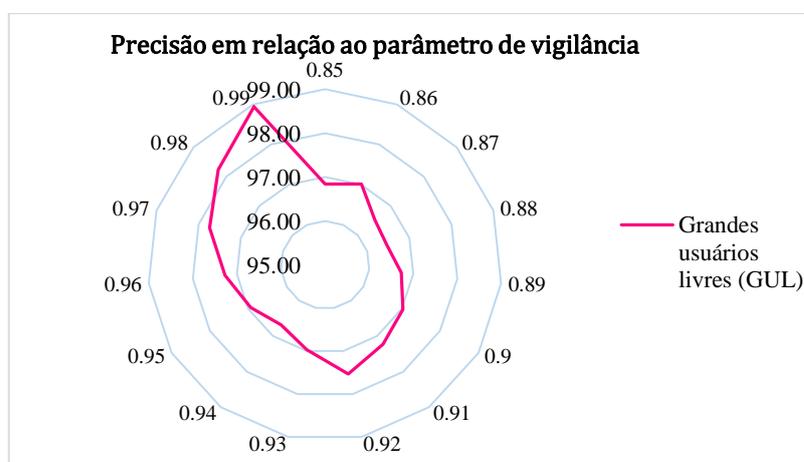
Figura 31 - Validação cruzada para a subárea sudoeste



Fonte: Elaborado pelo próprio autor.

Na Figura 32, apresenta-se a evolução do desempenho do modelo sob as variações do parâmetro de vigilância ρ_a para o nível grande usuário livre. Observa-se que a RNA ARTMAP *Fuzzy* modificada garante uma boa precisão na validação cruzada para os parâmetros de vigilância ρ_a iguais a 0,92, 0,96, 0,97, 0,98 e 0,99.

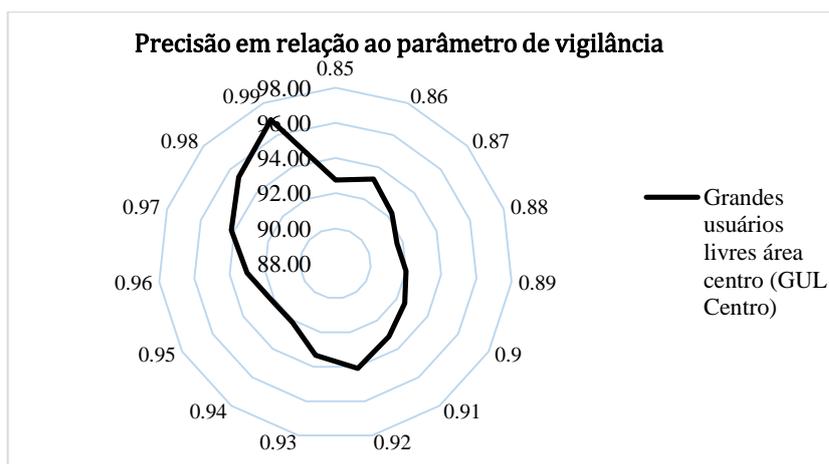
Figura 32 - Validação cruzada para grandes usuários livres



Fonte: Elaborado pelo próprio autor.

Na Figura 33, é apresentada a evolução do desempenho do modelo ARTMAP *Fuzzy* modificada para os grandes usuários livres na área central do SEIN, obtida a partir da validação cruzada. Os melhores resultados em termos de precisão foram obtidos para os parâmetros de vigilância ρ_a iguais a 0,92, 0,93, 0,97, 0,98 e 0,99.

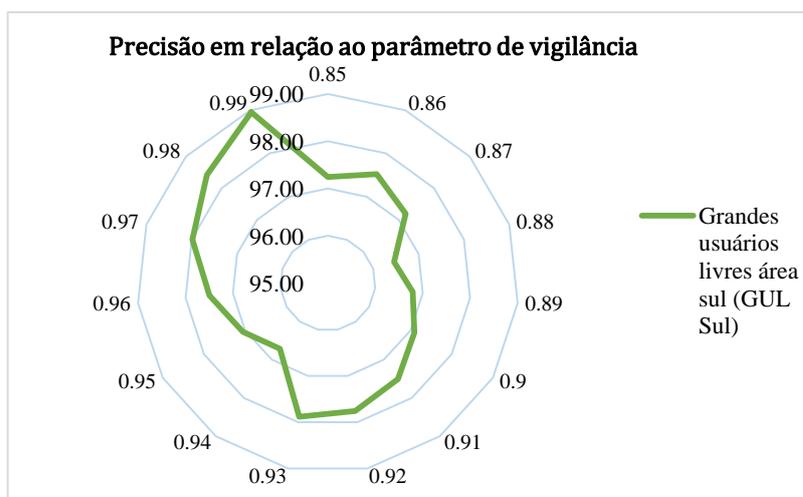
Figura 33 - Validação cruzada para grandes usuários livres área centro



Fonte: Elaborado pelo próprio autor.

No processo da validação cruzada para a subárea sudoeste, apresentado na Figura 34, observa-se que a RNA ARTMAP *Fuzzy* modificada alcança os melhores desempenhos em termos de precisão para os parâmetros de vigilância ρ_a iguais a 0,92, 0,93, 0,97, 0,98 e 0,99.

Figura 34 - Validação cruzada para grandes usuários livres área sul



Fonte: Elaborado pelo próprio autor.

Os parâmetros de vigilância encontrados nesta subseção serão aplicados no processo de treinamento das metodologias, com o objetivo de assegurar a procura dos melhores resultados

na previsão das cargas elétricas do dia seguinte para os diferentes níveis de desagregação propostos nesta tese.

7.2 RESULTADOS OBTIDOS DOS DIVERSOS NÍVEIS DE DESAGREGAÇÃO

Com base na análise dos resultados da validação cruzada, foram identificados parâmetros de vigilância que garantem altos desempenhos quando aplicados ao modelo ARTMAP *Fuzzy* modificado para cada nível de desagregação estabelecido no trabalho. Esses parâmetros foram utilizados no processo de treinamento para a previsão diária da semana de 12 a 18 de agosto de 2019. Os desempenhos das diferentes metodologias e configurações de parâmetros internos serão explicados a seguir.

A primeira metodologia foi treinada considerando os cinco melhores parâmetros de vigilância, ρ_a , estabelecidos pela validação cruzada. Para cada avaliação do diagnóstico, foram utilizados limites de corte do parâmetro de vigilância ρ_a no intervalo de 0,90 com intervalos de acréscimo de 0,01 até 0,99, a fim de criar a matriz que treine o modelo de regressão SVM.

Na segunda metodologia, foram utilizados os dados de aproximação para treinar a RNA ARTMAP *Fuzzy* modificada, também considerando os melhores parâmetros de vigilância, ρ_a . Os mesmos parâmetros usados no treinamento foram empregados para o diagnóstico. Além disso, no processo paralelo, foram utilizados os dados de detalhe no modelo MMQ_ARTMAP *Fuzzy*, para o qual foi estabelecido um parâmetro de vigilância ρ_a igual a 0,99 para o processo de treinamento. No diagnóstico foram estabelecidos limites de corte do parâmetro de vigilância no intervalo de 0,90 com intervalos de acréscimo de 0,01 até 0,99, a fim de criar a função de aproximação graças ao modelo de regressão MMQ.

Na terceira metodologia, foram utilizados os parâmetros estabelecidos na validação cruzada para treinar os dados de aproximação na ARTMAP *Fuzzy* modificada. No entanto, para o modelo paralelo que utiliza a SVM_ARTMAP *Fuzzy* foi estabelecido um parâmetro de vigilância ρ_a igual a 0,99 para o processo de treinamento. No diagnóstico foram estabelecidos limites de corte do parâmetro de vigilância no intervalo de 0,90 com intervalos de acréscimo de 0,01 até 0,99, a fim de criar a função de aproximação graças ao modelo de regressão SVM, utilizando os *kernels* linear, gaussiano e o processo otimizado.

Nos métodos de *benchmark* apresentados no Quadro 2 foram utilizados os cinco parâmetros de vigilância estabelecidos no processo de validação cruzada para cada nível de desagregação. Na fase de diagnóstico, o modelo ARTMAP *Fuzzy* modificada empregou os

respectivos parâmetros utilizados no treinamento. Contudo, para o MMQ_ARTMAP *Fuzzy*, foi utilizado um limite de corte do parâmetro de vigilância ρ_a de 0,90.

Os resultados da aplicação em cada nível de desagregação serão detalhados nas subseções a seguir. Para o processo de comparação no desempenho da previsão, utilizam-se as métricas de RMSE, Coeficiente de Correlação de Pearson (PEARSON, 1901) (PCC), MSE, MAE e, especialmente, o MAPE. Os resultados foram comparados para cada nível de desagregação, como apresentado nas Tabelas 1 a 6.

7.2.1 Previsão de carga global SEIN

Para comparar os resultados obtidos no nível de agregação SEIN, apresenta-se, na Tabela 1, o desempenho de cada metodologia. Cada modelo foi submetido a uma análise rigorosa dos parâmetros internos, visando identificar os melhores resultados possíveis. De modo geral, as metodologias propostas superaram significativamente os *benchmarks*, exceto no dia 17 de agosto, em que o *benchmark* 1 apresentou melhor generalização que o *benchmark* 2 e as metodologias propostas. Naquele dia, a melhor previsão entre as metodologias propostas foi alcançada pela metodologia 2. Apesar de manter uma diferença em relação ao MAPE, obteve um resultado de PCC comparável ao *benchmark* 1, fornecendo informações importantes sobre as características relevantes das curvas para a programação do despacho hidrotérmico, que evita problemas na operação e permite a utilização eficiente dos recursos energéticos.

Em geral, as metodologias 2 e 3 apresentam excelente desempenho em comparação com os *benchmarks* no nível de agregação global. No entanto, destaca-se que a metodologia 3, que utiliza o *kernel* Gaussiano com parâmetros de vigilância ρ_a de 0,98, para treinar os dados de aproximação, e 0,99 para os dados de detalhe, demonstrou um desempenho notável. Durante a fase de diagnóstico, um limite de 0,90 em ρ_a foi utilizado para o modelo de detalhe, o que gerou os melhores resultados em termos de generalização na metodologia. Como resultado, a metodologia 3 apresentou uma superioridade significativa em termos de MAPE e PCC na previsão das cargas elétricas no nível de agregação mencionado.

Tabela 1 - Resultados das metodologias (Previsão Global SEIN)

Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTSVMFAM Opt	WPTSVMFAM Linear	WPTSVMFAM Gauss
		0,93	0,98	0,98	0,97	0,98	0,92A/0,99D 0,91	0,98A/0,99D 0,91	0,98A/0,99D 0,90	0,98A/0,99D 0,90
12 de Agosto 2019	RMSE	110,60	97,98	97,35	119,97	95,53	98,87	104,67	105,86	103,65
	PCC	0,99	0,99	0,99	0,98	0,99	0,99	0,99	0,99	0,99
	MSE	12233,32	9600,33	9476,87	14393,20	9126,85	9775,94	10955,43	11205,51	10742,97
	MAE	83,51	77,92	76,26	95,10	75,06	72,56	82,76	83,54	79,79
	MAPE	1,51	1,36	1,33	1,68	1,31	1,31	1,45	1,46	1,41
13 de Agosto 2019	RMSE	107,88	154,57	149,33	118,76	145,61	107,38	108,25	105,30	97,86
	PCC	0,98	0,96	0,96	0,97	0,96	0,97	0,98	0,97	0,98
	MSE	11638,32	23893,32	22300,90	14104,65	21203,49	11530,71	11717,63	11087,34	9576,81
	MAE	85,07	118,00	111,56	97,48	116,85	85,29	87,80	87,10	79,31
	MAPE	1,42	1,94	1,83	1,63	1,91	1,43	1,47	1,45	1,32
14 de Agosto 2019	RMSE	136,09	143,85	136,74	153,55	139,35	128,90	117,46	117,57	122,27
	PCC	0,98	0,97	0,97	0,97	0,97	0,97	0,98	0,97	0,98
	MSE	18521,24	20693,31	18698,99	23576,54	19418,11	16614,40	13797,94	13822,64	14951,15
	MAE	112,73	117,12	110,13	130,79	105,30	102,47	92,08	93,14	99,78
	MAPE	1,88	1,98	1,86	2,20	1,78	1,72	1,54	1,56	1,68
15 de Agosto 2019	RMSE	94,03	109,42	110,32	97,38	108,28	105,83	85,49	85,23	84,25
	PCC	0,98	0,97	0,97	0,97	0,97	0,98	0,98	0,98	0,98
	MSE	8842,22	11972,21	12171,20	9483,57	11723,87	11199,87	7309,18	7264,90	7097,56
	MAE	72,96	85,52	87,09	73,34	82,26	87,50	70,56	69,84	67,16
	MAPE	1,20	1,40	1,43	1,20	1,35	1,46	1,18	1,17	1,12
16 de Agosto 2019	RMSE	104,56	92,56	95,17	112,15	94,60	71,35	59,31	59,78	63,47
	PCC	0,97	0,98	0,98	0,97	0,98	0,99	0,99	0,99	0,99
	MSE	10932,11	8566,66	9056,56	12576,66	8950,03	5090,62	3517,12	3573,13	4028,41
	MAE	85,36	76,59	78,28	90,14	81,00	54,60	49,42	49,69	50,64
	MAPE	1,41	1,26	1,29	1,49	1,34	0,89	0,81	0,81	0,82
17 de Agosto 2019	RMSE	71,80	156,16	158,99	150,68	152,41	106,04	123,88	126,01	129,44
	PCC	0,98	0,95	0,94	0,95	0,94	0,98	0,98	0,98	0,98
	MSE	5155,33	24385,94	25277,42	22705,23	23229,98	11245,53	15345,90	15878,37	16754,35
	MAE	53,22	131,39	133,60	124,40	126,15	85,67	107,59	109,89	111,61
	MAPE	0,89	2,19	2,23	2,08	2,11	1,44	1,78	1,82	1,84
18 de Agosto 2019	RMSE	94,34	106,51	107,27	91,91	99,63	93,20	98,71	95,76	94,52
	PCC	0,97	0,97	0,97	0,97	0,97	0,98	0,97	0,97	0,97
	MSE	8900,91	11345,20	11507,59	8448,00	9926,12	8686,20	9742,98	9170,90	8934,71
	MAE	76,60	87,30	88,10	71,20	79,63	79,34	79,52	76,35	76,91
	MAPE	1,39	1,60	1,61	1,30	1,45	1,44	1,46	1,40	1,41

Fonte: Elaborado pelo próprio autor.

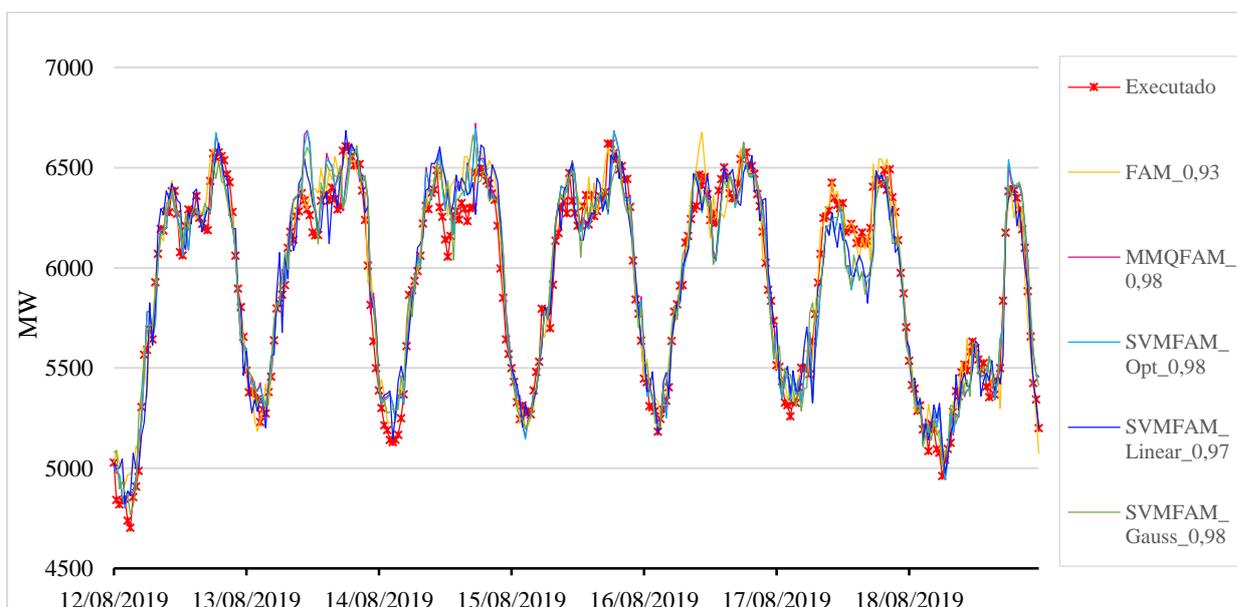
As curvas relativas aos resultados das previsões de carga diária são apresentadas nas Figura 35 e Figura 36, que foram divididas para uma melhor compreensão do comportamento em relação aos desempenhos apresentados na Tabela 1. Na Figura 35, observam-se as características da curva da demanda real do sistema elétrico comparadas com as curvas de previsão da metodologia 1 e dos *benchmarks* 1 e 2. Verifica-se que o *benchmark* 1 apresenta um desempenho aceitável em termos de MAPE e PCC, conseguindo acompanhar os valores reais das cargas elétricas. Nesse nível de agregação, somente a metodologia 1, que utiliza o

kernel Gaussiano, consegue superá-lo em termos de MAPE e PCC nos dias 12 e 14 de agosto. Porém, seu desempenho nos outros dias não é bom, o que faz com que seja deixada de lado sem ser considerada.

Por outro lado, na Figura 36, pode-se analisar as curvas de carga oriundas da metodologia 3, que apresentam um acompanhamento bastante aceitável, o que é refletido no PCC da Tabela 1, com exceção do dia 17 de agosto, no qual o *benchmark* conseguiu acompanhar notavelmente o valor executado da carga.

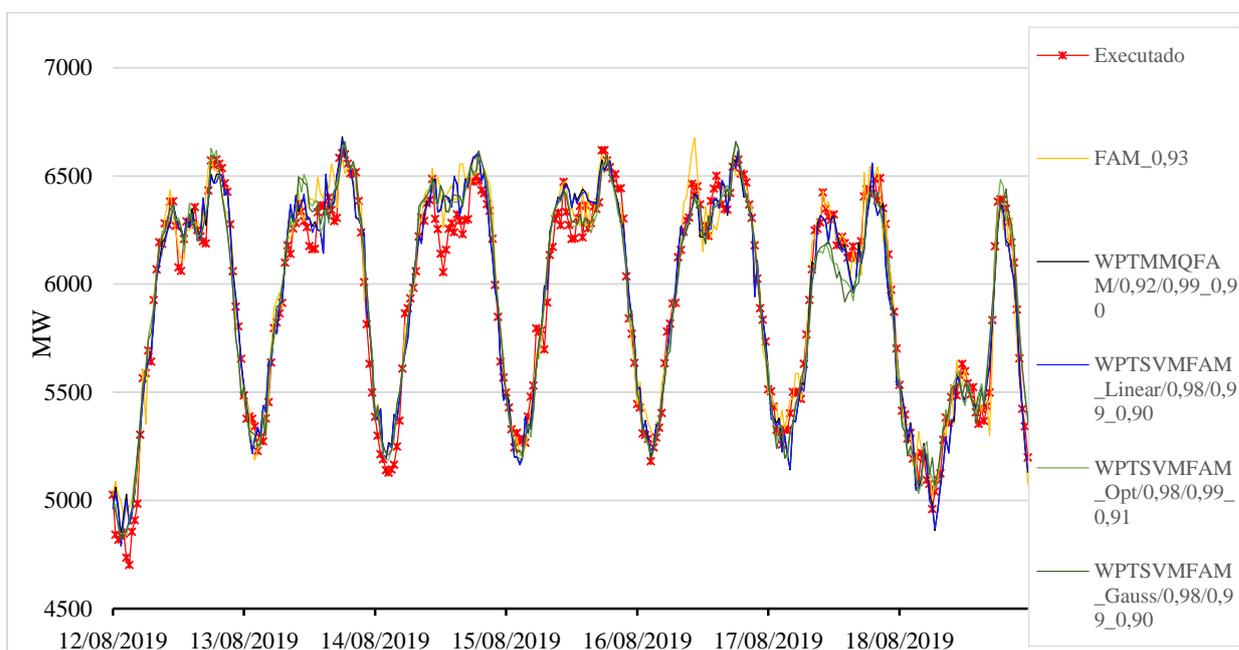
Por fim, utilizando os recursos da validação cruzada no processo de treinamento das metodologias, a metodologia 3, que utilizou o *kernel* Gaussiano, apresentou um desempenho notável na previsão da carga elétrica nesse nível de agregação global.

Figura 35 - Curvas da previsão da metodologia 1 versus *benchmark* 1 e 2 (Global SEIN)



Fonte: Elaborado pelo próprio autor.

Figura 36 - Curvas da previsão das metodologias 2 e 3 versus *benchmark* 1(Global SEIN)



Fonte: Elaborado pelo próprio autor.

7.2.2 Previsão de carga da área norte

No intuito de encontrar o melhor modelo predictor aplicado à área norte do SEIN, passou-se a comparar os resultados obtidos da aplicação das metodologias, incluindo-se os *benchmarks*, por meio da Tabela 2. Nesta, cada modelo foi submetido a uma análise exaustiva na procura dos melhores parâmetros internos utilizando a validação cruzada que consegue, desta forma, garantir bons níveis de generalização.

Tabela 2 - Resultados das metodologias (Previsão da área norte SEIN)

Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTSVMFAM Opt	WPTSVMFAM Linear	WPTSVMFAM Gauss
		0,96	0,96	0,96	0,96	0,96	0,99A/0,99D 0,90	0,99A/0,99D 0,94	0,99A/0,99D 0,90	0,99A/0,99D 0,90
12 de Agosto 2019	RMSE	38,34	40,32	42,88	42,13	35,98	23,60	23,91	24,71	20,22
	PCC	0,96	0,93	0,93	0,93	0,94	0,97	0,97	0,97	0,98
	MSE	1470,19	1625,58	1838,41	1774,79	1294,34	557,02	571,92	610,70	408,88
	MAE	31,35	33,23	34,74	34,10	28,30	18,73	18,76	19,43	15,98
	MAPE	3,93	4,12	4,31	4,23	3,56	2,40	2,39	2,47	2,05
13 de Agosto 2019	RMSE	45,68	45,13	49,22	49,69	46,95	46,86	45,89	46,77	46,79
	PCC	0,91	0,87	0,86	0,86	0,89	0,93	0,92	0,92	0,93
	MSE	2086,79	2036,37	2422,27	2468,84	2203,84	2195,53	2106,16	2187,40	2189,29
	MAE	36,38	33,28	35,65	36,19	36,47	38,51	37,42	38,21	38,25
	MAPE	4,26	3,86	4,11	4,17	4,21	4,43	4,30	4,39	4,40
14 de Agosto 2019	RMSE	31,51	40,10	41,04	42,31	38,65	40,71	38,85	39,21	36,54
	PCC	0,95	0,86	0,86	0,86	0,89	0,93	0,93	0,93	0,94
	MSE	992,78	1607,84	1683,96	1790,55	1493,49	1656,98	1509,60	1537,39	1335,07
	MAE	26,06	28,17	29,10	30,16	29,51	33,07	31,13	31,50	29,20
	MAPE	3,04	3,24	3,33	3,46	3,40	3,76	3,53	3,58	3,32
15 de Agosto 2019	RMSE	34,70	39,07	40,10	40,97	37,00	24,08	24,73	24,18	24,31
	PCC	0,95	0,90	0,90	0,90	0,89	0,95	0,95	0,95	0,95
	MSE	1203,99	1526,44	1608,35	1678,77	1369,34	579,94	611,64	584,73	590,76
	MAE	25,44	31,17	29,79	31,15	28,93	17,46	17,62	17,71	17,82
	MAPE	3,08	3,66	3,50	3,67	3,39	2,08	2,10	2,11	2,13
16 de Agosto 2019	RMSE	23,93	27,96	32,10	31,79	29,84	24,54	25,46	24,21	23,05
	PCC	0,96	0,93	0,93	0,93	0,94	0,96	0,95	0,95	0,96
	MSE	572,85	781,96	1030,24	1010,66	890,28	602,03	648,41	585,98	531,42
	MAE	18,39	20,97	22,24	22,17	21,52	19,91	20,11	19,43	18,92
	MAPE	2,14	2,39	2,52	2,52	2,48	2,30	2,32	2,24	2,20
17 de Agosto 2019	RMSE	36,50	37,78	41,41	41,51	44,60	26,97	26,18	25,01	23,66
	PCC	0,96	0,92	0,92	0,92	0,90	0,95	0,95	0,95	0,96
	MSE	1331,99	1427,54	1714,65	1722,99	1989,09	727,46	685,17	625,63	559,93
	MAE	30,29	28,24	31,74	31,76	36,46	21,82	21,13	19,90	19,25
	MAPE	3,55	3,25	3,65	3,65	4,25	2,52	2,45	2,31	2,25
18 de Agosto 2019	RMSE	47,88	51,71	51,05	50,64	49,90	32,34	33,61	33,93	31,17
	PCC	0,94	0,90	0,90	0,90	0,90	0,94	0,94	0,94	0,95
	MSE	2292,74	2673,43	2606,60	2564,00	2489,65	1045,65	1129,43	1151,01	971,86
	MAE	38,77	39,18	39,43	38,73	38,34	24,74	25,43	25,21	23,87
	MAPE	4,65	4,63	4,69	4,60	4,56	2,97	3,03	3,02	2,86

Fonte: Elaborado pelo próprio autor.

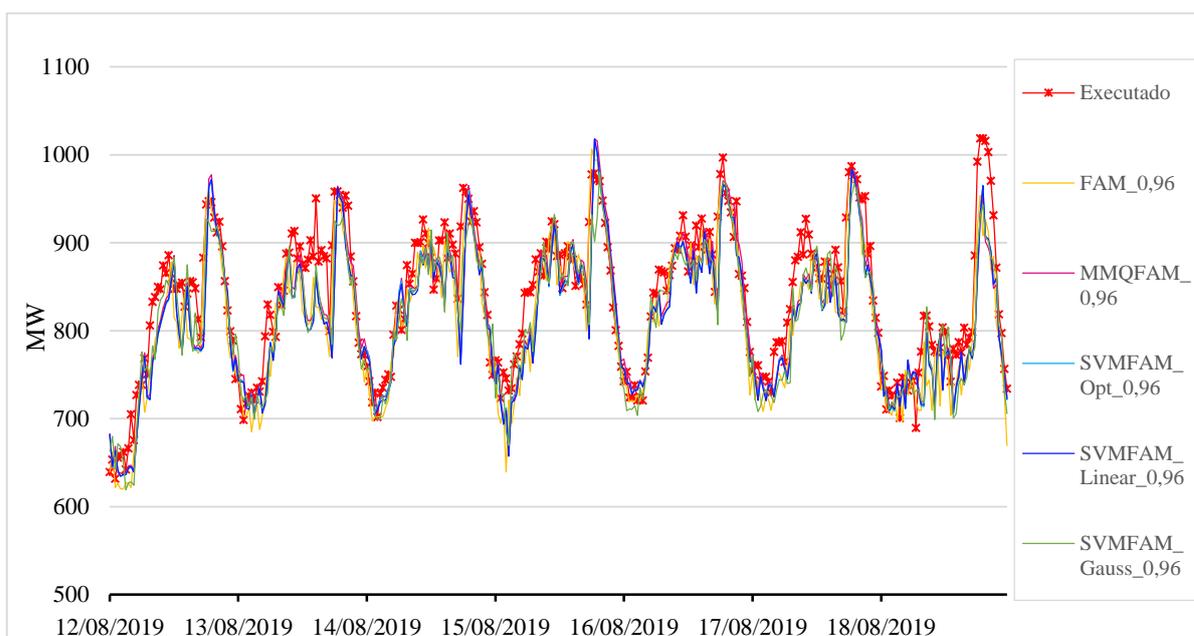
Na Tabela 2 são apresentados os resultados das metodologias propostas, bem como dos *benchmarks* 1 e 2. Os desempenhos das metodologias 2 e 3 foram excelentes, com reduções de cerca de 50% no MAPE das cargas elétricas dos dias 12, 15, 17 e 18 de agosto, em comparação com os *benchmarks* 1 e 2. Destaca-se que a metodologia 3, que utiliza o *kernel* Gaussiano com parâmetros de vigilância ρ_a de 0,99 para os dados de aproximação e de detalhe, apresentou um desempenho notável. Durante a fase de diagnóstico, um limite de ρ_a igual a 0,90 foi utilizado para o modelo de detalhe, o que gerou os melhores resultados em termos de generalização na

metodologia. Como resultado, a metodologia 3 apresentou uma superioridade significativa em termos de MAPE na previsão das cargas elétricas no nível de desagregação mencionado.

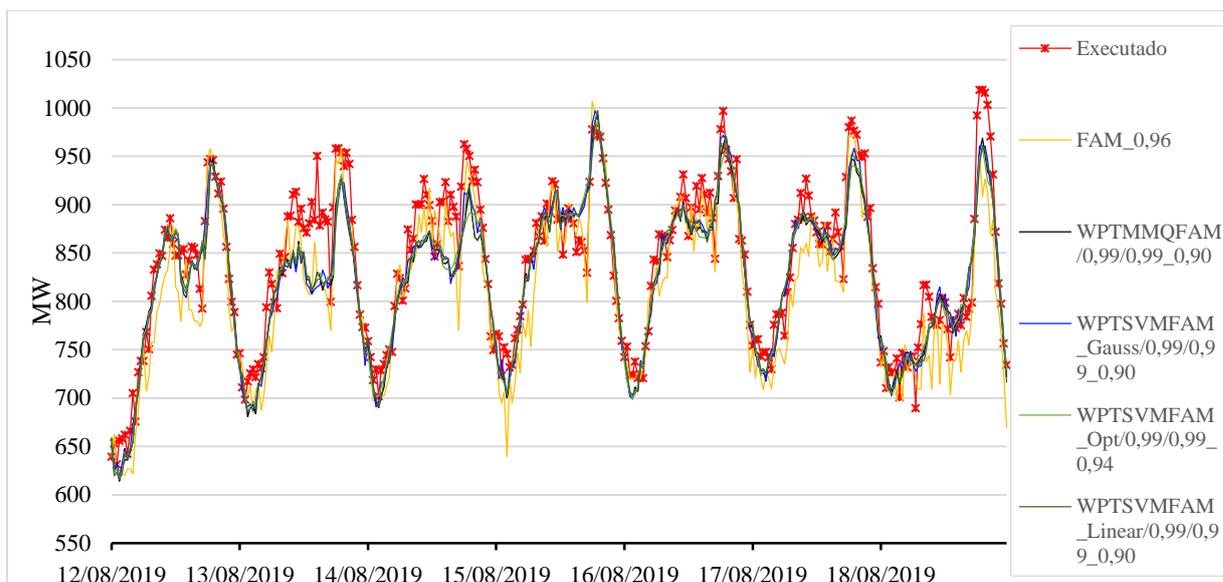
As curvas de previsão diárias são apresentadas nas Figura 37 e Figura 38, que foram divididas para uma melhor compreensão do comportamento em relação aos desempenhos apresentados na Tabela 2. Na Figura 37 observa-se o comportamento das cargas em relação à metodologia 1 e aos *benchmarks* 1 e 2. Percebe-se que a metodologia 1, que utiliza o *kernel* linear, é bastante competitiva em relação aos *benchmarks*. Entretanto, a metodologia só conseguiu ser superior parcialmente em dias específicos, como nos dias 12 e 18 de agosto.

Por outro lado, na Figura 38, observa-se um acompanhamento bastante aceitável da metodologia 3 em relação à curva de carga real. Esta metodologia com o *kernel* apresentou boa performance, o que é refletido no PCC da Tabela 2.

Figura 37 - Curvas da previsão da metodologia 1 versus *benchmark* 1 e 2 (Área norte SEIN)



Fonte: Elaborado pelo próprio autor.

Figura 38 - Curvas da previsão das metodologias 2 e 3 versus *benchmark* 1 (Área norte SEIN)

Fonte: Elaborado pelo próprio autor.

7.2.3 Previsão de carga da subárea sudoeste

Com o objetivo de encontrar um modelo predictor de alto desempenho para a subárea sudoeste peruana, foram realizadas comparações entre os resultados obtidos a partir das metodologias estabelecidas, incluindo os *benchmarks*, dispostos na Tabela 3. Cada modelo foi submetido a uma análise detalhada para determinar os melhores parâmetros internos por meio de validação cruzada, garantindo, assim, níveis aceitáveis de generalização.

A partir da análise da Tabela 3, é possível notar que a primeira metodologia apresentou bons resultados em termos de MAPE quando comparada aos modelos de *benchmark* 1 e 2. É importante destacar que essa melhoria foi significativa durante toda a semana.

Em relação às outras metodologias, a primeira não apresentou uma melhoria substancial em comparação com a segunda e a terceira. A terceira metodologia, que utiliza um *kernel* Gaussiano, apresentou resultados satisfatórios em geral, embora nem todas as avaliações tenham apresentado bom desempenho na previsão diária. Conforme indicado na Tabela 3, a metodologia 3, que utiliza um *kernel* Gaussiano com parâmetros de vigilância ρ_a de 0,99 para os dados de aproximação e detalhe, obteve um desempenho notável. Durante a fase de diagnóstico, um limite de ρ_a igual a 0,98 foi utilizado para o modelo de detalhe, o que resultou nos melhores resultados em termos de generalização da metodologia.

Em resumo, tanto a metodologia 2, quanto a metodologia 3 apresentaram melhorias significativas no desempenho em termos de MAPE na previsão das cargas elétricas no nível de desagregação mencionado. No entanto, é importante destacar que a metodologia 3 apresentou

um desempenho parcialmente aceitável. Portanto, pode ser considerada como uma opção promissora para previsões futuras.

Tabela 3 - Resultados das metodologias (Previsão subárea sudoeste)

Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTSVMFAM Opt	WPTSVMFAM Linear	WPTSVMFAM Gauss
		0,99	0,99	0,99	0,99	0,99	0,99A/0,99D 0,90	0,99A/0,99D 0,97	0,99A/0,99D 0,90	0,99A/0,99D 0,98
12 de Agosto 2019	RMSE	52,70	50,91	50,97	50,94	53,08	45,36	45,94	45,89	48,12
	PCC	0,73	0,76	0,76	0,76	0,73	0,82	0,82	0,82	0,80
	MSE	2777,79	2592,17	2597,60	2594,42	2817,00	2057,11	2110,79	2105,83	2315,67
	MAE	43,41	39,97	39,74	39,74	41,05	34,49	35,34	35,29	36,63
	MAPE	3,67	3,40	3,37	3,37	3,48	2,92	3,00	3,00	3,10
13 de Agosto 2019	RMSE	58,58	56,39	56,51	56,19	55,49	55,89	55,88	56,79	55,96
	PCC	0,19	0,01	0,01	0,02	-0,01	0,51	0,48	0,49	0,46
	MSE	3431,62	3180,35	3192,90	3157,34	3078,68	3123,85	3122,73	3225,57	3131,07
	MAE	47,77	44,58	44,65	44,35	44,08	47,64	47,36	48,53	47,73
	MAPE	4,04	3,77	3,78	3,75	3,73	4,02	4,00	4,09	4,03
14 de Agosto 2019	RMSE	58,24	56,03	53,46	56,18	56,17	57,76	57,11	58,18	55,90
	PCC	0,24	0,22	0,22	0,22	0,26	0,20	0,18	0,20	0,20
	MSE	3391,74	3139,00	2858,45	3156,17	3154,90	3336,37	3261,45	3384,44	3124,71
	MAE	50,19	48,52	46,01	48,54	47,22	50,67	50,44	50,61	48,39
	MAPE	4,38	4,23	4,02	4,23	4,12	4,43	4,40	4,42	4,23
15 de Agosto 2019	RMSE	108,97	103,43	103,77	103,37	105,08	121,15	119,86	120,57	121,14
	PCC	0,24	0,26	0,26	0,26	0,26	0,19	0,22	0,18	0,17
	MSE	11874,81	10697,21	10767,93	10685,78	11042,81	14677,45	14366,61	14535,96	14673,70
	MAE	77,39	73,51	73,72	73,28	75,63	97,51	97,02	96,63	97,40
	MAPE	7,25	6,90	6,92	6,88	7,09	9,07	9,02	8,99	9,06
16 de Agosto 2019	RMSE	63,04	55,94	56,81	55,98	56,87	61,69	62,33	62,15	62,17
	PCC	0,77	0,80	0,80	0,79	0,77	0,82	0,82	0,82	0,80
	MSE	3973,52	3129,39	3227,34	3133,26	3234,17	3805,53	3885,61	3862,08	3864,70
	MAE	52,88	46,86	47,28	46,69	47,56	54,42	54,44	54,47	53,54
	MAPE	4,48	3,97	4,00	3,96	4,04	4,61	4,61	4,61	4,54
17 de Agosto 2019	RMSE	44,16	39,46	39,80	39,67	40,24	31,97	31,30	31,24	31,36
	PCC	0,57	0,59	0,58	0,59	0,59	0,83	0,83	0,83	0,83
	MSE	1950,10	1556,93	1584,34	1574,02	1619,46	1022,10	979,68	975,83	983,43
	MAE	37,03	33,55	33,49	33,58	33,63	27,60	26,81	26,68	26,78
	MAPE	3,11	2,81	2,81	2,81	2,83	2,30	2,23	2,22	2,22
18 de Agosto 2019	RMSE	68,21	66,81	66,11	66,25	64,85	35,16	35,59	36,00	35,90
	PCC	0,28	0,10	0,10	0,10	0,10	0,49	0,47	0,48	0,47
	MSE	4652,06	4463,64	4370,25	4388,89	4205,63	1236,13	1266,36	1295,65	1288,71
	MAE	56,40	53,95	53,33	53,38	53,21	28,89	28,96	29,78	29,42
	MAPE	4,93	4,73	4,67	4,68	4,65	2,50	2,50	2,57	2,54

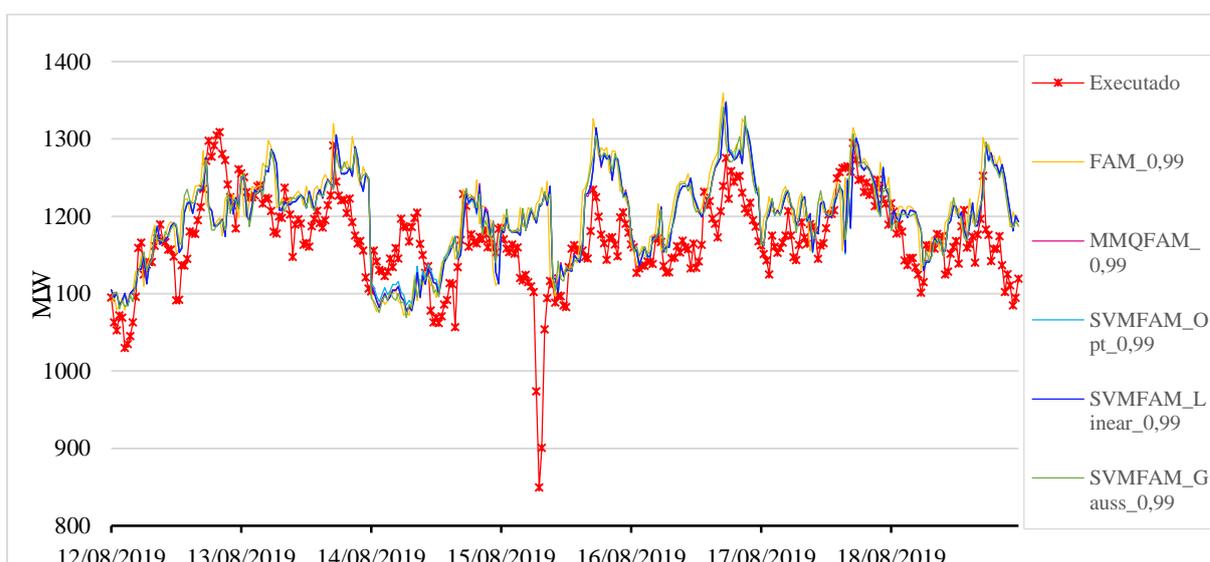
Fonte: Elaborado pelo próprio autor.

Para melhor compreensão do comportamento em relação aos desempenhos apresentados na Tabela 3, as curvas nas previsões diárias foram divididas em duas figuras distintas, Figura 39 e Figura 40. A Figura 39 apresenta o comportamento das cargas no que tange à metodologia 1 e aos *benchmarks* 1 e 2. É possível perceber que a metodologia 1, que utiliza o *kernel* Linear,

é bastante competitiva em relação aos *benchmarks* para os dias 12, 13, 16 e 17. Por outro lado, a Figura 40 mostra um acompanhamento bastante aceitável da metodologia 3 em relação à curva de carga real. Esta metodologia, com o *kernel* Gaussiano, apresentou um desempenho satisfatório, o que é refletido no PCC da Tabela 3.

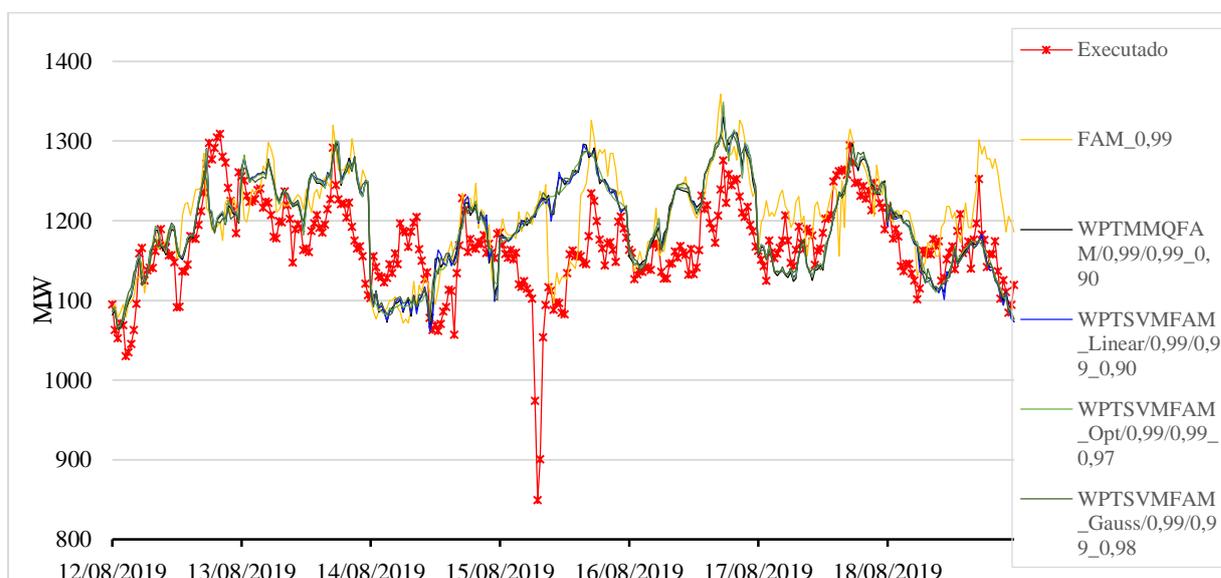
Em resumo, as figuras fornecem uma melhor visualização do desempenho das diferentes metodologias em relação aos *benchmarks* e à carga real, destacando a competitividade da metodologia 1 em alguns dias específicos e o bom acompanhamento da metodologia 3 ao longo do período analisado. Os resultados obtidos reforçam a importância de utilizar diferentes técnicas e *kernels* no processo de previsão, a fim de obter uma visão mais abrangente do comportamento das séries temporais.

Figura 39 - Curvas de carga das metodologias 1, 2 e 3 (Previsão subárea sudoeste)



Fonte: Elaborado pelo próprio autor.

Figura 40 - Curvas de cargas das metodologias 4 e 5 (Previsão subárea sudoeste)



Fonte: Elaborado pelo próprio autor.

7.2.4 Previsão de carga dos grandes usuários livres

Neste apartado compara-se os resultados obtidos da aplicação das metodologias de previsão, com o objetivo de encontrar o modelo que melhor se adapte ao nível desagregado de grandes usuários livres. Os resultados dessa comparação foram dispostos na Tabela 4. Cada modelo foi submetido a uma análise detalhada para determinar os melhores parâmetros internos por meio de validação cruzada, garantindo, assim, níveis aceitáveis de generalização.

Ao analisar a Tabela 4 é possível notar que a primeira metodologia apresentou uma melhoria leve e apenas em dias específicos em termos de MAPE, quando comparada aos modelos de *benchmark* 1 e 2. É importante destacar que essa melhoria foi pontualmente para o *kernel* Gaussiano nos dias 14, 16 e 17 de agosto.

Em relação às outras metodologias, a primeira não apresentou uma melhoria substancial em comparação com a segunda e a terceira. A terceira metodologia, que utiliza um *kernel* Gaussiano, apresentou resultados satisfatórios em geral, embora nem todas as avaliações tenham apresentado bom desempenho na previsão diária. Conforme indicado na Tabela 4, esta metodologia 3 com parâmetros de vigilância ρ_a de 0,92 para os dados de aproximação e 0,99 para os de detalhe, obteve um desempenho notável para a grande maioria das avaliações. Durante a fase de diagnóstico, um limite de ρ_a igual a 0,97 foi utilizado para o modelo de detalhe, o que evidenciou os melhores resultados em termos de generalização da metodologia.

Em resumo, tanto a metodologia 2, quanto a metodologia 3 apresentaram melhorias significativas no desempenho em termos de MAPE na previsão das cargas elétricas no nível de

desagregação dos grandes usuários livres. No entanto, é importante destacar que a metodologia 3 apresentou um desempenho aceitável para a grande maioria das avaliações, portanto, pode ser considerada como uma opção promissora para previsões futuras. Com base nisso, pode-se concluir que os resultados desta pesquisa fornecem importantes informações para aprimorar a gestão da energia elétrica neste nível de desagregação do SEIN, o que pode trazer benefícios significativos para a sociedade em termos de confiabilidade e eficiência do sistema.

Tabela 4 - Resultados das metodologias (Previsão dos grandes usuários livres SEIN)

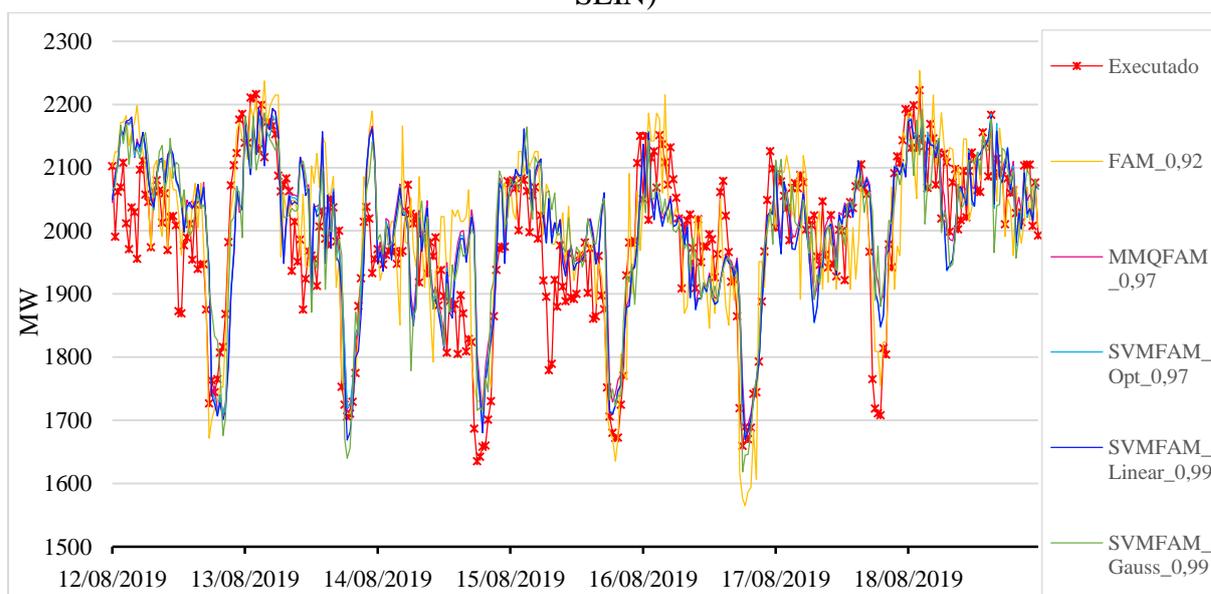
Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTSVMFAM Opt	WPTSVMFAM Linear	WPTSVMFAM Gauss
		0,92	0,97	0,97	0,97	0,97	0,92A/0,99D 0,90	0,92A/0,99D 0,90	0,92A/0,99D 0,90	0,92A/0,99D 0,97
12 de Agosto 2019	RMSE	79,65	110,09	108,48	111,63	104,20	81,65	79,76	81,00	82,67
	PCC	0,87	0,65	0,66	0,66	0,66	0,77	0,78	0,78	0,80
	MSE	6344,24	12118,91	11768,86	12461,76	10858,25	6667,02	6361,46	6560,81	6835,11
	MAE	61,14	93,24	91,66	95,24	89,07	60,20	59,73	60,58	58,83
	MAPE	3,10	4,73	4,65	4,84	4,55	3,03	3,01	3,05	2,96
13 de Agosto 2019	RMSE	86,68	88,16	86,50	86,78	83,32	80,00	83,55	84,38	73,06
	PCC	0,85	0,82	0,81	0,81	0,82	0,83	0,81	0,81	0,86
	MSE	7513,26	7772,86	7481,92	7530,82	6941,89	6399,26	6981,36	7120,61	5337,67
	MAE	61,86	66,75	67,30	69,45	63,08	63,06	63,81	63,54	59,01
	MAPE	3,11	3,39	3,42	3,52	3,22	3,21	3,24	3,22	2,98
14 de Agosto 2019	RMSE	104,35	104,33	100,57	98,95	102,55	58,03	59,68	60,22	58,41
	PCC	0,65	0,61	0,61	0,61	0,60	0,87	0,86	0,86	0,87
	MSE	10889,54	10885,64	10114,78	9791,11	10516,37	3366,93	3561,63	3626,22	3412,13
	MAE	86,89	82,11	79,20	78,53	81,91	46,22	47,20	47,64	47,13
	MAPE	4,63	4,45	4,27	4,22	4,39	2,44	2,50	2,52	2,48
15 de Agosto 2019	RMSE	88,44	91,25	87,69	87,95	97,61	79,00	77,99	79,22	81,40
	PCC	0,79	0,78	0,79	0,79	0,75	0,79	0,79	0,79	0,77
	MSE	7822,02	8326,97	7688,96	7735,18	9528,28	6240,30	6081,93	6275,57	6625,84
	MAE	64,01	72,91	68,28	68,47	77,09	61,49	60,64	62,20	62,37
	MAPE	3,34	3,86	3,60	3,61	4,06	3,22	3,18	3,26	3,27
16 de Agosto 2019	RMSE	87,84	74,55	77,38	76,94	73,23	86,96	79,93	76,75	81,46
	PCC	0,87	0,86	0,86	0,86	0,88	0,88	0,89	0,89	0,90
	MSE	7715,70	5557,55	5987,48	5919,78	5362,09	7561,62	6388,44	5889,87	6635,45
	MAE	75,21	60,12	63,10	62,64	60,57	73,80	66,99	63,83	69,96
	MAPE	3,88	3,02	3,16	3,13	3,04	3,71	3,39	3,22	3,54
17 de Agosto 2019	RMSE	81,23	86,19	83,50	83,47	83,60	90,76	91,00	90,67	84,17
	PCC	0,70	0,66	0,65	0,65	0,68	0,57	0,57	0,58	0,66
	MSE	6599,12	7429,01	6971,99	6967,58	6989,42	8238,22	8281,00	8220,69	7084,54
	MAE	61,88	62,77	62,07	61,96	57,28	70,40	70,03	70,07	64,30
	MAPE	3,11	3,28	3,21	3,21	3,01	3,64	3,63	3,63	3,33
18 de Agosto 2019	RMSE	61,10	49,09	60,83	58,95	61,06	58,34	52,79	56,99	56,88
	PCC	0,50	0,57	0,52	0,53	0,49	0,52	0,56	0,54	0,57
	MSE	3733,37	2410,21	3700,12	3475,28	3728,88	3403,62	2786,54	3247,58	3234,86
	MAE	48,01	40,79	48,44	46,81	48,52	49,34	44,16	47,50	48,20
	MAPE	2,31	1,96	2,32	2,24	2,32	2,37	2,12	2,28	2,31

Fonte: Elaborado pelo próprio autor.

Para facilitar a compreensão dos resultados apresentados na Tabela 4, foram geradas duas figuras distintas, a Figura 41 e a Figura 42. A Figura 41 mostra o comportamento das cargas executadas em relação à metodologia 1 e aos *benchmarks* 1 e 2. Observa-se que a metodologia 1, que utiliza os recursos de *kernel* e otimização, não conseguiu acompanhar a curva de carga real. Já a Figura 42 demonstra um acompanhamento bastante aceitável da metodologia 3. Essa metodologia, que utiliza o *kernel* Gaussiano, apresentou um desempenho satisfatório, o que é refletido no PCC da Tabela 4.

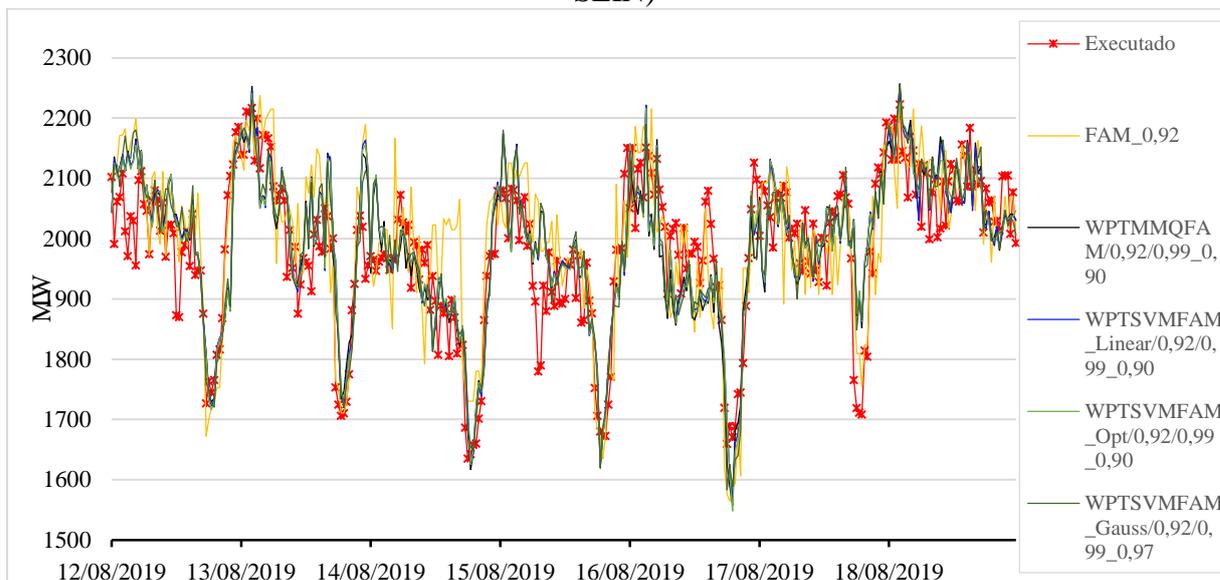
Em resumo, as figuras proporcionam uma melhor visualização do desempenho das diferentes metodologias em relação aos *benchmarks* e à carga real, evidenciando um excelente desempenho e acompanhamento da metodologia 3 ao longo do período analisado. Os resultados obtidos reforçam a importância de utilizar diferentes técnicas e *kernels* em análises de previsão, a fim de obter uma visão mais abrangente do comportamento das séries temporais.

Figura 41 - Curvas de carga das metodologias 1, 2 e 3 (Previsão dos grandes usuários livres SEIN)



Fonte: Elaborado pelo próprio autor.

Figura 42 - Curvas de carga das metodologias 4 e 5 (Previsão dos grandes usuários livres SEIN)



7.2.5 Previsão de carga dos grandes usuários livres da área centro

Visando encontrar o modelo de previsão mais eficiente para o nível de desagregação grandes usuários livres da área centro, foram comparados os resultados obtidos a partir da aplicação de diferentes metodologias, incluindo os *benchmarks*, dispostos na Tabela 5. Cada modelo foi submetido a uma análise detalhada para determinar os melhores parâmetros internos por meio de validação cruzada, garantindo, assim, níveis aceitáveis de generalização.

Durante a análise da Tabela 5, percebe-se que a primeira metodologia e a aplicação do *benchmark 2* apresentaram MAPEs de pobríssimo desempenho. Os modelos de regressão envolvidos nessas metodologias, *e.g.*, Métodos dos Mínimos Quadrados e *Support Vector Machine*, não conseguiram se adaptar de forma eficiente ao problema nesse nível, isto como consequência da ausência de categorias necessárias para a processo de regressão, como consequência forneceram funções de aproximação inexatas. É importante destacar a fraqueza dessas propostas e procurar uma forma de interromper esse processo frente à escassez de elementos necessários para uma boa generalização.

Portanto, a primeira não demonstrou qualquer melhoria em relação à segunda e a terceira metodologia. A terceira metodologia, que utiliza um *kernel* Gaussiano, apresentou resultados satisfatórios comparados com os *benchmarks*. Conforme indicado na Tabela 5, esta metodologia 3 com parâmetros de vigilância ρ_a de 0,93 para os dados de aproximação e 0,99 para os de detalhe, obteve um desempenho notável para a grande maioria das avaliações.

Durante a fase de diagnóstico, um limite de ρ_a igual a 0,99 foi utilizado para o modelo de detalhe, o que propiciou os melhores resultados em termos de generalização da metodologia.

Em resumo, a metodologia 3 apresenta melhorias significativas no desempenho em termos de MAPE na previsão das cargas elétricas no nível de desagregação dos grandes usuários livres da área centro. No entanto, é importante destacar que a metodologia 3 apresentou um desempenho aceitável para a grande maioria das avaliações fornecendo desempenhos que superaram os *benchmarks*, portanto, pode ser considerada como uma opção promissora para previsões futuras. Com base nisso, pode-se concluir que os resultados desta pesquisa fornecem importantes informações para aprimorar a gestão da energia elétrica neste nível de desagregação no SEIN, o que pode trazer benefícios significativos para a sociedade em termos de confiabilidade e eficiência do sistema.

Tabela 5 - Resultados das metodologias (Previsão dos grandes usuários livres área centro)

Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTVMFAM Opt	WPTVMFAM Linear	WPTVMFAM Gauss
		0,93	0,99	0,93	0,93	0,93	0,93A/0,99D 0,99	0,93A/0,99D 0,90	0,93A/0,99D 0,93	0,93A/0,99D 0,99
12 de Agosto 2019	RMSE	67,55	108,92	274,20	159,27	132,68	53,92	52,43	53,45	46,69
	PCC	0,83	0,78	0,48	0,36	0,17	0,90	0,90	0,89	0,92
	MSE	4563,15	11864,41	75183,54	25366,46	17602,86	2907,64	2748,95	2856,94	2179,92
	MAE	47,45	94,49	254,53	121,45	123,76	44,07	44,05	43,38	38,07
	MAPE	5,80	11,49	31,24	14,50	15,34	5,44	5,50	5,37	4,73
13 de Agosto 2019	RMSE	68,34	177,03	103,19	101,36	129,32	54,72	56,61	56,80	51,45
	PCC	0,82	0,45	0,63	0,63	0,02	0,83	0,82	0,82	0,85
	MSE	4670,64	31338,28	10648,96	10273,05	16722,94	2993,93	3204,34	3226,61	2647,07
	MAE	47,39	159,98	79,70	77,72	119,83	46,73	48,41	48,47	42,22
	MAPE	5,96	19,43	9,84	9,61	14,72	5,84	6,07	6,10	5,38
14 de Agosto 2019	RMSE	74,19	83,74	69,01	64,78	81,53	64,33	64,51	65,14	67,26
	PCC	0,79	0,75	0,74	0,74	0,05	0,85	0,85	0,85	0,85
	MSE	5504,00	7011,69	4761,73	4196,52	6647,37	4138,10	4161,09	4242,72	4524,39
	MAE	60,47	72,22	56,29	52,35	61,73	53,29	52,73	53,18	55,16
	MAPE	8,19	9,86	7,82	7,23	9,20	7,42	7,37	7,44	7,65
15 de Agosto 2019	RMSE	75,22	95,21	87,99	88,17	106,42	69,28	67,05	67,43	68,23
	PCC	0,68	0,68	0,62	0,62	0,11	0,69	0,70	0,71	0,71
	MSE	5657,31	9065,31	7741,42	7773,82	11325,11	4799,85	4495,28	4547,31	4655,71
	MAE	61,53	78,56	74,37	74,36	96,85	58,00	54,78	55,32	55,63
	MAPE	8,07	10,51	9,88	9,92	12,98	7,59	7,15	7,23	7,23
16 de Agosto 2019	RMSE	67,09	233,96	87,81	105,88	119,34	72,33	74,03	75,68	73,16
	PCC	0,80	0,71	0,71	0,69	0,04	0,86	0,86	0,86	0,88
	MSE	4501,46	54736,32	7710,19	11209,59	14242,70	5232,22	5480,10	5728,02	5351,75
	MAE	52,55	223,80	71,44	86,59	110,35	60,50	62,13	63,85	61,73
	MAPE	7,04	29,42	9,39	11,25	14,75	7,75	7,95	8,17	8,02
17 de Agosto 2019	RMSE	66,75	181,29	147,15	163,78	109,59	71,21	75,94	77,03	69,66
	PCC	0,70	0,41	0,58	0,55	-0,02	0,72	0,69	0,71	0,73
	MSE	4455,70	32867,84	21652,29	26823,43	12009,96	5070,47	5767,22	5933,83	4852,88
	MAE	52,41	163,36	130,36	147,39	97,96	58,61	61,53	64,49	55,18
	MAPE	7,12	20,95	17,05	19,28	13,12	7,71	8,15	8,40	7,25
18 de Agosto 2019	RMSE	72,11	138,44	78,07	92,37	76,62	62,02	60,83	63,36	63,28
	PCC	0,06	-0,05	-0,09	-0,09	0,09	-0,08	-0,14	-0,06	-0,07
	MSE	5200,44	19166,67	6094,68	8532,98	5870,74	3845,99	3700,12	4014,96	4004,97
	MAE	56,70	123,39	57,70	70,09	65,36	50,40	49,42	51,09	49,73
	MAPE	6,72	14,17	6,72	8,01	7,43	5,94	5,81	6,02	5,84

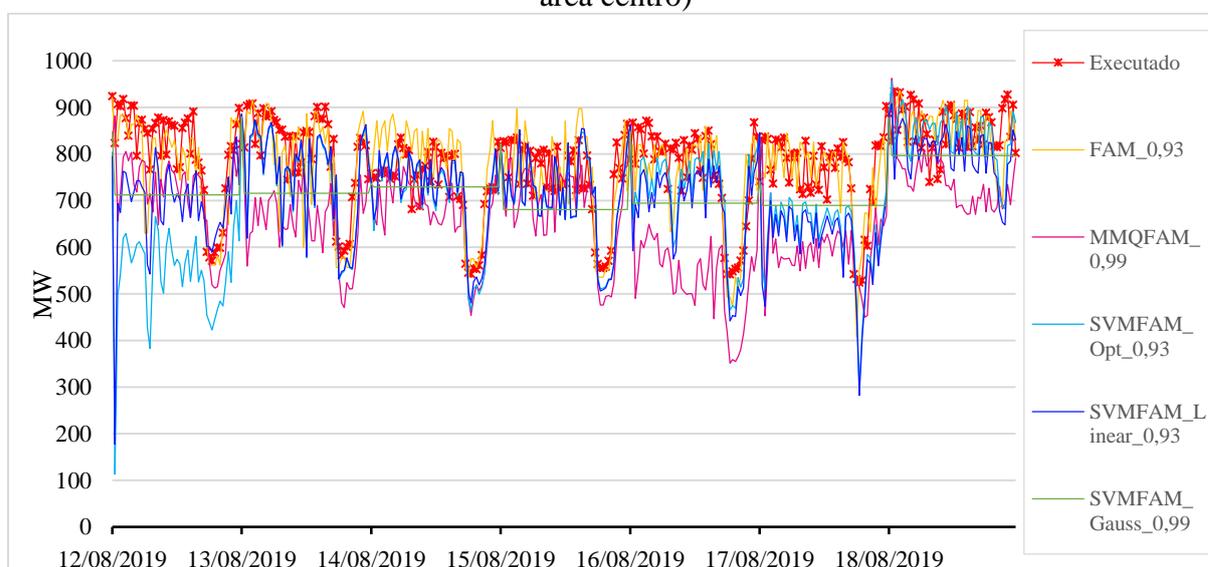
Fonte: Elaborado pelo próprio autor.

As curvas resultantes da comparação da performance exibido na tabela 5 são apresentadas em duas representações distintas, a Figura 43 e a Figura 44. A Figura 43 mostra como a metodologia 1 e os *benchmarks* 1, destaca-se que tanto a metodologia 1 quanto o benchmark 2 não conseguiram acompanhar a curva de carga real, sendo um dos piores modelos para se adaptar à previsão das cargas elétricas. Por outro lado, na Figura 44, a metodologia 3 apresentou um acompanhamento bastante aceitável em relação à curva de carga real. Essa

metodologia, que utiliza o *kernel* Gaussiano, demonstrou um desempenho satisfatório, como evidenciado pelo PCC apresenta na Tabela 5.

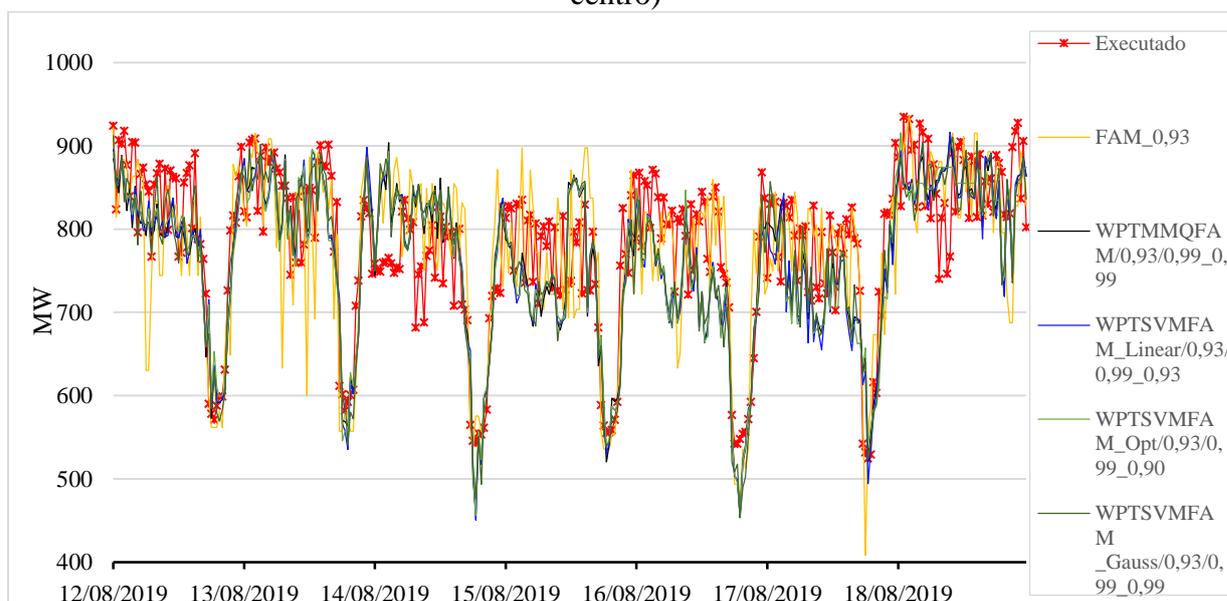
Em resumo, torna-se possível encontrar características particulares decorrentes da análise das curvas das diversas metodologias já que garantem uma melhor visualização do desempenho em relação aos *benchmarks* e à curva de carga real. Dessa análise, destaca-se um excelente desempenho e acompanhamento da metodologia 3 ao longo do período estudado. Os resultados obtidos enfatizam a importância de se utilizar diferentes técnicas e *kernels* em análises de previsão, a fim de obter uma visão mais abrangente do comportamento das séries temporais e, assim, providenciar previsões mais precisas e confiáveis.

Figura 43 - Curvas de carga das metodologias 1, 2 e 3 (Previsão dos grandes usuários livres área centro)



Fonte: Elaborado pelo próprio autor.

Figura 44 - Curvas de carga das metodologias 4 e 5 (Previsão dos grandes usuários livres área centro)



Fonte: Elaborado pelo próprio autor.

7.2.6 Previsão de carga dos grandes usuários livres da área sul

Com o objetivo de encontrar o modelo de previsão mais eficiente para o nível de desagregação grandes usuários livres da área sul, foram comparados os resultados obtidos a partir da aplicação de diferentes metodologias, incluindo os *benchmarks*, dispostas na Tabela 6. Cada modelo foi submetido a uma análise detalhada para determinar os melhores parâmetros internos por meio de validação cruzada, garantindo, assim, níveis aceitáveis de generalização.

Ao analisar a Tabela 6, é possível notar que a primeira metodologia apresentou uma notável melhoria para todos os dias da semana em termos de MAPE, quando comparada ao modelo de *benchmark* 1. É importante destacar que essa melhoria no resultado foi satisfatória para o processo otimizado aplicado a todos os dias da semana.

Em relação às outras metodologias, a primeira apresentou uma melhoria substancial em comparação com a segunda e terceira só para os dias 15, 16, 17 e 18 de agosto. No entanto, a terceira metodologia, que utiliza o processo otimizado, conseguiu reduzir notavelmente os MAPEs nos outros dias, fornecendo resultados em geral satisfatórios embora nem todas as avaliações tenham apresentado bom desempenho na previsão diária. Conforme indicado na Tabela 6, esta metodologia 3 com parâmetros de vigilância ρ_a de 0,99 para os dados de aproximação e 0,99 para os de detalhe, obteve um desempenho notável para a grande maioria das avaliações. Durante a fase de diagnóstico, um limite de ρ_a igual a 0,91 foi utilizado para o

modelo de detalhe, o que propiciou os melhores resultados em termos de generalização da metodologia.

Em resumo, tanto a metodologia 2 quanto a metodologia 3 apresentaram melhorias significativas no desempenho em termos de MAPE na previsão das cargas elétricas no nível de desagregação dos grandes usuários livres da área sul. No entanto, é importante destacar que a metodologia 3 apresentou um desempenho aceitável para a grande maioria das avaliações, portanto, pode ser considerada como uma opção promissora para previsões futuras. Com base nisso, pode-se concluir que os resultados desta pesquisa fornecem importantes informações para aprimorar a gestão da energia elétrica neste nível de desagregação no SEIN, o que pode trazer benefícios significativos para a sociedade em termos de confiabilidade e eficiência do sistema.

Tabela 6 - Resultados das metodologias (Previsão dos grandes usuários livres área sul)

Dia	Métrica	FAM	MMQFAM	SVMFAM Opt	SVMFAM Linear	SVMFAM Gauss	WPTMMQFAM	WPTSVMFAM Opt	WPTSVMFAM Linear	WPTSVMFAM Gauss
		0,99	0,99	0,99	0,99	0,99	0,99A/0,99D 0,90	0,99A/0,99D 0,91	0,99A/0,99D 0,93	0,99A/0,99D 0,93
12 de Agosto 2019	RMSE	83,71	82,38	83,76	83,71	77,22	72,92	72,54	72,38	72,67
	PCC	-0,10	-0,15	-0,15	-0,15	-0,06	-0,03	-0,03	-0,03	-0,01
	MSE	7007,12	6787,23	7015,57	7006,67	5963,30	5317,35	5262,14	5239,26	5280,70
	MAE	70,35	69,31	70,29	70,20	64,23	56,08	55,87	55,76	55,92
	MAPE	7,02	6,88	6,98	6,97	6,38	5,63	5,61	5,59	5,61
13 de Agosto 2019	RMSE	68,30	66,73	66,96	67,06	68,61	60,02	59,96	59,99	62,15
	PCC	0,70	0,70	0,69	0,69	0,70	0,76	0,76	0,77	0,75
	MSE	4665,28	4452,95	4483,39	4497,46	4707,84	3602,80	3595,45	3598,50	3862,63
	MAE	55,99	54,38	54,52	54,67	55,62	45,65	45,77	45,82	47,92
	MAPE	5,68	5,51	5,53	5,54	5,65	4,65	4,66	4,67	4,88
14 de Agosto 2019	RMSE	86,13	66,99	64,54	67,29	65,79	51,29	51,33	51,63	50,85
	PCC	0,09	0,34	0,34	0,33	0,33	0,48	0,49	0,49	0,50
	MSE	7418,55	4487,51	4164,77	4527,88	4327,95	2630,67	2634,77	2666,15	2585,51
	MAE	63,39	51,11	49,96	51,34	51,26	43,85	44,04	44,21	43,37
	MAPE	6,19	5,00	4,89	5,02	5,02	4,46	4,48	4,50	4,42
15 de Agosto 2019	RMSE	55,96	55,75	56,23	55,88	53,43	54,24	54,33	54,09	54,99
	PCC	0,61	0,60	0,60	0,60	0,61	0,62	0,62	0,62	0,61
	MSE	3131,54	3107,81	3161,96	3122,31	2855,14	2941,91	2951,32	2926,22	3024,07
	MAE	37,14	37,83	38,16	37,98	35,83	38,03	38,00	37,98	38,81
	MAPE	3,79	3,86	3,89	3,87	3,67	3,87	3,87	3,86	3,96
16 de Agosto 2019	RMSE	50,87	46,24	46,66	46,49	46,66	49,57	49,31	49,35	49,44
	PCC	0,45	0,54	0,54	0,54	0,54	0,50	0,50	0,50	0,51
	MSE	2587,48	2138,16	2177,22	2161,74	2176,84	2457,55	2431,76	2435,45	2443,85
	MAE	42,36	38,54	38,88	38,77	39,68	42,07	41,91	41,72	41,82
	MAPE	4,13	3,76	3,79	3,78	3,87	4,09	4,07	4,05	4,06
17 de Agosto 2019	RMSE	20,41	18,11	18,68	18,76	20,15	22,66	22,61	23,02	22,54
	PCC	0,72	0,72	0,71	0,71	0,68	0,52	0,53	0,54	0,55
	MSE	416,47	327,84	349,00	351,87	405,83	513,30	511,43	530,07	508,21
	MAE	17,19	14,67	15,18	15,23	16,21	17,67	17,51	17,69	17,73
	MAPE	1,62	1,38	1,43	1,44	1,53	1,67	1,66	1,67	1,68
18 de	RMSE	65,31	62,62	63,95	63,91	61,00	67,43	65,86	66,85	67,71
	PCC	-0,25	-0,32	-0,33	-0,32	-0,11	-0,15	-0,16	-0,15	-0,16

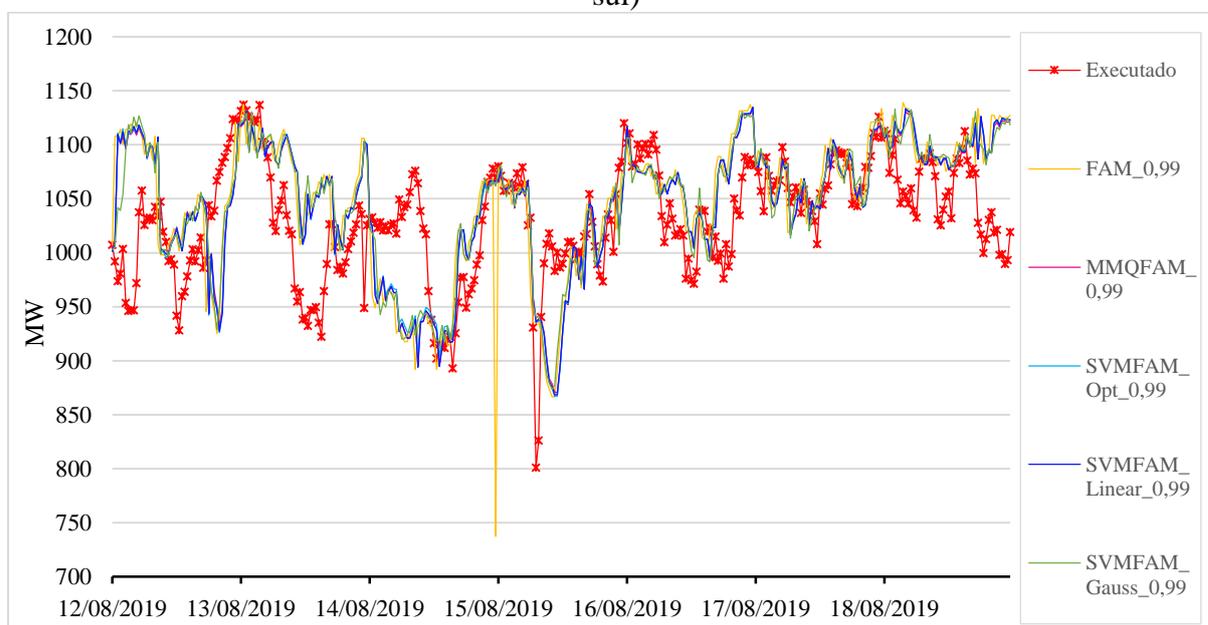
Agosto	MSE	4264,99	3921,36	4089,43	4084,75	3721,05	4546,92	4337,85	4468,56	4584,65
2019	MAE	51,22	48,56	49,66	49,65	48,60	54,36	52,70	53,80	54,61
	MAPE	4,97	4,72	4,83	4,83	4,72	5,27	5,11	5,21	5,29

Fonte: Elaborado pelo próprio autor.

Para facilitar a compreensão dos resultados apresentados na Tabela 6, foram geradas duas figuras distintas, a Figura 45 e Figura 46. A Figura 45 mostra o comportamento da curva de carga real em relação à metodologia 1 e aos *benchmarks* 1 e 2. Observa-se que a metodologia 1, que utiliza os recursos de *kernel* e otimização, conseguiu acompanhar as curvas das cargas relativas aos dias 15, 16 e 17 de agosto. Já a Figura 46 demonstra uma pequena melhoria da performance da metodologia 3 em relação aos *benchmarks*.

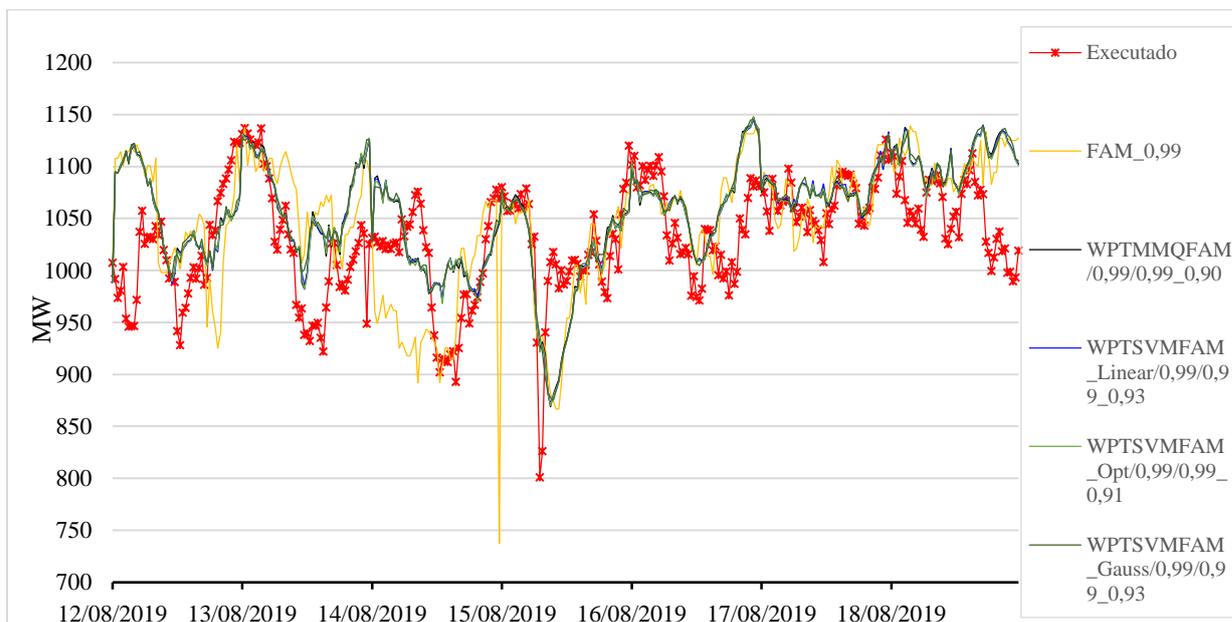
Em resumo, essas curvas de cargas elétricas possibilitam uma melhor visualização do desempenho das diferentes metodologias em relação aos *benchmarks*, evidenciando um desempenho algo significativo da metodologia 3 ao longo do período analisado. Os resultados obtidos reforçam a importância de utilizar diferentes técnicas e *kernels* em análises de previsão, a fim de obter uma visão mais abrangente do comportamento das séries temporais.

Figura 45 - Curvas de carga das metodologias 1, 2 e 3 (Previsão grandes usuários livres área sul)



Fonte: Elaborado pelo próprio autor.

Figura 46 - Curvas de carga das metodologias 4 e 5 (Previsão grandes usuários livres área sul)



Fonte: Elaborado pelo próprio autor.

7.3 DISCUSSÕES

Os resultados obtidos com a aplicação das metodologias foram comparados com os *benchmarks*. Dessa comparação, demonstrou-se que a metodologia 3 se destacou, entre as outras, por fornecer resultados altamente precisos na previsão de carga utilizando um *kernel* Gaussiano. Esses resultados foram provenientes de uma análise criteriosa dos dados e da validação cruzada nos seis níveis de desagregação estabelecidos.

Essa análise se baseou nas principais características extraídas da base de dados. Para garantir a precisão dos resultados, foram selecionados cuidadosamente os dados históricos de entrada para os modelos. Ressalta-se que foram identificados os dias com altos desequilíbrios energéticos e rotulados para uma melhor adaptação no processo de treinamento. Além disso, foi incluída uma análise dos dados meteorológicos para estabelecer a correlação direta e inversa à demanda elétrica de hora pico decorrente das características geográficas do Peru.

No processo de validação cruzada aplicado ao modelo da RNA ARTMAP *Fuzzy* modificada, foi realizado a busca pelo melhor parâmetro interno, *i.e.*, o de vigilância ρ_a . Desta forma, observou-se a existência de uma região geométrica que permitiu representar as melhores generalizações deste modelo predictor frente a dados de cargas em diferentes níveis de desagregação. Assim, este modelo simplifica consideravelmente as configurações internas na procura individual de avaliações quando não se tem certeza das características dos dados e garante, por meio da região geométrica, níveis altos em termos de generalização.

Desde o primeiro momento de avaliação das metodologias para diversos níveis de desagregação, já se tinha certeza dos parâmetros de vigilância a serem utilizados porque propiciam desempenhos significativos no que tange as métricas desejadas, que consistem nas mais comuns encontradas na literatura especializada e usadas por agentes de setor elétrico. Entretanto, para os parâmetros de limite nos cortes, MMQ e SVM, foram avaliados nos intervalos de 0,90 até 0,99 visando obter o melhor desempenho de cada metodologia, consequência da necessidade de incorporar dados nos modelos de regressão para obter as melhores funções representativas dos dados.

Embora se tenha constatado que a metodologia 1, proposta nesta pesquisa, conseguiu superar minimamente aos *benchmarks*. Percebe-se minimamente que os melhores desempenhos foram obtidos quanto mais desagregadas foram as entradas, como observa-se nas Tabela 3 e Tabela 6. Por esse motivo, abordou-se a hipótese de aplicar esta metodologia composta por dados mais desagregados e flutuante para procurar altos desempenhos.

Portanto, a característica da metodologia 1 foi incorporada à metodologia 3, que visou treinar dados de detalhes provenientes da análise multirresolução via *Wavelets*. De maneira similar, o *benchmark 2* foi introduzido na metodologia 2, visando assim buscar as melhores generalizações em relação aos dados com as características dos detalhes. Como resultado, foi observado um alto desempenho geral na aplicação dessas duas metodologias para todos os níveis de desagregação, especialmente na terceira metodologia utilizando o *kernel* Gaussiano. Isso demonstrou que este *kernel*, em conjunto com a metodologia 3, para dados de detalhe em um treinamento paralelo com a RNA ARTMAP *Fuzzy* modificada para os dados de aproximação, ofereceu um modelo de previsão robusto e com alto desempenho.

Durante a avaliação dos resultados, foi identificada uma anomalia na primeira metodologia, especificamente no treinamento dos dados para o nível de desagregação de grandes usuários livres na área centro, conforme apresentado na Figura 43. As curvas de previsão apresentaram flutuações e instabilidades. Após uma análise detalhada, concluiu-se que o problema estava relacionado à ausência de dados necessários no modelo de regressão, tanto no MMQ quanto no SVM. Esses modelos baseados em regressão exigem a identificação de dados relevantes para uma generalização adequada e uma boa proposta de função de aproximação do modelo.

Com base no que foi mencionado anteriormente, é essencial incluir um controle na programação da metodologia 1 para interromper o processamento quando o modelo não puder coletar os requisitos necessários. Isto ajudará a evitar resultados inadequados, flutuantes e

instáveis que possam tornar o sistema elétrico de potência instável e resultar em custos operacionais elevados.

Por fim, acentua-se que os melhores resultados apresentados neste trabalho foram obtidos por meio da metodologia 3 com o uso do *kernel* Gaussiano, que demonstrou um alto desempenho em ampla escala, sendo aplicável aos diferentes níveis de desagregação de cargas elétricas. Além disso, o modelo superou modelos convencionais utilizados na literatura, como a RNA ARTMAP *Fuzzy* modificada e o modelo híbrido MMQ_ARTMAP *Fuzzy*. Consequência disso, considera-se que este modelo previsor flexível é adequado para a previsão das cargas elétricas do dia seguinte para os níveis de desagregação mencionados.

8 CONCLUSÕES E SUGESTÕES PARA FUTUROS TRABALHOS

8.1 CONCLUSÕES

Esta pesquisa de doutorado teve como objetivo propor um modelo previsor flexível aplicado aos diversos níveis de desagregação de cargas elétricas, *e.g.*, as formas desagregadas das cargas elétricas nas diversas áreas de operação no Sistema Elétrico Peruano. Na sequência das avaliações, na busca da melhor modelagem, identificou-se algumas propostas que fizeram com que esta pesquisa se organize em três metodologias. Todas essas foram baseadas na RNA ARTMAP *Fuzzy* modificada. Porém a primeira metodologia utilizou um modelo híbrido entre essa RNA ARTMAP *Fuzzy* modificada e *Support Vector Machine* denominada como SVM_ARTMAP *Fuzzy*, na qual foram aproveitadas as categorias internas criadas no processo de diagnóstico para ajustar e criar funções de aproximação visando uma melhor adaptação das informações de cargas históricas. A segunda metodologia, utilizou o modelo híbrido MMQ_ARTMAP *Fuzzy* em um processo de treinamento paralelo com a RNA ARTMAP *Fuzzy* modificada, os dados de entradas utilizadas para o treinamento foram selecionados por meio de uma análise multirresolução via a transformada *Wavelet*. Finalmente, foi desenvolvida a metodologia 3, baseada em um processo de treinamento paralelo entre o modelo SVM_ARTMAP *Fuzzy* e a RNA ARTMAP *Fuzzy* modificada, da mesma forma que a metodologia 2, utilizou-se dados para o treinamento por meio de uma análise via a transformada *Wavelet*.

A partir de uma análise exaustiva, a metodologia 3, teve uma melhor performance nas avaliações nos diversos níveis de desagregação abordados na tese alcançado uma redução até um 60% em termos de MAPE no processo de comparação com as outras metodologias abordadas. Portanto, nesta tese foi proposta a metodologia 3, baseada no uso do *kernel* Gaussiano, como o modelo previsor flexível que conseguiu atingir altos desempenhos na previsão das cargas elétricas do dia seguinte aplicados aos diversos níveis de desagregação do Sistema Interligado Nacional Peruano. Dados meteorológicos, nesta pesquisa, como os dados da umidade relativa e a temperatura foram identificados como fatores-chave com correlações diretas e inversas com a demanda elétrica, dependendo da região avaliada, as quais foram utilizadas como entradas no intuito de melhorar os resultados. Finalmente, uma metodologia de pré-processamento baseado na modelagem matemática dos desvios de energias diárias históricos foi desenvolvida para identificar valores atípicos, incluindo dias festivos e anomalias na operação do sistema elétrico peruano, sem perder a sua natureza e características, resultando em previsões mais precisas.

Nesta tese visando atingir o objetivo geral, foram implementadas sequências de avaliação de desempenho das metodologias 1, 2 e 3, bem como dos *benchmarks*, em todos os níveis de desagregação estabelecidos, com uma análise cuidadosa dos parâmetros internos dos modelos. As métricas RMSE, PCC, MSE, MAE e MAPE foram utilizadas para avaliar os resultados das previsões. Com base na comparação e análise dos resultados, a metodologia 3 com o *kernel* Gaussiano foi escolhida como o modelo predictor flexível mais eficiente para se adaptar aos diferentes níveis de desagregação de cargas elétricas neste estudo.

Em resumo, esta tese oferece um modelo predictor flexível para a previsão das cargas elétricas, utilizando métodos de *Machine Learning*, no contexto de cargas em diversos níveis de desagregação. A metodologia proposta representa uma inovação no campo das previsões de carga, oferecendo maior precisão e eficiência na estimativa das demandas futuras.

As contribuições deste trabalho são:

- Implementação de um modelo matemático empírico baseado em desvios de grandes blocos de energia para a descoberta de dados atípicos. Essa abordagem se justifica pela necessidade de identificar e tratar dados atípicos que possam distorcer as previsões de carga. Ao implementar esse modelo, foi possível identificar e rotular esses dados, melhorando a qualidade das previsões.
- Análise de correlação e aplicação dos dados meteorológicos como *inputs* da RNA ARTMAP *Fuzzy* modificada. A inclusão de dados meteorológicos como inputs na rede se justifica pelo impacto que as condições climáticas podem ter no comportamento das cargas. Ao considerar esses dados na previsão de carga, foi possível melhorar a precisão das previsões, levando em conta as influências exógenas.
- Aplicação do processo de validação cruzada na RNA ARTMAP *Fuzzy* modificada com o intuito de gerar uma região geométrica oriunda dos melhores parâmetros de vigilância que se adequem aos dados em diversos níveis de desagregação. Essa abordagem se justifica pela necessidade de encontrar os melhores parâmetros de vigilância que se adequem aos dados de diferentes níveis de desagregação. A aplicação dessa, permitiu encontrar uma região geométrica ótima em cada nível de desagregação, melhorando a capacidade de adaptação da proposta aos diferentes padrões de carga.
- Análise multirresolução utilizando decomposição via transformada *Wavelet* aplicados aos diversos níveis de desagregação de carga. A aplicação dessa, permitiu identificar

e capturar informações dos dados em diversas escalas de frequência tendo um melhor controle dos *outliers* no processo de treinamento da proposta.

- Desenvolvimento e implementação das metodologias como a RNA ARTMAP *Fuzzy* modificada, a WPT-MMQ_ARTMAP *Fuzzy* e a WPT-SVM_ARTMAP *Fuzzy* aplicados aos distintos níveis de desagregação das cargas elétricas. Esses modelos foram desenvolvidos com o objetivo de melhorar a precisão das previsões de carga em diferentes níveis de desagregação. Cada modelo se justifica pela sua capacidade de lidar com os desafios específicos associados aos diferentes níveis de desagregação, proporcionando previsões mais acuradas e confiáveis.
- Aplicação e avaliação de resultados do modelo híbrido MMQ_ARTMAP *Fuzzy* (ALVES, 2019) em um processo de previsão de cargas aplicados a diversos níveis de desagregação. A aplicação desse modelo híbrido se justifica pela sua capacidade de modelar relações complexas e não lineares nos dados. Ao avaliar os resultados dessa rede em diferentes níveis de desagregação, foi possível verificar sua eficácia na previsão de cargas, proporcionando uma abordagem promissora para lidar com a complexidade das previsões em diversos níveis.

8.2 SUGESTÕES PARA TRABALHOS FUTUROS

Os resultados alcançados por esta tese podem ser considerados relevantes (Redução até um 60% do MAPE comparado aos *benchmarks*). No entanto, há sugestões de melhorias e procedimentos que podem ser considerados para tornar as metodologias ainda mais eficientes.

Algumas sugestões para futuras melhorias são:

- Avaliar e comparar os resultados obtidos por meio da utilização de outras funções *Wavelet* no processo de decomposição da base de dados, *i.e.*, *Molet*, *Coiflets*, *Meyer*, *Symlets* e *Splines* bi-ortogonais podem ser testadas nas metodologias 2 e 3.
- Ampliar o número de avaliações para níveis de desagregação com cargas ainda mais flutuantes, a fim de analisar o desempenho das seis metodologias apresentadas neste trabalho. Além disso, é importante comparar os resultados com a inclusão de variáveis exógenas de alta correlação, assim como foi feito com as variáveis meteorológicas, para alcançar resultados ainda melhores.
- Empregar algoritmos de treinamento por reforço, como o *Q-Learning*, na estrutura interna das metodologias com o intuito de melhorar ainda mais os resultados.

- Explorar métodos alternativos de *Machine Learning* e *Deep Learning* na base de dados, a fim de comparar o desempenho desses métodos com as metodologias propostas.
- Avaliar detalhadamente o problema da regressão de categorias, observado na previsão de cargas de grandes usuários livres na área central, e buscar soluções que propiciem confiabilidade à metodologia 1 e para o *benchmark 2*.
- Avaliar as metodologias abordadas nos problemas relacionados com temas de alta volatilidade e incertezas, como a previsão de preços de energia, a previsão de geração de energia renovável, previsão de falhas em equipamento elétricos e a estimação de perdas não técnicas nos sistemas de distribuição de energia.

REFERÊNCIAS

- ABREU, T. *et al.* Electrical customer profile using *Fuzzy* logic theory. **IEEE Latin America Transactions**, New York, v. 18, n. 8, p. 1353-1361, 2020.
- ALVES, M. **Previsão de cargas não residenciais mistas por redes neurais ARTMAP Fuzzy**. 2019. 92 f. Tese (Doutorado em Engenharia) - Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista, Ilha Solteira, 2019.
- AGUILAR, E.; ANTONIO, N. Short-Term electricity load forecasting with machine learning. **Information**, Basel, v. 12, n. 2(50), p. 1-21, 2021.
- AKANSU, A. N.; SERDIJN, W. A.; SELESNICK, I. W. Emerging applications of *Wavelets*: A review. **Physical communication**, Netherlands, v. 3, n. 1, p. 1-18, 2010.
- BOX, G. E.; JENKINS, G. M. **Times series analysis: forecasting and control**. San Francisco: Holden-Day, 1976.
- BURKOV, A. **Machine learning engineering**. Canada: True Positive Incorporated, 2020.
- CARPENTER, G. A; GROSSBERG, S. The ART of adaptive pattern recognition by a self-organizing neural network. **Computer**, Piscataway, v. 21, n. 3, p. 77-88, 1988.
- CARPENTER, G. A.; GROSSBERG, S.; REYNOLDS, J. H. ARTMAP: supervised real-learning and classification of non-stationary data by a self-organizing neural network. **Neural Network**, New York, v. 4, n. 5, p. 565-588, 1991a.
- CARPENTER, G. A.; GROSSBERG, S.; ROSEN D. B. *Fuzzy* ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. **Neural Network**, New York, v. 4, n. 1, p. 759-771, 1991b.
- CARPENTER, G. A. *et al.* *Fuzzy* ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. **IEEE Transactions on Neural Networks**, New York, v. 3, n. 5, p. 698-713, 1992.
- CORTÉS, J. A.; GARZÓN, H. B. C.; CHAVES, J. A. Del análisis de Fourier a las Wavelets-Transformada Continua Wavelet (CWT). **Scientia et technica**, Risaralda, vol. 1, n. 37, 2007.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Norwell, v. 20, p. 273-297, 1995.
- DEBNATH, L.; SHAH, F. A. **Wavelet transforms and their applications**. 2nd ed. Boston: Birkhäuser, 2015.
- DIMITROULIS, P.; ALAMANIOTIS, M. A Fuzzy logic energy management system of on-grid electrical system for residential prosumers. **Electric Power Systems Research**, Netherlands, v. 202, p. 107621, 2022.
- DONG, X. *et al.* A survey on ensemble learning. **Frontiers of Computer Science**, China, v. 14, p. 241-258, 2020.

DUA, S.; DU, X. **Data mining and machine learning in cybersecurity**. New York: CRC Press, 2011.

ESEYE, A. T. *et al.* Machine learning based integrated feature selection approach for improved electricity demand forecasting in decentralized energy systems. **IEEE Access**, Piscataway, v. 7, p. 91463-91475, 2019.

FALLAH, S. N. *et al.* Computational intelligence on short-term load forecasting: A methodological overview. **Energies**, Basel, v. 12, n. 3, p. 393-414, 2019.

FEINBERG, E. A.; GENETHLIOU, D. Load Forecasting. *In.*: CHOW, J. H.; WU, F. F.; MOMOH, J. A. **Applied mathematics for restructured electric power systems**. New York: Springer, 2005. p. 268-285.

FUMO, N.; BISWAS, M. R. Regression analysis for prediction of residential energy consumption. **Renewable and sustainable energy reviews**, Oxford, v. 47, p. 332-343, 2015.

GROSSMANN, A.; MORLET, J. Decomposition of Hardy functions into square integrable *Wavelets* of constant shape. **SIAM journal on mathematical analysis**, Philadelphia, v. 15, n. 4, p. 723-736, 1984.

HAHN, H.; Meyer-Nieberg, S.; Pickl, S. Electric load forecasting methods: Tools for decision making. **European Journal of Operational Research**, Amsterdam, v. 199, n. 3, p. 902-907, 2009.

HAYKIN, S. **Neural networks and learning machines**. 3rd ed. New Jersey: Prentice-Hall, 2008.

HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, Amsterdam, v. 22, n.4, p. 679-688, 2006.

ISLAM, M. *et al.* Chapter 5 - Energy demand forecasting. *In.*: HASANUZZAMAN, M.; ABD RAHIM, N. **Energy for sustainable development: demand, supply, conversion and management**. Cambridge: Academic Press, 2019.

ISLAM, S. M.; AL-ALAWI, S. M. Principles of electricity demand forecasting. Part 2: Applications. **Power Engineering Journal**, Herts, v. 11, n. 2, p. 91-95, 1997.

JAMES, G. *et al.* **An introduction to statistical learning**. New York: springer, 2013.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. **Electronic Markets**, Switzerland, v. 31, n. 3, p. 685-695, 2021.

JIMÉNEZ, C. *et al.* Aplicación de la transformada *Wavelet* para el análisis de señales de voz normales y patológicas, **Revista INGENIERÍA UC**, Carabombo, v. 15, n. 1, p. 7-13, 2008.

JIN, N. *et al.* Highly accurate energy consumption forecasting model based on parallel LSTM neural networks. **Advanced Engineering Informatics**, Amsterdam, v. 51, p. 101442-101459, 2022.

KARAYIANNIS, N.; VENETSANOPOULOS, A. N. **Artificial neural networks: learning algorithms, performance evaluation, and applications**. New York: Springer, 1993.

KAVOUSIFARD, A.; SAMET, H. Consideration effect of uncertainty in power system reliability indices using radial basis function network and fuzzy logic theory. **Neurocomputing**, Amsterdam, v. 74, n. 17, p. 3420-3427, 2011.

KING, R. D.; ORHOBOR, O. I.; TAYLOR, C. C. Cross-validation is safe to use. **Nature Machine Intelligence**, London, v. 3, n. 4, p. 276-276, 2021.

KOCAARSLAN, I.; ÇAM, E. *Fuzzy* logic controller in interconnected electrical power systems for load-frequency control. **International Journal of Electrical Power & Energy Systems**, Oxford, v. 27, n. 8, p. 542-549, 2005.

KONG, W. et al. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. **IEEE Transactions on Smart Grid**, New York, v. 10, n. 1, p. 841-851, 2019.

KROPOSKI, B. *et al.* Achieving a 100% Renewable Grid: Operating Electric Power Systems with Extremely High Levels of Variable Renewable Energy. **IEEE Power and Energy Magazine**, New York, v. 15, n. 2, p. 61-73, 2017.

KYRIAKIDES, E.; POLYCARPOU, M. Short Term Electric Load Forecasting: A Tutorial. *In.*: CHEN, K.; WANG, L. **Trends in Neural Computation**. Berlin: Springer, 2007. p. 391-419.

LAKEMEYER, G.; NEBEL, B. **Exploring Artificial Intelligence in the New Millennium**. Massachusetts: Morgan Kaufmann, 2002.

LI, L. *et al.* Short-term apartment-level load forecasting using a modified neural network with selected auto-regressive features. **Applied Energy**, London, v. 287, p. 116509-116523, 2021.

LORENA, A. C.; DE CARVALHO, A. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 14, n. 2, p. 43-67, 2007.

MALLAT, S. G. Multiresolution approximations and *Wavelet* orthonormal bases of $L^2(\mathbb{R})$. **Transactions of the American mathematical society**, Rhode Island, v. 315, n. 1, p. 69-87, 1989a.

MALLAT, S G. A theory for multiresolution signal decomposition: the *Wavelet* representation. **IEEE transactions on pattern analysis and machine intelligence**, Piscataway, v. 11, n. 7, p. 674-693, 1989b.

MAMLOOK, R; BADRAN, O; ABDULHADI, E. A *Fuzzy* inference model for short-term load forecasting. **Energy Policy**, Kidlington, vol. 37, n. 4, p. 1239-1248, 2009.

MERRY, R. J. E.; STEINBUCH, M. **Wavelet theory and applications**. 2005. 40f. Literature study - Eindhoven University of technology, Department of mechanical engineering, Control systems technology group, Eindhoven, 2005.

MEYER, Y. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. **Séminaire Bourbaki**, Marseille, v. 662, p. 1985–1986, 1985.

MORENO, A. L. **Análise de estabilidade transitória via rede neural ART-ARTmap Fuzzy Euclidiana modificada com treinamento continuado**. 2010. 143 f. Tese (Doutorado em Engenharia) - Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista, Ilha Solteira, 2010.

MORLET, J, et al. Wave propagation and sampling theory-Part II: Sampling theory and complex waves. **Geophysics**, Worldwide, v. 47, n. 2, p. 222-236, 1982.

MOSAVI, A. *et al.* State of the art of machine learning models in energy systems, a systematic review. **Energies**, Basel, v. 12, n. 7, p. 1301-1343, 2019.

MÜLLER, M.R. et al. Electrical load forecasting in disaggregated levels using *Fuzzy* ARTMAP artificial neural network and noise removal by singular spectrum analysis. **SN Appl. Sci.**, v.2, n. 1218, p.1-10, 2020.

MUNASINGHE, M. **Energy analysis and policy**: selected works. Oxford: Butterworth-Heinemann, 2013.

MURPHY, K. P. **Machine Learning**: A Probabilistic Perspective. Massachusetts: MIT press, 2012.

MURTO, P. **Neural network models for short-term load forecasting**. 1998. 90 f. (Dissertação de Mestrado) - Departamento de Física e Matemática Aplicada, Helsinki University of Technology, Finland, 1998.

NOURANI, V. *et al.* Multi-step-ahead solar irradiance modeling employing multi-frequency deep learning models and climatic data. **Applied Energy**, London, v. 315, p. 119069-119088, 2022.

OGIMOTO, K.; WANI, H. Making Renewables Work: Operational Practices and Future Challenges for Renewable Energy as a Major Power Source in Japan. **IEEE Power and Energy Magazine**, New York, v. 18, n. 6, p. 47-63, 2020.

PARK, D. C.; EL-SHARKAWI, M. A.; MARKS II, R. J.; ATLAS, E.; DAMBORG, M. J. Electric Load Forecasting Using an Artificial Neural networks. **IEEE Transactions on Power Systems**, New York, v. 6, n. 2, p. 442-449, 1991.

PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine**, v. 2, n. 11, p. 559–572, 1901.

PENG, Y. *et al.* Short-term Load Forecasting at Different Aggregation Levels with Predictability Analysis. *In*: 2019 IEEE INNOVATIVE SMART GRID TECHNOLOGIES, 2019, Chengdu, **Proceedings** [...] Chengdu: IEEE, 2019. p. 3385-3390.

PERU (País). Reglamento de Ley de las Concesiones Eléctricas (RLCE), Decreto supremo n° 009-93-EM, Lima, 1993.

SAJJAD, M. *et al.* A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting. **IEEE Access**, Piscataway, v. 8, p. 143759-143768, 2020.

SAMMUT, C.; WEBB, G. I. 2nd ed. **Encyclopedia of machine learning**. Berlin: Springer Science+Business Media, 2017.

SHENG, Y. **Wavelet transform**: The transforms and applications handbook Series. New York: CRC Press, 1996, p. 747-827.

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

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. *Statistics and computing*, New York, v. 14, p. 199-222, 2004.

SUGANTHI, L.; SAMUEL, A. A. Energy models for demand forecasting—A review. **Renewable and sustainable energy reviews**, Oxford, v. 16, n. 2, p. 1223-1240, 2012.

SUGANTHI, L.; INIYAN, S.; SAMUEL, A. A. Applications of *Fuzzy* logic in renewable energy systems—a review. **Renewable and sustainable energy reviews**, Oxford, v. 48, p. 585-607, 2015.

TUKEY, J. A. **Exploratory data analysis**. London: Pearson, 1977.

VAPNIK, V; CHERVONENKIS, A. **Theory of pattern recognition**. Moscow: Nauka, 1974.

VAPNIK, V; GOLOWICH, S; SMOLA, A. Support vector method for function approximation, regression estimation and signal processing. In: *Advances in neural information processing systems*, 9., 1996, Denver. **Proceeding** [...] Cambridge: The MIT Press, 1997. p. 281-287.

VETTERLI, M; HERLEY, C. *Wavelets* and filter banks: Theory and design. **IEEE transactions on signal processing**, Piscataway, v. 40, p. 2207-2232, 1992.

WANG, Y, *et al.* Short-term industrial load forecasting based on ensemble hidden Markov model. **IEEE Access**, Piscataway, v. 8, p. 160858-160870, 2020.

WANG, Z.; WANG, Y.; SRINIVASAN, R. S. A novel ensemble learning approach to support building energy use prediction. **Energy and Buildings**, Amsterdam, v. 159, p. 109-122, 2018.

WERON, R. **Modeling and forecasting electricity loads and prices**: A statistical approach. Hoboken: John Wiley & Sons, 2007.

WOOD, A. J.; WOLLENBERG, B. F.; SHEBLÉ, G. B. **Power generation, operation, and control**. Hoboken: John Wiley & Sons, 2013.

YAN, K. *et al.* A hybrid LSTM neural network for energy consumption forecasting of individual households. **IEEE Access**, Piscataway, v. 7, p. 157633-157642, 2019.

YANG, Y. *et al.* A novel decompose-cluster-feedback algorithm for load forecasting with hierarchical structure. **International Journal of Electrical Power & Energy Systems**, New York, v. 142, p. 108249-108262, 2022.

YING, X. An overview of overfitting and its solutions. In: Journal of physics - Conference series, 1168., 2018, Daqing, **Proceeding** [...] Bristol: IOP Publishing, 2019.

YOUNG, N. An introduction to Hilbert space. Cambridge: Cambridge university press, 1988.

ZADEH, L. A. *Fuzzy logic*. **Computer**, IEEE, vol. 21, no 4, p. 83-93, 1988.

ZHANG, X.; Y. WANG, Y.; HAN, R. P. S. *Wavelet transform theory and its application in EMG signal processing*. In: 2010 Seventh international conference on *Fuzzy systems and knowledge discovery*, Yantai, **Proceedings** [...] Yantai: IEEE, 2010. p. 2234-2238.

ZHOU, Z. H. Support Vector Machine. *In.*: **Machine Learning**. Singapor: Springer, 2021.