

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM MESTRADO EM ENGENHARIA**  
**ELÉTRICA E INFORMÁTICA INDUSTRIAL**

**HELLEN CRISTINA ANCELMO**

**DESENVOLVIMENTO DE BASE DE DADOS E APLICAÇÃO DO**  
**MÉTODO DE PRONY PARA EXTRAÇÃO DE CARACTERÍSTICAS E**  
**CLASSIFICAÇÃO DE CARGAS ELÉTRICAS**

**DISSERTAÇÃO**

**CURITIBA**

**2020**

**HELLEN CRISTINA ANCELMO**

**DESENVOLVIMENTO DE BASE DE DADOS E APLICAÇÃO DO MÉTODO DE PRONY PARA EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO DE CARGAS ELÉTRICAS**

**Development of Database and Application of the Prony's Method for Feature Extraction and Classification of Electric Loads**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, da Universidade Tecnológica Federal do Paraná (UTFPR), como requisito parcial para obtenção do título de "Mestre em Ciências" - Área de Concentração: Engenharia de Computação.  
Orientador: André Eugenio Lazzaretti.  
Coorientador: Elder Oroski.

**CURITIBA**

**2020**



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.



**Ministério da Educação  
Universidade Tecnológica Federal do Paraná  
Câmpus Curitiba**



HELLEN CRISTINA ANCELMO

**DESENVOLVIMENTO DE BASE DE DADOS E APLICAÇÃO DO MÉTODO DE PRONY PARA EXTRAÇÃO DE  
CARACTERÍSTICAS E CLASSIFICAÇÃO DE CARGAS ELÉTRICAS**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestra Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Engenharia De Computação.

Data de aprovação: 18 de Novembro de 2020

Prof Andre Eugenio Lazzaretti, Doutorado - Universidade Tecnológica Federal do Paraná

Prof.a Fabiana Pottker, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Miguel Moreto, Doutorado - Universidade Federal de Santa Catarina (Ufsc)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 19/11/2020.

## **AGRADECIMENTOS**

Aos professores que me guiaram nessa jornada com dedicação, apoio e amizade, meu orientador André Eugenio Lazzaretti, coorientador Elder Oroski, bem como os demais professores do laboratório Douglas Paulo Bertrand Renaux e Fabiana Pottker.

Aos meus amigos de laboratório Flávio Lori Grando, Bruna Machado Mulinari, Clayton Hilgemberg da Costa, Guilherme Dan Yamada e Marcelo Paludetto Rodrigues, por toda ajuda, conversas e risadas.

À toda minha família, em especial meus pais Ivone e Licindo Ancelmo, por serem meu porto seguro em todas as tempestades enfrentadas e o alento nos momentos de calmaria.

À Agência Nacional de Energia Elétrica (ANEEL) e à Companhia Paranaense de Energia Elétrica (COPEL) pelo suporte. Este trabalho teve financiamento via projeto de Pesquisa & Desenvolvimento dentro do programa da ANEEL, código PD2866-0464/2017.

Na vida, não existe nada a temer, mas a  
entender. (Marie Curie).

## RESUMO

ANCELMO, Hellen Cristina. **Desenvolvimento de Base de Dados e Aplicação do Método de Prony para Extração de Características e Classificação de Cargas Elétricas**. 2020. 82 f. Dissertação (Mestrado em Engenharia Elétrica e Informática Industrial) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Diante da crescente demanda por energia no mundo todo, pesquisadores vêm desenvolvendo diferentes soluções de gerenciamento e conservação de energia. Técnicas como as de monitoramento não intrusivo de cargas podem auxiliar na avaliação mais detalhada do consumo, permitindo aplicar medidas de eficiência energética, reduzindo o consumo de determinados equipamentos com maior impacto no consumo global de energia. Neste tipo de abordagem, quatro principais fases são consideradas: coleta de dados, detecção de eventos, extração de características e identificação da carga. Neste contexto, este trabalho tem como proposta duas frentes principais: o desenvolvimento de uma base de dados simulada baseada em cargas residenciais e a aplicação do método matemático de Prony para extração de características. Na base de dados, as cargas foram modeladas em um *software* de transitórios elétricos, buscando sanar as principais limitações em bases de dados existentes na literatura, como: identificação de eventos de forma precisa (a nível de amostra), classes balanceadas e inserção de ruídos e harmônicos na rede. Como resultado foi obtida uma base de dados com seis sub-bases, com um cenário ideal, com a presença de indutância parasita, harmônicos na rede e diferentes níveis de ruído. Em seguida, utilizando métodos de extração e classificação de cargas da literatura foi possível analisar que a presença de não idealidades interfere diretamente na atuação do extrator de características e na identificação da carga. No desenvolvimento do extrator de características a partir do método de Prony, cinco diferentes soluções matemáticas do método foram comparadas: polinomial, mínimos quadrados, mínimos quadrados totais, *Matrix Pencil* e com base em filtros com resposta ao impulso de duração infinita. Ao analisar os resultados de classificação a partir deste método utilizando diferentes classificadores (k-Vizinhos mais Próximos, Árvores de Decisão, *Ensemble*, Análise de Discriminantes Lineares e Máquina de Vetor Suporte), os resultados alcançaram valores superiores 90% de acurácia na grande maioria dos casos avaliados. Cabe ressaltar que essa análise foi feita com diferentes bases de dados públicas, além da base de dados aqui proposta, demonstrando a viabilidade do uso prático do método proposto no contexto de classificação de cargas por meio de monitoramento não intrusivo.

**Palavras-chave:** NILM. Base de Dados Simulada. Método de Prony. Classificação de Cargas Elétricas. Extração de Características.

## ABSTRACT

ANCELMO, Hellen Cristina. **Development of Database and Application of the Prony's Method for Feature Extraction and Classification of Electric Loads**. 2020. 82 p.  
Dissertation (Master's in Course Name) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

In face of the growing demand for energy worldwide, researchers have been developing different energy management and conservation systems. Techniques such as non-intrusive load monitoring can assist in a more detailed assessment of consumption, allowing the application of energy efficiency measures, reducing the consumption of certain equipment with a greater impact on global energy consumption. In this type of approach, four main phases are considered: data collection, event detection, feature extraction, and load identification. In this context, this work proposes two main contributions: the development of a simulated database based on residential loads and the application of the Prony's mathematical method to extract features. Regarding the database, loads are modeled in a software that allows electrical transients modeling, seeking to improve the main limitations in existing databases in the literature, such as: precise identification of events (at sample level), balanced classes, and insertion of noise and harmonics in the waveforms. As a result, a database with six sub-bases is obtained with: an ideal scenario; the presence of leakage inductance; harmonics in the network; and different noise levels. Then, using the extraction methods and load classification from the literature, it is possible to analyze whether the presence of non-idealities directly interferes on the performance of the feature extractor and load identification. In the development of the feature extractor using the Prony's method, five different mathematical solutions are compared: polynomial, least squares, total least squares, matrix pencil and based on infinite response filters. By analyzing the classification results using this method with different classifiers (k-Nearest Neighbors, Decision Trees, Ensemble, Linear Discriminant Analysis, and Support Vector Machine), the results reached values above 90% accuracy in the vast majority of evaluated cases. It is worth mentioning that this analysis was carried out with different public databases, in addition to the database proposed here, demonstrating the feasibility of the practical use of the proposed method in the context of load classification through non-intrusive monitoring.

**Keywords:** NILM. Simulated Database. Prony's Method. Classification of Electric Loads. Feature Extraction.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Descritivo de Regime Transitório e Regime Permanente. . . . .	20
Figura 2 – Corrente de <i>Inrush</i> da Furadeira Elétrica. . . . .	29
Figura 3 – Fases do método de Prony. . . . .	29
Figura 4 – Corrente de aquisição - Removedor de Tinha Elétrico. Seção A - Destaque para o momento em que a carga liga. Seção B - Destaque para o momento em que a carga desliga. . . . .	38
Figura 5 – Corrente de aquisição - Aspirador de Pó. . . . .	39
Figura 6 – Corrente de aquisição individual - Furadeira Velocidade 1. . . . .	41
Figura 7 – Corrente de aquisição dupla - Lâmpada Incandescente e Aquecedor a Óleo. Seção A - Destaque para o instante em que a Lâmpada Incandescente liga. Seção B - Destaque para o instante em que o Aquecedor a Óleo liga. Seção C - Destaque para o instante em que a Lâmpada Incandescente desliga. Seção A - Destaque para o instante em que o Aquecedor a Óleo desliga. . . . .	42
Figura 8 – Diagrama Geral da Base de Dados. . . . .	42
Figura 9 – Diagrama das cargas (a) Resistor, (b) Resistor Indutivo (c) Retificador a Diodo com Resistor, (d) Ponte Completa a Diodo com Resistor, (e) Retificador a Tiristor com Resistor, (f) Retificador a Tiristor com Resistor Indutivo. . . . .	43
Figura 10 – Modelo matemático do Motor Universal. . . . .	44
Figura 11 – Processo de Análise das Formas de Onda. . . . .	46
Figura 12 – Subtração de Cargas do Sinal de Corrente - Retificador e Ponte a Diodo. . . . .	47
Figura 13 – Treinamento de Classificação utilizando o Método de Prony. . . . .	49
Figura 14 – Procedimento de Teste utilizando o Método de Prony. . . . .	50
Figura 15 – Comparativo do Sinal de Tensão da Rede e o Sinal de Tensão da base LIT Simulada. . . . .	53
Figura 16 – Corrente da Ponte Completa a Diodo e Tensão da Rede. . . . .	54
Figura 17 – Comparação das Formas de Onda de Corrente Real e Simulada do Motor Universal. . . . .	56
Figura 18 – Forma de Onda da Corrente com Resistor Indutivo e Motor Universal. Seção A - Destaque para o instante em que o Resistor Indutivo liga. Seção B - Destaque para o instante em que o Motor Universal liga. Seção C - Destaque para o instante em que o Resistor Indutivo desliga. Seção D - Destaque para o instante em que o Motor Universal desliga. . . . .	57
Figura 19 – Análise das Frequências Relevantes (Conteúdo Harmônico). . . . .	58
Figura 20 – Reconstrução do método de Prony com 3, 5 e 10 ciclos. . . . .	58
Figura 21 – Reconstrução do sinal de corrente da Carga Base COOLL a partir do Método de Prony utilizando o método <i>Matrix Pencil</i> , ordem 10. . . . .	59
Figura 22 – Reconstrução da Carga Base LIT a partir do Método de Prony utilizando o método <i>Matrix Pencil</i> , ordem 10. . . . .	60
Figura 23 – Reconstrução do Sinal de Corrente da Carga Secador de Cabelo a partir do Método de Prony na Base LIT. . . . .	60

## LISTA DE TABELAS

Tabela 1 – Comparação entre as bases de dados reais. . . . .	26
Tabela 2 – Base de Dados COOLL. . . . .	37
Tabela 3 – Características do <i>subset</i> Sintético - LIT <i>Dataset</i> . . . . .	40
Tabela 4 – Ângulos de Disparo . . . . .	40
Tabela 5 – Combinação de Cargas . . . . .	40
Tabela 6 – Parâmetros das Cargas . . . . .	45
Tabela 7 – Configurações da Base de Dados . . . . .	46
Tabela 8 – Parâmetros dos componentes reais. . . . .	52
Tabela 9 – Parâmetros da Tensão da Rede. . . . .	53
Tabela 10 – Validação dos Parâmetros - Valores Máximos de Corrente e Tensão. . . . .	55
Tabela 11 – Validação dos Parâmetros (FP e EQM). . . . .	55
Tabela 12 – Parâmetros do Motor Universal. . . . .	56
Tabela 13 – Reconstrução do método de Prony. . . . .	59
Tabela 14 – EQM para a Base COOLL. . . . .	61
Tabela 15 – EQM para a Base LIT Sintética. . . . .	61
Tabela 16 – Acurácia do Método Ensemble vs Base de Dados. . . . .	62
Tabela 17 – Acurácia de Classificação para Diferentes Números de Ciclos. . . . .	63
Tabela 18 – Acurácia Média <i>Ensemble</i> - Base COOLL. . . . .	63
Tabela 19 – Acurácia Média <i>Ensemble</i> - Base LIT Sintética. . . . .	64
Tabela 20 – Comparativo das Abordagens do Método de Prony. . . . .	64
Tabela 21 – Comparativo de Acurácia de Classificação. . . . .	66
Tabela 22 – Comparativo de Acurácia por Classe - Base COOLL. . . . .	67
Tabela 23 – Comparativo de Acurácia por Classe - Base LIT Sintética. . . . .	68
Tabela 24 – Comparativo de Acurácia por Classe - Base LIT Simulada. . . . .	69
Tabela 25 – Acurácia para Testes de Generalização de Testes. . . . .	69
Tabela 26 – Tempo de execução dos Extratores de Características. . . . .	69
Tabela 27 – Tempo de Execução da Classificação por Métodos Extratores versus Classificadores. . . . .	70

## LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS

### SIGLAS

DCV	Dispositivos Continuamente Variáveis
DFT	Transformada Discreta de Fourier, do inglês <i>Discrete Fourier Transform</i>
DT	Árvores de Decisão, do inglês <i>Decision Tree</i>
DWT	Transformada Discreta de Wavelet, do inglês <i>Discrete Wavelet Transform</i>
ENS	<i>Ensemble</i>
EQM	Erro Quadrático Médio
FP	Fator de Potência
IoT	Internet das Coisas, do inglês <i>Internet of Things</i>
kNN	K Vizinhos mais Próximo, do inglês <i>k-Nearest Neighbors</i>
LDA	Análise de Discriminantes Lineares, do inglês <i>Linear Discriminant Analysis</i>
MEF	Máquinas de Estado Finito
MP	<i>Matrix Pencil</i>
MQ	Mínimos Quadrados
MQT	Mínimos Quadrados Totais
NTC	Semicondutor sensível à temperatura, do inglês <i>Negative Temperature Coefficient</i>
SVD	Decomposição de valor singular, do inglês <i>singular value decomposition</i>
SVM	Máquina Vetor Suporte, do inglês <i>Support Vector Machine</i>

### ACRÔNIMOS

ALM	Métodos de Monitoramento de Cargas, do inglês <i>Appliances Load Monitoring</i>
BLOND	<i>a Building-Level Office eNvironment Dataset</i>
BLUED	<i>Building-Level fUlly-labeled dataset for Electricity Disaggregation</i>
COOLL	<i>Controll On / Off Loads Library</i>
IIR	Resposta ao Impulso Infinita, do inglês <i>Infinitive Impulse Response</i>
ILM	Monitoramento Intrusivo de Cargas, do inglês <i>Intrusive Load Monitoring</i>
NILM	Monitoramento Não Intrusivo de Cargas, do inglês <i>Non Intrusive Load Monitoring</i>
PLAID	<i>Plug Load Appliance Identification Dataset</i>
REDD	<i>Reference Energy Disaggregation Data Set</i>
SynD	<i>Synthetic Energy Consumption</i>
UK-DALE	<i>United Kingdom recording Domestic Appliance-Level Electricity</i>
WHITED	<i>Worldwide Household and Industry Transient Energy Data Set</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
1.1	JUSTIFICATIVA	14
1.2	OBJETIVOS	15
1.2.1	Objetivo Geral	15
1.2.2	Objetivos Específicos	15
1.3	ESTRUTURA DO TRABALHO	16
<b>2</b>	<b>ESTADO DA ARTE</b>	<b>17</b>
2.1	MONITORAMENTO NÃO INTRUSIVO DE CARGAS	17
2.2	MÉTODOS DE EXTRAÇÃO	19
2.3	MÉTODOS DE CLASSIFICAÇÃO	22
2.4	BASE DE DADOS	24
2.5	CONTRIBUIÇÕES DESTE TRABALHO	27
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>28</b>
3.1	MÉTODO DE PRONY	28
3.1.1	Método de Prony Clássico	30
3.1.2	Método de Prony com Mínimos Quadrados	31
3.1.3	Método de Prony com Mínimos Quadrados Totais	32
3.1.4	Método de Prony usando <i>Matrix Pencil</i>	32
3.1.5	Método de Prony com Filtros IIR	33
3.2	CLASSIFICADORES	34
3.2.1	Método K Vizinhos mais Próximo, do inglês <i>k-Nearest Neighbors</i> (kNN)	34
3.2.2	Método Árvores de Decisão	35
3.2.3	Método <i>Ensemble</i>	35
3.2.4	Análise de Discriminantes Lineares	35
3.2.5	Máquina de Vetor de Suporte	36
<b>4</b>	<b>METODOLOGIA</b>	<b>37</b>
4.1	BASE COOLL	37
4.2	BASE LIT SINTÉTICA	38
4.3	BASE LIT SIMULADA	40
4.3.1	Cargas	41
4.3.2	Geração das formas de onda	44
4.3.3	Configurações da Base de Dados	45
4.4	ANÁLISE DAS FORMAS DE ONDA	46
4.4.1	Pré-Processamento e Extração de Características	47
4.5	CLASSIFICAÇÃO	48
<b>5</b>	<b>RESULTADOS</b>	<b>51</b>
5.1	BASE LIT SIMULADA	52
5.2	MÉTODO DE PRONY	56
5.2.1	Análise de Frequência de Amostragem e Número de Ciclos	57
5.2.2	Métodos de Prony Reduzido	59
5.3	RESULTADOS DE CLASSIFICAÇÃO	61

5.3.1	Base de Dados LIT Simulada . . . . .	61
5.3.2	Análise do Número de Ciclos . . . . .	62
5.3.3	Métodos de Prony Reduzido . . . . .	63
5.4	COMPARATIVO DAS ABORDAGENS DO MÉTODO DE PRONY .	64
5.5	COMPARAÇÃO COM MÉTODOS DA LITERATURA . . . . .	65
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>71</b>
6.1	TRABALHOS FUTUROS . . . . .	72
6.2	TRABALHOS PUBLICADOS . . . . .	72
	<b>REFERÊNCIAS . . . . .</b>	<b>75</b>

## 1 INTRODUÇÃO

Atualmente a conservação e gerenciamento de energia têm sido questões desafiadoras, visto a crescente demanda de energia em diversos setores do mercado e a popularização do conceito de *Smart Grids* (Redes de Geração, Transmissão e Distribuição Inteligentes de Energia). Neste contexto, os pesquisadores estão trabalhando para desenvolver soluções tecnológicas que supram esta necessidade (ZOHA *et al.*, 2012).

Um elemento fundamental das *Smart Grids* é a gestão de energia (TABATABAEI *et al.*, 2017), tanto do ponto de vista da produção (por meio da geração distribuída) quanto do consumo. Tal tecnologia pode ajudar as concessionárias de energia a controlar com eficiência a demanda de energia do consumidor nos momentos de pico de carga, causando impacto direto na economia e no uso eficiente de energia (ESA *et al.*, 2016).

Para realizar o monitoramento das cargas, existem dois tipos de métodos, do inglês *Appliances Load Monitoring* ALM: o monitoramento intrusivo de carga, do inglês, *Intrusive Load Monitoring* (ILM) e o monitoramento não intrusivo de carga, do inglês *Non Intrusive Load Monitoring* (NILM).

O método mais usual para monitoramento de cargas é a técnica não intrusiva, NILM, que consiste em um sistema que permite averiguar o consumo individual de energia de equipamentos conectados a um determinado barramento, medindo variáveis elétricas em um único ponto do próprio barramento (HART, 1992). Esse sistema desagrega os dados de consumo global, individualizando o consumo de cada equipamento conectado no mesmo barramento elétrico (ESA *et al.*, 2016).

Para resolver o problema de desagregação, duas abordagens são geralmente propostas na literatura (FIGUEIREDO *et al.*, 2012): métodos de otimização e reconhecimento de padrões. No caso de otimização, o problema de desagregação de carga é definido como uma minimização de uma função objetivo que compara (por meio de resíduos) a carga desconhecida composta com um conjunto de possíveis candidatos extraídos de um banco de dados conhecido que, quando combinados, fornecem o menor resíduo em comparação com a carga composta. Já no caso de reconhecimento de padrões, o banco de dados de assinatura de cargas conhecidas é usado para estimar um modelo capaz de classificar quais cargas estão presentes em um sinal composto desconhecido. Em ambos os métodos, é necessário extrair um conjunto de características gerais que podem ser obtidas por meio de medições convencionais para caracterizar a assinatura de

potência dos equipamentos. As formas mais usuais incluem (KAHL *et al.*, 2017): (i) formas de onda de corrente e tensão; (ii) potência ativa e reativa; (iii) admitância instantânea; (iv) análise harmônica; (v) transientes de comutação de carga (com base, por exemplo, na Transformada Wavelet e na Transformada S); (vi) Transformada de Fourier em janela); e (vii) trajetória tensão-corrente V-I. Normalmente, essas características são extraídas durante a ocorrência de instantes de chaveamento de carga (UKIL; ŽIVANOVIĆ, 2008).

Recentemente, propostas de classificação de cargas têm focado na combinação de extração de características transitórias e de regime permanente do sinal, como em Meziane *et al.* (2017), Nait-Meziane *et al.* (2019), Renaux *et al.* (2018), Mulinari *et al.* (2019). No entanto, essas abordagens fazem uso de uma combinação de diferentes técnicas computacionais para extrair características de cada estado (transitório ou regime permanente), elevando o custo computacional e a complexidade do método como um todo. Como exemplo, em Meziane *et al.* (2017), Nait-Meziane *et al.* (2019) é apresentado um estudo de classificação de cargas por meio de um método de extração de características baseado em uma modificação da Transformada de Fourier, incorporando um amortecimento exponencial. Os autores demonstram que este amortecimento permite obter desempenhos de identificação superiores a 90% para a base de dados proposta. Porém, a implementação do método apresenta uma instabilidade na etapa de regressão exponencial (HUA; SARKAR, 1990), alterando a forma de onda na porção senoidal, estimando conteúdo harmônico impreciso, comprometendo a classificação como um todo, particularmente para cargas múltiplas.

Por outro lado, a seleção e comparação dos métodos de extração e classificação de características depende muito da seleção do conjunto de dados (RENAUX *et al.*, 2018). Existem diferentes conjuntos de dados para análise de assinatura de energia, adquiridos em condições reais, como REDD (*Reference Energy Disaggregation Data Set*), COOLL (*Controll On / Off Loads Library*), UK-DALE (*United Kingdom recording Domestic Appliance-Level Electricity*), BLUED (*Building-Level fully-labeled dataset for Electricity Disaggregation*), WHITED (*Worldwide Household and Industry Transient Energy Data Set*), PLAID (*Plug Load Appliance Identification Dataset*), SynD (*Synthetic Energy Consumption*) e BLOND (*a Building-Level Office eNvironment Dataset*). Ainda, existem bases de dados simulados como apresentado em Venkatesh e Kumar (2008), Gillis *et al.* (2016), Lopez e Pouresmaeil (2019) e Collin *et al.* (2014). Essas bases de dados apresentam algumas limitações como falta de cargas múltiplas simultâneas e rótulos das cargas, baixa precisão na medida dos eventos de *ON* e *OFF*, ausência de análise de ruído e

conteúdo harmônico, o que dificulta o desenvolvimento e comparação de métodos de extração de características de assinatura de potência e classificação subsequente.

Tendo em vista as limitações dos métodos de extração de características, bem como das bases de dados, este trabalho se fundamenta em duas frentes. Inicialmente, com o objetivo de fornecer uma solução analítica para a extração de características transitórias e de regime permanente de um sinal, este trabalho propõe a estimativa de frequência, fase, amplitude e fator de amortecimento para sinais de corrente por meio do Método Prony (MARPLE, 1986). Usando essas componentes como características, diferentes classificadores de aprendizado de máquina são comparados, como *k-Nearest Neighbors* (kNN), Máquina Vetor Suporte, do inglês *Support Vector Machine* (SVM), a fim de realizar a classificação das formas de onda de diferentes bancos de dados, demonstrando a viabilidade do método proposto.

Adicionalmente, com base nas restrições dos conjuntos de dados existentes (reais e simulados), este trabalho apresenta um novo conjunto de dados simulado que controla o instante em que cada carga é trocada, permitindo extrair com precisão as características durante o transiente de cada carga. O conjunto de dados permite simular múltiplas cargas simultâneas, até sete tipos diferentes de cargas (linear, não linear com eletrônica de potência e motores universais), com quatro diferentes potências nominais para cada carga. Além disso, é possível incluir ruído e conteúdo harmônico nas simulações.

## 1.1 JUSTIFICATIVA

Segundo a Agência Internacional de Energia (IEA, 2019), a demanda mundial por eletricidade deve aumentar cerca de 2,1% ao ano e previsões apontam que ao menos 50% dessa energia será advinda de fontes renováveis. Com o intuito de adequar a demanda à produção de energia nos próximos anos, pesquisadores de diversas áreas têm trabalhado em soluções com foco em eficiência energética, que visem principalmente a redução de demanda de energia, reduzindo emissões de carbono e causando um menor impacto no meio ambiente.

Seguindo essa tendência, o Brasil tem apresentado estratégias de redução de consumo com base em programas de eficiência energética<sup>1</sup>, como substituição de aparelhos com alta consumo e construção de centrais elétricas a partir de fontes renováveis com a utilização de painéis fotovoltaicos, pequenas usinas hidrelétricas e turbinas eólicas.

<sup>1</sup> Destaca-se neste contexto o Programa de Eficiência Energética da Agência Nacional de Energia Elétrica, disponível em: <https://www.aneel.gov.br/programa-eficiencia-energetica>

Adicionalmente, o Protocolo Internacional de Medição e Verificação, desenvolvido em 1994, possui conceitos e opções para economia de energia e água. O documento sugere práticas para quantificação dos resultados de investimentos em eficiência energética, afim de incentivar as aplicações em gestão de demanda e projetos de energia renovável (Efficiency Valuation Organization, 2012).

Neste contexto, técnicas que abordem o monitoramento não intrusivo de cargas podem oferecer soluções para o desenvolvimento de processos para medição de economia de energia em projetos de eficiência, definidos pelo Protocolo Internacional de Medição e Verificação (Efficiency Valuation Organization, 2012), permitindo medições de energia por um período curto, com menor custo e mais representativo do que ocorre em uma dada instalação, além de também de atuar como ferramenta de gestão e previsão de demandas futuras (LAZZARETTI *et al.*, 2020).

## 1.2 OBJETIVOS

Dadas as limitações apresentadas na literatura e a necessidade técnica observada, os objetivos deste trabalho são detalhados a seguir.

### 1.2.1 Objetivo Geral

Implementar e analisar um método NILM para classificação de cargas elétricas, baseado no método de Prony, bem como propor uma base de dados simulada, para a extração de características e classificação de diferentes cargas.

### 1.2.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- Avaliar bases de dados públicas;
- Desenvolver uma base de dados simulada, levando em conta a limitação das bases públicas atuais;
- Avaliar diferentes formas de extração de características transitórias de formas de onda de tensão e corrente elétrica na literatura;

- Implementar as diferentes abordagens matemáticas do método de Prony para extração de características transitórias de formas de onda de tensão e corrente elétrica;
- Avaliar a melhor abordagem matemática do método de Prony para extração de características;
- Avaliar os métodos de Prony na classificação de cargas, utilizando bases de dados públicas;
- Avaliar os métodos de Prony na base de dados simulada;
- Comparar os métodos de Prony avaliados com o métodos do estado da Arte de classificação de cargas;
- Comparar complexidade computacional do método de Prony com os métodos do estado da Arte selecionados.

### 1.3 ESTRUTURA DO TRABALHO

Este trabalho é composto por seis capítulos. O segundo capítulo apresenta o estado da arte, com aplicações dos principais conceitos abordados neste trabalho. O terceiro e quarto capítulos apresentam os conceitos teóricos envolvidos e a metodologia das contribuições realizadas, respectivamente. Os resultados são apresentados no quinto capítulo, seguido pela conclusão do trabalho no sexto capítulo.

## 2 ESTADO DA ARTE

Neste capítulo serão apresentadas aplicações dos principais conceitos envolvidos neste trabalho. As seções seguem com: introdução aos conceitos de NILM, métodos de extração de características para assinatura de potência, métodos de classificação, bases de dados e principais contribuições deste trabalho.

### 2.1 MONITORAMENTO NÃO INTRUSIVO DE CARGAS

Devido ao crescente aumento pela demanda de energia, soluções tecnológicas vêm sendo desenvolvidas no âmbito da análise energética, levando ao foco pesquisas sobre métodos de análise e monitoramento de cargas de dispositivos elétricos. O objetivo é realizar a detecção detalhada de eventos de liga e desliga dos equipamentos residenciais, comerciais ou industriais, visando fornecer informações sobre consumo, permitindo assim, que sistemas automatizados realizem o planejamento de estratégias para economia de energia (ZOHA *et al.*, 2012).

Na literatura, existem duas principais abordagens para os Métodos de Monitoramento de Cargas, do inglês *Appliances Load Monitoring* (ALM): (i) Monitoramento Intrusivo de Cargas, do inglês *Intrusive Load Monitoring* (ILM); e (ii) Monitoramento Não Intrusivo de Cargas, do inglês *Non Intrusive Load Monitoring* (NILM). Ambos são métodos de monitoramento de cargas, sendo que a abordagem ILM exige um ou mais sensores por dispositivo para a detecção e a abordagem NILM exige apenas um único ponto de medição que abranja todos os dispositivos a serem detectados (ZOHA *et al.*, 2012).

Segundo Ruano *et al.* (2019), as técnicas NILM tornaram-se relevantes para a leitura de consumo e desagregação de energia, pois respeita a privacidade dos consumidores ao se utilizar um único ponto de medição, normalmente pelo uso de medidores inteligentes. Ainda, alguns domínios emergentes como Internet das Coisas, do inglês *Internet of Things* (IoT), (RAIKER *et al.*, 2020), (GHOSH *et al.*, 2020) e *Smart Grids* (MAKONIN *et al.*, 2018) promovem e aceleram o desenvolvimento de novas técnicas não intrusivas.

O conceito geral de NILM foi apresentado por Hart (1992) como um método para desagregar cargas elétricas, utilizando sinais de tensão e corrente elétrica de equipamentos, medindo o consumo de energia apenas no painel elétrico principal, fora do edifício ou residência.

O sinal de potência agregada  $P(t)$  no ponto de entrada do medidor pode ser definida por:

$$P(t) = p_1(t) + p_2(t) + \dots + p_n(t), \quad (1)$$

com  $p_i$  sendo o consumo de energia de aparelhos individuais que contribuem para a medição agregada, para  $i = 1, 2, \dots, n$  o número total de dispositivos ativos no período de tempo  $t$ .

As cargas elétricas possuem um padrão de consumo energético individual, denominado assinatura de potência, que permite aos algoritmos a identificação e desagregação dos eletroeletrônicos (ZOHA *et al.*, 2012). A desagregação de cargas pode ser definida como uma separação de cargas compostas (múltiplas ou agregadas) em componentes individuais utilizando abordagens baseadas em métodos de otimização, separação de fontes, processamento esparso de sinais ou algoritmos inteligentes (reconhecimento de padrões). Em Hart (1992) é proposta uma categorização das cargas em quatro diferentes tipos, baseados nos modos de operação:

- Tipo 1: aparelhos com dois estados de operação (*ON/OFF*);
- Tipo 2: aparelhos multi estados, também definidos como Máquinas de Estado Finito (MEF). Usualmente são equipamentos que apresentam mais de um estado de funcionamento, como uma máquina de lavar ou uma secadora.
- Tipo 3: aparelhos com consumo de energia variável, conhecidos como Dispositivos Continuamente Variáveis (DCV). DCVs não possuem um número de estados finito de funcionamento, seu consumo de energia varia sem um padrão específico, como um motor acionado por inversor;
- Tipo 4: aparelhos que consomem energia a uma taxa constante.

A técnica NILM pode ser dividida em quatro principais estágios: (i) Coleta de Dados; (ii) Detecção de Eventos; (iii) Extração das características; e (iv) Identificação da carga (RUANO *et al.*, 2019).

A coleta de dados é a primeira fase de um processo NILM e tem grande influência na escolha de quais tipos de algoritmos serão utilizados posteriormente. A aquisição usualmente está relacionada a um dispositivo ou sistema de coleta posicionado no barramento de energia, para medição de tensão e corrente em uma residência ou edifício e pode ser realizada em baixa frequência (menor que 1Hz) ou em alta frequência (kHz ou MHz). A aquisição realizada em alta frequência é recomendada para identificação de características transitórias e estacionárias, por

ser mais eficiente (SUN *et al.*, 2019). Porém, apresenta alto custo de instalação e manutenção (SUN *et al.*, 2019).

Após a aquisição de dados, a etapa de processamento dos dados é iniciada afim de calcular as métricas de energia, potência ativa e reativa, por exemplo. Em seguida, o processamento é realizado iniciando pela detecção de eventos de transição *ON* e *OFF* das medições (ZOHA *et al.*, 2012). Um evento pode ser definido como qualquer mudança em um sinal de um determinado estado estacionário para outro estado, também estacionário (gerando um regime transiente entre os dois estados estacionários), seja no domínio do tempo ou da frequência. É frequentemente associado a altas taxas de amostragem para garantir a uma maior precisão da detecção, além de ser identificado com maior clareza em sinais de corrente (RUANO *et al.*, 2019). Ainda, os eventos podem ser divididos em termos de estado transitório ou estacionário. Os métodos de estado estacionário são os mais simples (em termos computacionais) de implementar, mas os métodos de estado transitório são mais eficazes para lidar com cargas sobrepostas (ZOHA *et al.*, 2012).

Ao se detectar e definir os eventos, em geral é realizada a determinação e seleção de características que representam determinado dispositivo, chamada de extração de características. A maioria dos métodos disponíveis na literatura utilizam séries temporais de variáveis elétricas: tensão e corrente, potência aparente, ativa e reativa, ângulo da fase de potência e fator de potência (RUANO *et al.*, 2019).

A ultima etapa é referente à identificação das cargas na qual as características extraídas anteriormente são usadas para a classificação dos equipamentos, podendo ser de forma supervisionada ou não supervisionada (FIGUEIREDO *et al.*, 2012).

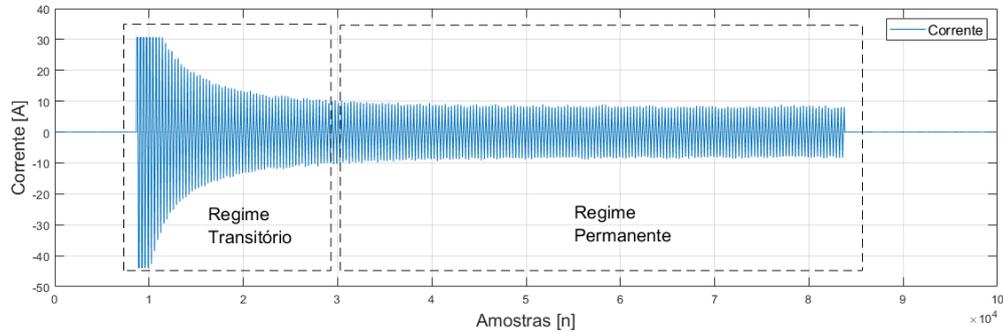
## 2.2 MÉTODOS DE EXTRAÇÃO

Na etapa de extração de características, os métodos podem ser divididos em (IKSAN *et al.*, 2016):

- Estado estacionário (regime permanente): realizam a extração das características em segmentos de onda em regime permanente, como pode ser visto na Figura 1;
- Estado Transitório: realizam a extração das características em segmentos de onda em regime transitório;

- Combinação de estados: realizam extração das características nos estados transitório e permanente das formas de onda.

**Figura 1 – Descritivo de Regime Transitório e Regime Permanente.**



**Fonte: Autorial Própria.**

Na literatura os principais exemplos de métodos estacionários são: (i) variação de potência (LEE *et al.*, 2005); (ii) harmônicos de corrente (SRINIVASAN *et al.*, 2006); e (iii) curvas V-I (WANG *et al.*, 2018).

Em Lee *et al.* (2005) é apresentado um método de estimativa de potência para *variable-speed drive*, que consiste em um retificador, um barramento de corrente contínua e um inversor. Diversos equipamentos funcionam por meio de acionamento de velocidade variável como motores, ventiladores e bombas. A proposta é estimar a energia consumida a partir de observações nas harmônicas mais elevadas dos sinais de corrente, subtraindo o sinal de energia agregado.

Já em Srinivasan *et al.* (2006), os harmônicos de corrente produzidos pelos equipamentos são considerados assinaturas exclusivas, usadas para fornecer informações precisas aos sistemas de gerenciamento e aprimoramento da qualidade de energia, para que as ações corretivas apropriadas possam ser tomadas. As características dos harmônicos alimentaram vários modelos de identificação de assinaturas baseados em redes neurais, incluindo o Perceptron de múltiplas camadas, rede de função de base radial e máquinas de vetores de suporte (SVM) com *kernels* lineares, polinomiais e funções de base radial (SRINIVASAN *et al.*, 2006).

O método das curvas V-I é baseado na extração de características do gráfico da tensão *versus* corrente do equipamento em um ciclo de rede. Em Wang *et al.* (2018) são propostas dez características dessa trajetória: área, amplitude da corrente, área com a direção do *loop*, assimetria, curvatura da linha média, auto-intersecção, pico do segmento médio, forma do segmento médio, área dos segmentos direito e esquerdo e variação da admitância instantânea.

Ainda, em Hassan *et al.* (2014) e Lam *et al.* (2006), Lam *et al.* (2007) são propostas diferentes características para a trajetória V-I.

Algumas desvantagens do uso de algoritmos de extração em regime estacionário são: (i) a similaridade das assinaturas quando se tem muitos tipos de carga combinadas; e (ii) a falta de informações extraídas durante os instantes de comutação de carga, como correntes de *inrush*<sup>1</sup> de motores elétricos (WANG *et al.*, 2018; NAIT-MEZIANE *et al.*, 2019).

Por outro lado, alguns métodos de extração de características baseada em regimes transitórios têm sido apresentados na literatura recente, como: Transformada de *Wavelet* (CHANG, 2012; CHANG *et al.*, 2014), a Transformada S (MARTINS *et al.*, 2012) e a análise do transitório da curva da corrente (MEZIANE *et al.*, 2015; NAIT-MEZIANE *et al.*, 2019).

Em Martins *et al.* (2012), um algoritmo de monitoramento de cargas baseado na transformada S é proposto. Segundo os autores, esse método é mais vantajoso em relação às outras representações no tempo, pois possui resolução progressiva, informações de fase referenciada e resposta de amplitude variante, auxiliando na caracterização transitória das cargas.

De forma similar, a transformada *Wavelet* é frequentemente usada para capturar o tempo de ocorrência e as características dos transitórios, como em Chang (2012), Chang *et al.* (2014). Nesses trabalhos, utiliza-se a distribuição de energia nos diversos níveis da transformada *Wavelet* do sinal transitório durante instantes de *ON* e *OFF* de diferentes cargas individuais.

Já em Meziane *et al.* (2015) é constatado que o comportamento transitório da corrente elétrica dos aparelhos é fortemente influenciado pela tarefa física que o aparelho realiza e que uso de características do regime transitório melhoram a precisão do reconhecimento de determinados aparelhos. Os autores utilizam a corrente transitória de ativação de aparelhos elétricos, corrente de *inrush*, para extrair características das variações temporais dos parâmetros do sinal como frequência, fase e amplitude, classificando os equipamentos de acordo com esses parâmetros.

No geral, o uso de métodos transitórios para extração de características e posterior classificação auxilia na análise de informações típicas de cada carga, pois assinaturas transitórias são mais curtas e diferem para cada tipo de equipamento. Mas para garantir a eficiência é preciso uma alta frequência de amostragem e grande capacidade de memória para armazenamento dos dados (FIGUEIREDO *et al.*, 2012). Além disso, informações de regime permanente que possuem características discriminatórias para as cargas são completamente descartadas nesse métodos.

Por fim, ainda é possível combinar métodos transitórios e estacionários para extrair

---

<sup>1</sup> Corrente transitória observada no momento que o motor é ligado, em função do seu circuito elétrico e magnético interno.

características do sinal. A literatura apresenta alguns métodos que seguem essa abordagem. Em Mulinari *et al.* (2019), são associadas características transitórias e permanentes da curva V-I e Nait-Meziane *et al.* (2019) apresenta a transformada de Fourier com uma modificação incorporando o amortecimento exponencial para conjunto de dados *Controll On / Off Loads Library* (COOLL). Em geral, mostra-se uma melhora na acurácia de classificação de cargas quando as características de ambos os estados (estacionário e transitório) são inseridas. Porém, em ambos os casos, somente bases de dados com cargas simples são avaliadas.

No caso de Meziane *et al.* (2017) e no seu trabalho subsequente em Nait-Meziane *et al.* (2019), os autores demonstram que este amortecimento permite incrementar os desempenhos de classificação para valores superiores a 90%. Porém a implementação do método apresenta uma instabilidade na etapa de regressão exponencial (HUA; SARKAR, 1990), alterando a estimativa da forma de onda na porção senoidal, resultando em estimação de conteúdo harmônico impreciso, comprometendo a classificação como um todo.

Ainda neste contexto, algoritmos extratores combinando os dois estados podem ser vistos em Renaux *et al.* (2018), o HCApP, que extrai características com base na potência ativa, reativa e aparente de um sinal, durante a comutação de carga e estado estacionário subsequente, para executar a classificação. Mais recentemente, em (LAZZARETTI *et al.*, 2020), foi demonstrado que o HCApP apresenta resultados comparáveis e, em alguns casos superiores à literatura recente, na acurácia de classificação de cargas em bases de dados públicas. No entanto, é necessário um grupo de técnicas (cálculos) para extração de características de cada estado (transitório e regime), o que pode elevar a complexidade computacional do processo como um todo.

### 2.3 MÉTODOS DE CLASSIFICAÇÃO

A última etapa do ciclo NILM é a identificação de cargas, utilizando as características extraídas anteriormente, combinadas com algoritmos de classificação. Os métodos mais utilizados para esta análise são algoritmos de aprendizado de máquina supervisionado e não supervisionado (RUANO *et al.*, 2019).

Na classificação supervisionada, o treinamento é realizado para obter um banco de dados de informações usadas para projetar os classificadores. Algumas técnicas comuns de aprendizado supervisionado são: as Redes Neurais Artificiais, principalmente o Perceptron Multicamada, Redes Neurais Convolucionais concatenadas, Redes Neurais Profundas, Máquina

Vetor Suporte, do inglês *Support Vector Machines (SVM)*, *K-Nearest Neighbours (kNN)* e recentemente, campos aleatórios condicionais de cadeia linear, do inglês *Conditional Random Fields* (RUANO *et al.*, 2019).

Na classificação não supervisionada não é necessário realizar o treinamento antes da classificação, sendo considerada como uma vantagem do método, pois demanda um esforço mínimo do usuário e a minimização da intrusão no ambiente envolvida na criação de um banco de dados é reduzida. As técnicas mais recentes são as que assumem que o sinal elétrico é a saída de um sistema estocástico, como cadeias ocultas de Markov (RUANO *et al.*, 2019).

Nos métodos supervisionados o processo de classificação é realizado em duas etapas: treinamento e teste. Usualmente 80% dos dados coletados são utilizados para treinamento do algoritmo, fase na qual os equipamentos recebem rótulos, ou seja, são relacionadas as características extraídas com a classe de cada equipamento. Os 20% restantes dos dados são utilizados para o teste do aprendizado feito anteriormente.

Na literatura, exemplos de aplicação de aprendizado de máquina supervisionado para classificação de cargas podem ser vistos em Aydinalp-Koksal e Ugursal (2008) e Venkatesh e Kumar (2008) com Redes Neurais, algoritmos genéticos em Egarter e Elmenreich (2013) e Zhang *et al.* (2015), SVM em Jiang *et al.* (2013), modelos de Markov em Kim *et al.* (2011), árvores de decisão em Liao *et al.* (2014), classificador de Bayes em Zeifman (2012) e Zeifman *et al.* (2011), e kNN em Berges *et al.* (2011) e Gopinath *et al.* (2020).

Em Zhuang *et al.* (2018) é apresentada uma aplicação do Perceptron Multicamada para a construção do algoritmo de classificação para cargas não lineares. Comparações entre os métodos de Máquina de Vetor Suporte, Redes Neurais Artificiais e *boost* adaptativo (aprendizado de máquina a partir de um algoritmo meta-heurístico) podem ser vistas em Hassan *et al.* (2014), indicando que o uso do *boost* adaptativo trouxe melhor desempenho entre os métodos (LIU, 2020).

Em Liu *et al.* (2019) é utilizado o classificador *Random Forest* alimentado com características baseadas em séries temporais. Ainda, são utilizados métodos como kNN em Berges *et al.* (2011), modelo Naive Bayes em Mathis *et al.* (2014) e árvores de decisão em Gillis *et al.* (2016) (LIU, 2020). De forma geral, métodos baseados em *Random Forest* apresentam resultados superiores em termos de acurácia de classificação.

Algoritmos supervisionados têm sido eficientes na identificação de equipamentos e desagregação de cargas (GOPINATH *et al.*, 2020), porém não se adaptam em tempo real às

mudanças de dispositivos presentes no ambiente. Para garantir a eficiência é preciso atualizar os rótulos para a etapa de aprendizado sempre que houver algum tipo de alteração ou inclusão de equipamentos.

Métodos não supervisionados não requerem o processo de rotular os dispositivos para treinamento do algoritmo. Aplicações de aprendizado não supervisionados a partir de modelos ocultos de Markov e agrupamento de *k-means* são amplamente utilizados na literatura NILM (BONFIGLI *et al.*, 2015). A abordagem não supervisionada é considerada competitiva em relação à supervisionada por não ser necessário a aquisição de bases de dados em alta frequência. Porém, não indica a classe ou tipo de equipamento específico, o que pode limitar sua aplicação (GOPINATH *et al.*, 2020).

## 2.4 BASE DE DADOS

A base de dados influencia diretamente na seleção e comparação na extração de características e na escolha dos métodos de classificação (RENAUX *et al.*, 2018). Na literatura existem diferentes base de dados, coletadas em condições reais e direcionadas para a análise de assinaturas de potência. Alguns exemplos de base de dados públicas são:

- REDD (*Reference Energy Disaggregation Data Set*);
- COOLL (*Controlled On/Off Loads Library*);
- UK-DALE (*United Kingdom recording Domestic Appliance-Level Electricity*);
- BLUED (*Building-Level fully-labeled dataset for Electricity Disaggregation*);
- WHITED (*Worldwide Household and Industry Transient Energy Data Set*);
- PLAID (*Plug Load Appliance Identification Dataset*);
- SynD (*Synthetic Energy Consumption*);
- BLOND (*Building-Level Office Environment Dataset*).

A base de dados COOLL, apresentada em Picon *et al.* (2016), é caracterizada por cargas residenciais operando em 50Hz, coletadas em um ambiente controlado no laboratório PRISME da Universidade de Orleans. As formas de ondas são individuais, com 42 tipos de classes com coletas a partir de 12 diferentes tipos de eletrodomésticos. Com uma alta taxa de

de amostragem de coleta de dados, 100kHz, essa base se torna aplicável para análise NILM residencial. Porém, apenas cargas individuais são apresentadas e não existe uma anotação precisa dos eventos *ON/OFF* nas formas de onda, ficando a cargo do usuário traçar a melhor forma de fazê-lo.

A base de dados REDD, realiza a aquisição de formas de ondas de tensão e corrente de diferentes residências e por várias semanas. Para cada casa são coletados os sinais de eletricidade residencial (monitoramento de corrente em duas fases de energia e um monitor de tensão em uma fase) gravado em alta frequência 15 kHz (KOLTER; JOHNSON, 2011). Com taxa de amostragem menor que a apresentada pela base COOLL, a REDD possui cargas múltiplas para análise, porém seus eventos não possuem anotação e sua estrutura de dados dispõe uma estrutura complexa.

A base de dados UK-DALE registra, a cada seis segundos, a demanda geral de energia e a demanda individual dos aparelhos de cinco casas. Ainda são coletadas as leituras de tensão de corrente em três das casas disponíveis (KELLY; KNOTTENBELT, 2015). Mesmo apresentando cargas múltiplas, não é determinado o número de equipamentos funcionando em cada casa, o que dificulta no processo de rotular cargas e definir classes.

A base de dados BLUED, em Anderson *et al.* (2012) apresenta um conjunto de dados que consiste em medições de tensão e corrente para uma residência nos Estados Unidos, amostradas a 12 kHz por uma semana. Nesta coleta, todas as transições de estado de cada dispositivo na casa, durante esse período, foram rotuladas com registro de data e hora. Cargas múltiplas também são coletadas nessa base, além de possuir precisão na anotação de eventos. Como principal dificuldade, sua estrutura de dados dispõe de uma organização complexa, que inviabiliza o uso da anotação precisa de eventos para extração de características transitórias.

A base de dados WHITED, Kahl *et al.* (2016), foi coletada em residências e pequenas indústrias de diferentes regiões do mundo. Neste trabalho é apresentado um conjunto de dados de medições de vários locais. Os aparelhos foram gravados com um medidor de “placa de som” personalizado de baixo custo. Com uma taxa de amostragem de 44,1kHz, não possui cargas múltiplas nem uma anotação de eventos *ON/OFF* das cargas.

A base de dados PLAID, Gao *et al.* (2014), é caracterizada pela coleta de dados de cargas residenciais, cada tipo de dispositivo é representado por dezenas de instâncias diferentes de marcas e modelos diferentes. Para cada aparelho, de três a seis medições foram coletadas em cada transição de estado. Essas medidas foram pós-processadas para extrair uma janela de

alguns segundos, contendo a operação em estado estacionário e o transiente de inicialização. No entanto, a anotação não é feita a nível de amostra do sinal, dificultando o seu uso para extração de características transitórias.

Com uma coleta de dados de um prédio comercial na Alemanha, a base de dados BLOND (KRIECHBAUMER; JACOBSEN, 2018) possui 53 equipamentos divididos em 16 diferentes classes, com uma frequência de 60 Hz e duração de 123 dias.

A base de dados SynD (KLEMENJAK *et al.*, 2020), da Áustria, é composta por dados residenciais de 21 eletrodomésticos monitorados num período de 180 dias com uma frequência de amostragem de 5 Hz.

A Tabela 1 apresenta um comparativo das bases de dados mencionadas. Pode se observar as principais características de cada base, como a frequência de amostragem, presença de cargas múltiplas, precisão de eventos e duração da coleta.

**Tabela 1 – Comparação entre as bases de dados reais.**

Base de Dados	Ano	Tipo	Tempo de Duração	Cargas Múltiplas	$f_s$	Resolução de Evento	Nº de Classes	Nº de Equip.
REDD	2011	Res.	119 dias (10 casas)	sim	15 kHz	3 s	8	24
BLUED	2012	Res.	8 dias (1 casa)	sim	12 kHz	640 ms	9	43
PLAID	2014	Res.	1094 formas de onda (1 s)	não	30 kHz	> 1 ciclo	12	235
UK-DALE	2015	Res.	655 dias	sim	16 kHz	6 s	16	54
COOLL	2016	Res.	840 formas de onda (6 s)	não	100 kHz	20 ms	12	42
WHITED	2016	Res.+ Ind.	5123 formas de onda (5 s)	não	44,1 kHz	-	47	110
BLOND	2018	Res.	50-213 dias	sim	50-250 kHz	-	16	53
SynD	2020	Res.	180 dias	sim	5 Hz	0,2 s	-	21

**Fonte: Adaptado de Renaux *et al.* (2020a).**

Para as análises de assinatura de potência e implementação de algoritmos em NILM ainda é possível utilizar bases de dados simulados. Dados simulados trazem vantagens como a facilidade de implementação de cargas simultâneas, identificação precisa e variação de ruídos e conteúdo harmônico.

Na literatura, em Venkatesh e Kumar (2008) e Gillis *et al.* (2016), o aplicativo PS-CAD/EMTDC é utilizado para modelagem de cargas elétricas não lineares de sistemas do-

mésticos e de pequena escala industrial. Os autores propõem um conceito baseado no design de *Wavelets* e no aprendizado de máquina aplicado ao NILM usando quatro cargas diferentes simuladas no software. Também em Lopez e Pouresmaeil (2019) e Collin *et al.* (2014), um simulador de carga residencial baseado no MATLAB é proposto para facilitar o estudo de sistemas de gerenciamento de energia residencial em redes inteligentes, com foco em sinais no estado estacionário.

Nos trabalhos que fazem uso de dados simulados anteriormente apresentados, uma limitação em comum é a ausência de problemas relacionados a ruído e conteúdo harmônico na simulação, o que pode impactar significativamente na análise de assinatura de energia.

## 2.5 CONTRIBUIÇÕES DESTE TRABALHO

Para a primeira frente deste trabalho, observa-se que nos métodos de extração de características, quando combinados estados transitórios e de regime permanente para cálculo de informações da assinatura de potência, a acurácia de classificação é incrementada. Sendo assim, propõe-se a utilização do método matemático de Prony, apresentado na seção 3.1, para a implementação de um algoritmo de estado transitório/permanente, que realiza a extração de características de sinais de corrente por meio de série de exponenciais complexas, a partir da estimativa da fase, amplitude, coeficiente de amortecimento e frequência. A abordagem aqui proposta apresenta algumas vantagens, tais como: (i) ser um método analítico que permite a estimativa de características transitórias e de regime permanente numa única abordagem matemática (sem necessidade de diferentes métodos ou etapas); (ii) pode ser utilizado no contexto de cargas múltiplas; e (iii) as características estimadas têm relação direta com o sinal modelado (harmônicos e decaimento exponencial).

Já para a segunda frente, as bases de dados para análise NILM possuem limitações como: (i) instantes transitórios não devidamente anotados; (ii) ausência de cargas simples e múltiplas simultaneamente; e (iii) ausência de harmônicos e diferentes condições de ruído. A proposta deste trabalho é desenvolver uma base de dados simulada baseada em cargas residenciais que, além dos eventos rotulados de forma precisa e da presença de cargas múltiplas, também apresenta a possibilidade da inserção de ruídos e harmônicos na rede.

### 3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os principais conceitos teóricos envolvidos neste trabalho. As seções seguem com a extração de características transitórias e de regime permanente, com base no Método de Prony e suas abordagens matemáticas, seguida por uma breve descrição dos modelos utilizados para classificação de cargas.

#### 3.1 MÉTODO DE PRONY

Gaspard-Clair-François-Marie Riche de Prony (1755-1839) estava entre os professores das principais escolas especializadas da França (BRADLE, 1994). No ano de 1795, este matemático desenvolveu um método, semelhante à série de Fourier, criada em 1807, que extrai informações de um sinal e constrói uma série de exponenciais complexas amortecidas. Assumindo as  $N$  amostras de dados  $x[1], \dots, x[N]$ , o método de Prony estima  $x[n] \in \mathbb{R}$ , sendo  $n \in \mathbb{N}$ , com um somatório de  $p$ -termos compostos por exponenciais complexas, conforme a seguinte equação:

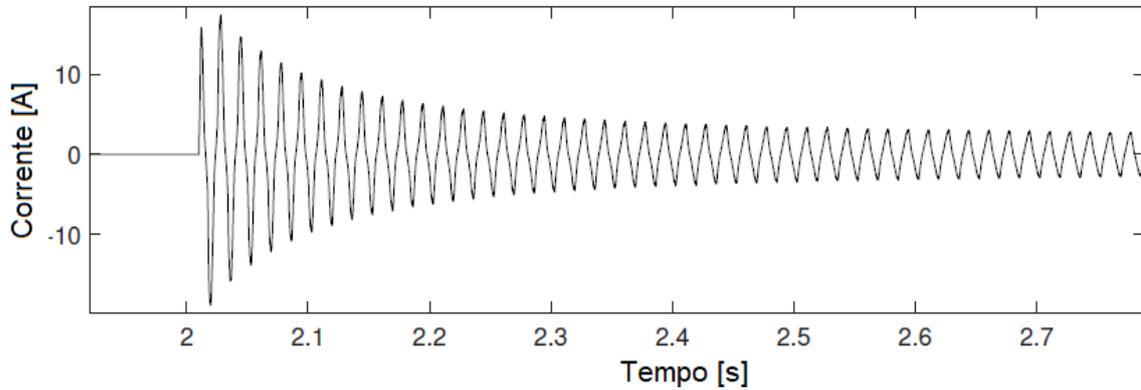
$$\hat{x}[n] = \sum_{k=1}^p A_k \exp[(\alpha_k + j2\pi f_k)(n-1)T_s + j\theta_k], \quad (2)$$

em que  $T_s$ ,  $A_k$ ,  $\alpha_k$ ,  $f_k$ , e  $\theta_k$  são o período de amostragem, amplitude, fator de amortecimento, frequência e fase, respectivamente, sendo  $k \in \mathbb{N}$ .

Em tal abordagem matemática, os parâmetros acima mencionados, podem ser utilizados para estimação de assinatura de potência e atuar na identificação de cargas em sistemas NILM. Em geral, somente o conteúdo harmônico dos sinais de assinatura de carga pode não ter todas as informações relevantes para diferenciar cargas com alta corrente inicial. Nesse sentido, pode-se usar uma assinatura de corrente transitória, como mostrado em Meziane *et al.* (2017). A corrente transitória de um equipamento elétrico, geralmente chamada corrente de *inrush*, é caracterizada por uma forma de onda de corrente que contém valores altos no início do consumo de energia e valores decrescentes à medida que o estado estacionário é atingido, conforme ilustrado na Figura 2.

O método de Prony pode ser utilizado exatamente neste contexto, realizando a estimação do sinal, neste caso, as assinaturas de cargas, com base nos parâmetros estimados  $A_k$ ,  $\alpha_k$ ,  $f_k$  e  $\theta_k$ . O método de Prony pode ser resumido em três etapas básicas, apresentadas na Figura 3, definidas

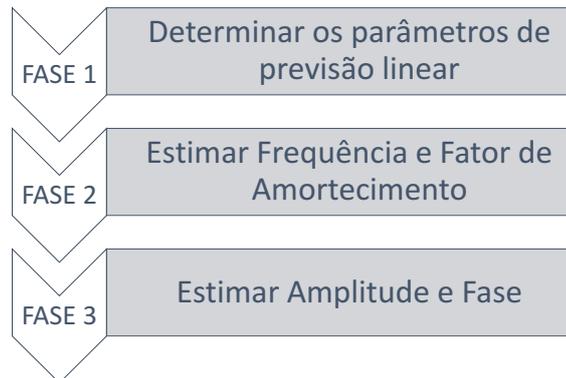
**Figura 2 – Corrente de *Inrush* da Furadeira Elétrica.**



Fonte: Adaptado de Ancelmo *et al.* (2019a)

como: (i) determinação dos parâmetros de previsão linear que se ajustam aos dados disponíveis; (ii) estimação das frequências de amortecimento e senoidais dos termos exponenciais; e (iii) estimação da amplitude e da fase inicial associada às exponenciais (MARPLE, 1986).

**Figura 3 – Fases do método de Prony.**



Fonte: Adaptado de Marple (1986).

Nas próximas seções serão apresentados cinco diferentes métodos de estimação dos parâmetros necessários ao modelo proposto por Prony: (i) Clássico (MARPLE, 1986); (ii) Mínimos Quadrados (MOHAMMAD *et al.*, 2019), (iii) Mínimos Quadrados Totais (GOLUB; LOAN, 1980); (iv) *Matrix Pencil* (HUA; SARKAR, 1988); e (v) Filtros Resposta ao Impulso Infinita, do inglês *Infinite Impulse Response* (IIR) (STORN, 1996).

### 3.1.1 Método de Prony Clássico

Um sinal  $x[n]$ , sendo  $n = 1, 2, \dots, N$  pode ser expresso como uma série de Laurent (ARFKEN *et al.*, 2013), com  $p$  termos complexos, ou seja,  $p$  é igual a ordem do sistema modelado, em que versões defasadas da resposta ao impulso de um sistema (capaz de gerar tais amostras), são defasadas no tempo discreto através do operador  $z$  e somadas entre si:

$$x[n] = \sum_{k=1}^p h_k z_k^{n-1}. \quad (3)$$

O sinal complexo  $h_k$  pode ser representado com um fasor expresso por:

$$h_k = A_k e^{j\theta_k}. \quad (4)$$

Já o operador de avanço/atraso discreto pode ser definido da seguinte forma:

$$z_k = e^{(\alpha_k + j2\pi f_k)T_s}, \quad (5)$$

sendo que  $k \in \mathbb{N}$  representa o indexador no somatório das variáveis envolvidas em (3). Se as  $n$  amostras de  $x[n]$  (várias equações escalares) forem concatenadas em forma vetorial, então a equação (3) pode ser reescrita na forma matricial, para  $p$  igual a ordem do sistema:  $\mathbf{Z}\mathbf{h} = \mathbf{x}$ :

$$\begin{pmatrix} z_1^0 & z_2^0 & \dots & z_p^0 \\ z_1^1 & z_2^1 & \dots & z_p^1 \\ \vdots & \vdots & & \vdots \\ z_1^{p-1} & z_2^{p-1} & \dots & z_p^{p-1} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \end{pmatrix} = \begin{pmatrix} x[1] \\ x[2] \\ \vdots \\ x[p] \end{pmatrix}, \quad (6)$$

sendo  $z_n^m$  para  $n = 1, 2, \dots, p$  e  $m = 0, 1, \dots, p-1$ .

A abordagem original proposta por Prony, chamada de clássica ou polinomial, define um polinômio que inclui os expoentes como suas raízes, como:

$$\phi(z) = \prod_{k=1}^p (z - z_k) = \sum_{m=0}^p a[m] z^{p-m}. \quad (7)$$

Essa representação também pode ser reescrita usando uma formulação de matriz como

$\mathbf{T}\mathbf{a} = -\mathbf{x}$ :

$$\begin{pmatrix} x[p] & x[p-1] & \dots & x[1] \\ x[p+1] & x[p] & \dots & x[2] \\ \vdots & \vdots & & \vdots \\ x[2p-1] & x[2p-2] & \dots & x[p] \end{pmatrix} \begin{pmatrix} a[1] \\ a[2] \\ \vdots \\ a[p] \end{pmatrix} = - \begin{pmatrix} x[p+1] \\ x[p+2] \\ \vdots \\ x[2p] \end{pmatrix}. \quad (8)$$

Portanto, a solução do método de Prony consiste em 3 fases, mencionadas na Figura 3: (i) solucionar o sistema da Equação (8) para encontrar os coeficientes  $a[m]$ , para então utilizando esses coeficientes: (ii) encontrar as raízes  $z_k$  que compõe a matriz de Vandermonde<sup>1</sup> da Equação (6). A última etapa consiste em (iii) solucionar o sistema para encontrar as amplitudes  $h_k$ .

Finalmente, o fator de amortecimento  $\alpha_k$  e a frequência senoidal  $f_k$  podem ser determinadas a partir das raízes  $z_k$  utilizando as seguintes relações:

$$\alpha_k = \frac{\ln|z_k|}{T_s}, \quad (9)$$

$$f_k = \frac{\tan^{-1} \left[ \frac{\text{Im}[z_k]}{\text{Re}[z_k]} \right]}{2\pi T_s}. \quad (10)$$

A amplitude  $A_k$  and fase inicial  $\theta_k$  podem ser determinadas a partir de cada parâmetro  $h_k$ :

$$A_k = |h_k|, \quad (11)$$

$$\theta_k = \tan^{-1} \left[ \frac{\text{Im}[h_k]}{\text{Re}[h_k]} \right]. \quad (12)$$

### 3.1.2 Método de Prony com Mínimos Quadrados

Na abordagem clássica ou polinomial do método de Prony, o número de amostras deve ser igual a  $2p$ , ou seja,  $N = 2p$ , sendo  $p$  o número de termos complexos do sistema modelado. Já na solução Mínimos Quadrados (MQ), a ideia é aumentar o número de amostras para se melhorar a estimativa dos parâmetros do modelo, resultando em  $N > 2p$ . Nesta situação, os sistemas lineares são sobre-determinados (mais equações do que incógnitas) e podem ser aproximados usando métodos baseados em Mínimos Quadrados, modificando a solução da Fase 1 na Figura 3, conforme exposto em Rodríguez *et al.* (2018).

Para sistemas lineares  $\mathbf{Z}\mathbf{h} = \mathbf{x}$  e  $\mathbf{T}\mathbf{a} = -\mathbf{x}$  quando  $N > 2p$ , a solução baseada no método dos Mínimos Quadrados  $\mathbf{x}_{MQ}$  de um sistema linear  $\mathbf{A}\mathbf{x} = \mathbf{b}$  é dado por:

$$\mathbf{x}_{MQ} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{b}, \quad (13)$$

em que  $\mathbf{A}^H$  corresponde ao Hermitiano<sup>2</sup> da matrix  $\mathbf{A}$ . Pode-se também usar a Decomposição de valor singular, do inglês *singular value decomposition* (SVD), para resolver o problema dos mínimos quadrados. Essa abordagem será tratada na próxima seção.

<sup>1</sup> Matriz em que os termos de cada linha estão em progressão geométrica.

<sup>2</sup> Um operador linear em um espaço vetorial com produto interno que é o adjunto de si mesmo.

### 3.1.3 Método de Prony com Mínimos Quadrados Totais

O método com Mínimos Quadrados Totais (MQT) pode ser considerado uma generalização da solução apresentada no MQ para  $\mathbf{Ax} = \mathbf{b}$ , particularmente desenvolvido para problemas de condicionamento numérico. Considerando que  $\mathbf{Zh} = \mathbf{x}$  e  $\mathbf{Ta} = -\mathbf{x}$  são construídos a partir dos sinais medidos, o algoritmo MQT consiste em aplicar, na Fase 1 da Figura 3, a SVD<sup>3</sup> de uma matriz  $\mathbf{C}$ :

$$\mathbf{C} = [\mathbf{A}|\mathbf{b}] = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \quad (14)$$

em que  $|\mathbf{b}$  corresponde a expansão de colunas de  $\mathbf{A}$ , usando o vetor  $\mathbf{b}$ .  $\mathbf{U}$  e  $\mathbf{V}$  são compostas pelos valores singulares de  $\mathbf{C} = [\mathbf{A}|\mathbf{b}]$ , compilados em matrizes que representam transformações lineares de rotação e  $\mathbf{\Sigma}$  é a matriz de escala com os valores singulares de  $\mathbf{C}$ .

Assim,  $\mathbf{x}_{MQT}$  pode ser substituído na Equação 8, resultando na solução MQT para o método:

$$\begin{pmatrix} \mathbf{x}_{MQT} \end{pmatrix} \begin{pmatrix} a[1] \\ a[2] \\ \vdots \\ a[p] \end{pmatrix} = - \begin{pmatrix} x[p+1] \\ x[p+2] \\ \vdots \\ x[2p] \end{pmatrix}. \quad (15)$$

sendo:

$$\mathbf{x}_{MQT} = - \frac{1}{\mathbf{v}_{(n+1),(n+1)}} [\mathbf{v}_{1,(n+1)}, \dots, \mathbf{v}_{n,(n+1)}]. \quad (16)$$

### 3.1.4 Método de Prony usando *Matrix Pencil*

Um problema padrão de decomposição em autovalores e autovetores é caracterizado por  $\mathbf{Ax} = \lambda\mathbf{x}$ , sendo  $\lambda$  um autovalor de  $\mathbf{A}$ . Já o problema generalizado de decomposição de  $\mathbf{A}$  em autovalores e autovetores, utilizado no método *Matrix Pencil*, é caracterizado por  $\mathbf{Ax} = \mathbf{B}\lambda\mathbf{x}$ , sendo que  $\mathbf{A}$  e  $\mathbf{B}$  são matrizes  $n \times n$  e  $\lambda$  e  $\mathbf{x}$  são os autovalores e autovetores a serem encontrados, respectivamente. Maiores detalhes podem ser encontrados em Pissanetzky (1984).

Ao solucionar o problema de autovalor generalizado, proposto em Rodríguez *et al.* (2018), pode-se definir as raízes do polinômio característico da Fase 1, Figura 3, aplicando o método *Matrix Pencil* (MP).

<sup>3</sup> Para maiores informações sobre a decomposição, consultar Brunton e Kutz (2019)

Dadas duas matrizes compostas pelo sinal a ser reconstruído  $\mathbf{Y}_1, \mathbf{Y}_2 \in \mathbb{C}^{(N-p) \times (p-1)}$ , o conjunto de matrizes  $\mathbf{Y}_2 - \lambda \mathbf{Y}_1$  é a *matrix pencil* e  $z_k$  pode ser calculada como:

$$z_k = \text{autovalores}[(\mathbf{Y}_1^H \mathbf{Y}_1)^{-1} \mathbf{Y}_1^H \mathbf{Y}_2], \quad (17)$$

sendo:

$$\mathbf{Y}_1 = \begin{pmatrix} x[1] & \dots & x[p] \\ \vdots & \vdots & \\ x[N-p] & \dots & x[N-1] \end{pmatrix}, \quad \mathbf{Y}_2 = \begin{pmatrix} x[2] & \dots & x[p+1] \\ \vdots & \vdots & \\ x[N-p+1] & \dots & x[N] \end{pmatrix}.$$

### 3.1.5 Método de Prony com Filtros IIR

A abordagem final do modelo de Prony é o uso de uma função de transferência que representam o filtros de Resposta ao Impulso Infinita, do inglês *Infinite Impulse Response* (IIR) para estimar os parâmetros como:

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + \dots + a_M z^{-M}}. \quad (18)$$

Esta equação representa uma convolução entre o sistema e um sinal impulso unitário, substitui a Fase 1 da Figura 3 e pode ser reescrita como uma multiplicação de matrizes. Para isso, pode-se definir  $H(z) = \sum_{n=-\infty}^{\infty} h[n]z^{-n}$ . Com o primeiro parâmetro  $K + 1$  da resposta ao impulso, sendo possível escrever a seguinte relação:

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_M \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_0 & 0 & \dots & 0 \\ h_1 & 0 & \dots & \\ h_2 & h_1 & \dots & \\ \vdots & & & \vdots \\ h_M & h_{M-1} & \dots & \\ \vdots & \vdots & & \\ h_K & \dots & & h_{K-N} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix}, \quad (19)$$

ou, em notação matricial:

$$\begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{h}_1 & \mathbf{H}_2 \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a} \end{bmatrix}. \quad (20)$$

Neste sistema,  $\mathbf{b}$  é o vetor de  $M + 1$  coeficientes do numerador,  $\hat{\mathbf{a}}$  é o vetor dos coeficientes do denominador concatenados com  $a_0 = 1$ ,  $\mathbf{h}_1$  é o vetor do último termo  $h_{K-M}$  da

resposta ao impulso,  $\mathbf{H}_1$  é a matriz  $(M + 1) \times (N + 1)$ ,  $\mathbf{H}_2$  é a parte restante  $(K - M) \times N$ , e  $\mathbf{H}$  a matriz de Toeplitz<sup>4</sup>  $\mathbf{T}$  do sinal de entrada  $\mathbf{x}$ . As equações inferiores cujos índices são  $K - M$  e superiores cujos índices são  $M + 1$  podem ser escritas como:

$$\mathbf{h}_1 = -\mathbf{H}_2 \mathbf{a}, \quad (21)$$

$$\hat{\mathbf{b}} = -\mathbf{H}_1 \mathbf{a}. \quad (22)$$

Dado  $\mathbf{a}$  da Equação (21) e  $\mathbf{b}$  da Equação (22), pode-se calcular o filtro IIR  $e$ , usando as fases (ii) e (iii), Figura 3, do método clássico, encontrar os polos  $z_k$ , as amplitudes  $h_k$  e os parâmetros  $\alpha_k$ ,  $f_k$ ,  $A_k$  e  $\theta_k$ , utilizando o mesmo procedimento presente nas Equações (9), (10) (11) e (12).

## 3.2 CLASSIFICADORES

Neste trabalho, diferentes métodos de aprendizado de máquina foram avaliados. Esses métodos são todos baseados no aprendizado supervisionado, o que significa que todos os rótulos de classe são conhecidos e definidos anteriormente durante o aprendizado do algoritmo (CHERKASSKY; MULIER, 1998). Cada método utilizado é detalhado conforme segue.

### 3.2.1 Método kNN

Na abordagem K Vizinhos mais Próximos, do inglês k-Nearest Neighbors, kNN, a classificação de um novo padrão é realizada de acordo com a distância euclidiana entre este novo padrão e os exemplos contidos na base de treinamento. Se os  $k$  vizinhos mais próximos deste padrão forem majoritariamente de uma dada classe, o novo padrão é classificado como sendo desta mesma classe (CHERKASSKY; MULIER, 1998), sendo  $k$  um parâmetro definido *a priori*. Como a classificação é simplesmente baseada em distâncias relacionadas a um conjunto de treinamento, esse método pode ser considerado um dos algoritmos mais simples de aprendizado de máquina.

<sup>4</sup> É uma matriz em que cada diagonal descendente da esquerda para a direita é constante.

### 3.2.2 Método Árvores de Decisão

O método de classificação a partir de Árvores de Decisão, do inglês *Decision Tree* (DT), tem como objetivo realizar classificações binárias e concatená-las em uma estrutura de árvore, sendo que cada nó representa cada variável (característica) a ser testada. A associação das diferentes combinações, normalmente expressa nos nós-folha da árvore, resulta no rótulo das classes. O processo de treinamento do algoritmo (construção automática da árvore) é definido com base no ganho de informação (CHERKASSKY; MULIER, 1998). O ganho de informação corresponde ao processo de minimizar a entropia<sup>5</sup> a cada novo nível da árvore. O tamanho máximo da árvore pode ser considerado um parâmetro a ser definido *a priori* para este modelo.

### 3.2.3 Método *Ensemble*

O método *Ensemble* (ENS) é composto por um conjunto de classificadores cujas decisões individuais são combinadas, normalmente utilizando a média, para classificar novos exemplos. Nos métodos baseados em combinação de classificadores, a classificação conjunta geralmente apresenta uma acurácia superior, quando comparada com os classificadores individuais que compõem a combinação (CHERKASSKY; MULIER, 1998). O *Ensemble* pode usar diferentes algoritmos para os classificadores individuais (normalmente baseados em árvores de decisão), tais como *AdaBoostM1*, *AdaBoostM2*, *Bag (Random Forest)*, *GentleBoost*, *LogitBoost*, *LPBoost*, *LSBoost*, *RobustBoost*, *RUSBoost*, *Subspace* e *TotalBoost* (SNOEK; LAROCHELLE H.AND ADAMS, 2012; BISHOP, 2006). A escolha deste classificador individual pode ser tratado como um parâmetro definido *a priori* para este modelo.

### 3.2.4 Análise de Discriminantes Lineares

O método de classificação a partir da Análise de Discriminantes Lineares, do inglês *Linear Discriminant Analysis* (LDA) aborda o problema de classificação assumindo que as funções densidade de probabilidade das classes envolvidas são normalmente distribuídas (Gaussianas) e que as matrizes de covariância das classes são idênticas (BISHOP, 2006). Com isso, o classificador se torna linear e dependente apenas das médias que caracterizam cada classe (BISHOP, 2006).

<sup>5</sup> Entropia, no contexto de árvore de decisão, corresponde à medida da homogeneidade de um conjunto de dados. Ou seja, se o conjunto é completamente homogêneo (somente dados de uma mesma classe), a entropia é zero. Por outro lado, se a amostra é dividida igualmente entre as classes, então ela tem entropia unitária.

Uma variação deste modelo, denominado Fisher's *Linear Discriminant*, pode ser usado para reduzir a dimensão das características, projetando os exemplos de treinamento em um espaço que maximiza a separação entre classes. Assim, a complexidade computacional do método pode ser reduzida, evitando a adaptação ao ruído ao invés de dados relevantes (CHERKASSKY; MULIER, 1998). Neste último caso, é necessário definir o número de dimensões a serem reduzidas *a priori*.

### 3.2.5 Máquina de Vetor de Suporte

As máquinas de vetores de suporte (SVM) foram originalmente desenvolvidas para resolver problemas de classificação usando o conceito de um hiperplano<sup>6</sup> de separação ideal, maximizando a margem de separação  $\rho$  entre classes. A motivação para maximizar  $\rho$  baseia-se em uma medição de complexidade conhecida como dimensão de Vapnik-Chervonenkis (VC) (VAPNIK, 1998), cujo limite superior é inversamente proporcional à  $\rho$ . No modelo também é incluído um termo referente ao erro de classificação (termo de penalização) para o conjunto de treinamento, considerando casos em que as classes são não-linearmente separáveis. O equilíbrio entre a maximização da margem e esse termo de penalização é parametrizada *a priori*. A definição da margem ótima e do hiperplano de separação são formulados por meio de um problema de otimização, que é normalmente solucionado utilizando o método dos multiplicadores de Lagrange (VAPNIK, 1998). Além disso, as entradas (vetores de características) entram na formulação usando somente produto interno. Com isso, o produto interno pode ser substituído pelo *kernel*, estendendo indiretamente a SVM para superfícies não-lineares de separação (VAPNIK, 1998). Existem vários tipos de *kernel* que podem ser escolhidos, levando em conta que todos devem obedecer às condições do teorema de Mercer. O tipo do *kernel* e seus respectivos parâmetros também devem ser definidos *a priori*.

---

<sup>6</sup> Hiperplano pode ser compreendido como a separação linear entre as classe do problema.

## 4 METODOLOGIA

Neste capítulo será apresentada a metodologia utilizada para o desenvolvimento do trabalho. As seções seguem explicando as bases de dados utilizadas neste trabalho, sendo elas: base de dados COOLL, a base de dados LIT<sup>1</sup> Sintética e por fim a base LIT Simulada, que é uma contribuição deste trabalho. Posteriormente, são definidos os procedimentos de pré-processamento das formas de onda das cargas coletadas, extração de características e o processo de classificação.

### 4.1 BASE COOLL

A base de dados COOLL (PICON *et al.*, 2016) é construída por formas de onda com amostragem de alta frequência (100 kHz) para sinais de tensão e corrente de várias cargas em diferentes cenários de uso, permitindo a análise de diferentes parâmetros dos métodos avaliados. Os cenários de uso especificados pelos criadores do conjunto de dados (*Université d'Orleans*) envolvem um único equipamento conectado à rede, cujos instantes de ativação e desativação são controlados. O conjunto de equipamentos compreende doze dispositivos e 42 aparelhos (classes utilizadas na classificação), conforme apresentado na Tabela 2.

**Tabela 2 – Base de Dados COOLL.**

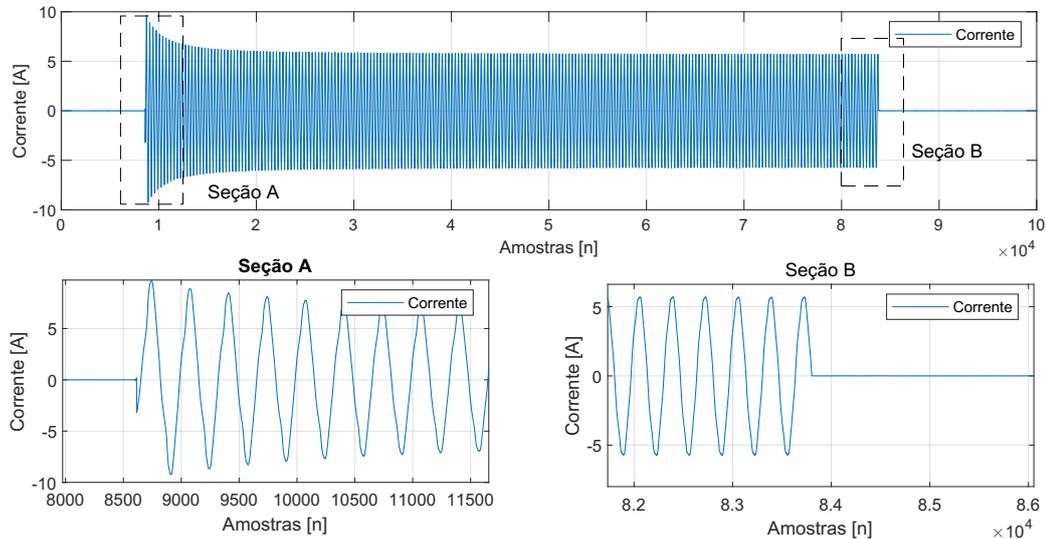
Classes	Equipamentos	Num. de Equipamentos	Num. de Sinais
1	Furadeira	6	120
2	Ventilador	2	40
3	Esmerilhadeira	2	40
4	Secador de Cabelo	4	80
5	Podador Elétrico	3	60
6	Lâmpada	4	80
7	Soprador Térmico	1	20
8	Plaina	1	20
9	Tupia	1	20
10	Lixadeira	3	60
11	Serra	8	160
12	Aspirador de Pó	7	140

**Fonte: Adaptado de Picon *et al.* (2016).**

Alguns exemplos de aquisição são apresentados nas Figuras 4 e 5, com a coleta de corrente do soprador térmico e aspirador de pó, respectivamente.

<sup>1</sup> Laboratório de Inovação Tecnológica em Sistemas Embarcados e Energia - UTFPR - Campus Curitiba - Centro.

**Figura 4 – Corrente de aquisição - Removedor de Tinha Elétrico. Seção A - Destaque para o momento em que a carga liga. Seção B - Destaque para o momento em que a carga desliga.**



**Fonte: Autoria Própria.**

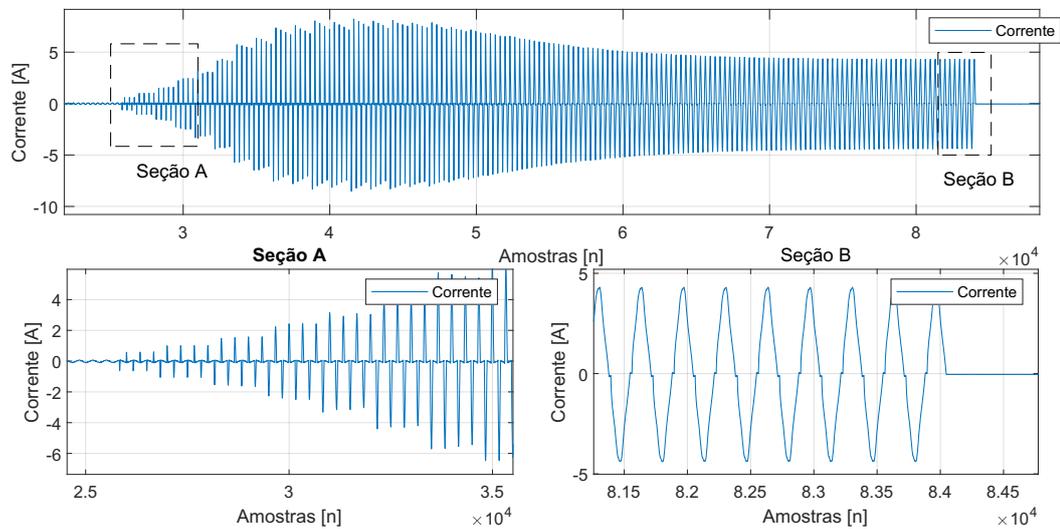
## 4.2 BASE LIT SINTÉTICA

A fim de desenvolver métodos mais precisos e eficientes para detecção e desagregação de carga, vários conjuntos de dados foram propostos, cada um com características particulares para avaliar uma ou mais etapas do processo de desagregação, conforme exposto na Seção 2.4. No entanto, os conjuntos de dados existentes não apresentam todos os requisitos relevantes para avaliar o processo de desagregação (da detecção à desagregação de carga), como: ser composto por cargas residenciais, comerciais e industriais; ter cargas de diferentes tipos (*ON* e *OFF*, ser baseado em máquina de estado, assimétrico, variável continuamente e aleatório) e possuir múltiplas cargas simultâneas. Além disso, é preciso também indicar de forma precisa o instante de ocorrência de eventos de chaveamento de carga, para que seja possível extrair as características transitórias de tensão e corrente no intervalo correto. Outras características desejáveis para bases de dados dessa natureza é ter uma taxa mínima de amostragem na ordem de kHz, garantindo a correta extração de características; modelagem de cargas múltiplas sintéticas<sup>2</sup> e naturais<sup>3</sup> e uma duração mínima de monitoramento para cargas com forma natural de 24 horas, para obter dados

<sup>2</sup> Base de dados coletada em bancadas de testes controladas.

<sup>3</sup> Base de dados coletada em ambientes não controlados, por exemplo: tomadas de residências.

**Figura 5 – Corrente de aquisição - Aspirador de Pó.**



**Fonte: Autoria Própria.**

suficientes para os testes.

Nesse sentido, o grupo de pesquisa ao qual este trabalho está vinculado, recentemente publicou um novo conjunto de dados, disponibilizado publicamente, que visa cumprir todos esses requisitos (RENAUX *et al.*, 2020b), definido como LIT *Dataset*<sup>4</sup>. Essa base de dados é composta por três diferentes *subsets*: Sintético, Natural e Simulado. O conjunto de equipamentos da base sintética compreende 16 dispositivos que equivalem as 16 classes utilizadas na classificação, conforme apresentado na Tabela 3. Vale ressaltar que esse conjunto de dados também apresenta várias cargas simultâneas. As combinações são mostradas na Tabela 5 e os diferentes ângulos na tensão da rede (ângulo da forma de onda no qual ocorre a comutação da carga) para coleta de diferentes equipamentos são apresentados na Tabela 4.

Nas Figuras 6 e 7 são apresentados alguns exemplos de aquisições do *subset* Sintético. Na primeira, uma aquisição de corrente de uma furadeira na velocidade 1, e na segunda, uma aquisição dupla de corrente, inicialmente é acionada uma lâmpada incandescente e depois um aquecedor a óleo.

<sup>4</sup> Disponível em: [http://dainf.ct.utfpr.edu.br/~douglas/LIT\\_Dataset](http://dainf.ct.utfpr.edu.br/~douglas/LIT_Dataset)

**Tabela 3 – Características do *subset* Sintético - LIT Dataset.**

Classes	Tipo de Cargas	Num. de Cargas Config.	Num. de Equip.	Num. de Formas de Onda LIT-SYN-1
1	Microondas ( <i>standby</i> e on)	2	1	32
2	Secador de Cabelo (duas velocidades)	4	2	64
3	Secador (duas potências)	2	1	32
4	Lâmpada de LED	2	2	32
5	Lâmpada Incandescente	1	1	16
6	Monitor CRT	1	1	16
7	Monitor LED	1	1	16
8	Extrator de Fumaça	1	1	16
9	Carregador Celular	2	2	32
10	Carregador Computador	2	2	32
11	Furadeira (duas velocidades)	2	1	32
12	Resistor	1	1	16
13	Ventilador	1	1	16
14	Aquecedor a Óleo (duas potências)	2	1	32
15	Estação de Solda	1	1	16
16	Aquecedor com Resistência	1	1	16

Fonte: Adaptado de Renaux *et al.* (2018).

**Tabela 4 – Ângulos de Disparo**

ID	Ângulo de Disparo (°)	ID	Ângulo de Disparo (°)
0	0	8	180
1	22.5	9	202.5
2	45	10	225
3	67.5	11	247.5
4	90	12	270
5	112.5	13	292.5
6	135	14	315
7	157.5	15	337.5

Fonte: Adaptado de Renaux *et al.* (2018).

**Tabela 5 – Combinação de Cargas**

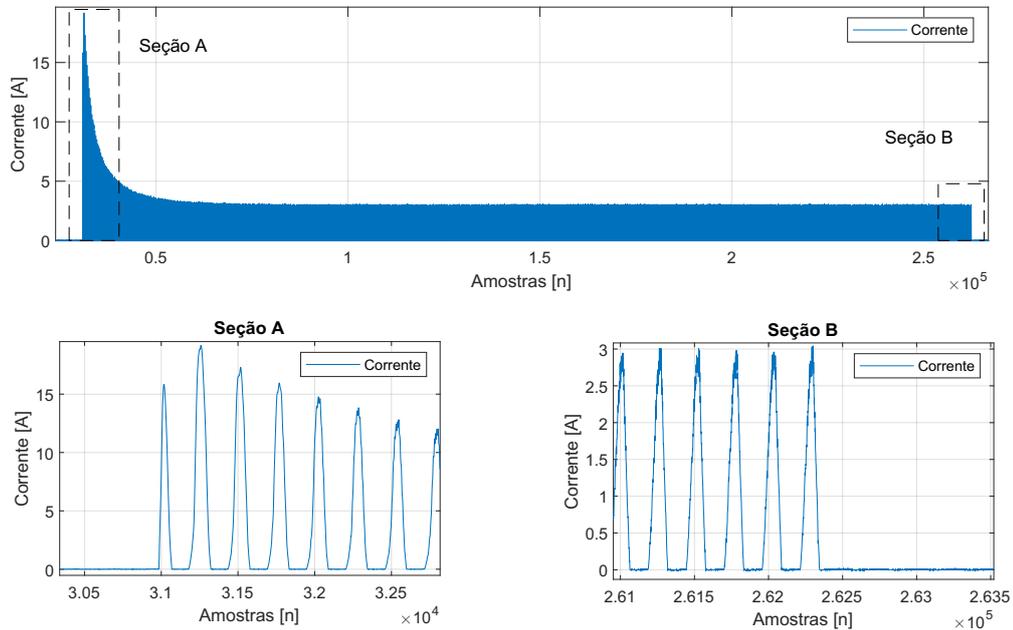
Cargas Combinadas (16 tipos de formas de onda)	Aquisições
Individual	26
2	42
3	30
8	6

Fonte: Adaptado de Renaux *et al.* (2018).

### 4.3 BASE LIT SIMULADA

O *subset* ou a base de dados LIT Simulada foi desenvolvida com sete diferentes tipos de circuitos elétricos, sendo que cada modelo contém quatro variações de potência, resultando em vinte e oito diferentes tipos de cargas. A seguir, a descrição das cargas, geração das formas de onda e as configurações da base de dados são detalhadas, já que a base LIT Simulada é uma

**Figura 6 – Corrente de aquisição individual - Furadeira Velocidade 1.**



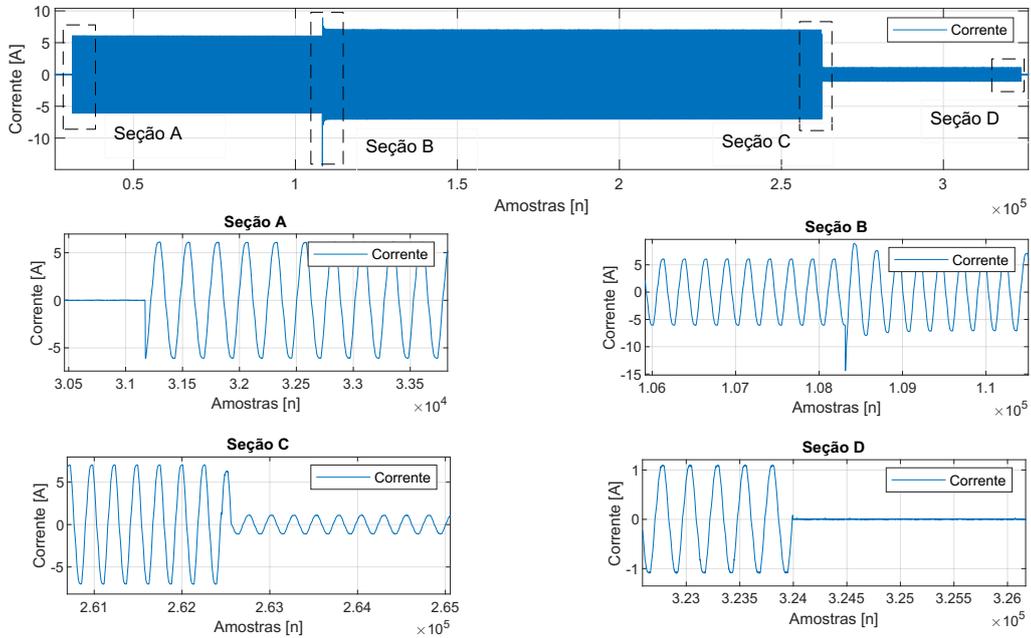
**Fonte: Autoria Própria.**

das contribuições principais do presente trabalho.

#### 4.3.1 Cargas

Os modelos de cargas disponibilizadas na base de dados são: (i) resistor; (ii) resistor e indutor; (iii) retificador a tiristor associado a um resistor; (iv) retificador a tiristor associado a um resistor e indutor; (v) retificador a diodo associado a um resistor; (vi) ponte completa a diodo associada a um capacitor; e (vii) motor universal. A escolha desses modelos de carga justifica-se porque contempla a maior parte do perfil de carga de aparelhos elétricos geralmente encontrados em unidades consumidoras (COLLIN *et al.*, 2014), como: furadeira, carregador de celular, ventilador, secador de cabelo, lâmpada LED, lâmpada incandescente, tupaia, e aspirador de pó. O diagrama esquemático da base está apresentado na Figura 8, cada bloco representa um diferente tipo de carga. O controle de comutação de cada carga é automatizado e o tempo de disparo pode ser ajustado anteriormente. Os circuitos de (i) a (vi), apresentados na Figura 9, foram implementados utilizando a Biblioteca de Eletrônica de Potência do software *MATLAB-Simulink*.

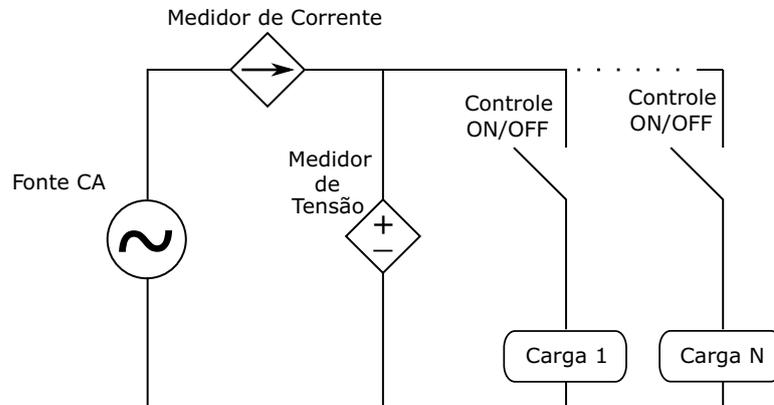
**Figura 7 – Corrente de aquisição dupla - Lâmpada Incandescente e Aquecedor a Óleo. Seção A - Destaque para o instante em que a Lâmpada Incandescente liga. Seção B - Destaque para o instante em que o Aquecedor a Óleo liga. Seção C - Destaque para o instante em que a Lâmpada Incandescente desliga. Seção D - Destaque para o instante em que o Aquecedor a Óleo desliga.**



**Fonte: Autoria Própria.**

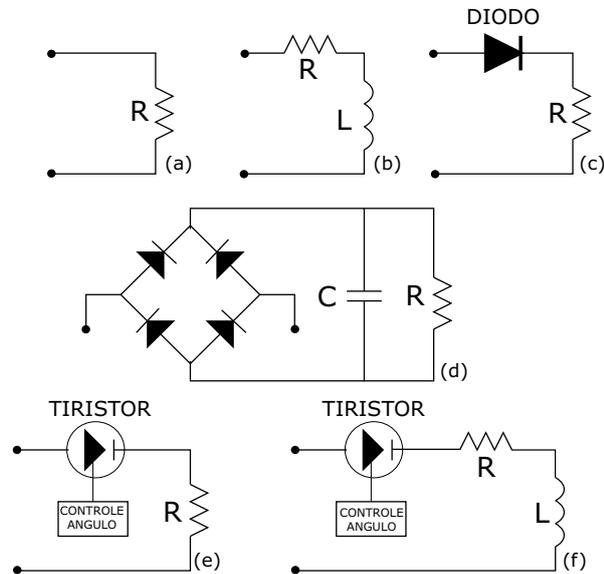
Particularmente no caso do motor universal, Figura 10, foi utilizado um modelo matemático apresentado em Ong (1998). As características do motor universal são alta velocidade sem carga e alto torque inicial. Sendo a corrente total de armadura  $I_a = I_m \cos(\omega t)$ , a corrente

**Figura 8 – Diagrama Geral da Base de Dados.**



**Fonte: Adaptado de Ancelmo et al. (2019b).**

**Figura 9 – Diagrama das cargas (a) Resistor, (b) Resistor Indutivo (c) Retificador a Diodo com Resistor, (d) Ponte Completa a Diodo com Resistor, (e) Retificador a Tiristor com Resistor, (f) Retificador a Tiristor com Resistor Indutivo.**



**Fonte: Adaptado de Ancelmo *et al.* (2019b).**

da bobina de armadura é dada por:

$$I_c = \frac{1}{a} I_m \cos(\omega t) [A], \quad (23)$$

sendo  $I_m$  a corrente de pico,  $a$  o número de espiras e  $\omega$  a frequência angular.

Em motores com suas bobinas associadas em série, o fluxo produzido pela corrente alternada  $I_a$  será também alternado no tempo. Logo, o fluxo por pólo pode ser expresso por:

$$\phi = \sum_{i=1}^{n_s} B(\theta, t) L \tau = \phi_m \cos(\omega t) \quad [Wb]. \quad (24)$$

Sabendo que o torque é definido por:

$$T_{em} = \frac{PZ}{2\pi a} \phi I_a \quad [Nm], \quad (25)$$

substituindo:

$$T_{em} = \frac{PZ}{2\pi a} \phi_m I_m \cos^2(\omega t), \quad (26)$$

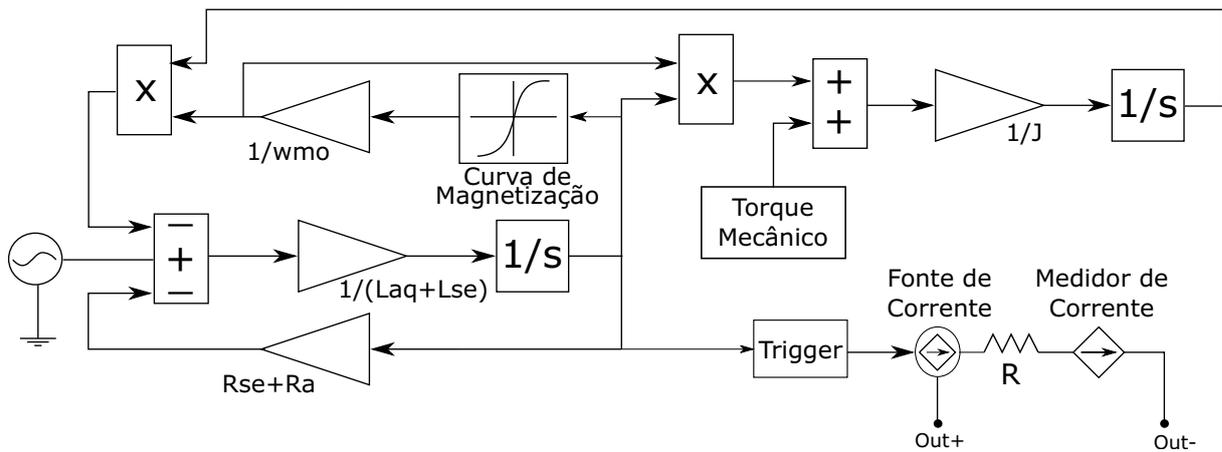
o valor médio do torque será:

$$T_{em} = \frac{PZ}{2\pi a} \frac{\phi_m I_m}{2}. \quad (27)$$

O modelo do motor universal, apresentado na Figura 10 pela implementação realizada no *MATLAB-Simulink*, é obtido a partir das equações de torque, Equações (25), (26) e (27) e dos seguintes parâmetros: potência nominal, tensão nominal do terminal, velocidade nominal,

indutância do enrolamento da armadura ( $L_{aq}$ ), indutância do enrolamento de campo em série ( $L_{se}$ ), frequência nominal da tensão de alimentação, resistência do enrolamento de armadura ( $R_a$ ), resistência do enrolamento de campo em série ( $R_{se}$ ), inércia do rotor ( $J$ ), velocidade na qual os dados da curva de magnetização foram obtidos ( $\omega_{mo}$ ). O sinal gerado pelo modelo é conectado a uma fonte geradora de corrente e em seguida conectado aos demais blocos do modelo no *MATLAB-Simulink*.

**Figura 10 – Modelo matemático do Motor Universal.**



Fonte: Adaptado de Ong (1998).

#### 4.3.2 Geração das formas de onda

Para gerar as formas de onda, de forma automática foi implementado um algoritmo que permite a variação dos parâmetros dos modelos elétricos e matemáticos. É possível configurar até quatro diferentes valores para cada componente elétrico, combinar até sete diferentes cargas na mesma forma de onda, escolher o tempo total de simulação de cada carga e alterar o ângulo de disparo de cada circuito para  $0^\circ$ ,  $45^\circ$  ou  $90^\circ$ . Como base para a simulação das cargas, quatro variações de parâmetros foram escolhidos, totalizando 28 equipamentos, que serão as 28 classes no processo de classificação e estão detalhados na Tabela 6. Os circuitos estão descritos como (i) Resistor, (ii) Resistor Indutivo (iii) Retificador a Tiristor com Resistor, (iv) Retificador a Tiristor com Resistor Indutivo, (v) Retificador a Diodo com Resistor, (vi) Ponte Completa a Diodo com Resistor e (vii) Motor universal;

**Tabela 6 – Parâmetros das Cargas**

(i)	$R_1 = 2,5\Omega$ $R_2 = 5\Omega$ $R_3 = 8\Omega$ $R_4 = 10\Omega$	(v)	$R_1 = 5\Omega$ $R_2 = 10\Omega$ $R_3 = 20\Omega$ $R_4 = 50\Omega$
(ii)	$R_1 = 10\Omega$ $R_2 = 20\Omega$ $R_3 = 30\Omega$ $R_4 = 30\Omega$ $L_1 = 10mH$ $L_2 = 10mH$ $L_3 = 10mH$ $L_4 = 100mH$	(vi)	$R_1 = 2k\Omega$ $R_2 = 1k\Omega$ $R_3 = 500\Omega$ $R_4 = 300\Omega$ $C_1 = 100\mu$ $C_2 = 100\mu$ $C_3 = 200\mu$ $C_4 = 330\mu$
(iii)	$R_1 = 10\Omega$ $R_2 = 12\Omega$ $R_3 = 15\Omega$ $R_4 = 20\Omega$	(vii)	$w_1 = 1500 rev/min$ $w_2 = 1200 rev/min$ $w_3 = 1000 rev/min$ $w_4 = 800 rev/min$
(iv)	$R_1 = 10\Omega$ $R_2 = 5\Omega$ $R_3 = 30\Omega$ $R_4 = 105\Omega$	$L_1 = 1mH$ $L_2 = 1mH$ $L_3 = 1mH$ $L_4 = 1mH$	

**Fonte: Adaptado de Ancelmo et al. (2019b).**

#### 4.3.3 Configurações da Base de Dados

Além da variação dos parâmetros de cada circuito, foram implementados seis diferentes cenários de simulação, definidos da seguinte forma: cenário ideal (DB-1); cenário com indutância parasita<sup>5</sup> (DB-2); cenário com indutância parasita e presença de harmônicos na rede (DB-3); e cenário com indutância parasita, presença de harmônicos na rede e diferentes níveis de ruído branco gaussiano, do inglês *Additive White Gaussian Noise* (AWGN)<sup>6</sup>, com relação de sinal ruído de 60 dB, 30dB e 10dB, associadas às bases DB-4, DB-5 e DB-6, respectivamente. A Tabela 7 apresenta as configurações para cada sub-base, que individualmente, possuem um total de 804 formas de onda.

No cenário ideal, DB-1, a indutância parasita e o conteúdo harmônico não são inseridos na forma de onda da tensão. No cenário DB-2 é adicionado uma indutância parasita com uma fonte de tensão ideal (sem conteúdo harmônico). Os valores do resistor e indutor,  $L = 1\mu H$  e  $R = 2n\Omega$ , foram escolhidos baseados nas características da rede do laboratório LIT. No terceiro cenário, além da indutância parasita, é adicionado conteúdo harmônico baseado em aquisições reais de tensão da rede. Os três últimos cenários de simulação consistem em uma rede com indutância parasita, conteúdo harmônico e diferentes níveis de presença de ruído AWGN com

<sup>5</sup> Indutância equivalente do sistema no ponto de medição.

<sup>6</sup> Modelo de ruído utilizado para imitar o efeito de processos aleatórios na natureza.

**Tabela 7 – Configurações da Base de Dados**

Base	Parâmetros
(DB-1)	ideal
(DB-2)	indutância parasita na rede
(DB-3)	indutância parasita na rede e harmônicos na rede
(DB-4)	indutância parasita na rede, harmônicos na rede e presença de ruído AWGN com SRN de 60dB
(DB-5)	indutância parasita na rede, harmônicos na rede e presença de ruído AWGN com SRN de 30dB
(DB-6)	indutância parasita na rede, harmônicos na rede and presença de ruído AWGN com SRN de 10dB

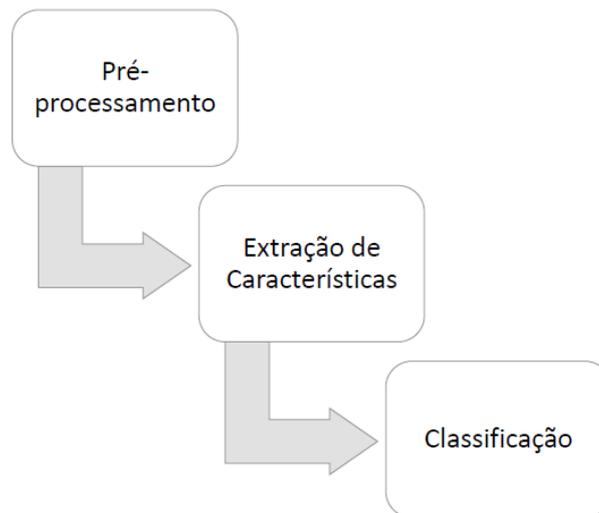
Fonte: Adaptado de Ancelmo *et al.* (2019b).

SRN.

#### 4.4 ANÁLISE DAS FORMAS DE ONDA

Para realizar a análise das formas de onda (classificação no contexto deste trabalho), o processo foi dividido em pré-processamento, extração de características e classificação, conforme mostra a Figura 11. A seguir, cada uma dessas etapas é detalhada.

**Figura 11 – Processo de Análise das Formas de Onda.**



Fonte: Autoria Própria.

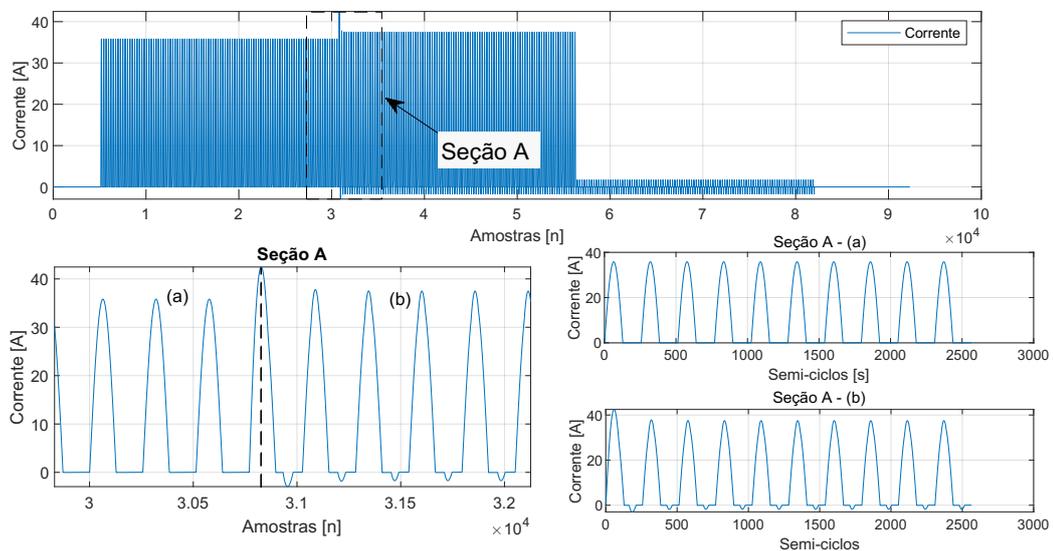
#### 4.4.1 Pré-Processamento e Extração de Características

Para a realização dos testes de extração de características e classificação de cargas, duas etapas de pré-processamento foram realizadas: a subtração de cargas e a normalização.

Em casos de cargas individuais, também foi determinado um número de ciclos de corrente para a extração das características. Essa escolha depende do método matemático de Prony utilizado. Em casos de cargas duplas ou com maior combinação, o mesmo método é utilizado. Porém, nesse segundo caso, além de determinar o número de ciclos, também é realizada a subtração a partir dos eventos de ON e OFF de cada carga.

A Figura 12 apresenta o processo de subtração de cargas. Nessa Figura é mostrado um sinal de corrente com aquisição de duas cargas, retificador a diodo e uma ponte completa a diodo. A transição das cargas está em *zoom* na Seção A. Após passar as amostras para semi-ciclos, as duas cargas são subtraídas uma da outra, resultando nos sinais das Seções A - (a) e A - (b). Dessa forma, inicialmente o sinal da Seção A passa pela extração de características, e na sequência, o sinal da Seção B.

**Figura 12 – Subtração de Cargas do Sinal de Corrente - Retificador e Ponte a Diodo.**



**Fonte: Autoria Própria.**

Convém ressaltar que o processo de subtração de formas de onda de corrente já vem sendo utilizado na literatura recente como alternativa para realizar a desagregação de cargas, conforme sugerido em Wang *et al.* (2018). Muito embora as cargas sejam não-lineares em sua maioria, Wang *et al.* (2018) mostraram que mesmo com a subtração, preservam-se as

características mais importantes das formas de onda de corrente que caracterizam cada carga.

Com base nas formas de onda subtraídas, calculam-se as características. No caso do método de Prony, essas características são: fator de amortecimento, frequência, amplitude e fase. A quantidade de valores de cada característica depende da solução utilizada para o método de Prony. No caso clássico, esse número é igual ao número de amostras do sinal. Ou seja, um sinal com 20 amostras contém 20 parâmetros, os quais resultam em 20 valores para fator de amortecimento, 20 para a frequência, 20 para a amplitude e 20 para a fase. Já para os demais métodos ele depende, da ordem do modelo, conforme mostrado em (RODRÍGUEZ *et al.*, 2018). Por exemplo, no Prony com mínimos quadrados, um sinal com 50 amostras e ordem 10, a quantidade de parâmetros e, conseqüentemente de características para fator de amortecimento, frequência, amplitude e fase, é de 10.

Após a subtração e extração das características, a etapa de classificação de cargas é iniciada. Para esta nova etapa, as características extraídas das cargas são normalizadas entre os valores de  $[-1,1]$  e inseridas no algoritmo de classificação. Esse processo é realizado pois alguns classificadores (por exemplo SVM) funcionam de forma mais eficiente quando as características estão normalizadas em determinados intervalos como  $[-1,1]$  (CHERKASSKY; MULIER, 1998).

Finalmente, vale destacar que foi considerado que as amostras de detecção são previamente definidas. Ou seja, todos os instantes de transição de carga (ON e OFF) são previamente anotados. Resultados individuais de métodos de detecção, correlatos a este trabalho, são apresentados em Lazzaretti *et al.* (2020).

#### 4.5 CLASSIFICAÇÃO

Neste trabalho, o processo de classificação é realizado em duas etapas: treinamento e teste. Durante o treinamento, é realizada a extração de características, bem como a avaliação dos parâmetros do método do Prony, levando em consideração o conjunto de dados selecionado. Depois que o classificador é treinado, um conjunto de teste é usado para avaliar o desempenho final.

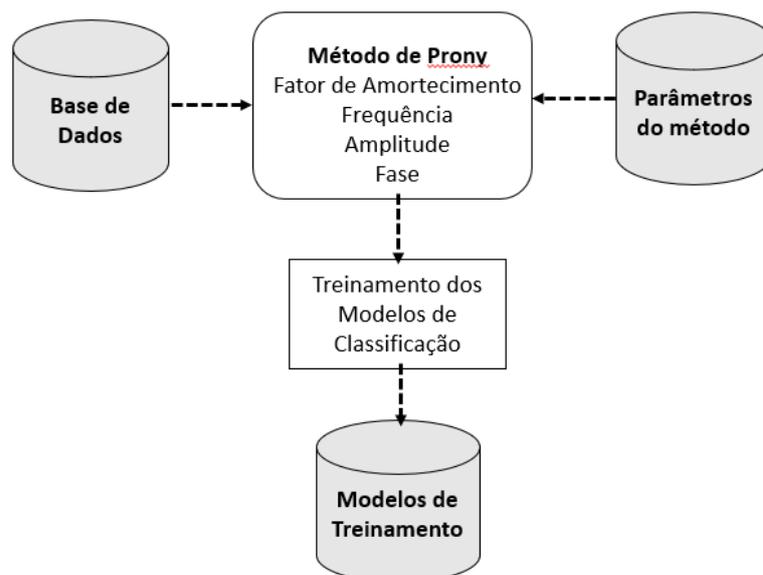
Cabe destacar que a definição dos parâmetros dos classificadores foi realizado com base no modelo de otimização Bayesiana disponível na caixa de ferramentas de aprendizagem de máquina e estatística do Matlab (SNOEK; LAROCHELLE H.AND ADAMS, 2012). Tal abordagem é equivalente ao procedimento de validação cruzada padrão (CHERKASSKY; MULIER, 1998) para selecionar parâmetros do classificador e, em geral, menos custoso computacionalmente,

conforme apresentado em Snoek e Larochelle H. and Adams (2012). Optou-se assim, por adotar essa metodologia para definição dos seguintes parâmetros dos classificadores:

- Número de vizinhos no kNN;
- Tamanho máximo das árvores de decisão;
- Classificadores individuais no método *Ensemble*;
- Número de dimensões para o discriminante linear;
- Parâmetro de regularização e do *kernel*<sup>7</sup> para a máquina de vetor suporte.

Ao final desta etapa, diferentes classificadores (modelos de classificação) são definidos para avaliação na próxima etapa, conforme apresentado na Figura 13.

**Figura 13 – Treinamento de Classificação utilizando o Método de Prony.**



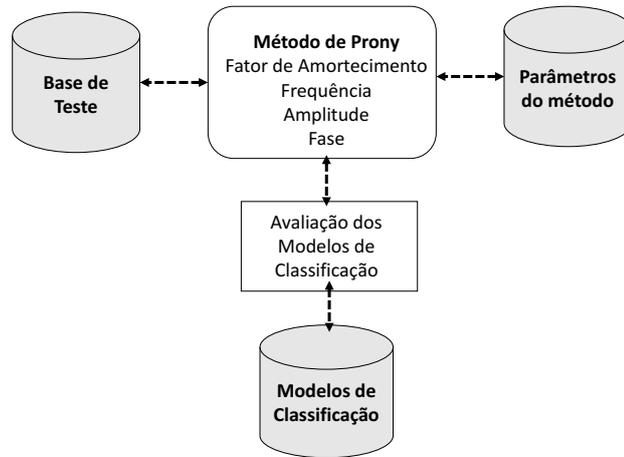
**Fonte: Adaptado de Ancelmo *et al.* (2018).**

Durante o estágio de teste, os modelos construídos na etapa anterior são avaliados. Para fazer isso, outro subconjunto do conjunto de dados é usado para realizar a avaliação (base de teste), conforme resume a Figura 14.

Por fim, é feita a avaliação geral do desempenho, que basicamente contempla o cálculo da acurácia da classificação. Para tanto, uma matriz de confusão contendo as  $Y$  classes pode ser construída, sendo  $N_{cargas}$  o número de cargas em cada *dataset* ou *subset*. Usando essa matriz

<sup>7</sup> Este parâmetro controla os erros de classificação que se busca evitar nos exemplos de treinamento.

**Figura 14 – Procedimento de Teste utilizando o Método de Prony.**



**Fonte: Adaptado de Ancelmo *et al.* (2018).**

de confusão, a acurácia para todas as classes ( $ACC$ ) pode ser calculada, utilizando a acurácia individual ( $ACC_i$ ) por classe  $i$ :

$$ACC = \left( \frac{1}{N_{cargas}} \right) \sum_{i=1}^Y ACC_i. \quad (28)$$

A acurácia individual por classe é calculada levando em conta o número de eventos corretamente classificados para a classe  $i$  ( $CCE_i$ ), dividido pelo número total de exemplos desta mesma classe ( $TNE_i$ ):

$$ACC_i = \frac{CCE_i}{TNE_i}. \quad (29)$$

## 5 RESULTADOS

Neste capítulo é apresentados os principais resultados obtidos no desenvolvimento da base de dados simulada e na extração de características a partir do método de Prony, combinada à classificação, para análise das cargas nas diferentes bases de dados utilizadas. De forma geral, os experimentos para a base de dados proposta, bem como seus objetivos, estão sumarizados a seguir:

- Validação das formas de onda adquiridas com cargas reais (Seção 5.1): o objetivo deste experimento é garantir que o conjunto de formas de onda simulado representa cargas reais e que pode ser generalizado para outros parâmetros e potências de carga;
- Resultados de classificação das formas de onda (Seção 5.3.1): nesta etapa são utilizados métodos de extração de características da literatura para posterior classificação com as seis sub-bases propostas na Base Simulada deste trabalho. O objetivo deste experimento é analisar o comportamento dos extratores e classificadores a partir da inserção de ruídos e harmônicos na rede, demonstrando quais tipos de análise essas sub-bases podem fornecer para a comunidade de NILM.

Já para o método de Prony, os experimentos e seus objetivos são:

- Análise de frequência de amostragem e número de ciclos (Seção 5.2.1): o objetivo deste experimento é avaliar o impacto da frequência de amostragem e, conseqüentemente, do número de amostras, na estimativa dos parâmetros do método de Prony;
- Análise dos métodos de Prony reduzidos (Seção 5.2.2): o objetivo é avaliar as cinco abordagens matemáticas do método de Prony para a estimação de parâmetros;
- Classificação e análise do número de ciclos (Seção 5.3.2): o objetivo é utilizar algoritmos de classificação para observar o impacto do número de ciclos na acurácia do método.
- Classificação e métodos de Prony reduzidos (Seção 5.3.3): o objetivo nessa seção é analisar a acurácia de classificação a partir das cinco abordagens do método de Prony;
- Comparativo entre as soluções de Prony (Seção 5.4): nesta seção é feito um resumo dos principais resultados obtidos com as cinco abordagens matemáticas do método de Prony;

- Comparativo com métodos da literatura (Seção 5.5): o objetivo nesta seção é avaliar o desempenho do método de Prony como extrator de características (melhor caso anteriormente avaliado com a comparação de frequência de amostragem e dos diferentes métodos de Prony reduzidos), comparando sua acurácia e complexidade computacional com métodos da literatura. Foram utilizadas as bases COOLL, LIT Sintética e LIT Simulada.

Nas seções subseqüentes, serão detalhados os experimentos acima descritos.

## 5.1 BASE LIT SIMULADA

Com o objetivo de validar as cargas simuladas e escolher possíveis variações de parâmetros para a base de dados, alguns testes foram realizados em uma bancada composta por componentes eletrônicos como diodo, resistores, indutores e Semicondutor sensível à temperatura, do inglês *Negative Temperature Coefficient* (NTC). Foram montados na bancada de testes os seguintes circuitos: (a) resistor; (b) resistor e indutor; (c) retificador a diodo associado a um resistor; (d) retificador a diodo associado a um resistor indutor; (e) ponte completa a diodo associada a um capacitor; (f) retificador a tiristor associado a um resistor; e (g) retificador a tiristor associado a um resistor indutor. Os valores dos componentes utilizados estão apresentados na Tabela 8.

**Tabela 8 – Parâmetros dos componentes reais.**

Circuito	Valores
(a)	$R = 50\Omega$
(b)	$R = 100\Omega \quad L = 1H$
(c)	$R = 100$
(d)	$R = 100\Omega \quad L = 1H$
(e)	$R = 300\Omega \quad C = 600\mu F$
(f)	$R = 100$
(g)	$R = 100\Omega \quad L = 1H$

**Fonte: Autoria Própria.**

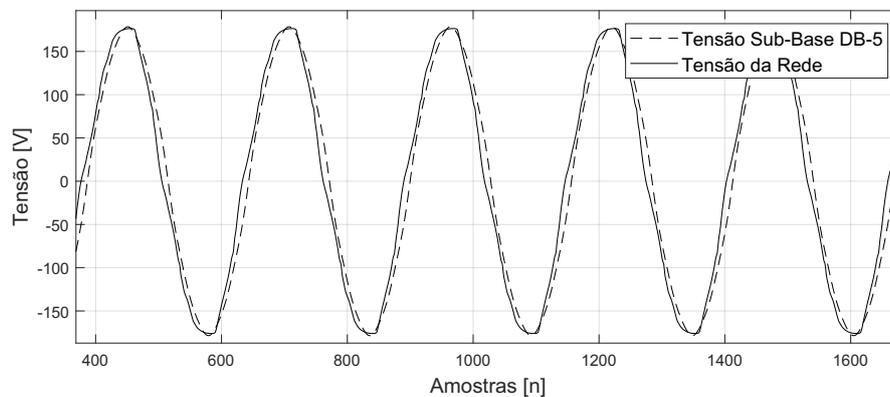
Analisando a Série de *Fourier* da forma de onda da rede de tensão, adquirida durante os testes, extraiu-se a amplitude e a fase de cada harmônico. Os valores obtidos estão apresentados na Tabela 9 e foram incluídos no bloco de fonte de tensão, usado em todas as simulações com conteúdo harmônico e na Figura 15 é apresentada uma comparação das curvas de tensão da rede real e simulada pela base.

Utilizando a tensão adquirida, os circuitos utilizados nos testes foram implementados no *MATLAB-Simulink* com os mesmos parâmetros apresentados na Tabela 8. Então, as formas

**Tabela 9 – Parâmetros da Tensão da Rede.**

Harm.	Amp. (V)	Fase (°)	Harm.	Amp. (V)	Fase (°)
1	179,000	-30,158	21	0,222	-79,499
3	3,587	17,677	23°	0,041	86,349
5°	4,503	22,734	25°	0,110	57,642
7°	1,048	179,630	27°	0,077	173,5639
9°	2,095	-143,576	29°	0,089	148,640
11°	0,242	-58,111	31°	0,097	-92,878
13°	0,492	-106,364	33°	0,015	38,334
15°	0,778	75,808	35°	0,054	16,602
17°	0,160	-10,185	37°	0,050	87,020
19°	0,027	126,329	39°	0,042	90,883

Fonte: Autoria Própria.

**Figura 15 – Comparativo do Sinal de Tensão da Rede e o Sinal de Tensão da base LIT Simulada.**

Fonte: Autoria Própria.

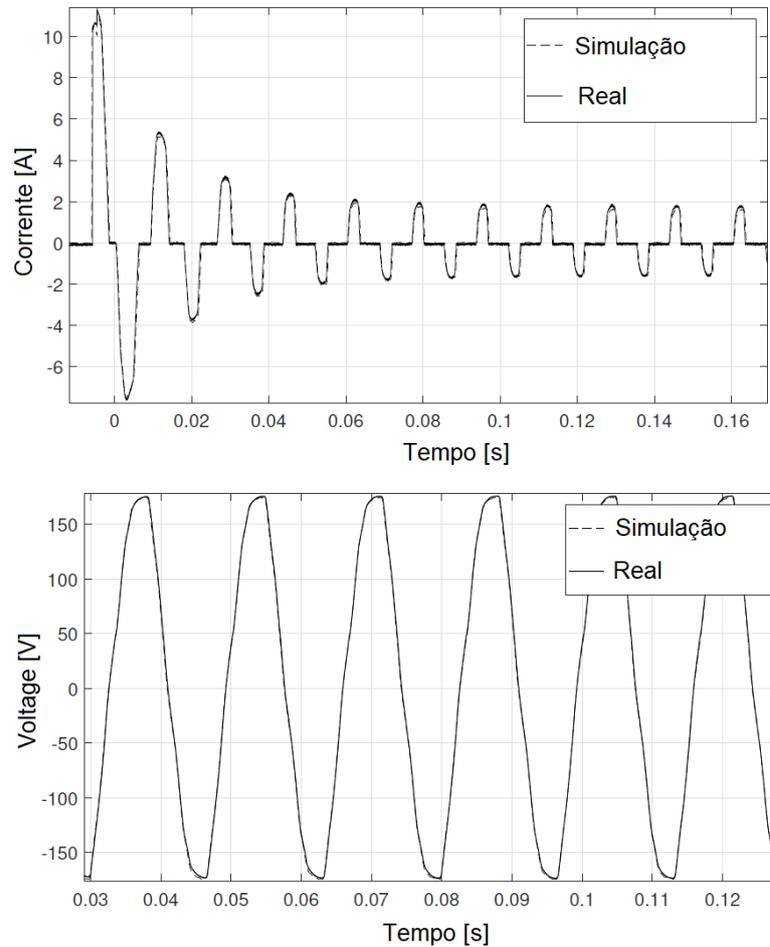
de onda de corrente e tensão reais e simuladas foram comparadas, como mostra a Figura 16.

Uma maneira de validar as formas de onda obtidas é comparar parâmetros, como valores máximos de corrente transitória e permanente e valores máximos de tensão, Fator de Potência (FP) e Erro Quadrático Médio (EQM) da forma de onda medida e simulada, conforme sugerido em Bacca *et al.* (2009) e resumido nas Tabelas 10 e 11 para as simulações propostas neste trabalho.

Pode-se observar nas Tabelas 10 e 11 que os valores de pico de tensão e corrente, EQM e o FP são próximos dos valores reais. Logo, pode-se considerar que a geração das formas de onda da simulação são aproximações representativas das formas de onda adquiridas empiricamente.

Para realizar a validação do modelo matemático correspondente ao motor universal, foi utilizado um motor real de uma furadeira de 750W com duas velocidades. Inicialmente foi realizada a aquisição das formas de onda de tensão e corrente do motor. Depois foram realizados

**Figura 16 – Corrente da Ponte Completa a Diodo e Tensão da Rede.**



**Fonte: Autoria Própria.**

testes com corrente alternada e contínua para obter os valores de resistências e indutâncias de campo e rotor. Os parâmetros obtidos foram aplicados no modelo matemático (ver Figura 10 na Seção 4.3.1), são apresentados na Tabela 12.

A Figura 17 apresenta o comparativo das formas de onda real e simulada do motor universal, com um  $EQM = 0,03A^2$ . Como pode-se observar, as formas de onda apresentam valores similares no transitório e em regime permanente.

A Figura 18 apresenta um exemplo de forma de onda de corrente gerada pela sub-base DB-3, composta pela combinação de duas cargas, resistor com indutor e motor universal, acionados em diferentes momentos.

Como pode ser visto na Figura 18, a carga possui uma combinação dupla e cada transitório das cargas, *ON* e *OFF* estão destacados nas seções A, B, C e D. Os instantes de *ON* e *OFF* são controlados no nível da amostra a uma frequência de amostragem de 15,360 kHz. Para cada evento (*ON* ou *OFF*), a carga também é identificada corretamente. Com isso,

**Tabela 10 – Validação dos Parâmetros - Valores Máximos de Corrente e Tensão.**

Circuito	Valor Máximo de Tensão		Valor Máximo de Corrente	
(a)	Real	174,200	Real	3,598
	Sim	174,300	Sim	3,599
(b)	Real	176,100	Trans. Real	0,592
	Sim	174,600	Trans. Sim	0,582
			Perm. Real	0,406
			Perm. Sim	0,402
(c)	Real	177,500	Real	1,831
	Sim	175,600	Sim	1,827
(d)	Real	176,700	Real	0,687
	Sim	174,500	Sim	0,684
(e)	Real	175,200	Trans. Real	11,260
	Sim	174,600	Trans. Sim	11,310
			Perm. Real	1,790
			Perm. Sim	1,740
(f)	Real	177,400	Real	1,699
	Sim	172,500	Sim	1,700
(g)	Real	174,800	Real	0,350
	Sim	172,300	Sim	0,340

**Fonte: Autoria Própria.**

**Tabela 11 – Validação dos Parâmetros (FP e EQM).**

	FP		EQM	
	Real	Sim	Tensão	Corrente
(a)	1,000	1,000	$2,460 \times 10^{-4}$	$1,800 \times 10^{-4}$
(b)	0,340	0,320	$4,260 \times 10^{-4}$	0,004
(c)	1,000	0,999	$3,350 \times 10^{-4}$	$1,010 \times 10^{-4}$
(d)	0,787	0,754	$5,230 \times 10^{-4}$	0,002
(e)	0,848	0,887	$1,490 \times 10^{-4}$	0,007
(f)	0,540	0,475	$3,640 \times 10^{-4}$	0,004
(g)	0,180	0,160	0,002	0,004

**Fonte: Autoria Própria.**

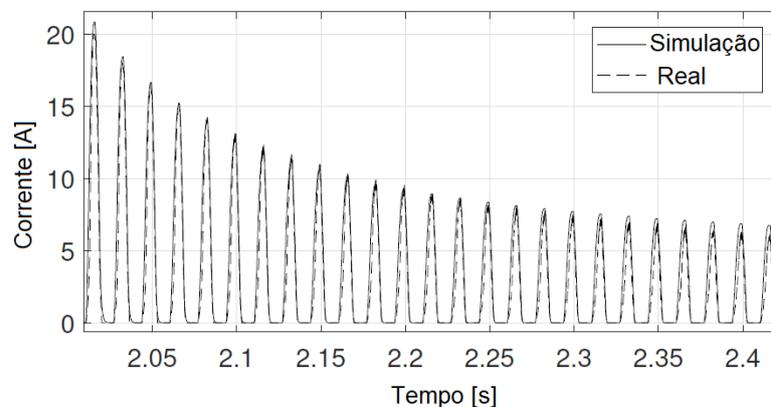
cada subconjunto do banco de dados fornece até sete tipos de circuitos elétricos, cada um contendo até quatro variações de potência, resultando em 28 cargas diferentes. Foi gerada uma média de 274 formas de onda por classe para cada subconjunto, incluindo a combinação de duas, três, quatro e cinco cargas na mesma forma de onda, além das cargas individuais. Essa combinação foi realizada por meio de um *script* no software *MATLAB* que automatiza a geração de eventos, usando instantes predefinidos e tipos de carga que são acionados em cada simulação. Os algoritmos e diagramas de blocos estão disponíveis no *GitHub* da autora<sup>1</sup>.

<sup>1</sup> <https://github.com/hellenancelmo>.

**Tabela 12 – Parâmetros do Motor Universal.**

Parâmetro	Valores
Potência Nominal	325 W
Tensão Nominal no Terminal	120 Vrms
Velocidade Nominal	2800 rev/min
Indutância do enrolamento de armadura $L_{aq}$	10 mH
Indutância do enrolamento de campo em série $L_{se}$	26 mH
Frequência Nominal da Tensão de Alimentação	60 Hz
Resistência do enrolamento de armadura $R_a$	0,6 $\Omega$
Resistência do enrolamento de campo $R_{se}$	0,1 $\Omega$
Inércia do Rotor $J$	0,0015 Kg $m^2$
Parâmetro Curva de Magnetização $\omega_{mo}$	$2\pi\omega/60$
Parâmetro Curva de Magnetização $\omega$	1500 rev/min.

Fonte: Fonte: Adaptado de Ong (1998).

**Figura 17 – Comparação das Formas de Onda de Corrente Real e Simulada do Motor Universal.**

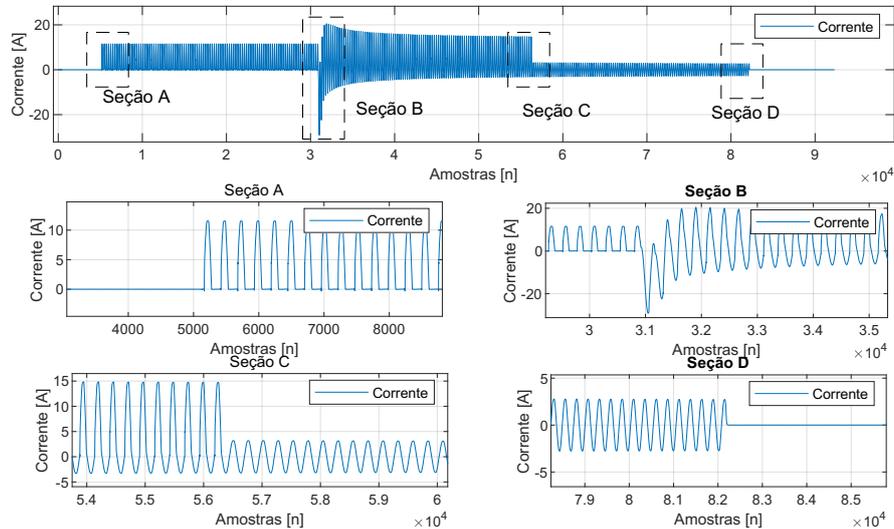
Fonte: Autoria Própria.

## 5.2 MÉTODO DE PRONY

No Algoritmo de Prony, a quantidade de amostras para reconstruir o sinal é proporcional ao número de características geradas pelo método (frequência, amortecimento, fase e amplitude). Quanto maior for o número, maior será o esforço computacional e mais complexo se torna o processo de classificação, de acordo com a problema da *maldição da dimensionalidade* (CHERKASSKY; MULIER, 1998). Ou seja, à medida que a dimensionalidade aumenta, o volume do espaço de características aumenta tão rápido que os dados disponíveis se tornam esparsos, como apresentado em Tax e Duin (2002). Essa dispersão é problemática para qualquer método que requeira significância estatística, como métodos de aprendizado de máquina. Para obter um resultado estatisticamente confiável, a quantidade de dados necessária para suportar o resultado geralmente cresce exponencialmente com a dimensionalidade.

Para minimizar o problema da dimensionalidade do método, reduzindo assim o número de características, duas abordagens foram utilizadas: análise do número de ciclos e modificação

**Figura 18 – Forma de Onda da Corrente com Resistor Indutivo e Motor Universal. Seção A - Destaque para o instante em que o Resistor Indutivo liga. Seção B - Destaque para o instante em que o Motor Universal liga. Seção C - Destaque para o instante em que o Resistor Indutivo desliga. Seção D - Destaque para o instante em que o Motor Universal desliga.**



**Fonte: Autoria Própria.**

da solução de Prony a partir da aplicação de diferentes métodos matemáticos, os Métodos de Prony Reduzidos.

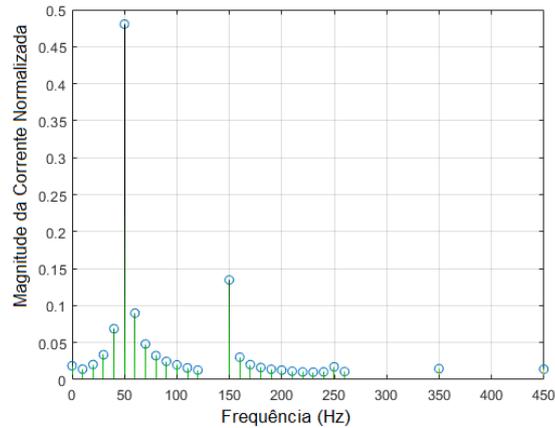
### 5.2.1 Análise de Frequência de Amostragem e Número de Ciclos

Para bases de dados com altas taxas de amostragem ( $> 15\text{kHz}$ ), como a COOLL, é necessário minimizar o número de amostras (reduzir a taxa), sem perder as informações principais de cada carga. Para tanto, optou-se por decimar<sup>2</sup> o sinal original. Para definir a taxa de decimação, foi realizada uma análise de harmônicos da Transformada Discreta de Fourier, do inglês *Discrete Fourier Transform* (DFT) das 840 cargas do banco de dados COOLL, determinando os componentes de harmônicos máximos do sinal. O critério era que a energia máxima do  $i$ -ésimo harmônico deveria ser maior que 1%. Como o valor máximo encontrado foi 500 Hz, uma taxa de decimação de 100 foi escolhida para que a taxa de amostragem do sinal se tornasse 1 kHz. A Figura 19 mostra as características típicas da carga com conteúdo de maior frequência no banco de dados COOLL (Aspirador de Pó).

Para avaliar o desempenho de reconstrução do método de Prony, foram analisados os

<sup>2</sup> Decimação é o processo de subamostragem e filtragem (passa-baixas) para reduzir a frequência de amostragem de um sinal.

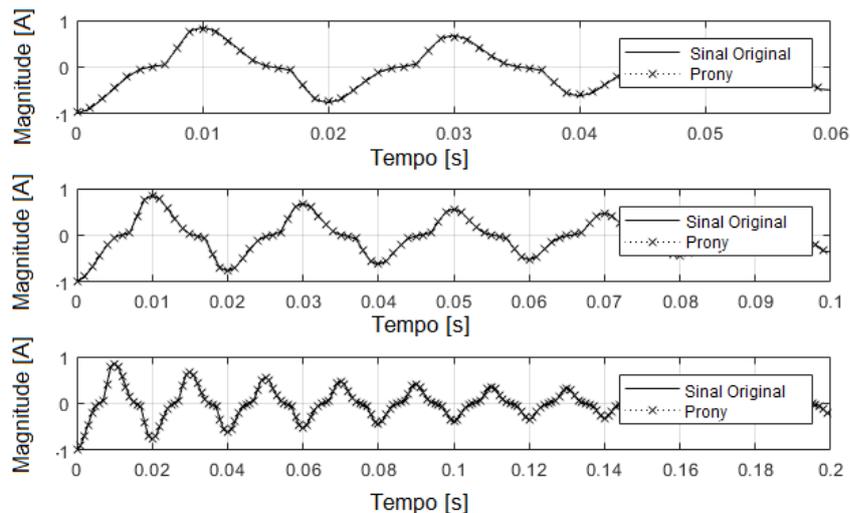
**Figura 19 – Análise das Frequências Relevantes (Conteúdo Harmônico).**



**Fonte: Autoria Própria.**

ciclos transientes de cada sinal, variando a janela de reconstrução entre 3, 5 e 10 ciclos da rede elétrica, conforme mostrado na Figura 20.

**Figura 20 – Reconstrução do método de Prony com 3, 5 e 10 ciclos.**



**Fonte: Autoria Própria.**

A Tabela 13 mostra o número de características estimadas pelo método de Prony o número de ciclos (tamanho da janela). Nesta Tabela, o EQM da reconstrução também é mostrado, para todas as 840 formas de onda do conjunto de dados COOLL. Como pode-se observar, quanto maior o número de ciclos, menor o erro de reconstrução. Porém, como o número total de características aumenta significativamente, decidiu-se realizar os testes com 3, 5 e 10 ciclos para classificação com a estimativa das características feita pelo método de Prony.

Tabela 13 – Reconstrução do método de Prony.

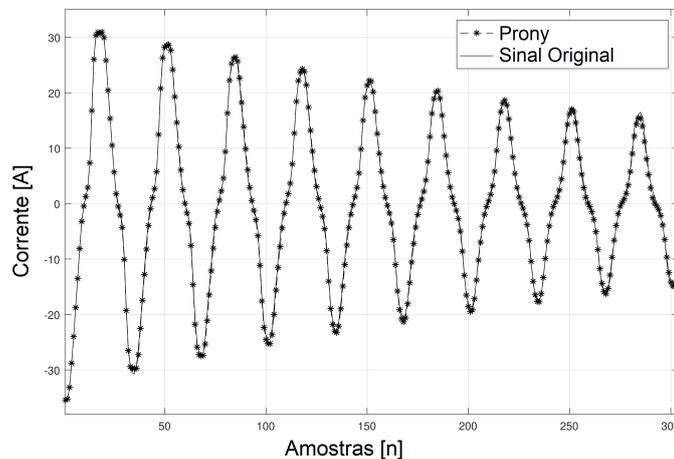
Número de Ciclos	Características Individuais	Total de Características	EQM
3	30	120	0,541
5	50	200	0,158
10	100	400	0,037

Fonte: Autoria Própria.

### 5.2.2 Métodos de Prony Reduzido

Além da análise de conteúdo harmônico, outra abordagem utilizada para redução da dimensionalidade nas características foi o uso dos métodos de Prony Reduzidos, a saber: Prony MQ, Prony MQT, Prony MP e Prony IIR. Na análise de reconstrução de cada abordagem do Prony reduzido, foram usados 20 ciclos para o banco de dados COOLL e 40 ciclos para o banco de dados LIT Sintético, a fim de representar as condições transitórias e de estado estacionário. Uma análise visual e o EQM foram utilizados para avaliar o desempenho de reconstrução do método de Prony. Nas Figuras 21 e 22, uma comparação entre o sinal original e a reconstrução do método de Prony nos bancos de dados COOLL e LIT Sintético é apresentada.

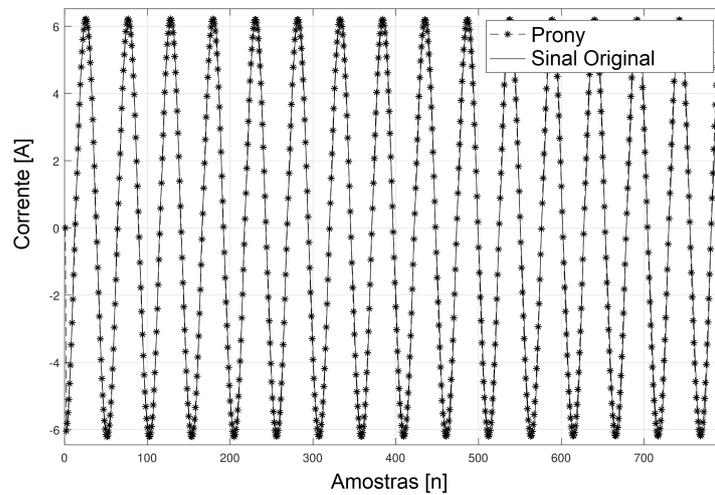
Figura 21 – Reconstrução do sinal de corrente da Carga Base COOLL a partir do Método de Prony utilizando o método *Matrix Pencil*, ordem 10.



Fonte: Autoria Própria.

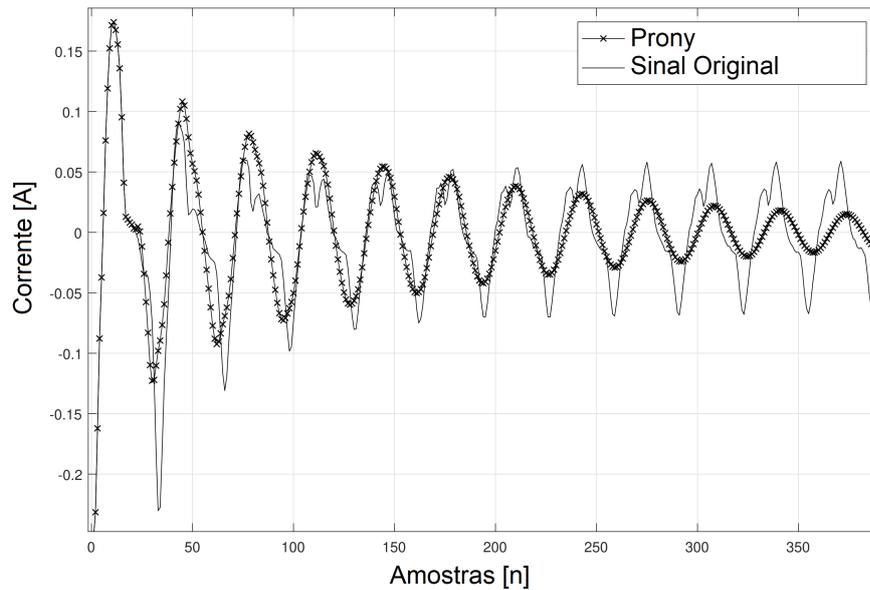
Como se pode observar, a reconstrução do sinal durante os transientes está de acordo com o sinal original. Esse resultado é normalmente observado para as abordagens de Prony na maioria das cargas das bases COOLL e LIT Sintética. Porém, para algumas cargas particulares com conteúdo harmônico relativamente alto, dependendo dos modelos e ordem utilizados, a reconstrução apresentou uma diferença entre o sinal original, como pode ser observado na Figura 23.

**Figura 22 – Reconstrução da Carga Base LIT a partir do Método de Prony utilizando o método *Matrix Pencil*, ordem 10.**



**Fonte: Autoria Própria.**

**Figura 23 – Reconstrução do Sinal de Corrente da Carga Secador de Cabelo a partir do Método de Prony na Base LIT.**



**Fonte: Autoria Própria.**

A reconstrução do método de Prony com uma diferença do sinal original foi observada principalmente para cargas de micro-ondas e monitor CRT no conjunto de dados LIT Sintética. Algumas características iniciais do regime transitório são mantidas, juntamente com o efeito de amortecimento, mas a distorção da forma de onda em estado estacionário é diferente, devido à ordem selecionada para o modelo (10). As Tabelas 14 e 15 apresentam o EQM da reconstrução do sinal de corrente das cargas, considerando um ciclo, para as diferentes abordagens do método.

Para os casos em que o método divergiu, o cálculo não foi realizado.

**Tabela 14 – EQM para a Base COOLL.**

Polinomial	Ordem	MQ (A <sup>2</sup> )	MQT (A <sup>2</sup> )	MP (A <sup>2</sup> )	IIR (A <sup>2</sup> )
33,519	10	3,736	25,963	3,681	0,668
	20	1,802	22,279	3,360	–
	30	0,920	–	3,750	–
	40	0,869	–	3,680	–
	50	0,837	–	3,690	–

**Fonte: Autoria Própria.**

**Tabela 15 – EQM para a Base LIT Sintética.**

Polinomial	Ordem	MQ (A <sup>2</sup> )	MQT (A <sup>2</sup> )	MP (A <sup>2</sup> )	IIR (A <sup>2</sup> )
6,398	10	0,680	159,000	0,567	0,130
	20	0,490	67,590	0,299	–
	30	0,272	4,029	0,224	–
	40	0,226	-	0,229	–
	50	0,180	-	0,361	–

**Fonte: Autoria Própria.**

Como se pode observar, nas bases de dados COOLL e LIT Sintética, os melhores EQM são observados para Filtro IIR de 10<sup>a</sup> ordem, resultando em 0,66 A<sup>2</sup> e 0,13 A<sup>2</sup>, respectivamente. Nos métodos MQ e MP, quanto maior a ordem, menor o EQM. O impacto na classificação será avaliada na seção 5.3.3.

### 5.3 RESULTADOS DE CLASSIFICAÇÃO

Os resultados de classificação serão apresentados em três partes. A primeira apresentará os resultados de classificação da base LIT Simulada utilizando extratores de características do estado da arte. Em seguida, os resultados de classificação do método de Prony com características determinadas a partir da análise harmônica. Por fim, o resultados de classificação dos métodos de Prony reduzidos.

#### 5.3.1 Base de Dados LIT Simulada

Com o intuito de avaliar o impacto do conteúdo harmônico e ruído propostos para a base simulada na classificação de cargas, optou-se por inicialmente demonstrar o resultado de classificação usando três diferentes extratores de características da literatura, combinados com um classificador *Ensemble*. Os três extratores de características utilizados são: Transformada Discreta de Fourier apresentada em Nait-Meziane *et al.* (2019), Transformada *Wavelet* Discreta

proposta em Chang *et al.* (2013) e Curvas V-I proposto em Hassan *et al.* (2014). O objetivo é calcular a acurácia de classificação para as 28 cargas em cada subconjunto. A Tabela 16 resume a precisão média para 10 combinações diferentes de conjuntos de treinamento e teste em cada subconjunto.

**Tabela 16 – Acurácia do Método Ensemble vs Base de Dados.**

Sub-base	V-I(%)	DFT(%)	DWT(%)
DB-1	92 ± 0,050	99 ± 0,002	99 ± 0,001
DB-2	89 ± 0,050	99 ± 0,002	99 ± 0,001
DB-3	91 ± 0,070	99 ± 0,005	99 ± 0,007
DB-4	81 ± 0,028	99 ± 0,001	99 ± 0,001
DB-5	70 ± 0,012	94 ± 0,002	96 ± 0,003
DB-6	41 ± 0,072	78 ± 0,012	70 ± 0,006

**Fonte: Autoria Própria.**

Como se pode observar, à medida que condições não ideais são inseridas nas bases, como harmônicos e principalmente ruído, o desempenho da classificação diminui. Particularmente, o método V-I é o mais afetado pela inserção de ruído. Tal resultado é justificado pela degradação das características extraídas das formas de onda, conforme sugerido em Hassan *et al.* (2014). Além disso, em cargas múltiplas, uma carga de alta potência seguida de uma carga de baixa potência, quando combinadas, a tendência é que a carga de baixa potência acabe perdendo características da sua forma de onda, principalmente durante o estado transiente. Vale ressaltar que este tipo de análise não é possível em outras bases de dados públicas, uma vez que o nível de ruído e conteúdo harmônico não é controlado nessas demais bases. Destaca-se que futuros usuários da base simulada poderão fazer uso desta característica para avaliar a robustez ou limitações dos seus respectivos métodos de extração de características.

### 5.3.2 Análise do Número de Ciclos

De forma complementar à análise apresentada na Subseção 5.2.1, aqui são apresentados os resultados de classificação quanto à análise harmônica na base COOLL. A Tabela 17 mostra um desempenho médio de classificação, que corresponde à pontuação média por classe, considerando as classes da base de dados COOLL, bem como diferentes números de ciclos usados na estimativa do Prony. É importante ressaltar que esta Tabela apresenta os resultados para o melhor conjunto de parâmetros estimados durante o procedimento de otimização dos parâmetros classificadores aplicado na etapa de treinamento.

Ao analisar a Tabela 17, pode-se ver que o número de ciclos possui uma relação inversa

**Tabela 17 – Acurácia de Classificação para Diferentes Números de Ciclos.**

Número de Ciclos	Média de Acurácia (%)		
	ENS	kNN	SVM
3	92 ± 0,058	85 ± 0,079	83 ± 0,041
5	86 ± 0,028	28 ± 0,082	33 ± 0,030
10	77 ± 0,015	42 ± 0,023	20 ± 0,045

**Fonte: Autoria Própria.**

à acurácia do classificador e diretamente proporcional à quantidade de características geradas pelo método de Prony. Portanto, inserir um número maior de características, mantendo a mesma quantidade de exemplos para treinamento e teste, provoca uma diminuição na acurácia, devido ao problema da dimensionalidade. O melhor cenário de classificação foi obtido com a reconstrução de 3 ciclos, inserindo 120 características no classificador Ensemble, com 92% de acerto médio nas classes analisadas.

### 5.3.3 Métodos de Prony Reduzido

Seguindo a análise feita na Subseção 5.2.2 em termos de erro de reconstrução, aqui são apresentados os resultados de classificação para os métodos de Prony reduzido com as duas bases de dados, COOLL e LIT Sintética. Com as formas de onda individuais e múltiplas de cada banco de dados, o desempenho de classificação para as cinco abordagens diferentes de Prony (Polinomial, MQ, MQT, MP e filtro IIR) foram avaliadas, usando um classificador *Ensemble* para identificar cada uma das cargas em cada subconjunto. As Tabelas 18 e 19 resumem a melhor acurácia média para dez diferentes combinações de treinamento e conjunto de teste para cada abordagem de Prony, com diferentes ordens de modelo, para bancos de dados LIT Sintética e COOLL, respectivamente.

**Tabela 18 – Acurácia Média *Ensemble* - Base COOLL.**

Método	Opções do Método	Acurácia (%)
Polinomial	-	95 ± 0,024
Mínimos Quadrados (MQ)	ordem: 10	95 ± 0,021
Mínimos Quadrados Totais (MQ)	ordem: 10	96 ± 0,013
<i>Matrix Pencil</i> (MP)	ordem: 10	98 ± 0,004
Filtro IIR	ordem: 50 subordem: 10	94 ± 0,007

**Fonte: Autoria Própria.**

Conforme mostrado na Tabela 18, a abordagem MP de ordem 10 foi a melhor combinação para o banco de dados COOLL, classificando corretamente 98% das formas de onda. O mesmo é observado para o banco de dados LIT Sintética, conforme mostrado na Tabela 19, com

**Tabela 19 – Acurácia Média *Ensemble* - Base LIT Sintética.**

Método	Opções do Método	Acurácia (%)
Polinomial	-	87 ± 0,009
Mínimos Quadrados (MQ)	ordem: 10	89 ± 0,016
Mínimos Quadrados Totais (MQT)	ordem: 10	88 ± 0,008
<i>Matrix Pencil</i> (MP)	ordem: 10	97 ± 0,013
Filtro IIR	ordem: 50 subordem: 10	83 ± 0,009

Fonte: Autoria Própria.

97% das formas de onda classificadas corretamente no conjunto de teste.

Os principais problemas de classificação incorreta foram observados para o banco de dados LIT Sintética com múltiplas condições de carga, ou seja, várias cargas conectadas simultaneamente, particularmente quando uma carga de alta potência é seguida por uma carga de baixa potência, resultando na predominância de características de carga de alta potência.

Ressalta-se que melhores casos de reconstrução apresentados na Subseção 5.2.2 não resultam, necessariamente, em melhores taxas de classificação. Por exemplo, a abordagem IIR apresenta os menores valores de EQM, mas o desempenho de classificação é de 94% e 83% para os bancos de dados COOLL e LIT Sintética, respectivamente. O desempenho da classificação é proporcional à ordem de reconstrução, pois um maior número de características pode degradar o desempenho do classificador, devido ao problema da dimensionalidade do método (ANCELMO *et al.*, 2018).

#### 5.4 COMPARATIVO DAS ABORDAGENS DO MÉTODO DE PRONY

Foram apresentadas cinco diferentes abordagens do método de Prony: Polinomial, Mínimos Quadrados, Mínimos Quadrados Totais, *Matrix Pencil* e Filtro IIR. Os parâmetros para avaliação dos métodos foram EQM e acurácia de classificação utilizando o classificador Ensemble. A Tabela 20 apresenta os parâmetros em função do método e da base de dados analisada.

**Tabela 20 – Comparativo das Abordagens do Método de Prony.**

Método	EQM (A <sup>2</sup> )		Acurácia (%)	
	COOLL	LIT Sint.	COOLL	LIT Sint.
Polinomial	35,519	6,390	95 ± 0,024	87 ± 0,009
Mín. Quadrados	3,730	0,680	95 ± 0,021	89 ± 0,016
Mín. Quadrados Totais	25,960	159,000	96 ± 0,013	88 ± 0,008
<i>Matrix Pencil</i>	3,680	0,560	98 ± 0,004	97 ± 0,013
Filtro IIR	0,660	0,130	94 ± 0,007	83 ± 0,009

Fonte: Autoria Própria.

Observando a Tabela 20 tem-se que o menor EQM obtido foi a partir do método de Prony

com abordagem do Filtro IIR. Ou seja, nessa solução os sinais de corrente foram reconstruídos com menor erro, com 0,66 e 0,13 para as bases COOLL e LIT Sintética, respectivamente. Mesmo não apresentando a melhor reconstrução, a maior taxa de classificação foi obtida pelo método *Matrix Pencil*. Para os casos analisados, as características obtidas pelo MP foram mais discriminativas para o classificador.

## 5.5 COMPARAÇÃO COM MÉTODOS DA LITERATURA

Ao analisar os resultados de classificação da seção 5.3, conclui-se que os melhores resultados são provenientes do método *Matrix Pencil* com ordem 10. Posto isso, os resultados de acurácia foram comparados com os seguintes métodos da literatura: Trajetória V-I (MULINARI *et al.*, 2019; LAZZARETTI *et al.*, 2020) e Transformada Discreta de Wavelet, do inglês *Discrete Wavelet Transform* (DWT) (CHANG *et al.*, 2013; LAZZARETTI *et al.*, 2020). As bases de dados utilizadas foram COOLL, base LIT Simulada (DB-5) e base LIT Sintética. Os classificadores aplicados foram LDA, kNN, SVM, DT e *Ensemble* (ENS). Para facilitar nas comparações, foi utilizado um algarismo significativo após a vírgula (exceto para os casos analisados por classe). Com isso, é possível obter uma visão geral das vantagens e limitações do Métodos de Prony aqui proposto.

A Tabela 21 apresenta os valores de classificação para cada base de dados em função do extrator de características e tipo de classificador. Os resultados da base de dados LIT Sintética foram separados em função das combinações de cargas, apenas cargas individuais (LIT-Sint-1), cargas duplas (LIT-Sint-2), cargas triplas (LIT-Sint-3), cargas óctuplas (LIT-Sint-8) e a junção de todas as combinações (LIT-Sint-T).

A partir dos dados Tabela 21 pode-se destacar que ao inserir novas cargas da base LIT Sintética, o desempenho de todos os métodos reduz e para os extratores utilizados, a melhor acurácia foi obtida com a combinação do método *Ensemble*. Ainda, o método de Prony a partir da abordagem de *Matrix Pencil* possui uma média geral de desempenho comparável aos outros métodos extratores da literatura.

Para avaliar o que acontece com mais detalhes, em termos de resultado de classificação, as Tabelas 22, 23 e 24 apresentam uma análise da acurácia por classe nas bases COOLL, LIT Sintética e LIT Simulada respectivamente.

Analisando os dados de acurácia por classe observa-se que em alguns casos que o método V-I e DWT não possuem uma boa acurácia, o método de Prony se destaca, como por

**Tabela 21 – Comparativo de Acurácia de Classificação.**

Extrator	Clas.	COOLL (%)	DB-5 (LIT -Sim) (%)	LIT-Sint -1 (%)	LIT-Sint -2 (%)	LIT-Sint -3 (%)	LIT-Sint -8 (%)	LIT-Sint -T (%)	Média (%)
V-I	LDA	98,3	87,8	99,3	98,9	89,3	85,7	72,7	90,3
	kNN	96,4	75,5	99,0	99,7	96,7	93,9	94,6	93,7
	SVM	96,3	91,1	98,1	98,4	97,5	92,5	93,3	95,3
	DT	98,4	92,3	98,6	99,4	95,6	91,0	95,8	95,9
	ENS	99,6	94,8	99,5	99,3	99,5	92,8	95,7	97,3
Prony MP	LDA	88,8	72,1	89,9	95,7	89,2	81,0	83,7	85,8
	kNN	84,6	53,9	88,2	94,6	86,3	83,3	82,8	82,0
	SVM	90,1	73,9	70,4	67,6	60,8	62,2	80,4	72,2
	DT	87,8	89,2	88,8	96,5	89,0	81,0	83,6	88,0
	ENS	98,4	93,7	96,1	98,5	95,2	90,8	94,4	95,3
DWT	LDA	91,5	86,3	82,8	70,0	61,1	60,6	60,4	73,2
	kNN	97,2	92,5	95,3	87,8	85,5	75,2	94,2	89,7
	SVM	93,1	90,4	93,2	88,7	83,6	77,8	95,1	88,8
	DT	96,2	93,5	92,2	90,1	90,4	90,2	90,3	91,8
	ENS	100,0	97,6	99,3	97,2	94,5	95,2	95,1	97,0

**Fonte: Autoria Própria.**

exemplo na Serra - Modelo 6 da base COOLL, Furadeira - Modelo 1 da base LIT Sintética e na Base LIT Simulada na Ponte a Diodo - Modelo 4. Essa análise demonstra a relevância da utilização do método de Prony e corrobora com a estratégia apresentada por Renaux *et al.* (2018), Lazzaretti *et al.* (2020) de extração de características por multi-agentes. Ou seja, ao utilizar diversos métodos para extração de características, um maior número de cargas pode ser identificado, mostrando a importância do desenvolvimento de métodos de extração com diferentes abordagens.

Ainda, para a base de dados LIT Sintética foi realizado o treinamento de modelos com uma e três cargas e avaliado (testado) o desempenho de classificação do modelo com duas, três e oito cargas para cada extrator e classificador de característica, conforme apresentado na Tabela 25. Com essa abordagem é possível observar a capacidade de generalização de um modelo treinado com nenhuma ou poucas cargas múltiplas, em um cenário onde mais cargas múltiplas podem ocorrer.

A Tabela 25 mostra que os resultados de classificação do método de Prony com o algoritmo *Ensemble* na maioria das bases estão próximos extrator baseado em V-I. A diferença é mais significativa é observada em relação ao método DWT. O valor mais baixo de classificação do método está relacionado a base LIT Sintética que, por conta da combinação com oito cargas simultâneas (LIT-Sint-1 - LIT-Sint-8), implica na atenuação da informação da corrente de *inrush*, reduzindo a acurácia do método. Observa-se que ao inserir mais cargas múltiplas no treinamento (LIT-Sint-3 - LIT-Sint-8), este erro diminui.

**Tabela 22 – Comparativo de Acurácia por Classe - Base COOLL.**

Classe	Prony MP (%)	V-I (%)	DWT (%)
Furadeira - Modelo 1	66	100	100
Furadeira - Modelo 2	100	100	100
Furadeira - Modelo 3	100	100	100
Furadeira - Modelo 4	100	100	100
Furadeira - Modelo 5	100	100	100
Furadeira - Modelo 6	100	100	100
Ventilador - Modelo 1	100	100	100
Ventilador - Modelo 2	100	100	100
Moedor - Modelo 1	100	100	100
Moedor - Modelo 2	100	100	100
Secador de Cabelo - Modelo 1	100	100	100
Secador de Cabelo - Modelo 2	100	100	100
Secador de Cabelo - Modelo 3	100	100	100
Secador de Cabelo - Modelo 4	100	100	100
Cortador de Grama - Modelo 1	100	100	100
Cortador de Grama - Modelo 2	100	100	100
Cortador de Grama - Modelo 3	100	100	100
Lâmpada - Modelo 1	100	100	100
Lâmpada - Modelo 2	100	100	100
Lâmpada - Modelo 3	100	100	100
Lâmpada - Modelo 4	100	100	100
Removedor de Tinta Plaina	100	100	100
Roteador	100	100	100
Lixadeira - Modelo 1	100	100	100
Lixadeira - Modelo 2	100	100	100
Lixadeira - Modelo 3	100	100	100
Serra - Modelo 1	100	100	100
Serra - Modelo 2	100	100	100
Serra - Modelo 3	100	100	100
Serra - Modelo 4	100	100	100
Serra - Modelo 5	100	100	100
Serra - Modelo 6	100	66	100
Serra - Modelo 7 - Modelo 1	100	100	100
Serra - Modelo 8	66	100	100
Aspirador de Pó - Modelo 1	100	100	100
Aspirador de Pó - Modelo 2	100	100	100
Aspirador de Pó - Modelo 3	100	100	100
Aspirador de Pó - Modelo 4	100	100	100
Aspirador de Pó - Modelo 5	100	100	100
Aspirador de Pó - Modelo 6	100	100	100
Aspirador de Pó - Modelo 7	100	100	100

**Fonte: Autoria Própria.**

Uma análise do tempo de execução de cada método de extração foi realizada a partir do software MATLAB. A Tabela 26 apresenta o tempo de processamento de cada método, em segundos, considerando dois algarismos significativos após a vírgula para facilitar o comparativo. Como pode ser observado na Tabela 26, o método de Prony a partir da solução por *Matrix Pencil* obteve o menor tempo de processamento para as bases COOLL e base LIT Sintética e na média, está próximo do valor do extrator V-I.

**Tabela 23 – Comparativo de Acurácia por Classe - Base LIT Sintética.**

Classe	Prony MP (%)	V-I (%)	DWT (%)
Microondas ( <i>standby</i> e on) - Modelo 1	95	97	92
Microondas ( <i>standby</i> e on) - Modelo 2	95	97	100
Secador de Cabelo (duas velocidades) - Modelo 1	100	100	100
Secador de Cabelo (duas velocidades) - Modelo 2	95	100	87
Secador de Cabelo (duas velocidades) - Modelo 3	91	91	95
Secador de Cabelo (duas velocidades) - Modelo 4	100	100	100
Secador (duas potências) - Modelo 1	76	55	33
Secador (duas potências) - Modelo 2	96	94	97
Lâmpada de LED - Modelo 1	88	82	91
Lâmpada de LED - Modelo 2	50	100	100
Lâmpada Incandescente	100	100	88
Monitor CRT	100	100	97
Monitor LED	100	93	100
Extrator de Fumaça	96	97	98
Carregador Celular - Modelo 1	100	100	100
Carregador Celular - Modelo 2	100	100	100
Carregador Computador - Modelo 1	97	100	95
Carregador Computador - Modelo 2	90	100	100
Furadeira (duas velocidades) - Modelo 1	91	76	85
Furadeira (duas velocidades) - Modelo 2	94	94	94
Resistor	100	100	100
Ventilador	85	100	95
Aquecedor a Óleo (duas potências) - Modelo 1	84	100	100
Aquecedor a Óleo (duas potências) - Modelo 2	100	100	100
Estação de Solda	84	100	100
Aquecedor com Resistência	100	100	100

**Fonte: Autoria Própria.**

Também foi verificado o tempo de execução dos algoritmos classificadores para cada tipo de extrator, conforme mostra a Tabela 27. Para facilitar as comparações foi utilizado um algoritmo significativo após a vírgula. A partir da análise da Tabela, vê-se que o tempo de classificação do método de Prony para o melhor classificador, *Ensemble*, obteve resultados mais elevados de tempo. Isso se deve ao fato do número de características inseridas no algoritmo é maior do que os outros métodos comparados.

Baseado nos resultados comparativos apresentados nesta seção, o método de Prony *Matrix Pencil* se mostra um método competitivo para extração de características quando combinado ao classificador *Ensemble*. Os valores obtidos de acurácia e tempo de processamento de extração foram próximos aos métodos do estado da arte. Obter diversos métodos extratores de características, com diferentes soluções e abordagens matemáticas, se torna interessante, pois, quando combinados em uma ação de Multiagentes, como apresentado por Renaux *et al.* (2018), Lazzaretti *et al.* (2020), é possível abranger um número maior de tipos e classes de cargas e aumentar a acurácia total do sistema.

**Tabela 24 – Comparativo de Acurácia por Classe - Base LIT Simulada.**

Classe	Prony MP (%)	V-I (%)	DWT (%)
Motor - Modelo 1	75	50	93
Motor - Modelo 2	37	56	56
Motor - Modelo 3	100	100	100
Motor - Modelo 4	35	76	94
Ponte a Diodo - Modelo 1	61	61	88
Ponte a Diodo - Modelo 2	83	77	88
Ponte a Diodo - Modelo 3	72	88	100
Ponte a Diodo - Modelo 4	100	93	93
Resistor - Modelo 1	100	100	100
Resistor - Modelo 2	94	100	100
Resistor - Modelo 3	100	100	100
Resistor - Modelo 4	76	100	100
Resistor + Indutor - Modelo 1	93	100	100
Resistor + Indutor - Modelo 2	88	100	100
Resistor + Indutor - Modelo 3	82	100	100
Resistor + Indutor - Modelo 4	100	100	100
Retificador a Tiristor com Resistor - Modelo 1	64	58	94
Retificador a Tiristor com Resistor - Modelo 2	100	100	100
Retificador a Tiristor com Resistor - Modelo 3	93	100	100
Retificador a Tiristor com Resistor - Modelo 4	93	93	93
Retificador a Tiristor com Resistor + indutor - Modelo 1	87	100	100
Retificador a Tiristor com Resistor + indutor - Modelo 2	100	100	100
Retificador a Tiristor com Resistor + indutor - Modelo 3	100	100	100
Retificador a Tiristor com Resistor + indutor - Modelo 4	93	93	93
Retificador a Diodo com Resistor - Modelo 1	95	95	100
Retificador a Diodo com Resistor - Modelo 2	93	86	93
Retificador a Diodo com Resistor - Modelo 3	88	82	100
Retificador a Diodo com Resistor - Modelo 4	46	92	100

Fonte: Autoria Própria.

**Tabela 25 – Acurácia para Testes de Generalização de Testes.**

Ext.	Base de Dados	LDA (%)	kNN (%)	SVM (%)	DT (%)	ENS (%)	Média (%)
V-I	LIT-Sint-1 to LIT-Sint-2	76,6	53,6	51,5	87,5	86,6	71,2
	LIT-Sint-1 to LIT-Sint-3	64,1	46,9	46,6	61,8	62,2	56,3
	LIT-Sint-1 to LIT-Sint-8	40,5	28,3	26,5	44,9	42,1	36,4
	LIT-Sint-3 to LIT-Sint-8	76,2	67,4	78,1	74,1	76,9	74,5
Prony	LIT-Sint-1 to LIT-Sint-2	71,1	65,3	62,7	50,2	83,4	67,1
	LIT-Sint-1 to LIT-Sint-3	51,8	53,2	44,8	35,5	59,4	48,9
	LIT-Sint-1 to LIT-Sint-8	31,4	40,2	34,8	25,5	45,3	35,4
	LIT-Sint-3 to LIT-Sint-8	51,3	55,3	63,6	50,2	64,2	56,9
DWT	LIT-Sint-1 to LIT-Sint-2	87,4	93,1	90,2	75,4	93,5	87,9
	LIT-Sint-1 to LIT-Sint-3	61,9	76,6	58,7	39,8	73,2	62,1
	LIT-Sint-1 to LIT-Sint-8	34,3	62,1	61,8	18,4	61,7	47,7
	LIT-Sint-3 to LIT-Sint-8	52,2	73,4	70,6	78,3	90,3	72,9

Fonte: Autoria Própria.

**Tabela 26 – Tempo de execução dos Extratores de Características.**

Método	COOLL (s)	SLIT5 (s)	LIT-Sint (s)	Média (s)
DWT	0,11	0,57	0,81	0,49
V-I	0,05	0,13	0,23	0,14
Prony MP	0,03	0,24	0,18	0,15

Fonte: Autoria Própria.

**Tabela 27 – Tempo de Execução da Classificação por Métodos Extratores versus Classificadores.**

Extrator	Classificador	COOLL (ms)	DB5 (ms)	LIT-Sint (ms)
V-I	ENS	19,0	268,2	12,5
	LDA	4,7	3,6	4,1
	kNN	4,9	4,9	5,4
	DT	2,2	2,4	2,7
	SVM	229,0	106,1	91,8
Prony MP	ENS	252,5	259,7	298,3
	LDA	4,4	4,6	3,8
	kNN	4,1	4,8	6,1
	DT	2,0	2,4	3,1
	SVM	227,8	108,1	90,4
DWT	ENS	9,8	119,7	35,5
	LDA	4,4	3,8	4,4
	kNN	3,9	4,0	5,1
	DT	2,3	2,3	2,1
	SVM	227,7	110,8	92,8

**Fonte: Autoria Própria.**

## 6 CONCLUSÃO

Diante das demandas de conservação e gerenciamento de energia, o desenvolvimento de técnicas para monitoramento de energia, como NILM, é de extrema relevância. Neste contexto, o trabalho proposto desenvolveu uma base de dados simulados baseado em sinais residenciais para o teste de técnicas de extração de características e classificação de cargas. Ainda, foi proposta a utilização do método matemático de Prony para extração de parâmetros dos sinais de corrente, para posterior classificação das cargas.

No desenvolvimento da base simulada, foram sanadas as principais limitações encontradas nos testes com bases disponíveis na literatura, como falta de cargas múltiplas, de anotação de eventos, dados balanceados e análise de harmônicas e ruídos da rede. Ainda, a partir da extração de características e classificação por métodos da literatura, foi possível mostrar que a inserção de ruídos e harmônicos na rede pode prejudicar a extração de características do método, e por conseguinte, a identificação da carga.

Na análise de extração de características, o método de Prony foi utilizado em sinais de corrente considerando o regime transitório e permanente do sinal. Considerando as melhores combinações obtidas neste trabalho, o Método de Prony com solução *Matrix Pencil* (MP) utilizando o classificador *Ensemble* obteve uma taxa de acurácia de 98% na base de dados COOLL, 93% na base LIT Simulada (DB-5) e 94% na base LIT Sintética, valores comparáveis aos principais métodos da literatura utilizados como referência. Em termos de custo computacional, o método de Prony MP também se mostra promissor, apresentando em diversos casos analisados, resultados superiores aos métodos do estado-da-arte comparados.

Os métodos de extração considerados como estados da arte foram apresentado na seção 2.2 e as bases de dados disponíveis na literatura foram avaliadas na seção 2.4, mostrando as principais limitações existentes. Nas seções 5.1 e 5.2 foram apresentados o desenvolvimento da base de dados e aplicação do método de Prony para extração de características. Na seção 5.5, os resultados obtidos foram comparados com os principais métodos da literatura, mostrando a viabilidade do método na aplicação NILM.

Todos os algoritmos utilizados para o desenvolvimento da base da dados simulada e para aplicação do método de Prony, com suas diferentes abordagens, como extrator de características, estão disponíveis no *Github* da autora<sup>1</sup>.

---

<sup>1</sup> <https://github.com/hellenancelmo>

## 6.1 TRABALHOS FUTUROS

Como continuação para o presente trabalho, propõe-se:

- Incluir na base de dados simulada cargas que apresentam o comportamento com diferentes estados, como um impressora em *stand by*, imprimindo, escaneando;
- Incluir na base de dados simulada motores com diferentes características transitórias, como motor de indução;
- Aplicar as abordagens do método de Prony para extração de características em cargas com diferentes estados;
- Embarcar em um sistema eletrônico o método de Prony;
- Avaliar técnicas de *deep learning* para extração de características e classificação de forma integrada;
- Explorar o desempenho por classe do método de Prony na arquitetura multiagente proposta em Lazzaretti *et al.* (2020).

## 6.2 TRABALHOS PUBLICADOS

Durante a realização deste trabalho foram publicados três artigos em congresso internacional e um artigo em congresso nacional, sendo todos diretamente relacionados ao presente trabalho:

**ANCELMO, H. C.; MULINARI, B. M.; POTTKER, F.; LAZZARETTI, A. E.; BAZZO, T. P. M.; OROSKI, E.; RENAUX, D. P. B.; LIMA, C. R. E.; LINHARES, R. P.; GAMBA, A. R. A.** A New Simulated Database for Classification Comparison in Power Signature Analysis. In: 20th International Conference on Intelligent Systems Applications to Power Systems (ISAP), 2019, Nova Deli.

**ANCELMO, H. C.; GRANDO, F. L.; MULINARI, B. M.; COSTA, C. H.; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; POTTKER, F.; LIMA, C. R. E.; LINHARES, R. P.** A Transient and Steady-State Power Signature Feature Extraction Using Different Prony's Methods

In: 20th International Conference on Intelligent Systems Applications to Power Systems (ISAP), 2019, Nova Deli.

**ANCELMO, H. C.;** GRANDO, F. L.; COSTA, C. H.; MULINARI, B. M.; OROSKI, E.; LAZZARETTI, A. E.; POTTKER, F.; RENAUX, D. P. B. Automatic Power Signature Analysis using Prony's Method and Machine Learning-Based Classifiers In: 2nd European Conference on Electrical Engineering & Computer Science.

**ANCELMO, H. C.;** MULINARI, B. M.; OROSKI, E.; LAZZARETTI, A. E.; LINHARES, R. P.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F. Método Não Intrusivo De Classificação de Cargas Baseadas em Power Signature In: Smart Energy, 2018, Curitiba.

Também foi feita a publicação de dois artigos para a revista *Energies* (fator de impacto 2.702), diretamente relacionados a este trabalho:

LAZZARETTI, A. E.; RENAUX, D. P. B.; LIMA, C. R. E.; MULINARI, B. M.; **ANCELMO, H. C.;** OROSKI, E.; POTTKER, F.; LINHARES, R. P.; NOLASCO, L. S.; LIMA, L. T.; OMORI, J. S.; SANTOS, R. R. A Multi-Agent NILM Architecture for Event Detection and Load Classification. *Energies MDPI*, 2020, 13(17).

RENAUX, D. P. B.; POTTKER, F.; **ANCELMO, H. C.;** LAZZARETTI, A. E.; LIMA, C. R. E.; LINHARES, R. P.; OROSKI, E.; NOLASCO, L. S.; LIMA, L. T.; MULINARI, B. M.; SILVA, J. R. L.; OMORI, J. S.; SANTOS, R. R. A Dataset for Non-Intrusive Load Monitoring: Design and Implementation. *Energies MDPI*, 2020, 13(20).

Ademais, também foram publicados quatro artigos, em congressos internacionais, que não incluem resultados do presente trabalho, mas que permeiam as mesmas áreas de conhecimento, sobretudo processamento de sinais e aprendizado de máquina aplicado à sistemas elétricos. Os artigos são os seguintes:

COSTA, C. H.; MORITZ, G. L.; LAZZARETTI, A. E.; **ANCELMO, H. C.;** MULINARI, B. M.; RODRIGUES, M. P.; OROSKI, E.; GOES, R. E. A Comparison of Machine Learning-Based

Methods for Fault Classification in Photovoltaic Systems In: Innovative Smart Grid Technologies (ISGT), 2019, Gramado.

MULINARI, B. M.; CAMPOS, D. P.; COSTA, C. H.; **ANCELMO, H. C.**; LAZZARETTI, A. E.; OROSKI, E.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F.; LINHARES, R. P. A New Set of Steady-State and Transient Features for Power Signature Analysis Based on V-I Trajectory In: Innovative Smart Grid Technologies (ISGT), 2019, Gramado.

LINHARES, R. R. ; LIMA, C. R. E. ; RENAUX, D. P. B. ; POTTKER, F. ; OROSKI, E. ; LAZZARETTI, ANDRÉ E. ; MULINARI, B. M. ; **ANCELMO, H. C.** One-millisecond low-cost synchronization of wireless sensor network In: IX Brazilian Symposium on Computing Systems Engineering, 2019, Salvador.

POTTKER, F. ; LAZZARETTI, ANDRÉ E. ; OROSKI, E. ; RENAUX, D. P. B. ; LINHARES, R. R. ; LIMA, C. R. E. ; **ANCELMO, H. C.** ; MULINARI, B. M. Non-Intrusive Load Monitoring: A Multi-Agent Architecture and Results In: 2nd European Conference on Electrical Engineering & Computer Science, 2018.

## REFERÊNCIAS

ANCELMO, H. C.; GRANDO, F. L.; COSTA, C. H.; MULINARI, B. M.; OROSKI, E.; LAZZARETTI, A. E.; POTTKER, F.; RENAUX, D. P. B. Automatic power signature analysis using prony method and machine learning-based classifiers. *In: 2nd European Conference on Electrical Engineering and Computer Science*. [S.l.: s.n.], 2018.

ANCELMO, H. C.; GRANDO, F. L.; MULINARI, B. M.; COSTA, C. H. da; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; POTTKER, F.; LIMA, C. R. E.; LINHARES, R. R. A transient and steady-state power signature feature extraction using different prony's methods. *In: 20th International Conference on Intelligent System Application to Power Systems (ISAP)*. [S.l.: s.n.], 2019. p. 1–6.

ANCELMO, H. C.; MULINARI, B. M.; POTTKER, F.; LAZZARETTI, A. E.; BAZZO, T. d. P. M.; OROSKI, E.; RENAUX, D. P. B.; LIMA, C. R. E.; LINHARES, R. R.; GAMBA, A. R. d. A. A new simulated database for classification comparison in power signature analysis. *In: 20th International Conference on Intelligent System Application to Power Systems (ISAP)*. [S.l.: s.n.], 2019. p. 1–7.

ANDERSON, K.; OCNEANU, A.n; BENITEZ, D.; CARLSON, D.; ROWE, A.; BERGES, M. BLUED: a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research. *In: Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*. Beijing, China: [s.n.], 2012.

ARFKEN, George B.; WEBER, Hans J.; HARRIS, Frank E. Chapter 11 - complex variable theory. *In: ARFKEN, G. B.; WEBER, H.J.; HARRIS, F. E. (Ed.). Mathematical Methods for Physicists (Seventh Edition)*. Seventh edition. Boston: Academic Press, 2013. p. 469 – 550. ISBN 978-0-12-384654-9.

AYDINALP-KOKSAL, M.; UGURSAL, V. I. Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. *Applied Energy*, v. 85, n. 4, p. 271 – 296, 2008. ISSN 0306-2619.

BACCA, I A.; MENDONÇA, M V. B.; TAVARES, C; GONDIM, I N.; OLIVEIRA, Jose. Atp-models language to represent domestic refrigerators performance with power quality disturbances. *Renewable Energy and Power Quality Journal*, v. 1, 04 2009.

BERGES, M.; GOLDMAN, E.; MATTHEWS, H. S.; SOIBELMAN, L.; ANDERSON, K. User-centered nonintrusive electricity load monitoring for residential buildings. *Journal of Computing in Civil Engineering*, v. 25, n. 6, p. 471–480, 2011.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738.

BONFIGLI, R.; SQUARTINI, S.; FAGIANI, M.; PIAZZA, F. Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview. *In: 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*. [S.l.: s.n.], 2015. p. 1175–1180.

BRADLE, M. Prony the bridge-builder the life and times of gaspard de prony, educator and scientist. **Centaurus**, v. 37, p. 230–268, 1994.

BRUNTON, S.; KUTZ, J. Chapter 1: Singular value decomposition. *In: \_\_\_\_*. [S.l.: s.n.], 2019. p. 3. ISBN 9781108380690.

CHANG, H.-H.; LIAN, K.-L.; SU, Y.-C.; LEE, W.-J. Energy spectrum-based wavelet transform for non-intrusive demand monitoring and load identification. *In: Proceedings of the Industry Application Society General Meeting*. [S.l.]: IEEE, 2013. p. 1–9.

CHANG, H.; LIAN, K.; SU, Y.; LEE, W. Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification. **IEEE Transactions on Industry Applications**, v. 50, n. 3, p. 2081–2089, May 2014.

CHANG, H.-H. Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses. **Energies**, v. 5, n. 11, p. 4569–4589, 2012.

CHERKASSKY, V. S.; MULIER, F. **Learning from Data: Concepts, Theory, and Methods**. 1st. ed. New York, NY, USA: John Wiley & Sons, Inc., 1998. ISBN 0471154938.

COLLIN, A. J.; TSAGARAKIS, G.; KIPRAKIS, A. E.; MCLAUGHLIN, S. Development of low-voltage load models for the residential load sector. **IEEE Transactions on Power Systems**, v. 29, n. 5, p. 2180–2188, 2014.

Efficiency Valuation Organization. **International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings**. 2012. Disponível em: <https://evo-world.org>.

EGARTER, D.; ELMENREICH, W. Evonilm -evolutionary appliance detection for miscellaneous household appliances. *In: .* [S.l.: s.n.], 2013.

ESA, N. F.; ABDULLAN, M. P.; HASSAN, M. Y. A review method in nonintrusive appliance load monitoring. **Renewable and Sustainable Energy Reviews**, v. 66, n. 3, p. 163–173, 2016.

- FIGUEIREDO, M.; ALMEIDA, A.; RIBEIRO, B. Home electrical signal disaggregation for non-intrusive load monitoring systems. **Neurocomputing**, v. 96, p. 66–73, nov. 2012. ISSN 0925-2312.
- GAO, J.; GIRI, S.; KARA, E.; BERGÉS, M. Plaid: A public dataset of high-resolution electrical appliance measurements for load identification research. p. 198–199, 11 2014.
- GHOSH, S.; CHATTERJEE, A.; CHATTERJEE, D. A smart iot based non-intrusive appliances identification technique in a residential system. *In: 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*. [S.l.: s.n.], 2020. p. 1–6.
- GILLIS, J. M.; ALSHAREEF, S.; MORSI, W. G. Nonintrusive load monitoring using wavelet design and machine learning. **IEEE Transactions on Smart Grid**, v. 7, p. 320–328, 2016.
- GOLUB, G. H.; LOAN, C. F. van. An analysis of the total least squares problem. **SIAM Journal on Numerical Analysis**, v. 17, n. 6, p. 883–893, 1980. Disponível em: <https://doi.org/10.1137/0717073>.
- GOPINATH, R.; KUMAR, M.; Prakash Chandra Joshua, C.; SRINIVAS, K. Energy management using non-intrusive load monitoring techniques – state-of-the-art and future research directions. **Sustainable Cities and Society**, v. 62, p. 102411, 2020. ISSN 2210-6707.
- HART, G. W. Nonintrusive appliance load monitoring. **Proceedings of the IEEE**, v. 80, n. 12, p. 1870–1891, 1992.
- HASSAN, T.; JAVED, F.; ARSHAD, N. An empirical investigation of V-I trajectory based load signatures for non-intrusive load monitoring. **IEEE Transactions on Smart Grid**, v. 5, n. 2, p. 870–878, 2014.
- HUA, Y.; SARKAR, T. Matrix pencil method and its performance. *In: ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*. Los Alamitos, CA, USA: IEEE Computer Society, 1988. p. 2476,2477,2478,2479. ISSN 1520-6149.
- HUA, Y.; SARKAR, T. K. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, v. 38, n. 5, p. 814–824, 1990.
- IEA. **World Energy Outlook 2019**. [S.l.], 2019. <https://www.oecd-ilibrary.org/content/publication/caf32f3b-en>.

IKSAN, N.; SEMBIRING, J.; HARYANTO, N.; SUPANGKAT, S. H. Appliances identification method of non-intrusive load monitoring based on load signature of V-I trajectory. **International Conference on Information Technology Systems and Innovation**, 2016.

JIANG, L.; LUO, S. H.; LI, J. M. Intelligent electrical appliance event recognition using multi-load decomposition. *In: Energy and Power Technology*. [S.l.]: Trans Tech Publications Ltd, 2013. (Advanced Materials Research, v. 805), p. 1039–1045.

KAHL, M.; HAQ, A. U.; KRIECHBAUMER, T.; JACOBSEN, H. Whited-a worldwide household and industry transient energy data set. *In: . [S.l.: s.n.]*, 2016.

KAHL, M.; HAQ, A. UI; KRIECHBAUMER, T.; JACOBSEN, H. A comprehensive feature study for appliance recognition on high frequency energy data. *In: Proceedings of the 8th International Conference on Future Energy Systems*. [S.l.: s.n.], 2017. p. 121–131.

KELLY, J.; KNOTTENBELT, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. v. 2, n. 150007, 2015.

KIM, H.; MARWAH, M.; ARLITT, M.; G., Lyon; HAN, J.; ANDERSON, K. Unsupervised disaggregation of low frequency power measurements. *In: . [S.l.: s.n.]*, 2011. v. 11, p. 747–758.

KLEMENJAK, C.; KOVATSCH, C.; HEROLD, M.; ELMENREICH, W. A synthetic energy dataset for non-intrusive load monitoring in households. **Scientific Data**, v. 7, 12 2020.

KOLTER, J.; JOHNSON, M. Redd: A public data set for energy disaggregation research. **Artif. Intell.**, v. 25, 01 2011.

KRIECHBAUMER, T.; JACOBSEN, H. Blond, a building-level office environment dataset of typical electrical appliances. **Scientific Data**, v. 5, 3 2018.

LAM, HY; TING, KH; LEE, W. K.; FUNG, GSK. An analytical understanding on voltage-current curve of electrical load. **Conference on Electrical Engineering**, n. June, 2006.

LAM, H Y; FUNG, G S K; LEE, W K. A Novel Method to Construct Taxonomy of Electrical Appliances Based on Load Signatures. 2007.

LAZZARETTI, A.E.; RENAUX, D.P.B.; LIMA, C.R.E.; MULINARI, B.M.; ANCELMO, H.C.; OROSKI, E.; PÖTTKER, F.; LINHARES, R.R.; NOLASCO, L.S.; LIMA, L.T.; OMORI, J.S.; SANTOS, R.B. A multi-agent nilm architecture for event detection and load classification. **Energies**, v. 13, n. 17, p. 1–37, 2020.

LEE, K. D.; LEEB, S. B.; NORFORD, L. K.; ARMSTRONG, P. R.; HOLLOWAY, J.; SHAW, S. R. Estimation of variable-speed-drive power consumption from harmonic content. **IEEE Transactions on Energy Conversion**, v. 20, n. 3, p. 566–574, Sep. 2005.

LIAO, J.; ELAFOUDI, G.; STANKOVIC, L.; STANKOVIC, V. Power disaggregation for low sampling rate data. **2nd international non-intrusive appliance load monitoring workshop.**, 2014.

LIU, H. Non-intrusive load monitoring, theory, technologies and applications. 01 2020.

LIU, H.; WU, H.; YU, C. A hybrid model for appliance classification based on time series features. **Energy and Buildings**, v. 196, p. 112 – 123, 2019. ISSN 0378-7788.

LOPEZ, J. M. Gonzalez; POURESMAEIL, E. Smart Residential Load Simulator for Energy Management in Smart Grids. **IEEE Transactions on Industrial Electronics**, v. 66, n. 2, p. 1443–1452, 2019.

MAKONIN, S.; WANG, Z.; TUMPACH, C. Rae: The rainforest automation energy dataset for smart grid meter data analysis. **Data**, MDPI AG, v. 3, n. 1, p. 8, Feb 2018. ISSN 2306-5729.

MARPLE, S. L. **Digital Spectral Analysis: With Applications**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1986.

MARTINS, J. F.; LOPES, R.; LIMA, C.; ROMERO-CADAVAL, E.; VINNIKOV, D. A novel nonintrusive load monitoring system based on the s-transform. *In: 13th International Conference on Optimization of Electrical and Electronic Equipment*. [S.l.: s.n.], 2012. p. 973–978.

MATHIS, M.; RUMSCH, A.; KISTLER, R.; ANDRUSHEVICH, A.; KLAPPROTH, A. Improving the recognition performance of nialm algorithms through technical labeling. *In: 2014 12th IEEE International Conference on Embedded and Ubiquitous Computing*. [S.l.: s.n.], 2014. p. 227–233.

MEZIANE, M. N.; HACINE-GHARBI, A.; RAVIER, P.; LAMARQUE, G.; BUNETEL, J. Le; RAINGEAUD, Y. Electrical appliances identification and clustering using novel turn-on transient features. *In: Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods*. [S.l.: s.n.], 2017. p. 647–654.

MEZIANE, M. N.; RAVIER, P.; LAMARQUE, G.; ABED-MERAIM, K.; BUNETEL, J. Le; RAINGEAUD, Y. Modeling and estimation of transient current signals. *In: 2015 23rd European Signal Processing Conference (EUSIPCO)*. [S.l.: s.n.], 2015. p. 1960–1964.

MOHAMMAD, H.; WAZIRI, M. Y.; SANTOS, S. A. A brief survey of methods for solving nonlinear least-squares problems. **Numerical Algebra, Control and Optimization**, v. 9, n. 1, p. 1–10, 2019.

MULINARI, B. Machado; LINHARES, R.; CAMPOS, D.; COSTA, C.; ANCELMO, H.; LAZZARETTI, A.; OROSKI, E.r; LIMA, C. Erig; RENAUX, D.; POTTKER, F. A new set of steady-state and transient features for power signature analysis based on v-i trajectory. *In: IEEE PES Innovative Smart Grid Technologies Conference - Latin America. [S.l.: s.n.]*, 2019. p. 1–6.

NAIT-MEZIANE, M.; RAVIER, P.; ABED-MERAIM, K.; LAMARQUE, G.; BUNETEL, J.s; RAINGEAUD, Y. Electrical transient modeling for appliance characterization. **EURASIP Journal on Advances in Signal Processing**, v. 55, p. 1–19, 12 2019.

ONG, C.-M. **Dynamic Simulation of Eletric Machinery Edition: Prentice Hall Ptr.** Upper Saddle River, New Jersey, USA: Prentice-Hall, Inc., 1998.

PICON, T.; MEZIANE, M. Nait; RAVIER, P.; LAMARQUE, G.; NOVELLO, C.; BUNETEL, J. Le; RAINGEAUD, Y. COOLL: Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. **arXiv:1611.05803**, 2016.

PISSANETZKY, S. Chapter 6 - sparse eigenanalysis. *In: PISSANETZKY, Sergio (Ed.). Sparse Matrix Technology. [S.l.]: Academic Press, 1984. p. 196 – 232. ISBN 978-0-12-557580-5.*

RAIKER, G. A.; UMANAND, L.; AGRAWAL, S.; THAKUR, A. S.; K, A.; BARTON, J. P.; THOMSON, M. Internet of things based demand side energy management system using non-intrusive load monitoring. *In: 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020). [S.l.: s.n.]*, 2020. p. 1–5.

RENAUX, D. P. B.; L., R. R.; POTTKER, F.; LAZZARETTI, A. E.; LIMA, C. E. E.de; COELHO-NETO, A.; CAMPANER, M. H. Designing a novel dataset for non-intrusive load monitoring. *In: VIII Brazilian Symposium on Computing Systems Engineering. [S.l.: s.n.]*, 2018.

RENAUX, D. P. B.; POTTKER, F.; ANCELMO, H. C.; LAZZARETTI, C. R. E. Lima A. E.; LINHARES, R. R.; OROSKI, E.r; NOLASCO, L. da S.; LIMA, L. T.; MULINARI, B. M.; SILVA, J. R. L. da; OMORI, J. S.; SANTOS, R. B. dos. A dataset for non-intrusive load monitoring: Design and implementation. **Energies**, v. 13, n. 20, p. 1–35, 2020.

RENAUX, D. P. B.; POTTKER, F.; ANCELMO, H. C.; LAZZARETTI, C. R. E. Lima A. E.; LINHARES, R. R.; OROSKI, E.; NOLASCO, L. da S.; LIMA, L. T.; MULINARI, B. M.; SILVA, J. R. L. da; OMORI, J. S.; SANTOS, R. B. dos. A dataset for non-intrusive load monitoring: Design and implementation. **Energies**, v. 13, n. 20, p. 1–35, 2020.

RODRÍGUEZ, A. F.; SANTIAGO, L.; GUILLÉN, E. L.; ASCARIZ, J. M. R.; MIGUEL-JIMÉNEZ, J. M.; BOQUETE, L. Coding prony's method in matlab and applying it to biomedical signal filtering. **BMC Bioinformatics**, v. 19, 12 2018.

RUANO, A.; HERNANDEZ, A.; UREÑA, J.; RUANO, M.; GARCÍA, J. Nilm techniques for intelligent home energy management and ambient assisted living: A review. **Energies**, v. 12, p. 2203, 06 2019.

SNOEK, J.; LAROCHELLE H.AND ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. *In: Proceedings of the 25th International Conference on Neural Information Processing Systems. [S.l.: s.n.]*, 2012. (NIPS'12), p. 2951–2959.

SRINIVASAN, D.; NG, W. S.; LIEW, A. C. Neural-network-based signature recognition for harmonic source identification. **IEEE Transactions on Power Delivery**, v. 21, n. 1, p. 398–405, Jan 2006.

STORN, R. Differential evolution design of an iir-filter. *In: Proceedings of IEEE International Conference on Evolutionary Computation. [S.l.: s.n.]*, 1996. p. 268–273.

SUN, M.; NAKOTY, F. M.; LIU, Q.; LIU, X.; YANG, Y.; SHEN, T. Non-intrusive load monitoring system framework and load disaggregation algorithms: A survey. *In: 2019 International Conference on Advanced Mechatronic Systems (ICAMechS). [S.l.: s.n.]*, 2019. p. 284–288.

TABATABAEI, S. M.; DICK, S.; XU, W. Toward non-intrusive load monitoring via multi-label classification. **IEEE Transactions on Smart Grid**, v. 8, n. 1, p. 26–40, 2017.

TAX, D.; DUIN, R. Uniform object generation for optimizing one-class classifiers. **Journal of Machine Learning Research**, v. 2, 04 2002.

UKIL, Abhisek; ŽIVANOVIĆ, Rastko. Adjusted haar wavelet for application in the power systems disturbance analysis. **Digital Signal Processing**, v. 18, n. 2, p. 103 – 115, 2008.

VAPNIK, V. N. **Statistical Learning Theory. [S.l.]**: Wiley-Interscience, 1998.

VENKATESH, C; KUMAR, Ds. Modelling of Nonlinear Loads and Estimation of Harmonics in Industrial Distribution System. *In: . [S.l.: s.n.]*, 2008. v. 2, n. December, p. 592–597.

WANG, A. L.; CHEN, B. X.; WANG, C. G.; HUA, D. D. Non-intrusive load monitoring algorithm based on features of V–I trajectory. **Electric Power Systems Research**, Elsevier B.V., v. 157, p. 134–144, 2018.

ZEIFMAN, M. Disaggregation of home energy display data using probabilistic approach. **IEEE Transactions on Consumer Electronics**, v. 58, n. 1, p. 23–31, 2012.

ZEIFMAN, M.; AKERS, C.; ROTH, K. Nonintrusive Appliance Load Monitoring (NIALM) for Energy Control in Residential Buildings: Review and Outlook. **IEEE Transactions on Consumer Electronics**, 2011.

ZHANG, G.; WANG, G.; FARHANGI, H.; PALIZBAN, A. Residential electric load disaggregation for low-frequency utility applications. *In: 2015 IEEE Power Energy Society General Meeting. [S.l.: s.n.], 2015. p. 1–5.*

ZHUANG, Z.; HAINAN, C.; LUO, X. A supervised event-based non-intrusive load monitoring for non-linear appliances. **Sustainability**, v. 10, p. 1001, 03 2018.

ZOHA, A.; GLUHAK, A.; IMRAN, M.; RAJASEGARAR, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. **Sensors (Basel, Switzerland)**, v. 12, p. 16838–16866, 12 2012.