



MONITORAÇÃO NÃO-INVASIVA DE CARGAS ELÉTRICAS RESIDENCIAIS

Werner Spolidoro Freund

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: José Manoel de Seixas

Rio de Janeiro
Dezembro de 2013

MONITORAÇÃO NÃO-INVASIVA DE CARGAS ELÉTRICAS RESIDENCIAIS

Werner Spolidoro Freund

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. José Manoel de Seixas, D.Sc.

Prof. Luiz Pereira Calôba, Dr.Ing.

Prof. Carlos Augusto Duque, D.Sc.

Dr. Charles Bezerra do Prado, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
DEZEMBRO DE 2013

Freund, Werner Spolidoro

Monitoração Não-Invasiva de Cargas Elétricas Residenciais/Werner Spolidoro Freund. – Rio de Janeiro: UFRJ/COPPE, 2013.

XXI, 212 p.: il.; 29,7cm.

Orientador: José Manoel de Seixas

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2013.

Referências Bibliográficas: p. 200 – 212.

1. Monitoração Não-Invasiva de Cargas Elétricas. 2. NILM. 3. Algoritmo Genético. I. Seixas, José Manoel de. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*À busca da compreensão,
convivência e/ou superação de
limites.*

Agradecimentos

Agradeço aos meus pais por tornar isso tudo possível. Vocês que tiveram o maior peso para que isso se transformasse em realidade. Agradeço de coração por todo esforço e trabalho que tiveram com o intuito de me verem chegar aqui. Aos meus avós por todo o carinho e apoio que sempre me deram, sempre me incentivando para conseguir atingir meus sonhos. Ao meu irmão mais novo que sacrificou seu tempo livre fazendo algumas de minhas tarefas enquanto eu estava trabalhando. Ao meu outro irmão pelas risadas.

Ao meu orientador, Seixas, que me deu suporte e me guiou neste trabalho, apesar de todas condições atípicas. Seus conselhos foram muito valiosos, sua ajuda foi muito mais do que essencial. Aos outros orientadores que tive, Torres e Damazio, suas orientações ajudaram a tornar-me o que sou hoje.

Aos engenheiros do CATE no Centro de Pesquisas de Energia Elétrica, João, Aroldo, Guilherme e Victor por toda a ajuda e compreensão durante o tempo de trabalho. A convivência com vocês foi muito agradável! Também à Beth por sempre ter sido prestativa. Ao Alvaro pela informação necessária para a construção dos gabaritos dos conjuntos de dados *Empilhado4* e *Empilhado7*.

Ao pessoal do LPS: Balabram, Junior, Moura, Graef, Rodrigo, João Victor, Diego e Hellen. Vocês fazem o laboratório ser o que ele é, sem vocês ele não tem graça alguma. À Ana e Talia, vocês fazem muita falta lá. Agradeço em especial ao João Victor pelo suporte providencial no momento em que precisei, espero ter retornado da maneira possível com o máximo de conhecimento em troca. Desculpas ao Graef por todo incomodo causado, muito obrigado por fazer o melhor que podia. Ao Pedro que ajudou com parte do material utilizado para o levantamento bibliográfico das técnicas empregadas no NILM. Aos meus orientados, Diego e Hellen, me desculpem se exagerei na cobrança: corram atrás de seus futuros que o *Walhalla* espera por vocês!

À toda ajuda, paciência e instruções oferecidas pela Dani do PEE/COPPE. Aos professores da COPPE/UFRJ por todo seu empenho, ética e dedicação.

Finalmente, agradeço a todos meus amigos que escutaram diversas vezes: “não posso, tenho que terminar o mestrado”, mas nunca deixaram de compreender o momento em que estava passando. Hoje chegou o dia!

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MONITORAÇÃO NÃO-INVASIVA DE CARGAS ELÉTRICAS RESIDENCIAIS

Werner Spolidoro Freund

Dezembro/2013

Orientador: José Manoel de Seixas

Programa: Engenharia Elétrica

Este trabalho foi realizado em colaboração com o CEPEL/Eletrobras no desenvolvimento da tecnologia conhecida como Monitoração Não-Invasiva de Cargas Elétricas (NILM). A monitoração não-invasiva pode ser aplicada para garantir a qualidade de energia, o diagnóstico de carga, identificação de aparelhos defeituosos ou com consumo excessivo de energia e eficiência energética. Este trabalho abordará a etapa de detecção de eventos de transitórios causados por aparelhos elétricos, o qual é necessária para a obtenção dos traços de informação deixados por equipamentos durante a alteração de seu estado de operação. A proposta do CEPEL/Eletrobras é a utilização de um filtro de derivada de Gaussiana que gera candidatos a eventos de transitórios, os mesmos ainda analisados por outros testes para eliminação de falsos eventos em seguida.

A contribuição do trabalho está na sistematização da determinação dos parâmetros necessários para o método proposto pelo CEPEL/Eletrobras, que permite o ajuste automático dos mesmos para diferentes cenários. Para isso, foi implementada uma adaptação de um algoritmo genético, que permite dinâmica para alocação de maior esforço computacional para configurações que melhor se adequam ao problema. Além disso, o otimizador faz parte de um extenso ambiente de análise que trouxe vantagens operacionais em termo de infraestrutura para o projeto.

Ao aplicar o algoritmo na base de dados, o mesmo mostrou capacidade de generalização para as condições avaliadas. Os resultados foram comparados com os trabalhos recentes desenvolvidos neste contexto. Obteve-se taxa de detecção superior a 80% para taxa de falso alarme inferior a 1 % em cenários com operação simultânea de diversos aparelhos e dinâmica de carga. Já para os conjuntos com operações simples de equipamentos as taxas são de 98,1 % e 0,6 %, respectivamente.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

NON-INTRUSIVE RESIDENTIAL LOAD MONITORING

Werner Spolidoro Freund

December/2013

Advisor: José Manoel de Seixas

Department: Electrical Engineering

This work was done in collaboration with CEPEL/Eletrobras at the development of the technology known as Non-Intrusive Appliance Load Monitoring (NILM). Non-intrusive monitoring may be applied to power quality, load diagnosis, faulty appliance detection or with excessive energy consumption and energy efficiency. This work will address the step of detecting events transients caused by appliance loads, which is required for obtaining the traces of information left by the appliance while changing its consumption status. The technique proposed by CEPEL/Eletrobras is a Gaussian derivative filter which generates the transient events candidates, also analyzed by other checks that further eliminate false alarm events.

The contribution of this work is on the systematic determination of the parameters required for the proposed CEPEL/Eletrobras's method, which allows their automatic adjustment for several scenarios. To this end, it was implemented an adaptation of a genetic algorithm, that has as its feature the dynamic allocation of more computational effort for those configurations which are best suited for solving the problem. Furthermore, the optimizer is part of an extensive analysis environment which led to operational advantages in terms of infrastructure for the project.

When applying the algorithm to distinct datasets, it showed generalization properties for the conditions evaluated. The results were compared with related recent work. Detection rate above 80 % meanwhile false alarm rate lesser than 1 % were achieved for practical scenario applications with several simultaneous appliances operation and load dynamics. Yet for datasets with naive appliance operation conditions, those rates were respectively of 98,1 % and 0,6 %.

Sumário

Sumário	ix
Lista de Figuras	xii
Lista de Tabelas	xv
Lista de Símbolos	xvi
Lista de Abreviaturas	xix
1 Introdução	1
1.1 Motivação	2
1.2 Objetivo	4
1.3 Contribuições do Trabalho	5
1.4 Estrutura Capítular	6
2 Eficiência Energética	8
2.1 Da Origem	9
2.2 Eficiência Energética no Brasil	11
2.3 Dificuldades para o Estudo em Eficiência Energética	15
3 A Eficiência Energética no Setor Residencial	18
3.1 Eletricidade e Potencial de Eficiência Energética no Setor Residencial	19
3.2 Expandindo o Potencial de Eficiência Energética através do Retorno de Informação de Consumo	22
3.2.1 Tipos de Retorno	24
3.2.2 Resultados por Tipo de Retorno	25
3.2.3 Indo Além dos Resultados	26
3.2.4 Tecnologias e Tendências	29
3.2.5 Aspectos Psicológicos	33
3.2.6 Aspectos Visuais da Informação	39

4	Monitoração Não-Invasiva de Cargas Elétricas	41
4.1	Aspectos Gerais	42
4.2	As Diversas Metodologias Utilizadas no Mundo	44
4.2.1	Modelos de Carga	45
4.2.2	Escolha da Metodologia	48
4.2.3	Cálculo da Eficiência	54
4.2.4	Técnicas Aplicadas por demais autores	59
4.2.5	Discussão	85
4.3	A Metodologia no CEPEL	90
4.3.1	Estudos Anteriores	90
4.3.2	Metodologia de Partida do Trabalho	92
4.3.3	Quanto à Escolha das Características Discriminantes	100
5	Ambiente de Análise	105
5.1	Da Necessidade	106
5.2	Leitura, Representação e Interação com os Dados	109
5.2.1	Dados do Medidor	110
5.2.2	Evento de Transitório	112
5.2.3	Equipamentos	116
5.3	Interação Gráfica com o Usuário	117
5.4	Análise dos Dados	120
5.5	Otimização dos Parâmetros	123
5.5.1	Algoritmo Genético de Estratégia Evolutiva	126
6	Metodologia	138
6.1	Descrição da base de dados	139
6.1.1	Conjunto de dados <i>NI00168</i>	140
6.1.2	Conjunto de dados <i>NI00171</i>	144
6.1.3	Conjunto de dados <i>NI00173</i> , <i>NI00174</i> , <i>NI00175</i> e <i>NI00177</i>	149
6.1.4	Conjunto de dados <i>Temporizado</i>	158
6.1.5	Conjunto de dados <i>Empilhado4</i>	169
6.1.6	Conjunto de dados <i>Empilhado7</i>	173
6.2	Aplicação do ES para Otimização do Detector de Eventos	177
7	Resultados	181
7.1	Otimização Automática dos Parâmetros	182
8	Conclusão	195
8.1	Trabalhos Futuros	198

Lista de Figuras

2.1	A matriz energética mundial em 1973 e 2010.	11
2.2	O efeito <i>leapfrogging</i>	12
2.3	A etiqueta do PBE.	13
2.4	Necessidades para o planejamento, implementação e monitoração dos programas de EE pelas entidades governamentais.	16
3.1	Matriz energética brasileira.	19
3.2	Consumo de eletricidade por setor em 2011.	20
3.3	Curva de carga média para a região sudeste, ano base 2005.	21
3.4	O potencial de consumo para cada tipo de retorno	27
4.1	Esboço dos passos envolvidos no projeto do NILM e da sua operação.	44
4.2	equipamentos com dinâmica de carga (C5): ar condicionado	47
4.3	Exemplo de equipamentos com características similares (C6)	49
4.4	Esboço da metodologia empregada pelo CEPTEL	93
4.5	Gráficos descrevendo a operação da metodologia proposta pelo CEPTEL/Eletrobras.	96
5.1	Esboço do ambiente de análise implementado.	109
5.2	Informação gráfica para o interação com dados do medidor	113
5.3	Alguns exemplos de dificuldades encontrados nos dados reais	114
5.4	Informação gráfica para o Módulo de Interação Gráfica com os Dados: Evento de Transitório com Sobreposição.	118
5.5	Informação gráfica para o Módulo de Interação Gráfica com os Dados: Evento de Transitório Corrigido.	119
5.6	Informação gráfica para o Módulo de Interação Gráfica com os Dados: Disposição da informação.	121
5.7	Exemplo de informação gráfica para o Módulo de Análise dos Dados.	124
5.8	Esboço de um algoritmo evolutivo genérico.	127
5.9	Função de <i>Ackley</i> em duas dimensões para $-30, 0 < x_i < 30, 0$	130
5.10	Evolução para o melhor indivíduo para a validação da versão original do ES	131

5.11	Competição interespécie.	133
5.12	Competição intraespécies sem intervenção na competição.	134
5.13	Pressão de seleção para competição intraespécie sem intervenção na competição.	135
5.14	Competição intraespécies com intervenção na competição.	136
5.15	Pressão de seleção para competição intraespécie com intervenção na competição.	137
6.1	Perfil de consumo agregado para o conjunto de dados <i>NI00168</i>	141
6.2	Informação no gabarito para o conjunto de dados <i>NI00168</i> - consumo temporal dos equipamentos.	142
6.3	Informação no gabarito para o conjunto de dados <i>NI00168</i> - gráfico circular do consumo dos equipamentos.	143
6.4	Perfil de consumo agregado para o conjunto de dados <i>NI00171</i>	145
6.5	Informação no gabarito para o conjunto de dados <i>NI00171</i> - consumo temporal dos equipamentos.	146
6.6	Informação no gabarito para o conjunto de dados <i>NI00171</i> - gráfico circular do consumo dos equipamentos.	147
6.7	Informação no gabarito para o conjunto de dados <i>NI00171</i> - envoltória para as diversas variáveis para o computador portátil.	148
6.8	Perfil de consumo agregado para o conjunto de dados <i>NI00173</i>	150
6.9	Informação no gabarito para o conjunto de dados <i>NI00173</i> - consumo temporal dos equipamentos.	151
6.10	Perfil de consumo agregado para o conjunto de dados <i>NI00174</i>	152
6.11	Informação no gabarito para o conjunto de dados <i>NI00174</i> - consumo temporal dos equipamentos.	153
6.12	Perfil de consumo agregado para o conjunto de dados <i>NI00175</i>	154
6.13	Informação no gabarito para o conjunto de dados <i>NI00175</i> - consumo temporal dos equipamentos.	155
6.14	Perfil de consumo agregado para o conjunto de dados <i>NI00177</i>	156
6.15	Informação no gabarito para o conjunto de dados <i>NI00177</i> - consumo temporal dos equipamentos.	157
6.16	Perfil de consumo agregado para o conjunto de dados <i>Temporizado</i>	160
6.17	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - con- sumo temporal dos equipamentos.	161
6.18	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - gráfico circular do consumo dos equipamentos.	162
6.19	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - en- voltória para as diversas variáveis para a geladeira.	163

6.20	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - envoltória para as diversas variáveis para a ventilador.	164
6.21	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - envoltória para as diversas variáveis para a lâmpada fluorescente 23W. . .	165
6.22	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - envoltória para as diversas variáveis para a lâmpada fluorescente 54W. . .	166
6.23	Informação no gabarito para o conjunto de dados <i>Temporizado</i> - envoltória para as diversas variáveis para a televisão.	167
6.24	Distúrbio recorrente presente no conjunto de dados <i>Temporizado</i> . . .	168
6.25	Perfil de consumo agregado para o conjunto de dados <i>Empilhado4</i> . . .	170
6.26	Informação no gabarito para o conjunto de dados <i>Empilhado4</i> : consumo temporal dos equipamentos.	171
6.27	Informação no gabarito para o conjunto de dados <i>Empilhado4</i> : gráfico circular do consumo dos equipamentos.	172
6.28	Perfil de consumo agregado para o conjunto de dados <i>Empilhado7</i> . . .	174
6.29	Informação no gabarito para o conjunto de dados <i>Empilhado7</i> : consumo temporal dos equipamentos.	175
6.30	Informação no gabarito para o conjunto de dados <i>Empilhado7</i> : gráfico circular do consumo dos equipamentos.	176
7.1	Falsos alarmes para a configuração <i>Ajuste CEPEL</i> no conjunto de dados <i>NI00171</i>	183
7.2	Convergência das espécies para o ajuste automático da configuração ES 1.	186
7.3	Convergência das espécies para o ajuste automático da configuração ES 2.	186
7.4	Convergência das espécies para o ajuste automático da configuração ES 3.	187
7.5	Falso alarme para a configuração <i>ES 1</i> e <i>ES 2</i> no conjunto de dados <i>NI00173</i>	188
7.6	Exemplos de falsos alarmes no conjunto de dados <i>Temporizado Gabarito 1</i>	190
7.7	Exemplos de perdas de alvos do distúrbio no conjunto de dados <i>Temporizado Gabarito 2</i>	191
7.8	Exemplos de perdas de alvo no conjunto de dados <i>Empilhado4</i> para a configuração <i>ES 2</i>	192

Lista de Tabelas

3.1	Empresas utilizando informação da concessionária e as oportunidades e incentivo de economia de energia oferecidas.	31
3.2	Especificações de <i>displays</i> domiciliares disponíveis.	32
3.3	Valores propostos por Rokeach e Maslow.	34
4.1	Taxonomias utilizadas por autores de estudos anteriores no CEPEL .	102
4.2	Taxonomias determinadas por LAM et al.: curvas I-V, decomposição SVD e variáveis padrão	103
7.1	Taxa de detecção de eventos de transitório e falso alarme para os três ajustes automáticos e os dois ajustes manuais.	184
7.2	Taxa de detecção de eventos de transitório e falso alarme para os três ajustes automáticos, os dois ajustes manuais e o Detector de Patamar Elaborado.	193

Lista de Símbolos

BMU	<i>Best Machine Unit.</i>
CR	razão de complementação.
E_{id}	energia corretamente identificada para o i -ésimo equipamento.
F^i	medida-F para o i -ésimo equipamento.
F_{en}^i	medida-eF para o i -ésimo equipamento.
IR	resíduo individual.
J_{max}	janela máxima de amostras para correlacionar a informação de análise com a do gabarito.
N_{ap}	número de equipamentos.
N_{ap}	número de eventos de transitórios.
N_{fp}	número de eventos devido a falsos positivos.
N_{id}	número de eventos corretamente detectados e classificados.
N_{ni}	número de eventos não identificados.
N_{reais}	número de eventos causados pelos equipamentos na rede.
Q_e	erro topográfico.
T_e	erro de quantização.
UR	resíduo unificado.
ΔD	variação de potência harmônica.
ΔQ	variação de potência reativa.
ΔS	variação de potência aparente.
$\eta_{En,prec}^i$	fração de energia corretamente identifica em relação ao total de energia detectado para o i -ésimo equipamento.
η_{En}	eficiência de reconstrução em energia do NILM.
η_{En}^i	eficiência de reconstrução em energia para o i -ésimo equipamento.
η_{class}	eficiência de classificação.

η_{det}	eficiência de detecção.
η_{total}	eficiência total.
λ	população da prole.
μ	população dos pais.
ρ_{En}	taxa de reconstrução em energia do NILM.
ρ_{En}^i	taxa de reconstrução em energia para o i-ésimo equipamento.
ρ_{red}	taxa de redundância de energia do NILM.
ρ_{red}^i	taxa de redundância de energia para o i-ésimo equipamento.
θ	ângulo de fase.
ε	energia redundante.
ε^i	energia redundante para o i-ésimo equipamento.
CO_2	gás carbônico.
ΔI	variação de corrente elétrica.
ΔP	variação de potência ativa.
A	unidade de corrente ampère.
BTU	unidade de energia.
C1	equipamentos de consumo permanente.
C2	FSM ou equipamentos de estados múltiplos.
C2a	FSM ou equipamentos de estados múltiplos com ciclos bem-definidos.
C2b	FSM ou equipamentos de estados múltiplos com ciclos aleatórios.
C3	FSM de dois estados ou equipamentos liga/desliga.
C4	equipamentos com níveis contínuos de consumo.
C5	equipamentos com dinâmica de carga.
C6	equipamentos com características similares.
D	potência harmônica.
FP	fator de potência.

Hz	unidade de frequência <i>hertz</i> .
I	corrente elétrica.
kcal	unidade de energia quilocalorias.
P	potência ativa.
Q	potência reativa.
S	potência aparente.
toe	unidade de energia toneladas equivalente de petróleo.
V	unidade de tensão <i>volt</i> .
VA	unidade de potência <i>volt-ampère</i> .
VAr	unidade de potência reativa <i>volt-ampère</i> reativo.
W	unidade de potência <i>watt</i> .
Wh	unidade de energia <i>watt</i> hora.

Lista de Abreviaturas

No caso de algumas abreviaturas internacionalmente conhecidas, optou-se por mantê-las em sua língua original.

ANEEL	Agência Nacional de Energia Elétrica.
ART	Teoria de Ressonância Adaptativa.
AVAC	Aquecimento, Ventilação e Ar Condicionado.
BEU	Balço de Energia Útil.
BLUED	<i>Building-Level fUlly-labeled dataset for Electricity Disaggregation.</i>
BNDES	Banco Nacional de Desenvolvimento Econômico e Social.
cdf	Função de Distribuição Acumulada.
CDM	Mecânismo de Decisão por Comissão.
CEPEL/Eletronbras	Centro de Pesquisas de Energia Elétrica.
CONPET	Programa Nacional de Racionalização do Uso dos Derivados do Petróleo e do Gás Natural.
COPPE/UFRJ	Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia.
CSE	<i>Fraunhofer Center for Sustainable Energy Systems.</i>
dNILM	NILM de Arquitetura Distribuída.
EE	Eficiência Energética.
Eletronbras	Centrais Elétricas Brasileiras S.A..
EMI	<i>Interferência Eletromagnética.</i>
EPE	Empresa de Pesquisa Energética.
ES	Algoritmo Genético de Estratégia Evolutiva.
EUA	Estados Unidos da América.

FEX	Extração de Características.
FFT	Transformada Rápida de <i>Fourier</i> .
FIR	Filtro de Resposta ao Impulso Finita.
FSM	Máquina de Estados Finitos.
GLP	Gás Liquefeito do Petróleo.
HMM	Modelo Oculto de <i>Markov</i> .
IBGE	Instituto Brasileiro de Geografia e Estatística.
ICT	Tecnologia de Comunicação e Informação.
INMETRO	Instituto Nacional de Metrologia, Qualidade e Tecnologia.
ISODATA	<i>Iterative Self-Organizing Data Analysis Techniques Algorithm</i> .
LPS	Laboratório de Processamento de Sinais.
LUR	Menor Resíduo Unificado.
Matriz-U	Matriz de distância Unificada.
MCO	Ocorrência Mais Comum.
MDIC	Ministério do Desenvolvimento, da Indústria e do Comércio Exterior.
MLE	Estimativa de Máxima-Verossimilhança.
MLP	<i>MultiLayer Perceptron</i> .
MME	Ministério de Minas e Energia.
MSE	Mínimo Erro Quadrático.
NILM	Monitoração Não-Invasiva de Cargas Elétricas.
OCDE	Organização para a Cooperação e Desenvolvimento Econômico.
ONU	Organização das Nações Unidas.
PBE	Programa Brasil de Etiquetagem.
PCA	Análise de Componentes Principais.
PCD	Análise de Componentes Discriminantes.
PDE	Plano Decenal de Energia.

pdf	Função de Distribuição de Probabilidade.
PED	País em Desenvolvimento.
PEE	Programa de Eficiência Energética.
PEE/COPPE	Programa de Engenharia Elétrica da COPPE/UFRJ.
Petrobrás	Petróleo Brasileiro S.A..
PME	Programa de Mobilização Energética.
PNE2030	Plano Nacional de Energia 2030.
PNEf	Programa Nacional de Eficiência Energética.
PPH	Pesquisa de Posse e Hábito de Eletrodomésticos.
PROCEL/Eletrabras	Programa Nacional de Conservação de Energia Elétrica.
PROESCO	Programa de apoio a Projetos de Eficiência Energética.
Proálcool	Programa Nacional do Álcool.
RBF	<i>Radial Basis Function.</i>
REDD	<i>The Reference Energy Disaggregation Data Set.</i>
RNA	Rede Neural Artificial.
ROC	<i>Receiver Operating Characteristic.</i>
SOM	Mapas Auto-Organizáveis.
SVD	Decomposição em Valores Singulares.
SVM	<i>Support Vector Machine.</i>
TED	<i>The Energy Detective.</i>
THD	Distorção Harmônica Total.
TI	Tecnologia da Informação.
TW	Transformada <i>Wavelet</i> .
VAST	Algoritmo de <i>Viterbi</i> com Transições Esparsas.

Capítulo 1

Introdução

A proposta de sustentabilidade, surgida como resposta à degradação do meio ambiente devido ao processo de industrialização, leva, em seu discurso, a questão das limitações tanto em termos de recursos presentes na natureza, ou quanto da sua capacidade de absorver resíduos de processos em ordens cada vez mais intensas. Sua vertente predominante traz a ideia do desenvolvimento sustentável, que busca regular a demanda de recursos do meio ambiente, de modo que seja possível o desenvolvimento sem que haja sua deterioração para as gerações futuras.

Adicionalmente, há uma correlação entre os impactos ambientais e a necessidade energética, onde os motivos prevaletentes para suas causas têm como fator de destaque a sua vinculação à cadeia da energia, desde a produção ao uso final. Entretanto, essa dependência não é decorrente apenas da intensidade, mas da eficiência energética relacionada com o seu consumo. Entende-se como eficiência energética a capacidade de produzir os mesmos resultados finais realizando o menor consumo possível de energia.

A eletricidade tem uma participação cada vez maior na matriz energética mundial devido à sua maior versatilidade, eficiência, limpeza, segurança e conveniência, quando comparada com as outras fontes energéticas. A eficiência energética do ponto de vista da eletricidade traz uma série de vantagens, além de benefícios ambientais:

- reduz ou posterga as necessidades de investimentos em geração, transmissão e distribuição de energia elétrica;
- reduz o custo de energia para o consumidor final;
- contribui para a confiabilidade do sistema elétrico;
- traz o aumento da atividade econômica com a redução da intensidade energética.

Além disso, o crescimento do consumo nos grandes centros urbanos tem levado ao surgimento de uma nova rede elétrica — as redes elétricas inteligentes —, que

procura descentralizar a geração devido à exigência excessiva da capacidade de transmissão e distribuição da versão usual centralizada. Dependendo da regulamentação dos medidores inteligentes — os medidores das novas redes elétricas —, uma série de novas possibilidades podem ser exploradas, como:

- maior versatilidade de operação e planejamento das redes elétricas devido à maior informação presente;
- tarifação variável conforme a demanda na rede, incentivando os consumidores a deslocarem cargas não-essenciais para operarem em horários fora de ponta de forma a reduzir ou postergar a necessidade de investimentos na expansão da rede — deslocamento de carga;
- o retorno da informação do consumo em tempo real nas residências pode ser explorado para obter eficiência energética, assunto melhor debatido a seguir.

1.1 Motivação

Uma tecnologia que tem adquirido maior interesse como forma de aliviar a pressão de consumo nos grandes centros urbanos e de atingir maior eficiência energética é a Monitoração Não-Invasiva de Cargas Elétricas (NILM, em inglês), seja no mundo corporativo — atraindo empresas como *Intel* e *Belkin* — como no meio acadêmico, em especial nos países desenvolvidos.

O monitoramento não-invasivo utiliza-se de um único medidor central no fornecimento de energia da residência para identificar o consumo dos equipamentos através dos distúrbios causados na rede elétrica pelos mesmos. Para isso, empregam-se técnicas de processamento de sinais, inteligência computacional e estatística para identificar os padrões dos distúrbios e correlacioná-los com o equipamento de origem. Dependendo da metodologia aplicada no NILM, isso pode ser realizado de maneira cega, ou seja, encontrando padrões recorrentes na rede elétrica e os identificando quando eles se repetem. A topologia mais comum pode ter sua operação resumida através destes passos:

- Aquisição de dados: a eficácia do NILM depende diretamente da capacidade do medidor de extrair informação da rede elétrica, sendo desejável amostragens com frequências elevadas ou com uma maior quantidade de representações, independentes entre si;
- Extração de características: com base na informação obtida pelo sistema de aquisição de dados, é possível transformá-la em características que serão utilizadas pelas etapas seguintes para obter o consumo desagregado por equipamento;

- Detecção de eventos de transitório na rede devido à mudança de operação de um equipamento: quando o NILM utiliza a informação no transitório na operação dos equipamentos para identificar a operação do equipamento e estimar o seu consumo — configuração mais comum —, é necessário detectar esses momentos e diferenciá-los de ruídos causados por demais oscilações no consumo da rede;
- Identificação da operação do equipamento e seu consumo: nessa etapa, processam-se as características do consumo de forma a reconhecer os padrões para estimar o consumo dos equipamentos. Geralmente essa tarefa ocorre somente quando é identificado um distúrbio na rede, reduzindo o processamento dos dados.

As aplicações do NILM são diversas, como monitoramento da qualidade de energia, diagnóstico de carga, identificação de equipamentos defeituosos ou com consumo excessivo de energia. No que se refere à eficiência energética, o mesmo pode auxiliar destas maneiras:

- Auxiliar na obtenção de dados para estudos de eficiência energética de eletricidade: no caso do setor residencial, os estudos de eficiência energética precisam de dados com informação desagregada por equipamento para obter melhor precisão e direcionar os esforços para sua obtenção. Atualmente, o levantamento para o setor residencial é realizado através de pesquisas de posse e hábito de consumo, havendo tanto uma demora para a sua obtenção, quanto sofre de certa degeneração devido à sua imprecisão. Essa tecnologia pode auxiliar no processo de obtenção dessa informação, melhorando a precisão e agilizando a obtenção de dados recentes;
- Fornecer o retorno de informação de consumo desagregado por equipamento para o consumidor: diversas pesquisas nos países desenvolvidos mostram que retornar uma informação mais detalhada para o consumidor — além daquela contida na conta de energia — é favorável no sentido de incentivá-lo a tomar ações para redução do consumo e, conseqüentemente, melhor eficiência energética. Esses estudos indicam que quanto maior for a quantidade de informação disponível melhor será essa capacidade, sendo o melhor caso a informação de consumo desagregado por equipamento em tempo-real, justamente a capacidade dessa tecnologia. Porém, nesse caso, não apenas se faz somente necessário o retorno da informação de consumo da mesma, mas também o envolvimento de outras áreas do conhecimento como psicologia e *design* da informação para que seu emprego seja eficiente e motive o consumidor para tomar ações sustentáveis no uso de energia.

O NILM pode aproveitar-se das redes elétricas inteligentes para obter a informação desagregada através da utilização dos medidores inteligentes — subordinado à capacidade de fornecer informação dos mesmos, onde uma taxa mínima de 1 Hz é indicada nesse sentido. Isso facilitaria a impregnação do método e, com isso, maximizaria seu potencial de eficiência energética ao atender ambos itens anteriores em maior escala. Porém, isso está sujeito a limitação da escolha da configuração dos medidores inteligentes no Brasil, de forma que o projeto do NILM, apesar de ser desejável sua impregnação, deve ser realizado sem a certeza de poder utilizar a futura estrutura provida pelas redes elétricas inteligentes no país.

A pluralidade de técnicas aplicadas na tecnologia envolvendo o tema revelam que o seu projeto, apesar de parecer simples, na verdade engloba diversos desafios. O maior desafio geralmente encontrado pelos autores é expandir a técnica para a aplicação em condições reais de operação das redes elétricas residenciais, aonde estarão presentes diversos equipamentos operando simultaneamente, alguns deles com dinâmica no seu consumo, o que torna complexa a identificação dos padrões deixados na rede por demais equipamentos.

1.2 Objetivo

O CEPEL/Eletrobras vem atuando no desenvolvimento de um NILM. Este trabalho tem como objetivo auxiliar no progresso desse dispositivo, sendo desenvolvido em parceria com o CEPEL/Eletrobras.

A principal intenção de aplicação da tecnologia pelo CEPEL/Eletrobras é para o auxílio nas Pesquisas de Posse e Hábito de Eletrodomésticos (PPHs) no setor residencial. Outro motivo para focar apenas nesse setor é a maior dificuldade de aplicação no setor comercial e industrial, que apresentam redes elétricas com um nível de desafio mais elevado quando comparado ao setor residencial. Não obstante, a rede residencial já apresenta obstáculos suficientes a serem superados, em especial nas condições de operação mais ativas da mesma. Nesses momentos, há uma maior quantidade de equipamentos operando, o que adiciona, potencialmente, dinâmica no consumo da rede e, com isso, torna a tarefa a detecção dos eventos de transitório dos estados de operação dos equipamentos não trivial pois há uma menor relação sinal-ruído. Essa dinâmica também tornará a identificação de padrões dos eletrodomésticos mais complexa, aonde o discriminador terá de lidar com distorções em seus padrões.

O trabalho atual irá expandir a metodologia proposta pelo CEPEL/Eletrobras para operação nessas condições, propondo uma abordagem sistemática que permita o ajuste da técnica aplicada de acordo com as condições presentes nas redes elétricas residenciais, limitando-se ao estudo da capacidade da metodologia proposta em de-

teccão de eventos de transitório. Porém, como será visto a seguir, o trabalho trouxe diversas outras contribuições.

1.3 Contribuições do Trabalho

Durante o levantamento bibliográfico, percebeu-se um forte apelo no exterior às questões de eficiência energética que vão além do intuito de aplicação do NILM nas PPHs. Diversos estudos nos Estados Unidos e Europa Ocidental citam a capacidade do consumidor de economizar energia ao retornar sua informação de consumo de energia elétrica. Com o objetivo de compreender melhor como isso pode ser realizado, este trabalho procurou explorar em maiores detalhes essas questões, trazendo uma compilação dos estudos que parecem ser de maior relevância envolvendo esse tema.

Como a própria questão das PPHs estão relacionadas com a eficiência energética, ficou evidente a necessidade de trazer no corpo do trabalho um levantamento da origem do tema — discurso ambiental e ecologia — e o que tem sido feito no mundo e no Brasil nesse sentido para tornar o assunto de mais fácil acesso para os leitores, que podem não estar familiarizados com o tema e importância das questões ambientais envolvidas.

Além disso, o NILM vem sendo desenvolvido desde 1992, sendo possível encontrar uma vasta quantidade de metodologias propostas por diversos autores. O trabalho centralizou e uniformizou, na medida do possível, a informação relevante para o projeto dessa tecnologia para facilitar seu desenvolvimento pelo CEPEL/Eletronbras e por demais autores quando levando em conta o tema de eficiência energética.

Já quanto à metodologia aplicada, como foi dito, o NILM constitui-se de diversas etapas, sendo necessário tratar de todas elas para que o projeto seja aplicável. Cada uma delas precisa ser estudada e compreendida, de tal modo que nem sempre é possível tratar de todos os pontos em um único trabalho. O CEPEL/Eletronbras propôs uma nova abordagem para a detecção de eventos de transitório na operação dos equipamentos, que utiliza como núcleo um filtro de derivada de Gaussiana. Essa abordagem será explorada e estendida pelo trabalho, porém, limitar-se-á à questão de detecção de eventos de transitório — apenas uma das etapas necessárias para a operação do NILM. Indo além, é importante notar que o objetivo do trabalho é estudar o comportamento da metodologia proposta pelo CEPEL/Eletronbras e sua extensão realizada no trabalho em cenários de aplicação prática. Os conjuntos de dados com essas condições foram fornecidos pelo próprio CEPEL/Eletronbras.

A evolução da metodologia proposta pelo CEPEL/Eletronbras foi realizada em termos de estruturação e sistematização. Também no levantamento bibliográfico se percebeu a capacidade de complementação das técnicas aplicadas, de forma que um

ambiente único de análise não apenas permite uma melhor compreensão das técnicas e rapidez no desenvolvimento do projeto, mas também de explorar a capacidade de suplementar outras técnicas, o que permite ao NILM explorar uma maior quantidade de equipamentos e/ou eficiência de desagregação do consumo por equipamento.

O ajuste dos parâmetros realizados pela metodologia do CEPEL/Eletrbras era feito empiricamente, aonde se viu a necessidade de sistematizar o processo. Foi implementado um sistema de otimização através de um algoritmo genético de estratégia evolutiva para auxiliar no ajuste dos parâmetros.

1.4 Estrutura Capítular

Os Capítulos 2 e 3 compilam a informação sobre as aplicações do NILM para eficiência energética. O Capítulo 2 tratará a origem do tema de eficiência energética (Seção 2.1) e prosseguir até as suas necessidades para sua melhor obtenção no Brasil (Seção 2.3), em especial ao que concerne energia elétrica para o setor residencial.

Por outro lado, o Capítulo 3 trata da expansão do potencial de eficiência energética através de um novo programa abrangendo o tema. Antes de tratá-lo, é realizado considerações de como são obtidos os potenciais de eficiência energética para o setor residencial na Seção 3.1. Somente em seguida, na Seção 3.2, é tratado como expandir esse potencial através do retorno de informação para o consumidor. É feito uma compilação das informações relevantes obtidas ao observar estudos no exterior envolvendo o tema.

O Capítulo 4 realizará uma introdução sobre a tecnologia na Seção 4.1, para então trazer um extenso levantamento da informação envolvida do ponto de vista de desenvolvimento, identificando os aspectos já resolvidos, as dificuldades e as tendências para o seu futuro na Seção 4.2. Em sequência, as técnicas aplicadas pelo CEPEL/Eletrbras e os trabalhos anteriores em colaboração com a COPPE/UFRJ serão tratados na Seção 4.3, aonde será realizado o levante das dificuldades e necessidades do projeto. Nessa seção, também está disponível a metodologia original do CEPEL/Eletrbras na Subseção 4.3.2.

As informações levantadas na Seção 4.3 serão resumidas e discutidas na Seção 5.1 do Capítulo 5. Esse capítulo se dedica à descrição do ambiente desenvolvido com o objetivo de melhorar a capacidade e sistematizar a análise, bem como unificar o projeto para facilitar sua continuidade. As alterações da metodologia original podem ser encontradas ao longo desse capítulo, no entanto, na Seção 5.5 é descrita a principal ampliação realizada na metodologia original, o Módulo de Otimização dos Parâmetros.

O Capítulo 6 irá descrever as condições simuladas na base de dados fornecida pelo CEPEL/Eletrbras (Seção 6.1), contendo tanto condições mais simples com

apenas acionamentos e desacionamentos, quanto condições mais próximas às reais nas redes elétricas residenciais, com operação simultânea de diversos equipamentos e dinâmica no consumo de alguns equipamentos. Também é apresentado a metodologia para a aplicação do algoritmo genético para ajuste dos parâmetros (Seção 6.2). Os resultados para a metodologia empregada por este trabalho estão no Capítulo 7. A conclusão do trabalho encontra-se no Capítulo 8.

Capítulo 2

Eficiência Energética

O trabalho atual está completamente imerso no escopo de Eficiência Energética (EE), sendo realizado neste capítulo uma abordagem ao tema. Apesar do discurso ambiental estar presente em nosso cotidiano, leitores não habituados com essas questões podem não ter real noção da importância do tema na atualidade. Grande parte da evolução do ser humano nos últimos dois séculos deu-se pela ampla expansão de sua capacidade de extrair da natureza energia e insumos, e transformá-la de forma a moldar o meio externo para operar em seu favor, tanto para melhorar sua qualidade de vida ou quanto para superar limitações biológicas. Computadores permitem-nos fazer milhões de cálculos em segundos, os meios de transporte permitem-nos deslocar centenas de quilômetros ou distâncias continentais em questões de horas, as telecomunicações permitem-nos entrar em contato e saber o que ocorre do outro lado do mundo praticamente em tempo real — nós superamos em muito a capacidade de nossos antepassados ao criarmos meios de transcender as limitações de nossos corpos.

A questão para a evolução humana seria apenas continuar extraindo mais energia e insumos do meio ambiente e transformá-los a nosso bel-prazer se não houvesse uma limitação da capacidade do meio de fornecê-los, bem como absorver os resíduos desses processos sem que ocorra sua degradação e desequilíbrio. E é justamente nesse ponto em que chegamos com os avanços dos últimos dois séculos, de maneira que agora a questão para nossa evolução não é apenas mais em como extrair mais do meio, mas em como realizar isso de modo eficiente e a respeitar suas limitações tanto de fornecer insumos ou quanto absorver resíduos. A Seção 2.1 aprofundar-se-á nesses assuntos, em especial quanto à cadeia energética, fator de destaque nas questões ambientais. Em seguida, reduz-se a abrangência para o Brasil (Seção 2.2), apresentando o que tem sido feito por órgãos governamentais para a busca de uma melhor EE.

Do ponto de vista de aplicação do NILM, o mesmo pode ser aplicado como uma solução à dificuldade para os estudos e planejamento em EE realizados por entidades

governamentais, que necessitam de dados recentes sobre a posse e hábitos de uso dos equipamentos. A técnica pode ser aplicada em conjunto ou substituir gradualmente o meio de coleta desses dados no setor residencial para o uso-final de eletricidade, assunto melhor debatido na Seção 2.3.

2.1 Da Origem

Estima-se que o ser humano consome entre 2.500 e 3.000 kcal/dia sob a forma de alimentos, onde apenas cerca de 20% dessa energia poderá ser reinvestida em atividades, de forma que a capacidade biológica humana de produção através de energia endossomática — provida por seu metabolismo — é de apenas 500 a 600 kcal/dia. É a particularidade do ser humano de utilizar fluxos de energia exossomáticos — fontes não provenientes de seu metabolismo — que o permite suplantar sua capacidade biológica de interagir com o meio, tornando possível uma melhor qualidade de vida [1]. Atualmente, o consumo energético exossomático médio per capita diário mundial supera 50.000 kcal [2]¹.

Os fluxos exossomáticos estão presentes em todas sociedades e são dependentes do nível de desenvolvimento das mesmas. Em sociedades primitivas, os resultados dos atos da humanidade exerciam pouca influência no meio ambiente, pois estes eram absorvidos por aquele. Entretanto, com o surgimento de sociedades industriais, ocorreu uma mudança radical no uso dos recursos naturais e nos seus efeitos ambientais. Com o crescimento populacional adicionados da produção e do consumo de massa, esses baseados no uso intensivo do petróleo e da eletricidade quanto fontes energéticas, a agressão humana ao meio ambiente tornou-se maior de maneira que o meio ambiente, em muitas situações, não consegue mais se reequilibrar. Os principais problemas ambientais que surgiram no cenário mundial tem como fator de destaque vinculados a cadeia da energia, desde a produção ao uso final [1, 3].

Indo mais além, a dependência desenvolvimento-energia não está somente atrelada ao patamar macroscópico da sociedade, mas também relacionada ao modo que o desenvolvimento está difundido nela. Diferentes classes sociais de uma determinada sociedade detêm modos de consumo, bem como formas de acesso a bens e serviços distintos. Nesse caso, aqueles com menor poder aquisitivo não apenas consomem menos energia, mas também tipos diferentes da mesma, geralmente mais poluentes e menos eficientes como aquelas oriundas de biomassa [1]. Percebe-se, portanto, que o avanço do perfil energético de uma sociedade não indica necessariamente a sua evolução como um todo, sendo necessário notar como a energia está distribuída nela.

¹Esse valor não está explícito na referência. É necessário realizar a divisão da produção bruta de energia pela população mundial e o número de dias no ano.

Em resposta à degradação do meio ambiente se dá a proposta da sustentabilidade. As primeiras objeções à poluição ambiental devido ao processo de industrialização, em especial após sua aceleração no início do século XX, foi da ecologia radical, dando atenção as questões de proteção e conservação da natureza. Há uma separação de territórios especiais para uma proteção integral e, numa visão em larga medida romantizada da natureza, muitas vezes sem permissão de nenhum uso antrópico [3].

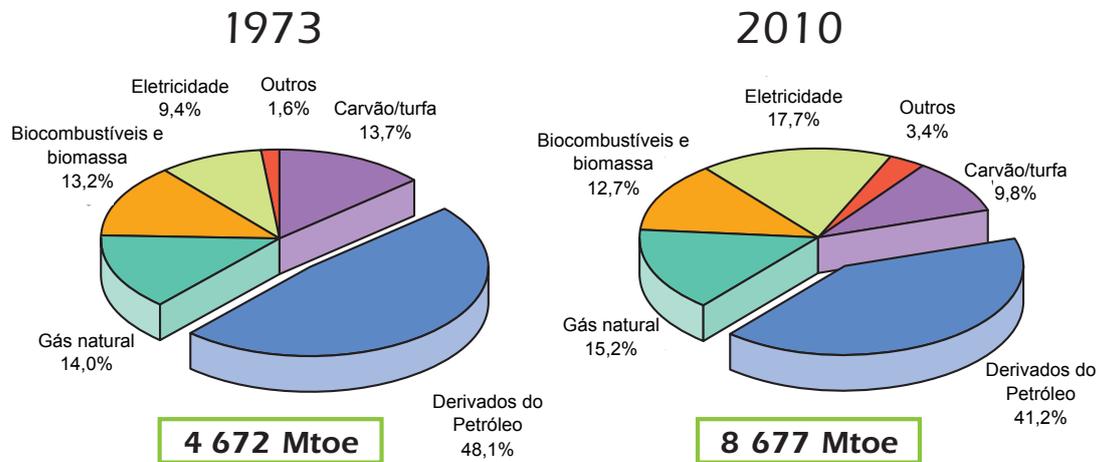
A política ambiental preponderante atualmente é o discurso do ambientalismo moderado, surgido em meio a Crise do Petróleo na década 1970, onde, percebendo-se a inviabilidade de sustentação do modelo econômico levando em conta o esgotamento progressivo dos recursos naturais do planeta, fez-se necessário a discussão de como colocar em prática as propostas da ecologia radical, mas tendo cautela de não necessariamente frear o crescimento econômico ou alterar substancialmente o modelo de desenvolvimento vigente — trazendo o conceito de desenvolvimento sustentável. Devido à Crise, houve uma busca mundial pela redução da dependência no petróleo e outras fontes fósseis na matriz energética, procurando fontes renováveis e menos poluentes que as fontes fósseis [1, 3, 4].

Vale aqui ressaltar que as questões ecológicas não estão ligadas apenas a economia, mas refletem uma intensa dinâmica econômica, social, cultural e política que são contraditórias entre si. Há também uma linha do discurso ambiental que destaca a justiça social, notando que os mais vulneráveis aos problemas ambientais são justamente os mais pobres, os quais serão os mais afetados na hipótese de agravamento da crise ambiental [3].

A maneira mais efetiva, eleita pela Organização das Nações Unidas (ONU), de reduzir os impactos ambientais locais e globais sem prejuízos ao desenvolvimento, refletindo a imagem de desenvolvimento sustentável, é através da EE [1, 5, 6]. O uso eficiente de energia deve ser entendido como o menor consumo possível de energia para obter uma mesma quantidade de produto ou serviço [7]. As ações de EE compreendem modificações ou aperfeiçoamentos tecnológicos ao longo da cadeia, mas podem também resultar de uma melhor organização, conservação e gestão energética por parte das entidades que a compõem [8].

A eletricidade como recurso energético vem adquirindo importância cada vez maior na matriz energética mundial — e assim nos debates englobando EE — devido a, geralmente, sua maior versatilidade, eficiência, limpeza, segurança e conveniência quando comparando com as outras fontes energéticas. O consumo de eletricidade no uso final é de 17,7% da matriz energética mundial em 2010, enquanto sua participação era de 9,4% em 1973, ver Figura 2.1. Estima-se para 2035 um crescimento de 80% da demanda de eletricidade quando utilizando o ano base de 2008, de modo a elevar a participação da eletricidade para uma parcela de 23%. O crescimento da demanda energética de eletricidade é liderado pelos países não pertencentes a

Organização para a Cooperação e Desenvolvimento Econômico (OCDE, em inglês), dentre eles o Brasil, que terão uma participação de cerca de 80% do crescimento da demanda de eletricidade mundial [9].



Obs: Biocombustíveis e biomassa estimados para 1973;
 Outros composto por energia: geotérmica, solar, eólica etc.;
 1 tonelada equivalente de petróleo (toe, em inglês) = 10^7 kcal.

Figura 2.1: A matriz energética mundial em 1973 e 2010. Extraído de [2, pp. 28, tradução própria].

Além dos benefícios ambientais, quando levando em conta o sistema elétrico, a EE [10, 11]:

- reduz ou posterga as necessidades de investimentos em geração, transmissão e distribuição de energia elétrica;
- reduz o custo de energia para o consumidor final;
- contribui para a confiabilidade do sistema elétrico;
- traz o aumento da intensividade econômica com a redução da intensividade energética.

A tendência nos países desenvolvidos é a realização de esforços cada vez maiores no sentido de aumentar a EE, estimando-se, por exemplo, um consumo 49% maior nos países da OCDE no período de 1973 a 1996 caso não houvessem sido adotadas medidas de racionalização e EE após a Crise do Petróleo [11, 12].

2.2 Eficiência Energética no Brasil

Os Países em Desenvolvimento (PEDs) não possuem uma capacidade de redução energética tão grande quanto os países da OCDE, uma vez que seu consumo

energético per capita é reduzido, justamente por ser necessário o desenvolvimento para aumentar o consumo. Por outro lado, é possível aliviar a pressão sobre a oferta energética através da melhoria dos níveis de EE nos diversos setores da sociedade, contrapondo ao pensamento de que para que haja desenvolvimento são necessários impactos ambientais e o crescimento no consumo total de energia — o chamado de efeito *leapfrogging* [12], Figura 2.2.

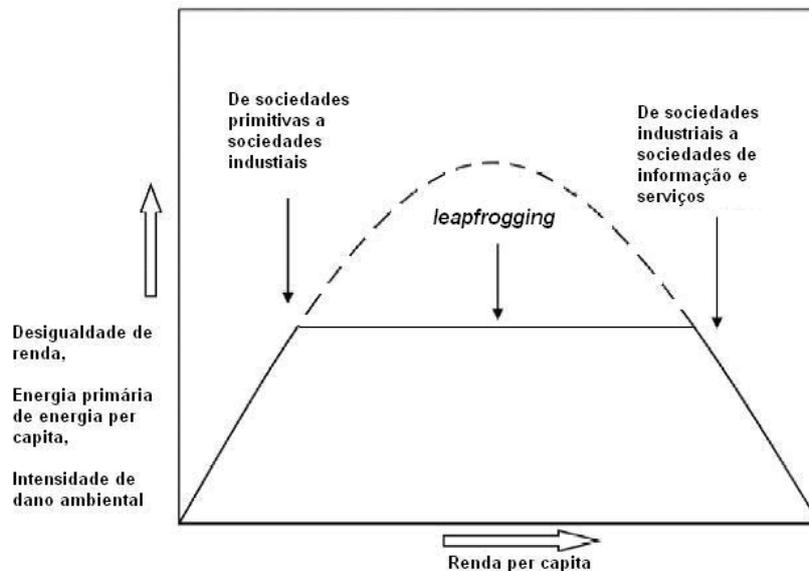


Figura 2.2: O efeito *leapfrogging*. Extraído de [12].

Por isso, a preocupação com EE se justifica mesmo no Brasil, que apresenta quase metade de sua matriz energética proveniente de fontes renováveis e preços de produção de energia economicamente competitivos. Diversas iniciativas vêm sendo empreendidas, dentre elas, destacam-se [7, 8, 11, 13]:

- 1931: primeiro horário de verão no Brasil;
- 1975: criado o Programa Nacional do Álcool (Proálcool), para a substituição em larga escala de veículos movidos a derivados de petróleo por álcool;
- 1981: criado o Programa CONSERVE, visando à promoção da conservação de energia na indústria, ao desenvolvimento de produtos e processos energeticamente mais eficientes, e ao estímulo à substituição de energéticos importados por fontes alternativas autóctones²;
- 1982: aprovada as diretrizes do Programa de Mobilização Energética (PME), conjunto de ações dirigidas à conservação de energia e à substituição de derivados de petróleo;

²Fontes de energia produzidas e consumidas localmente.

- 1984: o Instituto Nacional de Metrologia, Qualidade e Tecnologia (INMETRO) implementou o Programa Brasil de Etiquetagem (PBE)³ tendo por objetivo promover a redução do consumo de energia em equipamentos como refrigeradores, congeladores e condicionadores de ar domésticos. O PBE permite ao consumidor obter a informação de eficiência energética dos eletrodomésticos (ver Figura 2.3) que são indicados em classes representadas por letras — essas variando de “A” (mais eficiente) até a letra correspondente de menor consumo (dependendo do produto, pode haver três, quatro ou mais classes, por exemplo, forno microondas possui apenas três classes, enquanto lâmpadas possuem sete). Na cartilha disponibilizada pelo programa [14], estima-se que o consumidor, ao adquirir um modelo de ar condicionado *split* classe “A” de 9.000 BTU, irá economizar R\$ 124/ano em relação a um de mesma potência classe “E”;

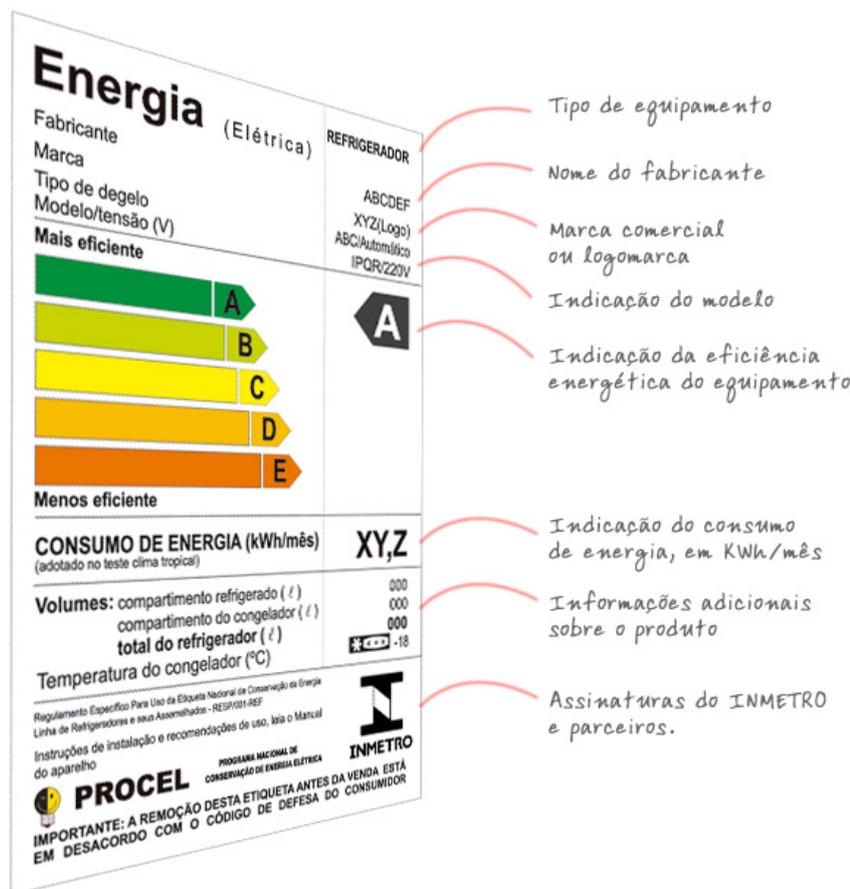


Figura 2.3: A etiqueta do PBE e sua informação contida. Extraído de [15].

- 1985: foi instituído o Programa Nacional de Conservação de Energia Elétrica (PROCEL/Eletrobras), com a finalidade de integrar as ações visando à con-

³Seu nome inicial foi Programa de Conservação de Energia Elétrica em Eletrodomésticos, sendo renomeado em 1992 para a nomenclatura atual.

servação de energia elétrica no país, sendo coordenado executivamente pela Eletrobras. O programa possibilitou uma economia de energia acumulada de 32,9 TWh entre 1986 e 2008, reduzindo a demanda de ponta em 9.538 MW. Somente essa energia economizada corresponde a investimentos evitados de R\$ 22,8 bilhões, enquanto os recursos investidos no programa foram de aproximadamente R\$1,2 bilhão, ambos para o período citado;

- 1991: foi instituído o Programa Nacional de Racionalização do Uso dos Derivados do Petróleo e do Gás Natural (CONPET), programa similar ao PROCEL/Eletrobras, mas destinado à parcela da matriz energética proveniente dos derivados do petróleo e do gás natural, com a coordenação executiva da Petrobrás. O CONPET evitou o consumo de 1030,2 milhões de litros diesel no período 2006–2010;
- 1998: iniciado o Programa de Eficiência Energética (PEE), regulamentado pela Agência Nacional de Energia Elétrica (ANEEL). As concessionárias, permissionárias e autorizadas do setor de energia elétrica devem realizar investimentos em pesquisa e desenvolvimento promovendo a conservação de energia elétrica. Atualmente esse valor é de 0,50% da receita operacional líquida das operadoras. Foi realizado um investimento total em PEEs pelas concessionárias do setor elétrico na quantia de R\$ 2 bilhões até o ano de 2006, onde se estima que o programa alcançou uma economia média de 4.000 GWh/ano e retirando uma carga de ponta de consumo na ordem de 1.140 MW no período de 1998–2005;
- 2001: Lei nº 10.295/2001 determina a instituição de “níveis máximos de consumo específico de energia, ou mínimos de EE, de máquinas e equipamentos consumidores de energia fabricados e comercializados no país”, sendo regulamentada pelo Decreto nº 4.059/2001;
- 2004: criada a Empresa de Pesquisa Energética (EPE), que tem entre suas diretrizes promover estudos e produzir informações para subsidiar planos e programas de desenvolvimento energético ambientalmente sustentável, inclusive, de EE;
- 2006: início do Programa de apoio a Projetos de Eficiência Energética (PROESCO), programa destinado a financiar projetos de EE, cuja coordenação executiva pertence ao Banco Nacional de Desenvolvimento Econômico e Social (BNDES);
- 2011: publicado as diretrizes do Programa Nacional de Eficiência Energética (PNEf) [8], que estabelece uma meta de 10% de redução no consumo por meio de ações de EE no período de 2015–2030.

Percebe-se uma evolução brasileira em EE, tanto na legislação, capacitação e conhecimentos acumulados, quanto na consciência da necessidade da EE nos diversos setores [8]. Esse patrimônio, no entanto, precisa ser constantemente atualizado e ter sua abrangência ampliada [7]. Através do planejamento pretende-se que recursos possam ser melhores aplicados e os resultados venham com maior velocidade, abrangência e amplitude. Esforços desse tipo podem ser vistos no PNEf [8], previsto pelo Plano Nacional de Energia 2030 (PNE2030) [7], e nos Planos Decenais de Energia (PDEs) [16].

A grande parcela do potencial de EE no Brasil, de acordo com as estimativas realizadas a partir do Balanço de Energia Útil (BEU) [17], encontra-se nos três maiores setores por consumo final de energia, sendo eles o industrial (37,3%), transportes (30,5%) e residencial (10,2%), que juntos representaram cerca de 78% do consumo energético do país em 2011 [4, 18].

O escopo deste trabalho é no setor residencial, no que tange o consumo de eletricidade como fonte energética. O consumo total dessa fonte no setor é de 41,7% [18], tendo a tendência de continuar se elevando com o crescimento do número de eletrodomésticos e a substituição da lenha (27,8% do consumo no setor) e carvão (2,1%) pela eletricidade, para iluminação e aquecimento durante o desenvolvimento do país.

2.3 Dificuldades para o Estudo em Eficiência Energética

Todavia, há uma grande deficiência na base de dados para os estudos de EE no Brasil, explicitamente indicado no PNE2030 [7, pp. 232] (ver esquema na Figura 2.4):

As entidades encarregadas de planejar, implementar e monitorar programas de eficiência energética necessitam de informações, tais como custos e rendimentos de equipamentos e veículos eficientes, estatísticas detalhadas de vendas de equipamentos e veículos e resultados de pesquisas de campo sobre posse e hábitos de usos dos equipamentos e veículos e sobre as respostas dos diversos grupos de consumidores às diferentes medidas de conservação. Bancos de dados contendo tais informações têm sido montados por diversos países desenvolvidos desde a década de setenta. Em função dos custos significativos das pesquisas de campo, isto não tem ocorrido na maioria dos países em desenvolvimento, inclusive no Brasil. Existe atualmente investimentos do PROCEL/Eletronbras na área, mas que carecem ainda de reforços e da abordagem dos demais programas nacionais. Sem uma base de dados consistente, que inclua o

levantamento de tecnologias disponíveis ou em estudo (e análise de sua potencialidade de mercado) e metodologia de resultados de projetos, não se podem modelar, de forma confiável, programas de eficiência energética no planejamento da expansão do setor energético brasileiro.

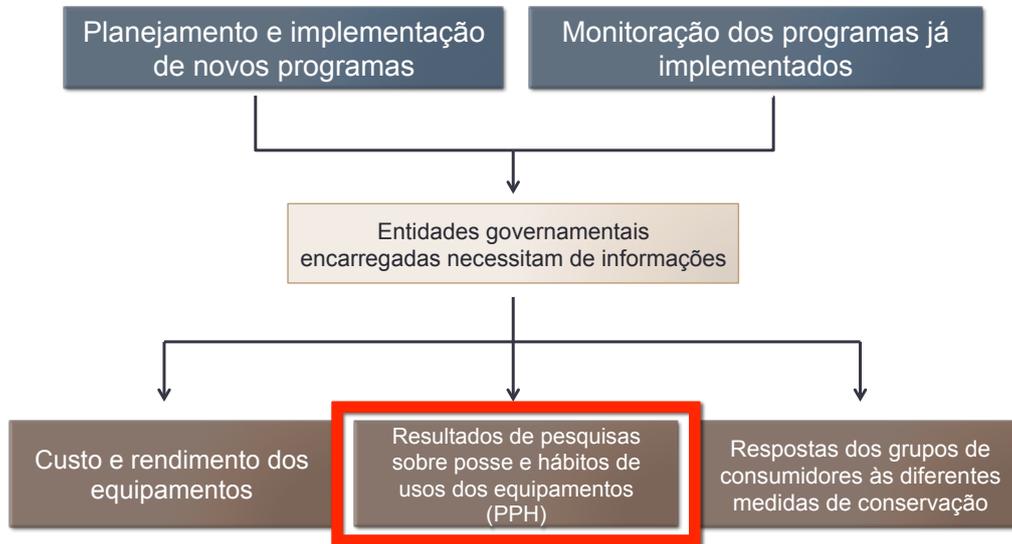


Figura 2.4: Necessidades para o planejamento, implementação e monitoração dos programas de EE pelas entidades governamentais.

Os esforços do PROCEL/Eletrabras citados englobam justamente o setor residencial, em especial detalhando seu consumo de energia elétrica de modo a fornecer informações para os estudos de EE no setor. Eles são o resultado da mais abrangente Pesquisa de Posse e Hábito de Eletrodomésticos (PPH) realizada entre 2004–2006 [19, 20]. Nela, utilizou-se cerca de 15 mil questionários em 18 estados e 20 concessionárias que abrangeram não apenas o setor residencial, mas inclusive o setor industrial e comercial, representando cerca de 92% do mercado consumidor de energia elétrica. Apenas no segmento residencial foram aplicados 9.847 questionários.

As informações na PPH para o setor residencial vão além da posse e hábito de consumo sobre eletrodomésticos, estando presentes informações sobre as características socioeconômicas e comportamentais do usuário, caracterização e identificação da residência, e detalhes sobre aquecimento de água sobretudo em relação ao banho. Desse modo, as PPHs podem ser utilizadas com outras finalidades, como:

- obter informações importantes tais como o consumo durante o horário de pico, fornecendo o fator de carga e fator de demanda;
- revelar causas de perda de energia elétrica devido a conexões ilegais a rede;
- informar a companhias de eletrodomésticos sua participação no mercado e o perfil de seus consumidores;

- fazer um paralelo do desenvolvimento e das necessidades das famílias brasileiras entre classes sociais e regiões do país, explanando as diferentes realidades existentes.

Atualmente, porém, os dados das PPHs dependem da precisão do usuário final e do entrevistador durante o preenchimento do questionário, uma vez que não são utilizadas medições no domicílio, podendo tornar as PPHs imprecisas. O consumidor pode não se lembrar ou não saber informar os dados referentes a uma determinada pergunta, assim como omitir certas informações, por não confiar no profissional desconhecido ou nos intuítos da pesquisa. Pode ocorrer também uma má manipulação por parte do entrevistador das informações, prejudicando a qualidade do material obtido.

Um outro recurso para se obter informações de posse e hábito é através da medição por meio de equipamentos específicos. As medições podem ser utilizadas para complementar, melhorando a precisão das PPHs, ou até mesmo substituí-las, nesse caso, depositando todo o peso de identificação de posse e hábito no sistema de medição. A fim de evitar um maior incômodo ao consumidor, é recomendável que o sistema não faça intrusão à sua propriedade, constituindo de um sistema de medição *não-invasivo*. Ainda, a utilização de um método intrusivo de medição pode alterar, mesmo que inconscientemente, os hábitos de consumo do usuário, prejudicando a qualidade dos dados. Novas técnicas, como o Monitoração Não-Invasiva de Cargas Elétricas (NILM, em inglês), podem ser utilizadas com essa finalidade (Capítulo 4), fornecendo as informações necessárias de posse e hábito do consumidor com maior precisão.

Ainda, vale ressaltar que as pesquisas citadas no texto do PNE2030, escrito em 2007, como uma exceção à precariedade de dados já se encontram defasadas atualmente, sendo um empecilho para a precisão dos estudos de EE, uma vez que os dados utilizados podem ser não fiéis a realidade brasileira atual. Esse dispositivo, além de fornecer melhor precisão, pode agilizar o processo de coleta no país, fornecendo dados com frequência mais adequada aos estudos de EE.

Capítulo 3

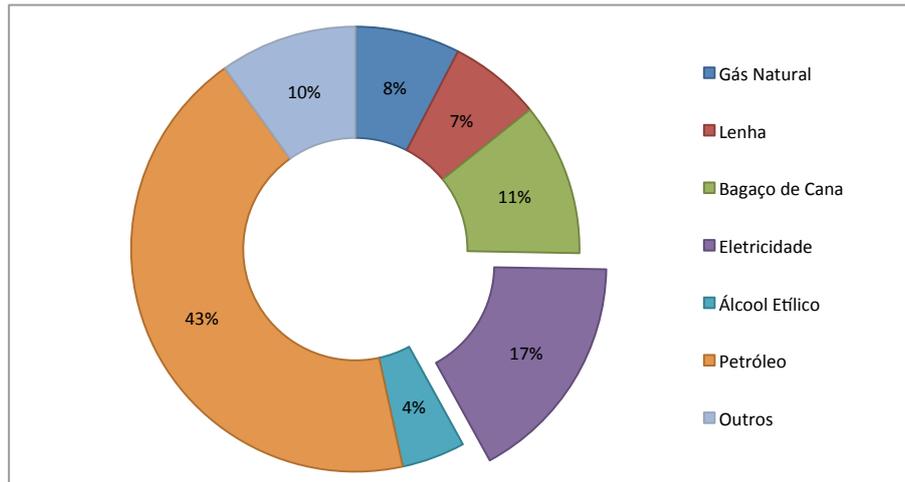
A Eficiência Energética no Setor Residencial

Durante o estudo das técnicas aplicadas no NILM, que serão descritas no próximo capítulo, percebeu-se um apelo no exterior — no caso, em países desenvolvidos como os EUA e outros países da Europa Ocidental — para a capacidade de aplicação do NILM diretamente como uma ferramenta na obtenção de uma melhor EE. Estudos nesses países indicam uma capacidade de redução de consumo de eletricidade acima de 10% no setor residencial ao retornar a informação de consumo em tempo real desagregada por equipamento. Na Seção 3.1, será visto que no Brasil o consumo nesse setor é também relevante, em especial quando levando em conta apenas a eletricidade. Além de detalhar o perfil do consumo no setor residencial, essa seção ainda se aprofundará na questão de eficiência energética, em especial quanto a possibilidade de aplicação do novo programa como um meio de expandir o potencial de EE no setor residencial.

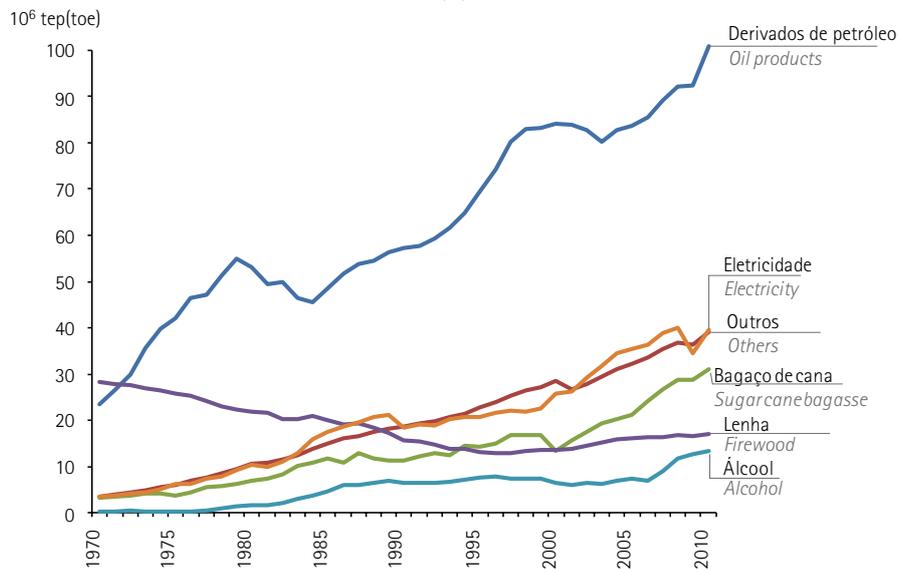
Os estudos utilizados como referências para a capacidade de redução e as considerações por eles levantadas são resumidos e comentados na Seção 3.2. Uma tendência no exterior é aliar os medidores inteligentes — os medidores das redes elétricas inteligentes — ao NILM para garantir o melhor custo-benefício para o programa de EE, garantindo o retorno da informação desagregada por equipamento com baixo custo. Nota-se, entretanto, que há limitações dessa tecnologia de fornecer essa informação, que serão detalhadas no próximo capítulo. Para o programa ser bem-sucedido, é necessário motivar o consumidor para reduzir seu consumo e não apenas retornar a informação. Assim, além das considerações feitas em torno da tecnologia necessária, chama-se atenção, baseando-se em referências, aos aspectos psicológicos e visuais da informação.

3.1 Eletricidade e Potencial de Eficiência Energética no Setor Residencial

A eletricidade é o segundo meio energético de maior participação na matriz energética brasileira, representando 16,7% da demanda, Figura 3.1a. Há uma tendência de crescimento dessa parcela na matriz indicado na Figura 3.1b, que deverá progredir devido à fatores como [9]:



(a)



(b)

Figura 3.1: A matriz elétrica brasileira em 2011 (a) e seu histórico (b). Baseado (a) e extraído (b) de [18].

- substituição da biomassa para eletricidade como meio de iluminação e aquecimento no setor residencial;

- nesse setor também há um aumento do número de eletrodomésticos como reflexo do desenvolvimento e melhor distribuição da riqueza;
- a expansão do setor comercial e de serviços utilizando uma quantidade maiores de equipamentos elétricos (como ar condicionado, iluminação e equipamentos de TI);
- transformação do setor industrial, gradualmente substituindo o carvão e aumentando a presença de dispositivos elétricos.

No que diz respeito ao consumo de eletricidade, Figura 3.2, o setor residencial possui uma posição de destaque no consumo, com uma demanda de 111,97 TWh e uma parcela equivalente à 26% do total no ano de 2011, exprimindo a importância desse setor para explorar a capacidade de economia de eletricidade através de EE.

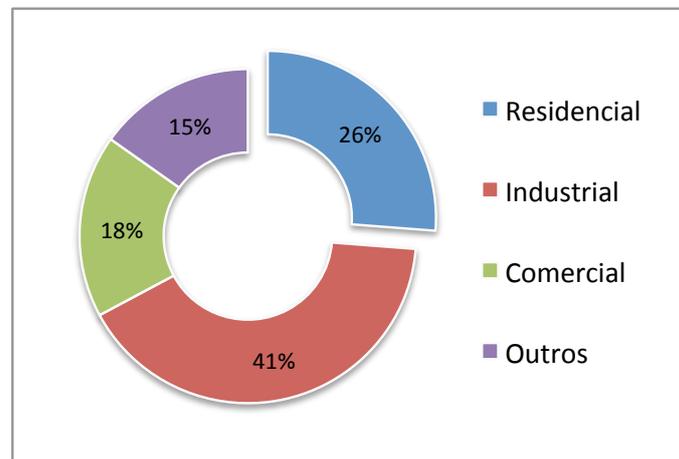


Figura 3.2: O consumo de eletricidade por setor em 2011. Baseado em [18, pp. 32].

De acordo com o último estudo do BEU¹ com o ano base de 2004 [17], o potencial de economia no consumo de eletricidade nesse setor chega a 13,60 TWh, ou 1,17 milhões de toe. Apenas como figura de mérito, quando considerando melhorias de eficiência em todas as fontes energéticas, o setor residencial possui o terceiro maior potencial de economia, após o setor industrial e de transporte, com uma capacidade de economia de 2,97 milhões de toe. Por sua vez, ressalta-se que esses valores de potenciais de economia apresentados são aproximados e reduzidos em relação ao valor real, onde se considera apenas a perda de energia na primeira transformação do processo produtivo. Também não são consideradas possíveis alterações no consumo de fontes energéticas, como a mencionada alteração de biomassa para eletricidade, uma fonte menos poluente e mais eficiente. Finalmente, deve-se acrescentar que esse potencial é calculado utilizando o rendimento de equipamentos no estado da arte entre aqueles normalmente comercializados, e não os possíveis de se alcançar quando

¹O BEU é realizado em intervalos decenais desde 1985.

considerando a literatura técnica. Os dados do BEU são normalmente utilizados para os estudos em EE para o setor industrial e comercial.

No entanto, os dados utilizados pela EPE para os estudos atuais de EE no setor residencial, como os PDEs e PNE2030, por sua vez, são das informações obtidas na última PPH realizada no Brasil entre 2004–2006 que possibilitam o estudo baseado em uma abordagem desagregada. Essa abordagem depende do número de domicílios, da posse média e hábito de consumo específico dos equipamentos eletrodomésticos — que estão implicitamente indicados na curva de carga ilustrada na Figura 3.3 para a região sudeste — e do rendimento médio desses equipamentos no país, informação baseada nas tabelas do PBE, coordenado pelo INMETRO. Também são utilizadas variáveis agregadas para o ajuste do modelo em questão, sendo elas a relação entre o número de consumidores residenciais e população (que permite a projeção do número de consumidores a partir da projeção da população), e consumo médio por consumidor residencial [4, 7, 16].

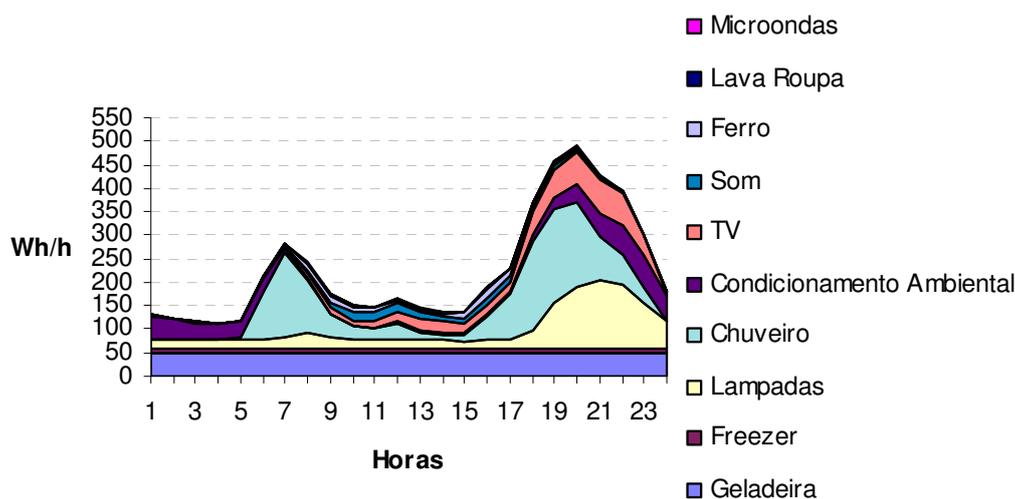


Figura 3.3: Curva de carga média para a região sudeste, ano base 2005. Extraído de [19].

Esses estudos exploram a conservação de energia no ganho de eficiência de equipamentos eletrodomésticos, mas não consideram melhorias possíveis em mudanças nos hábitos de consumo. Isso se justifica uma vez que a abordagem representa o ganho de conservação de energia através do progresso autônomo ou tendencial, progresso esse “que se dá por iniciativa do mercado, sem a interferência de políticas públicas de forma espontânea, ou seja, através da reposição natural do parque de equipamentos por similares novos e mais eficientes ou tecnologias novas que produzem o mesmo serviço de forma mais eficiente” [8, pp. 1], assim como por “efeitos de programas e ações de conservação já em execução no país” [16, pp. 247].

Por outro lado, o progresso induzido refere-se à “instituição de programas e ações adicionais orientados para determinados setores, refletindo políticas públicas; programas e mecanismos ainda não implantados no Brasil” [16, pp. 247]. As políticas de progresso induzido estão detalhadas no PNEf, e a estratégia adotada no tocante às mudanças nos hábitos de consumo é através da educação com os programas do PROCEL/Eletrobras e CONPET, com as linhas: Eficiência Energética na Educação Básica; Eficiência Energética na Formação Profissional; e Rede de Laboratórios e Centros de Pesquisa em Eficiência Energética [8, cap. 5]. No caso do PROCEL/Eletrobras, investiu-se um montante superior a R\$ 4,5 Mi em 2011 em projetos voltados para o desenvolvimento e aperfeiçoamento dessas três linhas [21], trabalhando na conscientização, sensibilização e informação para obter uma melhor EE, atuando nos três níveis de educação.

Entretanto, estudos no exterior mostram um potencial ainda a ser explorado de economia de energia elétrica no setor residencial, através do retorno de informação da utilização de energia para o consumidor, uma estratégia de progresso induzido que pode ser adicionada aos programas de EE no Brasil. Conforme será visto na próxima seção, esse potencial depende de aspectos, como a quantidade de informação retornada ao consumidor, essa inerentemente ligada à quantidade de investimento utilizada na tecnologia para permitir um maior retorno, mas ainda, dependem do modo em que esse retorno é fornecido para o mesmo no intuito de motivar ações sustentáveis de energia, sendo um problema complexo dependente de aspectos sociais, culturais e psicológicos.

Por fim, outros setores também podem reduzir seu consumo com políticas de EE aplicadas para o setor residencial devido a sua natureza similar a esse setor — nesse caso, não se referindo somente à política de EE sugerida. Os setores público e comercial, por exemplo, possuem prédios com natureza de consumo correspondente ao do setor residencial, de forma que estratégias desenvolvidas possam sinergicamente apresentar um potencial maior de economia de energia na matriz brasileira, em especial no que concerne estudos para melhor eficiência de equipamentos.

3.2 Expandindo o Potencial de Eficiência Energética através do Retorno de Informação de Consumo

As novas fontes energéticas, dentre elas a eletricidade e gás natural que atendem a demanda dos consumidores residenciais para a vasta variedade de serviços nas quais são utilizadas — desde cocção, condicionamento do ambiente, a lazer e entretenimento —, fluem invisível e silenciosamente para seus domicílios, sem deixar

qualquer traço notável de sua utilização além do efeito final desejado pelo consumidor. Para eles, o único retorno de seu consumo é informado na conta apresentada pela concessionária, fornecida em um longo período após o consumo (mensalmente, por exemplo). As informações nas contas são precárias, não informando muito além do total de energia consumido e o preço de energia. Os usuários não tem como inferir quais são os meios de uso final que demandam maior energia, nem a que ponto possíveis mudanças podem afetar sua demanda, seja através da mudança de seus hábitos ou na escolha de equipamentos mais eficientes. Atualmente, os usuários estão cegos quanto a essas mudanças, não é possível visualizar a energia que consomem. Além disso, as informações fornecidas não permitem o consumidor comparar seu consumo com o de outros, de modo que ele não é capaz de criar uma referência social para seu consumo. Sem uma referência, o consumidor tem dificuldades para determinar se o consumo é excessivo ou moderado e se é necessário algum tipo de intervenção [22].

Estratégias para intervir no comportamento podem ser classificadas de dois modos [22, 23]:

1. **Antecedentes**, que envolvem esforços para influenciar o que define um comportamento antes de sua realização;
2. **De consequência**, que buscam alterar o que determina o comportamento após a sua ocorrência.

Exemplos de estratégias antecedentes são campanhas de informação com o objetivo de aumentar o conhecimento público sobre o impacto de suas escolhas e das opções para economia de energia disponíveis — como as já citadas ações do PROCEL/Eletrobras e CONPET (ver pp. 22) —, engajar o indivíduo com um compromisso de mudança, criar metas de mudança comportamentais, ou modelar e demonstrar o comportamento desejado. Já para estratégias de consequência se pode citar recompensas, punições ou o retorno de informação [22, 23].

Iniciativas utilizando o retorno de informação mostraram-se altamente eficientes em mudanças comportamentais com relação ao consumo energético [22–29]. O uso do retorno de informação baseá-se em que tanto resultados positivos ou negativos podem modelar o comportamento. Resultados atribuídos como positivos irão torná-los em comportamentos mais atraentes, enquanto a atribuição de resultados negativos propiciam comportamentos ruins em menos desejáveis. Sempre que possível, a atribuição negativa deve ser evitada, pois ela tende a reduzir a motivação e não coloca nada no lugar do comportamento evitado [30].

A questão, por outro lado, não é apenas fornecer retorno do consumo ao usuário final — a própria conta de energia pode ser encarada como um meio de retorno —, mas como o retorno pode ser utilizado para efetivamente motivar pessoas para

reduzir o seu consumo. Algumas considerações devem ser tomadas: primeiro, quais são os tipos de retornos disponíveis (Subseção 3.2.1) e, dentre eles, quais tem mostrado resultados mais eficientes na redução do consumo energético (Subseção 3.2.2)? O que mais deve ser levado em consideração quando preparando tais programas e estudos de EE (Subseção 3.2.3)? Quais são as tecnologias disponíveis para fornecer essa informação e suas tendências (Subseção 3.2.4)? Ainda, pessoas possuem diferentes atitudes, crenças e valores, sendo motivadas de modos distintos. Uma breve consideração sobre a perspectiva psicológica que envolve mudança comportamental será realizada (Subseção 3.2.5) uma vez que esse aspecto é de principal relevância para o sucesso dos programas. Do mesmo modo, a apresentação visual da informação também irá influenciar no êxito, e por isso o tema também será colocado em pauta (Subseção 3.2.6). De nenhuma maneira as breves considerações realizadas nas subseções sobre os aspectos psicológicos e visuais devem substituir a análise de profissionais dessas áreas, servindo apenas para chamar atenção para a importância, bem como introduzir os leitores, aos temas.

3.2.1 Tipos de Retorno

A categorização dos tipos de retorno que será apresentada se iniciou em [31] e depois foi aprimorada por [23]. A primeira divisão toma em conta o modo no qual o retorno é fornecido, sendo possíveis o retorno direto ou indireto. O termo *indireto* é utilizado quando há alguma espécie de processamento antes de atingir o consumidor, enquanto *direto* determina o retorno instantaneamente² entregue ao usuário. Em seguida, realiza-se uma divisão em termos de frequência, diferenciando quatro tipos de retorno *indiretos*, e dois retornos *diretos*. A divisão é apresentada resumidamente em ordem crescente em termos de custo e quantidade de informação disponível, onde os dois retornos *diretos* estão no final da lista [22, 23]:

1. **Retorno por Faturamento Simples:** conta de energia contendo kWh consumido, o preço da tarifa unitária (R\$/kWh), o custo total e outros possíveis ônus. Nessa forma de retorno, normalmente se carece estatísticas comparativas ou qualquer informação detalhada sobre os aspectos temporais do consumo;
2. **Retorno por Faturamento Aprimorado:** fornece informações mais detalhadas sobre o padrão de consumo de energia, incluindo em alguns casos estatísticas comparativas, tanto comparando o maior consumo do mês atual e sua despesa aliados ao consumo histórico e/ou a comparação com outros domicílios pertencentes ao grupo do consumidor;

²A palavra instantâneo é empregada apesar da existência de atraso entre a informação e o consumo pois esse retardo, normalmente na ordem de segundos, é desprezível para os fins.

3. **Retorno Estimado:** essa abordagem utiliza geralmente de técnicas estatísticas para desagregar o total de energia baseado no tipo do domicílio do consumidor, informação de equipamentos e dados de faturamento. O retorno resultante fornece um relato detalhado do uso de eletricidade pelos equipamentos e dispositivos de maior importância. A forma mais comum é através de ferramentas de auditoria de energia residencial baseadas na internet, oferecida por um fornecedor de serviços a seus consumidores;
4. **Retorno Diário/Semanal:** esses relatórios utilizam a média de dados e frequentemente incluem estudos de leitura dos medidores pelos próprios consumidores, assim como estudos nos quais indivíduos são providos com relatórios mensais ou semanais do fornecedor de serviços ou entidade de pesquisa;
5. **Retorno em Tempo-Real:** fornecido por dispositivos que exibem o consumo em (praticamente) tempo-real e informações de custo da energia em nível agregado domiciliar;
6. **Retorno em Tempo-Real Desagregado:** nesse caso, as informações são exibidas desagregadas ao nível dos equipamentos.

3.2.2 Resultados por Tipo de Retorno

No exterior, há uma farta quantidade de estudos envolvendo o tema, com estudos iniciando na década de 1970 em resposta à Crise do Petróleo, que tiveram um declínio durante a década posterior. O interesse retornou a pauta recentemente devido à crescente preocupação com o meio ambiente e mudanças climáticas, assim como o aparecimento de novas possibilidades tecnológicas associadas a Tecnologias de Comunicação e Informação (ICT, em inglês) [22]. É possível encontrar pesquisas nas quais se compilam diversos estudos de retorno de informação para consumidores residenciais no intuito de generalizar resultados [22, 25–28].

Infelizmente, o mesmo não pode ser dito para o Brasil, onde não se encontrou pesquisas nesse sentido. Assim, este trabalho irá se guiar na pesquisa que mais se destacou com resultado de estudos no exterior [22]. Nela se revisou 57 estudos primários de retorno, realizados em países desenvolvidos incluindo Estados Unidos da América (EUA) (58% dos estudos), quatro países da Europa Ocidental (Países Baixos, Finlândia, Dinamarca e Reino Unido, com um total de 22%), Canadá (15%), Japão (5%) e Austrália (1 estudo). Em termos de retorno, metade dos estudos envolvem retorno *indireto*, dentre os quais 11 envolvem Faturamento Aprimorado, três estudos envolvem o uso de Retorno Estimado e 15 estudos consideram Retorno Diário/Semanal. Os remanescentes envolvem retorno *direto*, dos quais 23 exploram retorno agregado e outros seis estudos onde são fornecidos a informação em tempo

real no nível dos equipamentos. Salienta-se que pesquisas nesse sentido no Brasil são necessárias para validar os resultados, tomando em posse as diferenças culturais, sociais e econômicas quanto a nossa realidade.

Um resumo dos resultados obtidos nessa referência, para os estudos realizados entre 1995 e 2010 (cerca de dois terços dos revisados), estão na Figura 3.4. A redução do consumo mostrada leva em conta os resultados globais, ou seja, considerando a taxa esperada de adesão da população aos programas de EE, supondo que os mesmos terão participação voluntária.

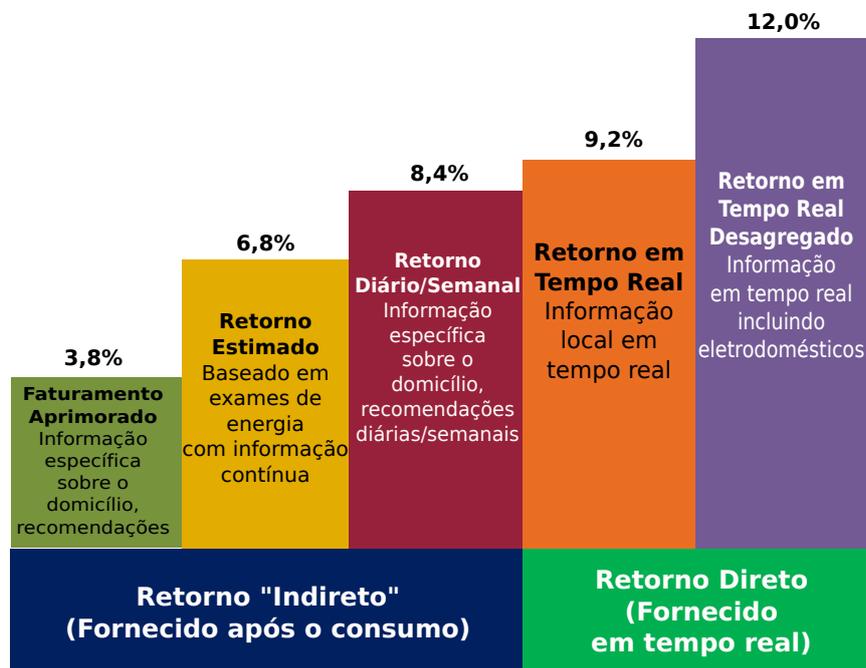
Identifica-se nos resultados que os tipos de retorno são tanto incrementais em custo e complexidade, quanto nos resultados de economia energética. Assim, é natural a implementação do sistema de retorno ser realizada de maneira contínua, aplicando sistemas já disponíveis enquanto se realiza investimento em tecnologia para o desenvolvimento do próximo nível de informação. Os desenvolvedores devem manter o sistema o mais flexível possível, sendo desenvolvidos sempre preparados para a mudança e considerando o surgimento de novos mecanismos de retorno com o avanço da tecnologia [22]. Por exemplo, por ora é possível obter economia de energia utilizando um sistema de baixo custo, como o Faturamento Aprimora, que informa melhor o consumidor, ou até mesmo, com mais ambição, fornecer o Retorno Diário/Semanal.

3.2.3 Indo Além dos Resultados

Ademais, outras considerações devem ser levadas quando no desenvolvimento de programas ou estudos de EE através do retorno do consumo de energia [22, 23, 26]:

- **Retorno Indireto versus Direto:** Os retornos indiretos são mais adequados que diretos para demonstrar o efeito de mudanças no condicionamento do ambiente, composição domiciliar e o impacto de investimentos em medidas de eficiência ou equipamentos de alto consumo. Já o retorno instantâneo se adéqua, geralmente, no fornecimento do impacto do consumo de equipamentos com usos de energia menores;
- **Participação Voluntária:** Programas nos quais os usuários tem de optar por não participar (*opt-out*) tiveram adesão significadamente maior (75%-85%) do que aqueles nos quais os usuários escolhem em colaborar (*opt-in*, participações menores a 10%), sendo assim recomendada a primeira abordagem para maximizar a participação dos consumidores;
- **Elementos Motivacionais:** A utilização de outros elementos para motivar a população aquém do financeiro e apelo ao meio ambiente mostram-se importantes para aumentar a eficiência dos programas de EE. São citados como

**Economia No Consumo Elétrico Médio Residencial (4-12%)
por tipo de retorno**



Baseado em 36 estudos ocorridos entre 1995-2010

Figura 3.4: O potencial de economia de consumo para cada tipo de retorno. Estudos no exterior e em países desenvolvidos. Adaptado de [22].

exemplo criar metas, compromissos, competições e normas sociais (tanto descritivas quanto injuntivas). A Subseção 3.2.5 irá tratar do tema com mais detalhes;

- **Contexto Regional:** Diferenças regionais e culturais afetam os resultados. Os resultados para a Europa Ocidental superam os obtidos nos EUA, podendo ser atribuídos as diferenças em como o discurso sobre as mudanças climáticas pelas lideranças políticas nas duas regiões é feito e assim a preocupação com o tema pela população. Nesse caso, chama-se atenção novamente às estratégias antecedentes no intuito de preparar a população para os programas e maximizar os resultados. Outro aspecto importante é a necessidade de estudos sobre o tema a fim de especificar como o brasileiro irá reagir em tais programas;
- **Duração do Estudo e Persistência dos Resultados:** Quando os estudos são de menor duração (< 6 meses) se obtém resultados mais eficientes (média de 10,1% de economia) que estudos mais longos (7,7%), discrepância essa atribuída a inaptidão de estudos curtos em observar variações sazonais na utilização de energia. Alguns estudos indicam que se faz necessário a pre-

sença do retorno em longo termo para que os resultados persistam, enquanto outros apontam a necessidade do retorno continuamente, enfatizando assim a necessidade na extensão dos programas de EE;

- **Tamanho do Estudo:** Estudos com grandes (> 100) amostragens domiciliares tendem a ter resultados mais modestos. Como esses estudos tem uma representatividade melhor das residencias, isso indica que programas de EE em larga escala também devem apresentar resultados mais modestos que aqueles apresentados na Subseção 3.2.2. Ainda, esses estudos mostram-se menos suscetíveis às oscilações quanto a duração dos estudos;
- **Resposta de Ponta e Demanda versus Economia Fora de Ponta:** Reduções de pico e demanda são de particular interesse das concessionárias que buscam atender essencialmente o mesmo nível de serviços mas com custos totais menores. Há dois modos de obter tal efeito: com uma melhoria em EE ou através do deslocamento de parte do consumo no horário de ponta para fora da ponta. O interesse em resposta de demanda, ou seja, em reduzir o consumo durante os horários de ponta difere dos programas de EE que focam em ter reduções eficientes economicamente durante todo o ano. Ainda que não seja desprezível, programas de resposta de demanda apresentam economia de energia bastante baixos quando em comparação aos de EE. Além disso, esses programas também são importantes para a questão ambiental, considerando que essa energia normalmente será proveniente de fontes fósseis. Por outro lado, os consumidores normalmente não percebem a diferença entre os dois programas, do mesmo modo que a integração dos programas é plausível e sinérgica, onde estudos mostram que a junção causa melhores resultados tanto em economia de energia quanto na redução de picos, por isso, sendo interessantes tanto do lado do consumidor quanto para a concessionária. Desta forma, a abordagem ótima ao tema deve ser conseguir todos os meios economicamente atraentes de reduzir o desperdício e ineficiências antes de procurar oportunidades restantes de reduzir cargas durante os picos;
- **Hábitos, Escolhas e Estilos de Vida:** Dentre os tipos de comportamentos de EE e conservação, os que aparecem mais frequentemente são investimentos em novos equipamentos e equipamentos em populações mais ricas, sendo geralmente empreendido em conjunto com mudança de residência ou melhoria no estilo (referido em oposição a funcional) do domicílio. Para a maioria da população, os domicílios obtém melhor EE através da mudança de hábitos e rotina, ou pela avaliação dos comportamentos relacionados a energia. Esses comportamentos de EE são motivados assim por uma variedade de fatores, incluindo interesse próprio (financeiro) e outros motivos altruístas e

preocupações cívicas. Desta forma, programas de EE que procuram apenas a instalação de equipamentos mais novos e eficientes irão desperdiçar o potencial relacionado à mudança comportamental, assim como programas que apelam apenas para o interesse financeiro não irão influenciar um largo grupo de fatores que motivam as pessoas para agir;

- **Segmentação Populacional:** Poucas pesquisas exploraram como o potencial de redução de consumo é afetado pelas diferentes classes sociais. Desses estudos, as descobertas sugerem grandes níveis de economia tendem estar associados a alto nível educacional e renda, grandes residenciais e dentre elas as com maior número de pessoas, consumidores jovens e/ou com grande tendência a valores ambientais.

3.2.4 Tecnologias e Tendências

Como constatado, os medidores atualmente utilizados pelas concessionárias, os medidores analógicos e eletrônicos, permitem fornecer um retorno de baixo custo mas com um potencial melhor de economia de energia, o Faturamento Aprimorado. As contas de energia de companhias como a Light, Ampla, Cemig e Eletropaulo fornecem o histórico de consumo dos últimos 12 meses, uma informação que pode auxiliar o consumidor, já podendo ser consideradas um Faturamento Aprimorado. Entretanto outras informações podem ser utilizadas, como referências do consumo acumulado e, em especial, comparações do consumo com o de vizinhos ou grupo pertencente. Também é possível estimar o uso energético por uso-final utilizando os valores médios de consumo das residenciais para cada uso no sentido de auxiliar o cliente. A ideia é transformar a conta de energia em uma espécie de relatório do consumo energético, com um visual mais atraente (ver Subseção 3.2.6), contendo gráficos e informações no sentido de atrair o consumidor a se preocupar com o tema, sendo esse o primeiro passo [23].

No entanto, o sistema elétrico atual está se tornando obsoleto para atender aos problemas de aumento de carga nos centros urbanos devido ao crescimento do setor de serviços e do consumo das residências. Há uma presença cada vez maior de cargas eletrônicas injetando harmônicos e a geração centralizada exige excessivamente da capacidade de transmissão e distribuição, sobrecarregando as linhas nesses grandes centros que nem sempre podem corresponder à necessidade de novas linhas. A falta de informações sobre o estado do sistema dificultam a operação e planejamento de uma rede cada vez mais sobrecarregada. As ICT revolucionaram as redes de telecomunicações e serão a tendência para a criação das redes elétricas inteligentes (*smart grids*), o novo sistema elétrico que tem como objetivo responder a essas dificuldades. Ainda não foram definidas todas as características desse sistema, contudo, as

principais características são o uso de comunicações em tempo real para o controle e informação, o uso massivo de sensores e medidores para monitoramento do sistema, faturamento com preços para o momento de uso, gestão pelo lado da demanda, a integração de componentes avançados como linhas de transmissão supercondutoras, armazenamento de energia, eletrônica de potência, geração distribuída etc. [22, 32]

Os equipamentos eletrônicos de medição utilizados nas redes inteligentes, referidos neste trabalho como medidores inteligentes (*advanced/smart metering*), irão fornecer uma gama maior de informações em tempo-real para as concessionárias, melhorando a operação e planejamento. Ao mesmo tempo, será possível a concessionária comunicar-se com o cliente, oferecendo incentivos (como descontos) para reduções de carga durante os horários de ponta, outros planos de tarifação com preços dinâmicos de acordo com os horários, aumentando a interação da concessionária com o cliente. Essas especificações estão regulamentadas na Resolução Normativa nº 502 da ANEEL [33].

Por outro lado, pelo ponto de vista da demanda (ou dos consumidores) essa informação também estará disponível, trazendo uma gama de novas oportunidades para os usuários participarem ativamente. Mais especificamente, na abordagem do tema atual, os medidores inteligentes oferecem uma base a ser explorada para fornecer o retorno em larga escala para os consumidores, tanto o direto quanto indireto. Nos medidores utilizados nos EUA foram apontadas algumas dificuldades técnicas para esse fornecimento, sendo elas: a necessidade de grande quantidade de energia para enviar um sinal frequente ao consumidor e a atender a necessidade do sinal ser enviado em intervalos frequentes, atendendo a taxa escolhida de 7 s. Com um custo adicional, um estudo na indústria estadunidense mostrou que é possível dos medidores terem seu *hardware* substituídos no futuro para que possam fornecer medições de pequena energia e *chips* de comunicação para habilitar dados de equipamentos específicos, assim como automação para grandes cargas, como unidades de condicionamento ambiental, bombas etc. Desta forma, sendo possível fornecer tanto a tecnologia para o Retorno em Tempo Real com a utilização de *displays* dentro do domicílio, quanto o Retorno em Tempo Real Desagregado [22].

Algumas empresas se estabeleceram no novo mercado para informar o consumidor sobre o seu uso de energia e em auxiliá-lo nas atitudes para reduzir o seu consumo, antes mesmo que estivessem disponíveis os medidores inteligentes. Elas fornecem o retorno indireto em alguns países desenvolvidos, dentre eles o EUA, Austrália, Nova Zelândia, Reino Unido. Dentre essas empresas, faz-se referência a *Positive Energy* [34] e *C3 Energy* [35] que disponibilizam seus serviços, organizados na Tabela 3.1, utilizando os dados da concessionária, independente quando presentes na residência os medidores convencionais ou inteligentes. É importante notar que as abordagens utilizadas por empresas nesse ramo utilizarão análises mais complexas

conforme a presença de dados mais detalhados, frequentes e desagregados estejam disponíveis.

Empresa	Tecnologia de Retorno	Princípios Comportamentais
Positive Energy [34]	Dependendo da concessionária, envia correspondências mensais ou trimestrais e/ou fornecem um portal na internet com novas redes sociais	<i>Tipo de Retorno:</i> Retorno indireto incluindo informação sobre o domicílio e conselhos, auditorias de energia através do uso da <i>web</i> , análise de faturamento, consumo estimado por equipamento, gás carbônico (CO_2), kWh e \$. <i>Princípios Comportamentais:</i> Comparações sociais, metas, comparações pessoais e plano de ações.
C3 Energy [35]	Portal de comunidade social com retorno de consumo de energia e água	<i>Tipo de Retorno:</i> Retorno indireto incluindo informação sobre o domicílio e conselhos, auditorias de energia através do uso da <i>web</i> , análise de faturamento, consumo estimado por equipamento, CO_2 , kWh, \$ e outras unidades. <i>Princípios Comportamentais:</i> Comparações sociais, metas, competições redes sociais, comparações pessoais e plano de ações.

Tabela 3.1: Empresas utilizando informação da concessionária e as oportunidades e isentivo de economia de energia oferecidas. Extraído e atualizado de [22, tradução própria].

Já o retorno direto pode ser encontrado através de *displays* de energia no domicílio. A Tabela 3.2 identifica alguns dos *displays* oferecidos atualmente e suas propriedades. Muitas vezes as companhias oferecem também análises e estimativas do consumo específico de equipamentos, comparações sociais e outros princípios para motivar os consumidores a economizar energia. A informação de consumo de equipamentos específico ou é estimada, ou realizada através de sensores nos equipamentos.

Uma outra maneira para fornecer o Retorno em Tempo Real Desagregado é através do NILM (Capítulo 4). Essa técnica — ainda em desenvolvimento, mas de crescente interesse no mundo devido a todas questões aqui levantadas, com ofensivas pesquisas no tema por grandes empresas como *Intel* e *Belkin* — coloca o peso da desagregação da informação no *software*, reduzindo a necessidade de investimento em sensores e *hardware*, sendo assim um método potencialmente mais favorável economicamente para a implementação de programas de EE fornecendo esse tipo de retorno, no entanto isso irá depender de suas limitações e eficácia. No Brasil, os esforços da ANEEL em regulamentar as bases para os novos medidores inteligentes dão poder ao consumidor de exigir à concessionária acesso às medições de tensão e corrente de cada fase, como rege no art. 3º da Resolução Normativa nº 502 [33]. Ainda não se especificou a taxa de amostragem na qual essas leituras serão dispo-

	The Energy Detector TED [36]	Wattson [37]
Descrição da Tecnologia	<i>Software</i> de suporte, aplicativos para celular	<i>Software</i> de suporte com acesso a comunidades
Mecanismos de Retorno	Mostradores em tempo real de kW, \$/hr, CO_2 , consumo e gastos diários, conta estimada em kWh e \$, pico de consumo, voltagem min/max e custo/demanda projetada	Mostradores em tempo real aproximado do consumo em W, kW, conta estimada. Leituras entre 3 a 20 s. Brilha conforma o consumo: azul para consumo baixo; roxo para médio; vermelho para alto.
<i>Retorno de Informação:</i>		
Princípios Comportamentais	Retorno direto incluindo conselhos, auditorias de energia baseadas na <i>web</i> , análise do consumo, estimativa de consumo por equipamentos, CO_2 e \$.	
	<i>Motivações Oferecidas:</i> Comparações sociais, metas, comparações pessoais e etapas de ações.	
	PowerCost Monitor [38]	Efergy Elite [39]
Descrição da Tecnologia	<i>Software</i> de suporte, aplicativos para celular	<i>Software</i> de suporte, aplicativos para celular
Mecanismos de Retorno	Mostradores em tempo real aproximado do consumo em kW e \$/hr, pico de consumo nas últimas 24 horas, contagem de k/Wh (reiniciável), recurso para medição de equipamentos específicos.	Mostradores em tempo real aproximado do consumo em kW e \$/hora (leituras em 6, 12 ou 18 s), informação de consumo média por hora, semanal, mensal. Alarmes para consumo alto.
<i>Retorno de Informação:</i>		
Princípios Comportamentais	Retorno direto incluindo conselhos, auditorias de energia baseadas na <i>web</i> , análise do consumo, estimativa de consumo por equipamentos, CO_2 e \$.	Retorno direto, análise de consumo, estimativa de consumo em \$.
	<i>Motivações Oferecidas:</i> Metas e comparações pessoais	

Tabela 3.2: Especificações de *displays* domiciliares disponíveis. Extraído de [22, tradução própria].

nibilizadas, por sua vez, pode-se utilizar toda a infra-estrutura das coleitas de medidas e comunicação oferecida pelos medidores inteligentes no intuito de maximizar o custo-benefício. Caso a amostragem deles seja baixa, ou de interesse aumentá-la para obter uma maior capacidade de identificação dos equipamentos, o NILM pode fazer mão de um *hardware* próprio de medição.

Percebe-se que ainda é incerto se os medidores inteligentes são a melhor alternativa para fornecer retorno de informação, contudo parece natural sua utilização. Diversas tecnologias podem ser utilizadas envolvendo, ou não, as concessionárias. Apenas com o desenvolvimento dessas tecnologias será possível determinar as limitações, custos e vantagens para definir o que é economicamente mais atraente.

Finalmente, uma outra tendência é o uso de automação da rede doméstica. A automação, além de melhorar a qualidade de vida dos consumidores, pode aumentar o potencial de redução de consumo. Com uma maior capacidade de administração de sua demanda sem grande esforço, facilita-se aos consumidores de realizarem a mudança de hábitos no sentido de um comportamento sustentável, simplificando a

conduta do sistema elétrico de um modo mais econômico.

3.2.5 Aspectos Psicológicos

A tecnologia apenas concebe as possibilidades de informação a serem repassadas ao consumidor, no entanto, a questão ainda está em como apresentar essa informação e motivar o usuário para a mudança. O conselho de profissionais nos campos de psicologia, sociologia, *marketing*, mudança e economia comportamental serão críticos para motivar, habilitar e continuamente empreender consumidores na gestão de sistemas de energia residenciais [22]. Um exemplo é a empresa *Positive Energy*, que utiliza psicólogos para auxiliar no desenvolvimento de suas tecnologias, disponibilizando ferramentas sociais e estratégias persuasivas para um engajamento maior de seus clientes.

Um levantamento das noções básicas da psicologia motivacional foi realizada em [30], assim como uma estrutura para os desenvolvedores da tecnologia aplicada nos programas de retorno de informação no sentido de motivar a mudança comportamental para uma melhor EE e um mundo sustentável. Será realizado um resumo desses tópicos a seguir guiado nessa referência, mas vale enfatizar que os profissionais nessas áreas devem analisar o tema e escolher a melhor abordagem a ser utilizada nas tecnologias desenvolvidas.

O objetivo é motivar o consumidor para a mudança através de recomendações levando em conta o processo de mudança comportamental do consumidor. Define-se a motivação como [30, pp. 927-928, tradução própria]:

Motivação é um questionamento ao porquê do comportamento. Ela é um estado interno ou condição (as vezes descrita como uma necessidade, desejo ou querer) que serve para ativar ou energizar o comportamento. Motivação está fortemente ligada a processos emocionais. Emoções podem estar envolvidas na iniciação comportamental (como a emoção de solidão pode motivar a ação de procurar companhia). Em alternativa, o desejo para viver uma emoção em particular pode também motivar para a ação (como a decisão de correr uma maratona pode ser motivada pelo desejo de experimentar a sensação de realização de um feito).

Ela é influenciada por ideais psicológicos que foram aprendidos pelos indivíduos. Nota-se que diferentes indivíduos tem ideais psicológicos distintos, estes estando apresentados em ordem decrescente quanto a possibilidade de sofrerem alteração:

- *Atitudes* são predisposições aprendidas quanto a respostas para uma pessoa, objeto ou ideia em um modo favorável ou desfavorável. Por exemplo, o ato

de tomar banho curtos devido a uma atitude favorável em respeito ao meio ambiente;

- *Crenças* são os meios nos quais as pessoas estruturam seu entendimento da realidade, refletindo a ideia do que é certo e o que é errado. A maioria das crenças são baseadas em experiências passadas, como a reciclagem faz bem ao meio ambiente.
- *Valores* são os fundamentos para o conceito de um indivíduo sobre si mesmo. Podem ser conceituadas como ideais comportamentais ou preferências por vivências. No caso dos primeiros, valores funcionam como conceitos duradouros de bem e mal, certo e errado, enquanto para preferências por vivências os valores guiam indivíduos para vivenciarem situações nas quais as proporcionam certos tipos de emoções. A Tabela 3.3 contém um subgrupo de valores definidos por Rokeach e Maslow, onde se propõe que pessoas possuem uma estrutura hierárquica ou prioritária de valores individuais. Rokeach acredita que as diferenças no comportamento ocorrem devido a diferenças na classificação de importância de valores, enquanto Maslow fornece uma ordem de valores em níveis que serão priorizados pelos indivíduos pelos níveis mais baixos antes dos níveis superiores.

Ideais Comportamentais (Rokeach)	Preferências por Experiências (Rokeach)	Preferências por Experiências (Maslow, níveis em ordem crescente)
Capaz Competente, eficiente. Prestativo Trabalhando para o bem-estar dos outros. Honestidade Sinceridade e confiável. Imaginativo Ousado e criativo. Independente Autoconfiante, autossuficiente. Intelectual Inteligente e pensativo. Lógico Consistente e racional. Obediente Atencioso e respeitoso. Responsável Fidedigno e confiável.	Vida Confortável Uma vida próspera. Liberdade Independência e liberdade de escolha. Saúde Bem-estar psicológico e físico. Harmônia Interna Ausência de conflitos interiores. Sentimento de Realização Uma contribuição duradoura. Reconhecimento Social Respeito e admiração. Sabedoria Um entendimento maduro da vida. Um Mundo Belo Beleza da natureza e das artes.	Psicológico Homeostase e apetites. Segurança Segurança do corpo, emprego, recursos, família, saúde, propriedade. Amor/Aceitação Afeição e ser aceito. Estima Respeito próprio, autoestima, estima dos outros Realização Pessoal Encontrar satisfação pessoal e compreender seu potencial.

Tabela 3.3: Valores propostos por Rokeach e Maslow. Tradução própria de [30, pp. 928].

Outra questão importante é a persistência ou durabilidade do comportamento, ou seja, da capacidade do comportamento manter-se, sem a necessidade de intervenções. Para atingir esse objetivo é aconselhável motivação intrínseca, que é a

realização de uma atividade pelas satisfações que a tarefa oferece, enquanto o seu oposto, a motivação extrínseca, está ligada à realização para obter uma consequência separável. Exemplos para o primeiro são: curiosidade, competência e satisfação; no outro caso: incentivos materiais e reconhecimento social.

Utilizou-se nessa referência o Modelo Transteórico³, onde se considera que o processo da mudança ocorre em uma série de estados, sendo a motivação a força motriz para se deslocar entre os estágios. O objetivo em cada etapa, assim como recomendações para o atingir estão a seguir:

- Estrutura Exemplo — *Etapa*: explicação.
 1. Objetivo nessa etapa.
 - (a) Recomendação para atingir o objetivo.
- *Pré-contemplação*: o indivíduo está desencorajado, relutante em manter atitudes em prol da mudança, mal-informado ou desconhece o problema comportamental. Não há previsão de ação no futuro, esse medido normalmente como os próximos 6 meses.
 1. Apresentar informação em moderação com o propósito de plantar a semente no sentido dos indivíduos tomarem conhecimento de seus comportamentos (de consumo de energia) atuais problemáticos. A moderação é importante pois uma maior intensidade irá, geralmente, produzir menores resultados nessa etapa.
 - (a) Fornecer retorno de informação personalizado notando tanto os prós e contras de um comportamento *não-sustentável*. Apresentar os benefícios e consequências em relação aos valores pessoais, de modo não tendencioso.
 - (b) Utilizar normas sociais a respeito de comportamentos de consumo sustentáveis, combinando o uso de normas injuntivas e descritivas. As normas sociais são as regras ou expectativas por um comportamento apropriado em uma situação social em particular. Elas “podem levar pessoas a dizer coisas que sabem que não são verdade, utilizar drogas ilícitas ou deixar de reagir a uma ameaça eminente” [22, pp. 51, tradução própria]. Normas descritivas são percepções de comportamentos normalmente realizados (ex. 85% da sua vizinhança reciclam), apelando ao valor de Maslow de *Amor/Aceitação*.

³A referência levanta outros modelos comumente utilizados e seus prós e contras. Também há uma preocupação quanto a modelar a mudança através de estados discretos, que podem não representar bem a realidade do processo. A utilização do modelo é justificada por seu valor heurístico, usado como um modelo simplificado de uma mudança ideal.

Normas injuntivas são percepções e comportamentos que são normalmente aceitos ou aprovados (ex. um sinal de polegar positivo com o texto “Gere menos resíduos”). Essas apelam para o valor de Rokeach *Obediente*. Ao juntar ambas normas descritivas e injuntivas, há uma chance ainda maior de sucesso quando em comparação da aplicação delas isoladamente.

- (c) Fornecer uma variedade de pequenas ações de consumo que, se realizadas, podem ter impactos positivos no meio ambiente. Isso irá trabalhar em duas barreiras para a motivação, que são: não se sentir competente e não acreditar que suas ações irão levar a um resultado positivo. Apresentar uma variedade de ações apela ao valor de Rokeach de *Liberdade* e aumenta o senso de controle pessoal assim como a motivação intrínseca.
- *Contemplação*: há conhecimento de seu problema comportamental e planeja-se uma mudança no futuro. Contempladores são abertos a informação sobre o problema, ainda que estão distantes do compromisso de mudança devido ao sentimento de ambivalência;
 1. Pender a balança no sentido de mudança. Ambivalência⁴ é o problema chave que precisa ser resolvido, uma vez que a avaliação dos prós e contras tem mais ou menos o mesmo peso.
 - (a) Nessa etapa, deve-se fornecer o retorno de informação enfatizando os prós de um comportamento sustentável e os contras do comportamento não-sustentável. É importante auxiliar o consumidor em perceber os prós, uma vez que os mesmos irão resistir às mudanças enquanto perceberem isso como um fator redutor de sua qualidade de vida, em especial àquelas que salientam o sacrifício pessoal pelo bem comum. Os contras devem enfatizar nos custos de comportamentos não-sustentáveis, numa perspectiva de perda em detrimento de ganho. O foco nos valores pessoais podem ser extremamente persuasivos nesse estágio.
 - (b) Lembrar o indivíduo de sua atitude em favor do meio ambiente, informá-lo da discrepância de suas atitudes e o comportamento correspondente, encorajar a mudança. Essa técnica apela para a dissonância cognitiva⁵. Como as pessoas mudam de atitudes mais fácil que de comportamento, é importante encorajar o comportamento

⁴Presença de sentimentos/pensamentos conflitantes perante uma coisa ou pessoa.

⁵Um estado desconfortável que ocorre quando a pessoa possui uma atitude e um comportamento que são psicologicamente inconsistentes.

sustentável.

- (c) Fornecer incentivo para pequenas mudanças de consumo (independente da intenção do consumidor original era a utilização sustentável de energia) para fomentar maiores mudanças no futuro.
 - (d) Vincular a tecnologia de retorno a um portal de uma comunidade virtual e incentivar o indivíduo para procurar e ler a informação de experiências de outros usuários com consumos sustentáveis na comunidade. Isso apela para normas sociais de um modo vivo e personalizado, explorando a abertura dos contempladores ao tema, mas, ao mesmo tempo, sem forçar nenhum tipo de ação.
- *Preparação*: momento em que o indivíduo está pronto para ação no futuro iminente (medido normalmente como 1 mês), e tem como objetivo desenvolver e engajar-se a um plano. Pelo menos uma tentativa de mudança foi realizada no último ano;
 - 1. Ajudar os usuários a desenvolver um plano que seja aceitável, acessível e efetivo. Esses planos podem se relacionar a ações extraordinárias (compra de um refrigerador eficiente) ou diárias (tomar banhos mais curtos). Um objetivo é definido como uma representação interna de um resultado desejado. Usuários na fase de preparação podem ter objetivos abstratos mas não saber necessariamente como os alcançar.
 - (a) Ajudar na criação de metas pessoais específicas e quantitativas (o nível de dificuldade deve se elevar conforme o sucesso, começando com metas simples e partindo para mais complicadas). Metas difíceis, pessoais e específicas tem maior engajamento quando comparadas a tarefas “faça o seu melhor”, fáceis ou que lhe foram atribuídas.
 - (b) Desenvolver modos múltiplos para os consumidores atingirem suas metas e incentivá-los para utilizar sua habilidade e experiência pessoal nesses planos.
 - (c) Fornecer no portal da comunidade virtual aos usuários a opção de ter um “Conselheiro/Tutor”. Os conselheiros seriam pessoas exemplares na fase de ação ou manutenção. Essa conexão fornece um maior nível de engajamento.
 - *Ação*: ocorre a manifestação de modo evidente da mudança comportamental, usualmente dentro dos últimos 6 meses;
 - 1. Reforçar positivamente ações sustentáveis de energia. O reforço positivo é a técnica mais efetiva para motivar a maior ocorrência de um

comportamento desejado, ela tende a aumentar a motivação intrínseca. Técnicas como a punição ou reforço negativo evitam um comportamento não desejado, mas não o substitui por nada, além de reduzir a motivação intrínseca.

- (a) Fornecer o reforço positivo imediatamente após o comportamento desejado ocorrer e em múltiplos modos com o intuito de aumentar sua eficácia.
2. Desenvolver motivações intrínsecas para o comportamento sustentável.
 - (a) Permitir uma exploração interativa, personalizável e anotações na interface oferecida.
- *Manutenção, Recaída e Reciclagem*: trabalho no sentido de manter o comportamento alterado e a luta para prevenir recaídas. Se aquela ocorrer, o indivíduo regressa a um dos estágios anteriores e o processo recomeça.
 1. Manter o comportamento sustentável permanente. Em algum momento as mudanças irão se tornar sustentáveis por si próprias, sendo possível a saída do indivíduo do ciclo de mudança. Enquanto isso, o objetivo deve ser fazer o indivíduo apenas um pouco mais consciente e informado.
 - (a) Apoiar ações sustentáveis para que elas virem hábitos, lembrando os usuários para realizar ações específicas.
 - (b) Fornecer a opção para usuários nessa etapa de se tornarem “Conselheiros/Tutores” para indivíduos na etapa de preparação. Essa técnica aplica dissonância cognitiva, uma vez que indivíduos que tentaram persuadir alguém irão racionalizar internamente o seu comportamento, e assim estarão propensos a intensificar seu engajamento. Esse método apela para o valor de Rokeach *Reconhecimento Pessoal e Sabedoria*, e, em retorno, pode gerar satisfações intrínsecas de competência e satisfação.
 - (c) Encorajar os usuários para reforço e reflexões pessoais em suas experiências através de um diário. A reflexão de suas atitudes em relação a energia e percepção de seu progresso podem trazer a tona satisfações intrínsecas de interesse, competência e satisfação. O reforço pessoal (na forma de orgulho ou senso de realização) também irá trazer satisfações intrínsecas, no caso de competência, e ainda levar a percepções de auto-eficácia. Isso é importante pois, para um indivíduo vivenciar sucesso de longo-termo, eles precisam de auto-eficácia e atribuições intrínsecas do comportamento.

- (d) Manter um ciclo de motivação intrínseca de interesse, curiosidade, desafio ótimo, competência e satisfação. A motivação intrínseca é um ciclo de dois passos. Primeiro, estímulos como novidade, complexidade e mudança atraem a atenção, curiosidade e interesse, o que convida para a investigação, exploração e manipulação dos estímulos. Segundo, desempenhos de competência em tarefas são desfrutados, enquanto o crescimento da satisfação aumenta tanto o desejo na atividade quanto a capacidade de confrontar desafios parecidos no futuro.

3.2.6 Aspectos Visuais da Informação

A apresentação na qual a informação é realizada também irá afetar o entendimento e engajamento do consumidor nos programas de EE. Novamente, os especialistas e profissionais da área, no caso de *design* e da informação, serão importantes nesse intuito, frisando que o problema jaz além de apenas disponibilizar a tecnologia para fornecer a informação: também será necessário tornar a informação atrativa ao consumidor.

O processo do fluxo de informação ocorre entre o remetente e o receptor, onde este analisa a informação necessária e a apresenta na mensagem através de textos e imagens, enquanto aquele realiza escolhas entre as informações disponíveis nela e opta se irá processá-las mentalmente. A visão é o sentimento mais importante para a compreensão e vivência humana do meio externo. Cerca de 70% de nossas células sensoriais estão nos olhos, assim, a visualização é um meio bastante efetivo de realizar a comunicação da informação e de dados. Tornar a mensagem interessante visualmente depende de vários elementos [40]. Alguns exemplos são: fornecer para a mensagem uma estrutura na qual serão guiados os princípios para a inserção das representações; clareza para a simplicidade de compreensão e legibilidade da informação; ênfase para atrair, direcionar ou manter a atenção; unidade da mensagem, com uma coerência e união global [41].

Em [40], apresentou-se a jovens de um colégio sueco uma visualização do consumo de energia em um *display* em tempo real na sala de aula. Os jovens tiveram participação na criação da interface⁶, onde desenharam e escolheram uma imagem para informar se o consumo estava elevado, médio ou baixo, assim como informaram o que entenderam das informações nos *displays*. Esse estudo revelou cinco aspectos importantes no desenvolvimento de sua visualização nos *displays*:

⁶A participação dos usuários no desenvolvimento facilita à tecnologia atender às especificações e necessidades dos mesmos. No caso, a participação dos usuários utilizada está de acordo com a recomendação de [42], na qual há a descrição de uma abordagem orientada ao usuário para o retorno do consumo de energia. Nela, o usuário apenas fornece inspiração aos desenvolvedores profissionais, que desenvolvem as soluções.

1. Deve chamar a atenção dos usuários, realizando o uso de cores brilhantes, contrastes e quando possível uma exibição dinâmica;
2. Mostrar comparações entre o consumo de modo a deixar evidente aos usuários resultados positivos de um esforço;
3. Fornecer o consumo em tempo real para estimular a mudança direta no comportamento;
4. Deve conter um tom positivo e encorajador, potencializando a positividade de um comportamento correto;
5. Ser explicativa, com pequenos textos instrutivos que fazem aos usuários simples de entender o que eles estão vendo.

Capítulo 4

Monitoração Não-Invasiva de Cargas Elétricas (NILM)

Este capítulo trata as particularidades envolvidas no desenvolvimento da tecnologia conhecida como NILM¹. Diferente dos capítulos anteriores, que se dedicaram à contextualizar as aplicações do NILM em torno do tema de EE, este capítulo se dedica à descrição do projeto do NILM em termos técnicos e das abordagens utilizadas observadas durante o levantamento bibliográfico realizado para este trabalho, bem como o estado do projeto e a metodologia proposta pelo CEPEL/Eletronbras que foi expandida e sistematizada pelo trabalho atual, que será descrita nos dois próximos capítulos.

Antes, porém, o NILM é abordado trazendo seus aspectos gerais na Seção 4.1, aonde é realizado uma sinopse da informação dos capítulos anteriores e citadas outras possíveis aplicações para o NILM — suas aplicações não se limitam à EE. Em seguida, a Seção 4.2 apresenta as convenções adotadas por este trabalho, estando disponíveis os conceitos e nomenclaturas que serão utilizados em diante no trabalho. Essa seção apresenta, em seguida, um extenso levantamento de técnicas empregadas por demais autores para realizar a desagregação do consumo que será resumida e debatida em sua subseção final. Por sua vez, a metodologia proposta pelo CEPEL/Eletronbras está disponível na Seção 4.3, aonde também serão tratados os estudos anteriores realizados pela colaboração COPPE/UFRJ–CEPEL/Eletronbras e considerações sobre a escolha das características discriminantes, que utilizará como base os demais estudos de NILM abordados por este trabalho.

Apesar deste capítulo trazer as técnicas para o NILM e tratá-lo de uma maneira geral, vale ressaltar que a metodologia aplicada pelo trabalho se restringiu apenas para a detecção de eventos de transitório. O tratamento do NILM como um todo

¹NIALM e NALM são outras abreviaturas utilizadas na literatura com o mesmo significado àquela aqui utilizada, Monitoração Não-Invasiva de Cargas Elétricas, tradução própria do termo em inglês *Non-Intrusive Appliance Load Monitoring*.

foi realizado para elucidar e facilitar, da melhor maneira possível, o projeto dessa técnica, não se limitando a compreensão apenas do ponto atacado pela metodologia do trabalho, mas permitindo uma visão do projeto de uma maneira global.

4.1 Aspectos Gerais

O NILM é uma alternativa para fornecer a informação de consumo de energia elétrica desagregada por equipamento ou dispositivo. Ao invés das técnicas normalmente utilizadas — onde se lança mão de sensores dispostos em cada equipamento, esses enviando informações para uma central encarregada de decodificá-las e, assim, identificar os equipamentos que estão demandando consumo na rede —, o NILM é um método em que não ocorre a intrusão na propriedade do usuário (ou intrusão em escala mínima), contendo geralmente apenas um medidor central no fornecimento de energia dessa propriedade. O peso da identificação dos equipamentos é transferido da utilização de uma grande quantidade de sensores e *hardware* complexo para um *software* e algoritmos de processamento de sinais no intuito de realizar uma análise profunda das medições e desagregar as informações.

Ainda que o peso da identificação esteja no *software*, a aptidão dos algoritmos implementados dependem da capacidade do medidor de extrair informações da rede e das distorções causadas pelos equipamentos na mesma, de forma que quanto mais avançado for o *hardware* de medição, com capacidade de enviar uma quantidade e/ou frequência maior de informação aos algoritmos encarregados de realizar a discriminação dos rastros deixados na rede elétrica pelos equipamentos, também maior será a capacidade dos mesmos de desagregar o consumo específico dos equipamentos.

Possíveis aplicações para esse dispositivo são:

- auxiliar ou substituir as PPHs sem que seja necessário causar incômodo ao consumidor devido a presença de medidores na residência, fornecendo assim dados com maior fidelidade em carácter desagregado por equipamento e maior frequência para estudos de EE no consumo de eletricidade (Seção 2.3), em especial para o setor residencial;
- disponibilizar a informação desagregada para o fornecimento do Retorno Indireto Diário/Semanal e/ou Retorno em Tempo Real Desagregado em programa de EE. Os programas de EE utilizando retorno de informação são possíveis fontes de redução de consumo nos grandes centros urbanos. No entanto, são necessários estudos, no Brasil, para determinar seu potencial (Seção 3.2). O retorno de informação desagregada por equipamento pode ser utilizado por empresas que oferecem serviços de redução de consumo (Subseção 3.2.4), ou

mesmo por consumidores que precisam ter algum tipo de controle sobre seu consumo de energia;

- no interesse das concessionárias, a informação desagregada pode auxiliar os clientes a identificarem consumo não essencial durante o horário de ponta, auxiliando no deslocamento de carga (ver item *Resposta de Ponta e Demanda versus Economia Fora de Ponta* na Subseção 3.2.3), assim como identificar clientes com maior potencial de redução de consumo durante esses períodos, para oferecer incentivos nesse sentido (Subseção 3.2.4);
- identificar equipamentos defeituosos ou com consumo excessivo de energia;
- informar aos fabricantes de eletrodomésticos o perfil de seus consumidores (caso seja possível obter essa informação);
- permitir a identificação de perdas não-técnicas através de discrepâncias entre a informação de consumo agregado obtido pelo NILM e aquele obtido pela concessionária, o que indicaria incoerências na tarifação;

As possibilidades de aplicação como um meio de melhoria da EE e redução da intensidade elétrica tem elevado o interesse nesse assunto, em especial nos países desenvolvidos, como uma forma de aliviar a pressão de consumo nos grandes centros urbanos e na redução de emissões de CO_2 [43]. Isso tem levado a gigantes no setor eletrônico, como *Intel* e *Belkin*, a investirem fortemente no desenvolvimento dessa tecnologia. O crescente interesse na evolução por parte da academia levou a organização do primeiro *workshop* especificamente para o tema em 2012 [44]. Essa alternativa é mais simples quando comparando com a automação residencial por não requerer mudanças na produção dos eletrodomésticos — a automação requer comunicação nos dois sentidos (entre a interface e o equipamento), tal como a capacidade de controle do equipamento, de forma que se faz necessário adaptar equipamentos antigos e a produção dos novos equipamentos com essas capacidades — juntamente com o fato de haver um relutância social quanto às residências automatizadas, apesar de esforços governamentais e da mídia local [45] (a referência estudou a falta de interesse na automação residencial nos EUA). Não bastasse, também se pode citar o nascimento de *start-ups* procurando espaço na corrida por esse novo mercado, como *GetEmme* [46] e *Navetas* [47].

Porém, é importante aqueles envolvidos no projeto terem em mente a questão da ética: a informação do consumo não deve sair da residência sem o consentimento do consumidor. Além disso, o projeto deve ser realizado em um sistema autocondido, ou seja, deve ser garantido ao consumidor a segurança da informação e sua privacidade, de forma que esses itens também devem ser incluídos para o projeto e proliferação da técnica.

4.2 As Diversas Metodologias Utilizadas no Mundo

A ideia de desagregar a informação não é nova, sendo possível encontrar pesquisas nesse sentido datando da década de 1980. Apesar disso, um extenso levantamento bibliográfico [29, 48–77] realizado em Zeifman e Roth [43] expõe que as técnicas aplicadas em NILMs ou não são robustas no sentido de atenderem especificamente a um grupo limitado de equipamentos estudados, ou apresentam acurácia marginal, mostrando que o processo de desagregação da informação não é trivial. Demais referências que não constam em [43] citadas nesta seção contendo abordagens de NILMs são [78–92] (ver Subseção 4.2.4).

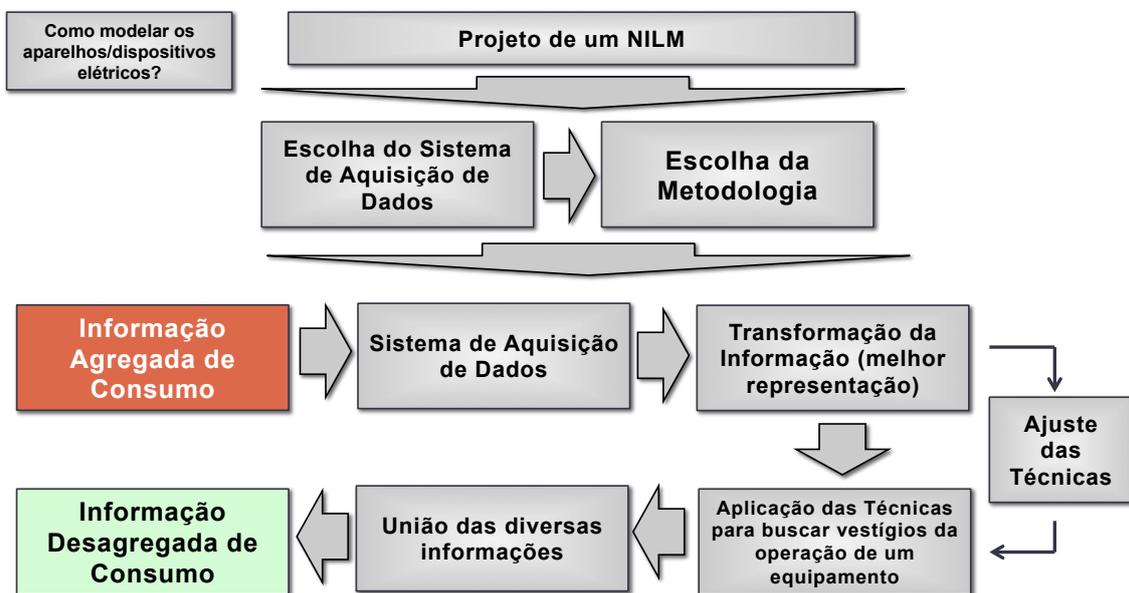


Figura 4.1: Esboço dos passos envolvidos no projeto do NILM e da sua operação.

Na Figura 4.1 há um esboço dos passos envolvidos no projeto do NILM e de sua operação, baseando-se nas diversas técnicas observadas nas referências citadas, apenas quando levando em conta o processo para a obtenção da informação desagregada do consumo. Vale ressaltar que a informação do consumo desagregado ainda precisa ser tratada levando em conta os aspectos psicológicos (Subseção 3.2.5) e visuais (Subseção 3.2.6) para motivar os consumidores a reduzirem seu consumo quando aplicando o NILM como um meio de economia de energia e intensificação da EE. Outro ponto do projeto para o projeto do NILM é a questão da segurança da informação e privacidade do consumidor, sendo necessário garantir que a mesma não deixará a residência sem seu consentimento. Dito isso, o projeto começa com a questão de como modelar os equipamentos e dispositivos presentes nas redes elétricas, sendo divididos de acordo com o comportamento de suas cargas, descritos na Subseção 4.2.1. Em seguida, Subseção 4.2.2, será tratado a escolha do sistema

de aquisição de dados e da metodologia a ser adotada. Em um projeto sem limitações, a metodologia escolhida implica na escolha do sistema de aquisição de dados a ser utilizado, porém, podem haver limitações no mesmo, nesse caso a metodologia tem de ser ajustada conforme as capacidades dele em fornecer informação. Com o crescente interesse pelo desenvolvimento do NILM, os autores perceberam a dificuldade de comparar as diferentes técnicas já presentes devido ao emprego de diferentes medidas para o cálculo de eficiência ou utilização de conjuntos com características diferentes entre si, de forma que o assunto merece atenção e será tratado na Subseção 4.2.3. Em seguida, a Subseção 4.2.4 irá apresentar as diversas abordagens utilizadas, sendo guiada na proposta de divisão das técnicas e características feita por [43]. Essa subseção contém levantamento bastante técnico, sendo de interesse somente daqueles que desejam se aprofundar no tema. A discussão das informações levantadas é realizada em 4.2.5, sendo sua leitura suficiente para aqueles que desejam as informações mais relevantes.

4.2.1 Modelos de Carga

Os equipamentos podem ser modelados devido às características de comportamento de suas cargas. A seguir, encontram-se possíveis características de carga elétrica dos eletrodomésticos. Os quatro primeiros itens são modelos de cargas mutuamente exclusivos, enquanto os dois seguintes podem ser incluídos, dependendo das propriedades dos equipamentos ligados à rede, para caracterizar equipamentos potencialmente dificultadores do processo de desagregação (baseado em [43, 48, 58, 73, 77, 78, 81]):

- **Equipamentos de consumo permanente (C1):** equipamentos que permanecem ligados 24 h/dia, 7 dias/semana, com consumo de energia praticamente constante. Ex.: detectores de fumaça, fontes de alimentação constantemente ligadas à rede;
- **FSM ou equipamentos de estados múltiplos (C2)²:** essa categoria é utilizada para identificar equipamentos que contém um conjunto definido de patamares discretos de consumo. Pode-se dividir esses equipamentos em dois conjuntos:

²Nas referências, não há separação na categoria C2, que é utilizada para descrever exclusivamente equipamentos modelados como C2a. Nelas, a categoria C2b é incluída em C4, uma vez que em ambos os casos as Máquinas de Estados Finitos (FSMs, em inglês) podem alterar o seu estado de operação para qualquer outro estado independente tanto de qual estado anterior ele estava operando como o tempo de operação nesse estado. A separação da FSM com estados discretos em duas categorias parece mais familiar do que incluir FSM com patamares discretos e aleatórios em uma categoria que permite patamares contínuos.

- **FSM ou equipamentos de estados múltiplos com ciclos bem-definidos (C2a)²**: os estados do equipamento repetem-se em padrões definidos temporalmente, garantindo que os seus ciclos serão observados frequentemente durante intervalos diários ou semanais. Ex.: máquina de lavar roupas, máquinas de lavar louças;
- **FSM ou equipamentos de estados múltiplos com ciclos aleatórios (C2b)²**: nesses equipamentos não há um padrão para os seus ciclos de operação. A operação por uma fonte externa, como o consumidor, altera o seu padrão de consumo sem ser possível encontrar uma regra operativa para o ciclo através da busca de repetições de suas trocas de estado na rede, os estados mudam aleatoriamente depois de quantidades de tempo também aleatórios. Ex.: ventilador de múltiplas velocidades, liquidificador — ambos dependendo de como operados pelo consumidor: se apenas ligados e desligados, irão comportar-se como C3s, enquanto se operados de modo padronizado, irão se comportar como C2as;
- **FSM de dois estados ou equipamentos liga/desliga (C3)**: um caso particular das C2s ocorre quando o equipamento pode ser modelado como tendo apenas dois estados: ligado/desligado. Ex.: lâmpadas, bombas de água;
- **Equipamentos com níveis contínuos de consumo (C4)**: uma generalização das C2s, onde há uma infinidade de estados para os quais o equipamento pode operar. Essa categoria pressupõe que o equipamento irá estabilizar o seu consumo em um patamar após um período de tempo. Sua operação pode ser dividida em dois grupos: operação manual do operador, equipamentos autocontrolados. Estes são de mais simples detecção quando comparados com aqueles, uma vez que seus ciclos de mudanças de estado são distribuídos uniformemente no tempo. Ainda assim, essa categoria é o maior desafio para as técnicas empregadas nos NILMs, sendo raramente tratada por elas. Ex.: lâmpadas com *dimmer*, ferramentas elétricas (furadeiras, serras etc.), bomba de aquário.
- **Equipamentos com dinâmica de carga (C5)³**: equipamentos que causam distúrbios na rede continuamente devido à dinâmica durante sua operação, se referindo a possíveis oscilações inerentes à característica operativa do mesmo. Alguns exemplos observados pela equipe do Centro de Pesquisas de Energia Elétrica (CEPEL/Eletrobras): televisores, onde a variabilidade de brilho, co-

³As referências optaram por não criar essas categorias uma vez que essas são apenas características das cargas. Por sua vez, as mesmas são citadas, no mínimo, em [43, 77] como dificultadores no processo de desagregação e, por esse motivo, preferiu-se adicionar diretamente essas categorias para enfatizar e facilitar a identificação de cargas com essas características.

res e som, alteram seu consumo (também observado em [84]); computadores, que alteram sua potência conforme a demanda dos processadores e *coolers*, consumindo mais quando o usuário está realizando tarefas como, por exemplo, executando algoritmos do mestrado, escutando música etc.; alguns outros equipamentos como o ar condicionado — esse quando com o compressor ligado — apresentam uma dinâmica com oscilações em frequências sub-harmônicas, os mesmos são potenciais dificultadores à identificação de rastros deixados por outros equipamentos, particularmente os de menor consumo, caso estudado em [77]. A Figura 4.2 demonstra a dinâmica sub-harmônica causada na envoltória para esse equipamento. Outros exemplos de equipamentos com motores que também geram oscilações — mas em ordem inferior ao ar condicionado — são: microondas, geladeira, desumidificador [77];

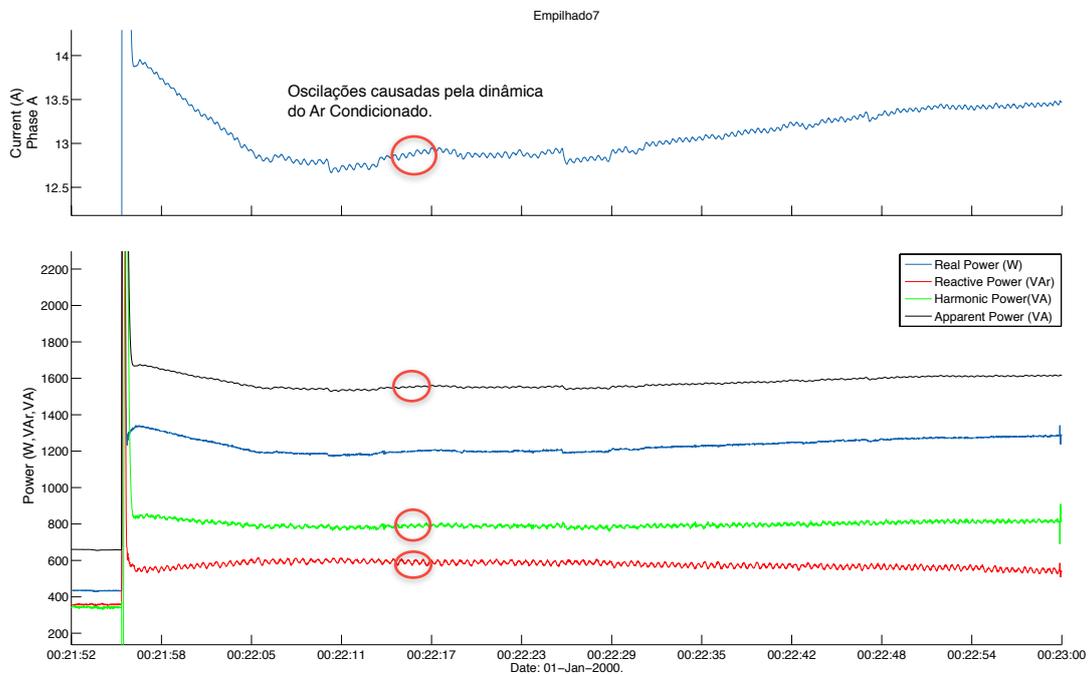


Figura 4.2: Exemplo de equipamentos com dinâmica de carga (C5): ar condicionado. Também é possível observar uma alteração gradual no nível de consumo, que irá dificultar o processo de reconstrução de energia.

- **Equipamentos com características similares (C6)³:** apesar de não ser uma característica de um equipamento *per se* e nem constituir um modelo de carga elétrica, estudos de performance de NILMs podem considerar quais equipamentos serão potencialmente vistos como se fossem um mesmo equipamento por possuírem os mesmos padrões. Essa categoria varia conforme os equipamentos presentes e quais são as características sendo extraídas. Por exemplo,

um computador e uma lâmpada incandescente possuem consumos semelhantes quando procurando padrões no plano $\Delta P \times \Delta Q$ [62], enquanto equipamentos com motores de potências distintas podem não ser desagregados quando apenas olhando para seus transitórios — em especial quando normalizados, caso que possivelmente ocorreria se utilizando Rede Neural Artificial (RNA). Um exemplo de ocorrência de C6 pode ser observado na Figura 4.3 para a característica ΔP . [73] propõe a equação 4.1⁴ como uma maneira de medir a similaridade entre dois equipamentos, onde $s_{ia,ib}^j$ é a relação entre a j -ésima característica entre os equipamentos ia e ib , N é o tamanho da dimensão dessa característica e $y_{k|(ia,j)}$ é o ponto k para o equipamento ia e j -ésima característica. Quanto mais próximo for $s_{ia,ib}^j$ da unidade, mais difícil será desagregar ia e ib utilizando a característica s^j . Uma última consideração, a presença de similaridade entre dois equipamentos não necessariamente é um problema. Deseja-se para certos equipamentos, como por exemplo, lâmpadas do mesmo tipo, ou equipamentos de marcas diferentes, que eles sejam vistos como um mesmo equipamento pelo NILM. Há certos casos em que as características são tão discriminantes que dificultam a aplicação do NILM em larga escala (ver metodologia na pp. 75), sendo necessário encontrar um equilíbrio entre a capacidade discriminante das características. No entanto, utiliza-se o termo C6s exclusivamente para identificar os casos em que a similaridade dos equipamentos não é desejada causando erros de classificação.

$$s_{ia,ib}^j = \frac{\sum_{k=1}^N y_{k|(ia,j)}^2}{\left[\sum_{k=1}^N y_{k|(ia,j)}^2 + \left(\sum_{k=1}^N y_{k|(ia,j)}^2 - \sum_{k=1}^N y_{k|(ib,j)}^2 \right)^2 \right]} \quad (4.1)$$

4.2.2 Escolha da Metodologia

A escolha da metodologia do NILM pode estar limitada à capacidade do sistema de aquisição de dados de extrair informação do consumo da rede elétrica, caso que ocorre quando desejando aplicar o NILM a uma estrutura já existente, ou livre, aonde o custo do sistema de aquisição de dados também torna a escolha de uma metodologia com o menor custo e capaz de atender as necessidades do projeto como a ideal a ser aplicada. Além disso, em ambos os casos há uma limitação da capacidade de processamento dos dados em tempo real, já que a informação desagregada deve ser informada com pequena defasagem após a ocorrência do consumo. Como no esboço apresentado na Figura 4.1, normalmente realiza-se a transformação da informação

⁴Nota-se, aqui, que qualquer outra divergência representativa pode ser empregada com esse intuito. Divergências são generalizações das métricas e medem a quase-distância entre duas medidas não-negativas [93, cap. 2]. No caso de algumas divergências, como a métrica Euclidiana, é possível aplicar valores negativos.

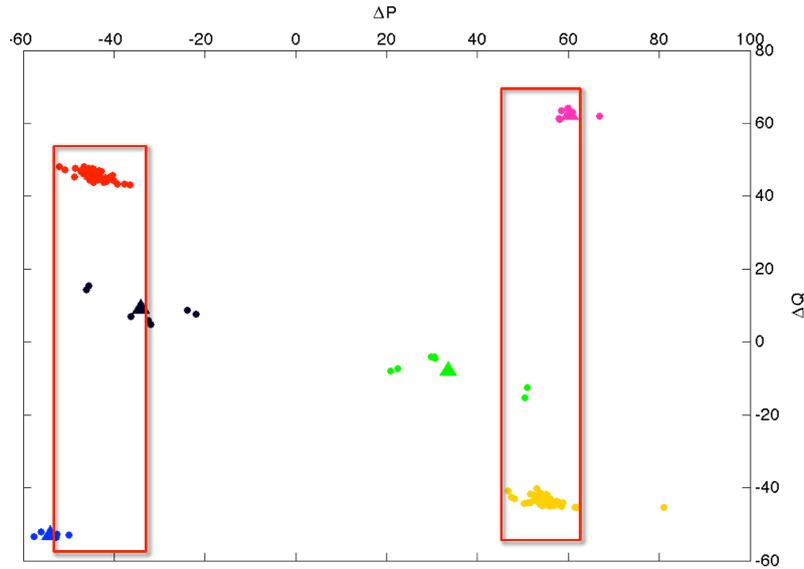


Figura 4.3: Exemplo de equipamentos com características similares (C6). As cores amarelo e vermelho constituem respectivamente acionamentos e desacionamentos de uma geladeira; rosa e azul análogo para um ventilador; e verde e preto análogo para lâmpadas fluorescentes no plano $\Delta P \times \Delta Q$. Caso apenas observando ΔP , percebe-se que os equipamentos possuem uma região de confusão o que prejudicará o processo de discriminação se apenas empregando essa característica.

em características que apresentam uma melhor representação para a realização da discriminação e desagregação do consumo. Há ainda a questão de quando extrair a informação, se em regime permanente ou apenas quando ocorre a transição do estado operativo de um equipamento. É necessário optar pelas técnicas para realizar a discriminação ou se serão empregadas múltiplas técnicas e a estratégia para detecção de transitórios quando extraindo informação nesses momentos. As técnicas precisam ser ajustadas para a operação, seja o treinamento de técnicas supervisionadas, a obtenção de dados para o emprego de discriminadores estatísticos ou quais informações devem ser trabalhadas para a otimização, assim cabe a questão de quando e com que informação realizar o ajuste. Essas questões são esquematizadas de maneira genérica a seguir, levando em conta as diversas metodologias observadas e que serão detalhadas na Subseção 4.2.4 e debatidas em seguida na Subseção 4.2.5.

A procura por padrões de operação dos equipamentos

Quando extrair informação para a procura por padrões de operação dos equipamentos?

- **Em regime transitório:** identifica o novo consumo dos equipamentos apenas quando ocorre alteração do estado de um ou mais deles. Para isso, se faz necessário a detecção de eventos de transitório aonde são diferenciados os distúrbios na rede devido a ruído daqueles causados por transições no consumo

dos equipamentos. Isso possibilita o tratamento de menos informação pelas técnicas de discriminação e a independência da informação tratada em relação à presença de outros equipamentos (exceto quando há a presença da operação de C5s, que irão interferir na informação obtida);

- **Em regime permanente:** identifica o consumo durante a operação contínua dos equipamentos em um mesmo estado de operação. Isso traz como necessidade altas frequências de amostragem (superiores a frequência da rede), porém permite a identificação de C1. Da maneira que essa informação foi extraída pelos autores nas referências observadas, a exploração dessa informação é limitada para a presença de poucos equipamentos, uma vez que há um crescimento exponencial da complexidade problema de acordo com o número de estados operativos dos equipamentos presentes na rede.

Definição das representações da informação

Qual representação da informação vai ser utilizada como característica para as técnicas empregadas para discriminação do consumo dos equipamentos?

- Macroscópicas - informação simplificada dos ciclos da rede, ex.:
 - Corrente eficaz/pico;
 - Tensão eficaz/pico;
 - Potência Ativa, Reativa, Harmônica, Aparente;
 - Distorção Harmônica Total (THD, em inglês).
- Intermediária - informação contida nas envoltória das ondas (corrente, tensão, potência etc.) — geralmente relevante apenas para acionamentos de equipamentos ou eventos de transitório com acréscimo de consumo;
- Microscópicas - informação detalhada de um ciclo da rede, ex.:
 - Forma de onda (sem tratamento);
 - Decomposição harmônica resultante da Transformada de Fourier;
 - Informação nos níveis de detalhes da Transformada Wavelet;
 - Autovetores mais relevantes da Decomposição em Valores Singulares (SVD, em inglês) das formas de onda (de corrente, por exemplo);
 - Curvas I-V.
- Outros - demais informações que podem ser exploradas não diretamente extraídas do consumo de energia dos equipamentos na rede, ex.:

- Informação de estatística de uso - tempo e horários que os equipamentos costumam operar em cada um de seus possíveis estados operativos;
- Interferência Eletromagnética (EMI, em inglês) - assinaturas deixadas pelos equipamentos devido ao chaveamento, seja por sua operação (ex.: fontes chaveadas, escova de motores) ou durante seu acionamento e desacionamento (ex.: conectar o equipamento na tomada, alteração do estado de um interruptor);
- Informação do ambiente - correlacionar informação proveniente no ambiente (ex.: movimento, som) com o uso dos equipamentos. Necessário a utilização de outros sensores, constituindo uma abordagem semi-invasiva (ver pp. 84).

Abordagens (baseado em [43, 73])

Quais técnicas empregar para a discriminação das características?

Abordagem 1 - **Abordagem por reconhecimento de padrões**⁵: as técnicas de reconhecimento de padrões podem ser tanto técnicas de aprendizado de máquina quanto técnicas estatísticas ajustadas (ou treinadas) em conjuntos de dados similares aos quais eles irão operar. O reconhecimento de padrões pode ocorrer apenas para os eventos de transitórios detectados pela Etapa 2 ou para cada ciclo da rede, dependendo de qual informação está sendo tratado para a identificação de padrões. Algumas técnicas utilizadas nessa abordagem podem ser robustas aos equipamentos desconhecidos, sendo capaz de destacar seu padrão dos outros já conhecidos e adicioná-lo ao catálogo de padrões. Assim, quando esses padrões ocorrerem novamente, eles serão identificados como o mesmo equipamento — chamado de aprendizado em tempo real. Em [61] é observado a importância dessa estratégia para tornar possível o crescimento do catálogo, que, tendo o novo equipamento nomeado pelo consumidor, torna possível a criação de um catálogo universal de equipamentos. Essa tarefa é impraticável, se não impossível, de ser realizada em laboratório. Exemplos de técnicas:

- Técnicas de Aprendizado de Máquina:

⁵A referência [43] trabalha a diferença das abordagens em termo da quantidade de dados abordados, onde a Abordagem 1 trata a informação uma a uma, ou seja, para cada evento de transitório, enquanto a Abordagem 2 opera com todo o período sendo otimizado, o que não é necessariamente verdade. Geralmente ambos os casos passam por um período de otimização antes de serem empregados e, depois de otimizados, são utilizados para a detecção dos padrões dos equipamentos na rede. A distinção está que o primeiro otimiza a capacidade de discernir os padrões — sendo a reconstrução consequência disso —, enquanto o segundo a capacidade de reconstruir com maior fidelidade possível o sinal original — obtendo os padrões como resultado.

- Redes Neurais (*MultiLayer Perceptron* (MLP) e *Radial Basis Function* (RBF));
- *Support Vector Machine* (SVM).
- Mapeamento em grupos:
 - Mapas Auto-Organizáveis (SOM, em inglês);
 - *Iterative Self-Organizing Data Analysis Techniques Algorithm* (ISODATA);
 - Teoria de Ressonância Adaptativa (ART, em inglês);
 - *K-means*.
- Discriminadores estatísticos:
 - Vizinho próximo;
 - *Naiïve Bayes*.

Abordagem 2 - **Abordagem por otimização**⁵: concentra a capacidade de suas técnicas na otimização, onde é realizada a procura por uma combinação de equipamentos cujo o sinal agregado resultante é a melhor aproximação do possível do sinal observado. Em alguns casos, utiliza-se a concentração dos dados em longos períodos de tempo para identificar o consumo desagregado, retornando a operação dos diversos equipamentos no final do processo. Nesses casos, a informação final pode ser utilizada como padrões a serem identificados posteriormente. Para manter os equipamentos atualizados, novos períodos (possivelmente menores ao período inicial) podem ser utilizados para garantir a resposta adequada a possíveis alterações na presença ou utilização de equipamentos. Na outra possibilidade, a otimização é realizada no nível de ciclo da rede, onde, sabendo o padrão dos possíveis equipamentos presentes, busca-se a melhor combinação operativa que reflitam o sinal observado. Exemplos de técnicas:

- Programação por Inteiros;
- Algoritmo Genético;
- Força Bruta.

Abordagem 3 - **Abordagem empregando múltiplas técnicas**: ao perseguir melhor eficiência, é possível aplicar múltiplas técnicas, tanto de otimização quanto de reconhecimento de padrões, que irão em paralelo para complementarem as deficiências entre si. Por outro lado, isso torna necessário mais uma etapa no processo de discriminação, aonde será unida a informação das diversas técnicas aplicadas no NILM (o assunto é tratado em mais detalhes na pp. 80).

Ajuste das técnicas

- **Pré-ajuste:** os padrões são obtidos via laboratório ou em testes realizados anteriormente. As técnicas podem vir pré-ajustadas antes de serem aplicadas na residência, estando prontas para operação assim que o NILM for instalado. Porém, isso pode ter algumas limitações uma vez que o pré-ajuste dificilmente irá representar a realidade exata da rede elétrica residencial. Mesmo que o catálogo seja extenso, é necessário ajustá-lo posteriormente para evitar a presença de C6, limitando o catálogo apenas aos equipamentos existentes na residência;
- **Ajuste na residência:** obtenção dos padrões através de submedição ou operação individual dos equipamentos (realizado pelo próprio consumidor ou com auxílio). Esse seria o ideal, se não fosse uma tarefa árdua. A exigência da participação do consumidor pode reduzir o interesse dos mesmos no produto, bem como eles podem não conseguir realizar o ajuste de maneira adequada do NILM. Mesmo que o processo seja realizado por técnicos (o que tornaria o produto mais caro), é interessante evitar a intrusão da residência em longos períodos de tempo para tornar o NILM em um produto mais atraente;
- **Auto-ajuste:** a técnica aplicada é capaz de descobrir os padrões a serem encontrados. Esforços foram feitos para a identificação e modelagem automática dos equipamentos de maneira cega durante a operação. Geralmente o NILM passa por uma fase maior de coleta de dados (ex. 1 semana), aonde é realizado a identificação dos modelos a serem encontrados, para então operar e identificar esses padrões. Porém, apenas equipamentos que tem padrões recorrentes podem ser encontrados (são utilizadas adaptações do algoritmo de *Viterbi* que encontra a sequência mais provável para as alterações de estado). Além disso, esse ajuste pode ser utilizado em conjunto com o pré-ajuste, para permitir a operação do NILM assim que instalado na residência, para então o mesmo se autoajustar conforme a coleta de dados for sendo realizada.

Etapas durante operação (baseado em [61])

Etapa 1 - **Extração de Características (FEX, em inglês):** são extraídas informações das amostragens realizadas. A diversidade de características que podem ser extraídas depende da capacidade do sistema de aquisição de dados, em especial no medidor, mas podendo sofrer devido a outras limitações (como capacidade de transferência de dados etc.). As características são utilizadas nas etapas seguintes, podendo haver reaproveitamento. Em alguns casos, a Etapa 2 utiliza o sinal sem processamento e a FEX é realizada somente para

a Etapa 3, o que permite reduzir o esforço de processamento;

Etapa 2 - **Detecção de eventos de transitório**: identificar alterações causadas por equipamentos na rede. Essa etapa é necessária para extrair a informação em regime transitório. Pode-se empregar limiares estáticos ou dinâmicos para a detecção dos eventos. Os limiares dinâmicos permitem o ajuste de operação, reduzindo ou aumentando a sensibilidade do detector conforme a presença de equipamentos C5;

Etapa 3 - **Discriminação da informação**: utilizar as características pertinentes para o reconhecimento de padrões, deduzindo, assim, qual foi o equipamento que causou o distúrbio na rede e qual seu novo estado de consumo. A discriminação da informação pode ser realizada por técnicas de aprendizado de máquinas, estatística ou via otimização. Possivelmente, múltiplas técnicas podem ser empregadas nesta etapa a fim de maximizar a eficiência. É desejável que a técnica seja capaz de identificar a ocorrência de novos padrões e reconhecê-los em suas próximas aparições — ajuste automático das técnicas — pois a construção de um catálogo com todos os possíveis eletrodomésticos é impraticável, se não impossível. Tal tarefa só será possível com a capacidade dos NILMs de incluírem novos equipamentos ao catálogo;

Etapa 4 - **União da informação e refinamento dos resultados**: quando empregando múltiplas técnicas, faz-se necessário a união da informação discriminante fornecida pelas técnicas na Etapa 3. Além disso, durante a união da informação, pode-se procurar por possíveis erros ou melhorias que possam ser feitas na informação desagregada. Por exemplo, ao correlacionar a nova informação obtida com a passada, pode ser possível identificar um equipamento que remanesce consumindo energia da rede por tempo demasiado, enquanto sua operação normalmente ocorre em intervalos curtos, o que possivelmente foi causado por falhas na Etapa 2, onde o desacionamento do equipamento não foi encontrado, ou na Etapa 3, na qual o desacionamento foi identificado como causado por outro equipamento. As estratégias corretivas podem ser meramente remediativas, ou seja, simplesmente remover incoerências sem realizar um tratamento procurando sua causa, ou aplicar técnicas complementares para reanalisar a informação disponível.

4.2.3 Cálculo da Eficiência

O assunto aqui tratado é de grande importância para permitir uma comparação justa entre os NILMs de diferentes autores e, por isso, levado em consideração nesta

subseção. Porém, este trabalho não teve a oportunidade de empregar as medidas de eficiência debatidas nesta subseção pois trabalhou a detecção de eventos de transitório — foram utilizados a taxa de detecção e falso alarme dos eventos nos resultados divulgados no Capítulo 7 —, não estando disponíveis, portanto, a capacidade discriminante e a estimativa do consumo para empregar as medidas aqui debatidas.

Padronização

O estudo bibliográfico realizado por [43] teve dificuldades ao tentar comparar as diferentes técnicas aplicadas nos NILMs. O primeiro empecilho está na variedade das base de dados utilizadas, possuindo equipamentos e estados de operações bastante distintos, criando condições que podem privilegiar a eficiência de um determinado NILM. Para a unificação dos dados estudados e permitir a comparação de performance entre os algoritmos empregados nos NILMs, foi disponibilizado por [94] um conjunto de dados públicos para a análise, o BLUED. Um outro conjunto de dados público utilizado nas referências é o REDD [95], onde há dados gravados tanto em alta frequência (15 kHz) e baixa frequência (0,5 Hz e 1 Hz). Os conjuntos de dados foram construídos para representar a realidade de residências nos EUA e por isso podem não corresponder a realidade brasileira. Apesar disso, o conjunto de dados podem servir como base para comparação da performance dos NILMs aqui desenvolvidos com os do exterior, assim como nada impede o emprego em paralelo de dados próprios mais representativos para demonstrar a realidade de aplicação e capacidade da abordagem utilizada.

Outra dificuldade foi o fato de autores utilizarem uma medida própria para o cálculo das taxas de eficiência. Além disso, normalmente os autores não reportaram as taxas de falsos positivos na Etapa 2, apenas a capacidade dos algoritmos de detectarem os eventos (exceções observadas são [57, 77]). Em outros casos, os autores concentraram-se apenas na capacidade dos algoritmos da Etapa 3 de discriminarem equipamentos, reportando medidas representativas para essa eficiência.

Por isso, [43] recomenda a utilização das medidas apresentadas por [73], no qual se apresentou considerações metódicas para o tema. Foram apresentadas três medidas. A primeira medida, eficiência de detecção (η_{det}), considera a capacidade do NILM de desagregar a informação nos eventos que foram detectados⁶. Quando interessado apenas em estudar a capacidade do classificador para os eventos detectados, a medida eficiência de classificação (η_{class}) deve ser utilizada. Finalmente, a eficiência total (η_{total}) é dada por 4.2c, levando em conta apenas a capacidade do NILM de corretamente classificar os eventos reais, causados pelos equipamentos na

⁶Empregada por [48] quando não disponível a medição paralela de energia dos equipamentos e por [68].

rede⁷.

$$\eta_{det} = \frac{N_{id}}{N_{real} + N_{fp} - N_{ni}} \quad (4.2a)$$

$$\eta_{class} = \frac{N_{id}}{N_{real} - N_{ni}} \quad (4.2b)$$

$$\eta_{total} = \frac{N_{id}}{N_{real}} \quad (4.2c)$$

onde:

N_{id} são eventos identificados, ou seja, corretamente detectados e classificados pelo NILM;

N_{real} é o número de eventos realmente causados pelos equipamentos na rede;

N_{fp} são eventos devido a falsos positivos, ou seja, evento erroneamente identificados;

N_{ni} são eventos não identificados, ou perdas de alvo.

Segmenta-se 4.2c para obter a eficiência do NILM por equipamento conforme:

$$\eta_{total}^i \approx \frac{N_{id}^i}{N_{real}^i} \quad \forall \quad i = 1, 2, \dots, N_{ap} \quad (4.3)$$

onde N_{id}^i e N_{real}^i são os respectivos N_{id} e N_{real} para o i -ésimo equipamento dos N_{ap} disponíveis.

O grande favorecimento para essas medidas é sua simplicidade de serem obtidas, porém algumas considerações podem ser feitas sobre elas. Primeiro, a medida com maior sensibilidade à capacidade do NILM é a 4.2a, uma vez que os valores por ela representados levam em conta as perdas de alvo e os falsos positivos. Segundo, as mesmas não levam em conta o consumo de energia dos equipamentos, dando importância análoga para equipamentos com parcelas pequenas ou grandes de consumo. Além disso, a correta identificação dos N_{id} não significa que a energia será corretamente reconstruída, dependendo da capacidade do NILM de unir essas informações para gerar a informação do consumo desagregado. Ainda, como apontado por [43], elas apenas representam a eficiência no ponto de operação, não sendo possível observar como o NILM se portaria para outros pontos. Indo além, elas não permitem comparações de técnicas utilizadas exclusivamente para as Etapas 2 e 3, impedindo a contraposição de técnicas onde os autores se limitaram a uma dessas etapas.

⁷Empregada por [67, 76] pois a Etapa 2 foi realizada manualmente e, geralmente, por demais estudos que estudaram apenas a Etapa 3.

Outras representações

Por isso, além das medidas apontadas, outras maneiras de representar a eficiência podem ser utilizadas para complementar o estudo do comportamento da abordagem utilizada. Uma técnica para representar o compromisso entre a capacidade de detectar eventos e a quantidade de falsos positivos encontrados é a curva *Receiver Operating Characteristic* (ROC), também recomendada por [43]. A ROC além de ser utilizada para expressar de maneira geral a capacidade do algoritmo de detectar e identificar em função dos falsos positivos, pode ser utilizada para estudar a eficiência específica da Etapa 2. Já para a Etapa 3, a matriz de confusão permite entender quais equipamentos ou classes de equipamentos são confundidos em outras classes, assim como a eficiência de classificação em uma única representação.

As outras medidas utilizadas na literatura levantada são: o percentual de classificações corretas por ciclo da rede, ou seja, a correta classificação do estado de operação para cada ciclo dividido pelo número total de ciclos [66, 74]; porcentagem de detecção de transitórios [67]⁸; taxa de reconstrução em energia para o i -ésimo equipamento (ρ_{En}^i) [48, 53, 54, 56, 57]; taxa de reconstrução em energia do NILM (ρ_{En}) [29]; desvio do tempo em que o equipamento foi identificado operando em relação ao tempo que ele realmente estava operando [56]⁹; porcentagem de detecção de eventos de transitório para ligado perdidos [56]⁹; erro médio absoluto de reconstrução de energia e outras estatísticas por equipamento [55].

A medida mais utilizada pelas referências, a ρ_{En}^i , embora por elas não definida matematicamente, concebe-se que seja dada por:

$$\rho_{En}^i = \frac{E_{det}^i}{E_{real}^i} \quad \forall \quad i = 1, 2, \dots, N_{ap} \quad (4.4a)$$

$$\rho_{En} = \frac{\sum_i^{N_{ap}} E_{id}^i}{\sum_i^{N_{ap}} E_{real}^i} \quad (4.4b)$$

onde E_{det}^i , E_{real}^i é o consumo detectado e consumo real do i -ésimo equipamento, respectivamente, o último sendo obtido por submedição ou por um estimador. A ρ_{En}^i pode ser generalizada para calcular a ρ_{En} através de 4.4b. Essas medidas levam em consideração o consumo detectado pelo NILM, mas perdem a capacidade das medidas 4.2 de representar a informação que foi corretamente identificada. Por exemplo, se um equipamento é considerado como ligado em um espaço de tempo em que o mesmo está desligado, isso irá contribuir para corrigir possíveis erros que seriam atribuídos quando o estado estimado e a operação estiverem na lógica oposta.

⁸O estudo reportou eficiência para as Etapas 2 e 3 separadamente. Como foi dito na nota 7, a Etapa 3 utilizou eventos recortados manualmente.

⁹O estudo focou na identificação de grandes cargas elétricas como ar condicionado e aquecedores de água, modelados por C3. Essa medida seria limitada para outros modelos.

Assim, fica evidente que essas medidas precisam ser refinadas para identificar os momentos nos quais a energia foi corretamente reconstruída. Para isso, aqui se sugere o uso de 4.5 com o intuito de determinar a energia corretamente identificada para o i -ésimo equipamento (E_{id}^i). A ideia é representar que identificações do equipamento em outros estados, mas com pequena diferença de energia, não irão afetar tanto na resolução de energia, assim como resguardar que identificações em estados de consumo maiores para os quais os equipamentos realmente operam não arremeterão na conta de energia corretamente identificada.

$$E_{id}^i = E_{det}^i - \varepsilon^i \quad (4.5)$$

A energia redundante (ε) representa a ideia, em energia, para tanto falsos positivos ou quanto identificações errôneas para estados de maior consumo, ou seja, a parcela de E_{det}^i que excede àquela obtida através de um submedidor — medição diretamente da energia do equipamento — ou estimada E_{real}^i . Ela pode ser descrita por:

$$\varepsilon^i = \begin{cases} E_{det}^i - E_{real}^i & \text{se } E_{det}^i > E_{real}^i \\ 0 & \text{o.c.} \end{cases} \quad \forall i = 1, 2, \dots, N_{ap} \quad (4.6)$$

Isto posto, para obter a eficiência de reconstrução em energia para o i -ésimo equipamento (η_{En}^i) e sua generalização, eficiência de reconstrução em energia do NILM (η_{En}), basta empregar:

$$\eta_{En}^i = \frac{E_{id}^i}{E_{real}^i} \quad \forall i = 1, 2, \dots, N_{ap} \quad (4.7a)$$

$$\eta_{En} = \frac{\sum_i^{N_{ap}} E_{id}^i}{\sum_i^{N_{ap}} E_{real}^i} \quad (4.7b)$$

E para as taxas de redundância:

$$\rho_{red}^i = \frac{\varepsilon^i}{E_{real}^i} \quad \forall i = 1, 2, \dots, N_{ap} \quad (4.8a)$$

$$\rho_{red} = \frac{\sum_i^{N_{ap}} \varepsilon^i}{\sum_i^{N_{ap}} E_{real}^i} \quad (4.8b)$$

Posteriormente, descobriu-se que o próprio autor de [43] criou uma medida que, depois de modificada por [90], permite também exprimir a questão de precisão de reconstrução de energia, bem como reflete, de outra maneira, a quantidade de energia redundante. A alteração de [90] será referida como medida-eF (F_{en}^i). Ela pode ser descrita pelo quadrado da média geométrica normalizado pela média aritmética (4.10) de duas grandezas:

- parcela de energia atribuída ao equipamento que foi realmente consumida pelo mesmo em relação ao seu consumo pelo NILM atribuído, um parâmetro que também reflete a ideia de energia redundante, descrita por 4.9;
- parcela de energia que foi corretamente identificada em relação ao consumo total do equipamento, a própria η_{En}^i aqui descrita em 4.7a.

$$\eta_{En,prec}^i = \frac{E_{id}^i}{E_{det}^i} \quad (4.9)$$

$$F_{en}^i = \frac{2 \eta_{En,prec}^i \eta_{En}^i}{\eta_{En,prec}^i + \eta_{En}^i} \quad (4.10)$$

Entretanto, a medida-F (4.11) original apresentada por Zeifman [82], como observado pelo próprio autor, não é ótima como medida de desagregação, por não levar em consideração erros na detecção de energia. Ela fica aqui referenciada por ter sido utilizada em diversos estudos desse autor.

$$F^i = 2 \frac{\left(\frac{N_{id}}{N_{id} + N_{fp}} \right) \eta_{total}}{\left(\frac{N_{id}}{N_{id} + N_{fp}} \right) + \eta_{total}} \quad (4.11)$$

4.2.4 Técnicas Aplicadas por demais autores

As abordagens aplicadas desde o início dos estudos ao tema e utilizadas como referências fizeram mão de ostensivas técnicas para a desagregação do consumo. Cada vertente buscou extrair características ou inovar aplicando outras técnicas, de forma que é possível observar uma grande diversidade de abordagens. As abordagens serão agrupadas em relação à FEX realizada. A capacidade de extrair características dos sinais é correlacionada com a frequência de amostragem e, em vista disso, dividir-se-ão os métodos aplicados de acordo com a taxa de amostragem utilizada. A ideia de subdivisão aqui seguida foi de autoria da referência [43].

Em seguida, um outro tópico específico para as topologias onde há semi-intrusão da residência, utilizando submedição em pequenas escalas, tanto energeticamente, quanto correlacionando a utilização de energia com outros medidores, como temperatura, movimento, som etc.

1. Medição com Baixa Amostragem

A utilização de características macroscópicas de consumo do equipamento, como alterações no patamar de consumo da rede, foi a primeira abordagem encontrada ao

tema. As mesmas podem ser obtidas sem grande granularidade na taxa de amostragem, por isso, esse tipo de abordagem beneficia-se de medidores de baixo custo, amplamente disponíveis no mercado. No entanto, [61, 75] alertam para discrepâncias entre medidores na ordem de 10%-20%, bem maiores que aquelas alegadas, de 3%. Os medidores testados no caso foram *Brand Meter I*, *Watts Up? PRO* e *EnerSure*.

A taxa de amostragem mais frequentemente utilizada é 1 Hz, entretanto alguns estudos fizeram mão taxa de amostragem ainda menores por desejarem identificar equipamentos que se ressaltam dentre os outros devido ao seu relativo alto consumo, como ar condicionado, aquecedores de água e geladeira. Exemplos de medidores utilizados no exterior são *The Energy Detective* (TED) [36] e *Watts up? PRO* [96], o último sendo capaz de informar o consumo de potência reativa.

1.1 Potência ativa e Potência reativa

A referência inicial de grande destaque no tema, Hart [48]¹⁰, ocorreu em 1992. Nela, utilizam-se medições de potência ativa (P) e potência reativa (Q) com uma taxa de amostragem de 1 Hz. A abordagem aplica uma normalização para reduzir flutuações no consumo devido a alterações na tensão de acordo com 4.12 com o intuito de reduzir dispersões nos dados. O estudo de Hart limitou-se a identificar apenas cargas com potência maior a 150 W. A essência da metodologia ainda pode ser encontrada em NILMs mais atuais, sendo esta:

$$P_{\text{norm}}(t) = \left[\frac{120}{V(t)} \right]^2 P(t) \quad (4.12)$$

- 1 Detectam-se transitórios de consumo na rede devido a mudança de estado de um equipamento através de alterações no consumo que devem superar um limiar específico (15 W/VAr) para os sinais normalizados como em 4.12 para P e Q. As amostragens dentro de um regime permanente são normalizadas para sua média com o objetivo de tirar o ruído. Para a FEX, utiliza-se o degrau entre o regime permanente posterior e anterior (já no valor de suas médias) ao evento transitório para P e Q;
- 2 Os eventos de transitório são analisados por um algoritmo de agrupamento que irá gerar os centroides das mudanças de estado possíveis causadas pelos equipamentos no plano $\Delta P \times \Delta Q$;
- 3 Centroides com simetria em relação aos eixos são tomados em pares e com eles são criados modelos C3. Para os centroides remanescentes, além de regras

¹⁰Sultanem [69] publicou seu estudo em 1991, antes de Hart que o publicou em 1992. Aparentemente, nenhum dos autores teve conhecimento do trabalho do outro. Apesar do trabalho de Sultanem ter sido publicado primeiro, Hart retrata em seu trabalho que sua descoberta ocorreu em 1982, tendo diversos trabalhos realizados pelo seu grupo em relação ao tema desde essa data, resultando em um material bem mais extenso que o de Sultanem.

heurísticas como a junção de centroides próximos que permitam o pareamento com um outro refletido nos eixos — etapa conhecida como resolução de conflitos —, determinam-se possíveis combinações de centroides que possam formar uma C2 utilizando uma adaptação do algoritmo de *Viterbi* [49, 50]¹¹. Assim que é determinada uma combinação que permite a criação de uma C2, os centroides da mesma são removidos, e o processo continua até que todas as FSMs tenham sido construídas. A adaptação utilizada [49, 50] permite várias operações para consertar corrupções e retornar uma estimativa ótima das alterações de estados operativos da FSM original. Como a reconstrução depende da estatística do processo, é necessário que as mudanças de estado das FSMs observadas tenham um comportamento recorrente para que a FSM original seja reconstruída, e por isso, restringe-se apenas às C2as. As C2bs podem ser reconstruídas se houver conhecimento prévio da presença das mesmas, de modo que elas sejam medidas operando em cada um de seus estados e então inseridas manualmente no catálogo do NILM;

- 4 Em seguida é levantado o comportamento dos equipamentos, montando o estados de consumo para cada equipamento. É utilizado um algoritmo de força-bruta para corrigir ocorrências de dois acionamentos ou desacionamentos de um mesmo equipamento encontrados em seguida. A causa desses erros é, geralmente, a ocorrência de um evento simultâneo de dois equipamentos. Assim, o algoritmo busca por eventos não-usuais cuja soma é o valor de dois outros eventos perdidos;
- 5 Por fim, é levantada a estatística detalhando o comportamento de consumo, como o tempo ligado e desligado de cada equipamento. Essa informação, junto com a potência do equipamento é utilizada para auxiliar a identificar o equipamento.

O método é robusto para o desagregação de cargas C3s (> 150 W) e as adaptações [49, 50] do algoritmo de *Viterbi* parecem resolver o problema das C2as. Outro problema das metodologias envolvendo algoritmos de agrupamento é a lenta alteração da resistência conforme a operação do equipamento. Geralmente, ao interromper a operação, o equipamento tem alterações no consumo na margem de 5%-10% em relação ao início de operação [69]. Hart observa o degrau geralmente é menor em valor absoluto para os desacionamentos nos casos de equipamentos com motores, que reduzem o consumo conforme seu aquecimento.

Uma estratégia bastante parecida é realizada por Cole e Albicki [53, 54], onde são feitas considerações em relação as características de bordas e inclinações. Aqueles

¹¹A versão original pode ser encontrada em [97].

são definidos como o auge atingido de potência durante o acionamento e estas variações lentas de mudança no consumo. Apesar de definir as bordas como o pico de potência, as referências empregam as bordas apenas como os eventos de transitório de consumo, não empregando essa informação para classificação. A abordagem aplicada, ao invés de agrupar os dados para depois procurar por possíveis equipamentos como feito por Hart, primeiro busca temporalmente por ciclos fechados nos eventos de transitório (ou bordas, como na nomenclatura da referência), que depois serão adicionadas aos centroides no espaço $\Delta P \times \Delta Q$. Se o centroide não existir, será criado um candidato a centroide. Conforme a quantidade de ciclos dos centroides aumenta, o mesmo irá se tornar um candidato a uma carga. Para as cargas C3s, a carga será aceita apenas se a detecção das bordas ocorrerem repetidamente. Já para as cargas C2as, foi realizado um estudo da probabilidade dela ter sido originada pela sobreposição de duas bordas geradas por equipamentos distintos. A conclusão foi que se forem encontrados mais de um ciclo de três bordas em um período de 6 horas é suficiente para aceitá-lo como uma C2a. Finalmente, o envelope só foi considerado para a melhoria em resolução de energia e, no entanto, a referência indica que a utilização da média de consumo entre as bordas apresenta melhores resoluções.

1.2 Potência ativa, Potência reativa e Transitório

Um trabalho paralelo ao de Hart foi realizado pelo mesmo instituto para aplicar o NILM no setor comercial e industrial, sendo realizado por Norford e Leeb [51], Leeb, Shaw, e Kirtley [52]. No setor comercial são encontrados equipamentos com características diferentes ao setor residencial, geralmente com transitórios mais lentos (podendo chegar a cerca de centenas de segundos [51]), menor consumo reativo devido às preocupações com a qualidade de energia e conseqüentemente correção do fator de potência, e a presença de C5, como exemplo, na referência foi observada uma bomba com picos periódicos de 20 kW. Tipicamente há também uma maior presença de equipamentos com cargas variáveis, C4s, como motores de velocidade variável.

Assim, foi adicionado a informação do transitório da envoltória em amostragens de 1 Hz, que por serem maiores, ainda podem ser observados em amostragens baixas, suprindo, ao mesmo tempo, a menor capacidade de discriminação da variável Q nesse setor. Para redução dos ruídos, utilizou-se um filtro de mediana com 11 pontos, esse sendo mais indicado para a eliminação dos picos quando comparado aos filtros lineares, que terão dificuldades de distinguir os picos e os degraus, uma vez que eles tem espectros de frequência parecidos. Os trabalhos utilizam o conceito de seções-v (tradução própria de *v-sections*), que são subdivisões do transitório em segmentos de variação significativa. Para isso, o grupo utilizou uma janela para identificar mudanças na média e criar as seções-v, onde os valores ótimos são determinados

por um processo de treinamento pré-instalação ou em laboratório. É aplicada uma medida de distância entre as seções-v observadas e os transitórios característicos que, ao estarem dentro de um limiar, serão identificados como um determinado equipamento. Para o tratamento das C4s, a referência indica o emprego de variáveis de controle, quando disponíveis, correlacionadas com o seu consumo para estimá-las, como o caso para os equipamentos de Aquecimento, Ventilação e Ar Condicionado (AVAC) em geral.

1.3 Unicamente potência ativa

A medição de potência reativa adiciona custo ao NILM — ainda que não tão oneroso quanto medições em altas frequências — e, para detectar certos equipamentos com assinaturas de destaque na rede, essa variável pode ser desnecessária. Em outros casos, medidores que disponibilizam essa informação podem não estar disponíveis, sendo possível operar apenas com a potência ativa.

1.3.1 Separação dos principais equipamentos por uso-final

Exemplos do primeiro caso são os estudos de Powers, Margossian, e Smith [55], Farinaccio e Zmeureanu [56], Marceau [57], para os quais os autores se preocuparam em identificar apenas equipamentos de maior uso-final. [55, 56] utilizaram somente regras heurísticas, enquanto [57] também empregou os degraus em potência real e um filtro para a detecção dos acionamentos/desacionamentos dos equipamentos.

Em [55], foram reportadas a capacidade de reconstrução para ar condicionado e aquecedores de água. A amostragem é realizada a cada 15 minutos e os arquivos são analisados dia a dia. Por ser proprietário, as regras não são detalhadas (é utilizada uma árvore de decisões, embora o estudo considere a aplicação de redes neurais), mas o algoritmo procura por picos no consumo, assim como sua duração, tempo e magnitude, que são utilizados pelas regras para determinar se os mesmos foram utilizados para os usos-finais cobçados. Posteriormente, eles são ajustados conforme verificações de consistência. Para o ar condicionado, é relatado que o valor de pico estimado médio para as residências difere cerca de apenas 5% do valor original médio, enquanto o consumo fica na margem de 10% e observa-se boa capacidade de estimar os horários de consumo.

Já os estudos [56, 57], realizados por outro grupo, empregaram amostragem de P a cada 16 segundos. Os equipamentos estudados foram: geladeira, aquecedor de água e aquecedores de ambiente (este somente em [57]¹²). [43] expõe a arbitrariedade e não intuitividade das regras utilizadas em [56], que precisam ser estudadas para cada caso de equipamento. Foram determinadas 8 regras para cada equipamento (algumas regras são reaproveitadas entre equipamentos), divididas em duas

¹²O algoritmo da referência [57] também leva em consideração a máquina de lavar roupa, mas os resultados focaram apenas nos outros três equipamentos.

etapas: determinar o conjunto de eventos de transitório e a duração do consumo. Em seguida, a duração de consumo é multiplicada pela demanda média do equipamento durante a fase de treinamento para obter o consumo estimado. A fase de treinamento, período em que há medição paralela dos equipamentos, e, por isso, ocorrendo intrusão da propriedade do consumidor, é feita para um período de uma semana. A reconstrução de energia diária para os equipamentos é na margem de $-10,5\%$ a $15,9\%$.

O estudo em sequência aperfeiçoou o anterior com uma abordagem única para determinar os equipamentos em operação. Ele compara, em ordem decrescente em termos de demanda média operativa, se a magnitude do evento é próxima à média do nível operativo de um dos equipamentos almejados, empregando como limiar de corte dois desvios padrão. Ainda assim, a referência emprega diversas regras de pré/pós-processamento determinadas empiricamente para melhorar a resolução em energia, assim como também necessita do período de treinamento através de medição paralela de 1 semana, limitando a aplicabilidade do método para uma gama maior de equipamentos. Por outro lado, o método serve para o seu propósito, obtendo reconstruções na faixa de 10% para a maioria das análises realizadas.

1.3.2 Potência reativa não disponível

- *A abordagem de Baranski e Voss*

Com o intuito de possibilitar a impregnação da aplicação de NILMs na Alemanha, Baranski e Voss [58, 59, 60] recorreram a leitura ótica dos medidores eletromecânicos (o trabalho [58], realizado em 2003, indica que mais de 99% dos medidores desse país possuem essa configuração) para obter as medições com frequência de 1 Hz . Por isso, apenas P estava disponível para esses estudos.

Apesar das limitações, essa abordagem é uma referência de destaque devido às diversas contribuições feitas. Para melhorar a capacidade de discriminação entre os equipamentos, além da potência ativa, o estudo adicionou como característica o pico de consumo para o evento de transitório, bem como o período que o mesmo leva para estabilizar (apresentado em [59]), diferente de [53, 54] que observou essas propriedades, mas empregou somente a última com o intuito de melhorar a capacidade de reconstrução de energia. A grande contribuição foi uma abordagem por otimização para a identificação dos modelos de carga a serem encontrados que trabalham como uma adaptação do algoritmo de *Viterbi*, que podem ser resumidos da seguinte maneira:

- i. criação das FSM por algoritmos genéticos para reduzir o tempo de otimização;
- ii. em seguida essas são otimizadas por Algoritmo Genético de Estratégia Evolutiva (ES, em inglês) para obter os parâmetros da FSM (tempo e consumo em cada estado);

iii. e finalmente, os modelos utilizam lógica *fuzzy*, permitindo que dois ou mais modelos sejam criados para um mesmo distúrbio na rede, sendo depois escolhido o modelo que melhor se aplica.

Tratar-se-ão de detalhes das técnicas aplicadas pelos autores por três motivos:

- a técnica teve bons resultados apesar de utilização de pouca informação, podendo ter melhores resultados quando alimentada com mais informação e, por isso, sendo um possível caminho a ser percorrido;
- outros autores [79–84] se basearam nessas ideias;
- os artigos não são de compreensão trivial, em especial para a elucidação da adaptação do algoritmo de *Viterbi*¹³.

A estratégia começa com o agrupamento dos dados em centroides, limitando-se a degraus acima de um limiar mínimo de potência (valores aplicados de 50 W em [58] e 80 W [59]). As abordagens em [58, 59] se basearam no agrupamento utilizando lógica *fuzzy*, mas na última referência [60], além desse método, cita-se o emprego de SOM para essa etapa. A fim de reduzir a complexidade do problema (a estimativa de eventos é de 16.000 por dia), o autor desconsidera os centroides com poucas ocorrências, limitando-se a identificar apenas equipamentos com padrões recorrentes.

Na primeira abordagem [58], Baranski e Voss segmentaram a etapa de modelar os equipamentos. A primeira modela as C3s simplesmente encontrando pares de centroides no espaço. Para validar os modelos C3s encontrados, é gerado a matriz de correlação cruzada utilizando o estado de operação para os equipamentos em cada instante de tempo. Se o equipamento i e o equipamento j forem na realidade uma FSM, espera-se $r_{ij} \approx 1$, onde r_{ij} indica a frequência de operação do equipamento j quando i está operando, associando, assim, j com a operação i (o corte utilizado é de 0,8). Já para as C2, são criados todos os modelos que juntos somam aproximadamente zero e seus centroides tem frequência de eventos também próximos. Esses modelos são então validados temporalmente, e junto com as C3s são comparados com um possível catálogo antigo a fim de atualizá-lo. Em seguida, uma rede neural é treinada com os padrões encontrados para os equipamentos (o autor cita como exemplo: tempo médio de consumo, consumo médio, número de estados) para encontrar esses padrões na residência.

Essa abordagem é aprimorada em [59, 60], que ao invés de encontrar todos possíveis modelos de FSM, faz a otimização das possíveis máquinas através de algoritmo genético. Nessa abordagem não há a discriminação para a criação de C2 ou C3, a abordagem única utiliza N_{ap} (o número de equipamentos deve ser maior que o número de centroides, no entanto, não é especificado um bom valor a ser utili-

¹³Para os leitores que desejarem se aprofundar, recomenda-se a leitura de [79] antes dos artigos de Baranski e Voss.

zado) FSMs para os quais e o algoritmo genético fica encarregado de alterar valores binários em uma matriz \mathbf{X} representando se um determinado centroide pertence, ou não, à FSM. É possível que um mesmo centroide pertença a mais de uma FSM. São utilizados três critérios para otimização:

- minimização do valor absoluto de potência da soma dos centroides pertencentes a FSM;
- o item anterior, mas levando em conta a frequência de eventos em cada centroide;
- e minimização do número de centroides em cada FSM (priorizando equipamentos com menos estados).

Uma adaptação do algoritmo de *Viterbi* é utilizada para encontrar os modelos de FSM. Com os modelos resultantes, é criada as sequências de estado para elas supondo que as mesmas são C2a. Mais precisamente, os autores consideram que as sequências de estados devem ser recorrentes com seus parâmetros em uma área limitada dentro de seu valor esperado (os autores citam dois exemplos de parâmetros: tempo de duração no determinado estado e a capacidade de reconstrução de energia para o consumo estimado nos estados da FSM em relação ao consumido nos caminhos percorridos; mas não dá detalhes de quais empregou) e apenas visitados uma vez em cada ciclo. Para isso, é realizada uma otimização em dois tempos. Primeiro, encontra-se o melhor caminho para aqueles que obedecem as restrições (consumo de potência positiva e permanência em um estado por um tempo não muito longo), juntando os estados da máquina em um caminho de operação com a melhor qualidade em relação aos parâmetros escolhidos. Os parâmetros podem ser iniciados com os valores da mediana para todos os eventos acoplados a FSM. A qualidade é avaliada pela entropia de *Shannon*, 4.13. Isso é repetido iterativamente utilizando ES que irá alterar os parâmetros até a convergência da qualidade.

$$Q_{shannon} = -\Delta e_i \log |\Delta e_i| \quad (4.13)$$

Os melhores caminhos operativos para as FSMs ainda precisam ser resolvidos quanto aos centroides que pertencem a mais de um equipamento. Para isso, [60] cita resumidamente um algoritmo de força bruta que irá investigar para cada sobreposição qual caminho tem a melhor qualidade.

Os autores revelam que o método necessita de 5 a 10 dias para encontrar os modelos dos equipamentos típicos, enquanto dados diários são suficientes para atualizar o catálogo de equipamentos detectados em cada residência. Os resultados mostram que os equipamentos de maiores consumo, como geladeira, aquecedor elétrico (de fluxo) e fogão podem ser detectados com eficiência.

- *NILM de Arquitetura Distribuída (dNILM, em inglês)*

A abordagem de Baranski e Voss é a base empregada para o trabalho de Bergman,

Jin, Juen, Tanaka, Gunter, e Wright [79], contando com a mesma sequência de criação através de algoritmo genético e otimização das FSM. Diferenças podem ser notadas apenas para a técnica de agrupamento, que não utilizam as informações de tempo nem o pico atingido no transitório, contudo, podem utilizar a Q se o medidor da residência realizar essa medida. Além disso, o agrupamento é realizado em tempo real, atualizando a média e desvio padrão de cada centroide conforme os eventos ocorrem. Se o evento não for atribuído a nenhum centroide, um novo centroide é criado. Ainda, o método utiliza outra otimização para fazer a identificação do equipamento que alterou o estado operativo, um algoritmo de otimização de mochila (do inglês *Knapsack problem*), enquanto a abordagem anterior citou a utilização de redes neurais na referência [58]¹⁴.

Entretanto a maior contribuição do trabalho em questão é uma nova arquitetura, distribuída, para o NILM. Nesse caso, diferente dos medidores eletromecânicos disponíveis para Baranski e Voss, o trabalho opera com medidores inteligentes (os medidores das redes inteligentes descritos na Subseção 3.2.4) para a coleta de dados. Os medidores inteligentes também servem como pontos de processamento local, mas devido às limitações de processamento, apenas a detecção e identificação dos eventos é realizado no mesmo. Por isso, a geração das FSMs é realizada em uma central, com maior capacidade de processamento, no qual este envia os eventos de transitório para que aquela os processe e retorne um catálogo com as FSMs e seus padrões a serem encontrados. O catálogo é chamado de tabela estática. Assim, o medidor fica encarregado apenas de comparar, localmente, os distúrbios encontrados com o catálogo, identificando assim os estados operativos dos equipamentos. O estados operativos de cada equipamento é chamado de tabela dinâmica, e é preenchida por um algoritmo adaptado para otimização do problema da mochila. A etapa de criação da tabela estática, chamada de aprendizagem, é realizada devido a critérios do controlador (iniciação pró-ativa) ou do medidor (reativamente). No primeiro caso, o controlador atualiza as tabelas do medidor se as mesmas expirarem. Já o medidor inteligente pode requisitar um novo treinamento de acordo com um dos critérios:

- a diferença absoluta entre a soma da demanda real e estimada está superior a um patamar;
- uma FSM muda de estado frequentemente, onde os patamares aplicados para determinar se a mudança de estado é frequente dependem do equipamento (é mais aceitável observar mudanças frequentes no ar condicionado ou aquecedor do que em um carregador de bateria veicular);

¹⁴ Baranski e Voss praticamente não levaram o assunto em consideração, a descrição dos artigos desses autores gira em torno do tema de modelagem dos equipamentos, mas tirando a informação aqui citada, não há outra informação sobre como se tratou o reconhecimento dos estados operativos através dos modelos.

- mais de um determinado número de FSMs alteram de estado em um único evento.

Uma das dificuldades do projeto está em ajustar o fluxo de dados. A referência considera armazenar os dados em períodos de maior atividade nas residências, enviando as alterações de estado posteriormente conforme a rede de comunicação não estiver congestionada. Diversas outras considerações são feitas em relação ao processo de aprendizagem.

Os resultados reportados são em comparação com um NILM centralizado. O trabalho reservou-se a detectar equipamentos com consumo superior a 1000 W. O NILM centralizado recebe uma tabela estática otimizada para todo o período, enquanto o dNILM recebe uma tabela treinada para o primeiro dia, podendo atualizá-la de acordo com o critérios anteriormente citados. As diferenças de acurácia entre o dNILM e o NILM centralizado ficaram entre 60% e 90%.

- *Otimização dos estados operativos*

O problema para a otimização da tabela dinâmica — construção temporal dos estados operativos dos equipamentos — foi abordado por Egarter, Sobe, e Elmenreich [85] (aparentemente sem conhecimento do trabalho de BERGMAN *et al.*). Ao invés de um algoritmo de força bruta adaptado, aplicou-se um algoritmo genético para realizar a otimização do problema da mochila. O artigo limitou-se a estudar a performance do algoritmo em resolver o problema da mochila e, por isso, considerou-se que são conhecidos os momentos de transição e os consumos de cada equipamento.

A referência utilizou simulações de 2 horas, gerando aleatoriamente o tempo de operação dos equipamentos e suas potências. Foram simuladas diversas condições, variando o número de equipamentos, transições, a presença de ruído e de equipamentos desconhecidos. As observações foram:

- quanto maior o número de eventos de transitórios (N_{ap}) e número de equipamentos (N_{eq}), menor é a eficiência de detecção. Casos com pequenos números de N_{ap} e N_{eq} obtiveram 100% de eficiência de detecção;
- o algoritmo se comportou bem na presença de ruído, onde houve pouca deterioração na eficiência de detecção;
- no entanto, a presença de equipamentos desconhecidos ou equipamentos com sobreposição de potências deterioraram bruscamente a performance do algoritmo.

1.4 Potência ativa e estatística de uso

Também com o objetivo de possibilitar a impregnação do método, mas no caso para os EUA, Zeifman propôs um novo método [80–84] para abordar o tema em cima de mostradores de energia domiciliares. Esses mostradores normalmente disponibilizam apenas a potência ativa, com uma taxa de amostragem na faixa de 1 Hz. A fim de melhorar a capacidade de desagregação do consumo que é limitada pela

pouca informação oferecida pela sistema de aquisição de dados, o estudo empregou o conhecimento prévio de uso dos equipamentos através de uma abordagem estatística para melhorar a capacidade de seu NILM. O autor observa que os estudos [56, 57] também utilizaram a estatística de uso, servindo, provavelmente, de inspiração para o seu trabalho. Será tratado da abordagem desse autor com detalhes pois os trabalhos fornecidos pelo mesmo parecem serem a mais próximas de possibilitar o NILM em um programa de EE em larga escala.

Devido à ausência da potência reativa, a sobreposição dos centroides ocorrerá com frequência maior, aumentando as ocorrências de erros decorrentes de C6. O autor criou o método Algoritmo de *Viterbi* com Transições Esparsas (VAST, em inglês) para empregar a informação estatística no processo de discriminação, Etapa 3. Diferente das outras adaptações do algoritmo de *Viterbi* que se encontram no levantamento bibliográfico [49, 50, 58–60], onde o objetivo é modelar as FSMs presentes na rede, a ideia, aqui, é modelar os estados operativos dos equipamentos como sendo dependentes, além de seu próprio estado operativo, no estado operativo de outros equipamentos. Indo além, criar essa dependência para auxiliar aonde a informação é precária, ou seja, quando há sobreposição de potência. O estudo observa que, normalmente, isso só ocorre para trios de equipamentos, ou seja, supondo que os equipamentos da residência estejam distribuídos em ordem crescente de consumo, o equipamento i só deverá ter sobreposição de consumo com o equipamento $i - 1$ e $i + 1$. Assim, ao invés do algoritmo de *Viterbi* de força bruta que tornaria a aplicação inviável conforme o número de equipamentos (ou estados operativos) aumentasse, já que todos os estados dos outros equipamentos podem influenciar no equipamento i , somente será utilizado os estados dos equipamentos $i - 1$ e $i + 1$ para determinar o seu estado operativo. Com isso, reduz-se a complexidade da versão original de *Viterbi*, que é exponencial conforme o número de estados possíveis, para uma complexidade linear.

Para tornar mais simples a tarefa, considera-se que os equipamentos $i - 1$ e $i + 1$ também serão independente entre si, sendo necessário, assim, determinar para cada equipamento apenas dois pares de cadeias de *Markov* (exclusive o equipamento de menor e maior consumo, que terão apenas uma cadeia): $\{i - 1, i\}$ e $\{i, i + 1\}$. O preenchimento das matrizes com as probabilidades de transição é feita com a informação obtida *a priori*, aplicando as estatísticas de tempo ligado e desligado, distribuições dos momentos em que os equipamentos trocam de estado operativo no dia e as distribuições dos eventos de transições para cada equipamento. Outras características podem ser utilizadas, tais como: as distribuições para as informações de tempo e amplitude para o pico de acionamentos (influenciado por [58–60]) e características “finas” específicas de cada equipamento (o autor deu exemplos em [84] e serão cita-

dos mais adiante neste tópico). Essas probabilidades podem ainda ser melhoradas considerando distribuições para períodos diurnos ou noturnos e sazonalidade.

Como o ciclo de observações sofre de alterações como falta de transições (no caso de perda de alvo para um evento de transitório) ou a presença de intrusos (tanto falso positivo, ou quanto a presença de transitório é causado na cadeia $\{i - 1, i\}$ por um equipamento $i + 1$, por exemplo), é necessário que a matriz de probabilidade considere mudanças não esperadas para garantir a solução do problema. Por exemplo, no caso de um intruso, é necessário dar a possibilidade dos equipamentos continuarem no mesmo estado reconhecendo sua presença, enquanto também é necessário dar a possibilidade de um equipamento i , que estava desligado, modificar seu estado para ligado, enquanto o evento de transitório foi de desacionamento, caso que ocorreria excepcionalmente para um evento de transitório perdido de acionamento para o equipamento i e o evento sendo julgado ser intruso. Obviamente essas probabilidades serão pequenas, sendo utilizadas constantes para preencher essas possibilidades. Essas constantes podem utilizar taxas esperadas de perdas de alvo, falso positivos ou intrusão por outro equipamento, entretanto a determinação dos valores parece estar deixada em aberto pelo autor.

O algoritmo calcula, então, a probabilidade máxima de observar um estado após cada evento de transitório para todas observações das cadeias. Ao chegar no final delas, é escolhido a configuração com maior probabilidade e então se rastreia a sequência de transições no sentido oposto até chegar no estado inicial dos sistemas (algoritmo de *Viterbi*¹⁵). A fusão para as duas cadeias de *Markov* de um mesmo equipamento i é feita utilizando o conceito de máxima verossimilhança, empregando a configuração de maior probabilidade final para esse equipamento.

Para descobrir os equipamentos e suas estatísticas, o algoritmo realiza a coleta de dados por um longo período de tempo (ex. duas semanas) e, em seguida, realiza o agrupamento dos dados por centroides (foi utilizado ISODATA, mas o próprio autor pretende empregar outras técnicas). A fim de se corrigir possíveis problemas com os centroides, são feitas operações de separação e agrupamento dos mesmos de acordo com regras que indicam a necessidade das mesmas. Os centroides finais positivos e negativos são casados com a mesma abordagem utilizada por Hart (ver o passo 3 desse autor na pp. 60, limitando-se a parte para C3). A montagem das cdfs com as características de tempo é realizada através do modelo empírico — constrói a cdf através dos dados amostrados, sem ajuste —, enquanto para as pdfs de tempo e potência, ajusta-se a distribuição através da mistura de duas componentes de distribuições de Laplace¹⁶.

¹⁵A última versão do código utilizou um modelo adaptado para a cadeia de *Markov* de segunda ordem.

¹⁶A primeira versão realizou o ajuste com duas componentes Gaussianas.

Por fim, após o processo de levantamento de estatística, o algoritmo é aplicado em tempo real, limitando a janela de dados para um tamanho aceitável, citando como exemplo 1 dia. O tamanho de janela pode ser ajustado experimentalmente. A base de dados utilizadas para a estatística também é atualizada conforme mais dados são recolhidos. Assim, o método não necessita de treinamento, identificando os modelos de carga e suas propriedades. O autor considera depois comparar a informação encontrada com as estatísticas de uso e consumo típicas para identificar o equipamento, mas havendo comunicação com o usuário, é possível que o próprio realize a identificação, dando a comparação apenas como sugestão. Outra consideração é trabalhar com trios de cargas ao invés de duplas, já que alguns casos essa configuração parece ser mais realista devido à sobreposição de vizinhos aquém daqueles imediatamente subsequentes, reduzindo a presença de intrusos nas cadeias de *Markov*.

O autor empregou dados simulados e reais para provar a eficácia de seu algoritmo, onde foi comparada a capacidade do mesmo em relação a um classificador *Naïve Bayes*, obtendo resultados em termos de F^i (4.11) na ordem de 0,83 a 0,97 em relação à medição paralela, em geral superiores ao classificador de *Bayes*. Nos dados simulados, o autor demonstra que as outras abordagens normalmente empregadas para 1 Hz são deficientes para as configurações de sobreposição de consumo, enquanto sua versão consegue reconstrução praticamente perfeita. Entretanto, atualmente o algoritmo só atende equipamentos modelados por C3.

Já no estudo mais recente [84], o autor não empregou o VAST. Como nas abordagens [56, 57], limitou-se a identificar os equipamentos de maior uso-final. No caso, os equipamentos de interesse foram: *a*) ar condicionado; *b*) aquecedores de ambiente; *c*) aquecimento de água doméstica; *d*) iluminação; *e*) geladeira; *f*) secadores de roupa elétricos; *g*) equipamentos eletrônicos; que representam em média 80% do consumo residencial, no caso, para os EUA. Nesse configuração, utilizou-se o conceito de máxima entropia para selecionar modelos estatísticos mais adequados para os dados. Uma vez que a média e alcance das variáveis são conhecidos, as distribuições Beta são as mais indicadas por esse conceito. Foram feitas considerações em relação a iluminação, que tem dependência em relação aos ambientes de uso. Ainda, para melhorar a performance, adiciona-se como característica as assinaturas específicas dos equipamentos. O exemplo é a televisão em comparação com uma lâmpada, onde aquele varia o seu consumo conforme flutuações na imagem e som, enquanto este tem o consumo bastante estável. Essas características “finas” dos equipamentos podem ser modeladas matematicamente e empregadas em conjunto com o conhecimento prévio de utilização dos equipamentos. Já para o caso dos secadores de roupa elétricos (uma FSM), foi utilizado duas distribuições Beta, uma

para cada estado.

Os autores utilizaram o classificador *Naïve Bayes* com base nas distribuições conjuntas resultantes obtidas para os sete uso-finais indicados, adicionados de uma classe “outros”. Os resultados empregaram a F^i , obtendo valores de acurácia para os uso-finais mais desafiantes na ordem de 0,65 e 0,70, sendo os mesmos os itens *d)* e *g)*, respectivamente. Já quando utilizando a informação de características finais, essas mesmas acurácias se elevam para 0,92 e 0,90. Para os outros equipamentos, os resultados ficaram na margem de 0,9, atingindo a melhor marca para os aquecedores de água doméstica: 0,99.

2. Medição com Alta Amostragem

A abordagem para os NILMs operando com aquisição de dados obtidos com alta amostragem possibilita a obtenção de características possivelmente mais discriminantes devido à alta resolução. A maior granularidade na amostragem permite o acesso aos harmônicos e formas de onda do sinal, chamadas de características microscópicas. Vale ressaltar aqui que a utilização de taxas de amostragem alta no sistema de aquisição não implica na extração de características microscópicas. Por exemplo, ao extrair a diferença de consumo de potência ativa entre o pré e pós transitório com 100 MHz, isso não define essas características como microscópicas caso elas possam ser obtidas com frequência.

Devido à maior amostragem, é inerente que os autores, em muitos casos, buscam empregar técnicas de processamento de sinais para o tratamento dos dados, onde a bem conhecida Transformada Rápida de *Fourier* (FFT, em inglês) foi a primeira abordagem. Outras abordagens vão além da FFT, procurando outras maneiras de representar a informação, em alguns casos trabalhando até mesmo com a informação crua, sem tratamento. A criatividade dos autores leva a tona uma gama de características para descrever o sinal e buscar através de diversas técnicas as assinaturas dos equipamentos. As características e técnicas empregadas pelos diversos autores para essa amostragem é tratada a seguir.

Antes disso, considerações quanto aos requerimentos mínimos de amostragem deve ser realizada [61]. Quando explorando as características no domínio da frequência, o teorema de *Nyquist-Shannon* formula que a frequência de amostragem deve ser no mínimo duas vezes superior à frequência do maior harmônico a ser estudado. Já no domínio do tempo, uma regra de boa-prática [61] formula que a amostragem deve ser no mínimo de cinco vezes a frequência de interesse (no caso, o mínimo necessário para representar a fundamental no Brasil é uma amostragem em 300 Hz). Entretanto, é preciso ter em mente que certos distúrbios só poderão ser observados em frequências maiores, em especial a caracterização dos regimes

transitórios, sendo necessários nesses casos frequências bem maiores. As frequências utilizadas pelos autores variaram nas ordens de ~ 100 Hz até ~ 1 MHz.

2.1 Harmônicos e FFT

A decomposição através de FFT fornece informação além daquelas obtidas por P e Q, cuja única informação presente é das componentes fundamentais e sua defasagem. A informação presente na decomposição auxilia na discriminação de cargas não-lineares, separando as mesmas entre si e das cargas lineares. Por outro lado, é importante notar que a decomposição não fornece mais informação discriminante além da informação já contida em P e Q para desagregar cargas lineares.

- *Harmônicos para identificação de C4*

O grupo que trabalhou com o intuito de levar o NILM para o setor comercial e industrial [51, 52] utilizava amostragem em frequências superiores a 8 kHz para construir o envelope espectral do sinal [62]. O grupo se referia dessa maneira ao envelope pois ele era construído através da resposta da FFT. Como foi dito durante o levante desses trabalhos, o grupo subdividia os transitórios em seções-v que eram utilizadas para classificar os equipamentos.

Uma vez que a decomposição harmônica já estava sendo extraída, os trabalhos posteriores procuraram empregar essa informação para tratar um problema presente em maior escala nesses setores, as C4s. Os trabalhos de LEE *et al.* [63], WICKAKOOL *et al.* [64], SHAW *et al.* [65] avaliaram como identificar possíveis maneiras de reconstruir o consumo de cargas variáveis em cenários previamente conhecidos. Um dos cenários incluiu inclusive a aplicação do NILM em um carro [65]. Foi observado nas C4s estudadas a presença de correlação entre sua demanda e os harmônicos de potência aparente. Assim, é ajustado uma função polinomial utilizando Mínimo Erro Quadrático (MSE, em inglês) para determinar a sua mudança de consumo. A função polinomial é utilizada e comparada com as envelope espectral das seções-v para determinar se a mudança foi causada por alterações no estado operativo da C4.

Apesar do método se comportar bem para as aplicações propostas, há a necessidade de fazer o levantamento caso a caso, sendo interessante apenas para a aplicação no setor comercial e industrial de grande porte, onde dificilmente será possível aplicar um NILM genérico devido as diversas peculiaridades das cargas presentes. Além disso, não se sabe como a presença de outros equipamentos irão afetar a correlação, podendo limitar sua aplicação.

- *Extração harmônica ciclo a ciclo*

A abordagem de SRINIVASAN *et al.* [66] alterou a estratégia do NILM ao não realizar a extração dos eventos de transitório. Ao invés disso, ele realiza diretamente para cada ciclo a desagregação da informação de consumo por equipamento.

Para isso, utilizou-se técnicas de aprendizado de máquina (Redes Neurais Artificiais (RNAs) com as arquiteturas MLP e RBF, além de SVM) treinadas com a informação das componentes harmônicas ímpares até o 15º harmônico. O caso analisado continha oito equipamentos. Em cada possível combinação operativa desses equipamentos ($2^{N_{ap}} = 256$), obtiveram-se dados para treinar e testar as técnicas aplicadas. Nessa configuração, o método obteve reconstrução praticamente perfeita para a MLP e RBF. Em seguida, para melhorar a aplicabilidade dos métodos, utilizou-se as amostras coletadas com apenas um dos equipamentos ligados por vez que foram, então, somadas para reproduzir os outros estados operativos. Nessas condições, houve deterioração para quatro equipamentos para taxas em torno de 70%, enquanto os outros equipamentos continuaram com taxas em próximas a 100%, tanto para a MLP e RBF (com algumas discrepâncias entre as duas). Também se testou condições com 10 equipamentos. Para esse caso, como seriam necessárias medições em 1024 situações, utilizou-se o método de obter apenas as operações de cada equipamento sozinho, e somá-las para obter as outras configurações. Outras configurações testadas incluíram adição de ruído e teste para rede com topologia trifásica, onde se apresentou que o método é robusto para essas configurações. O SVM não teve respostas compatíveis com as da MLP e RBF para as diversas configurações testadas.

Todavia, conforme o número de equipamentos cresce, a quantidade de combinações necessárias para treinar a rede cresce exponencialmente. Uma residência com cerca de 30 a 50 equipamentos (valores típicos citados por [43]) necessitariam do treinamento para números na ordem de 10^7 a 10^{15} configurações — isso sem considerar possíveis estados que serão adicionados por equipamentos C2s —, onde possivelmente diversas dessas configurações seriam similares e, por isso, o método não conseguiria reproduzir as capacidades para os exemplos citados. Além disso, é necessário saber *a priori* todos equipamentos pertencentes na residência, assim como o autor não levantou como as respostas seriam afetadas devido à presença de equipamentos não-conhecidos. Mais adiante, também é necessário estimar o consumo por equipamento, assunto deixado em aberto pelo autor. Para isso, é possível realizar a identificação de transitórios em cima da informação de discriminação, como considerando uma janela para a qual se o equipamento mudar de estado operativo por um tempo maior àquele da janela, então se confirma a mudança operativa e calcula o consumo desse equipamento no novo estágio através da diferença pós e pré-transitório. Possíveis adaptações podem ser feitas para obter melhor resolução em energia. Uma abordagem mais elegante seria realizar as duas etapas em paralelo e comparar as informações obtidas em cada uma para validar se houve realmente a mudança de estado. Já com o intuito de tornar essa tarefa trivial, pode-se evitar a utilização da Etapa 2 e simplesmente empregar um valor de consumo estimado para

cada equipamento.

- *Ruído espectral de tensão*

Patel, Robertson, Kientz, Reynolds, e Abowd [67], Gupta, Reynolds, e Patel [68] encontraram características com capacidade discriminantes *sui generis*. A metodologia utilizada permite diferenciar equipamentos iguais mas em diferentes cômodos. De acordo com os autores, em alguns casos é possível identificar qual interruptor de uma lâmpada *three-way* teve seu estado alterado. Para isso, ao invés de utilizar o consumo agregado, os autores mudaram o conceito de como realizar a aquisição de dados. No caso, seu sistema pode ser conectado em qualquer tomada da residência, aonde serão analisados a EMI causado por chaveamentos, sendo necessário medir somente a *tensão*.

As características internas do equipamento, bem como a configuração da rede elétrica entre o equipamento e o medidor, alteram em como o ruído será observado. Os ruídos podem ser gerados continuamente ou apenas durante os eventos de transitório. Exemplos de equipamentos que geram ruídos contínuos são: equipamentos com funcionamentos a base de motores (como secadores de cabelo, ventiladores) que irão gerar ruídos de tensão em frequências síncronas àquelas da rede causados pelo chaveamento das escovas dos motores (os equipamentos desse tipo que não tiverem escova, não vão produzir *emi* continuamente); e os equipamentos eletrônicos com fontes baseadas a chaveamento, nesse caso gerando ruído síncrono ao oscilador de sua fonte. Já eventos de ruído EMI podem ser observados ao colocar um equipamento na tomada, alterar o estado de um interruptor ou ligar uma televisão.

Na primeira abordagem [67], o objetivo foi encontrar os ruídos transitórios de chaveamentos na rede. São gerados 2048 pontos igualmente espaçados com os resultados da decomposição da FFT (apenas magnitude) no alcance de 0 a 50 kHz. Essa informação é repassada para uma janela deslizante que calcula a média para 1 μ s de dados. Com os resultados das janelas, é tirada a distância Euclidiana em relação a janela anterior, e se esse valor superar um limiar considera o início de um transitório. O transitório será registrado até a distância Euclidiana ter outra mudança drástica. A resolução de 1 μ s é importante pois alguns transitórios tem duração de apenas alguns poucos μ s. Tendo o transitório isolado, obtém-se a média para os 2048 pontos em relação ao número de amostras temporais. Essa informação alimenta uma SVM. A capacidade de isolamento de transitórios (taxa de detecção apenas para a Etapa 2) obtida foi entre 88% a 98%, embora a taxa de falsos positivos não tenha sido informada. Os acertos para a quando considerando a Etapa 3 ficaram entre 85-90% quando a SVM é treinada com 5 instâncias para cada evento. Caso seja coletado apenas 2 instâncias por evento, há uma deterioração para cerca de 80%.

Já a segunda abordagem tratou da abordagem para ruídos gerados continuamente na rede [68]. Os autores perceberam que os transitórios contínuos não são dependentes da configuração da rede, onde os equipamentos terão as mesmas características independente das residenciais e tomadas em questão. Nesse caso, o espectro analisado de *Fourier* é de 36-500kHz para a mesma quantidade de pontos igualmente espaçados. A mesma ideia de janela móvel e corte através de um patamar é realizada para essa abordagem, no entanto, a frequência de geração de vetores é de 244 Hz, sendo utilizados 25 deles para calcular a média da janela deslizante (a frequência resultante é bem menor que aquela utilizada para estudar eventos de transitório). A extração de características é realizada em cima da diferença entre o espectro do ruído atual e anterior, aonde são ajustados uma Gaussiana. As características são a média, amplitude e variância, que são comparadas utilizando a técnica de vizinho mais próximo. A acurácia reportada é de 91,75%.

Apesar das boas capacidades, ainda há uma série de dúvidas quanto a capacidade de aplicar essa metodologia nos NILMs. Em [68], os autores levantam uma série de limitações quanto à abordagem por detecção de transitórios, devido ao custo computacional necessário para avaliar ruído transitório, a necessidade de treinar para cada eletrodoméstico, a dependência das características na configuração da rede elétrica do consumidor e os eventos tem deposição de energia pequena devido a seu longo espectro de frequência. Percebe-se que a grande sensibilidade das características presentes nos transitórios traz um problema inverso àqueles carregados pelas C6s: ao simplesmente trocar o equipamento de tomada altera como o mesmo é visto pelo NILM. Como levantado por [43], a segunda abordagem é mais robusta, no entanto, diversas questões ficam em aberto. Ela é limitada para equipamentos que tenham chaveamento, por exemplo, fornos e secadores de roupas elétricos não puderam ser identificados. Os autores não consideraram como a presença de EMI nos vizinhos, em especial para o caso de apartamentos, irá interferir no processo de detecção. Na construção do catálogo, é necessário considerar todas as mudanças de fontes chaveadas que forem realizadas pelos fabricantes, assim como as diferenças entre as fontes utilizadas por eles. A sobreposição de assinaturas EMI também irá ocorrer, sendo necessário tratar esse problema. Enfim, os autores apenas se preocuparam em identificar os equipamentos, não sendo obtido o consumo desagregado dos mesmos.

- *Fourier como função de ajuste para Vizinho mais Próximo*

Os trabalhos de [29, 75, 76] buscaram apenas reproduzir as eficácias dos NILMs presentes na literatura, não tendo como objetivo apresentar melhorarias para as técnicas presentes nos NILM, mas apenas fornecer essa tecnologia para o consumidor ter acesso em baixo custo para uma ferramenta que o permita compreender suas despesas com energia. O projeto começou desde o sistema de aquisição de dados, onde se observou problemas com medidores disponíveis no mercado estadunidense

que apresentavam discrepâncias de 10-20% entre si, enquanto a taxa reportada era de 3%. O seu sistema de aquisição de dados atual coleta amostras na frequência de 100 kHz utilizados para gerar características em 20 Hz. Devido às dificuldades para rotular os momentos exatos para o treinamento das técnicas, eles criaram um detector *wireless* de eventos de transitório, diretamente colocado no cabo do equipamento.

Ele utiliza um modelo probabilístico para a detecção de eventos alimentando um discriminador de vizinho mais próximo utilizando função de ajuste como a FFT na sua melhor configuração. As eficiências de detecção obtida estão em torno de 67% e 100%. A discrepância em energia do NILM para o valor total de energia foi de 14,8%.

2.2 Transformada *Wavelet* (TW)

Uma das diferenças mais marcantes entre a FFT e a TW é a possibilidade que a última oferece de resolução tanto de frequência como temporal. Para isso, a TW reduz o tamanho da janela conforme há o aumento das frequências para permitir resolução temporal. A decomposição é realizada em funções chamadas *wavelets* ao invés de senoides. A TW tem diversas aplicações em Sistemas de Potências, o levantamento realizado por [98] em 2002 mostra que as aplicações mais comuns eram proteção, qualidade de energia e detecção de transitório. A larga aplicação da TW em detecção de transitório em sistema de potência mostra que elas podem ser utilizadas para, no mínimo, realizar a Etapa 2 do NILM, sendo um substituto potencial para as métricas simples adicionados de um corte linear comumente aplicadas pelos NILMs nessa etapa.

Indo além, o levantamento de [43] encontrou um estudo (Chan, So, e Lai [70]) realizado em 2000 que mostrou a capacidade de aplicação da TW para a identificação de cargas não-lineares. Entretanto, esse estudo não se aprofundou no tema, apenas levantando a capacidade da TW de diferenciar três equipamentos, seguindo uma ideia similar àquela explicada para SRINIVASAN *et al.*¹⁷. Para isso, explorou-se os níveis de detalhes da TW discreta para todas as combinações operativas dos três equipamentos, salientando que essa informação era discriminante e seria possível aplicar RNA para discriminar essa informação, ainda que isso não tenha sido realizado.

Um trabalho no Brasil, realizado por Rodrigues [88], seguiu exatamente essa metodologia. A taxa de amostragem empregada foi de 15.360 kHz, onde se testou duas configurações, uma aonde os equipamentos elétricos operavam isolados na rede e outra com as diversas combinações operativas. Os testes foram feitos com três e

¹⁷No caso, o estudo de CHAN *et al.* foi realizado depois.

quatro equipamentos e obtiveram capacidade de identificação de 100%. Em seguida, o trabalho mostrou a importância da TW quando em comparação com a FFT, onde a FFT obteve 90% mesmo utilizando 16 componentes do espectro (linearmente espaçado) de frequência, enquanto a TW foi decomposta para 7 níveis de detalhes. Entretanto, essa abordagem sofre das mesmas desvantagens referidas para o trabalho de Srinivasan, Ng, e Liew (ver pp. 73).

Apesar de no exterior ter sido encontrado apenas um estudo preliminar mostrando a possibilidade de utilizar TW [70], um outro trabalho brasileiro também seguiu o caminho do emprego de TW para a identificação de equipamentos. O trabalho de DO NASCIMENTO [87] utilizou uma amostragem em 256 Hz. É empregado um detector de eventos onde se calcula o incremento de corrente em relação a amostra passada, no caso de o valor ultrapassar o patamar de 32 mA, é criado um evento de transitório com 8 ciclos e um total de 2048 pontos a serem analisados.

Uma pré-classificação é realizada utilizando o *hardware* para equipamentos com valores específicos de ΔP e ΔQ . Caso não seja um dos equipamentos de simples discriminação, modifica-se a informação em uma série de etapas, nesta ordem: aplica-se Transformada de *Hilbert* no módulo da corrente, TW com cinco níveis e por último o Método de Burg, esse adicionado para suavizar e sintetizar a informação contida nos níveis de detalhes obtidos na decomposição do sinal pela TW. As características utilizadas são os picos para os níveis de detalhes do Método de Burg adicionados do fator de potência (FP).

A discriminação é feita em duas etapas, primeiro se seleciona um grupo genérico ao qual o equipamento pertence, para então selecionar o grupo específico. A seleção é feita através do vizinho mais próximo com métrica Euclidiana entre o valor no banco de dados e o resultado do espectro de Burg para os níveis de detalhe da TW. É realizada uma votação (em cada etapa) para cada nível de detalhe, o equipamento que tiver maior votação é o resultado da etapa do processo de discriminação.

Os resultados para acionamentos individuais reportados pelo autor foram de 99,7% para a seleção na primeira etapa e 88,8% na segunda. Foi realizado um ensaio para verificar o comportamento do algoritmo em casos empilhados, porém o mesmo foi um teste inicial, adicionando uma lâmpada incandescente — sem dinâmica de consumo — antes de acionar os outros equipamentos. Assim, não é possível deduzir a capacidade do método para a aplicação em condições reais, onde é necessário a robustez das técnicas aplicadas para ao acúmulo de diversos equipamentos, injetando ruídos e dificultando a desagregação da informação. O autor identificou, também, que é necessário trabalhar no detector de eventos transitórios, que talvez tenham sido o motivo para a deterioração da eficácia dos motores no teste empregado. Ainda, as condições operativas normais em uma residência podem revelar a

presença de falsos positivos, sendo também um levantamento necessário pelo trabalho. Por fim, os autores preocuparam-se apenas com a capacidade de identificação dos equipamentos, sendo necessário, pela perspectiva de aplicação do NILM, ainda o passo de transformar essa informação qualitativamente para termos de consumo.

2.3 Curvas I-V

As características geométricas das formas de onda para diversos equipamentos foi explorado pelos estudos de LEE *et al.* [71], LAM *et al.* [72]. Conforme [43], a contribuição foi a investigação de uma nova características, as curvas I-V¹⁸. As curvas são formadas ao distribuir as amostragens temporais de corrente contra tensão no plano cartesiano. No caso de cargas lineares, elas formam elipses, podendo transformar-se em uma reta caso o equipamento seja puramente resistivo. Nesse caso, o sentido de rotação da curva dependente se a carga é indutiva ou capacitiva, mais precisamente estando relacionada ao ângulo de carga. De acordo com a presença de harmônicos, pode ocorrer distorções nas elipses de modo que as mesmas enlacen a si mesmas em n pontos de interseção. Assim, o estudo propôs diversas características a serem extraídas dessas curvas, cada uma representando informações que permitam discernir as características dos equipamentos entre si.

São utilizados os valores normalizados para que todos os equipamentos tenham a mesma escala, removendo as discrepâncias causadas por equipamentos de diferentes marcas. A fim de estudar a capacidade discriminativa, a metodologia empregada foi a geração de dendrograma para compreender a taxonomia resultante e compará-la com aquelas normalmente obtidas quando utilizando variáveis padrão, no caso o autor considerou: valor eficaz da corrente em um ciclo; potência reativa; FP; número total de harmônicos ímpares; número total de harmônicos pares. O estudo também comparou com a taxonomia obtida ao realizar SVD da corrente em um período. Os primeiros 24 autovetores foram utilizados como representação da informação, contendo 99% da energia, para serem formados os agrupamentos.

Os resultados da separação taxonômica está na Tabela 4.2, (ver pp. 103, estando mais adiante simplesmente para agrupar as taxonomias com as do CEP-EL/Elektrobras em um único lugar neste trabalho). Constatou-se a presença de grupos mais representativos para as curvas I-V em relação aos outros dois casos, obtendo grupos bem definidos para os tipos resistivos, motores e eletrônicos. A representação com variáveis padrão teve dificuldade em distinguir entre os equipamentos eletrônicos. Ocorreu também a união de equipamentos resistivos, eletrônicos e equipamentos de operação através de motores. Já no caso da decomposição, os aglomerados formados permitiram a classificação dos equipamentos em relação ao

¹⁸Provavelmente os leitores já tiveram a oportunidade de observar essas curvas no osciloscópio, caso oposto, diversas curvas podem ser observadas na referência [72].

formato de suas ondas, no entanto, não é possível inferir como a afinidade dos equipamentos foi gerada.

2.4 Ondas sem processamento

A metodologia empregada por INAGAKI *et al.* [74] também opera ciclo a ciclo na rede, como o caso das abordagens de [66, 88], mas alterando a lógica quando em comparação com esses outros métodos. Ao tratar o problema como sendo de otimização, foi resolvido o caso do paradoxo da quantidade de configurações para o treinamento das técnicas supervisionada. O próprio autor observa que outros autores japoneses tentaram resolver o problema da mesma maneira — através de redes neurais — ressaltando a quantidade de dados necessários por essa abordagem. A técnica consiste-se em ajustar a melhor configuração das ondas cruas de corrente — sem pré-processamento — que se justaponha àquela observada pela medição centralizada. Isso é realizado através de Programação por Inteiros (tradução própria de *Integer Programming*), modelando tanto C1, C2 quanto C3, desde que se saiba da existência dos mesmos *a priori*. São obtidos 666 pontos por ciclo, resultando em uma frequência de amostragem de 40 kHz.

A abordagem foi avaliada em uma residência, onde os residentes realizaram suas tarefas normalmente, mas anotando os momentos em que eles utilizavam os equipamentos durante 6 dias. Os resultados mostraram que a fase 2 com 7 equipamentos obteve taxas de detecção na ordem de 96,8%, enquanto a fase 1 com 15 equipamentos, sendo eles diversas lâmpadas, obteve para essa acurácia, valor de 79,0%. É esperado para esse método a mesma dificuldade operativa das abordagens [66, 88], já que a quantidade de configurações possíveis cresce exponencialmente com o número de estados possíveis nos equipamentos da residência. Como o próprio autor observa, os erros na fase 2 foram causados por configurações onde os estados de consumo agregado são semelhantes, no caso “Torradeira + Secador (Apenas Ventilação)” \approx “Secador de Cabelo (Com aquecimento no fraco)”. Segundo o crescimento de equipamentos, é esperado que a ocorrência desses casos se torne maior. Os autores planejam trabalhar em uma contramedida para possibilitar a modelagem de C4, bem como conseguir uma maneira automática para unir equipamentos similares. Para isso, aqui se nota que utilizar um patamar de corte em 4.1 parece ser suficiente. Outra questão ressaltada por [43] é que a utilização das ondas cruas pode ter reduzido a eficácia da abordagem, considerando que a extração através de outras representações da mesma são geralmente mais robustas. Por outro lado, nada impede o emprego de Programação por Inteiros em cima de outras representações.

2.5 Aplicação de múltiplas técnicas

Os estudos Liang, Ng, Kendall, e Cheng [73, 77] trataram o tema de desagregação

do consumo de maneira metódica, definindo matematicamente diversas grandezas que muitos autores tentaram exprimir por palavras. No entanto, a maior contribuição foi mostrar como combinar as técnicas presentes na vasta extensão de abordagens já realizadas ao tema. A primeira questão já foi explicitada na Subseção 4.2.1 quando se referindo às C6s, sendo o ponto de similaridade entre as características. A seleção de características que não apresentem similaridade para os mesmos equipamentos possibilita que as técnicas de discriminação possam obter padrões relevantes para a solução do problema. O mesmo pode ser dito quanto a capacidade de complementação das técnicas empregadas na Etapa 3, os autores propuseram a razão de complementação (CR) em 4.14, sendo Z_a e Z_b dois discriminadores quaisquer empregados.

$$CR_{z_a, z_b} = \frac{P\{Z_a = false | Z_b = true\}}{P\{Z_a = true\}} \quad (4.14)$$

Naturalmente, essa abordagem requer mais um subpasso para combinar as respostas dos diversos discriminadores para produzir uma resposta única. Essa técnica é um problema conhecido tratado através pela ideia de um Mecanismo de Decisão por Comissão (CDM, em inglês), que dita as regras de como será tratado o valor de cada decisão para realizar a fusão de informação. [73], levanta algumas possibilidades:

Ocorrência Mais Comum (MCO, em inglês): escolha do candidato mais comum entre os membros da comissão. É o processo mais trivial de ser executada computacionalmente, realizando apenas a contagem de votos. Essa abordagem pode criar soluções não-únicas devido ao empate na votação;

Menor Resíduo Unificado (LUR, em inglês): seleciona o melhor candidato através da escolha daquele que obtém o menor valor LUR. Para isso, são definidos o resíduo individual (IR), onde $y_{(k|j)}$ representa a característica desconhecida, que é comparada com a assinatura conhecida $\hat{y}_{(k|(i,j))}$. A grandeza resíduo unificado (UR) é representada por 4.15b, onde M é o número de candidatos sendo consideradas. E, finalmente, o LUR sendo definido por 4.15c, onde Θ é a população de candidatos.

$$IR_{(i,j)} = \frac{\left(\sum_{k=1}^N y_{(k|j)} - \hat{y}_{(k|(i,j))}\right)^2}{\sum_{k=1}^N y_{(k|j)}^2} \quad (4.15a)$$

$$UR_i = \prod_{j=1}^M IR_{(i,j)} \quad (4.15b)$$

$$LUR = \min (UR_i | \forall i \in \Theta) \quad (4.15c)$$

O LUR considera assim, qual candidato tem o menor resíduo em relação às decisões tomadas pelas diversas técnicas. Esse CDM é mais pesado computacionalmente quando comparado com MCO, por outro lado, sempre permite soluções, e frequentemente únicas.

Estimativa de Máxima-Verossimilhança (MLE, em inglês): outra maneira é utilizar o levantamento de estatística. Ao simular diversas condições com casos conhecidos, é possível determinar a resposta mais provável para a k -ésima característica utilizando a j -ésima técnica, sendo ela determinada pela função $K(l, j)$. Assim, calcula-se a probabilidade marginal de ter ib como resposta verdadeira enquanto a j -ésima técnica com a k -ésima característica dão como candidato a resposta i_a (θ_{ia}). Um conjunto de verossimilhanças é dado por

$$\rho(ib) = \prod_{j=1}^M \prod_{l=1}^L p(\theta_{ia}, ib, j, l | ia = K(l, j), \forall i_b) \quad (4.16)$$

onde M e L são o número de características e técnicas, respectivamente. A máxima verossimilhança pode ser determinada através de

$$\Lambda(\rho) = \max\{\rho(ib) | \forall ib \in \Theta\} \quad (4.17)$$

A MLE é o método mais intensivo computacionalmente dos citados pelo autor.

Já [43] propõe a utilização de técnicas adaptando a teoria de *Dempster-Shafer* para realizar tal tarefa, e cita [99] como exemplo. Outra abordagem possível é o emprego de uma RNA.

Em [77] os autores criaram um simulador de Monte-Carlo, capaz de gerar situações de operação simultânea de equipamentos, ruído devido a dinâmica de carga (C5) ou fontes externas de ruído, dando uma capacidade maior para as possibilidades dos testes das técnicas a serem empregadas pelos NILM. Com base nos dados simuladores por esse gerador, eles mostraram que o uso de múltiplas características e múltiplas técnicas é benéfico para o desagregador, aumentando a eficiência em cerca de 10% quando comparado com a melhor resultado de qualquer outra técnica quando aplicada em isolado, obtendo valores na ordem de 90% para a η_{class} . A facilidade do simulador também permitiu uma série de outros levantamentos:

- a) as CDM são tanto incrementais em peso computacional quanto em performance, onde a MLE obteve a melhor performance para todos os casos simulados. A LUR, por outro lado, apresenta praticamente o mesmo resultado, sendo talvez uma abordagem com melhor custo benefício computacional;

- b) a opção com múltiplas técnicas se apresentou superior a ruído que as opções singulares;
- c) as simulações para o ar condicionado mostraram que sua presença deteriora a eficácia do desagregador;
- d) há deterioração conforme a maior demanda na rede, bem como uma presença maior de equipamentos operando;
- e) quanto maior for o incremento de potência em relação à demanda da rede, maior será a facilidade do NILM desagregar essa informação; e
- f) a análise por similaridade permitiu identificar os equipamentos com maior confusão.

O autor de [43], motivado pelas considerações de [77], realizou o estudo [81] mostrando a intensão de unir o seu método para baixas amostragens [80] com a combinação de múltiplos métodos para altas amostragens. Baseado nos estudos de seu levantamento bibliográfico, o autor utilizou amostragens de 500 kHz para trabalhar com análise de transitórios. O detector de transitório observa por mudanças estatisticamente significantes na forma de onda, que ao serem encontradas causam a coleta de diversas formas de onda antes e depois da mudança. Realiza-se a subtração entre o valor pós e pré transitório para a geração da diferença no formato da onda decorrente do transitório. Essa informação alimenta duas técnicas para discriminação: uma utilizando os componentes harmônicos e outra os dados crus para alimentar uma técnica multivariável. Calcula-se o valor de pdf para os dados de entrada em relações às pdfs obtidas experimentalmente para todos os equipamentos. A classificação final de cada algoritmo é o valor máximo de pdf.

Um teste simples foi realizado para comparar os dois métodos. Foram alterados estados operativos de quatro equipamentos, obtendo desagregação perfeita para o caso das ondas cruas, e um erro para a decomposição harmônica. O método combinado também obteve reconstrução perfeita. Em seguida, coletaram-se dados para 8 horas com os quatro equipamentos, obtendo novamente 1 erro para a decomposição harmônica e 100% para os outros dois casos. O autor retrata que seria ainda possível combinar essa informação com o VAST, que também apresenta saída probabilística. No entanto, parece que o autor descontinuou o projeto em virtude da melhor capacidade prática do método desenvolvido para baixas frequências que tem se mostrado capaz de cumprir com as necessidades de projeto para o NILM ser aplicado em um programa de EE.

3. Utilização de outros sensores

Diferente das propostas anteriores, alguns autores consideraram uma abordagem não centralizada. A ideia é empregar outros sensores aquém do medidor central na alimentação da residência para permitir maior capacidade de desagregação, seja por eles já estarem disponíveis, ou, em alguns casos, através de sua instalação. Mesmo que esses sensores já estejam disponíveis, dificilmente se terá acesso a sua informação sem intrusão da propriedade, constituindo assim uma abordagem intrusiva na maioria dos casos.

Já se referiu a um caso onde os autores consideraram a utilização de outros sensores. A abordagem de [51] para o setor comercial foi bem sucedida ao empregar variáveis de controle para possibilitar a obtenção do consumo de equipamentos C4.

Em [92], os autores mostram a intenção de seguir uma nova abordagem. A ideia é utilizar a informação de múltiplos sensores no ambiente, como iluminação, som etc. e correlacionar a informação por eles obtidas automaticamente com o consumo de equipamentos, permitindo, por exemplo, desagregar o consumo de uma lâmpada na cozinha com a de um banheiro devido à um sensor de iluminação nesse ambiente.

Um projeto na Suíça por [91], utiliza diversos submedidores de energia conectados via *wireless* a uma central. A ideia é reduzir o acúmulo de equipamentos no sinal a ser desagregado ao colocar os medidores mais próximos ao consumo, facilitando assim a identificação dos equipamentos. Cada módulo de submedição realiza medições em 3,2 kHz de tensão e corrente, que são utilizados para gerar a potência ativa, potência reativa e os harmônicos ímpares de corrente com índices de 1-11. Quando é detectado um evento de transitório, a informação é repassada para uma central, que procura no catálogo o equipamento que melhor se corresponde com o transitório. Reporta-se eficiência de detecção de 95% para submedidores com um único equipamento.

Uma outra abordagem bastante interessante realizada por Nambi, Papaioannou, Chakraborty, e Aberer [90], altera o conceito para uma perspectiva semi-intrusiva, onde a intenção é reduzir a necessidade monetária associada com custos em sensores fazendo a seleção ótima do número de sensores ao mesmo tempo que se maximiza a capacidade em desagregar a informação. Para isso, eles criaram uma regra heurística para encontrar o conjunto ótimo de equipamentos a serem monitorados por submedição. Os dados foram obtidos na frequência de amostragem de 1 Hz para o medidor central, e $\frac{1}{3}$ Hz para os medidores por equipamento. Essa informação foi reduzida para uma taxa de 1 amostra a cada 20 segundos através de um filtro de mediana. Também se utiliza o algoritmo de *Viterbi* para modelar as FSM, no caso utilizando Modelo Oculto de *Markiv* (HMM, em inglês). Conseguiu-se experimentalmente eficiências na ordem de 95% por equipamento em termos de F_{en}^i (4.10). Em um trabalho futuro, os autores também desejam utilizar outros sensores no

ambiente, como em [92].

4.2.5 Discussão

A primeira questão que chama atenção no levantamento bibliográfico é o quesito de escala. O NILM deve ser robusto para a aplicação independente da quantidade de equipamentos disponíveis na residência do consumidor. Naturalmente irá ocorrer a presença de equipamentos desconhecidos, bem como de C5 que, conforme o acúmulo de diversos equipamentos, irá tornar a rede um ambiente muito mais complexo para operação das técnicas. É justamente nesse ambiente que as abordagens devem ser analisadas, demonstrando como será o seu comportamento nas condições adversas presentes nas redes elétricas residenciais.

Geralmente os NILMs utilizam eventos de transitório para buscar por alterações no estado operativo dos equipamentos na rede — Etapa 2. Essa estratégia reduz a análise de informação pelo discriminador, uma vez que a busca por padrões só é realizada quando distúrbios são encontrados na rede. Apesar disso, as técnicas empregadas para encontrar os eventos devem ser capazes de separar distúrbios causados por ruídos daqueles causados pelos equipamentos, uma vez que a eficiência do NILM será diretamente afetada pela sua capacidade. Isso também evidencia que os discriminadores precisam ser robustos à presença de intrusos, tanto causado por ruídos ou por equipamentos desconhecidos. Poucos estudos trataram a capacidade do detector de transitório ou como um transitório falso é tratado pelo discriminador.

Outra capacidade desejada nos NILMs é a criação dos modelos de consumo dos equipamentos de maneira automática. Diversos autores fizeram mão de modificações do algoritmo de *Viterbi* [49, 50, 58–60, 79, 80, 82, 83, 90] operando em cima de técnicas de agrupamento de dados para a modelagem de C2s e C3s, no entanto, por modelar utilizando o comportamento mais provável, esta técnica não permite observar C2b e C4. É compreensível que tal tarefa não seja possível de ser realizada cegamente, justamente por não existir um padrão a ser identificado. Possivelmente esses equipamentos terão seus estados parcialmente modelados por C3s e/ou C2, sendo necessária intervenção humana para corrigir sua modelagem. As C4s representam um desafio ainda maior para os NILMs, podendo ser mais uma fonte de ruídos para aquelas configurações que realizam a Etapa 2, bem como um empecilho para a reconstrução de energia por não alterarem seu consumo de maneira discreta. Deste modo, até o momento as técnicas para criação de modelos trabalharam apenas com as características macroscópicas e estatística de uso. Normalmente, os autores utilizaram em conjunto com os modelos construídos cegamente técnicas de otimização [79] ou discriminadores estatísticos [80, 82, 83, 90], o que seria a continuação natural ao problema. Ainda assim, a extração dos modelos necessita de

uma grande quantidade de estatística — 5 dias à 2 semanas —, possibilitando a utilização dessa estatística obtida na construção dos modelos para alimentar diretamente o treinamento de técnicas supervisionadas.

Além do mais, há uma discussão em relação a como transformar a informação dos modelos construídos cegamente em equipamentos. Os casos observados [48, 79, 82] tendem a comparar as informações obtidas com um catálogo, mas como observado por [61], a melhor estratégia ao tema seria ter uma plataforma de comunicação com o consumidor, deixando a tarefa de identificação para o mesmo. O catálogo poderia ser utilizado como valor inicial ou sugerido para o equipamento, e se nenhum valor fosse encontrado no catálogo, um valor temporário como “desconhecido01” poderia ser utilizado até o usuário modificar seu valor. Com todas as facilidades oferecidas pelas ICTs, a comunicação com o usuário pode ser realizada por uma interface *web*, um aplicativo para *smart phones* ou alterar os mostradores domiciliares para permitir interação com o consumidor.

Nesse caso, também seria interessante a identificação de novos equipamentos conforme a utilização da rede. A identificação de equipamentos não mais em uso pode ser benéfica para técnicas que sofrem com C6, reduzindo a possibilidade de confusão se um equipamento não estiver mais sendo utilizado. Aproveitando a possibilidade de comunicação com o consumidor, o projeto pode permitir do mesmo alterar possíveis erros na modelagem dos equipamentos, unindo ou separando estados de FSM erroneamente identificados. Isso também pode ser expandido para o caso de erros de discriminação dos equipamentos. Já para o caso das abordagens onde o NILM não tem a aptidão para modelar automaticamente os equipamentos, fica evidente a necessidade de priorizar os equipamentos de maior uso-final para, assim, maximizar sua reconstrução de energia. Ainda assim, o sucesso da técnica depende de alta eficiência, onde a intervenção do consumidor não pode ser um ponto necessário no projeto. Em especial para o caso da aplicação da técnica como um programa de EE, onde a tarefa do NILM, além de desagregar energia, é informar o consumidor sobre como utilizar a energia de maneira sustentável através da redução do consumo e resposta em demanda.

Do ponto de vista técnico, o emprego das abordagens devem ser realizadas tendo em mente que a tarefa principal do NILM é a desagregação da informação de energia, independente de como a técnica empregada entenda o problema. Por exemplo, técnicas de separação cega de fontes irão ver o problema como um problema de desmistura de fontes, técnicas de aprendizado de máquina como a capacidade de entender as diferenças das características apresentadas e técnicas de otimização como obter a melhor qualidade/aptidão ou menor desvio com aquilo que é oferecido. As técnicas têm seus próprios parâmetros para informar o quão bem realizaram suas tarefas, no entanto, pela perspectiva do NILM e do consumidor, a eficiência deve

ser informada em termos da capacidade de reconstruir a energia por equipamento. Isso se torna ainda mais evidente quando o NILM opera em tempo real, de tal modo que sua eficiência deve analisar *temporalmente* o acerto de energia identificando erros devido à redundância ou subestimação do consumo por equipamento como foi tratado na Subseção 4.2.3.

Quanto à extração de características e técnicas empregadas, observam-se três vertentes. Buscando a impregnação do método, uma vertente limita-se ao emprego de técnicas em cima de características macroscópicas obtidas com baixa amostragem [48, 51–60, 79, 80, 82–85]. Essa vertente apresenta a configuração de menor embargo financeiro e pode aproveitar de uma infraestrutura que está sendo disponibilizada pela rede elétrica inteligente, os medidores inteligentes (no Brasil, sujeito a possíveis restrições descritas na Subseção 3.2.4). Os autores de [79] realizaram uma arquitetura distribuída para permitir desagregar informação de consumo sem qualquer alteração da estrutura fornecida nos EUA. [58–60] mostraram que é possível realizar essa tarefa em medidores eletromecânicos. No entanto, a presença apenas das informações macroscópicas irão deteriorar a eficiência do NILM, onde C6 irão afetar a capacidade de discriminação dessa vertente.

Por isso, as tendências nessa vertente são o emprego de estatística de uso tanto de maneira cega [80, 82, 83], como através de levantamentos de dados por institutos [84]. No último caso, a técnica irá se concentrar nos equipamentos de maior uso-final para maximizar a capacidade de reconstrução de energia total. Por outro lado, não se sabe, no momento, se é possível seguir esse caminho no Brasil uma vez que são necessários dados recentes e representativos das diversas regiões do país. Como foi articulado na Seção 2.3, a concepção inicial para o NILM no Brasil foi justamente o intuito de facilitar a coleta e precisão de dados, reduzindo o peso das PPHs. Nada obstante, o levantamento de estatística na própria residência, como mostrado por [82, 83], parece ser suficiente para atender os pontos necessários para a aplicação dessa tecnologia em um programa de EE. A única questão dessa abordagem seria como rotular os modelos encontrados com os nomes dos equipamentos, assunto debatido previamente nesta subseção.

Outra vertente, ao procurar por melhores capacidades de discriminação, seguiu o caminho de elevar a taxa de amostragem [29, 63–68, 70–77, 81, 87, 88] para tornar possível a avaliação de características microscópicas. A melhor capacidade discriminativa das características a serem exploradas nesse universo levaram aos autores geralmente a se contentarem em utilizar apenas uma técnica para a desagregação do consumo. Em alguns casos [67, 68], a resolução das características é tão fina que as próprias propriedades da rede interferem nos padrões dos equipamentos. Logo, os discriminadores teriam a inconveniência de ser alimentados com a informação para cada residência, bem como a utilização de um catálogo global seria impossível,

limitando a aplicação dessa abordagem se não houver uma maneira automática de realizar esse processo.

Os estudos realizados com características microscópicas geralmente utilizaram do conhecimento prévio dos equipamentos presentes na rede e analisaram a capacidade de desagregar a informação somente com esses equipamentos presentes, ignorando a necessidade de escalabilidade do projeto. Ainda, por analisar características com maior resolução, essas técnicas são possivelmente mais sensíveis a ruído, já que os mesmos distorcem com mais facilidade as características microscópicas tornando as mesmas um padrão não similar àquele procurado. Por isso, ainda é necessário descobrir se a maior capacidade de discriminação também se aplica para essas condições.

A presença de informações discriminativas no nível de ciclo da rede possibilitou abordagens onde a Etapa 2 não é realizada antes da Etapa 3. Uma das capacidades dos métodos que exploram o regime operativo permanentemente é a identificação de CIs. Porém, eles sofrem com a necessidade de que todos os padrões possíveis sejam conhecidos. No caso, só foram observadas três abordagens que operam desse modo [66, 74, 88]. As abordagens de [66, 88] trabalham com todos os possíveis estados operativos da rede (combinação de todos os estados operativos possíveis para os equipamentos), tornando essas abordagens frágeis para a aplicação quando a quantidade de estados na rede é grande, uma vez que é necessário treinar as técnicas supervisionadas (no caso foram aplicadas RNAs) com um número de padrões que cresce exponencialmente proporcional ao número total de estados dos equipamentos. Já a abordagem de [74], por trabalhar o problema através de otimização, conseguiu uma maior possibilidade prática para a aplicação de sua técnica, ainda estando sujeito à possíveis problemas de patamares de consumo parecidos que possam ser determinados por diversas combinações de equipamentos diferentes. Além disso, não se sabe como a presença de equipamentos desconhecidos afetarão essas abordagens. Ainda ficou em aberto como determinar o consumo dos equipamentos através da informação retornada pela etapa de discriminação, sugestões foram realizadas na pp. 74. Finalmente, os resultados mostram que para configurações onde há poucos equipamentos conhecidos essas técnicas solucionam o problema bem-sucedido.

Uma possibilidade — e *quixá* tendência — para essa vertente é a combinação de múltiplas características e técnicas para o reconhecimento de padrões [73, 77, 81]. A pluralidade de características independentes é benéfica para a capacidade de discriminação, sendo ainda um fato pouco explorado. Alguns autores chegaram a utilizar/considerar as medidas macroscópicas para auxiliar no processo discriminativo em conjunto com as medidas microscópicas. Por exemplo, [87] empregou uma etapa de pré-seleção para os equipamentos que mais se destacavam nas características macroscópicas facilitando no processo discriminativo. Nessa mesma abordagem, como se utilizou apenas da magnitude da corrente como característica

microscópica, observou-se que a adição da variação do FP era benéfico para a discriminação. Porém, geralmente as características microscópicas conseguem representar as características macroscópicas, sendo necessário buscar por outras características microscópicas representativas. Além disso, o emprego de múltiplas técnicas independentes em paralelo (em cima, ou não, das mesmas características) permite uma melhor discriminação, já que poderá acontecer complementação das lacunas individuais de cada técnica. Há uma variedade de possibilidades para a fusão dos resultados apresentados em paralelo pelos discriminadores, que além de RNAs, podem ser realizadas como proposto por [73] ou [43] (ver pp. 81). No entanto, é importante frisar que isso também trará uma maior carga computacional ao NILM, de forma que a quantidade de características e técnicas aplicadas será limitada pela capacidade de processamento disponível, particularmente quando operando em tempo real. Ressalta-se aqui que, além da discussão para aplicação de múltiplas características microscópicas e técnicas, o emprego de estatística do uso, característica até o momento *sui generis* na vertente explorando características macroscópicas, também pode ser realizada em quaisquer topologias do NILM.

A última vertente quebra a ideia de um medidor central para desagregar toda a informação de consumo, percebendo que a questão do projeto é reduzir o custo com sensores de forma a torná-lo viável em larga escala, o que não implica, necessariamente, na utilização de apenas um medidor. Para essa versão, o NILM deixa de ser não-intrusivo, para ter uma configuração *semi-intrusiva* [90–92]. O objetivo para essa configuração pode ser descrito como minimizar o custo com sensores sem deteriorar a capacidade de desagregação. Essas abordagens podem optar por submedidores mais próximos ao nível de consumo, reduzindo o acúmulo de operação de equipamentos que, intrinsecamente, torna a tarefa mais simples; utilizar outros sensores que deem sensibilidade para a discriminação, como correlacionar o nível sonoro, de iluminação e movimento com certos equipamentos; ou submedição diretamente nos equipamentos que as técnicas empregadas tiverem maior dificuldade. Porém, deve-se considerar a estética quando na utilização de submedidores, onde os consumidores tenderão a rejeitar configurações que reduzam a harmônia do ambiente em seus domicílios. Isso também é verdade para qualquer configuração que reduza a privacidade ou façam-nos ter a sensação de que isso possivelmente irá ocorrer. Uma outra percepção em vista das dificuldades das abordagens que operam ciclo a ciclo [66, 74, 88], uma maneira de superar a dificuldade do crescimento exponencial seria através da aplicação de submedidores atuando em cima de uma informação agregada até um número máximo de equipamentos.

Portanto, as técnicas devem ser robustas para aplicação em residências com a presença de equipamentos desconhecidos e para os momentos onde haverá maior operação, causando a operação simultânea de diversos equipamentos que, possi-

velmente, injetarão ruídos na rede através de sua dinâmica. Devido a presença de equipamentos desconhecidos, um aspecto desejável para o NILM é conseguir detectá-los de maneira automática e de algum modo possibilitar a correlação dos modelos identificados com seus nomes. Outra questão é como obter informação discriminante garantindo que o projeto seja viável. A utilização de amostragem baixa deteriora a capacidade de discriminação, por outro lado é possível que a estrutura para essas abordagens já esteja disponível. Essa abordagem parece ser a mais próxima de tornar-se realidade. Enquanto a elevação da amostragem aumenta a eficiência, dificulta a aplicação do projeto por necessitar de *hardware* específico. Além disso, as técnicas presentes atualmente para alta amostragem ainda precisam ser testadas quanto a sua robustez. Todavia, um outro caminho para obter maiores eficiências pode ser através da escolha ótima de submedidores que, ao mesmo tempo, maximizam a eficiência e minimizam o custo do projeto.

4.3 A Metodologia no CEPEL/Eletrabras

Esta seção se dedica à descrição do projeto no CEPEL/Eletrabras e as colaborações realizadas pela COPPE/UFRJ no mesmo. Os trabalhos de autores precedentes ao presente trabalho é realizado na Subseção 4.3.1. A contribuição realizada pelo grupo de engenheiros do CEPEL/Eletrabras posterior ao último trabalho em colaboração com a COPPE/UFRJ encontra-se na Subseção 4.3.2, sendo essa metodologia a base para o trabalho atual. Na Subseção 4.3.3 é realizada considerações quanto a escolha da FEX em relação a discriminação, levantamento deixado para contribuir na escolha das mesmas em trabalhos futuros. Ver-se-á nesta seção que há uma série de dificuldades a serem abordadas antes de retornar a questão de discriminação dos equipamentos.

4.3.1 Estudos Anteriores

Os estudos envolvendo a aplicação de técnicas não-invasivas para a desagregação do consumo no CEPEL/Eletrabras tiveram seu início com o estudo de Aguiar [100], em 1996. Esse estudo, bem como os seus sucessores: DO PRADO [101] (2006) e Cañón [102] (2009), utilizaram a informação da envoltória da onda de corrente no regime transitório para identificar o acionamento dos equipamentos domiciliares. Diferente dos trabalhos realizados pelo mesmo grupo de Hart para o setor comercial e indústria [51, 52]¹⁹ onde a extração da envoltória obtida através da decomposição harmônica é realizada em amostras espaçadas em 1 Hz, os trabalhos utilizaram a

¹⁹Esses são apenas alguns dos trabalhos realizados, a literatura disponível pelo grupo é muito mais extensa.

envoltória em 60 Hz. Essa diferença é justificável pela diferença dos dois ambientes, onde os transitórios de muitos equipamentos disponíveis para os trabalhos [51, 52] tem características lentas (ordem de segundos a poucos minutos), tanto que os autores tiveram que explorar o conceito de seções-v (ver pp. 62), enquanto os explorados no setor residencial apresentam toda a informação relevante de transitório em até 1 segundo ([100] observou para seus equipamentos analisados 417 ms). Como foi notado por diversos autores [48, 53, 54, 69], há um lento transitório em alguns equipamentos que tenuemente alteram seu consumo — por exemplo, devido ao aquecimento de motores —, mas aqui se refere essencialmente a informação discriminante do transitório.

Todos os trabalhos utilizaram RNA para discriminar a envoltória de corrente, entretanto algumas peculiaridades podem ser observadas para cada estudo. Devido a maior limitação, Aguiar [100] utilizou 16 pontos da envoltória (normalizados por 100 A) espaçados não-linearmente para conseguir uma melhor representação da informação do transitório — distribuiu-se os pontos para obter uma maior granularidade próxima ao transitório, que gradualmente vão se esparsando conforme eles ficam distantes do centro —, o valor de corrente em regime permanente (normalizados por 50 A) e ângulo de carga. Apesar de todas as limitações técnicas, o método conseguiu eficiência de classificação superiores à 90%, mas sofrendo deterioração para o grupo de eletrônicos.

Já DO PRADO [101], trabalhou em cima de um sistema de aquisição mais adequado, onde sua melhor configuração empregou 60 amostras da envoltória que são transformadas utilizando Análise de Componentes Discriminantes (PCD, em inglês), resultando em 98% de eficiência de classificação. Esse autor também realizou outros estudos, sendo eles: *a*) consideração entre a utilização ou não da informação de regime permanente para classificação, onde ela se mostrou desnecessária e inclusive obtendo resultados piores que a sua melhor configuração; *b*) análise sobre como a sobreposição de transitórios afetaria a classificação, mostrando que esses eventos são críticos para a capacidade discriminativa; e *c*) implementação de diversos cenários possíveis para o projeto, considerando inclusive a implementação de um dNILM.

Finalmente, o sistema de aquisição disponível para Cañón [102] possibilitou a análise de uma gama maior de equipamentos, sendo capaz observar equipamentos com consumo superior à 150 mA, enquanto Aguiar [100] operava somente para equipamentos consumindo mais de 1 A. Houve também uma melhoria no levantamento do banco de dados, havendo uma amostragem bastante superior aos trabalhos anteriores, com um total de 1324 acionamentos observados. O trabalho implementou dois detectores especialistas para a identificação dos momentos de acionamentos (a Etapa 2, porém apenas para acionamentos), e um baseado em correlação. Foram utilizadas 150 amostras da envoltória de corrente, obtendo taxa de detecção de 86 %.

Grande parte da inexatidão ocorreu devido aos erros no grupo de eletrônicos, obtendo taxas de classificação bastante inadequadas para esse grupo, na ordem de 30 %.

Apesar da evolução dos trabalhos, diversos pontos podem ser levantados. Primeiro, a principal característica utilizada pelos estudos só é discriminante para os acionamentos de eventos, não podendo utilizar a mesma na identificação de padrões de desacionamentos ou mudanças para estados de menor consumo, que na maioria dos casos são apenas uma queda abrupta no consumo. Além disso, não é possível inferir se essa informação também seria discriminante para mudanças de estado positivas, já que os estudos não trataram o caso de C2 — sendo uma outra dificuldade a ser tratada. Não bastasse, os estudos trataram apenas de casos de equipamentos sendo acionados isoladamente. Como foi notado na Subseção 4.2.5, a maior dificuldade para os NILMs ocorrem devido à operação simultânea de equipamentos na rede, em especial quando eles se constituem de C5. Os esforços para realizar a Etapa 2 envolveram apenas a detecção de acionamentos, bem como o trabalho de [102] não considera a questão de classificação do desacionamento. Indo adiante, a taxonomia levantada pelos autores realiza a discriminação em grupos bastante genéricos de equipamentos. Não se sabe até que ponto essa informação pode contribuir para a desagregação do consumo de energia residencial, em especial para a aplicação em programa de EE, onde os consumidores desejam saber o consumo por equipamento. Continuando, a técnica se mostrou ineficiente para um grupo de equipamentos cada vez mais presente nas residências, os eletrônicos, sem ser possível saber o porquê isso ocorre. Por fim, não se tratou o tema de equipamentos desconhecidos, sendo ainda necessário considerar como adaptar o método para identificar novos padrões recorrentes.

Em vista das dificuldades de tornar a informação da classificação em informação relevante para os NILMs, o grupo do CEPTEL/Eletrabras empenhou-se no sentido de melhorar a capacidade de detecção de eventos de transitório, como será visto a seguir. Quanto à questão do problema da envoltória como característica para discriminação, serão realizadas considerações mais adiante, na Subseção 4.3.3.

4.3.2 Metodologia de Partida do Trabalho

A equipe do CEPTEL/Eletrabras comprometeu-se desde o trabalho de Cañón para tornar o projeto mais próximo das necessidades do NILM. Reparou-se na necessidade de um método para a detecção de eventos de transitório que fosse capaz de identificar tanto mudanças para estados através do acréscimo de energia, quanto no decréscimo. Os detectores fornecidos por esse autor são especialistas para detecção de acionamentos, sendo dois deles determinados empiricamente e um outro utili-

zando correlação entre o sinal e formas de onda pré-determinadas. Exceto o último, que não teve bons resultados, eles não são aplicáveis para detecção de decréscimo de demanda causado pelos equipamentos, sendo necessário adaptar os detectores empíricos para os casos de desacionamento, ou criar um novo filtro.

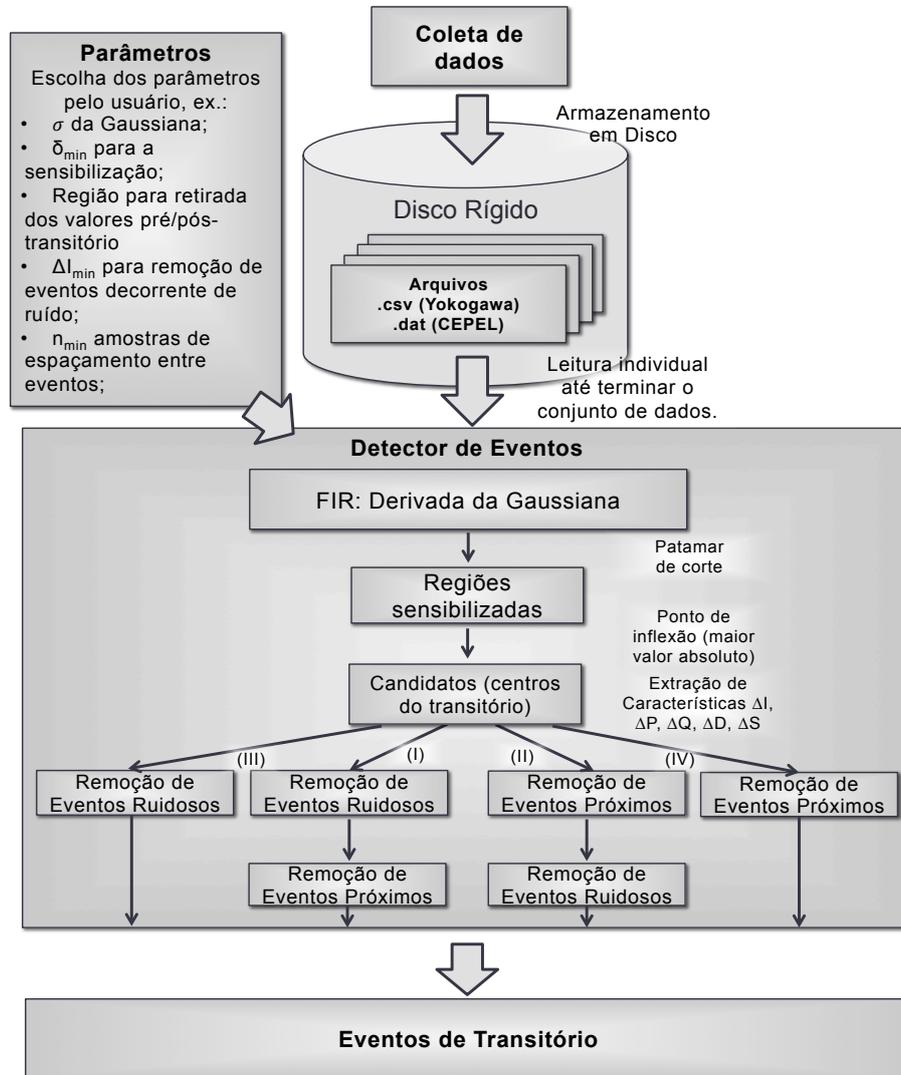
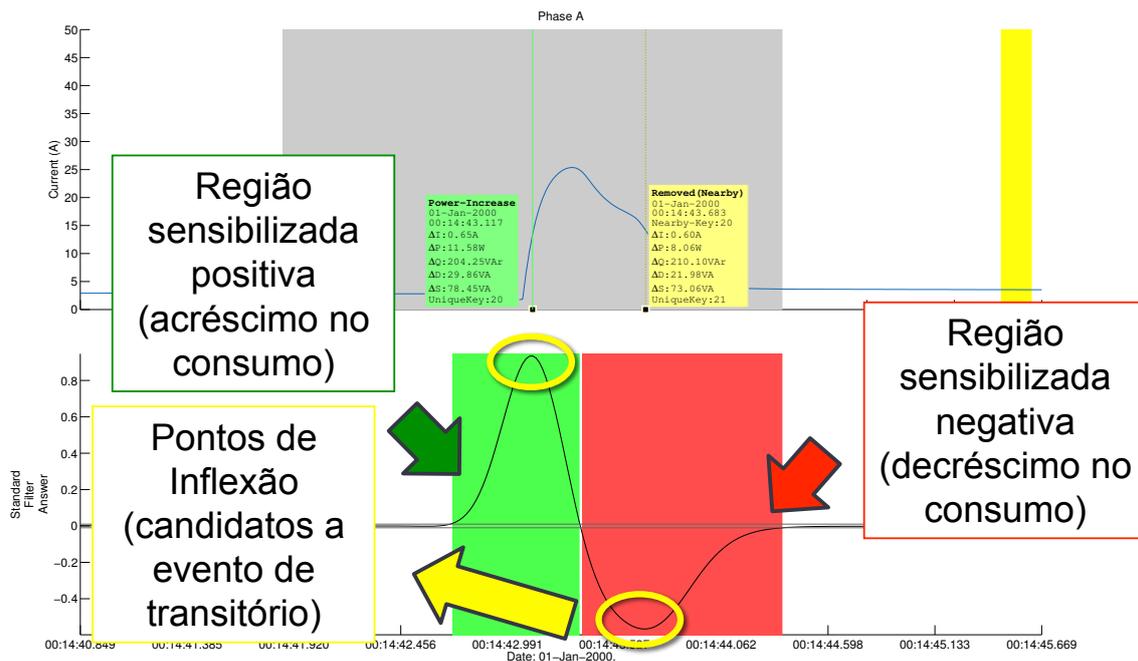


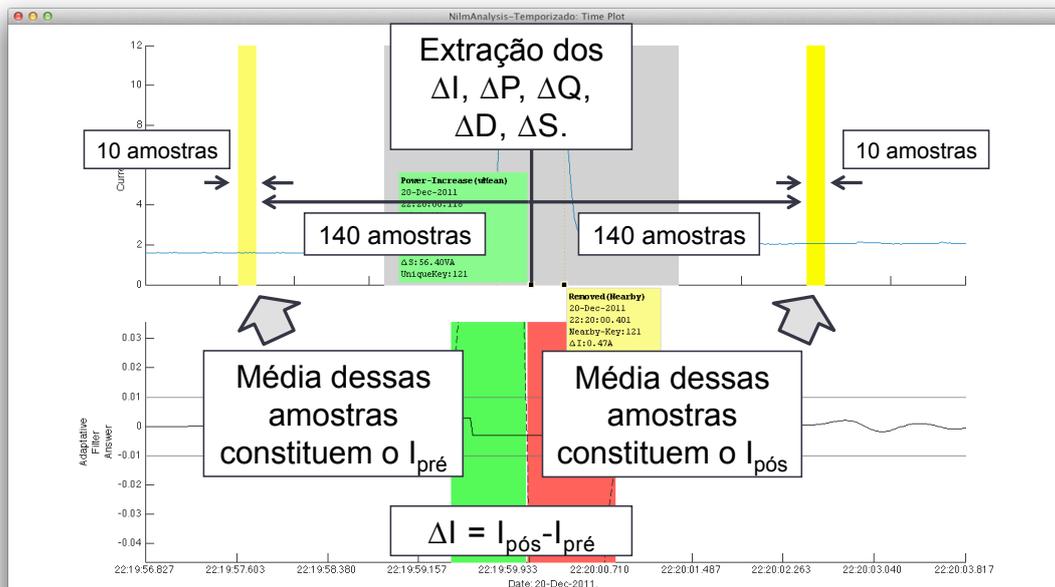
Figura 4.4: Esboço da metodologia empregada pelo CEPEL/Eletrabras, ponto de partida para este trabalho.

A opção tomada pelo grupo do CEPEL/Eletrabras foi a segunda, a equipe formada por Ribeiro, Navarro, da Silva, Borba, e Aguiar [103] decidiu utilizar a derivada de Gaussiana, baseando-se na literatura de processamento de imagens para a detecção de bordas. Um esboço da metodologia empregada pelo CEPEL/Eletrabras está disposto na Figura 4.4. No caso, a diferença na aplicação na derivada de Gaussiana é que, ao invés de operar em cima de *pixels*, os valores são amostras de corrente e, por isso, a mesma opera em apenas uma dimensão. Para realizar a convolução, aplicou-se um FIR com uma janela inicialmente com o mesmo número de amostras que havia sendo utilizado, de 150 amostras a uma taxa de 60 Hz, o que implica em

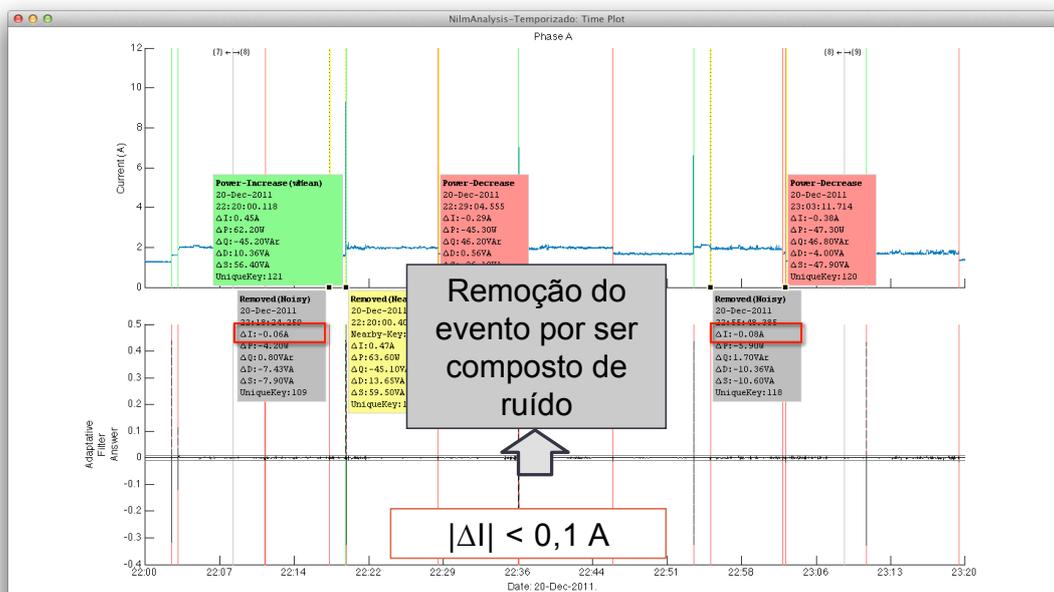
2,5 s. Vale ressaltar que na verdade é o desvio padrão da Gaussiana que determina o tamanho relevante da janela, já que ao se distanciar do centro os valores irão tender a zero — o mesmo sendo verdade para sua derivada —, sendo assim, o número de 150 amostras é apenas um corte para a quantidade máxima de informação contida no filtro e utilizadas para a convolução. O valor de resposta alimenta um corte linear no qual, se a resposta for superior em absoluto, irá gerar uma região sensibilizada para ser analisada até a resposta do filtro normalizar, ou continuar sensibilizando mas com sinal oposto, nesse caso já criando uma nova região sensibilizada. Para determinar o centro do transitório, simplesmente é realizado o cálculo do ponto de inflexão da resposta do filtro, e se houver mais de um, o ponto de inflexão com maior valor absoluto é utilizado. Esse processo está ilustrado na Figura 4.5a.



(a) Exemplo de resposta do FIR, de regiões sensibilizadas e dos candidatos resultantes proposta pelo CEPEL/Eletrbras. Na subfigura superior está a envoltória de corrente, que é utilizada como entrada do FIR e tem sua resposta representada na subfigura inferior. Observe que a resposta do FIR, ao ultrapassar o limiar de 0,003, gera uma região sensibilizada, representada em verde e vermelho, respectivamente para um candidato a acréscimo e decréscimo de consumo. Os candidatos são gerados no centro do transitório, o ponto de inflexão da resposta do filtro com maior valor absoluto, que estão circutados em amarelo na imagem.



(b) A realização da FEX. O cálculo é feito para todas as variáveis ΔI , ΔP , ΔQ , ΔD e ΔS .



(c) Exemplo de remoção de eventos ruidosos. A eliminação desses candidatos é importante para a redução do falso alarme. No caso, os candidatos possuem valores de ΔI menores ao limiar de 0,1, não sendo aceitos por esse motivo.

Figura 4.5: Gráficos descrevendo a operação da metodologia proposta pelo CEPEL/Eletrobras. Os gráficos foram gerados utilizando o ambiente de análise implementado por este trabalho descrito no Capítulo 5.

A fim de avaliar a eficácia, o relatório utilizou de novos dados (descritos na Seção 6.1), onde algumas configurações já apresentavam operação justaposta de equipamentos operando em diversos estados, no entanto sem ruído. A ideia aplicada foi de gradualmente aumentar a dificuldade a ser estudada e adaptar a abordagem até chegar em condições similares às aquelas encontradas nas residências brasileiras. A análise para as condições com pouca presença de ruído (os conjuntos de dados NI00***) revelou que a derivada de Gaussiana obteve melhor eficiência que os filtros apresentados no trabalho anterior, sendo a técnica escolhida desde então para a detecção de eventos de transitório. A versão do algoritmo que foi o ponto de partida deste trabalho ainda considerou mais alguns detalhes. Depois de ser determinado o centro, é calculado o pré e o pós-transitório para as variáveis macroscópicas disponíveis. No caso, elas são variação de corrente elétrica (ΔI), variação de potência ativa (ΔP), variação de potência reativa (ΔQ), variação de potência harmônica (ΔD) e variação de potência aparente (ΔS), obtidos através de 4.18.

Ao acompanhar a Figura 4.5b, dedicada para a ilustração desse processo, observa-se que os valores determinados inicialmente pelo CEPEL/Eletrabras para obter os valores de pré e pós transitório são a média de 10 amostras afastados de 140 pontos em relação ao centro do transitório. A variável ΔI é empregada para realizar um corte linear para um valor mínimo ($\delta_{min} = 0,003$) em que o candidato a evento — o centro selecionado da região sensibilizada pelo filtro — com o intuito de remover possíveis regiões sensibilizadas por ruído, Figura 4.5c ($\Delta I_{min} = 0,1$ A). Também se realiza um outro corte levando em conta a proximidade de candidatos a evento de transitório. Caso os mesmos estejam próximos entre si abaixo de um valor mínimo de amostras, os eventos temporalmente posteriores ao primeiro evento são rejeitados ($n_{min} = 150$ amostras). Um exemplo pode ser observado na Figura 4.5a, aonde o candidato proveniente da região sensibilizada negativa foi eliminado por estar próximo ao candidato da região positiva. Ainda foi considerado se o corte por ruído deveria ser realizado antes do corte por evento próximo, caminho (I) na Figura 4.4; vice-versa, caminho (II); ou apenas um dos cortes, caminhos (III) e (IV). O CEPEL/Eletrabras conseguiu excelentes resultados para a configuração sem ruído (os resultados estão disponíveis no capítulo de resultados na Tabela 7.1), entretanto, não havia uma maneira sistemática para a determinação os valores de corte, bem como visualizar o comportamento do algoritmo para entender suas nuances. Ao analisar a metodologia em novas configurações, o processo para análise sofria com a falta de infraestrutura.

$$\Delta I = I_{pós} - I_{pré} \quad (4.18a)$$

$$\Delta P = P_{pós} - P_{pré} \quad (4.18b)$$

$$\Delta Q = Q_{\text{pós}} - Q_{\text{pré}} \quad (4.18c)$$

$$\Delta D = D_{\text{pós}} - D_{\text{pré}} \quad (4.18d)$$

$$\Delta S = S_{\text{pós}} - S_{\text{pré}} \quad (4.18e)$$

Cabe definir o que são as características representadas em 4.18 com mais detalhes. As mesmas são bastante conhecidas e amplamente utilizadas em análise de sistemas elétricos operando em condições não-senoidais [104], sendo definidas para sistemas *monofásicos* no domínio da frequência por 4.19. Também se define uma grandeza já mencionada anteriormente no texto, o FP, através de 4.20. Uma observação importante quanto ao D, o mesmo é obtido exatamente como a equação 4.19f, ou seja, ele não é uma variável amostrada, e sim *estimada*. É importante notar que de acordo com 4.19f essa variável só pode assumir valores positivos, diferente de ΔQ que assume valores em ambas direções, sendo, na realidade, a única característica que apresenta essa propriedade. Obviamente, erros de medição irão ocorrer, de modo que o valor de D pode acabar sendo pertencente ao conjunto de números complexos. Para evitar isso, nesses casos o valor de D é determinado como zero.

Apesar de se ter conhecimento das deficiências para as variáveis em 4.19d–4.19f de representar grandezas físicas [104, cap. 2], não se levará o assunto em consideração pois:

1. mesmo que a representação física dessas variáveis não seja clara, não se está interessado na mesma, apenas na capacidade delas de servirem como característica para a detecção de eventos e discriminação, este trabalho focando na primeira parte;
2. essas variáveis são representações possíveis de serem obtidas com medidores de baixo custo.

$$v(t) = \sqrt{2}V_1\text{sen}(\omega_1 t + \phi_{1v}) + \sqrt{2}V_2\text{sen}(\omega_2 t + \phi_{2v}) + \sqrt{2}V_3\text{sen}(\omega_3 t + \phi_{3v}) + \dots \quad (4.19a)$$

$$i(t) = \sqrt{2}I_1\text{sen}(\omega_1 t + \phi_{1v} - \theta_1) + \sqrt{2}I_2\text{sen}(\omega_2 t + \phi_{2v} - \theta_2) + \sqrt{2}I_3\text{sen}(\omega_3 t + \phi_{3v} - \theta_3) + \dots \quad (4.19b)$$

$$P_n = V_n I_n \cos \theta_n \quad , \quad P = \sum_{n=1}^{\infty} P_n \quad (4.19c)$$

$$Q_n = V_n I_n \text{sen} \theta_n \quad , \quad Q = \sum_{n=1}^{\infty} Q_n \quad (4.19d)$$

$$S = VI \quad (4.19e)$$

$$D^2 = S^2 - P^2 - Q^2 \quad (4.19f)$$

onde:

V_n é a n-ésima componente da decomposição harmônica em valor eficaz de tensão;

I_n é a n-ésima componente da decomposição harmônica em valor eficaz de corrente;

ω_n é velocidade angular da n-ésima componente harmônica, podendo ser escrita como $2n\pi f$ (f frequência da fundamental);

ϕ_n é o ângulo de defasagem para a n-ésima componente de tensão;

θ_n é o ângulo de defasagem entre tensão e corrente para n-ésima componente da decomposição harmônica;

P_n é a potência ativa para o n-ésimo harmônico;

Q_n é a potência reativa para o n-ésimo harmônico;

$$FP = \frac{P}{S} \quad (4.20)$$

Em especial quanto ao segundo item, há um interesse por parte do CEPEL/Eletrabras de procurar uma solução menos complexa, mesmo que isso deteriore — até níveis aceitáveis — a capacidade do NILM, mas que resulte em uma resposta mais rápida em tempo de projeto. Por esse mesmo motivo, o CEPEL/Eletrabras considera a utilização de medidores presentes no mercado, dando preferência para os de baixo custo para facilitar o projeto, já que o desenvolvimento do *hardware* adiciona peso e reduz a velocidade do projeto. Por enquanto, dois medidores estão sendo utilizados, um proprietário do CEPEL/Eletrabras e um medidor da empresa *Yokogawa* [105], estando alguns dados disponíveis em uma configuração, e outros na outra.

As variáveis anteriormente citadas — adicionadas de suas envoltórias — são as que estão disponíveis atualmente para serem utilizadas (em amostragem de 60 Hz). Este trabalho foca na detecção de eventos, buscando nessa informação para identificá-los, e não entrando em detalhes sobre a classificação. Na subseção a seguir serão feitas algumas considerações com o intuito de auxiliar trabalhos futuros na escolha de características. Como é possível perceber, há uma série de dúvidas a serem tratadas antes de chegar no mérito da discriminação. Assim, focou-se em fornecer uma plataforma para dar maleabilidade ao projeto e laborar a generalização do detector de eventos para as condições de ruído que irão estar presentes nas redes elétricas domiciliares. Em especial, pretende-se que isso seja realizado de maneira automática,

encontrando uma boa configuração sem ser necessário realizar exaustivos testes em busca da mesma, independente da estrutura dos dados sendo testados. A plataforma consiste-se de um ambiente de análise *a posteriori*, estando descrito no Capítulo 5. O conjunto de dados disponibilizados pelo CEPEL/Eletrbras e o modo como o ambiente será aplicado para a detecção de eventos de transitório estão disponíveis no Capítulo 6. Os seus resultados podem ser encontrados no Capítulo 7. Porém, antes de chegar nessas questões, a seguir serão realizadas considerações quanto à escolha das características discriminantes, uma vez que as características anteriormente utilizadas pelos estudos em conjunto com o CEPEL/Eletrbras possuem informação discriminante apenas para mudanças de estados com acréscimo no consumo.

4.3.3 Quanto à Escolha das Características Discriminantes

Como foi observado anteriormente, há dúvidas quanto a utilização da envoltória de corrente como característica discriminante. A primeira consideração tratada leva em conta a taxonomia empregada para discriminação, sendo essa muito abrangente. Esta subseção começa tratando exatamente esse assunto.

O levantamento da taxonomia permite ter uma concepção do comportamento da distribuição dos dados nas características utilizadas. Elas oferecem um ponto de partida para evitar erros causados por C6, ou seja, é possível estimar quais equipamentos terão padrões parecidos utilizando tais características, começando assim, com uma configuração onde esses equipamentos representam apenas uma única informação a ser discriminada. Obviamente, a aglutinação de partículas não resolve o problema quanto à discriminação específica para os elementos aglutinados, mas possibilita obter uma informação com qualidade, ainda que mais genérica. Se essa generalização for um problema, será necessário procurar por características mais finas ou utilizar um discriminador especializado para realizar a segregação posteriormente. Em outros casos, a aglutinação pode ser realizada apenas para reduzir a dimensão do problema, procurando apenas encontrar assinaturas que permitam generalizar os equipamentos elétricos em grupos mais abrangentes que são suficientes para a abordagem em questão. No ponto de vista do NILM quando aplicado com o objetivo de explorar a EE e/ou melhorar a precisão das informações das PPHs, a generalização não é desejada já que é necessário a informação por equipamento.

A taxonomia encontrada pelos estudos do CEPEL/Eletrbras foi determinada empiricamente observando a forma das envoltórias e estão na Tabela 4.1. Apenas como referência para outros autores, foi adicionado também um outro levantamento empírico realizado por Aguiar [100], que se baseou nas considerações de [69], para o plano $I \times \theta$ (sendo θ o ângulo de fase). As principais diferenças nas taxonomias entre os estudos no CEPEL/Eletrbras foram devido às mudanças no sistema de aquisição

de dados²⁰. Percebe-se que, apesar de haver avanços na capacidade discriminante dos grupos, eles ainda permanecem bastante genéricos, em especial para a aplicação do NILM em um programa de EE. Avanços poderiam ser realizados ao empregar discriminadores especializados, no entanto, não é possível inferir qual seria a capacidade de discriminação para os diversos grupos, nem até que ponto a envoltória pode ser discriminante para disponibilizar informação desagregada de consumo.

Outro fato que aqui deve ser levado em conta é o aspecto levantado na Subseção 4.3.1, onde se percebeu a dificuldade da abordagem para discriminação do grupo de equipamentos eletrônicos. As considerações feitas por Cañón revelam que há dissimilaridade entre os padrões encontrados nos equipamentos. Infelizmente o autor não detalhou a causa para a dissimilaridade, não sendo possível saber se isso foi devido à aglutinação dos equipamentos eletrônicos ser indevida por eles terem padrões diferentes, e assim causando uma maior dificuldade ao discriminador para encontrar um padrão nesse grupo, ou se os acionamentos desses equipamentos ocorrem de maneira não bem-definida, por isso não sendo possível encontrar um padrão. O único fato revelado é que há dissimilaridade entre os acionamentos nesse grupo.

Independente da resposta para a questão anterior, as envoltórias podem ser consideradas como características macroscópicas quando levando em questão a *informação contida em um período de onda*, sendo um nível intermediário entre a informação de alteração no nível de consumo e aquela contida dentro dos ciclos de onda. Percebe-se que os equipamentos eletrônicos, por serem não-lineares, tem grande parte de sua informação discriminante dentro do ciclo de onda. Essas características — na literatura foram utilizados harmônicos obtidos da decomposição por FFT, níveis de detalhes das TWs, curvas I-V e, até mesmo, as ondas cruas — poderiam ser utilizadas para explorar a capacidade de discriminação dos equipamentos eletrônicos e outras cargas não-lineares. Diferente da limitação das envoltórias, onde a informação discriminante está disponível apenas para acionamentos (ver Figuras 6.19–6.23 para exemplos), a informação discriminante para essas características está disponível em todos os casos, aonde é necessário apenas extrair a informação da diferença causada pela mudança de estado dos equipamentos em pós e pré-transitório. Um exemplo seria a própria taxonomia levantada por LAM *et al.* [72] (para mais detalhes ver pp. 79), onde se percebe a distribuição dos equipamentos eletrônicos em um grupo bem definido (Grupo 7) e alguns outros equipamentos com grupos únicos para si. Vale ter em mente que os autores não estavam interessados em generalizar por equipamento, somente mostrar que é possível encontrar grupos de equipamentos com características elétricas que façam sentido fisicamente, e mesmo assim, alguns equipamentos tiveram assinaturas *sui generis*.

²⁰A aglutinação realizada por DO PRADO para equipamentos resistivos e de ventilação foi realizada com o intuito de comparar o seu método com outro utilizado pelo CEPEL/Eletronbras.

Grupo	Equipamento
<i>Envoltória de corrente por Aguiar [100]</i>	
1	Equipamentos de refrigeração em geral - refrigeradores, freezers e condicionadores de ar
2	Equipamentos para aquecimento (resistivos) - chuveiros, <i>boilers</i> , ferros de passar roupas, fornos em geral
3	Equipamentos de agitação/movimento/mistura - liquidificadores, batedeiras, máquinas de lavar, furadeiras
4	Equipamentos de ventilação - ventiladores e circuladores
5	Equipamentos eletrônicos - televisores, microcomputadores, equipamentos de som, videocassetes
6	Lâmpadas incandescentes
7	Lâmpadas fluorescentes
<i>Envoltória de corrente por DO PRADO [101]</i>	
1	Equipamentos de refrigeração em geral - refrigeradores, freezers e condicionadores de ar
2	Equipamentos resistivos (aquecimento) e ventilação
3	Equipamentos de agitação/movimento/mistura - liquidificadores, batedeiras, máquinas de lavar, furadeiras
4	Equipamentos eletrônicos - televisores, microcomputadores, equipamentos de som, videocassetes
5	Lâmpadas incandescentes
6	Lâmpadas fluorescentes
7	Microondas
<i>Envoltória de corrente por Cañón [102]</i>	
1	Equipamentos de refrigeração em geral - refrigeradores, freezers e condicionadores de ar
2	Equipamentos para aquecimento (resistivos) - chuveiros, <i>boilers</i> , ferros de passar roupas, fornos em geral
3	Equipamentos de agitação/movimento/mistura - liquidificadores, batedeiras, máquinas de lavar, furadeiras
4	Equipamentos de ventilação - ventiladores e circuladores
5	Equipamentos eletrônicos - televisores, microcomputadores, equipamentos de som, videocassetes
6	Lâmpadas incandescentes
7	Lâmpadas fluorescentes de reator eletrônico
8	Lâmpadas fluorescentes de reator eletromagnético
9	Máquinas de lavar
10	Monitores CRT
11	Monitores LCD
<i>Plano I-θ por Aguiar [100] baseado em [69]</i>	
1	Refrigeradores e <i>freezers</i>
2	Equipamentos resistivos (aquecimento)
3	Lâmpadas incandescentes
4	Motores universais (liquidificadores, batedeiras, aspiradores de pó etc.)
5	Ventiladores e circuladores
6	Equipamentos eletrônicos e lâmpadas fluorescentes
7	Ar condicionado

Tabela 4.1: Taxonomias encontradas em estudos anteriores no CEPEL/Eletronbras. Todas foram determinadas empiricamente.

Grupo	Equipamento
<i>Variáveis padrão</i>	
1	Aparelhos eletrônicos em modo <i>stand by</i> - computadores <i>desktop</i> , monitores <i>LCD</i>
2	<i>CD players</i> , geladeiras, carregadores de bateria, desumidificadores
3	Mistura de equipamentos resistivos, eletrônicos e com motores - jarros elétrico, lâmpadas fluorescentes de reator eletrônico, lâmpadas incandescentes, ventiladores, computadores portáteis, projetores, televisão de plasma, liquidificadores, exaustores
4	Lâmpadas fluorescentes de reator eletromecânicos, desumidificador
5	Cargas resistivas de potência média - secadores de cabelo (potência média), aquecedores, ar condicionado, fogão de indução (potência média)
6	Cargas resistivas de potência elevada e outros equipamentos ligeiramente não-lineares - ferro de passar, aquecedor de cabelo (potência alta), fogão de indução (potência alta)
7	Aspiradores de pó e forno microondas
8	Aspiradores de pó
9	Termo-ventilador (alta potência)
10	Computadores <i>desktop</i> e impressora a jato de tinta em <i>stand by</i>
11	Computadores <i>desktop</i> e projetores
12	Lâmpadas de economia de energia, computadores <i>desktop</i> , impressora de tinta a jato em <i>stand by</i> , televisores, <i>video cassette</i> , <i>scanner</i> , impressora a <i>laser</i> em <i>stand by</i> , <i>DVD player</i> ativo, televisores <i>stand by</i>
Outros	Secadores de cabelo operando na metade de consumo e televisores em <i>stand by</i>
<i>Curvas I-V</i>	
1	Lâmpadas incandescentes, ventiladores, aspiradores de pó, aquecedores, ar condicionado
1.1	Lâmpadas incandescentes, ar condicionado
1.2	Ventiladores
1.3	Aspiradores de pó
1.4	Televisores <i>LCD</i>
2	Lâmpadas fluorescentes com reator eletromagnético, <i>CD player</i> , carregadores de bateria, geladeiras, desumidificadores
2.1	Lâmpadas fluorescentes com reator eletromagnético, geladeiras, desumidificadores
2.2	<i>CD player</i> , carregadores de bateria
3	<i>CD player</i> e televisores <i>LCD</i> em <i>stand by</i>
4	Secadores de cabelo operando em baixa potência
5	Forno microondas
6	Lâmpadas fluorescentes com reator eletrônico, computadores portáteis, forno de indução
6.1	Lâmpadas fluorescentes com reator eletrônico
6.2	Computadores portáteis
6.3	Fornos de indução
7	Computadores <i>desktop</i> , televisores, <i>video cassette</i> , <i>scanner</i> , impressoras a <i>laser</i> , carregadores de bateria de telefones celulares
8	Aparelhos em modo <i>stand by</i> - computadores <i>desktop</i> , monitores <i>LCD</i> , forno de indução
Outros	projetores, máquinas de lavar
<i>SVD</i>	
1	Mistura de equipamentos resistivos, indutivos e outras ligeiramente não-lineares
1.1	Lâmpadas incandescente, aquecedores, ar condicionado
1.2	Fornos microondas
1.3	Secadores de cabelo (potência baixa)
1.4	Lâmpadas fluorescente com reator eletromagnético, geladeiras, desumidificadores
1.5	Ventiladores, exaustor
1.6	<i>CD player</i> em <i>stand by</i> , carregadores de bateria
1.7	<i>CD player</i>
2	Carregadores de bateria para telefones celulares
3	Computadores, monitores <i>LCD</i> , televisão, <i>scanner</i> , impressora de tinta a jato
4	Aparelhos eletrônicos em modo <i>stand by</i> - computadores, monitores <i>LCD</i> , impressora de tinta a jato

Tabela 4.2: Taxonomias determinadas através de dendrograma por LAM *et al.* [72] (ver pp. 79), tradução própria. O grupo *Outros* determina equipamentos com um grupo único para si. Em alguns casos, o autor realizou a separação em subgrupos, aqui indicados por uma barra vertical para facilitar a identificação visual dos mesmos.

Pode ser do interesse do CEPEL/Eletrbras alterar a amostragem dos medidores dependendo, por exemplo, se for possível utilizar os medidores inteligentes para garantir a impregnação do método, ou de ser necessário melhores eficácias, nesse caso, sendo necessário aumentar a frequência de amostragem. No primeiro caso, ter-se-á acesso apenas a variáveis macroscópicas, como ΔI , ΔP e ΔS , possivelmente outras variáveis podem ser empregadas, como ΔQ e ΔD , entretanto isso dependerá da capacidade do medidor utilizado. Já para alta amostragem, frequências maiores à 60 Hz permitem a obtenção da informação contida dentro do ciclo, possibilitando o acesso às informações descritas no parágrafo anterior e, assim, uma maior capacidade de discriminação. Nota-se que, como levantado na pp. 80, a escolha não precisa se limitar a uma única característica ou técnica, podendo abranger-se para um conjunto delas para obter melhor capacidade discriminativa. Nessa alternativa, pode-se também explorar a envoltória das ondas uma vez que elas mostram uma capacidade de, no mínimo, auxiliar na discriminação de cargas lineares para os eventos de transitório de equipamentos durante o acionamento. Essa capacidade pode estender-se para as cargas não-lineares, porém não será tratado neste trabalho se isso é verdadeiro, limitando-se aqui em apenas mostrar que a questão está aberta. Além disso, a informação de estatística de uso, como foi visto na discussão realizada para o levantamento bibliográfico (ver Subseção 4.2.5) é bastante discriminante e pode auxiliar tanto na vertente de baixas amostragens, onde seu uso chega a ser praticamente uma necessidade, ou quanto ao optar por uso de diversas características, que mesmo não sendo necessária, a sua utilização é fortemente indicada tomando como argumento os resultados mostrados em [80–84]. Finalmente, existe a possibilidade de utilizar submedidores, seja para reduzir a operação justaposta dos equipamentos a uma quantidade aceitável, ou para obter diretamente a informação desagregada de equipamentos complexos e importantes em termos de uso-final [90–92].

Capítulo 5

Ambiente de Análise

Inicialmente, no capítulo anterior, foi feito um levantamento dos aspectos envolvidos no projeto do NILM, levando em consideração diversos os pontos técnicos para a aplicação do NILM em um programa de EE. Em seguida, foi apresentado a evolução do projeto dessa tecnologia no CEPEL/Eletrbras. Observou-se a necessidade de detecção de eventos de transitório para a operação do NILM quando o mesmo explora a informação presente na alteração de um estado operativo de um equipamento. Essa etapa recebeu atenção do CEPEL/Eletrbras atualmente, onde se decidiu aplicar um filtro de derivada de Gaussiana, metodologia explicada em maiores detalhes na Seção 4.3.2. Aqui apenas será recapitulado que além da aplicação de um limiar na resposta do filtro, há a remoção de candidatos ruidosos e de candidatos próximos.

O trabalho atual firmou-se com o intuito de sistematizar o processo de ajuste de parâmetros da metodologia proposta pelo CEPEL/Eletrbras e estudar o comportamento da metodologia em uma nova base de dados que contém a presença de ruídos causados pela dinâmica de carga de equipamentos alguns aparelhos operando na rede — os equipamentos C5 (ver definições na Subseção 4.2.1). Além disso, ao observar nos capítulos anteriores a complexidade do projeto, percebeu-se a necessidade da implementação de uma infraestrutura que permitisse unificar o projeto e facilitar o seu desenvolvimento. Este capítulo se dedica à descrição dessa infraestrutura — o ambiente de análise. As necessidades para sua implementação serão debatidas mais profundamente na Seção 5.1. As seções seguintes (5.2–5.5) irão detalhar os módulos do ambiente implementado. Houve diversas alterações na versão original, trazendo benefícios operacionais e melhor capacidade de compreensão e interação com os dados e análises, especialmente no aspecto gráfico. Além disso, foram adicionados duas maneiras de remoção de eventos (Subseção 5.2.2) e a construção do gabarito (Seção 5.3) — estimativa da informação desagregada a ser obtida — que aqui são resumidos:

- Remoção de eventos próximos utilizando a média dos candidatos: os candidatos dentro de uma janela deslizante são substituídos pelo valor de sua média;
- Remoção de eventos inconsistentes: remoção do evento caso os sinais de entrada e da resposta do filtro e da variável ΔI sejam opostos;
- Construção do gabarito: a informação que se deseja obter é o consumo desagregado de energia por equipamento, porém conseguir essa informação para o ajuste das técnicas não é simples, sendo necessário realizar submedição em cada equipamento e anotar os instantes das mudanças de estados de operação de cada equipamento. Em alguns casos — como os dados coletados pelo CEPEL/Eletrobras (descritos no próximo capítulo) —, apenas a segunda informação está disponível, de forma que é necessário estimar o consumo desagregado dos equipamentos a ser obtido nos arquivos. Para melhorar a estimativa de consumo a ser obtida, contida no arquivo chamado de gabarito, possibilitou-se a interação gráfica para a construção dessa estimativa.

A Seção 5.5 apresenta a descrição da versão do algoritmo genético implementado, uma adaptação de um ES para possibilitar maior capacidade computacional às configurações que melhor se adequam ao problema sendo otimizado. Os detalhes da aplicação dessa versão adaptada do ES para o ajuste de parâmetros será debatido mais adiante, no Capítulo 6.

5.1 Da Necessidade

No capítulo anterior, foram observados diversos aspectos envolvidos para o desenvolvimento da tecnologia do NILM e a configuração do projeto no CEPEL/Eletrobras. Com base nisso, as seguintes dificuldades podem ser destacadas:

- Item 1 - era necessário melhorar a capacidade de interpretação dos dados e da análise;
- Item 2 - o CEPEL/Eletrobras pode julgar necessário alterar a sua abordagem conforme as necessidades do projeto, como as técnicas aplicadas, frequência de amostragem e o medidor;
- Item 3 - a larga gama de técnicas encontradas no levantamento bibliográfico, e em especial a chamada de atenção para o fato de sua utilização em paralelo ser benéfica para a capacidade de desagregação do NILM, mostra que o projeto deve ter aptidão de agregar em um único ambiente tudo aquilo que for desenvolvido, pois mesmo que uma técnica não seja ótima, sua operação em paralelo pode ser benéfica para o sistema de desagregação como um todo;

Item 4 - ainda que o item anterior não seja de interesse, é importante manter todas as abordagens já desenvolvidas em um único ambiente para garantir uma melhor evolução do projeto;

Item 5 - o levantamento bibliográfico mostrou que há uma dificuldade dos autores em obter dados onde a informação desagregada em energia também esteja disponível¹. Tem-se uma necessidade de tanto obter um meio para armazenar os momentos de transição dos estados operativos para treinar (se existentes) técnicas supervisionadas, quanto ter o consumo desagregado para avaliar a eficiência do NILM. No exterior disponibilizaram-se dois conjuntos de dados (ver Subseção 4.2.3) justamente com o intuito de possibilitar dados para autores e permiti-los compararem suas técnicas. Para obter a informação desagregada, pelo menos uma das informações seguintes deve estar disponível, no entanto, as duas se complementam, sendo desejável trabalhar com ambas: as marcas caracterizando os momentos de alteração de estados; e informação de consumo temporal dos equipamentos através de submedidores. Porém, nem sempre é possível obter as duas informações juntamente. Devido às dificuldades como o fato de não ser fácil monitorar os estados operativos de alguns equipamentos, por não ter acesso nem controle de seus ciclos de maneira trivial, bem como, nem sempre ser possível realizar a submedição de todos equipamentos, estando essa informação parcialmente ou até mesmo não disponível.

Viu-se a necessidade do desenvolvimento de um ambiente de análise que atendesse os seguintes pontos:

1. Versatilidade: em vista dos itens 2–4, o ambiente deve ser capaz de se adequar às mudanças no projeto conforme elas ocorram, sem que outras partes do ambiente sejam afetadas. Devido a isso, optou-se por uma implementação orientada a objeto, mas se limitando a escolha para uma linguagem de amplo conhecimento no campo da engenharia. Por isso, optou-se por desenvolver o ambiente no *Matlab*, que disponibiliza desde sua versão *R2008a* essa capacidade. Ainda que o *Matlab* e, em especial, sua linguagem orientada a objeto sofra em relação a sua performance, o ambiente de análise é *a posteriori* à coleta de dados, sendo aceitável essa desvantagem. O interesse no ambiente é de explorar a capacidade das técnicas, antes de implementá-las para operação em tempo real. O ambiente foi organizado procurando modularizar os componentes, para que, se fosse necessário o desenvolvimento ou adaptação de código, simplificasse o processo para que apenas o módulo em questão seja atacado;

¹Para fugir dessa dificuldade, [77] chegou a criar um simulador de Monte-Carlo (ver pp. 80).

2. Capacidade de Interpretação dos Dados e Resultados: o tempo investido neste ponto retorna em capacidade de interpretação, de forma que o projeto irá ter um melhor andamento. Um dos aspectos para atingir isso, é através de uma boa visualização [41], já que a mesma é um meio bastante efetivo para a comunicação da informação (aqui se referindo a informação presente nos dados, análise etc.). No caso, uma visualização dinâmica permite ainda melhor compreensão das nuances contidas na informação;
3. Otimização dos Parâmetros: também considerando o Item 1, seria interessante obter configurações ótimas de maneira automática, sem a necessidade do usuário ficar alterando parâmetros empiricamente até obter um valor considerado bom. Ademais, que o algoritmo seja capaz de realizar isso encontrando uma das melhores configurações possíveis para os parâmetros (não necessariamente a melhor);
4. Estimativa da Informação a ser Desagregada pelos Algoritmos (Construção dos Gabaritos): já quanto ao Item 5, os dados disponíveis no CEPEL/Eletrbras não foram amostrados com submedição, tendo apenas acesso às marcas de mudanças de estado operativo e, por isso, é necessário uma maneira para estimar a informação desagregada contida nos mesmos. Com esse intuito, aproveitou-se o ponto “Capacidade de Interpretação dos Dados e Resultados”, e adicionou-se essa capacidade na visualização dinâmica oferecida ao usuário. Essa estimativa da informação desagregada será referida neste trabalho como *gabarito*.

O resultado foi um ambiente de análise com ~ 29.000 linhas de código distribuídas em ~ 190 arquivos. Um esboço de sua arquitetura pode ser observado na Figura 5.1. As palavras em inglês representam as classes mais importantes dos módulos como referidas no ambiente. Observa-se que o Módulo de Leitura, Representação e Interação com os Dados (Seção 5.2) é a base do ambiente, sendo utilizado para análise e otimização. O Módulo de Análise dos Dados (Seção 5.4) é executado pelo Módulo de Otimização dos Parâmetros (Seção 5.5), que realiza diversas análises iterativamente em busca de uma configuração ótima para os dados alimentados.

É importante notar que apesar do esboço mostrar toda a cadeia para a obter os parâmetros otimizados, essa não é sua única operação. O usuário irá determinar sua operação, ex.: seja só para realizar uma análise, obter os resultados e explorar graficamente a sua resposta, construir um gabarito para um novo conjunto de dados, explorar os dados a procura de alguma informação etc.

Um detalhe, a implementação foi realizada em inglês por opção do autor do trabalho com o intuito de que o programa também seja compreensível no exterior.

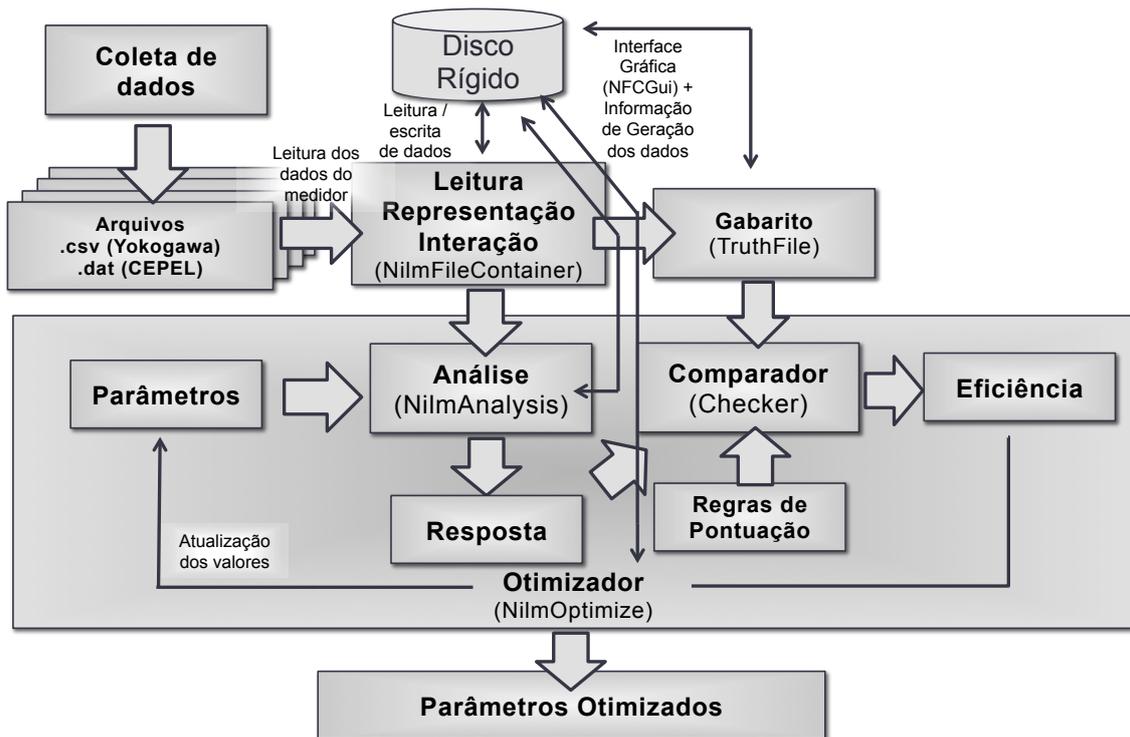


Figura 5.1: Esboço do ambiente de análise implementado.

Nem todas as informações puderam ser traduzidas para colocá-las no trabalho, nesses casos será realizado a tradução e referência aos elementos no texto do trabalho.

5.2 Leitura, Representação e Interação com os Dados

O Módulo de Leitura, Representação e Interação com os dados é o mais complexo em termos de estrutura no ambiente implementado. Ele conta com os seguintes segmentos:

- Dados do Medidor (Subseção 5.2.1): representação em memória transitória dos dados do medidor. Uma série de aspectos tiveram de ser tratados neste segmento;
- Evento de Transitório (Subseção 5.2.2): contém a informação de eventos de transitório. Essa representação pode ser criada tanto pelo usuário durante a criação de um gabarito, ou seja, informar um evento de transitório e suas propriedades a serem encontradas para avaliar a performance de análise ou otimizar os parâmetros baseando-se nessa informação, ou quanto pelo Módulo de Análise, que irá gerar essa informação através de sua metodologia;

- Equipamentos (Subseção 5.2.3): contém a informação do estado de consumo dos equipamentos, seu consumo temporal estimado bem como a estimativa de seu consumo total para o conjunto de dados. Apesar deste trabalho ainda não ter tratado do problema de geração da informação dos equipamentos (e por isso essa informação só ser gerada pelo usuário para o gabarito), é interessante em termos de continuidade do projeto que essa informação já fosse gerada nos gabaritos, para que eles não precisem ser revisados no futuro, contando com toda informação necessária para a otimização e avaliação de performance. Como foi visto na Subseção 4.2.3 e frisado diversas vezes em tal capítulo, é necessário dar a eficiência do NILM em termos de energia. Outro aspecto importante para a motivação da criação dessa informação foi da capacidade de compreensão dos dados, a informação por equipamento é muito mais intuitiva que os eventos de transitório, constituindo em um nível mais alto informativo para a compreensão dos dados, bem como facilitando a geração do gabarito;

A seguir, entrar-se-á em mais detalhes para cada um deles.

5.2.1 Dados do Medidor

Para atender as necessidades do projeto, a implementação da interface para leitura e representação dos dados do medidor abordou os seguintes tópicos:

- Transformação dos dados em um formato único: atualmente o NILM utiliza dados de dois medidores diferentes, sendo necessário representar essa informação de uma única maneira para atender a questão de Versatilidade. Um efeito colateral decorrente da transformação foi a compressão dos dados, que estavam em formatos de texto e ao serem armazenados em formato binário sofreram compressões de $30\times$ a $40\times$ dependendo do número de fases;
- Robustez: a leitura e transformação dos dados para o formato único deve ser robusta a possíveis erros durante a aquisição de dados, sendo eles: descontinuidade da amostragem, seja por intervenção humana ou algum problema no medidor; ou sobrecarga devido ao consumo excessivo na rede, geralmente causado pelo acionamento de um equipamento a motor de maior consumo, como o ar condicionado. Para o primeiro caso, implementou-se um algoritmo capaz de identificar esses momentos, e no caso da descontinuidade ser pequena (ex. 10 s, determinado pelo usuário), a informação entre as bordas dos arquivos é completada com amostras geradas através de um ajuste linear. Essas amostras são marcadas para identificar que foram criadas e não medidas. Enquanto para a sobrecarga, é grampeado o valor de consumo máximo para as variáveis em que

isso ocorre, bem como as amostras são marcas para identificar os momentos em que isso ocorre;

- Segmentação da memória persistente: alguns dados contém dias de amostragens, sendo impossível analisar toda essa informação de uma vez só em memória transitória. Por isso, segmentou-se os dados em diversos arquivos com um tamanho pré-definido (ex. 1 hora). Anteriormente era necessário segmentar a informação manualmente. A segmentação automática foi realizada de maneira transparente durante a leitura da base de dados, ou seja, a leitura ocorre sem que o usuário precise se preocupar em como o conjunto de dados está representado e compreendido no ambiente;
- Redução da necessidade de leitura de disco: devido a segmentação em memória, era necessário garantir que informações nas bordas dos arquivos estivessem disponíveis para os algoritmos de análise sem que eles tivessem de requisitar a troca da informação mantida em memória transitória. Para isso, uma quantidade de amostras nas bordas dos arquivos segmentados é mantida em memória transitória sempre disponível, evitando que seja necessário uma navegação excessiva entre a informação segmentada, reduzindo drasticamente a velocidade dos algoritmos já que a leitura em disco é lenta;
- Leitura de redes elétricas com até três fases: era necessário compatibilidade de leitura de dados de redes monofásicas, bifásicas e trifásicas, representando essa informação de uma maneira universal. Um dos fatores que influenciou também na compressão dos dados foi armazenar para as fases com pouca atividade somente os momentos em que havia consumo;
- Informação gráfica: representar graficamente a informação contida nos dados. Essa funcionalidade é utilizada como base pela interface gráfica para realizar a interação com os dados.

Um exemplo de dados trifásicos em uma residência *real*² está disponível na Figura 5.2. As três primeiras subfiguras são a injeção de corrente (medido em valor eficaz) no sistema, enquanto a última figura é o fluxo em potência trifásico para as variáveis descritas na pp. 98. O fluxo de potência é informado para o consumo trifásico porque o medidor *Yokogawa* utilizado nessa coleta de dados só permite o acesso a essa informação, sendo necessário operar com um nível mais agregado de consumo que a corrente. As barras verticais cinzas indicam como está realizada a segmentação dos dados em disco. Apesar de não se utilizar esse conjunto de dados

²A palavra *real* é empregada para identificar amostragens não geradas em condições de laboratório.

para análise — nesses dados não há como construir o gabarito, não sendo possível a otimização dos valores, nem calcular sua eficiência —, ele revela uma série de aspectos importantes para a compreensão do problema envolvido na desagregação do NILM, bem como algumas necessidades durante a implementação da parte de representação dos dados no ambiente de análise.

Quanto a questão da descontinuidade, há uma falha na medição próximo às 07:15 h do dia 31, mostrando que o algoritmo foi capaz de perceber essa falha e montar a descontinuidade. Já próximo às 19:20 h, ocorreu uma outra descontinuidade menor, de 30 s, onde o algoritmo identificou e uniu as bordas através de um ajuste linear para simular a continuidade e recuperar a informação perdida. Enquanto isso, o ajuste para a descontinuidade às 07:15 h não pode ser feito porque houve alterações de estados operativos dos equipamentos³. Por isso, esse conjunto de dados seria analisado em duas partes, uma partindo do início até às 07:15 h, e outra do fim da descontinuidade até o fim da medição.

Também há nesse período a ocorrência de sobrecarga do medidor próximo às 21:50 do dia 30 (provavelmente causada por um ar condicionado, ver Figura 5.3a), onde o algoritmo foi capaz de identificá-la e alterar os valores dessa amostras para a capacidade máxima de medição.

Outras condições que se referiu no Capítulo 4 também podem ser observadas e melhores compreendidas nesse conjunto de dados. Por exemplo, é possível observar na fase A Figura 5.3c⁴ há a ocorrência de uma C5 a partir das 20:00 h injetando ruído na rede elétrica devido a sua dinâmica (provavelmente esse equipamento é uma televisão). Já um exemplo típico de dinâmica causada pela máquina de lavar roupa se encontra na Figura 5.3b. Apenas como curiosidade, também é possível observar no período de menor atividade da rede — 08:00 h–18:00 h do dia 31 — nitidamente a operação da geladeira (ou outro equipamento similar) na fase B.

5.2.2 Evento de Transitório

Eventos são gerados tanto pelo usuário quando criando o gabarito ou quanto pelo algoritmo de análise. A informação nos eventos de transitório são de mais baixo nível àquelas contidas no objeto que representa o equipamento. Suas capacidades são:

- FEX para classificação: Realiza o cálculo das variáveis ΔI , ΔP , ΔQ , ΔD e ΔS e extrai uma janela da envoltória dessas variáveis durante o transitório;
- Remoção de eventos ruidosos: se ΔI , ΔP ou ΔS forem menores a um limiar,

³Estimar essas alterações sem nenhuma informação é mais complexo do que a tarefa do próprio NILM, que realiza isso tendo a informação agregada de consumo.

⁴Foi necessário reduzir a qualidade da Figura 5.3c para permitir a navegação na versão digital deste trabalho, a figura vetorizada exigia grande capacidade computacional.

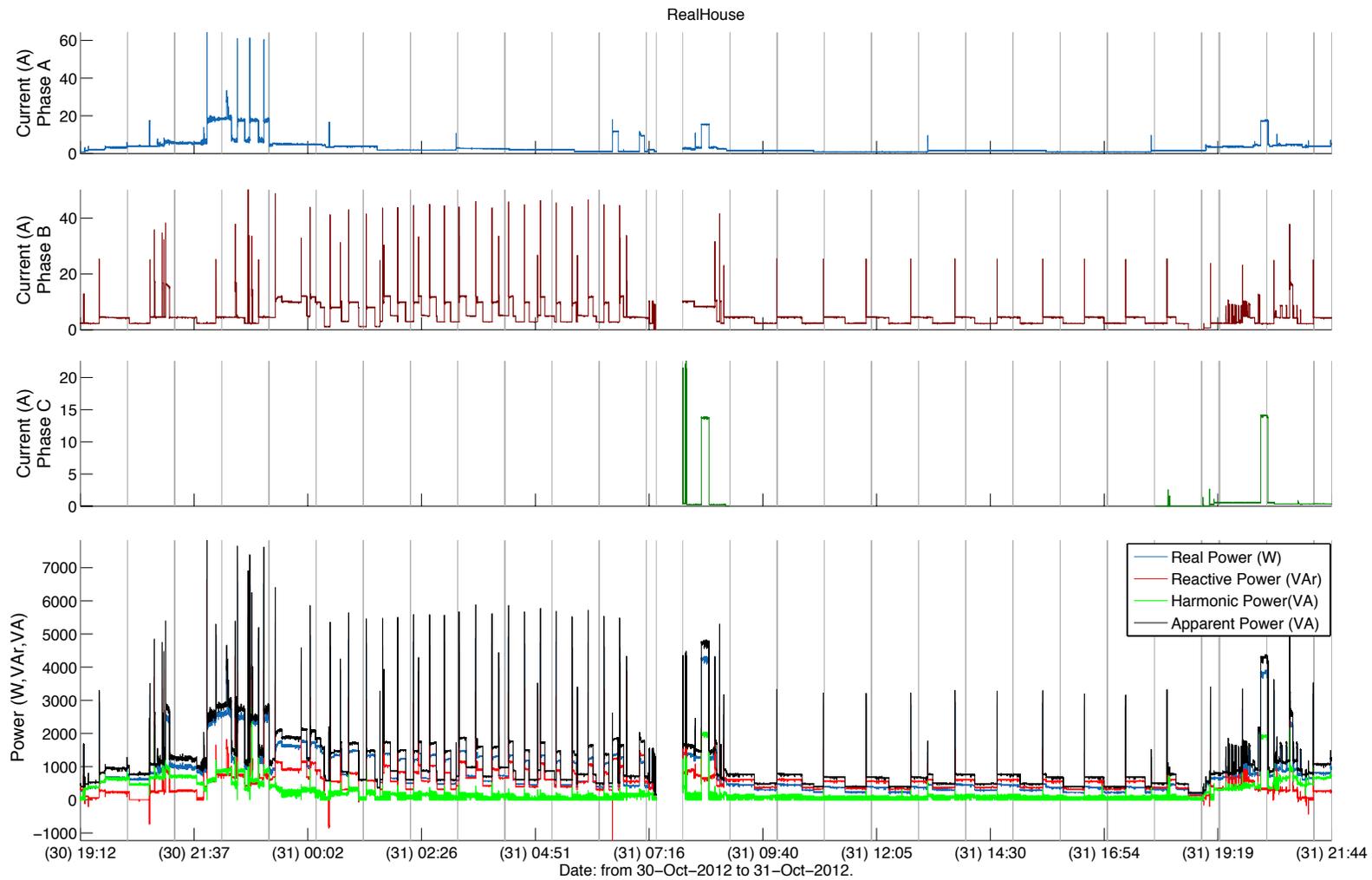
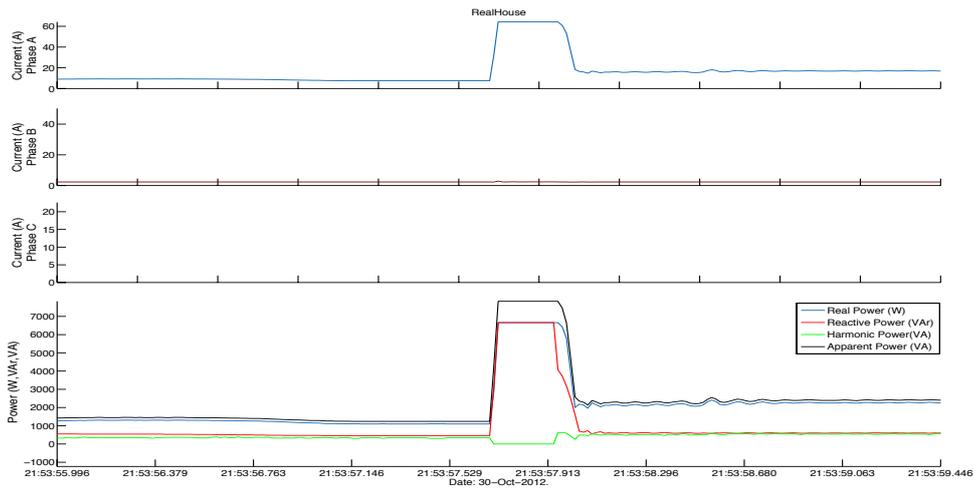
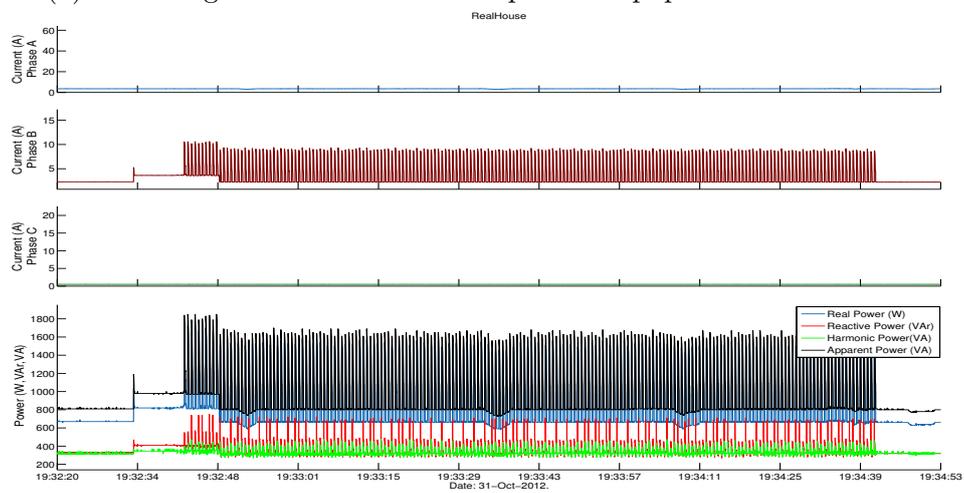


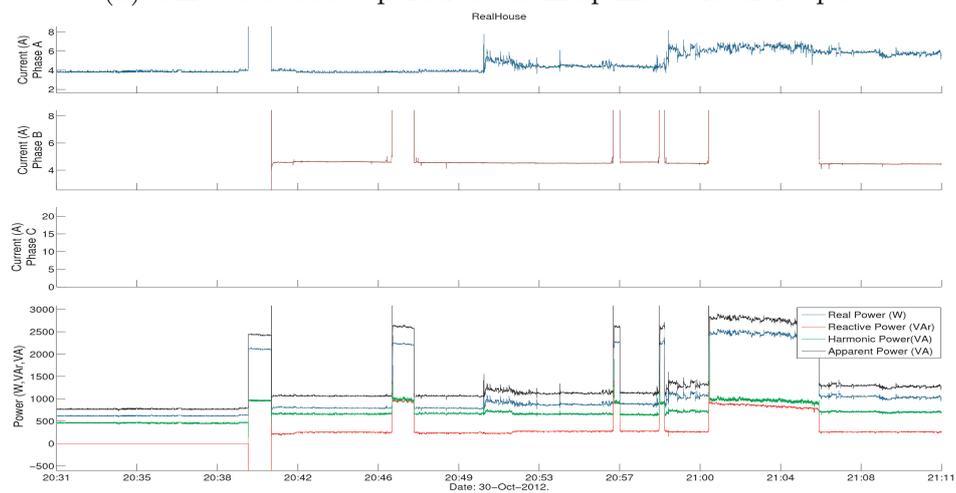
Figura 5.2: Informação gráfica para a interação com os dados do medidor. Gráfico gerado através do ambiente de análise para um conjunto de dados com amostragem em 60 Hz de uma rede trifásica em uma casa *real* durante aproximadamente um dia de coleta. A injeção de corrente para cada uma das três fases encontra-se nas subfiguras superiores, enquanto o fluxo trifásico de potência entrando na rede elétrica é representado na subfigura inferior. São utilizados as cores azul, vermelho, verde e preto para as potências ativa, reativa, harmônica e aparente, respectivamente.



(a) Sobrecarga do medidor causado por um equipamento na fase A.



(b) Um dos estados operativos da máquina de lavar roupa.



(c) Dinâmica de uma carga C5 na fase A. Observe a diferença entre a relação de sinal ruído dessa fase em relação com a fase B.

Figura 5.3: Alguns exemplos de dificuldades encontrados nos dados reais da Figura 5.2.

o evento é considerado como ruidoso. O corte pode ser realizado em apenas uma dessas variáveis, no momento o corte é apenas em ΔI .

- Remoção de eventos próximos pelo valor da média dentro de uma janela deslizante: Para a remoção de eventos próximos, adicionou-se uma outra maneira de realizar a mesma. Ao invés de simplesmente ignorar a informação de outros candidatos dentro de uma janela após um número determinado de amostras, como realizado na metodologia proposta pelo CEPTEL/Eletrobras, decidiu-se avaliar uma nova maneira de realizar essa tarefa, empregando a média dos valores das amostras desses candidatos. A inspiração para isso se deu pelo fato da maioria dos eventos próximos removidos serem causados por um acionamento de equipamento com um pico de consumo, assim, ao substituir o evento causada pelo acréscimo de consumo e seu decréscimo logo em seguida por sua média, obtém-se um evento mais próximo ao centro do distúrbio causado pelo equipamento na rede. Um exemplo pode ser observado na Figura 5.7;
- Remoção de eventos inconsistentes: ao observar que grande parte dos eventos que eram removidos por serem eventos próximos na verdade eram causados por eventos criados após um pico de consumo devido ao acionamento de um equipamento, que gerava um evento na decréscimo de consumo após o pico, decidiu-se adicionar um novo tipo de corte. Esse corte remove os candidatos que tiverem o sinal da resposta do filtro de derivada de Gaussiana invertido em relação ao degrau de consumo causado pelo evento. Para esses eventos citados, a resposta do filtro é negativa, enquanto o degrau é positivo, havendo assim inconsistência entre os dois;
- Estado dos eventos: para melhorar a capacidade de análise, eventos removidos não são excluídos, tendo apenas seus estados alterados. Os possíveis estados dos eventos são:
 - Em bom estado;
 - Removido devido a evento próximo, nesse caso indicando qual evento que causou sua remoção;
 - Evento ruidoso;
 - Inconsistente (ver item anterior);
 - Quantidade de amostras insuficientes, se não houver amostras suficientes para construir o evento;
 - Ainda não preenchido, quando o evento é criado mas ainda é necessário preencher a FEX e realizar os cortes para determinar o seu estado.

- Mudança de estado e equipamento: os eventos também armazenam a informação de qual equipamento pertencem e qual foi a mudança de estado por eles representadas. Por enquanto essa informação só é preenchida pelo usuário durante a criação do gabarito;
- Informação gráfica: os eventos são representados conforme o seu estado (observar exemplos na Figura 5.7). São utilizadas linhas verdes verticais para indicar eventos de transitório onde houve acréscimo no consumo, e linhas vermelhas para decréscimo. Eventos removidos possuem linhas cinza tracejadas. Dependendo de como a geração gráfica é realizada, será criado uma área cinza para a região onde será realizada a extração da envoltória e duas faixas amarelas indicando as amostras para as quais será calculado o valor pré/pós-transitório utilizados para calcular o degrau, como indicado nas equações 4.18.

5.2.3 Equipamentos

A informação contida no equipamento une toda aquela contida nos eventos. Nela, reconstrói-se todos os estados operativos do equipamento temporalmente e seus consumos. No momento, essa informação só é gerada pelo usuário quando preenchendo a informação do gabarito. Para ter uma ideia de como o processo é realizado observe a Figura 5.6. As capacidades desse elemento são:

- Detecção automática de estados: quando gerando o gabarito, faz-se uma análise de dendrograma para agregar os eventos de transitório com informações de ΔI , ΔP , ΔQ , ΔS próximas. O usuário só necessita alterar o nome dos estados pré/pós transitório dos eventos agrupados, simplificando o processo de geração do gabarito;
- Informação gráfica: a capacidade de geração de informação dos equipamentos é bem mais extensa que a dos eventos. Para evitar redundância de informação, refere-se diretamente as figuras onde são mostradas as características dos dados que foram gerados através do método de informação gráfica de cada equipamento. Estes são os possíveis gráficos de serem gerados são:
 - gráfico do consumo temporal por equipamento para os dados, Figura 6.17. Esse gráfico representa a energia desagregada estimada a ser encontrada;
 - gráfico circular de consumo dos equipamentos, Figura 6.18;
 - gráfico das envoltórias para todos os eventos de transitório, Figura 6.19. Essa informação auxilia a encontrar eventos anômalos, como o evento destacado na Figura 6.20.

5.3 Interação Gráfica com o Usuário

A interação gráfica com o usuário oferece uma melhor compreensão dos dados. Além disso, este módulo também permite a capacidade de geração do gabarito onde está toda informação considerada como alvo para o NILM, desde os momentos onde ocorreu os transitórios, até a estimativa de informação de energia. Por enquanto o módulo só opera com a informação dos dados do medidor, contudo sua expansão para realizar a interação com a informação de análise não é complexa e se pretende realizar sua implementação no futuro. A seguir estão suas capacidades:

- Informação da amostra próxima ao ponteiro: funciona de maneira muito similar à um medidor, mostrando os valores das amostras I (essa para cada fase), P, Q, D, S para a amostra mais próxima ao ponteiro (ver Figura 5.4, na região superior à esquerda denominada de *Cursor Info*, ou informação do ponteiro em português). Assim, o usuário pode comparar os valores em cada amostra com facilidade;
- Geração do gabarito: a dinâmica para a geração do gabarito ocorre da seguinte maneira: o usuário seleciona a opção “Adicionar Evento” (na figura estando em inglês: *Add Event*) e então seleciona a amostra que deseja ser o centro do transitório. Nessa amostra é criado um evento, onde são calculados os ΔI , ΔP , ΔQ , ΔD , ΔS . Essa informação é disponibilizada já calculada para o usuário no canto inferior da esquerda, contendo tanto o valor do patamar operativo do pré/pós-transitório dessas variáveis. O algoritmo detecta e agrupa os possíveis estados automaticamente por uma análise em dendrograma. Conforme o usuário vai preenchendo a informação do gabarito, a interface gráfica mostra-lhe o consumo desagregado por equipamento e mantém-na atualizada para cada alteração realizada, possibilitando o usuário saber se o gabarito está sendo preenchido de maneira correta ou não. Cabe ao usuário determinar o nome do equipamento e o nome de seus estados. Para equipamentos C3 a interface gráfica já os cria automaticamente com o estados *on* (ligado) e *off* (desligado). É possível selecionar qualquer evento e alterar suas características, bem como arrastar as regiões aonde são calculadas as variáveis (ver Figura 5.5) para corrigir possíveis sobreposições de eventos ou regiões inicialmente mal selecionadas.
- Escolha da informação disponível: há a opção de escolher a informação disponível na tela (observar figuras 5.6a e 5.6b), podendo controlar a disponibilidade gráfica das seguintes informações: informação de consumo temporal dos equipamentos; centros dos eventos de transitório (no caso, dos eventos

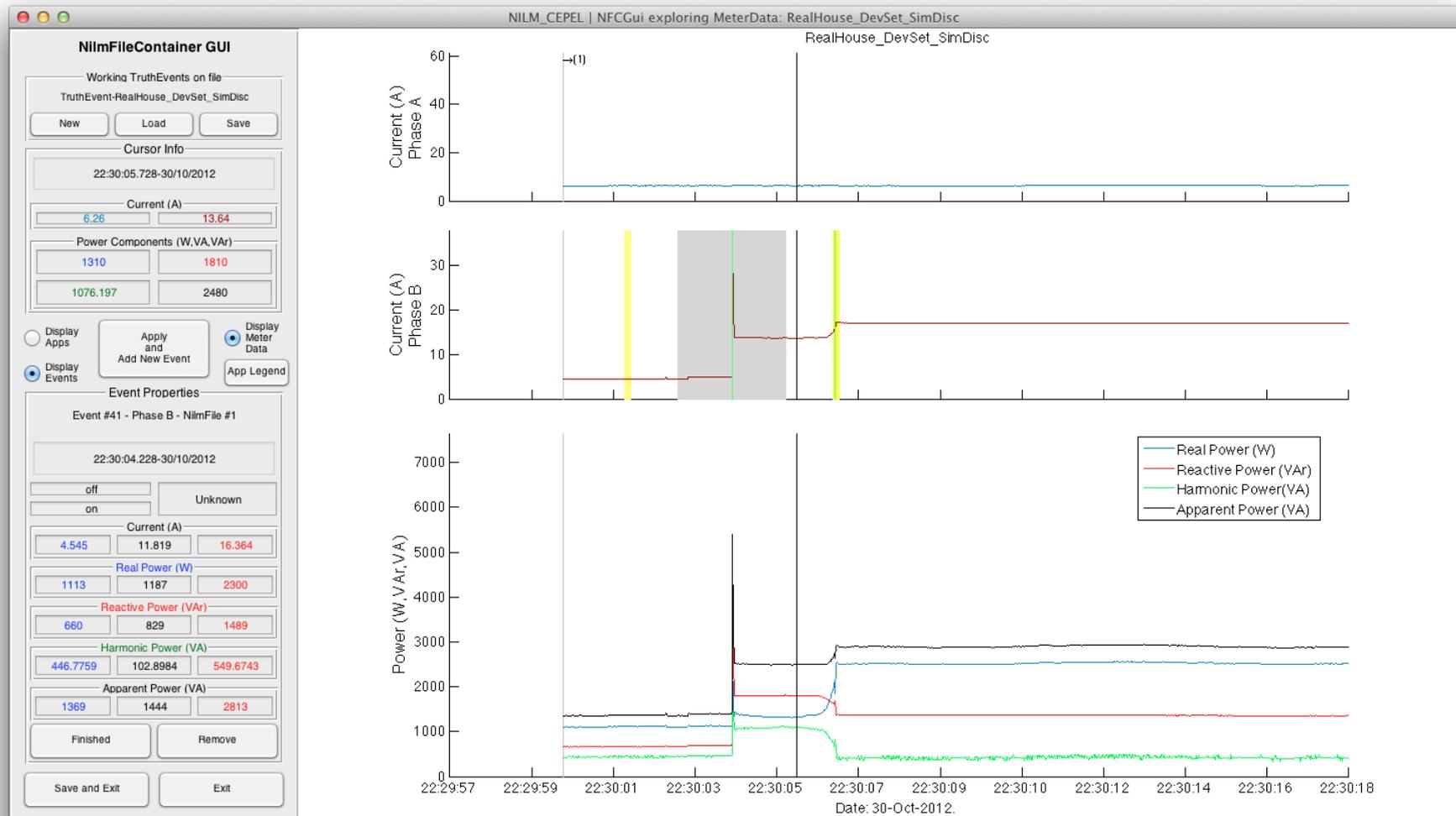


Figura 5.4: Informação gráfica para o Módulo de Interação Gráfica com os Dados: Evento de Transitório com Sobreposição. A região com as amostras para o cálculo da média de pós transitório está sobrepondo com outro evento.

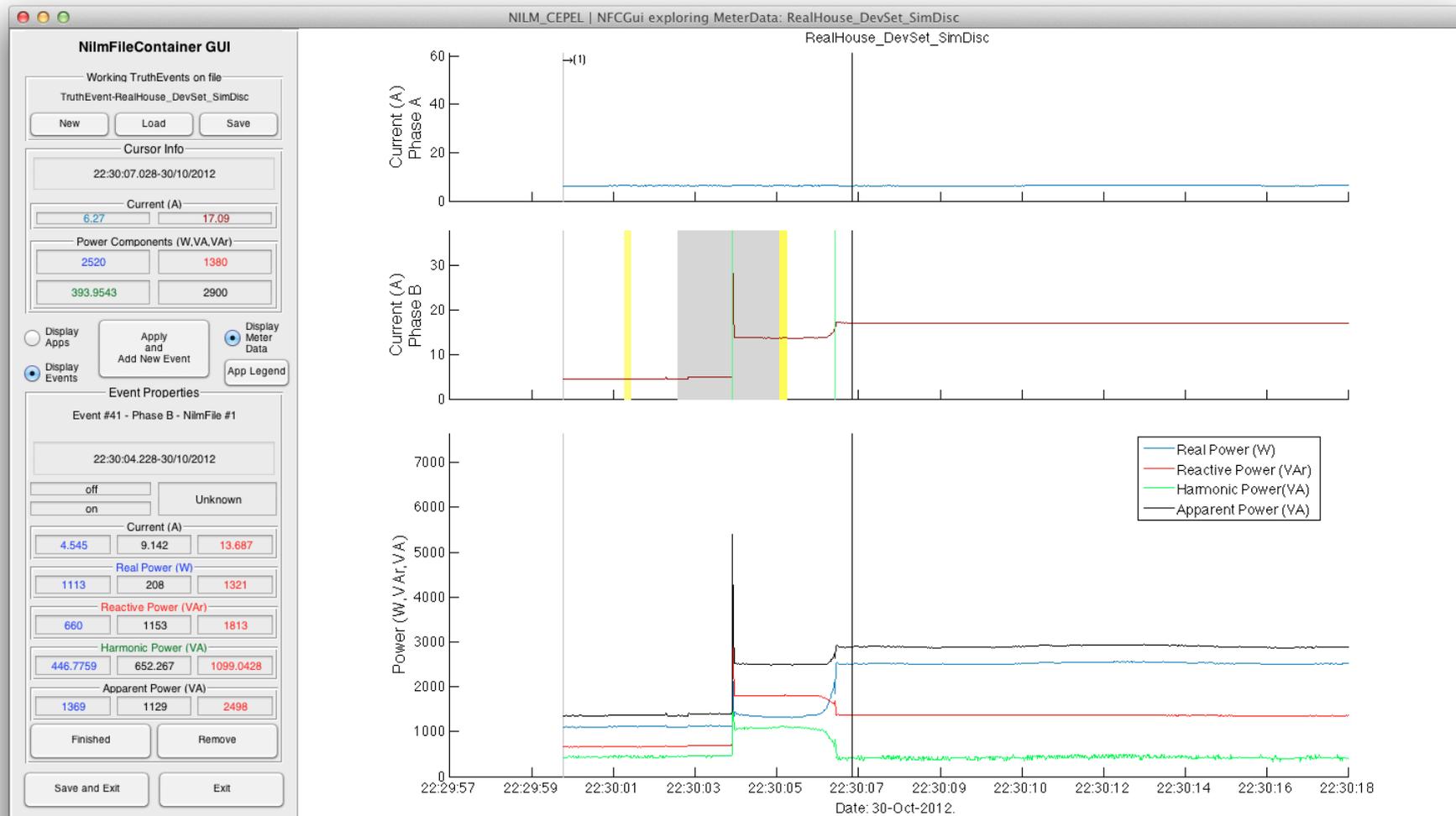


Figura 5.5: Informação gráfica para o Módulo de Interação Gráfica com os Dados: Evento de Transitório Corrigido. A sobreposição foi corrigida ao arrastar a região com o ponteiro, o que resulta em uma estimativa de consumo mais fiel.

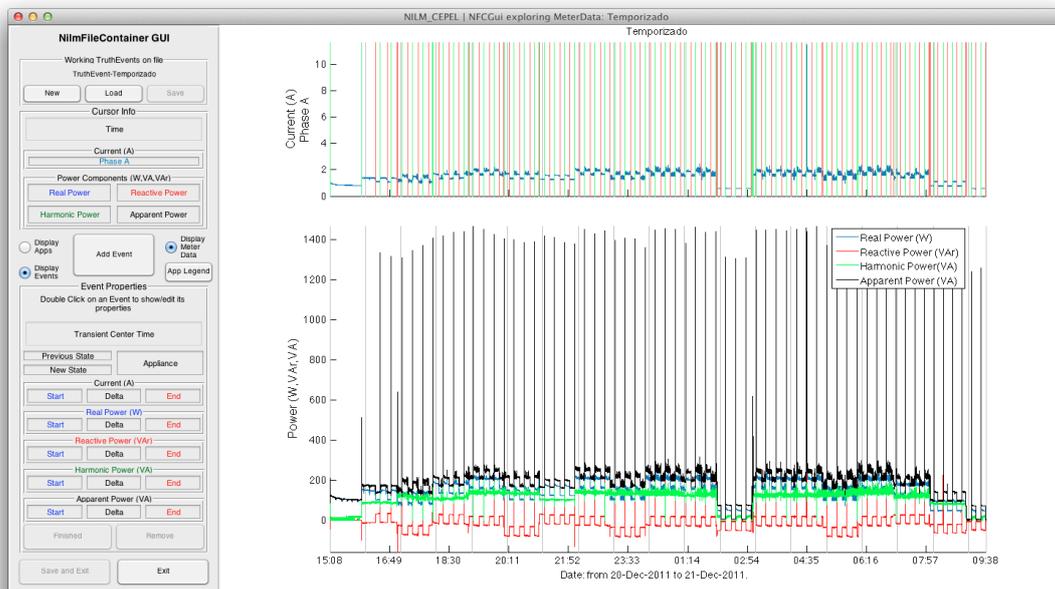
criados pelo usuário no arquivo em questão); e consumo amostrado pelo medidor. Isso é realizado porque a junção de toda essa informação de uma só vez acaba criando confusão e dificuldade para a sua interpretação, assim, com essas opções o usuário tem controle sobre elas, permitindo que haja comparação das informações sem que haja sobreposição delas;

- Armazenar e carregar arquivos: como o gabarito na verdade é uma estimativa da informação desagregada, é possível gerar diversos gabaritos para um mesmo conjunto de dados. Por isso, há um mecanismo de proteção da memória persistente e transitória. Ou seja, enquanto o usuário altera a informação do gabarito, o mesmo permanece intacto em disco, assim como se o usuário tentar encerrar a interface gráfica ou carregar um outro gabarito e houver assincronia entre a informação na memória persistente e transitória, ele será perguntado se deseja descartar suas alterações ou sincronizar a informação;
- Janela de legenda para o consumo dos equipamentos: a informação do consumo temporal possui uma cor única para cada equipamento, porém adicionar essa legenda na própria figura com os dados do medidor e informação do gabarito não era possível sem gerar dificuldade para sua interpretação, bem como não haveria espaço suficiente para colocar todos os nomes dos equipamentos (espera-se cerca de 30 a 50 equipamentos nas residências) na tabela sem comprometer a figura. Por isso, o usuário tem a opção de abrir uma janela que irá conter essa legenda (Figura 5.6b), que será atualizada automaticamente enquanto o usuário preenche o gabarito.

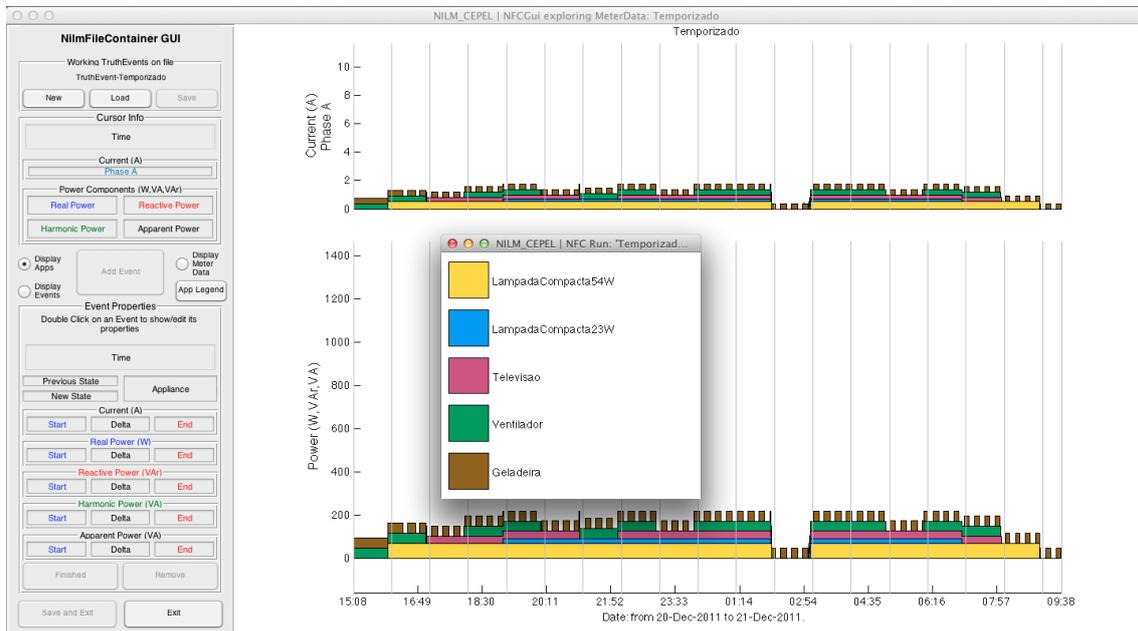
5.4 Análise dos Dados

O Módulo de Análise dos Dados foi implementado para refletir o ponto inicial da metodologia empregada pelo CEPEL/Eletrabras, mas adaptando o mesmo para o ambiente de análise, corrigindo possíveis pequenos equívocos no código original e expandindo o mesmo. Suas aptidões são:

- Continuidade da análise: a versão de ponto de partida gerava um evento na leitura de cada novo arquivo. Isso ocorria porque o consumo inicial em cada arquivo raramente é zero, e a inicialização das condições do filtro pelo *Matlab* é realizada considerando que ele estava em repouso recebendo entradas nulas e respondendo valores também nulos. Para dar a continuidade entre as segmentações dos dados o próprio *Matlab* fornece o vetor de condições finais \underline{z}_f no final da leitura de um arquivo para ser aplicado na leitura do próximo arquivo como condições iniciais \underline{z}_i . Porém, no caso do conjunto de dados e nas



(a) Informação disponível ao selecionar amostragem do medidor e eventos de transitório para o conjunto de dados *Temporizado*.



(b) Informação disponível ao selecionar a informação estimada de consumo desagregado por equipamento e a janela de legenda para o conjunto de dados *Temporizado*.

Figura 5.6: Informação gráfica para o Módulo de Interação Gráfica com os Dados: Disposição da informação.

possíveis reinicializações da análise devido à descontinuidade das amostras, era necessário determinar como simular a condição onde o filtro estava em repouso (com respostas nulas em estado permanente) para a entrada com o valor da amostra inicial. Por exemplo, no caso da Figura 5.2 há uma descontinuidade às 07:15 do dia 31 causada por alguma falha na medição, que, ao executar a convolução do filtro no período posterior a essa falha, irá gerar um evento de transitório falso uma vez que haverá sensibilização na resposta do FIR pelo motivo referido.

Seja, assim, a determinação da resposta $y(m)$ para um FIR dada por 5.1a [106, pp. 311-312]⁵, onde $b(k)$ é o k -ésimo coeficiente do filtro de ordem $n-1$. Deseja-se encontrar o vetor \underline{z}_i que, dado o vetor de amostras de entrada \underline{x} , gerará uma resposta \underline{y} constante e nula para toda janela do filtro. Ao representar 5.1a em forma matricial e igualando $\underline{y} = 0$, obtém-se 5.1b. No caso, quer-se simular que o filtro estava em repouso para a primeira amostra do arquivo, bastando fazer $\underline{x} = a_1$, onde a_1 é a primeira amostra do conjunto de dados ou a primeira amostra após uma descontinuidade. Assim, utiliza-se 5.1b com o valor de a_1 para todo o vetor \underline{x} ;

$$\begin{aligned}
 y(m) &= b(1)x(m) + z_1(m-1) \\
 z_1(m) &= b(2)x(m) + z_2(m-1) \\
 \vdots &= \quad \quad \quad \vdots \quad + \quad \quad \quad \vdots \\
 z_{n-2}(m) &= b(n-1)x(m) + z_{n-1}(m-1) \\
 z_{n-1}(m) &= b(n)x(m)
 \end{aligned} \tag{5.1a}$$

$$\underline{z} = \underbrace{\begin{bmatrix} b(n-1) & b(n-2) & b(n-3) & \dots & b(1) \\ & b(n-2) & b(n-3) & \dots & b(1) \\ & & b(n-3) & \dots & b(1) \\ & \mathbf{0} & & \ddots & \vdots \\ & & & & b(1) \end{bmatrix}}_{\mathbf{B}'} \underline{x} \tag{5.1b}$$

- Pareamento da resposta do FIR com os dados: como a análise é *a posteriori*, para facilitar a interpretação, remove-se o atraso na resposta do FIR para que a mesma fique alinhada com a amostra sendo analisada, facilitando a compreensão do problema;
- Janela do filtro com apenas com valores relevantes: o tamanho da janela do filtro tem influência direta no tempo de execução do algoritmo, uma vez que

⁵Referência utilizada pelo *Matlab* na implementação do FIR.

a convolução será realizada com uma janela maior de pontos para cada uma das amostras analisadas. Diminuir a janela em duas amostras significa que, para cada amostra no arquivo, serão realizados pelo menos dois cálculos a menos. Assim, além de dar um valor limite para o tamanho da janela do filtro, é necessário podar o mesmo para conter apenas valores relevantes, desconsiderando pontos com grandeza irrelevante. Para isso, realiza-se um corte para pontos com ordem de grandeza 10^{-3} menores que o ponto de maior relevância do filtro. Apenas com esse corte já foi possível obter grandes reduções no tamanho da janela e conseqüentemente no tempo de execução do algoritmo;

- Reconhecimento de múltiplas análises com mesma configuração: suponha que se deseje realizar quatro análises, duas delas com um filtro com $\sigma = 2$ e valores de corte $\Delta I_{min} = 0,1$ e $\Delta I_{min} = 0,2$, enquanto os outros dois filtros irão ter $\sigma = 3$ com os mesmos cortes. Nesse caso, não é necessário gerar a respostas dos filtros de derivada de Gaussiana quatro vezes, apenas duas vezes e então aplicar os dois cortes em cada uma dessas respostas. O ambiente de análise realiza a identificação dessas ocorrências quando executando múltiplas análises e as executa apenas uma vez, partilhando essa memória para os algoritmos seguintes;
- Informação gráfica: a Figura 5.7 contém um exemplo de gráfico gerado para uma análise. É possível observar a resposta do filtro de derivada de Gaussiana, as regiões sensibilizadas, os eventos, suas janelas para cálculo das FEX e os seus estados (ver pp. 115) informando se os mesmos foram aceitos ou foram eliminados devido a algum dos cortes.

5.5 Otimização dos Parâmetros

Retornando agora para o esboço do ambiente (Figura 5.1), observa-se que o Módulo de Otimização dos Parâmetros utiliza o Módulo de Análise iterativamente para obter parâmetros ótimos (não necessariamente o ótimo global). A otimização é feita em posse do gabarito, onde os parâmetros da análise são alterados pelo otimizador em busca de valores que melhor adequem a resposta da análise em relação ao gabarito. Os seguintes tópicos resumem como é realizada a otimização dos parâmetros:

- Redução de memória transitória: um detalhe operativo, as análises geradas pelo usuário mantém a informação da resposta do filtro de Gaussiana após o término da análise pois essa informação é necessária para gerar a informação gráfica, porém, no caso do otimizador diversas análises serão geradas, sendo interessante limitar ao máximo o consumo de memória em cada uma delas. Por

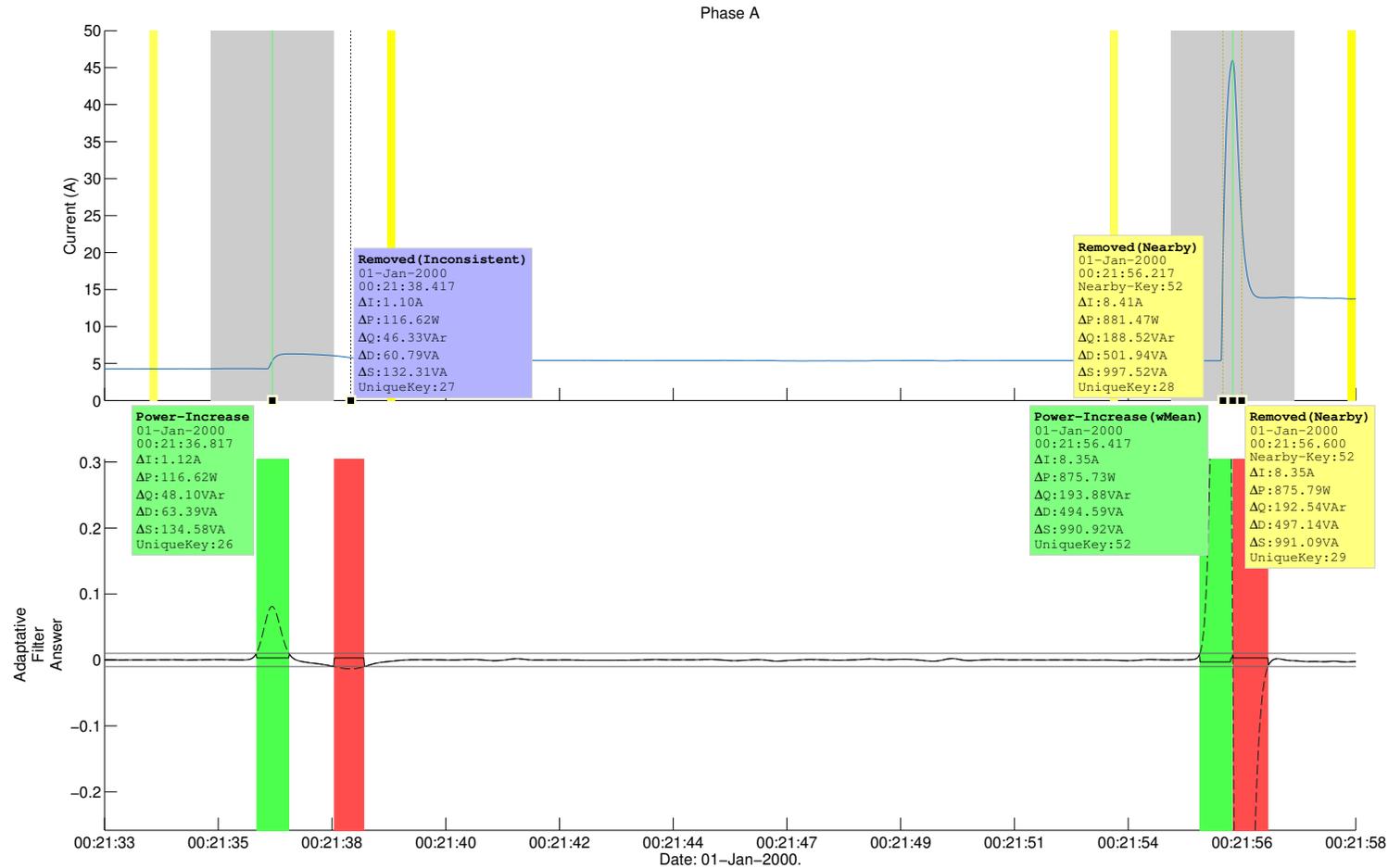


Figura 5.7: Exemplo de informação gráfica para o Módulo de Análise dos Dados. Na subfigura inferior, as regiões verdes e vermelhas indicam regiões sensibilizadas por respostas positivas e negativas, respectivamente. A resposta para o filtro de derivada de Gaussiana é representado pela linha pontilhada, enquanto a linha horizontal cinza é o limiar de corte para a geração de uma região sensibilizada. É possível observar um caso de evento inconsistente e outro removido devido a evento próximo. Para o caso do evento inconsistente, em azul, seu degrau de potência é positivo enquanto sua resposta é negativa, revelando sua inconsistência. Já os eventos próximos representados pelas caixas amarelas foram removidos por estarem próximos, sendo substituídos pela sua média (a linha verde). Caso esses eventos não fossem removidos, eles seriam constituídos de falsos alarmes. Na subfigura superior é possível observar as regiões que serão utilizadas para a extração do transitório (região cinza) e as regiões utilizadas para calcular o degrau de potência (regiões amarelas pré/pós-transitório).

isso, as análises geradas pelo algoritmo remove qualquer informação irrelevante durante a execução da análise como a resposta do filtro e qualquer outros elementos não necessários, mantendo apenas os eventos de transitório;

- Regras de Pontuação: para realizar a otimização, é necessário haver uma regra para o otimizador avaliar aquilo que é desejável do que não é. É natural que essa regra tome como recompensa identificar corretamente um evento de transitório e como punição gerar um evento de transitório aonde essa informação não existe. O primeiro caso é referido de detecção, enquanto o segundo consiste de um falso positivo ou falso alarme. Também é possível adicionar outras punições, por exemplo, não é desejável que sejam gerados muito candidatos para serem removidos, isso irá reduzir a velocidade de processamento e se uma solução consegue realizar a detecção com acurácia parecida mas gerando menos candidatos, é preferível optar por essa opção, mesmo que ela tenha uma performance ligeiramente reduzida para obter uma performance de execução quando aplicando o algoritmo em um NILM operando em tempo real. Porém, essa punição deve ser pequena, por não ser a grande questão a ser otimizada. A regra de pontuação utilizada está representada em 5.2. Ainda, os eventos detectados pelo Módulo de Análise não estarão exatamente na mesma posição que os eventos gerados pelo usuário no gabarito, já que o usuário não necessariamente irá marcar a amostra que constitui o ponto de inflexão da resposta do filtro e, mesmo que isso ocorra, os eventos podem sofrer deslocamento quando empregando a remoção de eventos próximos por sua média. Assim, é preciso definir uma janela de um número máximo de amostras J_{max} para os quais se aceitará o evento de detecção. O casamento do evento da resposta da análise com o do gabarito é feito quando os dois se encontram no máximo à J_{max} amostras de distância. Caso existam mais de um evento dentro dessa janela, o casamento é realizado com o evento mais próximo;

$$\text{Aptidão} = \gamma_{det}N_{det} + \gamma_{fa}N_{fa} + \gamma_{rem}N_{rem} \quad (5.2)$$

onde:

Aptidão mede a capacidade de resposta da análise realizada, sendo de interesse maximizar esse valor.

γ_{det} é a pontuação que a análise recebe para cada evento de detecção;

N_{det} é a quantidade de eventos detectados;

γ_{fa} é a pontuação que a análise recebe para cada ocorrência de falso alarme;

N_{fa} é a quantidade de ocorrências de falso alarme;

γ_{rem} é a pontuação que a análise recebe para cada ocorrência de candidatos removidos;

N_{rem} é a quantidade de candidatos removidos.

- Comparação da resposta da análise com o gabarito: em posse da regra de pontuação 5.2 e a janela máxima de amostras para correlacionar a informação de análise com a do gabarito (J_{max}), realiza-se a comparação entre as duas informações e retorna-se a eficiência em termos de Aptidão. Como dito no item anterior, J_{max} é necessário para casar os eventos de transitório da resposta do filtro com aqueles presentes no gabarito;
- Escolha do algoritmo: a função a ser otimizada não é diferenciável e portanto não é possível utilizar os métodos convencionais de otimização. É necessário empregar algum método de tentativa e erro para realizar essa tarefa. Neste trabalho, optou-se pela utilização de um ES, porém outras estratégias de otimização podem ser utilizadas, como Inteligência de Enxame. Aprofundar-se-á nas características do algoritmo implementado na Subseção 5.5.1;
- Capacidade de recuperação do processo caso ocorra alguma falha: como as otimizações podem levar dias para serem executadas, viu-se a necessidade de armazenar informações relevantes do processo enquanto ele evolui, para garantir que se ocorresse algum problema na máquina em execução, não houvesse de recomeçar o processo desde a primeira geração. Assim, a versão do algoritmo é capaz de armazenar, se o usuário requerir, a evolução do processo de otimização e retornar caso o processo seja interrompido — por exemplo, devido à falta de energia.

5.5.1 Algoritmo Genético de Estratégia Evolutiva (ES)

Foi realizada a implementação de uma versão própria de um ES com base em [107, cap. 4], mas antes de entrar em detalhes sobre a versão implementada cabe introduzir o tema sobre algoritmos de estratégia evolutiva.

Dispõe-se na Figura 5.8 a sequência de otimização de um algoritmo evolutivo genérico. Ela conta com uma população inicial, onde se recomenda que a mesma seja iniciada aleatoriamente. Essa população inicial passará por um processo de seleção parental aonde serão obtidos os espécimes ou indivíduos para compor a população dos pais (μ), população que irá realizar a propagação de sua informação genética para a população da prole (λ). Porém, a população da prole recebe o material perturbado através de operações de recombinação e mutação. A mutação é uma perturbação que ocorre somente levando em conta o material de um único pai, enquanto a recombinação ocorre no mínimo com dois pais. A capacidade de

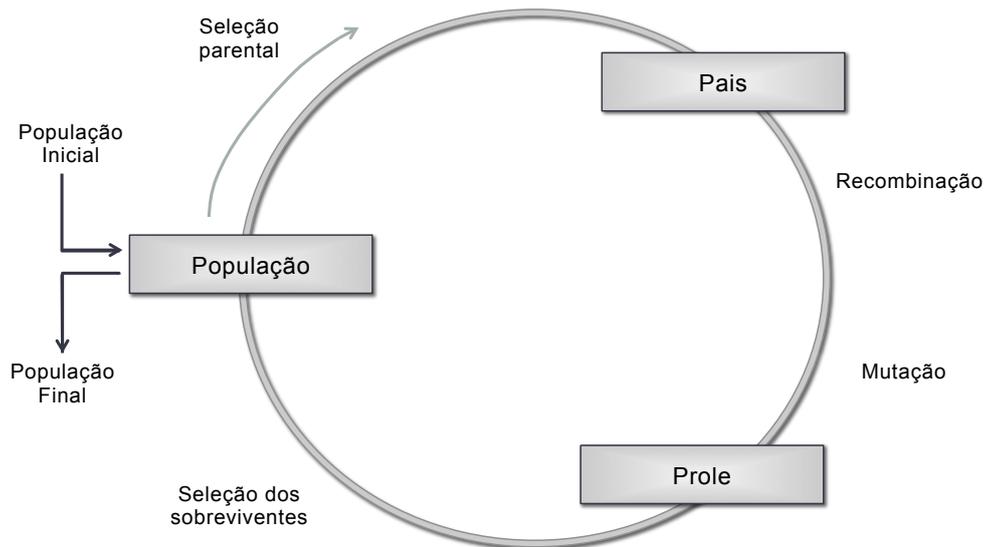


Figura 5.8: Esboço de um algoritmo evolutivo genérico. Baseado em [107, pp. 17].

encontrar novas regiões promissoras (genes alelos) — descoberta — no espaço de busca de soluções e, com isso, adquirir informação no problema é dada pela mutação através de perturbações aleatórias. Já a recombinação realiza a otimização dentro de uma área promissora (no caso, dentro da informação genética dos pais) — exploração —, dando grandes pulos para uma região dentro de duas áreas. O resultado do material genético perturbado constitui a base para a geração da prole. Irá ocorrer, então, a seleção dos sobreviventes, que pode levar em consideração a prole e os pais nesse processo $(\mu+\lambda)$ ou apenas a prole (μ,λ) . Os indivíduos selecionados são a nova geração da população, e o ciclo será repetido gerando novas populações até o processo atender uma condição de parada. Geralmente essa é através de um número máximo de gerações.

Cabe ainda definir a diferença entre genótipo e fenótipo do ponto de vista computacional. A representação em genótipo é aquela que sofre as perturbações e codifica a representação do espécime, enquanto o fenótipo representa a informação como é demonstrada pelo espécime para o problema em questão, ou seja, o espaço de solução. Pode haver diferença entre as duas representações ou não, por exemplo, a representação {Norte,Leste,Sul,Oeste} seria as possíveis representações do fenótipo, e sua codificação em genótipo poderia ser feita em {1,2,3,4}. Nota-se a importância de não só saber representar a informação em genótipo para realizar as operações de perturbação no material, como possuir uma maneira de decodificá-la novamente no espaço do fenótipo. Para isso, cada solução do fenótipo deve ser mapeável, bem como cada genótipo tenha a sua decodificação em apenas uma solução. Ainda, a escolha da representação irá afetar o problema: no exemplo citado a representação escolhida não parece ter o significado do fenótipo, uma vez que Norte e Oeste são vizinhos entre si, enquanto na representação eles estão distantes de três unidades.

Uma escolha em variáveis cíclicas parece representar o problema fidedignamente $\{(\text{sen}(\frac{1\pi}{2}), \text{cos}(\frac{1\pi}{2})), (\text{sen}(\frac{2\pi}{2}), \text{cos}(\frac{2\pi}{2})), (\text{sen}(\frac{3\pi}{2}), \text{cos}(\frac{3\pi}{2})), (\text{sen}(\frac{4\pi}{2}), \text{cos}(\frac{4\pi}{2}))\}$, porém, nesse caso, o genótipo seria representado em duas variáveis.

Versão original

O ES é um algoritmo genético cuja especialidade é a autoadaptação de sua estratégia evolutiva. O ES concentra sua capacidade evolutiva na mutação, enquanto outros algoritmos genéticos costumam focar na recombinação. Todos os indivíduos passam por perturbações Gaussianas em seu material genético, no entanto, a ordem dessas perturbações são ajustadas conforme a evolução da espécie, aumentando ou diminuindo sua ordem de acordo com as necessidades de evolução. Assim, o ES irá aumentar a ordem de suas perturbações quando distante de um valor ótimo — sujeito a capacidade de perceber uma tendência no espaço de solução apontando na direção de um ótimo — e reduzir as perturbações conforme se aproxima desse valor para realizar o ajuste fino. Há também uma pequena taxa de recombinação para aumentar a velocidade de convergência, normalmente na faixa de 10%⁶. A seleção parental não é influenciada pela aptidão dos indivíduos, quando utilizando taxa de recombinação a escolha dos pais para haver troca de material genético é realizado de maneira aleatória uniforme. A melhoria gradual das gerações é realizado pela seleção dos sobreviventes, que é realizada através de (μ, λ) com pressão alta de seleção, aonde apenas os indivíduos mais aptos permanecem para a próxima geração. Entende-se como pressão de seleção dos sobreviventes a grandeza descrita por 5.3. É importante a seleção através de (μ, λ) em comparação com a $(\mu + \lambda)$ para evitar ótimos locais, bem como garantir que não haverá propagação de indivíduos com estratégias mal-adaptadas através das gerações. O mesmo se dá para versões que utilizam elitismo — manter ao menos uma cópia do membro mais apto da população na próxima geração —, não sendo recomendado no ES pelo mesmo problema da seleção parental através de $(\mu + \lambda)$. É importante também que a seleção de sobreviventes aplique uma alta pressão de seleção, garantindo a capacidade adaptativa do ES, normalmente utilizando $\frac{\lambda}{\mu} = 7$.

$$\text{Pressão de Seleção} = \frac{\lambda}{\mu} \tag{5.3}$$

Optou-se pela mutação descorrelacionada com n tamanhos de passo [107, pp. 76–78]. Nessa configuração, cada variável representado no genótipo tem sua própria perturbação, sendo descrita por 5.4a, enquanto sua perturbação é adaptada anteriormente de acordo com 5.4b. A taxa de aprendizado é dividida em dois parâmetros, τ e τ' , onde aquele é a base de aprendizado, que garante uma mudança global na muta-

⁶As taxas de recombinação e mutação são dadas em termos de probabilidade de ocorrência.

bilidade para preservar os graus de liberdade do problema, e este é uma mutação específica por coordenada, fornecendo flexibilidade para empregar diversas estratégias de mutação em diferentes direções. Os valores indicados para ambos são $1/\sqrt{2n}$ e $1/\sqrt{2\sqrt{n}}$, respectivamente. O alcance da perturbação no material genético, dado uma determinada probabilidade de ocorrência fixa, formam elipsoides no espaço de solução alinhadas com os eixos da representação escolhida.

$$x'_i = x_i \sigma'_i N_i(0, 1) \quad (5.4a)$$

$$\sigma'_i = \sigma_i e^{\tau' N(0,1) + \tau N_i(0,1)} \quad (5.4b)$$

onde:

$N(0, 1)$ e $N_i(0, 1)$ são uma perturbação Gaussiana com média zero e σ unitário, a primeira sendo um único valor para todas as representações, enquanto a segunda uma para cada representação i ;

x_i e x'_i é a i -ésima representação e a mesma após sofrer a perturbação;

σ_i e σ'_i é a i -ésima estratégia evolutiva e a mesma após sofrer a perturbação.

Para a recombinação, implementou-se a versão local da mesma por ela ser mais simples de elaborar, pretendendo alterar no futuro para a recombinação global que é mais indicada para o caso do ES. Porém, a recombinação utilizada não é o quesito principal do algoritmo, servindo apenas para melhorar a velocidade de convergência. No caso implementado, o material genético é mesclado entre apenas dois indivíduos. A informação genética que representa o espaço de solução é misturada através de 5.5a — recombinação discreta —, enquanto a versão para a estratégia evolutiva é realizada através de 5.5b — recombinação intermediária.

$$\begin{cases} x'_{i,1} = x_{i,1} \text{ ou } x_{i,2} & \text{escolhidos aleatoriamente} \\ x'_{i,2} = x_{i,o.c.} \end{cases} \quad (5.5a)$$

$$\begin{cases} \sigma'_{i,1} = \frac{(\sigma_{i,1} + \sigma_{i,2})}{2} \\ \sigma'_{i,2} = \frac{(\sigma_{i,1} + \sigma_{i,2})}{2} \end{cases} \quad (5.5b)$$

onde:

$x_{i,1}$ e $x_{i,2}$ são a i -ésima representação para o primeiro e segundo pai, respectivamente;

$x_{i,o.c.}$ é a i -ésima representação para o pai não selecionado para $x'_{i,1}$;

$\sigma_{i,1}$ e $\sigma_{i,2}$ são a i -ésima estratégia evolutiva para o primeiro e segundo pai, respectivamente;

$x'_{i,1}$ e $x'_{i,2}$ são a i -ésima representação para o primeiro e segundo pai após a perturbação, respectivamente;

$\sigma'_{i,1}$ e $\sigma'_{i,2}$ são a i -ésima estratégia evolutiva para o primeiro e segundo pai após a perturbação, respectivamente.

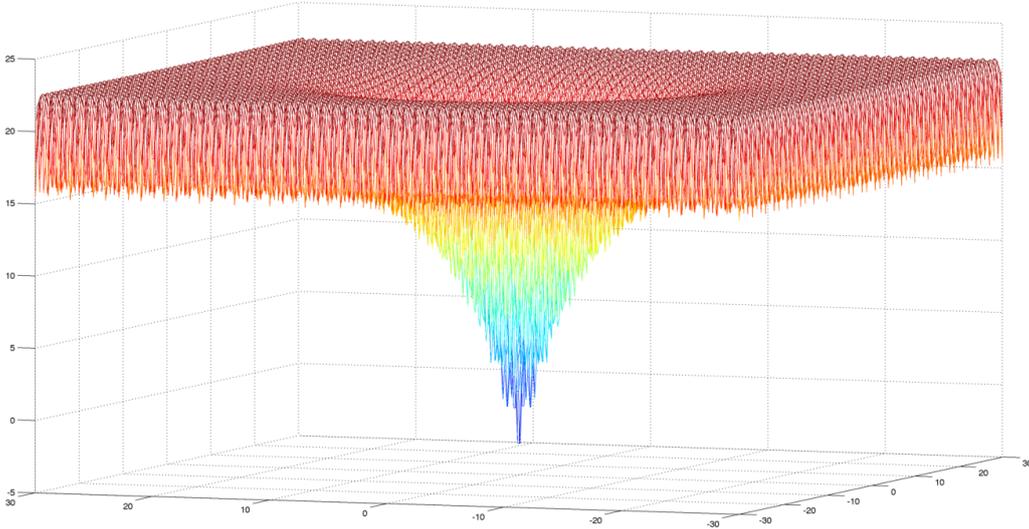


Figura 5.9: Função de *Ackley* em duas dimensões para $-30, 0 < x_i < 30, 0$.

A referência [107, pp. 84] cita um exemplo de um outro autor que aplicou o ES para a função de *Ackley*, onde foi utilizado $\mu = 30$, $\lambda = 200$, x_i inicial entre $-30, 0 < x_i < +30, 0$ e um total de 200.000 avaliações da função de aptidão. Para um total de 10 execuções, o exemplo citado obteve a melhor solução com valor da função de $7,48 \times 10^{-8}$. A função de *Ackley* (Figura 5.9) é altamente multimodal, com um grande número de mínimos locais, mas com apenas um máximo global $\bar{x} = 0$ e seu valor $f(\bar{x}) = 0,0$. Para validar o algoritmo implementado, executou-se o ES para essas configurações, obtendo uma ocorrência de convergência para mínimo local com o valor de $1,34$, e todas as outras ocorrências dão uma aptidão de média de $1,87 \times 10^{-7} \pm 2,56 \times 10^{-7}$, onde a melhor solução obtém a aptidão $1,34 \times 10^{-8}$. A evolução da execução com melhor convergência está na Figura 5.10, mostrando o valor médio de aptidão na geração da população na convergência. Apesar do exemplo citado informar que todos os mínimos encontrados foram os globais, na versão aqui implementada, ocorrem casos em que não há a convergência para o mínimo global, ainda que em outras execuções seja possível encontrar todas as 10 minimizações convergindo para o mínimo global. É importante ter em mente que a convergência não ocorre necessariamente para o mínimo global, sendo uma propriedade bem conhecida dos algoritmos genéticos. De qualquer forma, os valores obtidos estão próximos da referência e existem parâmetros não informados como o valor mínimo de perturbação σ_{min} e o valor inicial para as perturbações $\sigma_{inicial}$ que

podem influenciar na resposta. Apenas como referência, os valores utilizados para esses casos foram $\sigma_{min} = 1 \times 10^{-9}$ e $\sigma_{inicial} = 1$.

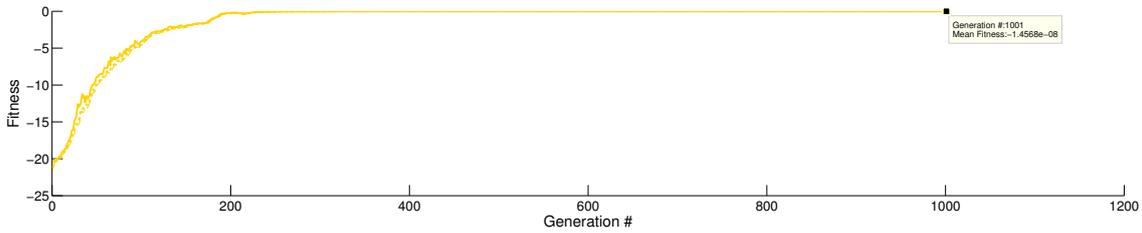


Figura 5.10: Evolução para o melhor indivíduo para a validação da versão original do ES. O objetivo é minimizar a função de *Ackley*, que pela visão do ES funciona como maximizar a função com seus valores opostos. Por isso, os valores mostrados são negativos.

Versão Multiespécie

Poderia ser utilizada a versão original do ES para realizar a otimização dos parâmetros necessários na abordagem do problema. No entanto, devido à decorrência da dúvida quanto a qual caminho percorrer para a análise (ordem de remoção de eventos e quais delas empregar), são necessárias diversas execuções do algoritmo para otimizar os valores de maneira individual. Ao invés de executar cada uma delas individualmente, motivado pela ideia de otimização multiobjetivo [107, cap. 9], decidiu-se utilizar a ideia de subpopulações — que serão referidas por espécies, por poderem ter cromossomos diferentes dependendo da configuração utilizada —, mas aplicando a mesma para um outro conceito. No caso, ao invés de utilizar espécies para otimização de múltiplas funções objetivo, esse conceito irá ser utilizado para criar uma dinâmica entre as várias otimizações sendo realizadas no problema.

A ideia da dinâmica é reservar o esforço computacional para aquelas abordagens que estão mostrando capacidade de resolver o problema com maior aptidão, revelando-se uma espécie mais adequada para o *habitat* em que os espécimes estão sendo avaliados. Porém, isso deve ser realizado sem comprometer a evolução de espécies que, por algum motivo, sofreram desvantagem durante o processo evolutivo — seja por uma inicialização em condições desprivilegiadas, ou por uma demora maior para ajustar sua estratégia evolutiva. Assim, a proposta é executar apenas uma otimização para as diferentes maneiras de tratar o problema, aonde todas as configurações desejadas irão competir entre si de modo que o algoritmo irá reservar maior esforço computacional para otimizar mais profundamente aquela que se melhor adéqua ao espaço de solução, diferente da versão aonde se executaria para cada espécie, reservando esforço computacional igual para espécies que não têm se mostrado adequadas para a solução do problema.

Assim, propuseram-se duas configurações para as competições dos espécimens:

- **Interespécie:** nesse caso há cooperação entre os indivíduos de uma mesma espécie. É calculada a aptidão para cada espécie (5.6a) a ser utilizada como parâmetro na competição das mesmas e determinar a parcela da população global que elas tem direito. Uma vez determinado o tamanho da população de cada espécie, seus indivíduos irão competir entre si para determinar os sobreviventes. É necessário escolher um método para avaliar a aptidão das espécies e como determinar suas populações através dele;
- **Intraespécie:** os espécimes disputam entre si na população global independente de qual espécie pertencem. Nessa configuração não há cooperação entre os indivíduos de uma mesma espécie, apenas os melhores da população global sobrevivem;

Para evitar que espécies não tenham a oportunidade de se desenvolverem antes que sua população seja drasticamente reduzida ou até mesmo extinta, tratou-se cada um dos casos individualmente. No caso da seleção interespecie, é necessário escolher uma função que privilegie espécies mais aptadas, mas que uma diferença de aptidão muito grande — que irá ocorrer em especial durante o início da evolução devido a espécies condicionadas em ambientes mais propícios que outras — não elimine toda a diversidade das populações antes que elas adequem sua estratégia evolutiva. Para isso, escolheu-se empiricamente a função 5.6b, para suavizar a pressão aplicada em espécies menos aptas. A população reservada para uma espécie para a próxima geração é dada por 5.6c. Entretanto, como se utiliza a função de transformação em inteiro *floor*, ao somar μ'_i para cada espécie pode acabar resultando em uma população menor que μ . Assim, distribui-se aleatoriamente os indivíduos faltantes nas espécies de forma que a soma dos μ'_i seja o mesmo que μ .

Para a avaliação da configuração intraespécie, bem como interespecie, será utilizada a função de *Ackley* contendo 10 espécies. As espécies têm de 20 a 29 dimensões, sendo rotuladas sequencialmente de 01 a 10. As espécies de menores índices tem uma carga genética mais simples de ser adaptada através da estratégia evolutiva por possuir menos parâmetros livres que aquelas com índices maiores, porém todas espécies tem a mesma capacidade de resolver o problema — todas podem obter o mínimo em zero. Ao observar a Figura 5.11, percebe-se que as espécies inferiores tendem a tomar conta da população nas primeiras gerações confirmando essa hipótese. Por outro lado, conforme a evolução ocorre, há uma tendência para o equilíbrio, uma vez que, nesse caso, todas as espécies são capazes de resolver igualmente bem o problema. Percebe-se que a espécie 05 (marrom) é aquela que mais sofre durante as gerações iniciais, porém, por volta da 30ª geração ela alcança o nível de desenvolvimento das outras espécies, crescendo gradualmente sua população.

$$\text{Aptidão}_{(i)} = \sum_{j=1}^{\lambda_i} \text{Aptidão}_{(i,j)} \quad (5.6a)$$

$$f_{inter}(i) = \log_2(\text{Aptidão}_{(i)} - \min(\text{Aptidão}_{(i)} | \forall i \in \Gamma) + 2) \quad (5.6b)$$

$$\mu'_i = \text{floor} \left(\frac{f_{inter}(i)}{f_{inter,norm}} \right) \quad (5.6c)$$

$$f_{inter,norm} = \sum_{\forall i \in \Gamma} f_{inter}(i) \quad (5.6d)$$

onde:

λ_i é o tamanho da população da prole da i -ésima espécie;

Γ é o espaço contendo todas as espécies;

μ'_i é o tamanho da população dos pais da i -ésima espécie para a próxima geração.

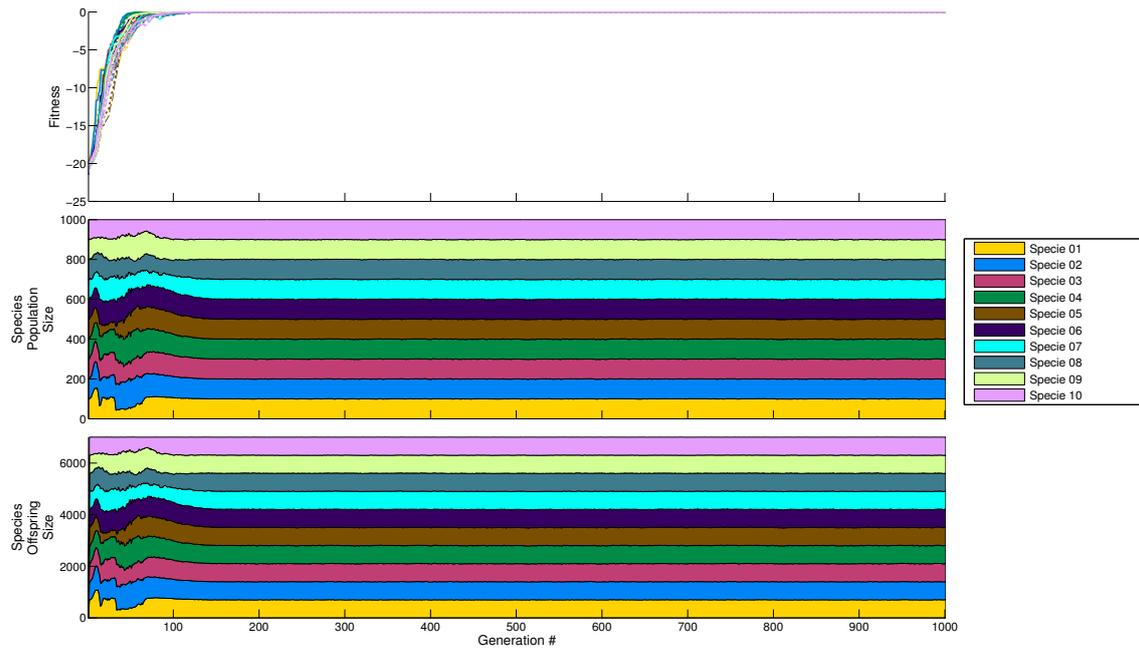


Figura 5.11: Competição interespecie. As linhas contínuas na subfigura superior indicam os indivíduos mais aptos de cada espécie. As subfiguras inferiores indicam a população para cada espécie, sendo a mais inferior a população para a prole, e o outro caso a população dos pais da espécie.

Um outro mecanismo foi implementado para impedir a extinção de uma espécie. Ele funciona como um órgão de proteção da diversidade de espécies, limitando que espécies próximas de entrarem em extinção tenham sua população reduzida, independente do quão mal esses indivíduos se adequam ao *habitat* para o qual estão sendo avaliados.

Isso foi especialmente importante para o caso de competição intraespécie, onde uma espécie ao encontrar um material genético de melhor qualidade em comparação com as outras, rapidamente tomava conta da população global por espalhar esse material entre sua população com grande velocidade. Na Figura 5.12, observa-se que, se não fosse esse mecanismo, a espécie azul ou amarela iriam extinguir todas as outras tomando conta da população global. Fica evidente também que apenas um órgão de proteção da diversidade de espécies não é suficiente para garantir a evolução das espécies, é necessário suavizar a competição, de modo que uma espécie que por algum motivo se tornou mais apta não extermine outras espécies rapidamente acabando com sua diversidade e não as dê a oportunidade para evoluir, já que os indivíduos que sobraram possivelmente ainda não ajustaram sua estratégia evolutiva. A Figura 5.13 mostra as pressões de seleção em ordens muito além daquelas que deveriam ocorrer, obtendo valores na ordem de 20 logo no início da evolução para as espécies que foram iniciadas em condições menos favoráveis, eliminando toda sua diversidade em poucas gerações. Já as espécies que conseguiram se desenvolver, observa-se que as mesmas ao conseguirem uma aptidão melhor que a da outra espécie rapidamente tomam conta da população global, acontecimentos marcados pelos picos na pressão de seleção das espécies antes dominantes. Observa-se também que apenas as espécies mais simples conseguiram evoluir o suficiente para convergir para o mínimo global.

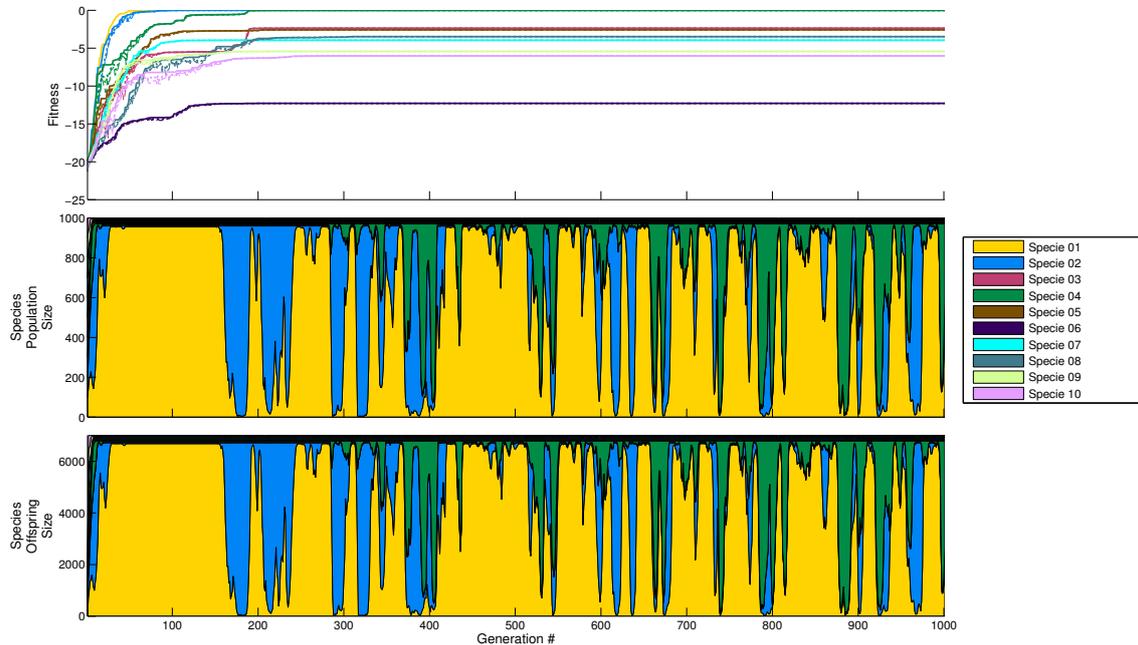


Figura 5.12: Competição intraespécies sem intervenção na competição.

Em vista disso, implementou-se um mecanismo de controle de pressão de seleção por espécie. Esse mecanismo aceita um valor máximo de pressão para cada espécie, se o valor ultrapassar o limiar, então se reduz sua pressão ao alocar espaço da população para essa espécie retirando das espécies com maior alocação de população

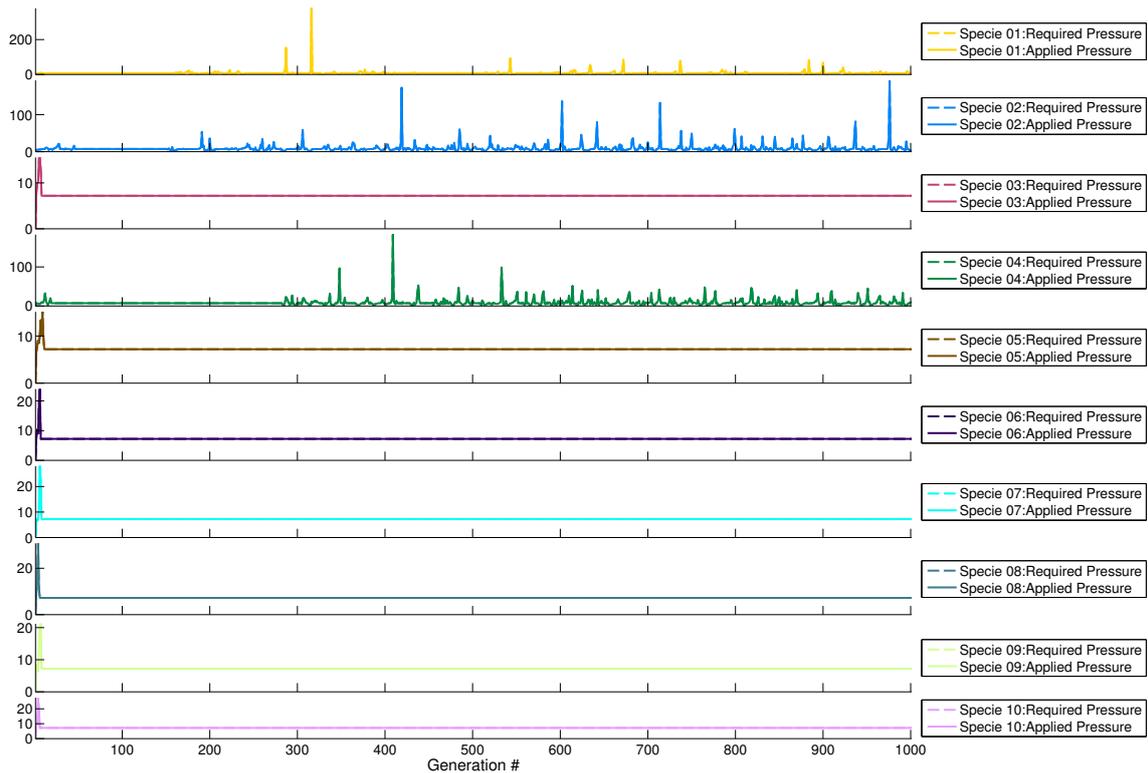


Figura 5.13: Pressão de seleção para competição intraespécie sem intervenção na competição.

até que a alocação de população dessas espécies que ultrapassaram o corte máximo de pressão resultem em um valor aceitável da mesma. A escolha de retirada da alocação de indivíduos é feita da seguinte maneira:

- Inicia-se reduzindo a alocação de população da espécie com maior população;
- Caso o valor da espécie de maior população atingir o tamanho da população de uma outra espécie, adiciona-se essa espécie para a redução de população e continua o processo. Caso isso ocorra novamente, a próxima espécie também será adicionada para redução de população e o processo continua até que seja determinado quais espécies irão ceder espaço para que a população das espécies com altas pressões satisfaça o critério mínimo;
- Quando há o agrupamento de espécies para redução e a necessidade de reduzir um número não inteiro de população em cada espécie agrupada, a escolha do resíduo da divisão é feita aleatoriamente. Ex. se houver de suprir 100 espécimes para garantir que a pressão de seleção seja feita em níveis aceitáveis, e houver 3 espécies tendo sua população reduzida, retira-se 33 indivíduos de cada uma delas, porém a escolha da espécie que perderá mais um indivíduo será realizada aleatoriamente uniforme.

A execução para uma pressão máxima de 7,3 pode ser visualizada na Figura 5.14, onde fica evidente a convergência da população para a máxima aptidão (ou mínimo da função da *Ackley*). Também se observa uma mudança menos brusca quando comparado à versão sem intervenção na competição, mostrando que o controle é importante para garantir mudanças mais suaves na configuração da população. Uma observação importante pode ser realizada quanto ao órgão de proteção de diversidade de espécies: não fosse sua operação, diversas espécies teriam sido extintas durante o processo evolutivo. Na Figura 5.15, observa-se a pressão de seleção requerida pela competição natural, e aquela aplicada pelo sistema de intervenção. Em certos casos, a competição natural chega a exigir pressões de até 500 unidades, o que acabaria com a diversidade de uma espécie em uma única geração. Assim, ao optar pela versão de competição intraespécies faz-se necessário interferir na competição para que todas as espécies tenham chances de evoluir. Ainda, as espécies de índices mais altos precisam de mais tempo para chegarem em seu auge. Isso ocorre pela maior dificuldade para o ajuste da estratégia de evolução devido a maior presença de parâmetros livres nessas espécies, bem como a menor alocação de população para sua evolução uma vez que essas espécies irão tender a se inicializar probabilisticamente com menor aptidão — a dispersão nas dimensões a mais tenderão a aumentar o erro e, conseqüentemente, reduzir a aptidão dessas espécies.

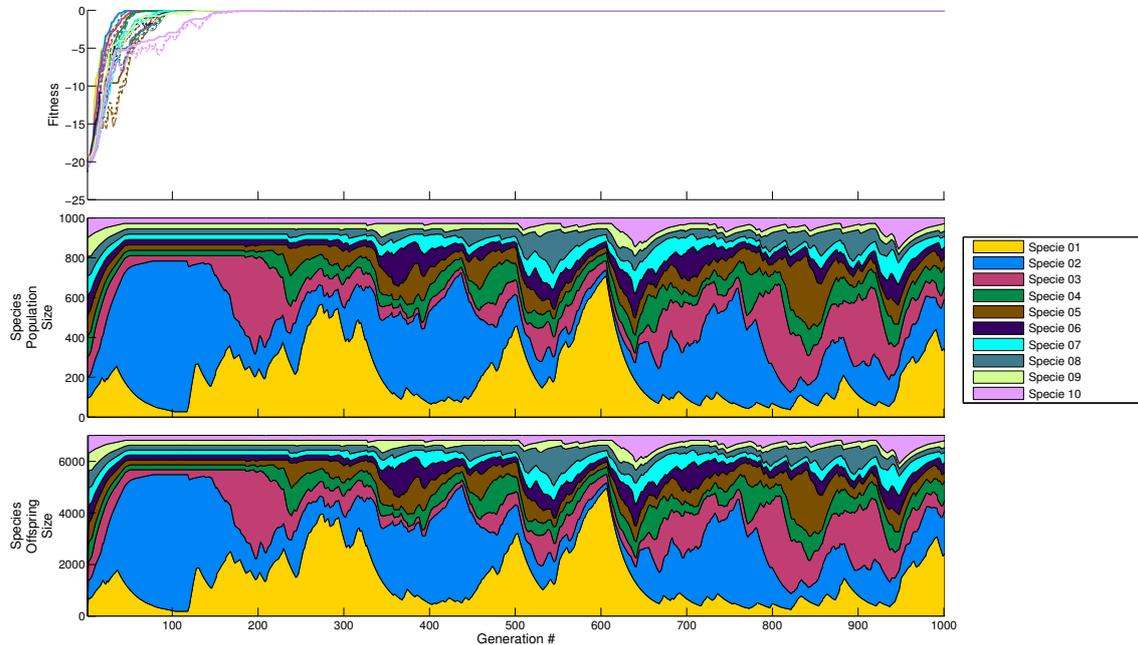


Figura 5.14: Competição intraespécies com intervenção na competição.

Há uma nítida diferença entre como as duas seleções se comportam. Ao comparar as figuras 5.11 e 5.14, percebe-se que o caso de competição interespecie tem uma mudança bastante tênue na configuração da população, privilegiando as espécies com melhor aptidão proporcionalmente à sua aptidão como espécie. No caso, a escolha

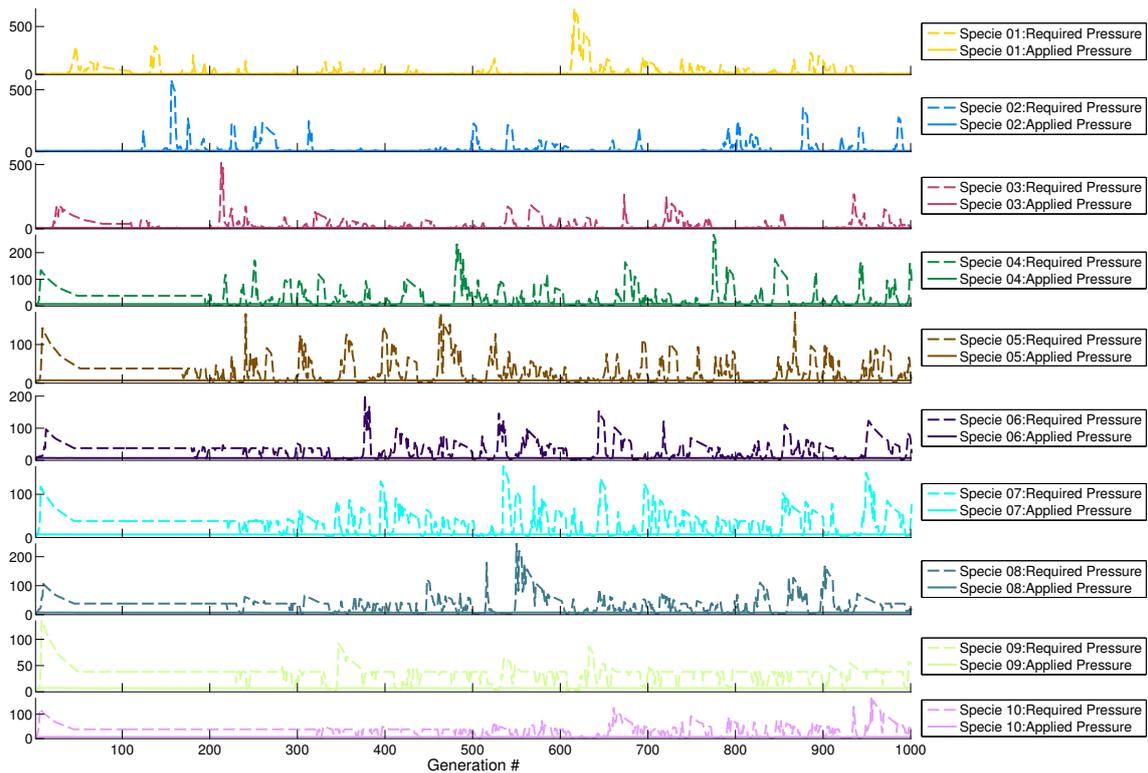


Figura 5.15: Pressão de seleção para competição intraespécie com intervenção na competição.

da função torna a vantagem pequena entre elas, já que se utiliza uma atenuação logarítmica. Enquanto isso, na competição intraespécie com intervenção (a versão sem intervenção não é recomendada) observa-se que o crescimento da população da espécie ocorre gradualmente conforme seus indivíduos ocupam posições privilegiadas no espaço de solução. Seria interessante adiar a competição, dando um número de gerações iniciais para os quais as espécies evoluiriam sem competir, deixando as mesmas ajustarem sua estratégia evolutiva e estarem em um estágio pré-evoluído para então começarem a competir entre si. Comparando essas mesmas figuras novamente, observa-se que o método de competição intraespécies, com a implementação desse novo mecanismo, favorece espécies com carga genética mais simples de ser otimizada ou que começaram em condições mais favoráveis.

Capítulo 6

Metodologia

Uma vez definido o ambiente de análise implementado e suas capacidades, cabe detalhar como o mesmo será aplicado para determinar os parâmetros da abordagem final para detecção de eventos de transitório e a base de dados que será trabalhada. Este capítulo começa realizando a descrição das características presentes nos conjuntos de dados trabalhados, que podem ser separados em dois grandes grupos: um contendo a operação de poucos equipamentos atuando em conjunto e praticamente nenhuma injeção de ruído devido à equipamentos C5 (ver Subseção 4.2.1) — os dados puros contidos nos arquivos NI00***¹; enquanto o outro conta com a operação de múltiplos equipamentos atuando simultaneamente (cerca de 5 a 8, em geral) e da presença de equipamentos C5, constituindo, por isso, de dados com menor relação sinal-ruído — conjuntos *Temporizado*, *Empilhado4* e *Empilhado7*. Todos os conjuntos de dados possuem apenas um gabarito, exceto o *Temporizado*, que possui dois gabaritos. O motivo para isso é a presença de um distúrbio durante a operação da televisão LCD nesse conjunto de dados, que *não* é considerado como alvo na configuração *Temporizado Gabarito 1*, enquanto a configuração *Temporizado Gabarito 2* contém essa informação como eventos de transitórios desejados como alvo.

Em seguida, detém-se atenção para quais configurações e parâmetros presentes na metodologia proposta serão otimizados pelo ES descrito na Seção 5.5.1 do capítulo anterior, aonde detalhará como isso foi realizado. Como referência rápida, estes foram os parâmetros e as configurações otimizadas pela versão multiespécie do ES implementado:

- σ_{gauss} : O valor do σ da Gaussiana empregado para retirar sua derivada e compor o FIR que é realiza a sua convolução com a envoltória de corrente;
- δ_{min} : O valor mínimo que a resposta do FIR deve ter, em módulo, para as

¹O símbolo “*” é utilizado com o mesmo significado do metacaractere para os sistemas operacionais *Linux* na computação, ou seja, podem significar qualquer outro caractere — o análogo à carta coringa nos baralhos de cartas.

respectivas amostras constituam de uma região sensibilizada;

- ΔI_{min} : Mínimo valor para a característica ΔI na qual o candidato a evento não será eliminado como ruído
- n_{min} (sujeito à utilização na configuração otimizada da remoção por eventos próximos): Quantidade mínima de amostras que um candidato deve estar distante de outro para não ser removido por evento próximo. A estratégia para a remoção depende da configuração escolhida, conforme próximo item;
- Configurações otimizadas (em competição intraespécie), todas com remoção devido a eventos inconsistentes:

Espécie (I) - Com remoção de eventos próximos sem deslocamento, essa *depois* de remover eventos ruidosos;

Espécie (II) - Com remoção de eventos próximos utilizando a média, essa *depois* de remover eventos ruidosos;

Espécie (III) - Com remoção de eventos próximos sem deslocamento, essa *antes* de remover eventos ruidosos;

Espécie (IV) - Com remoção de eventos próximos utilizando a média, essa *antes* de remover eventos ruidosos;

Espécie (V) - Sem remoção de eventos próximos (não otimiza n_{min}).

6.1 Descrição da base de dados

O CEPEL/Eletrbras propôs uma evolução gradual para a análise de sua técnica. As configurações mais simples incluem a coleta de dados com equipamentos sendo acionados e desacionados sequencialmente operando individualmente em apenas um único estado. Em seguida, simulou-se condições de operação em conjunto de equipamentos dois a dois, e equipamentos C2 operando em seus diversos estados. Essas condições mais simples e com dados bastante limpos constituíram a base de dados dos arquivos NI00*** que serão detalhados adiante nesta seção. Para todos os conjuntos utilizados neste trabalho, foi construído o gabarito com a infraestrutura implementada descrita no Capítulo 5 e de posse de registros das alterações nos estados operativos dos equipamentos fornecidos pelo CEPEL/Eletrbras.

Já os conjuntos de dados *Temporizado*, *Empilhado4* e *Empilhado7* constituem de simulações em laboratório mais complexas, contendo a operação de diversos aparelhos simultaneamente que têm sua detecção dificultada pela presença da dinâmica de aparelhos C5. O conjunto *Temporizado* foi criado pelo CEPEL/Eletrbras mais recentemente com o intuito de explorar os resultados de sua técnica em condições

ruidosas, enquanto os conjuntos *Empilhado4* e *Empilhado7* foram elaborados durante a época do estudo de Cañón, que não teve a oportunidade de analisar mais a fundo tais conjuntos de dados. Os registros das alterações nos estados operativos dos equipamentos nesses conjuntos de dados foram fornecidos pelo mesmo, o que possibilitou a recuperação desses dados com a informação necessária para a construção do gabarito.

6.1.1 Conjunto de dados *NI00168*

Este conjunto de dados contém apenas a operação de equipamentos individualmente. Há a presença de dois equipamentos desconhecidos, que não foram possíveis de ter sua informação recuperada ao correlacionar com os registros de alterações. Como não era prioridade neste trabalho ter essa informação — esses equipamentos não-rotulados não apresentam características especiais nos eventos de transitório decorrentes de seus acionamentos e desacionamentos, e essa informação é mais importante para a etapa de discriminação —, trabalhou-se com os dados sem a rotulação desses equipamentos.

O consumo agregado neste conjunto de dados pode ser observado na Figura 6.1, enquanto a Figura 6.2 contém a estimativa do consumo desagregado no gabarito.

Os equipamentos presentes neste conjunto de dados são, todos com dois eventos de transitório, exceto quando especificado:

- Lâmpada incandescente 100 W, 60 W (com *dimmer*, o que constitui uma C4);
- Dois equipamentos desconhecidos;
- Lâmpada fluorescente 15 W, 16 W, 18 W, 20 W, 21 W (circular), 40 W (tubular), 32 W (tubular), 28 W (tubular);
- Secador de cabelo (4 eventos: Desligado → Fraco → Forte → Fraco → Desligado);
- Televisão (3 eventos: Desligado → Ligado → Ligado e exibindo imagem → Desligado)²;
- Geladeira;
- Ar condicionado (4 eventos: Desligado → Ventilação → Ventilação e compressor → Ventilação → Desligado, constitui uma C5).

Todos os arquivos NI00*** foram coletados utilizando o medidor do CEP/PEL/Eletronbras.

²A televisão pode ser ligada em um canal que não apresenta imagem.

NI_00168



Figura 6.1: Perfil de consumo agregado para o conjunto de dados *NI00168*.

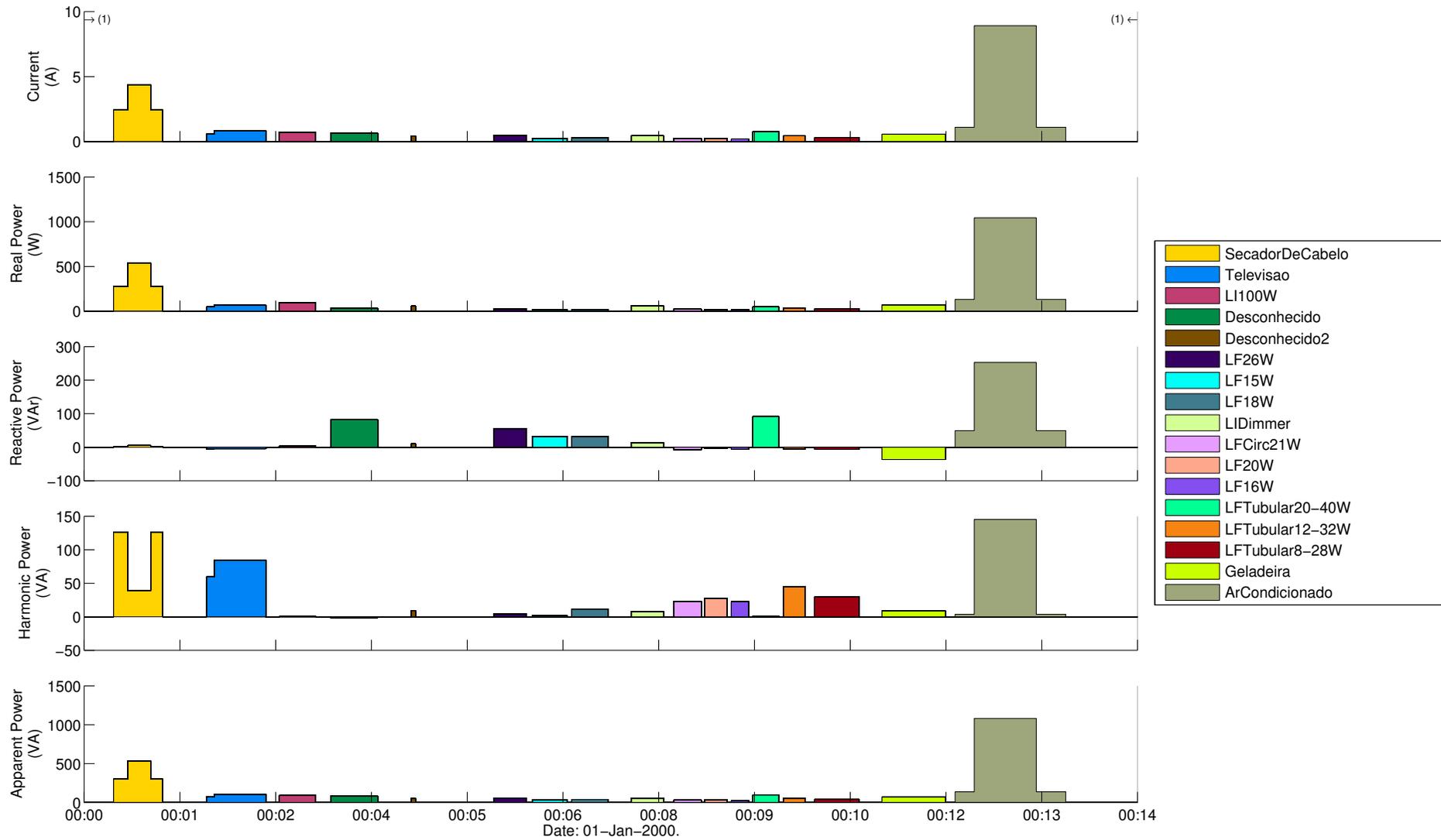


Figura 6.2: Informação no gabarito para o conjunto de dados *NI00168* - consumo temporal dos equipamentos.

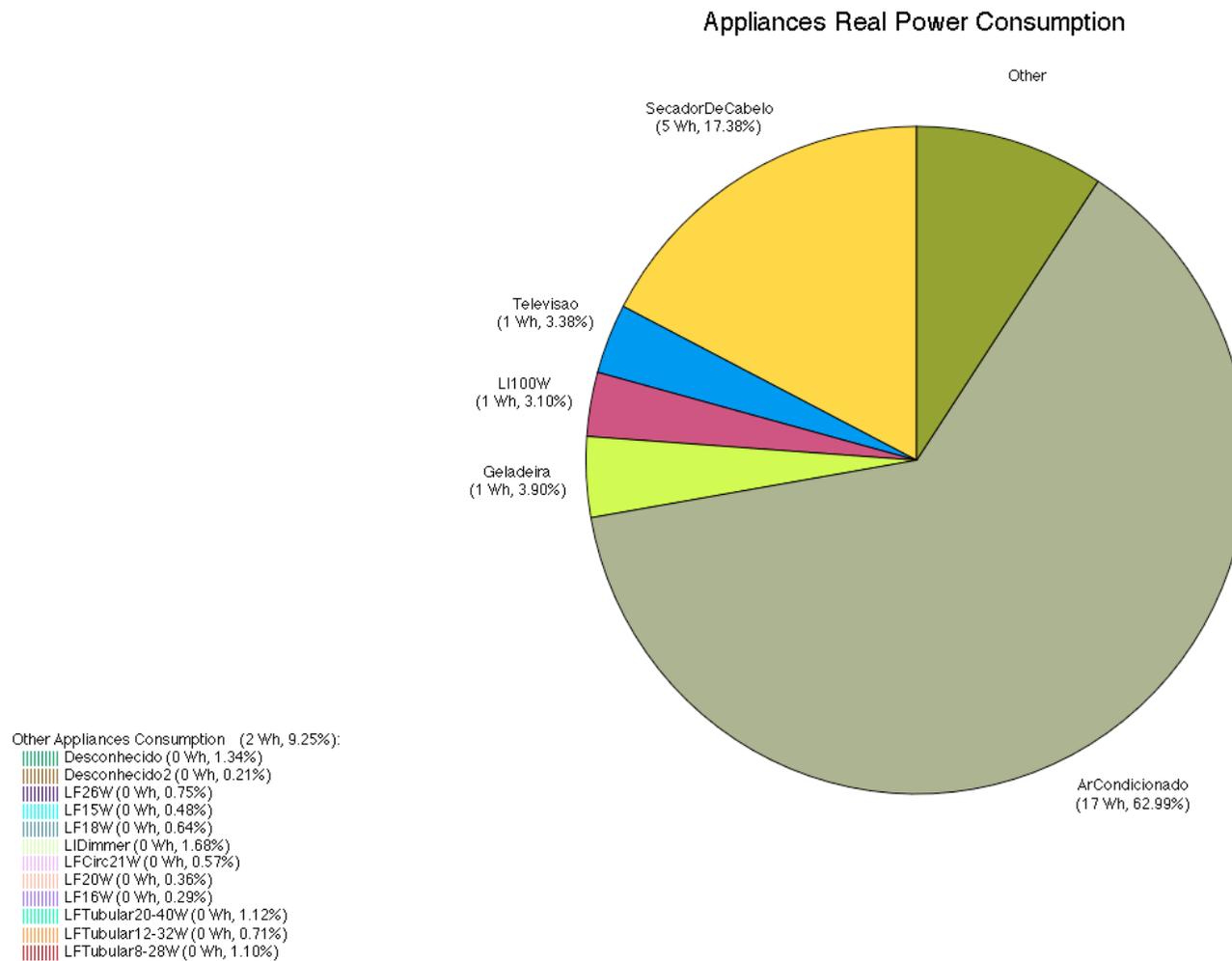


Figura 6.3: Informação no gabarito para o conjunto de dados *NI00168* - gráfico circular do consumo dos equipamentos.

NI_00171

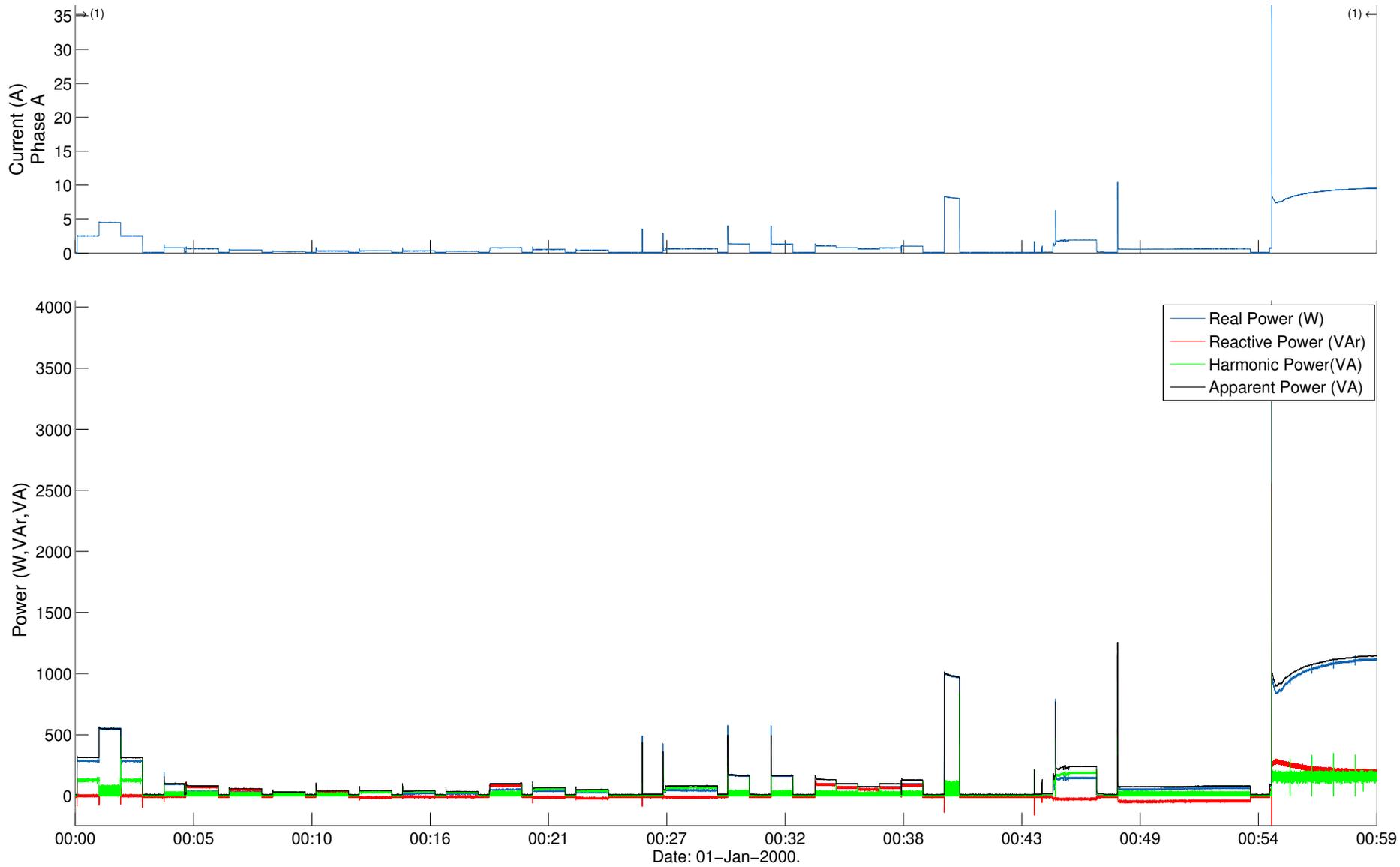


Figura 6.4: Perfil de consumo agregado para o conjunto de dados *NI00171*.

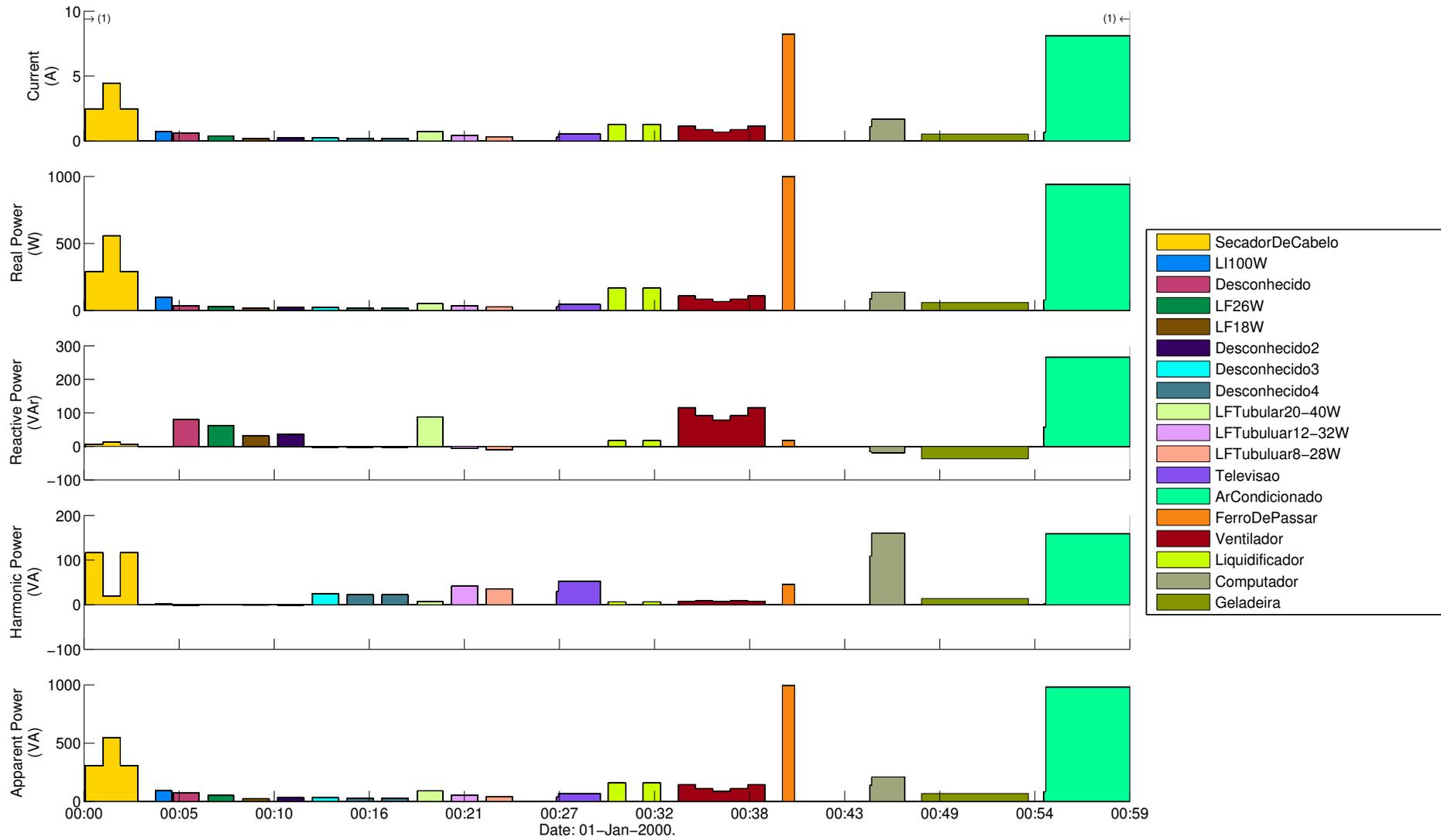


Figura 6.5: Informação no gabarito para o conjunto de dados *NI00171* - consumo temporal dos equipamentos.

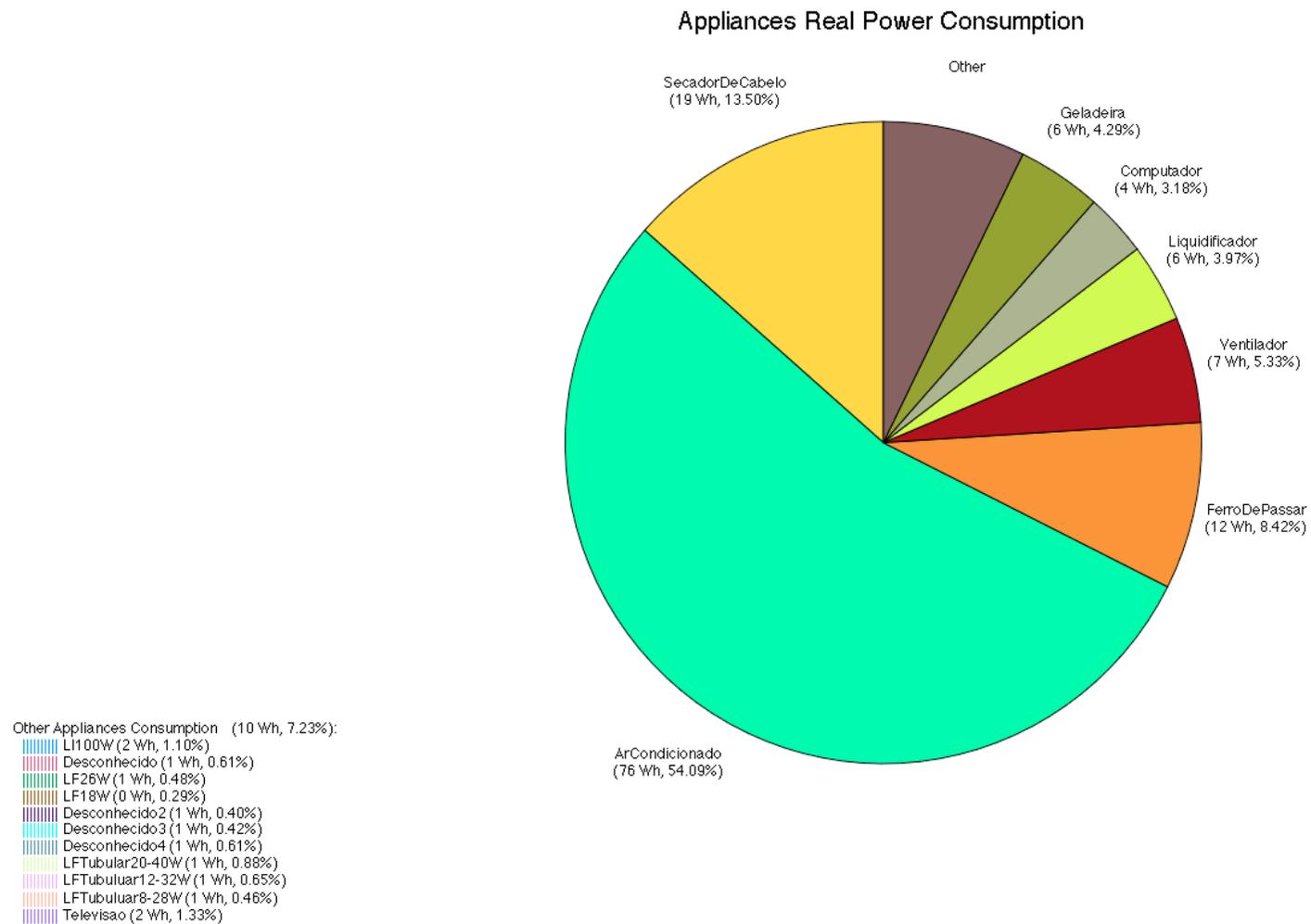


Figura 6.6: Informação no gabarito para o conjunto de dados *NI00171* - gráfico circular do consumo dos equipamentos.

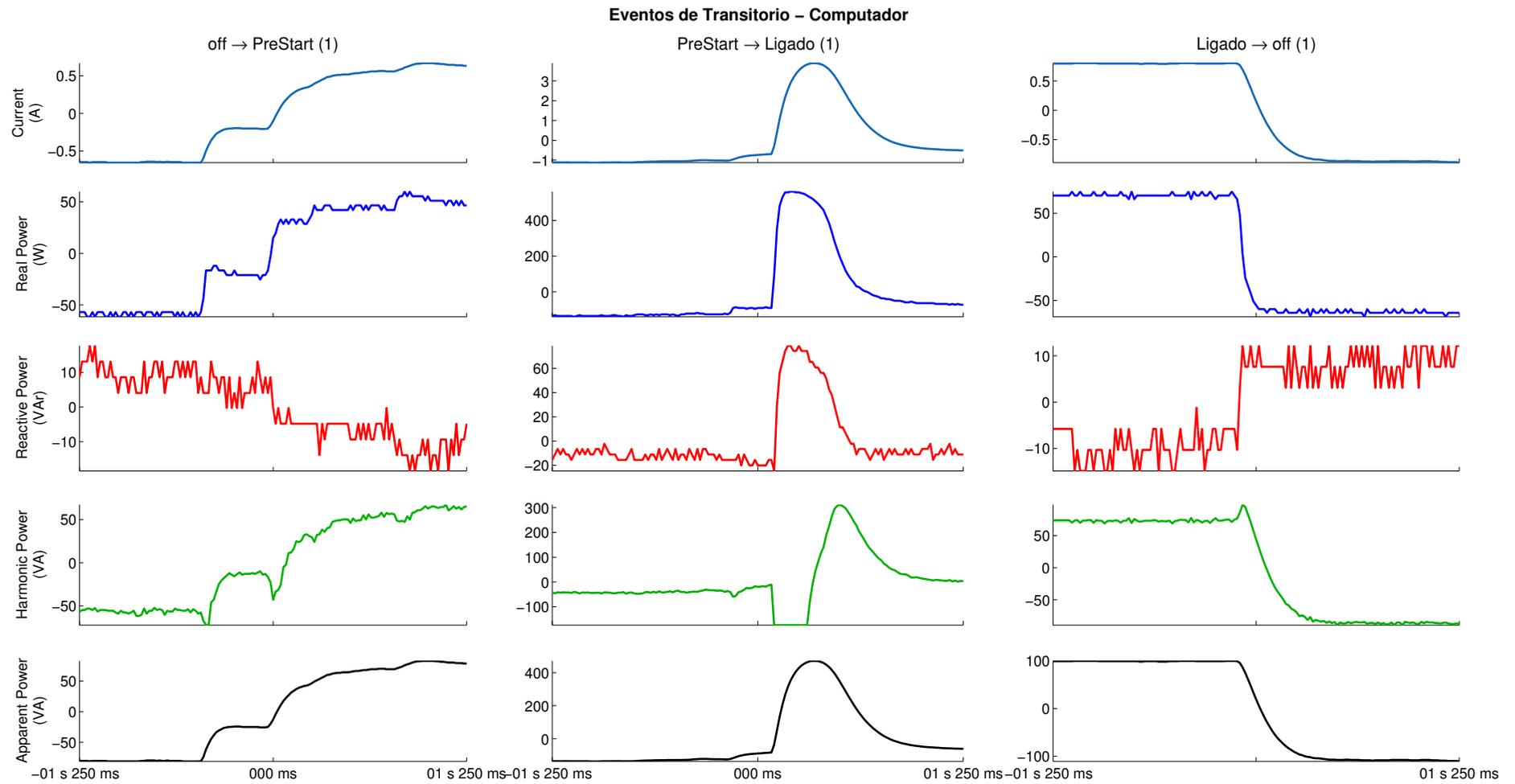


Figura 6.7: Informação no gabarito para o conjunto de dados *NI00171* - envoltória para as diversas variáveis para o computador portátil.

6.1.3 Conjunto de dados *NI00173*, *NI00174*, *NI00175* e *NI00177*

Estes conjuntos de dados apresentam acionamentos de equipamentos simultaneamente, porém isso é realizado de maneira simples, apenas com a operação de dois equipamentos atuando ao mesmo tempo. Em ambos os casos é utilizado uma lâmpada incandescente (100 W) em conjunto com um outro aparelho para realizar a operação conjunta de equipamentos. No caso do conjunto de dados *NI00173*, o equipamento operando em conjunto com a lâmpada incandescente é um ventilador acionado em um único estado (pode ser modelado como uma C3, apesar de ser uma C2), enquanto nos demais conjuntos o equipamento é um secador de cabelo que opera em dois estados (Forte e Fraco, e, portando, atua como uma C2). A única diferença entre os arquivos *NI00174*, *NI00175* e *NI00177* é a ordem na qual o secador de cabelo e a lâmpada têm seus estados de operação alterados. Para o conjunto *NI00174*, a lâmpada permanece ligada enquanto o estado de operação do secador é alterado. Os conjuntos *NI00175* e *NI00177* alteram a operação dos estados desses equipamentos de diversas modos diferentes, tentando simular diversas configurações em que esses aparelhos poderiam ter seus estados alterados em uma residência. As Figuras 6.8-6.15 permitem observar esses comportamentos, contendo o consumo agregado nesses conjuntos de dados e a informação estimada do consumo desagregado em seus respectivos gabaritos.

Todos os arquivos NI00*** foram coletados utilizando o medidor do CEP/EL/Eletrobras.

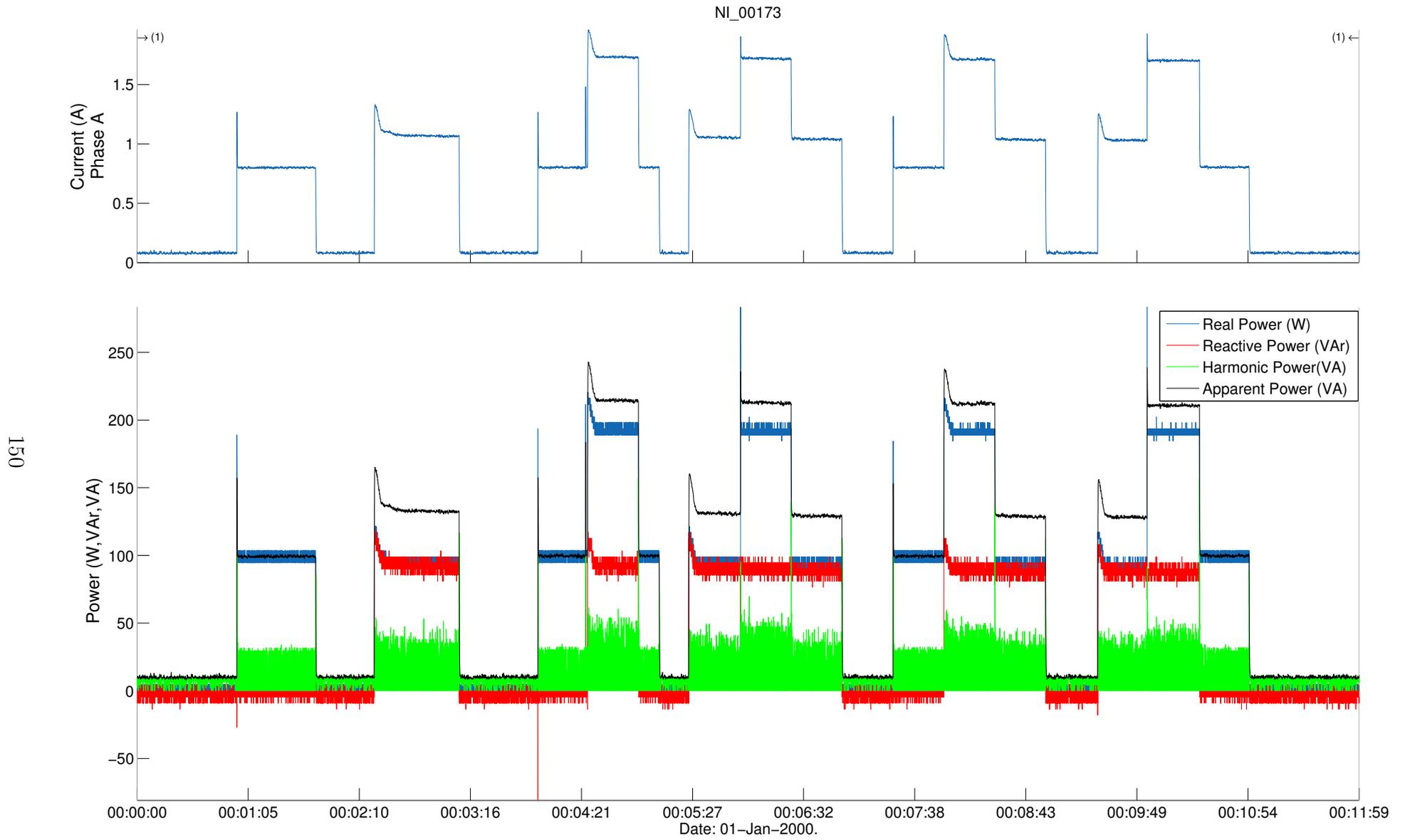


Figura 6.8: Perfil de consumo agregado para o conjunto de dados *NI00173*.

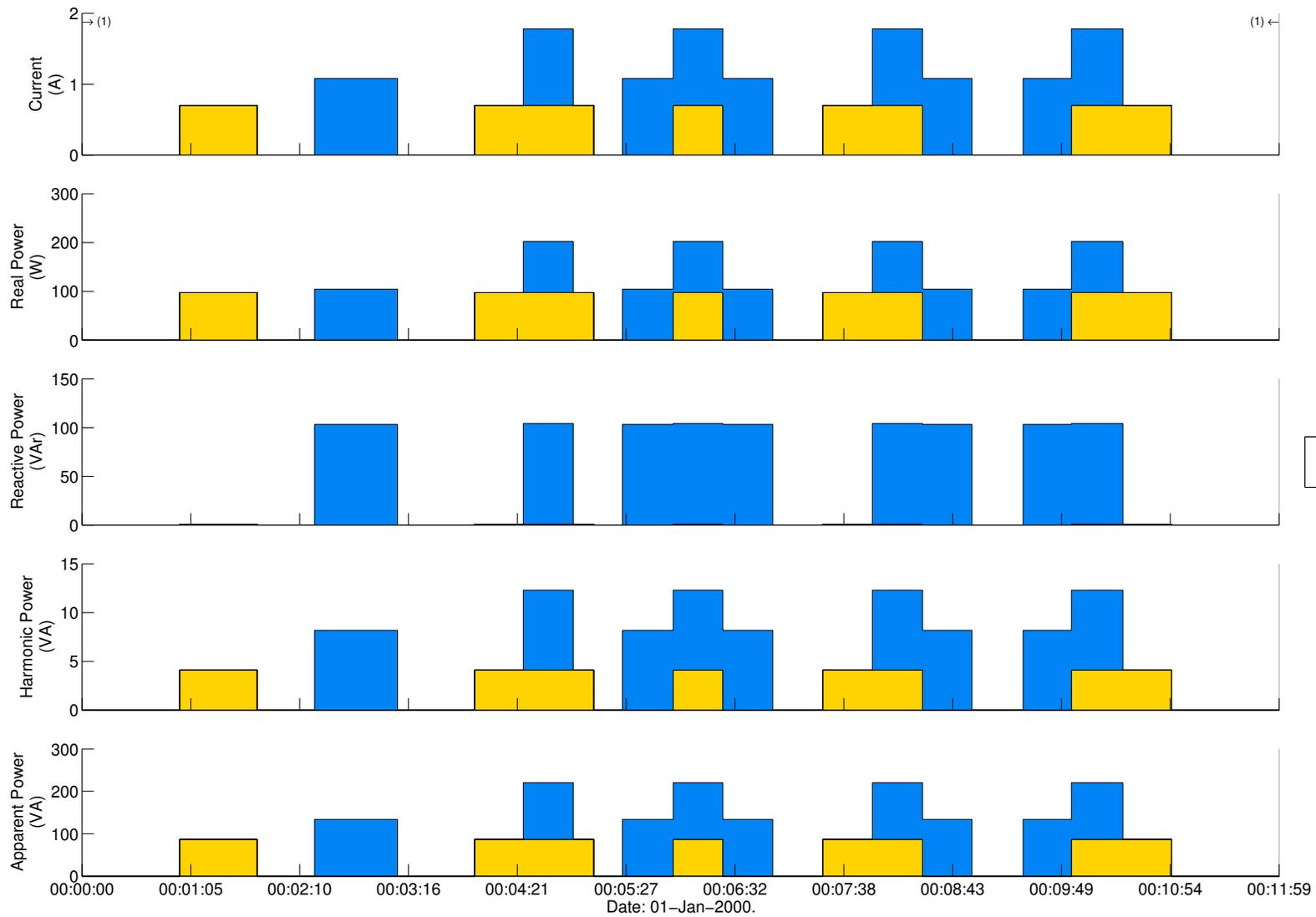


Figura 6.9: Informação no gabarito para o conjunto de dados *NI00173* - consumo temporal dos equipamentos.

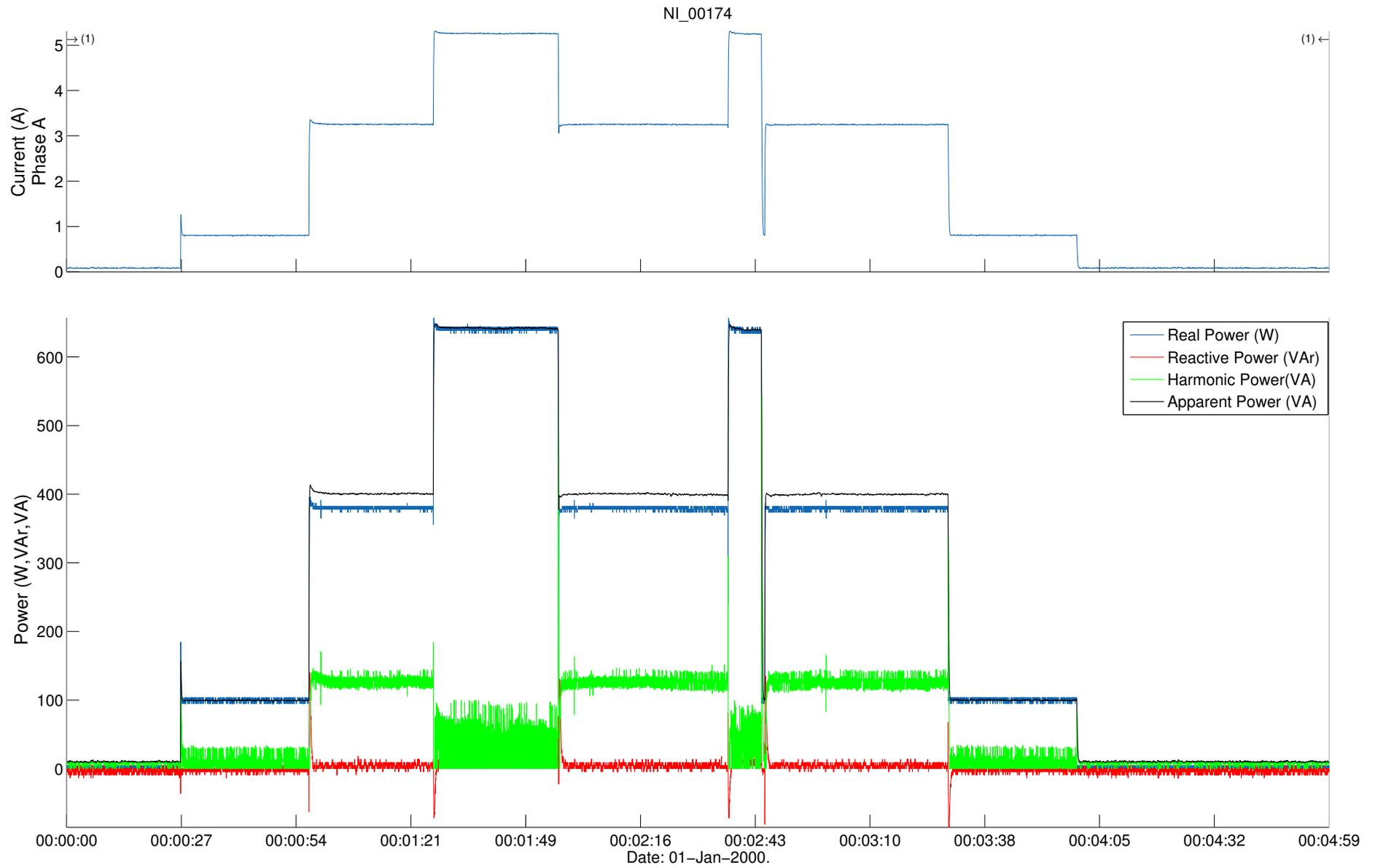


Figura 6.10: Perfil de consumo agregado para o conjunto de dados *NI00174*.

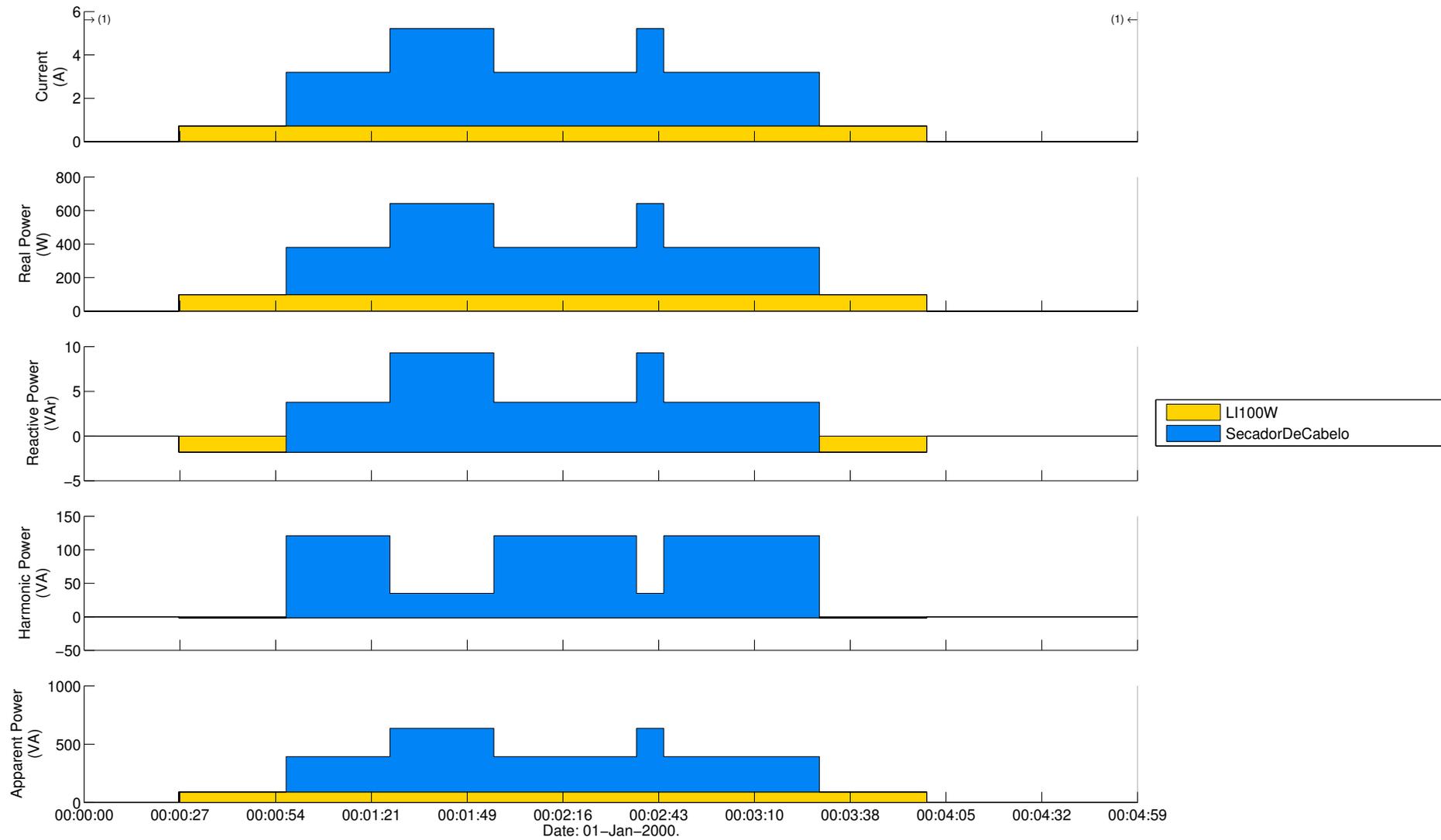


Figura 6.11: Informação no gabarito para o conjunto de dados *NI00174* - consumo temporal dos equipamentos.

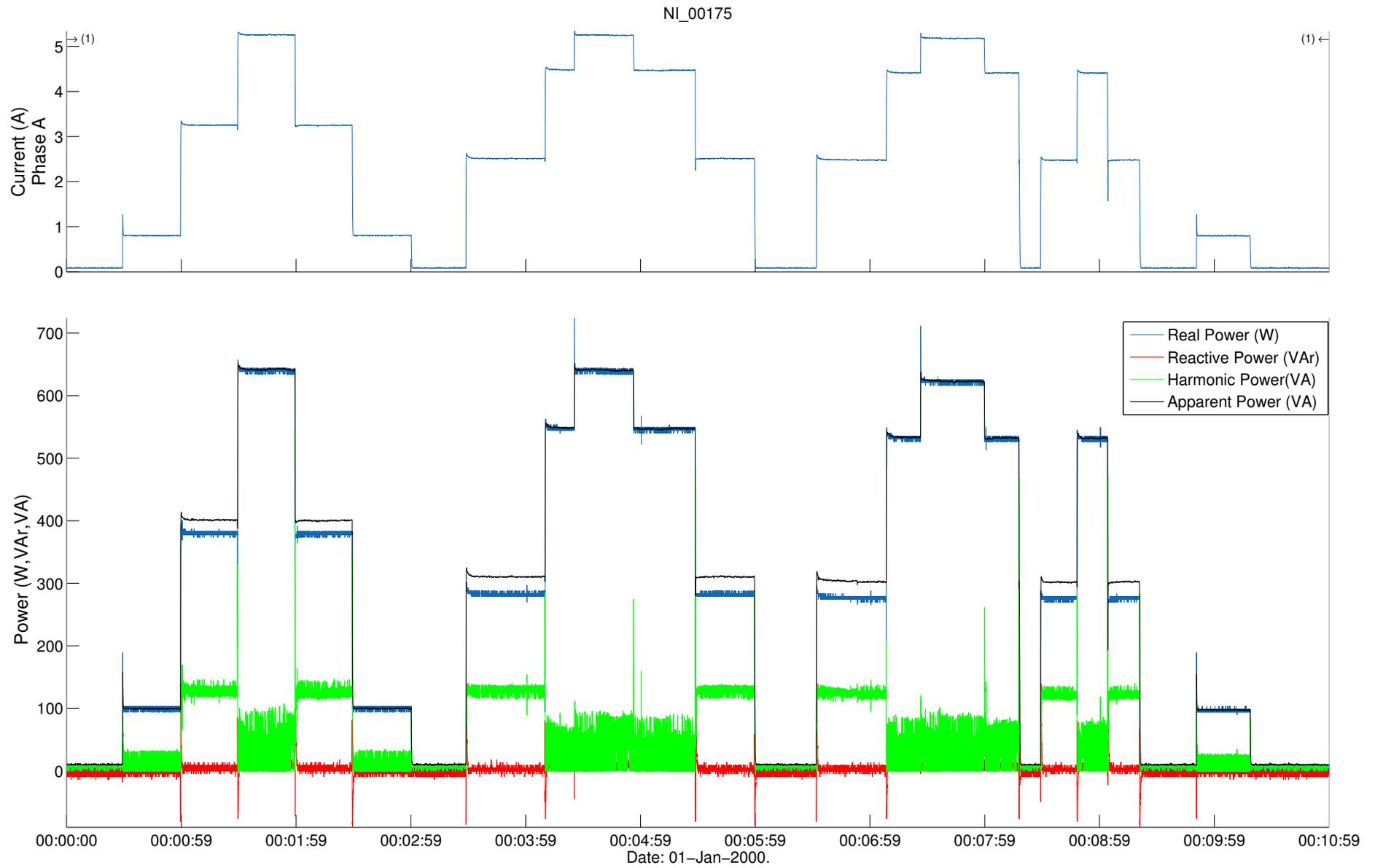


Figura 6.12: Perfil de consumo agregado para o conjunto de dados *NI00175*.

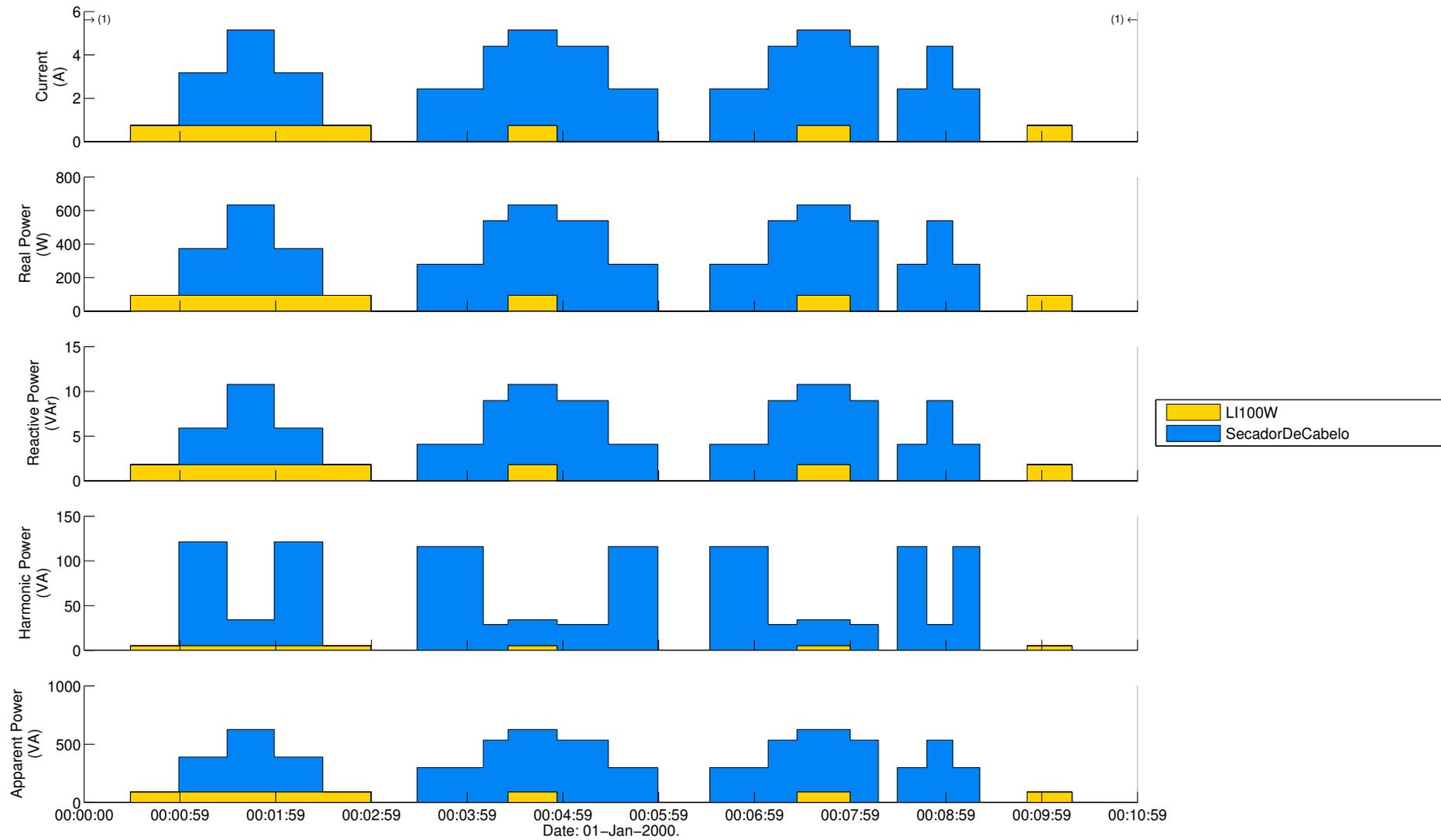


Figura 6.13: Informação no gabarito para o conjunto de dados *NI00175* - consumo temporal dos equipamentos.

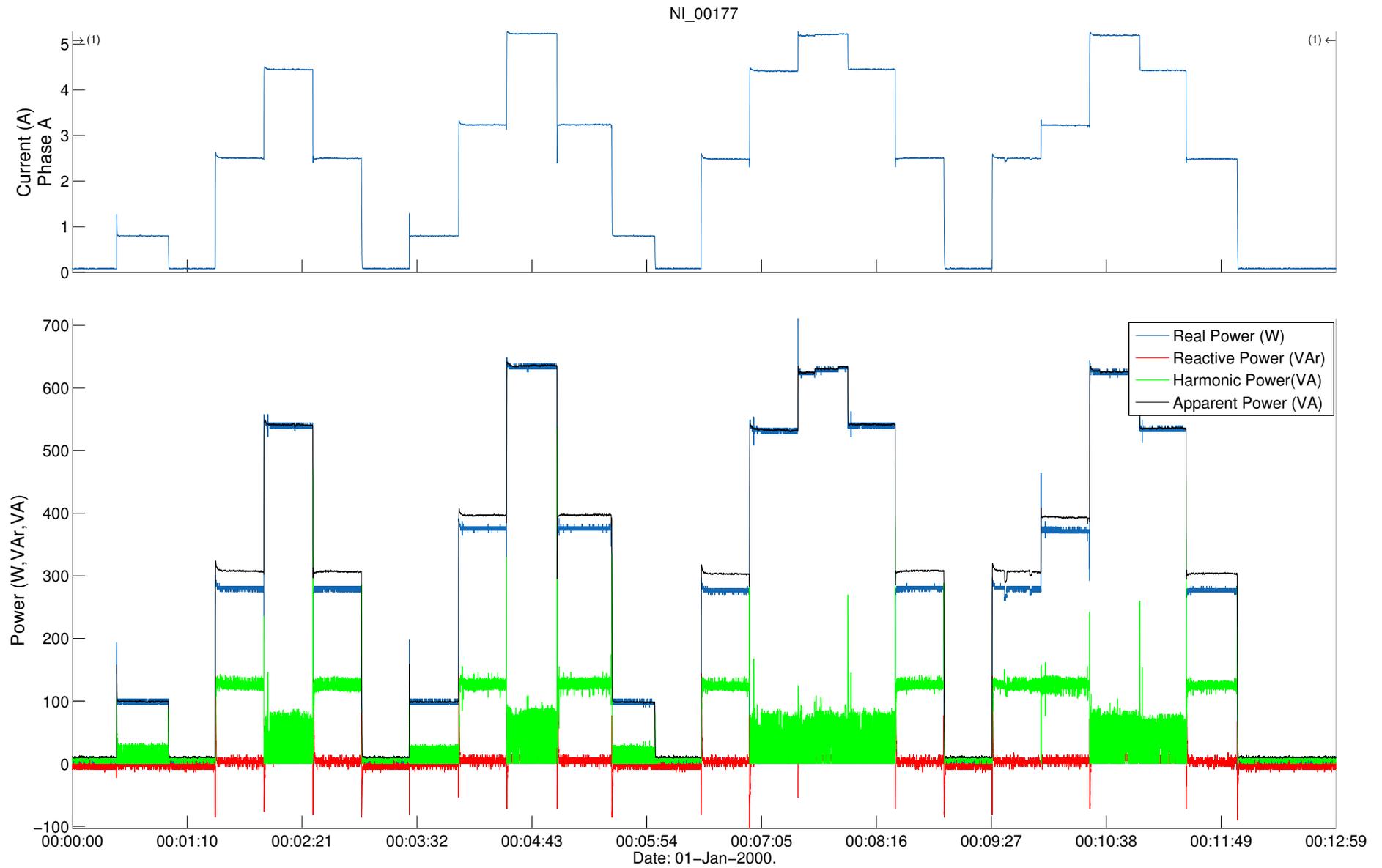


Figura 6.14: Perfil de consumo agregado para o conjunto de dados *NI00177*.

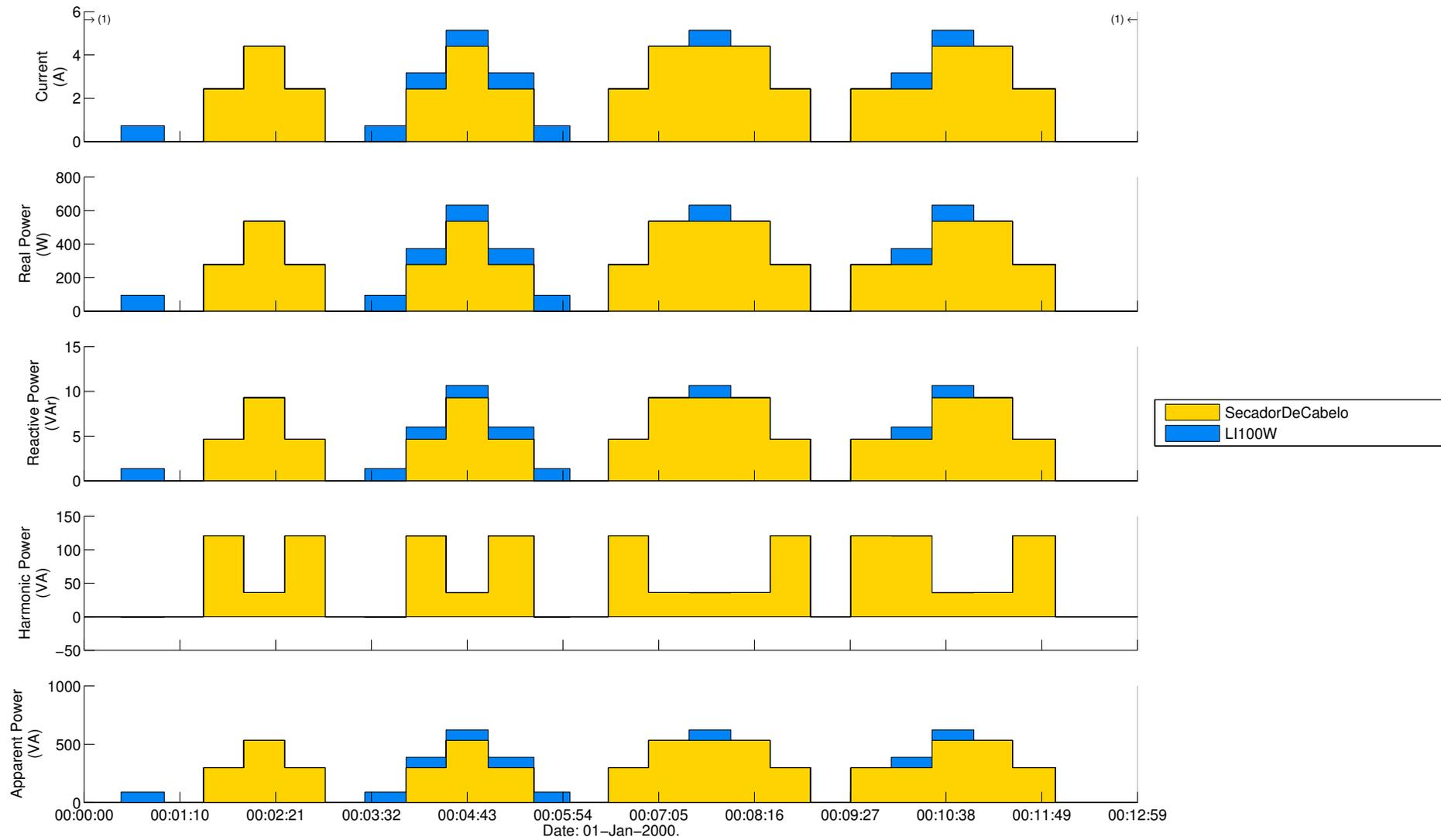


Figura 6.15: Informação no gabarito para o conjunto de dados *NI00177* - consumo temporal dos equipamentos.

6.1.4 Conjunto de dados *Temporizado*

Este conjunto possui a operação de cinco aparelhos que podem ser modelados como C3 por atuarem apenas em um único estado quando operando. Uma característica deste conjunto é a presença de um alto nível de ruído durante a operação de uma C5, a televisão LCD. Outra peculiaridade deste conjunto é a sua longa duração, cerca de 18 horas foram coletadas. Além disso, ele possui dois gabaritos, motivo detalhado um pouco mais a diante. A atuação dos equipamentos foi feito através de chaveamento automático, onde tentou-se gerar eventos próximos de acionamentos e desacionamentos dos equipamentos. Um total de 149 eventos de transitório estão presentes no conjunto, causados por estes equipamentos (sempre na sequência Desligado → Ligado → Desligado):

- Televisão LCD (6 eventos de transitório), constitui-se de uma C5;
- Geladeira (123 eventos, o número ímpar de eventos é causado pelo fato da geladeira já estar operando quando a medição do conjunto é iniciada);
- Lâmpada fluorescente 23 W (4 eventos), 54 W (4 eventos);
- Ventilador (12 eventos).

O consumo agregado pode ser observado na Figura 6.17. A informação contida no gabarito para este conjunto de dados pode ser observada nas figuras 6.17–6.23. A Figura 6.17 contém a informação do consumo temporal dos equipamentos, enquanto a Figura 6.18 contém o gráfico circular do consumo estimado no gabarito para os equipamentos. As figuras 6.19–6.23 contém os transitórios dos equipamentos marcados pelo usuário durante a criação do gabarito. Todos os eventos são movidos para obterem média zero, de forma que os eventos de transitório estejam centrados no mesmo patamar e seja possível compará-los. A informação contida nesse gráfico auxilia a identificar eventos no gabarito que fogem do padrão, seja por erro do usuário no preenchimento, ou por caracterizar um evento excêntrico, facilitando a identificação desses casos. Um exemplo pode ser observado na Figura 6.20, onde há a ocorrência de um evento, marcado por uma etiqueta (a mesma torna possível a identificação do evento através de sua chave), que foge do padrão dos outros coletados. Essas figuras também permitem observar a quantidade de eventos para cada alteração de estado (indicado entre parênteses no título das subfiguras). Sua medição foi realizada com o medidor *Yokogawa*.

Este conjunto é o que tem a maior presença de ruído, causado pela televisão LCD, em especial para os períodos das 00:00 às 02:00 e 04:30 às 08:00 do dia 21. Durante esses períodos, há a ocorrência de uma perturbação que se assemelha com a operação de um equipamento C3, como exibido na Figura 6.24. Durante a

aplicação da metodologia descrita na Seção 6.2 e ao observar os resultados presentes no Capítulo 7, observou-se uma alta presença de falsos alarmes nesse conjunto para todas os parâmetros examinados, o que levou a descoberta dessa perturbação. Fica então a questão de como tratá-la, se essa perturbação deve ser considerado como um estado interno de operação da televisão LCD, ou se esse distúrbio deve ser ignorado e tratado como falso alarme. Para permitir a análise de ambas situações, gerou-se o gabarito para este conjunto de dados com os dois casos, o primeiro gabarito contendo apenas os acionamentos e desacionamentos registrados pelo CEPEL/Eletrobras (149 eventos de transitório), e o segundo contém, também, os distúrbios observados na Figura 6.24 (211 eventos).

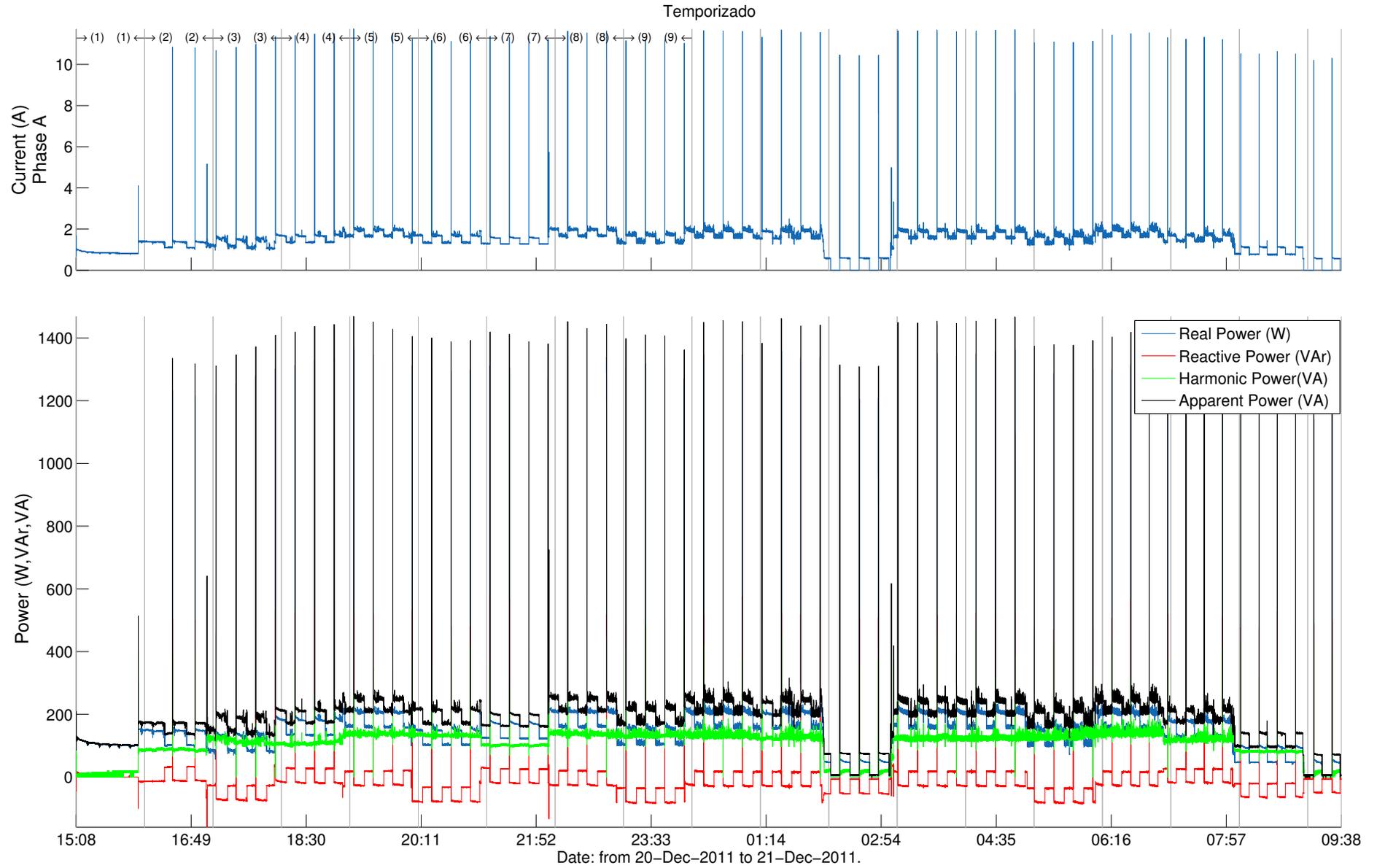


Figura 6.16: Perfil de consumo agregado para o conjunto de dados *Temporizado*.

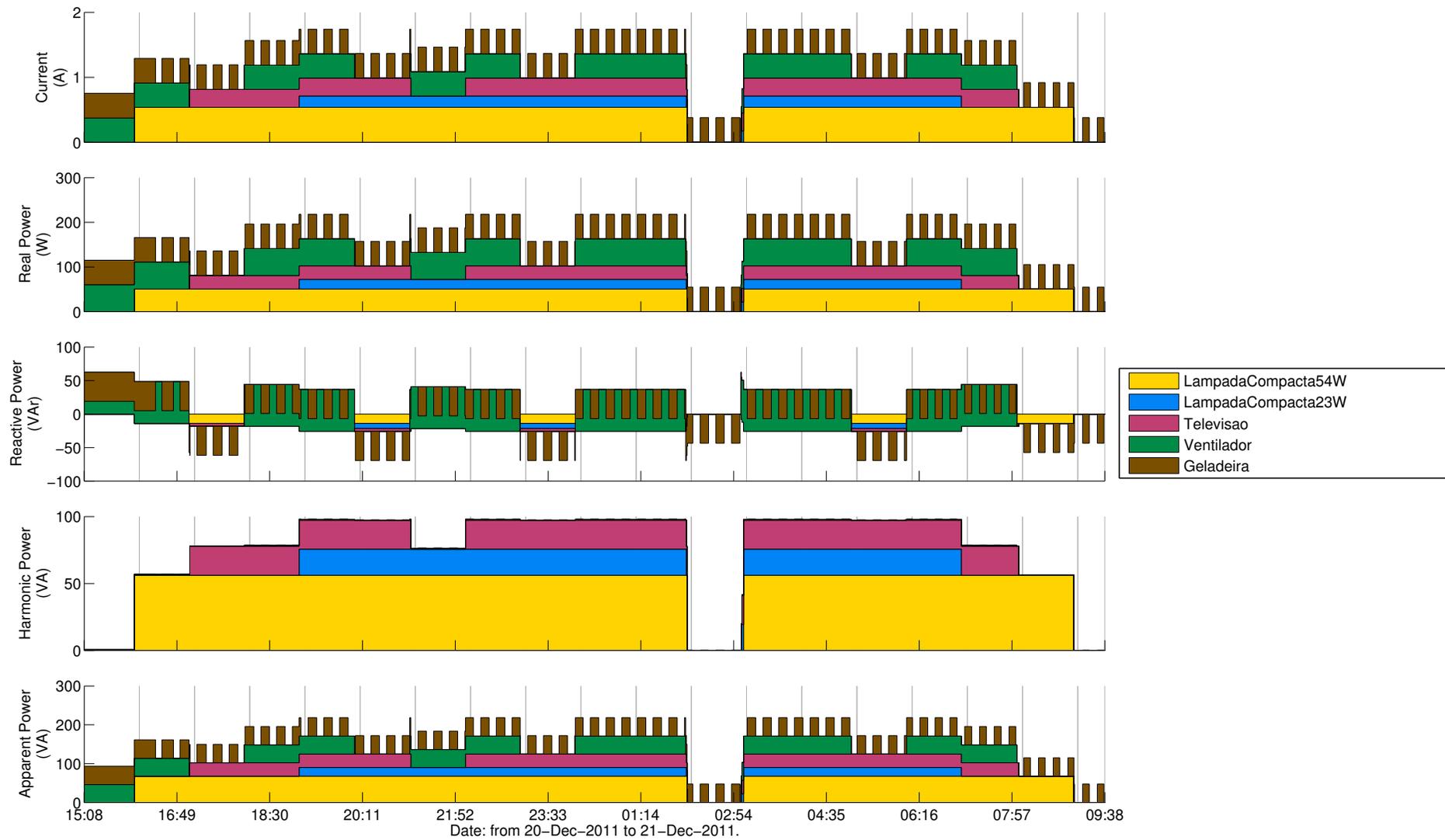


Figura 6.17: Informação no gabarito para o conjunto de dados *Temporizado* - consumo temporal dos equipamentos.

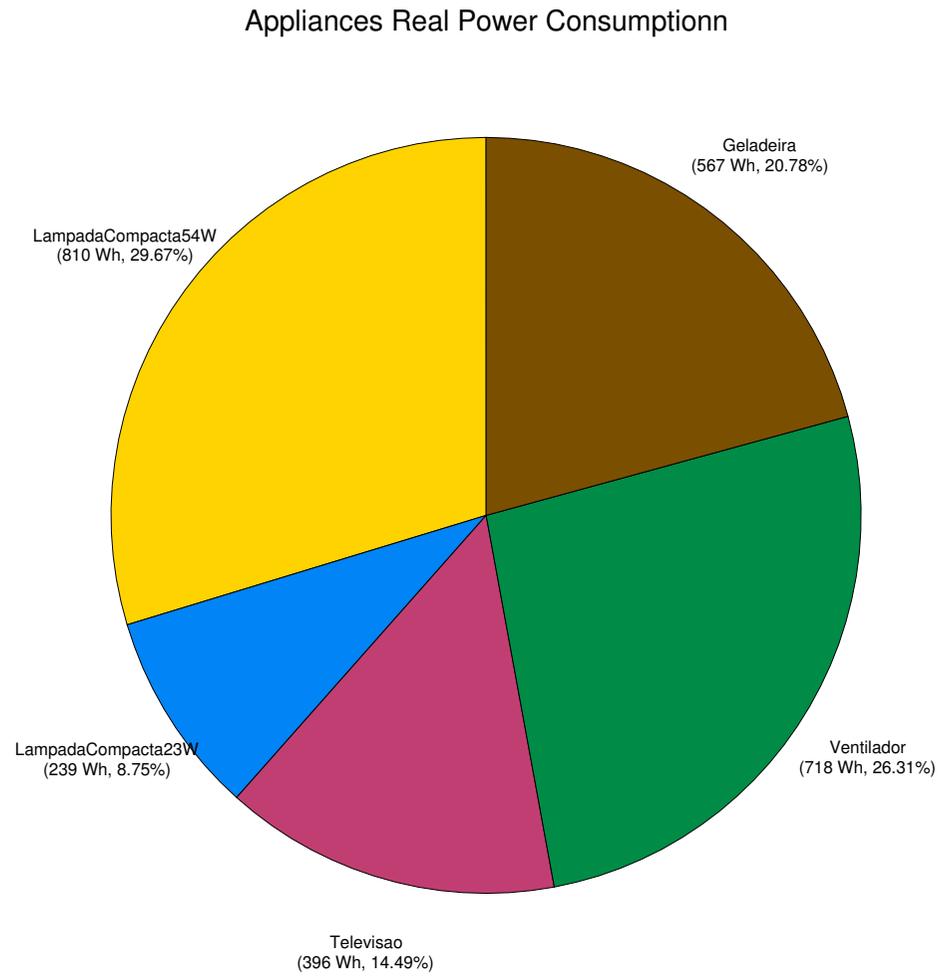


Figura 6.18: Informação no gabarito para o conjunto de dados *Temporizado* - gráfico circular do consumo dos equipamentos.

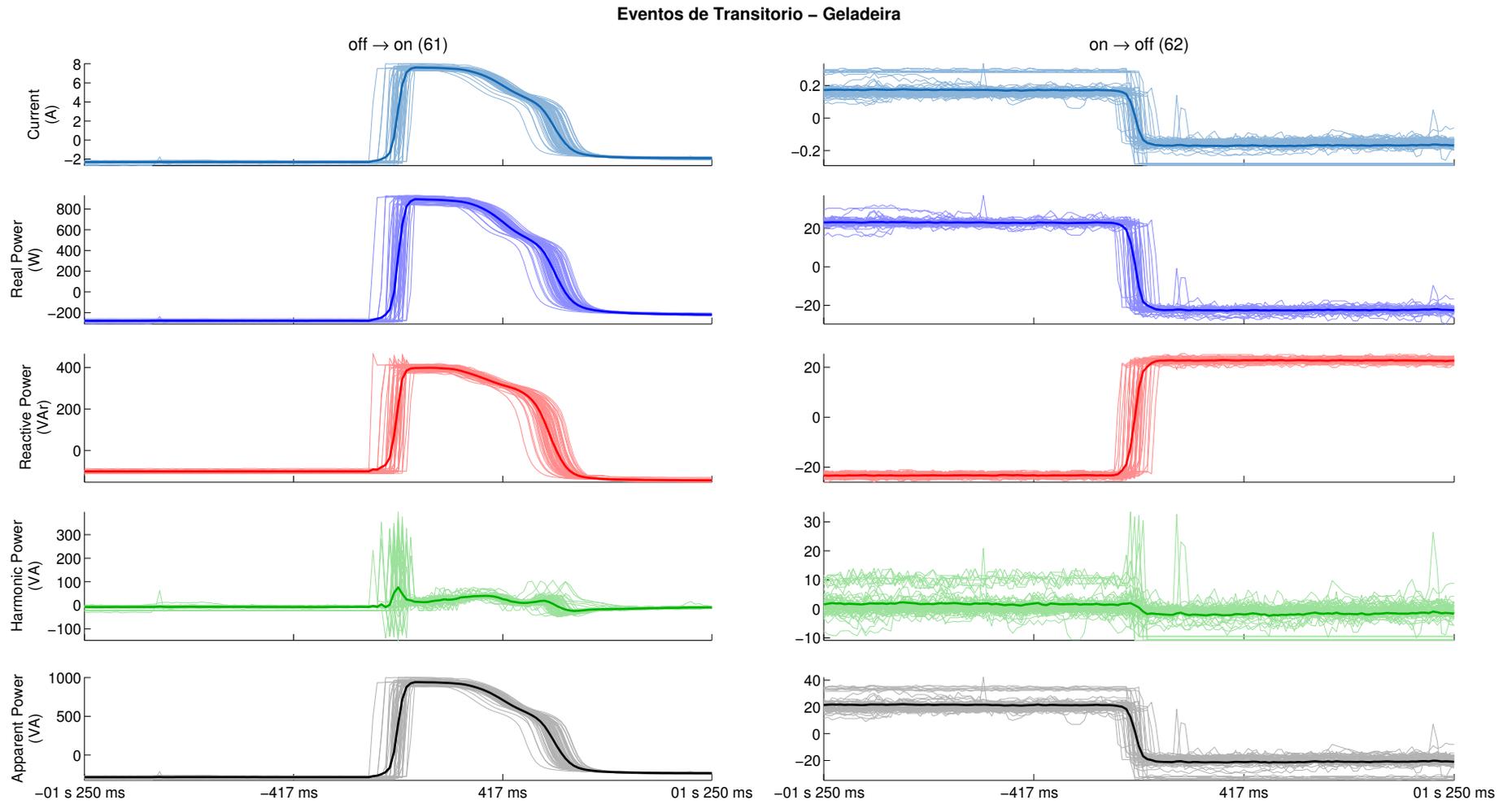


Figura 6.19: Informação no gabarito para o conjunto de dados *Temporizado* - envoltória para as diversas variáveis para a geladeira.

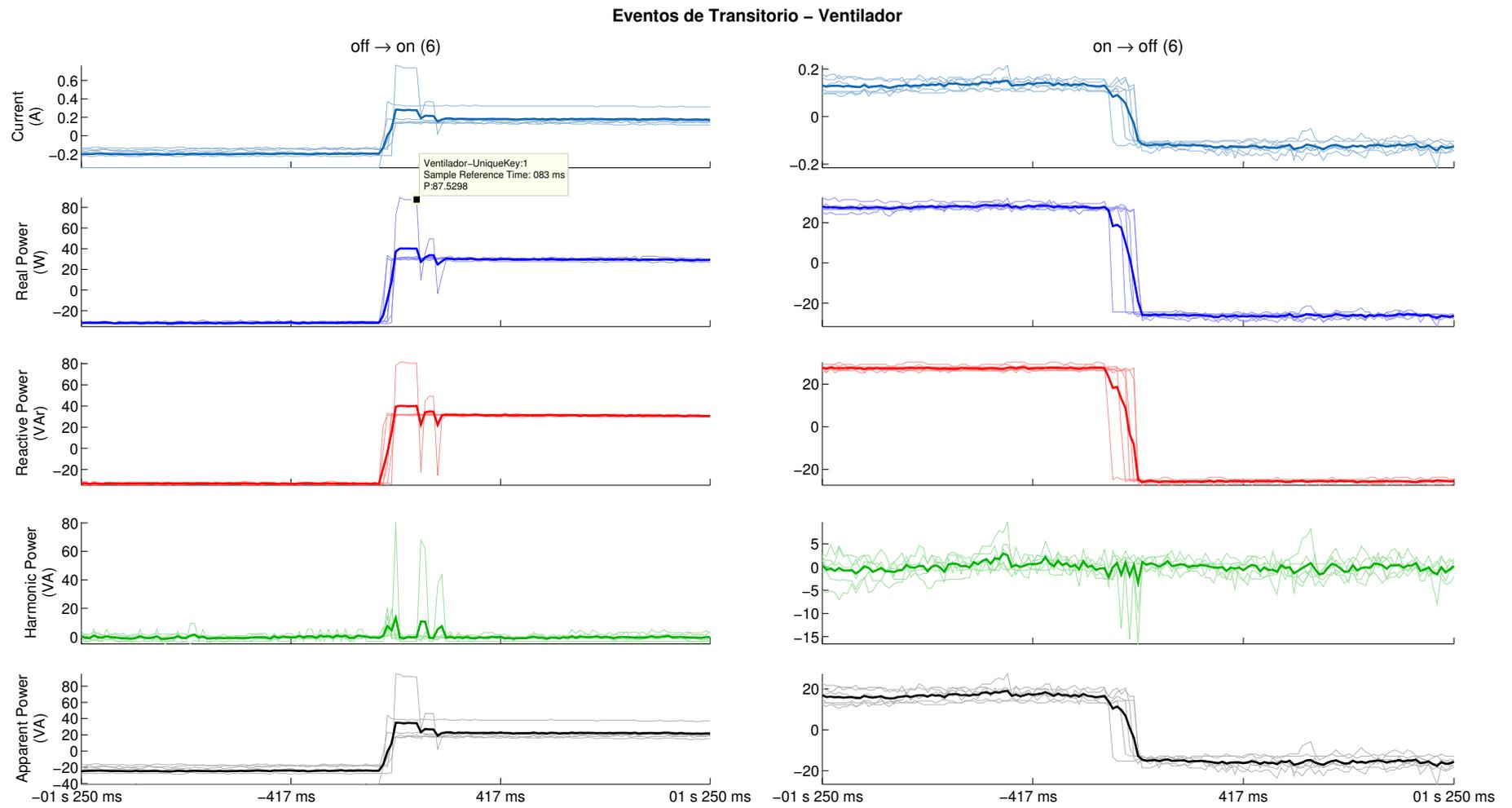


Figura 6.20: Informação no gabarito para o conjunto de dados *Temporizado* - envoltória para as diversas variáveis para a ventilador.

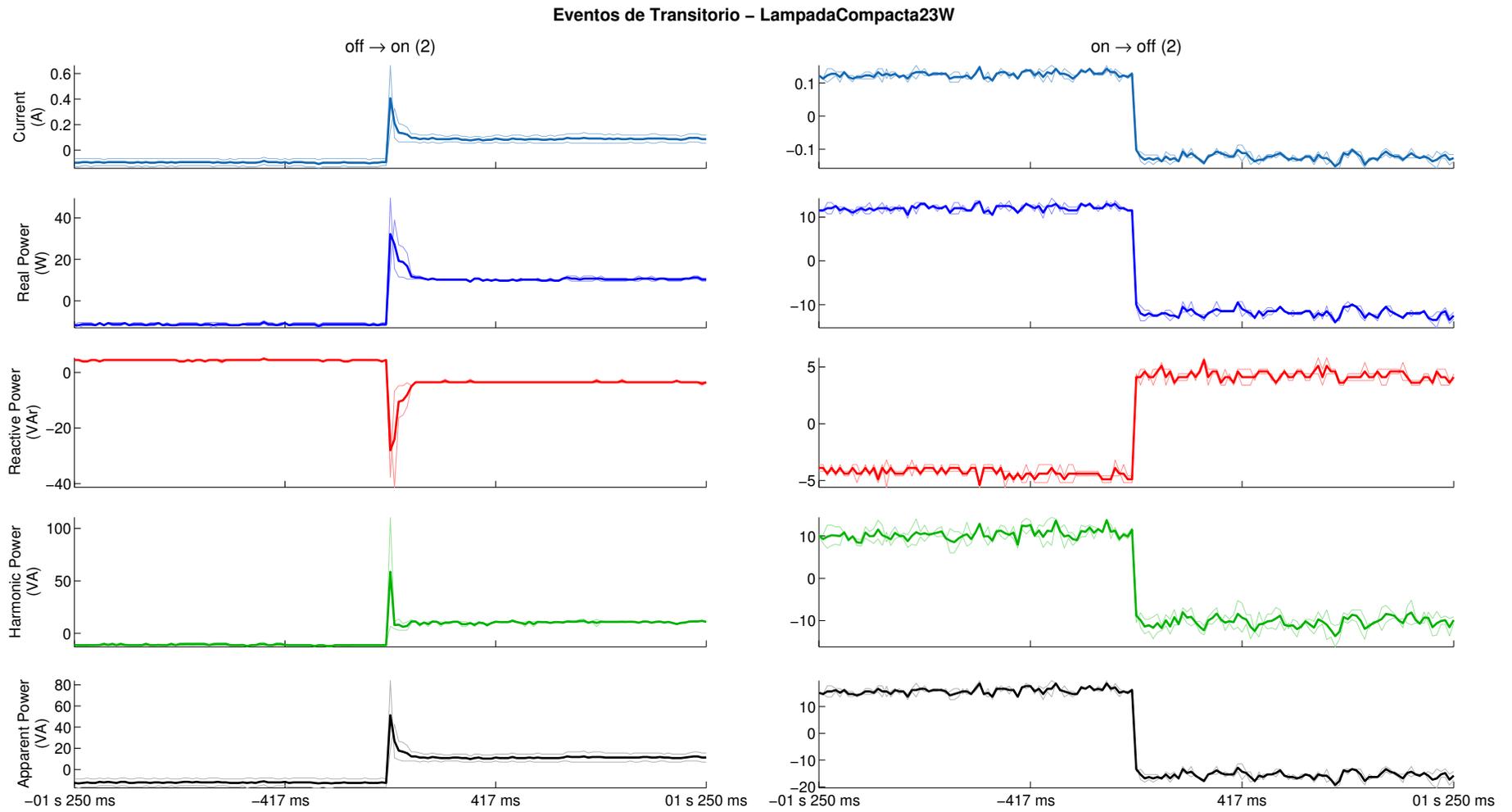


Figura 6.21: Informação no gabarito para o conjunto de dados *Temporizado* - envoltória para as diversas variáveis para a lâmpada fluorescente 23W.

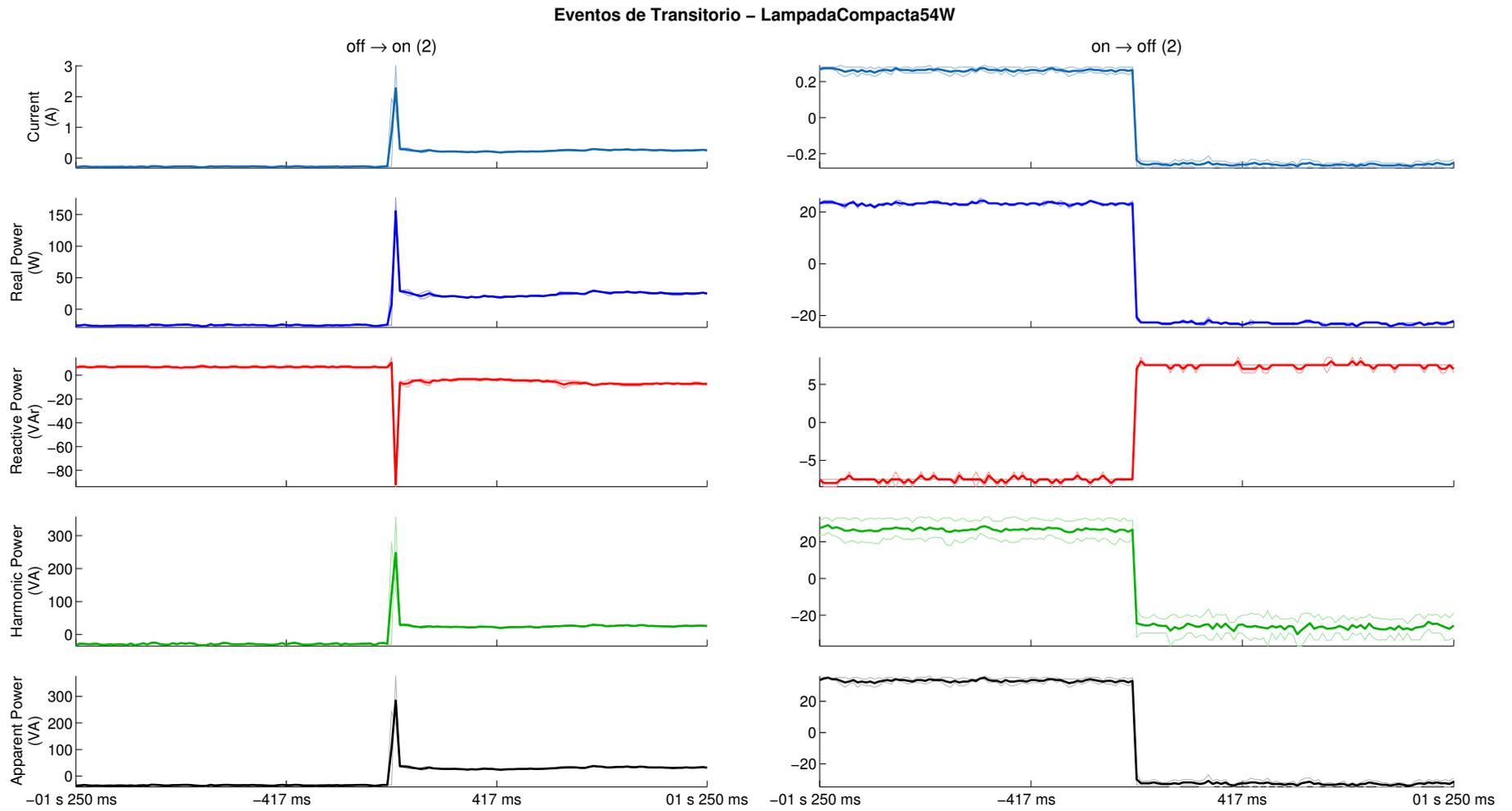


Figura 6.22: Informação no gabarito para o conjunto de dados *Temporizado* - envoltória para as diversas variáveis para a lâmpada fluorescente 54W.

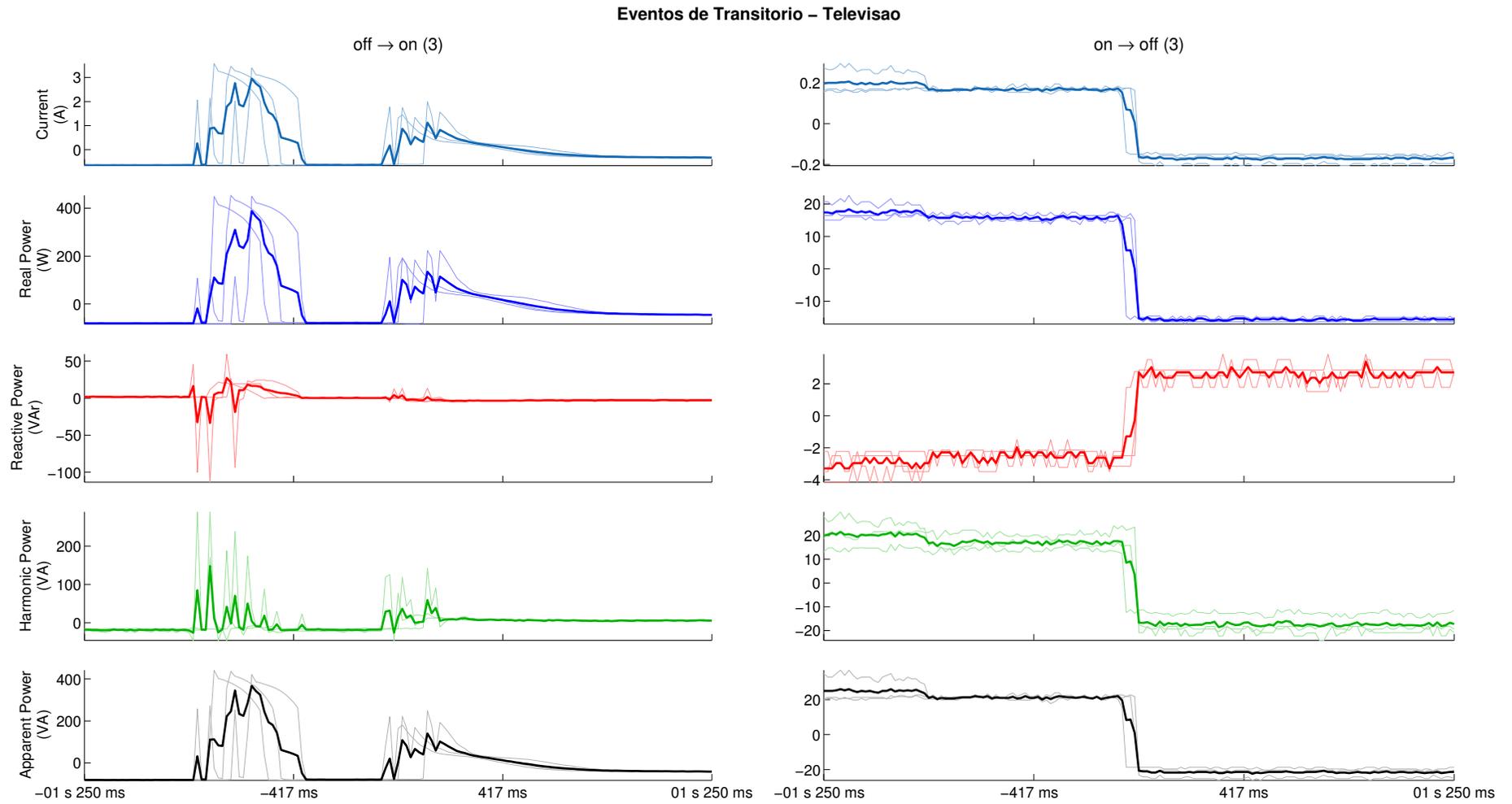


Figura 6.23: Informação no gabarito para o conjunto de dados *Temporizado* - envoltória para as diversas variáveis para a televisão.

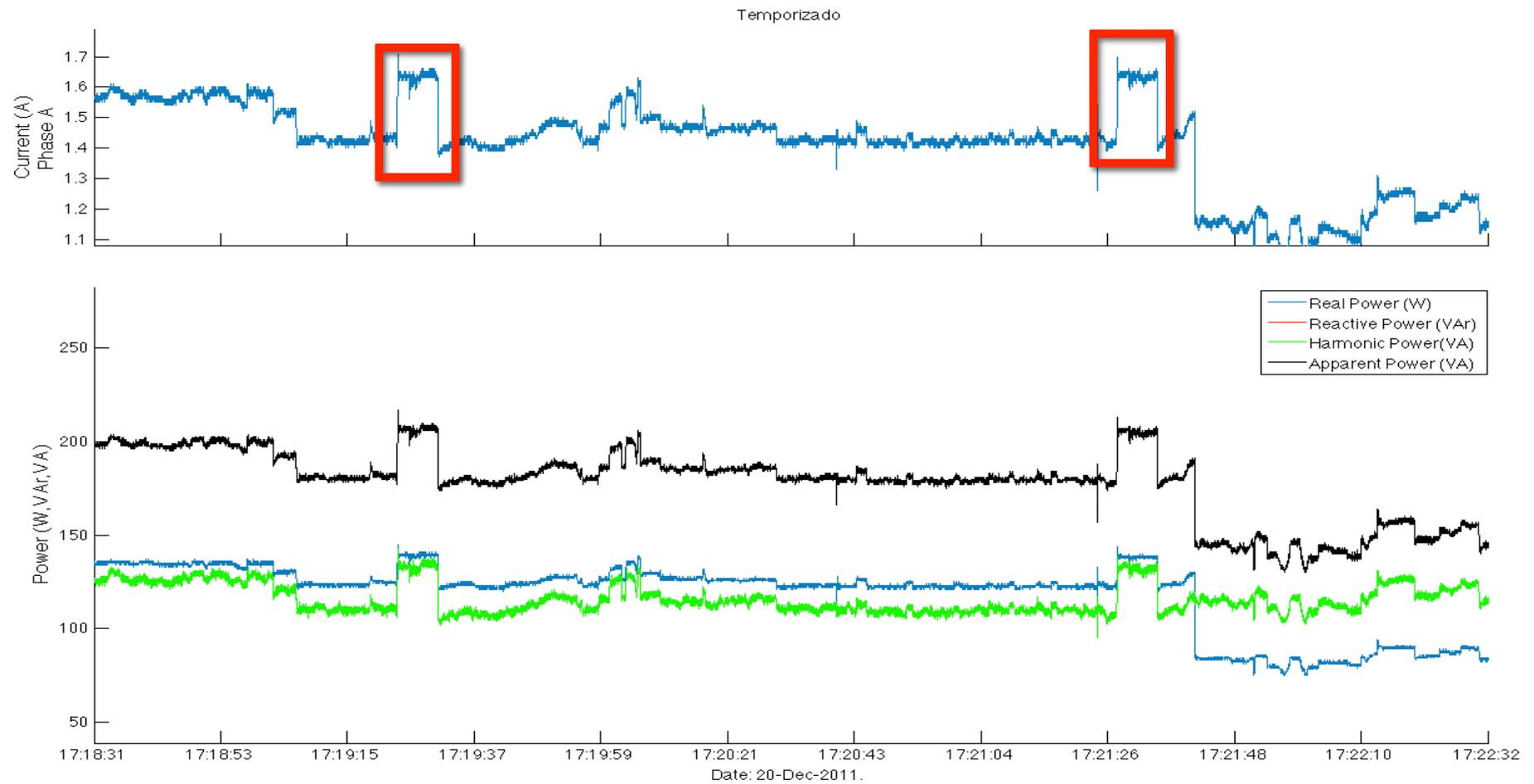


Figura 6.24: Distúrbio recorrente presente no conjunto de dados *Temporizado*. O distúrbio apresenta características semelhantes a um aparelho C3 com consumo de 15 W e opera durante cerca de 3 s.

6.1.5 Conjunto de dados *Empilhado4*

A principal característica deste conjunto é a operação simultânea de lâmpadas fluorescentes e incandescentes de diversos consumos em conjunto com outros três equipamentos: forno elétrico, chuveiro elétrico e televisão CRT. Esses aparelhos operam um a um em conjunto com as lâmpadas. Este conjunto também é o que contém a maior quantidade de eventos de transitório de baixo consumo causado pelas lâmpadas fluorescentes de baixo consumo. A seguir segue a descrição dos equipamentos presentes neste conjunto de dados:

- Forno elétrico: dinâmica de consumo após acionamento (queda lenta do consumo até estabilizar);
- Chuveiro elétrico: dinâmica de consumo após acionamento (flutuações após acionamento, tendência a estabilizar, porém durante a operação observada neste conjunto se trata de uma C5 durante toda sua atuação);
- Televisão CRT: dinâmica de consumo permanente;
- Lâmpada fluorescente (LF) 25 W (2 unidades), 22 W (2 unidades), 15 W (3 unidades), 9 W;
- Lâmpada incandescente (LI) 40 W (2 unidades).

Sua medição foi realizada com o medidor do CEPEL/Eletrobras. O perfil de seu consumo pode ser visualizado na Figura 6.25. A informação no gabarito pode ser observada nas figuras 6.26 e 6.27. A Figura 6.26 contém a informação do consumo temporal dos equipamentos, enquanto a Figura 6.27 contém o gráfico circular do consumo estimado no gabarito para os equipamentos.

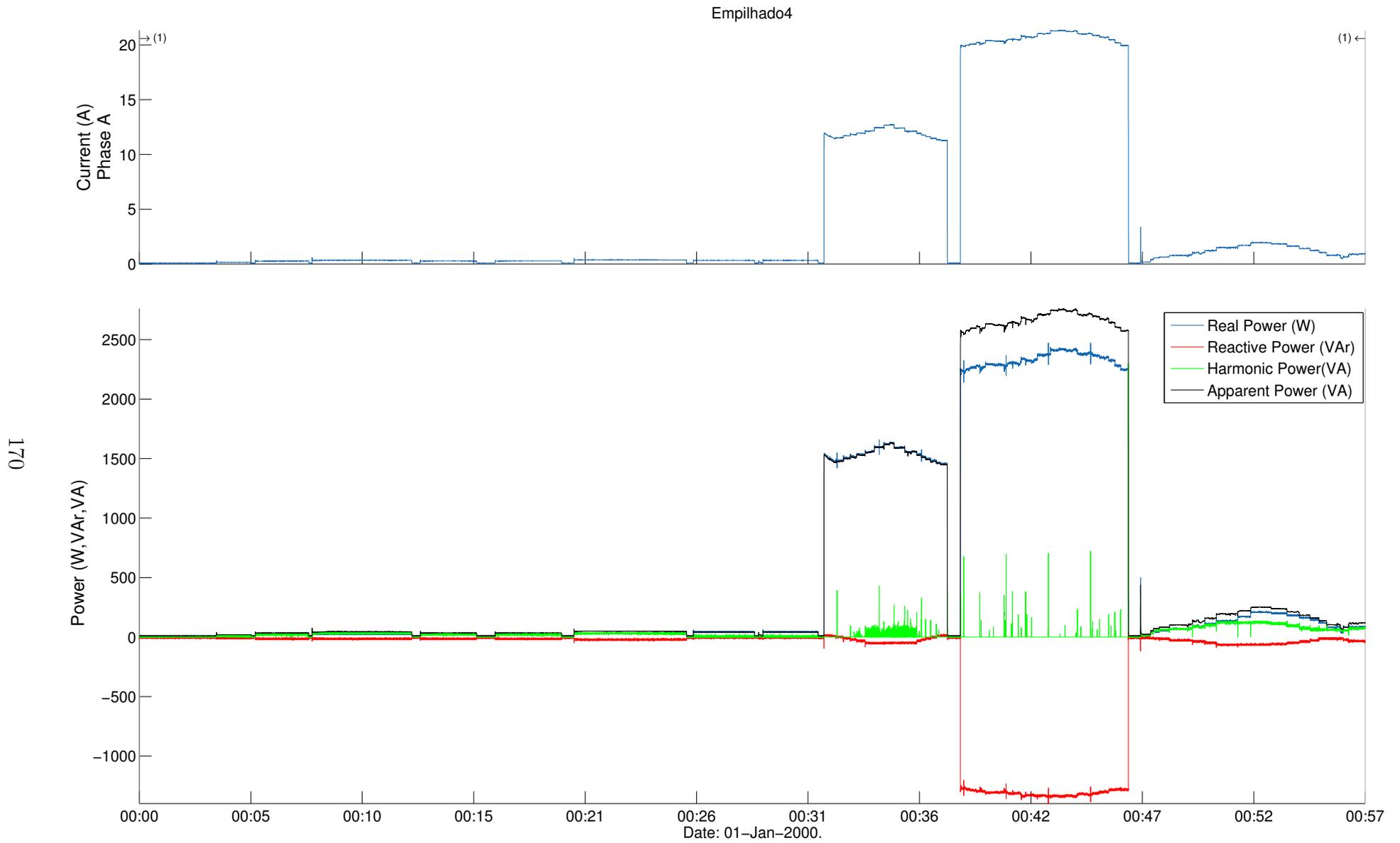


Figura 6.25: Perfil de consumo agregado para o conjunto de dados *Empilhado4*.

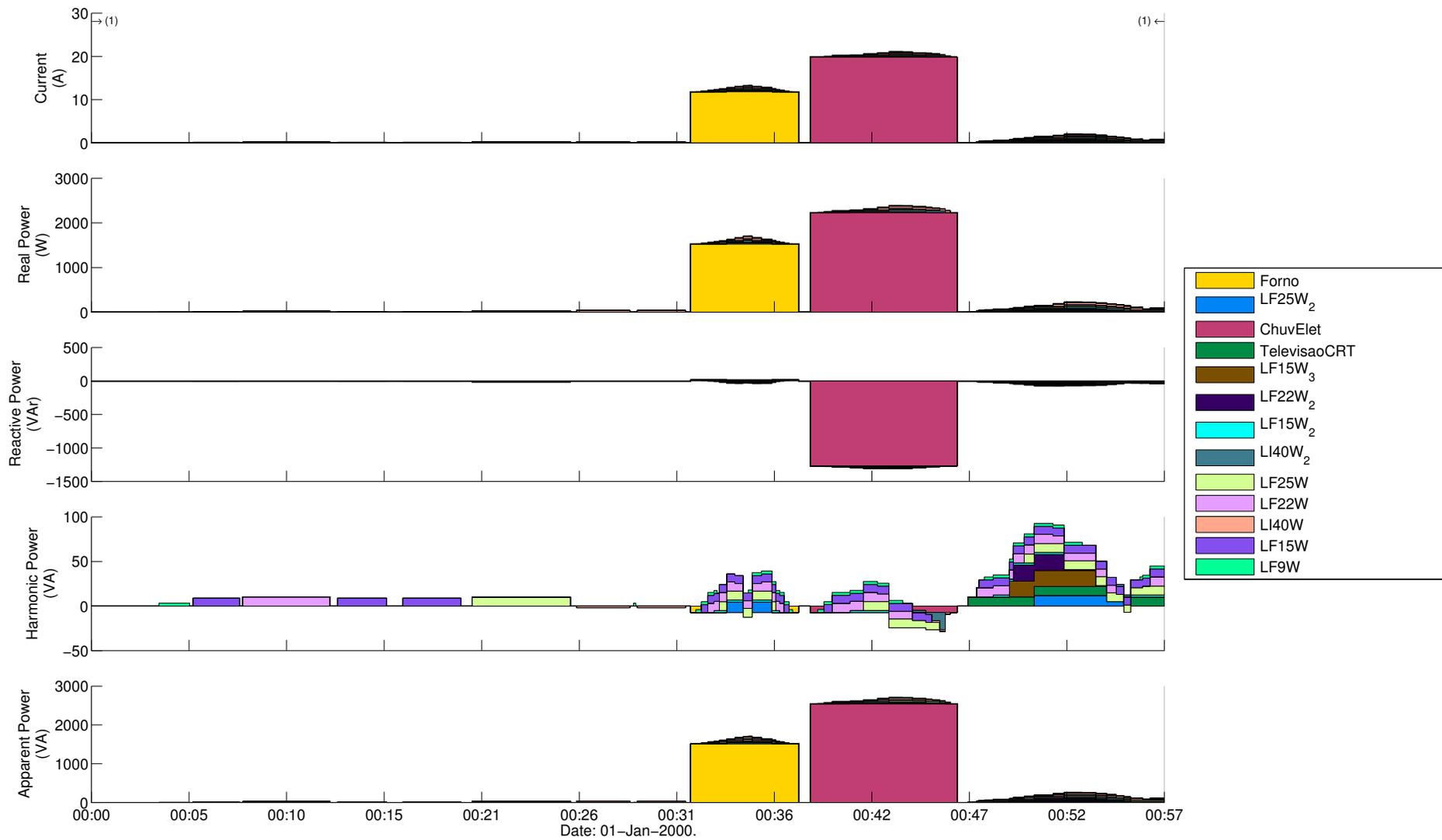


Figura 6.26: Informação no gabarito para o conjunto de dados *Empilhado4*: consumo temporal dos equipamentos.

Appliances Real Power Consumption

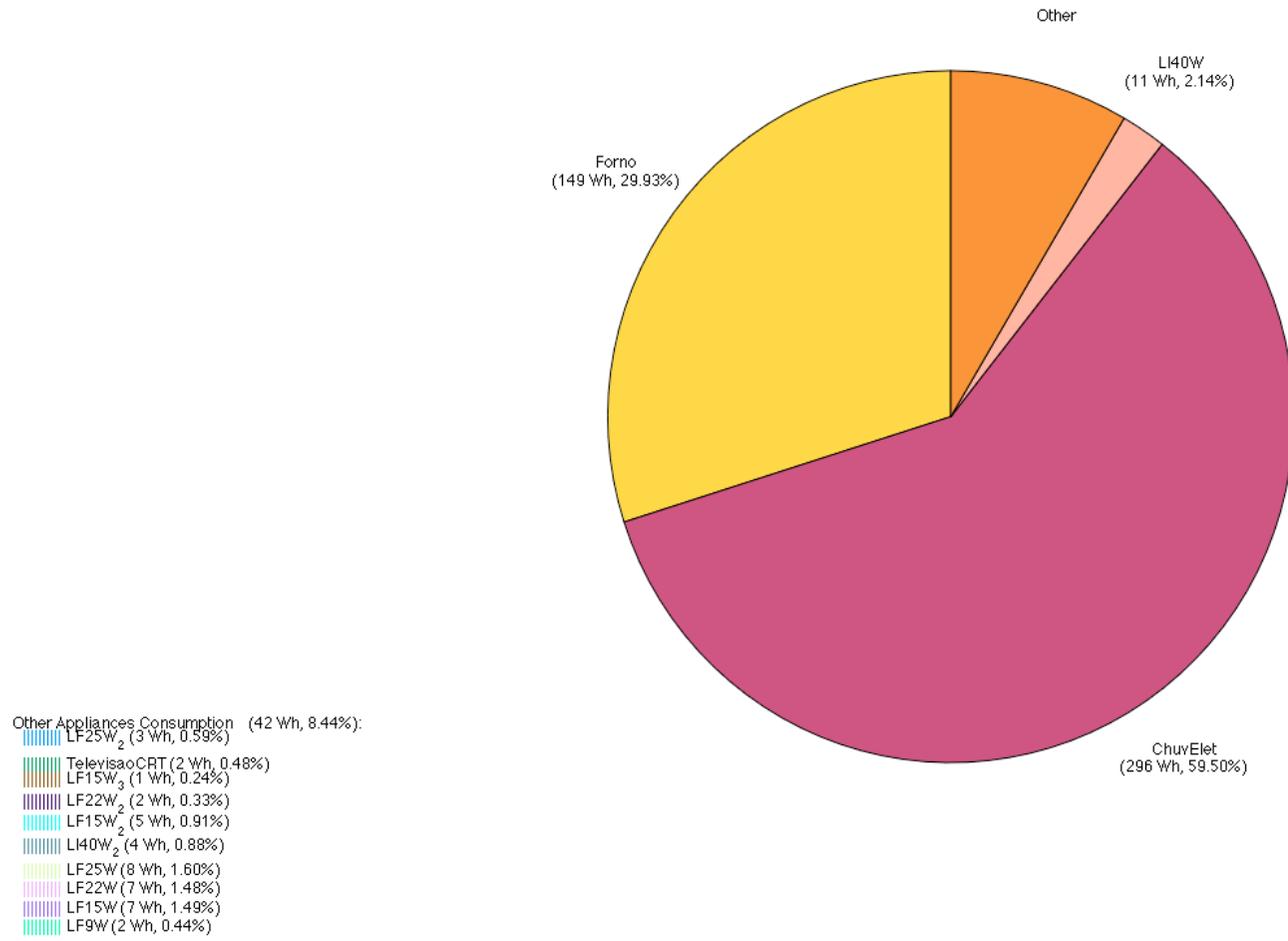


Figura 6.27: Informação no gabarito para o conjunto de dados *Empilhado4*: gráfico circular do consumo dos equipamentos.

6.1.6 Conjunto de dados *Empilhado7*

Este é o conjunto com a maior operação simultânea de equipamentos de diferentes uso-finais. Há a ocorrência de um evento de transitório aonde ocorre o desacionamento de quatro equipamentos em conjunto (geladeira, uma lâmpada incandescente 100 W, e duas lâmpadas fluorescentes 20 W e 24 W). Outra peculiaridade deste arquivo é a distorção dos eventos de transitório durante o momento que o ar condicionado está operando (um total de 13 eventos presentes). Os equipamentos presentes neste conjunto são:

- Lâmpada incandescente (LI) 60 W, 100 W;
- Lâmpada fluorescente (LF) 20 W, 21 W (circular), 24 W, 26 W, 28 W, 40 W;
- Secador de cabelo;
- Ar condicionado (constitui de uma C5);
- Sanduicheira;
- Geladeira (obs: essa geladeira tem consumo bastante superior àquele utilizada no arquivo *Temporizado*);
- Televisão CRT (constitui de uma C5).

Sua medição foi realizada com o medidor do CEPEL/Eletronbras. A informação no gabarito pode ser observada nas figuras 6.29–6.30. A Figura 6.29 contém a informação do consumo temporal dos equipamentos, enquanto a Figura 6.30 contém o gráfico circular do consumo estimado no gabarito para os equipamentos.

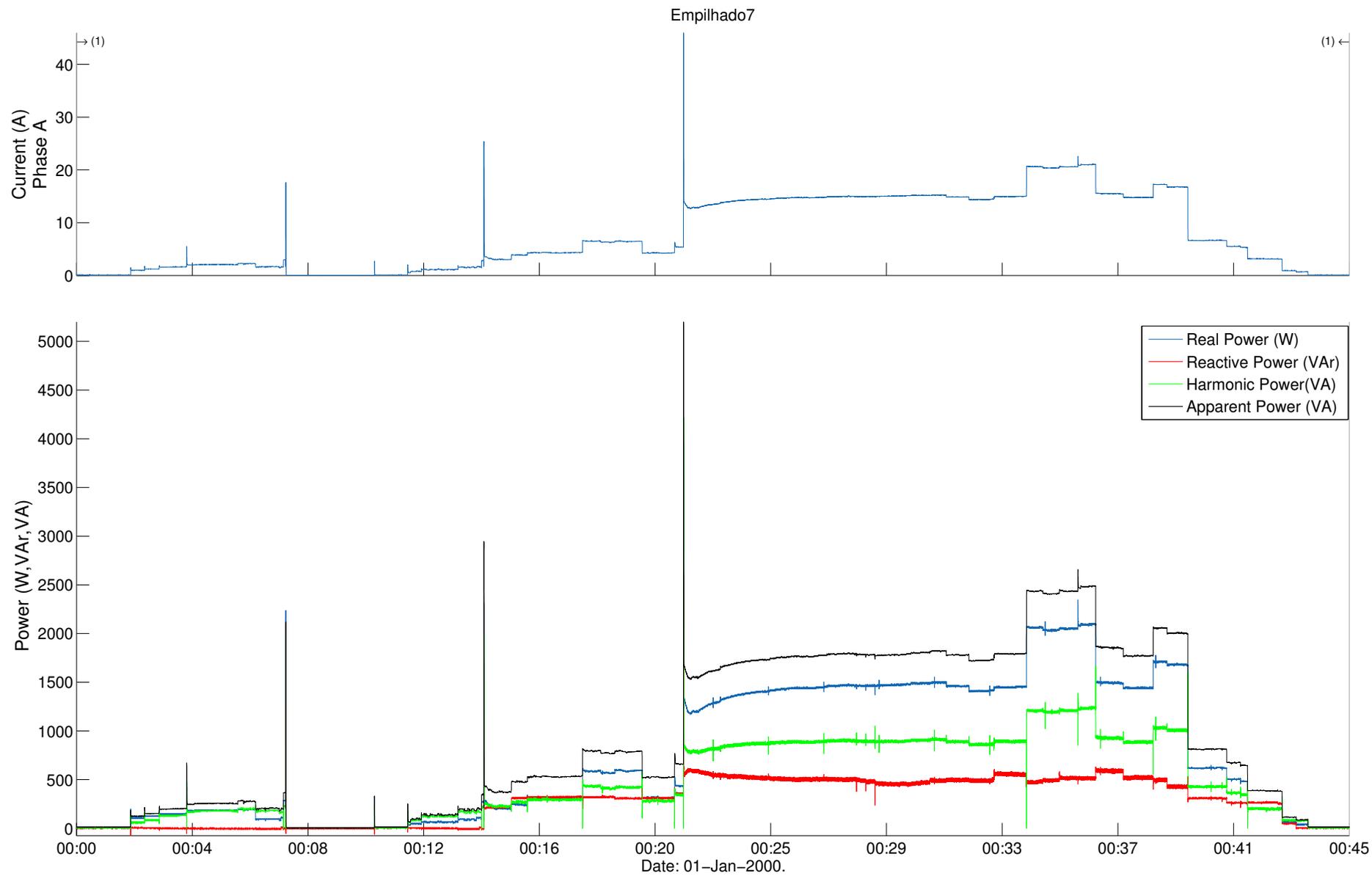


Figura 6.28: Perfil de consumo agregado para o conjunto de dados *Empilhado7*.

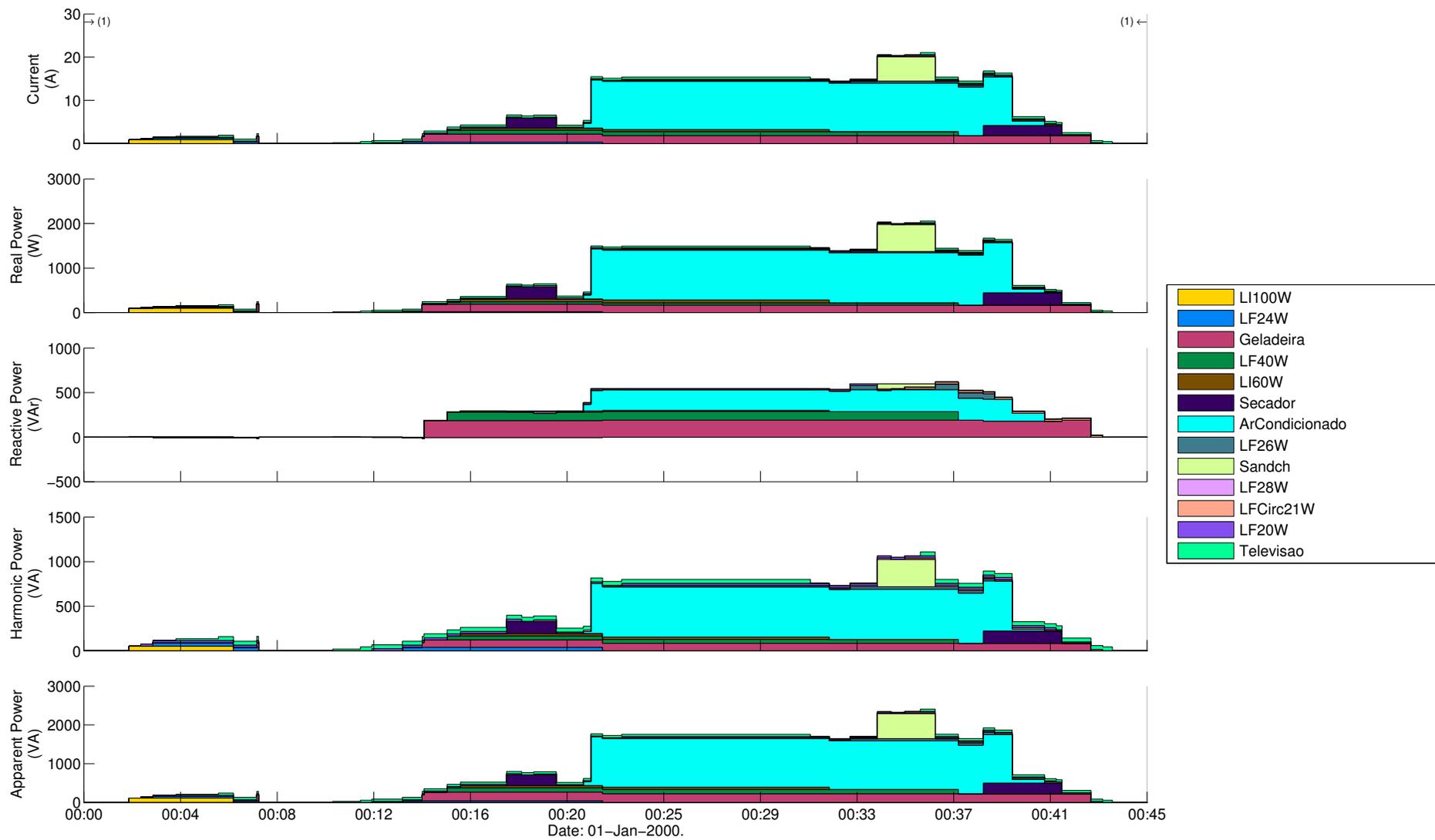


Figura 6.29: Informação no gabarito para o conjunto de dados *Empilhado7*: consumo temporal dos equipamentos.

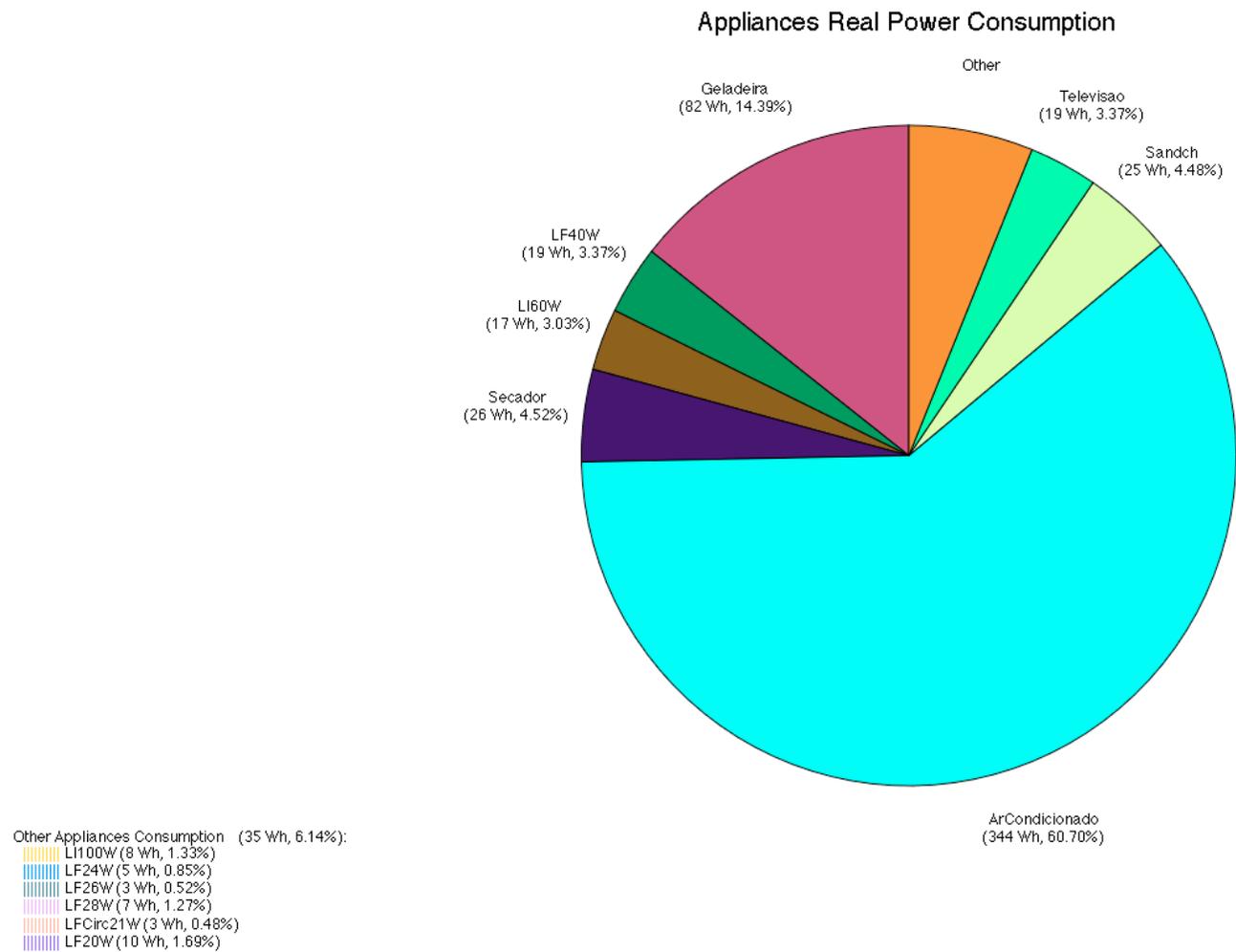


Figura 6.30: Informação no gabarito para o conjunto de dados *Empilhado7*: gráfico circular do consumo dos equipamentos.

6.2 Aplicação do Algoritmo Genético de Estratégia Evolutiva (ES) para Otimização do Detector de Eventos

Esta seção se dedica a descrição da metodologia final adotada para a detecção de eventos por este trabalho que foi aplicada na base de dados descrita na seção anterior. A metodologia base é aquela explicada na Seção 4.3.2, proposta pelo CEPEL/Eletrobras para a detecção de eventos de transitório. Resumidamente, a metodologia utiliza um filtro para convoluir a derivada de Gaussiana com o sinal de corrente, que irá gerar regiões sensibilizadas caso a resposta do filtro ultrapasse um limiar. O ponto de inflexão dessas regiões representam candidatos a evento de transitório, que serão aceitos caso não tenham sido gerados por ruído ou estejam demasiadamente próximos a outros candidatos. Porém, a mesma foi reimplementada de acordo com as considerações realizadas no Capítulo 5 que, além de diversas alterações técnicas, trouxeram três alterações:

- Remoção de eventos próximos utilizando sua média (ver pp. 115);
- Remoção de eventos inconsistentes (ver pp. 115);
- Ajuste automático dos parâmetros através de um otimizador, que neste trabalho se trata da versão multiespécie do ES implementado e detalhado na Seção 5.5.

Antes de entrar no mérito do ajuste automático, este trabalho também realizou um ajuste empírico dos parâmetros como era feito pelo CEPEL/Eletrobras na sua metodologia original, que será referido como *Ajuste Manual*. Isso foi realizado para permitir uma melhor afinidade com o ambiente de análise, uma vez que a utilização unicamente de parâmetros otimizados automaticamente não permite a compreensão de nuances da operação e capacidade da metodologia. Apesar de ser realizada manualmente, algumas diferenças podem ser citadas em relação ao ajuste realizado pelo CEPEL/Eletrobras. A primeira, e mais importante, é que o ajuste foi realizado para os conjuntos de dados *Temporizado*, que contém a presença mais forte de ruídos. Além disso, diferente do ajuste automático de parâmetros que será explicada a seguir e opera ajustando vários parâmetros citados no início deste capítulo (pp. 138), o *Ajuste Manual* foi realizado apenas para o δ_{min} , enquanto os demais parâmetros são os mesmos empregados pela metodologia do CEPEL/Eletrobras. Contudo, a versão do algoritmo já permitia a utilização de novas maneiras de remoção de evento, que foram analisadas e optou-se empregar a versão com remoção de eventos inconsistentes e utilização da média para remoção de eventos próximos.

A metodologia original do CEPEL/Eletrbras já apresentava cinco diferentes maneiras para a remoção de candidatos, indicadas na Figura 4.4. A adição de duas novas maneiras de remoção de candidatos aumentou a quantidade de possibilidades de configurações a serem analisadas com relação a quais exames empregar para a remoção de candidatos. Não se espera que a questão dos exames realizados para a remoção seja um fator chave no desempenho da metodologia, e sim o valor dos parâmetros empregados. Porém, não se sabe *a priori* qual é a configuração mais indicada para a operação, sendo necessário realizar comparações entre elas para determinar qual é a mais adequada. Ao invés de realizar a otimização de cada configuração individualmente e então determinar qual teve a melhor convergência, implementou-se a versão multiespécie do ES (ver pp. 131) que permite dinâmica na quantidade de esforço computacional reservado para cada configuração. Assim, analisar-se-á diversas configurações, cada uma representada como uma espécie competindo durante a evolução no ES, que terão maior ou menor esforço computacional dedicado conforme sua melhor ou pior evolução.

Ainda assim, a avaliação de uma estratégia em força bruta — aqui se referindo a otimizar todas as possíveis configurações e identificar a melhor convergência delas — não parece ser a melhor maneira de abordar o problema, mesmo que houvesse capacidade computacional para realizar tal tarefa em tempo hábil. Percebeu-se a necessidade de realizar a escolha de algumas configurações a serem testadas para reduzir a quantidade de caminhos possíveis. Durante a realização do *Ajuste Manual*, percebeu-se que os eventos removidos devido à inconsistência eram sempre eventos de falso alarme, mas não havia a ocorrência de perda de eventos de detecção causados por esse tipo de remoção. Por isso, determinou-se que a remoção de eventos de inconsistentes seria sempre realizada durante o ajuste de parâmetros. Por outro lado, a remoção de eventos devido à ruído era importante para remoção de perturbações rápidas geradas na rede que não constituíam na mudança do patamar operativo da rede e, com base nisso, determinou-se que a remoção de eventos ruidosos também sempre seria realizada. Assim, a variável ΔI_{min} sempre estará presente no material genético das espécies, bem como o σ da Gaussiana a ser utilizada no filtro de derivada de Gaussiana e o seu valor de corte δ_{min} . Já para o caso da remoção de eventos próximos, não era possível determinar se sua utilização era necessária, nem qual das versões de remoção — por média, ou sem deslocamento — era o que mais se adequava ao problema. No caso de não usar corte, a variável n_{min} não estará presente no material genético das espécies, porém no caso oposto as espécies há a ocorrência de um gene a mais, cujo fenótipo é inteiro. Ainda assim, sua representação no genótipo é realizada por números reais, sendo preciso determinar como codificar a informação. Simplesmente se escolheu arredondar o valor na representação para obter o valor do fenótipo. Finalmente, também não se sabia *a priori* determinar se

a ordem de aplicação das remoções pode causar alguma influência na capacidade de detecção e, apesar de não se esperar grandes diferenças devido à mudança da ordem em que eles seriam removidos, decidiu-se testar ambas configurações.

Assim, decidiu-se empregar cinco configurações a serem analisadas, que determinam as espécies utilizadas no ES. As mesmas estão descritas no início deste capítulo, constituindo as espécies (I)–(V) na pp. 139.

Além da questão das configurações, era necessário determinar a inicialização. Para garantir melhor convergência, empregou-se fronteiras mínimas e máximas para cada um dos valores representados pelos genes, sendo estas:

$$0,003 < \sigma_{gauss} < 0,5 \quad (6.1a)$$

$$0,009 < \delta_{min} < 0,5 \quad (6.1b)$$

$$0,0 < \Delta I_{min} < 1,0 \quad (6.1c)$$

$$5,0 < n_{min} < 500,0^3 \quad (6.1d)$$

A inicialização foi aleatória uniforme dentro das fronteiras descritas em 6.1. Para o parâmetro da estratégia evolutiva, porém, iniciou-se esse valor de acordo com 6.2a, garantindo que inicialmente cada indivíduo tenha, em média, 95% de chance de explorar uma região no interior a 10% do espaço de solução permitido. Os limites inferiores e superiores foram determinados por 6.2b e 6.2c, respectivamente. Os valores de $\sigma_{min,i}$ e $\sigma_{max,i}$ garantem que, no mínimo, o material genético sofrerá perturbações dentro da região de 0,01% da região de solução em 67% dos casos, enquanto o mesmo será perturbado no máximo em 30% da região para a mesma probabilidade:

$$2\sigma_{init,i} = 0,1\Delta x_i \quad (6.2a)$$

$$\sigma_{min,i} = 1 \times 10^{-4}\Delta x_i \quad (6.2b)$$

$$\sigma_{max,i} = 3 \times 10^{+3}\sigma_{min,i} \quad (6.2c)$$

onde Δx_i é a região do espaço dentro das fronteiras para a i -ésima variável.

Além disso, há de escolher os parâmetros livres no cálculo da aptidão explicitada em 5.2. Os parâmetros γ_{det} e γ_{fa} permitem priorizar uma melhor taxa de detecção ou menor taxa de falso alarme, em detrimento de um acréscimo na taxa de falso alarme ou decréscimo na taxa de detecção, respectivamente. Isso pode ser feito dando valores maiores ou menores para a recompensa de encontrar corretamente

³Apesar do n_{min} ser inteiro na representação no fenótipo, sua representação é por número reais no genótipo por estar utilizando-se do algoritmo genético por ES. Também se reitera, aqui, que esse gene só estará presente nas configurações que realizam a remoção de eventos próximos.

um evento de transitório ou a penalidade de aceitar um evento que se constitui de falso alarme. Os valores utilizados foram a unidade para γ_{det} e -0,9 para γ_{fa} . O objetivo dessa diferença é uma leve priorização para a detecção de eventos em detrimento de uma maior taxa de falso alarme. O valor para γ_{rem} foi de -0,05. Esse parâmetro é importante para garantir a obtenção de soluções mais simples, que exploram a capacidade do filtro de Gaussiana em gerar o mínimo possível de candidatos, o que reduz a carga da remoção de eventos através de cortes em ΔI_{min} ou n_{min} . Ao mesmo tempo, a geração de menos candidatos significa que a solução exige menor esforço computacional, e é mais adequada para a aplicação na operação em tempo real. Ainda, o parâmetro J_{max} foi definido como 50 amostras.

Finalmente, cabe ainda decidir qual dos métodos de competição para as espécies será utilizado e os tamanhos das populações. Como a implementação do método de competição intraespécies, da maneira atualmente disponível neste trabalho, acaba favorecendo espécies com carga genética mais simples de serem otimizadas ou que começaram em condições privilegiadas — quando sem o mecanismo de adiamento da competição —, irá optar-se pela competição intraespécie. Para μ , utilizou-se o valor de inicial de 30 indivíduos para cada uma das espécies. Como há cinco espécies que serão otimizadas, a população global é de 150 indivíduos. Cada indivíduo gerará 7 filhos, de forma que o valor de λ global é de 1050 indivíduos.

Capítulo 7

Resultados

A metodologia (Seção 6.2) foi aplicada na base de dados (Seção 6.1), ambos descritos no capítulo anterior, para verificar a capacidade da versão multiespécie implementada do ES (Subseção 5.5.1) de realizar o ajuste automático de parâmetros para os diferentes cenários simulados nos conjuntos de dados. Como dito durante o detalhamento da metodologia final aplicada neste trabalho, antes do ajuste automático dos parâmetros, também se realizou neste trabalho um ajuste empírico para os novos cenários presentes no conjunto de dados *Temporizado Gabarito 1*, que representa um cenário com atuação simultânea de cinco equipamentos e operação de um equipamento com dinâmica de carga — os equipamentos C5 (ver definições na Subseção 4.2.1). Esse ajuste será referido como *Ajuste Manual*, enquanto o ajuste empírico realizado pelo CEPEL/Eletronbras, que foi feito levando em consideração os conjuntos de dados *NI00****, será referido como *Ajuste CEPEL* (mais detalhes na discussão da metodologia final, na Seção. 6.2). Por sua vez, foram realizados três ajustes automáticos, todos seguindo a metodologia descrita na Seção 6.2. Estas são as características e objetivos de cada um desses ajustes:

- *ES 1*: a otimização foi realizada alimentando o ES com os conjuntos de dados *NI00****. O objetivo é permitir uma comparação entre o otimizador e o *Ajuste CEPEL*;
- *ES 2*: neste caso os conjuntos de dados alimentados para o otimizador foram o *Temporizado Gabarito 1*, *Empilhado4* e *Empilhado7*. Assim, é possível observar o comportamento da metodologia quando otimizada para os dados que representam uma simulação mais real das características presentes em uma rede residencial;
- *ES 3*: já a última configuração alimentou o ES apenas com o conjunto *Temporizado Gabarito 1* e *Empilhado4* para analisar a capacidade de generalização do algoritmo.

Não obstante, aplicou-se o Detector de Patamar Elaborado desenvolvido pelo trabalho de Cañón [102] para verificar o comportamento do mesmo nos novos dados. Porém, não foi realizado um novo ajuste de seus parâmetros, de forma que os resultados mostrados para esse detector não representam sua real capacidade.

Um detalhe que aqui deve ser recapitulado é a presença de dois gabaritos para o conjunto *Temporizado*. Isso se deve à presença de um distúrbio que tem características similares a uma C3 exibido na Figura 6.24, que está presente apenas na informação do gabarito *Temporizado Gabarito 2* (ver Subseção 6.1.4 para mais detalhes).

7.1 Otimização Automática dos Parâmetros

Na Tabela 7.1, encontram-se os resultados os ajustes comentados anteriormente no início deste capítulo, exceto o Detector de Patamar Elaborado que será apresentado mais adiante.

Fica evidente a maior sensibilidade da escolha dos parâmetros para o caso *Ajuste CEPTEL*, para o qual ocorre uma maior capacidade de detecção, em especial para o *Empilhado4* que contém eventos de equipamentos de menor consumo (lâmpadas fluorescentes). Porém, isso também reflete em uma taxa de falso alarme excessiva para o conjunto *Temporizado* aonde, mesmo na sua versão com o gabarito contendo o distúrbio descrito, ocorrem 234 eventos de falso alarme. A alta sensibilidade dos valores também é refletida no conjunto *Empilhado7*, que, assim como o *Temporizado*, contém a presença de duas C5. Porém, o arquivo é mais curto e conta com a televisão operando durante apenas 30 min e ar condicionado por cerca de 20 min, enquanto a televisão no conjunto de dados *Temporizado* opera por ~ 12 h. Já os conjuntos *NI00****, para os quais foi realizado o ajuste dessa configuração, apresentam reconstrução praticamente perfeita, exceto pela presença de dois falsos alarmes no conjunto *NI00171*. A Figura 7.1 apresenta a ocorrência desses dois falsos alarmes, que ocorrem justamente após o acionamento do computador portátil aonde há diversas oscilações até a estabilização do consumo. As taxas globais de detecção e falso alarme para os conjuntos *NI00**** foram respectivamente de 100 % e 1,2 %, enquanto esses valores de 90,9 % e 111,7 % para os conjuntos *Temporizado Gabarito 1*, *Empilhado4* e *Empilhado7*. No caso do *Temporizado* ser utilizado com a segunda versão de seu gabarito, as taxas são respectivamente de 90,2 % e 82,9 %.

A versão de *Ajuste Manual* foi a menos sensível, na qual se obteve baixa taxa de detecção para o conjunto de dados *Empilhado4* que contém as lâmpadas de baixo consumo, mas ao mesmo tempo garantindo um falso alarme total de apenas 29 ocorrências (28 ocorrências no *Temporizado Gabarito 1* e 1 no *Empilhado7*). Levando em conta que o seu ajuste foi realizado levando em consideração apenas

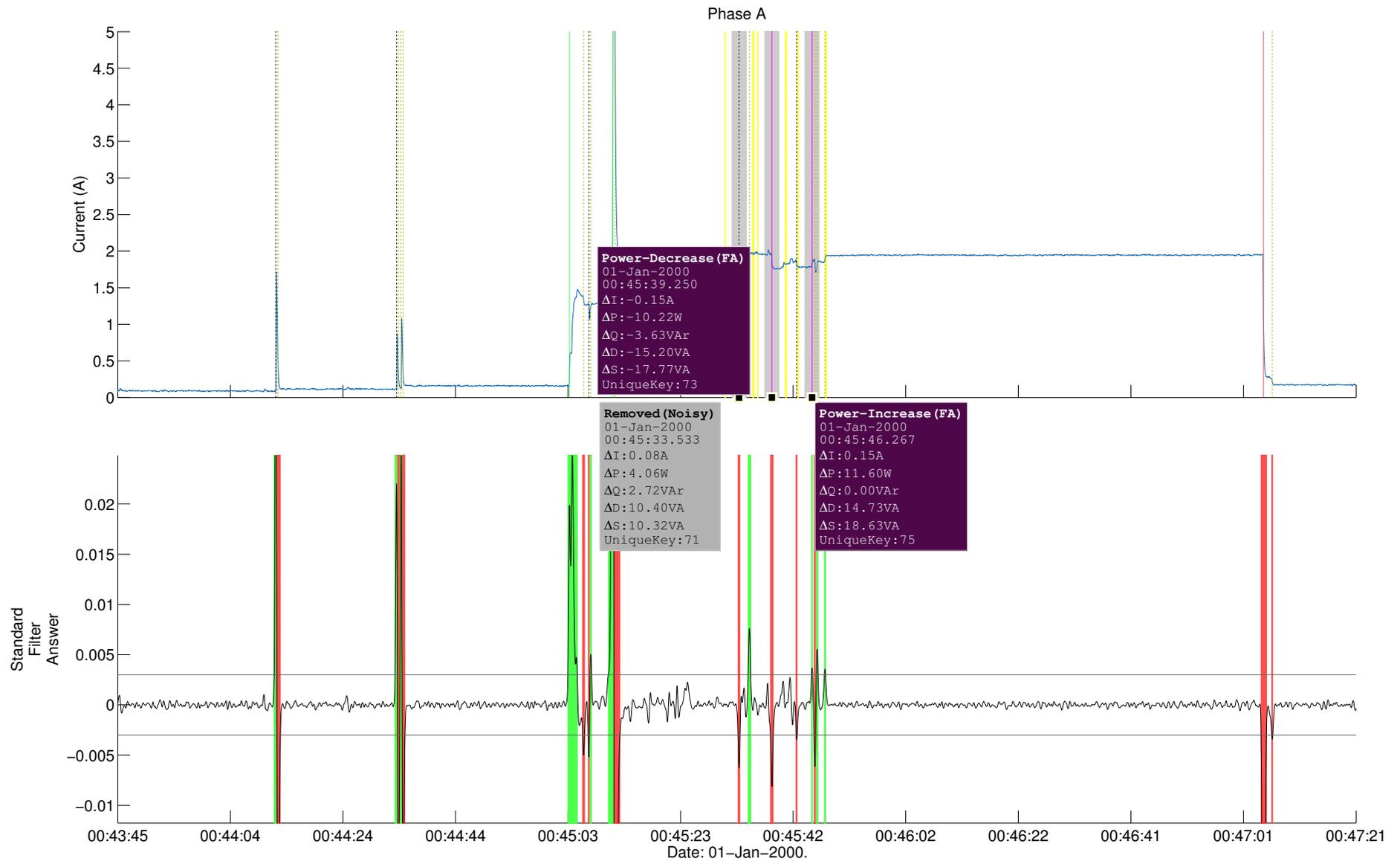


Figura 7.1: Falsos alarmes para a configuração *Ajuste CEPEL* no conjunto de dados *NI00171*. Os eventos marcados por etiquetas roxas indicam as ocorrências dos falsos alarmes durante o acionamento do computador portátil antes da estabilização de seu consumo. Observe que há a remoção de diversos outros candidatos devido à eventos próximos ou ruído que causariam ocorrência ainda maior de falsos alarmes caso não houvessem suas remoções.

Conjunto de Dados		ES 1		Manual		CEPEL		ES 2		ES 3	
		DET	FA	DET	FA	DET	FA	DET	FA	DET	FA
<i>NI00168</i>	Ocorr.	38	0	33	0	39	0	36	0	23	2
(39 eventos)	Taxa (%)	97,4	0,0	86,8	0,0	100,0	0,0	92,3	0,0	59,0	5,1
<i>NI00171</i>	Ocorr.	47	1	42	0	48	2	43	0	39	2
(48 eventos)	Taxa (%)	97,9	2,1	89,4	0,0	100,0	4,2	89,6	0,0	83,0	4,2
<i>NI00173</i>	Ocorr.	20	1	19	1	20	0	20	1	20	0
(20 eventos)	Taxa (%)	100,0	5,0	95,0	5,0	100,0	0,0	100,0	5,0	100,0	0,0
<i>NI00174</i>	Ocorr.	8	0	7	1	8	0	8	0	7	0
(8 eventos)	Taxa (%)	100,0	0,0	87,5	12,5	100,0	0,0	100,0	0,0	87,5	0,0
<i>NI00175</i>	Ocorr.	23	0	23	0	23	0	23	0	23	0
(23 eventos)	Taxa (%)	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0
<i>NI00177</i>	Ocorr.	24	0	24	0	24	0	24	0	24	0
(24 eventos)	Taxa (%)	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0
<i>Temp. Gab. 1</i>	Ocorr.	147	52	147	28	147	259	148	43	148	13
(149 eventos)	Taxa (%)	98,7	34,9	98,7	18,8	98,7	173,8	99,3	28,9	99,3	8,7
<i>Temp. Gab. 2</i>	Ocorr.	191	8	174	1	201	234	188	3	161	1
(211 eventos)	Taxa (%)	90,5	3,8	82,5	0,0	95,3	110,9	89,1	1,4	76,3	0,5
<i>Empilhado4</i>	Ocorr.	42	1	24	0	56	10	37	0	34	0
(74 eventos)	Taxa (%)	56,8	1,4	32,4	0,0	78,0	12,3	50,0	0,0	45,9	0,0
<i>Empilhado7</i>	Ocorr.	39	1	35	1	38	27	37	0	34	1
(42 eventos)	Taxa (%)	92,9	2,4	83,3	2,4	90,5	64,3	88,1	0,0	80,9	2,4

Tabela 7.1: Taxa de detecção de eventos de transitório e falso alarme para os três ajustes automáticos e os dois ajustes manuais. As configurações Manual e CEPEL referem-se respectivamente aos casos determinados empiricamente pelo autor do trabalho e pelo grupo do CEPEL/Eletrabras, sendo o primeiro ajustado para os dados *Temporizado Gabarito 1* e o segundo para os dados *NI00****. As configurações ES 1, ES 2 e ES 3 referem-se ao ajuste automático do ES alimentado respectivamente pelos conjuntos: *NI00****; *Temporizado Gabarito 1*, *Empilhado4* e *Empilhado7*; *Temporizado Gabarito 1* e *Empilhado7*. As abreviaturas DET e FA referem-se a taxa de detecção e falso alarme, respectivamente.

o conjunto de dados *Temporizado Gabarito 1*, isso mostra que a capacidade de visualização permite encontrar um valor que reduziria a sensibilidade do detector, sem que fosse analisado vários casos até a obtenção dos parâmetros. Porém, para os conjuntos sem a presença de ruídos, essa menor sensibilidade se reflete em perdas de detecção nos conjuntos *NI00168* e *NI00171* que possuem equipamentos de menor consumo.

Antes de discutir a questão das taxas de detecção e falso alarme para os ajustes automáticos, detém-se a convergência de suas espécies em cada um dos casos. No ES 1, a espécie que obteve maior aptidão foi a (III), enquanto nas configurações ES 2 e 3, foi a Espécie (II). As Figuras 7.2–7.4 mostram a evolução das espécies em termo de aptidão para os ES 1, 2 e 3, respectivamente. Observa-se que a Espécie (III) nas otimizações ES 2 e 3, não consegue acompanhar o processo evolutivo das demais espécies, porém isso não ocorre na configuração ES 1, onde todas as espécies convergem. Há indícios de que a espécie (III) não é adequada para os conjuntos

de dados ruidosos, porém é necessário a realização de mais execuções do algoritmo nessas condições para obter mais estatística e confirmar isso, uma vez que esses casos podem ter ocorrido devido à convergência de um máximo local de aptidão. Porém, não foi possível aumentar essa estatística devido à demora no tempo de execução das configurações ES 2 e 3, que levam cerca de uma semana para convergir.

Indo além da questão da Espécie (III), observa-se que não há grande diferença na aptidão das demais espécies na convergência. Isso indica que quaisquer outras configurações além da espécie (III) é capaz de resolver igualmente bem o problema. Um motivo para isso ocorrer é a presença de eventos próximos apenas nos acionamentos com um pico de consumo, que causa a sensibilização e geração de um candidato na descida de consumo após o pico. Esses candidatos são devidamente removidos por inconsistência, método de remoção adicionado justamente para o tratamento desses casos. Ao observar o material genético de convergência dos ajustes automáticos, observa-se que o valor de n_{min} foi de 6, 5 e 495, para as melhores espécies na convergência dos ES 1, 2 e 3, respectivamente. Exceto para o ES 3, ocorrência *sui generis* que será comentado mais adiante quando se referindo a sua taxa de detecção e falso alarme, os valores de convergência para n_{min} são próximos ao limite inferior, o que mostra que esse parâmetro não é relevante para o problema. Assim, para os conjuntos de dados analisados, a remoção de eventos inconsistentes é suficiente para a remoção de eventos de falso alarme que antes eram removidos por constituírem de eventos próximos, porém, a remoção de eventos inconsistentes é mais recomendada uma vez que ela não adiciona tempo morto de resposta do filtro, que pode causar perda de alvo.

Retornando à questão das taxas de detecção e falso alarme, a configuração ES 1 obteve eficiência compatível ao *Ajuste CEPEL* nos conjuntos *NI00****, com umas poucas diferenças. Há a perda de alvo de um evento nos conjuntos *NI00168* e *NI00171* que é causada durante o desacionamento de uma lâmpada com *dimmer*, que ocorre em um processo lento e, portanto, que necessita de uma maior sensibilidade do filtro. A ocorrência de falso alarme no conjunto de dados é devida a um ruído próximo ao acionamento do ventilador, causado por uma má manipulação da chave que seleciona a velocidade do ventilador e indicado na Figura 7.5. Porém, é ao observar as taxas para os conjuntos *Temporizado Gabarito 1*, *Empilhado4* e *Empilhado7* que se percebe a real vantagem do ajuste automático dos parâmetros. Apesar do mesmo não ter tido acesso à esses conjuntos de dados, sua capacidade de generalização foi bastante superior àquela apresentada pelo *Ajuste CEPEL*. Isso é causado pela presença do parâmetro γ_{rem} (ver equação 5.2, para o cálculo da aptidão na pp. 125), que causa uma tendência maior para a exploração dos limites do limiar de sensibilização. Ao invés da excessiva ocorrência de falso alarme do *Ajuste CEPEL* com 259, 10 e 27 casos para esses conjuntos, obteve-se apenas 52, 1 e 1

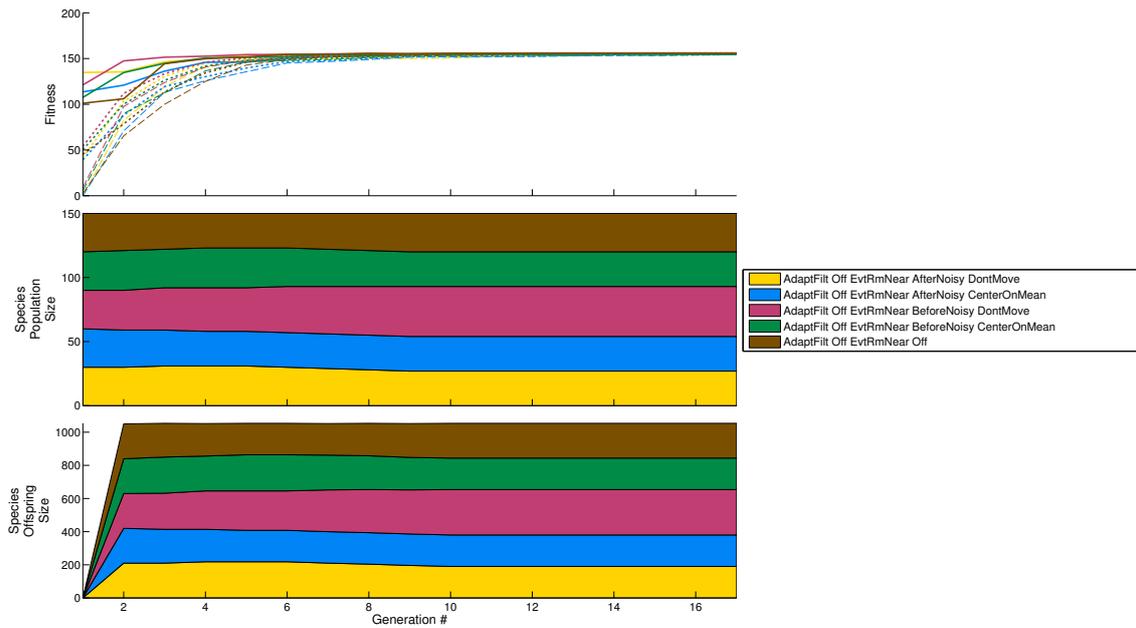


Figura 7.2: Convergência das espécies para o ajuste automático da configuração ES 1. As espécies e suas cores são: Espécie (I): amarelo; Espécie (II): azul; Espécie (III): rosa; Espécie (IV): verde; Espécie (V): marrom. Observa-se que praticamente não há diferença na aptidão das espécies na convergência.

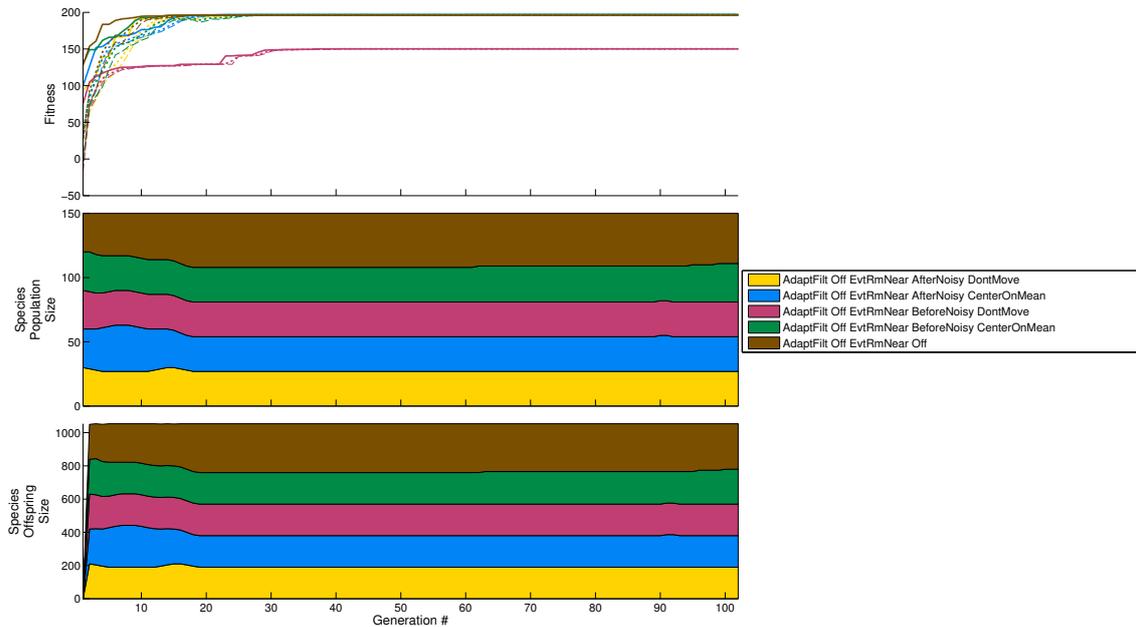


Figura 7.3: Convergência das espécies para o ajuste automático da configuração ES 2. As espécies e suas cores são: Espécie (I): amarelo; Espécie (II): azul; Espécie (III): rosa; Espécie (IV): verde; Espécie (V): marrom. A Espécie (III) não obteve convergência como as demais espécies.

casos de falso alarme. Isso, por sua vez, trouxe uma menor detecção no conjunto *Empilhado4* que contém os aparelhos de menor consumo. A taxa de detecção e falso alarme do ES 1 para os conjuntos *NI00**** é de 98,8 % e 1,2 %, respectivamente.

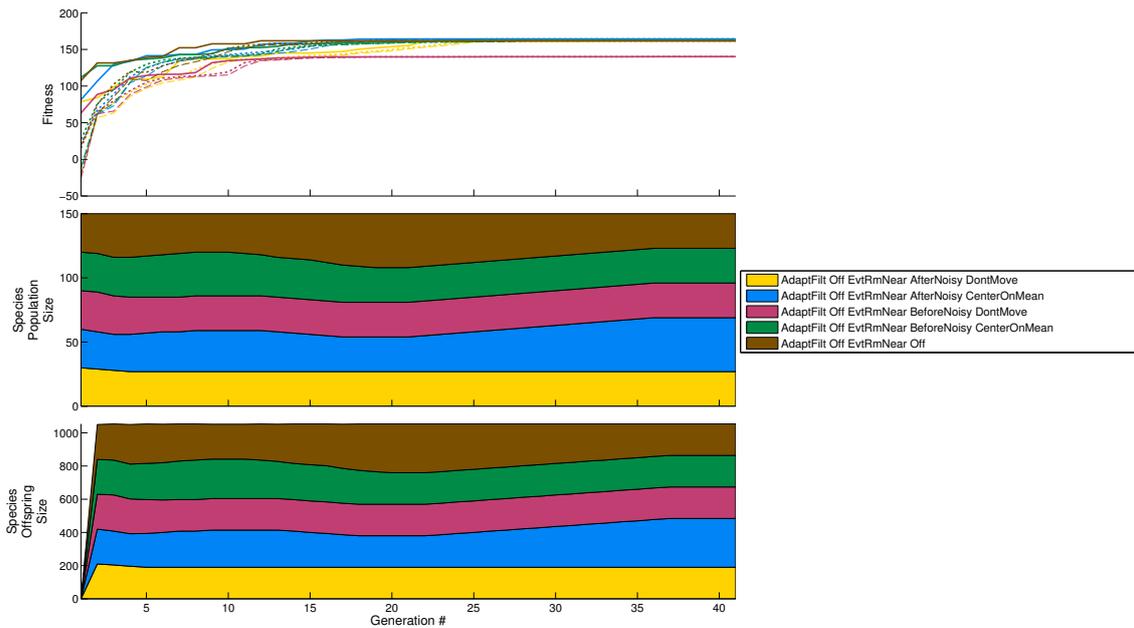


Figura 7.4: Convergência das espécies para o ajuste automático da configuração ES 3. As espécies e suas cores são: Espécie (I): amarelo; Espécie (II): azul; Espécie (III): rosa; Espécie (IV): verde; Espécie (V): marrom. A Espécie (III) novamente não obteve convergência como as demais espécies.

Para os conjuntos *Temporizado Gabarito 1*, *Empilhado4* e *Empilhado7* essas taxas são de 86,03 % e 16,51 %, enquanto esses valores são de 83,2 % e 3,0 % quando empregando o Gabarito 2.

Em seguida, alimentou-se o otimizador para realizar a otimização dos parâmetros recebendo os conjuntos de dados ruidosos, na esperança de que isso trouxesse a redução na taxa de falso alarme ainda elevada no conjunto *Temporizado Gabarito 1*. Na época, ainda não havia sido observada a perturbação presente nesse conjunto de dados. A resposta para essa configuração está na coluna ES 2 da Tabela 7.1. Ao analisar sua taxa de eficiência, percebeu-se que a taxa de falso alarme, apesar de sofrer redução, não foi na ordem esperada: o valor se manteve alto, 43 ocorrências de falso alarme nesse conjunto de dados, o que necessitou de uma investigação mais profunda. A descoberta foi a perturbação presente na Figura 7.6, e observada em mais detalhes na Figura 6.24. Como foi dito durante a descrição desse conjunto de dados, essa perturbação tem uma característica muito marcante, contendo cerca de 2-3s e consumo de 15 W. Em vista disso, gerou-se o segundo gabarito para esse conjunto de dados, que para o ES 2 apresentou apenas 3 falsos alarmes. Os mesmos são causados pelo acionamento da televisão, que, como pode ser observado na Figura 6.23, possuem dois estágios durante o seu acionamento, o falso alarme sendo causado justamente durante a primeira crista de consumo. Por outro lado, ao considerar essas perturbações como alvo, observa-se que nem todas são detectadas, o que causa uma queda na taxa de detecção razoável. A Figura 7.7 indica algumas per-

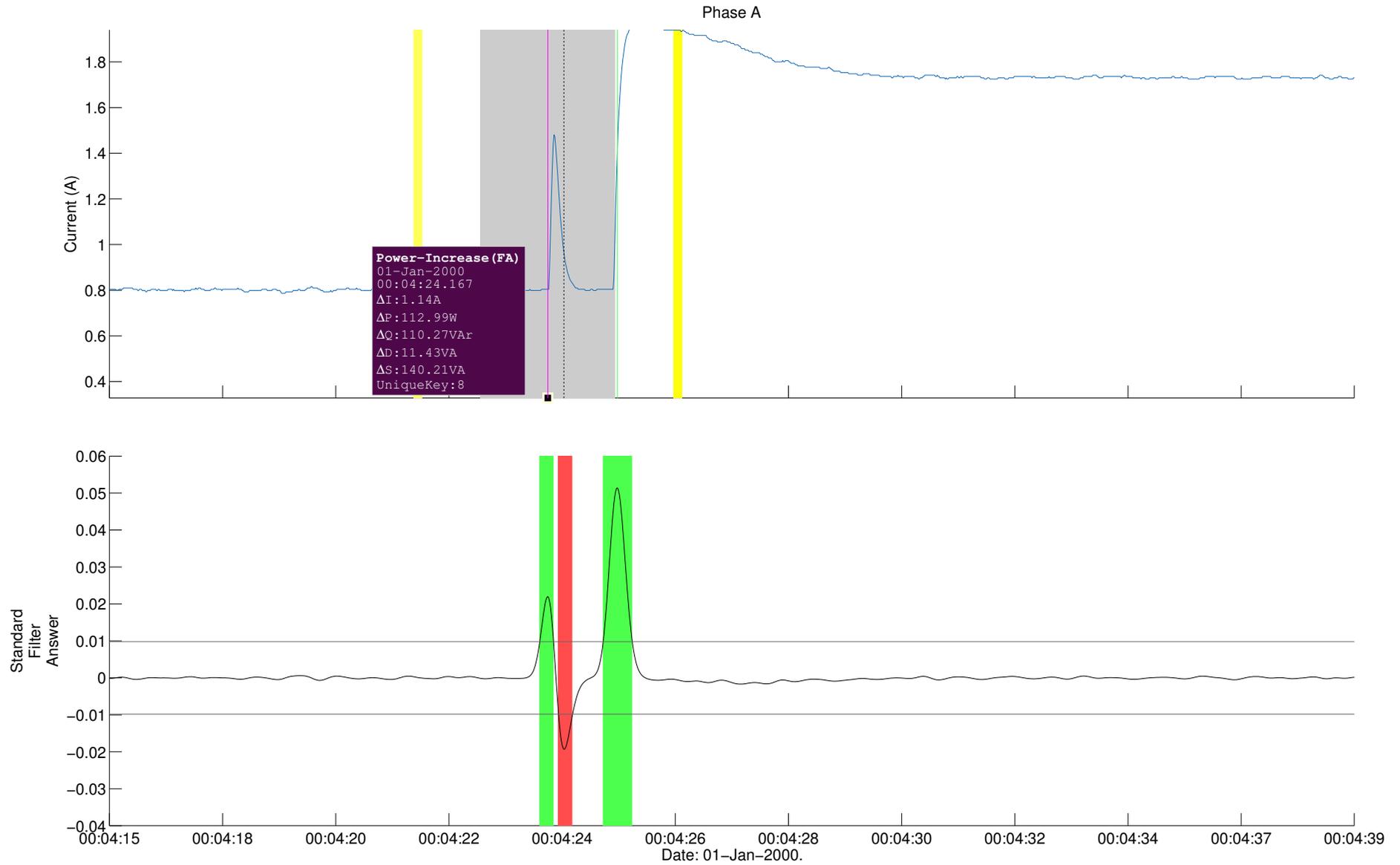


Figura 7.5: Falso alarme para a configuração *ES 1* e *ES 2* no conjunto de dados *NI00173*. O mesmo é causado por uma má manipulação da chave que seleciona a velocidade do ventilador.

das de alvo quando considerando esse distúrbio como de detecção desejável. Assim, esse é um problema ainda a ser resolvido para a aplicação do NILM, uma vez que da maneira que o mesmo está configurado, não é possível obter a detecção de uma taxa razoável de eventos, nem de evitá-los. Porém, o caminho parece ser tratar a ocorrência desses eventos através de sua detecção. A taxa de detecção e falso alarme do ES 2 para os conjuntos *Temporizado Gabarito 1*, *Empilhado4*, *Empilhado7* é de 83,8 % e 16,2 %, respectivamente, enquanto as mesmas taxas são de 80,12 % e 0,9 % quando considerando o *Temporizado Gabarito 2*. Já suas taxa nos conjuntos *NI00**** são de 98,1 % e 0,6 %.

Também se deteve atenção às perdas de alvo no conjunto de dados *Empilhado4*, o conjunto em que há maior dificuldade de detectar eventos. Na Figura 7.8, observa-se que os valores absolutos dos pontos de inflexão, por volta de 0,006, estão na ordem do ruído do arquivo *Temporizado* (ver Figura 7.6), ou seja, para detectar esses eventos é necessário descer o patamar do filtro a um valor que irá causar uma excessiva ocorrência de falso alarme nesse conjunto de dados — caso do *Ajuste CEPTEL*, cujo patamar é de 0,003. A ocorrência de um grande número de falsos alarmes irá acarretar numa perda de aptidão, tornando essas soluções como indivíduos não desejáveis e, portanto, os mesmos serão eliminados durante a evolução do ES.

Outra otimização foi realizada, desta vez alimentando o algoritmo genético com os conjuntos *Temporizado Gabarito 1* e *Empilhado4* para verificar a capacidade de sua generalização para as condições ruidosas. Os resultados estão representados na coluna ES 3. Ao observar os valores obtidos para os conjuntos para os quais essa configuração foi alimentada, percebe-se que houve uma drástica redução do falso alarme para o conjunto *Temporizado Gabarito 1*. Esse fato se deve ao parâmetro utilizado pelo ES para a remoção de eventos próximos, de 495 amostras, como dito anteriormente. O valor desse parâmetro indica que há uma taxa de tempo morto de resposta de aceitação de outro candidato de no mínimo 8,3 s. Como geralmente não há a ocorrência de eventos causados por acionamentos de equipamentos próximos nesses dois conjuntos, isso não se mostrou relevante para o algoritmo genético, que percebeu uma brecha na fronteira desse parâmetro e descobriu uma configuração que reduz em quase 30 ocorrências a taxa de falso alarme para o *Temporizado Gabarito 1* no qual a aptidão de seus indivíduos estava sendo analisada.

Porém, esse alto tempo morto pode causar a perda de alvo, como observado nos conjuntos *NI00168* e *NI00171*, aonde o desacionamento é seguido de um acionamento de outro equipamento em curtos espaços de tempo, como pode ser observado nas figuras 6.1 e 6.4, e está refletido na perda de eficiência nesses conjuntos, especialmente no *NI00168*. Assim, fica evidente a necessidade de análise pós-convergência das configurações resultantes do ES, aonde o valor obtido de convergência representa o melhor valor obtido pelo ES para o conjunto de dado analisado, porém, pode não

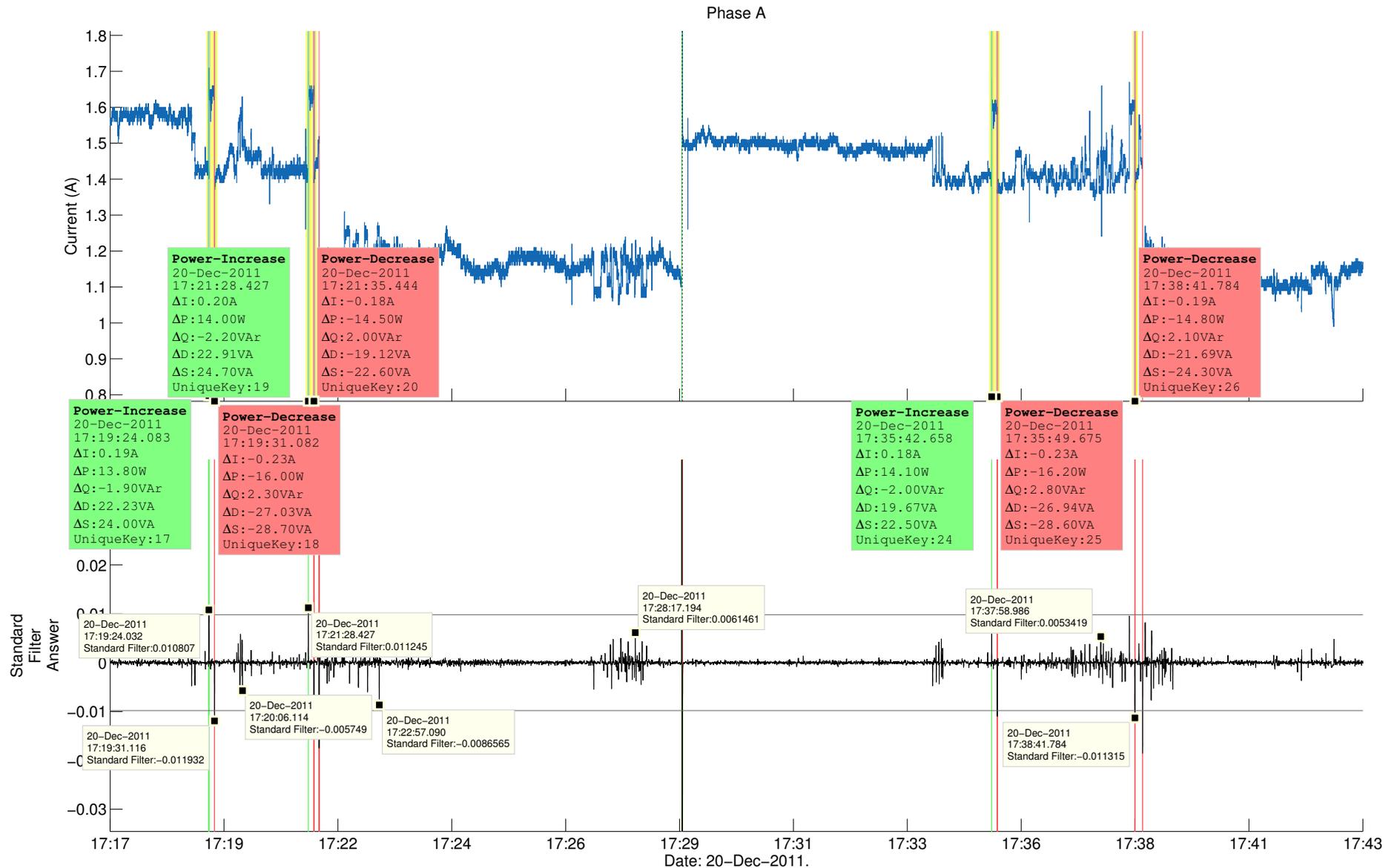


Figura 7.6: Exemplos de falsos alarmes no conjunto de dados *Temporizado Gabarito 1*. Os falsos alarmes são as caixas verdes e vermelhas exibidas. Na subfigura inferior, são mostrados alguns pontos de inflexão e seus valores em unidades da resposta do filtro, bem como valores da resposta do filtro para ruídos.

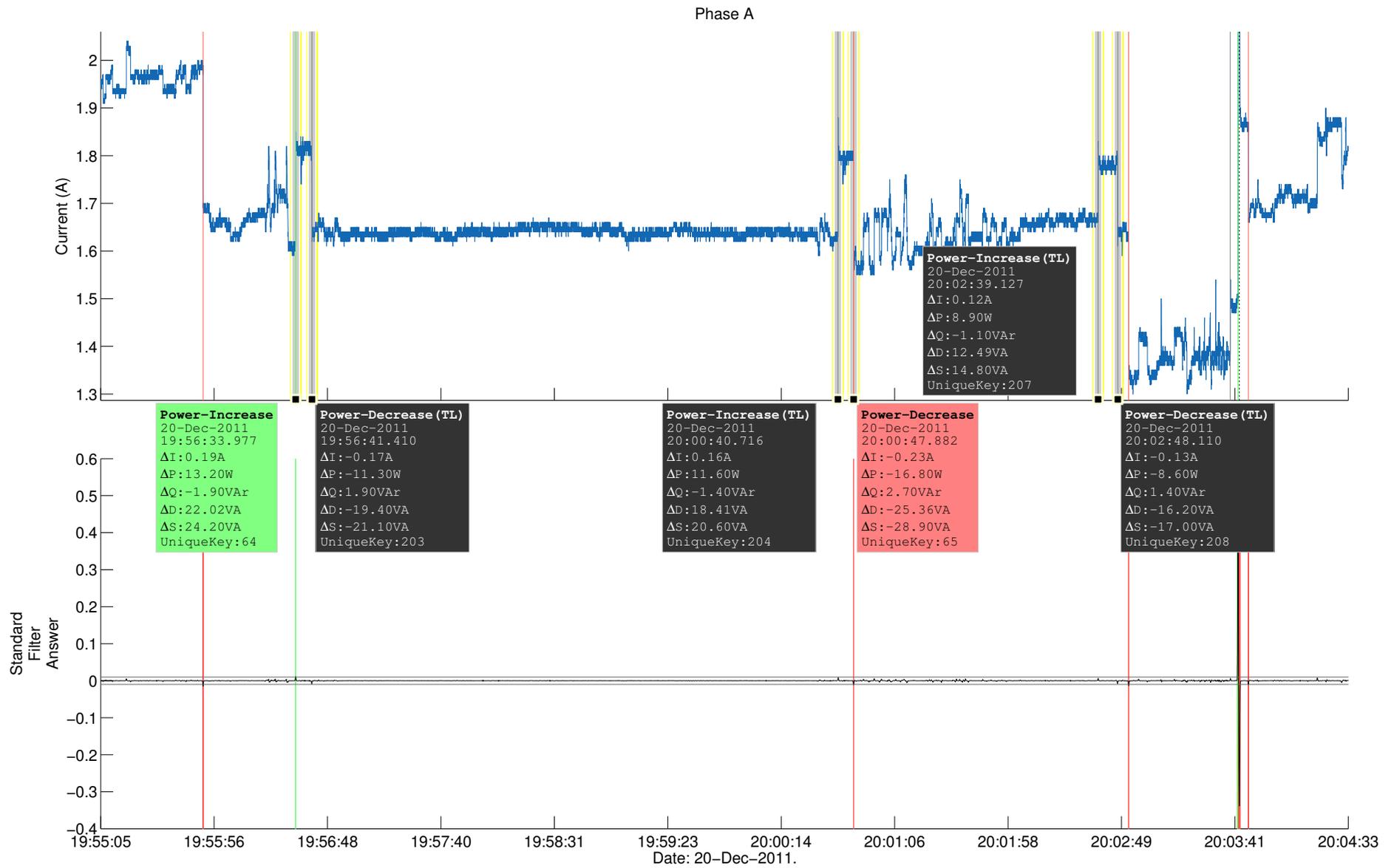


Figura 7.7: Exemplos de perdas de alvos do distúrbio no conjunto de dados *Temporizado Gabarito 2*.

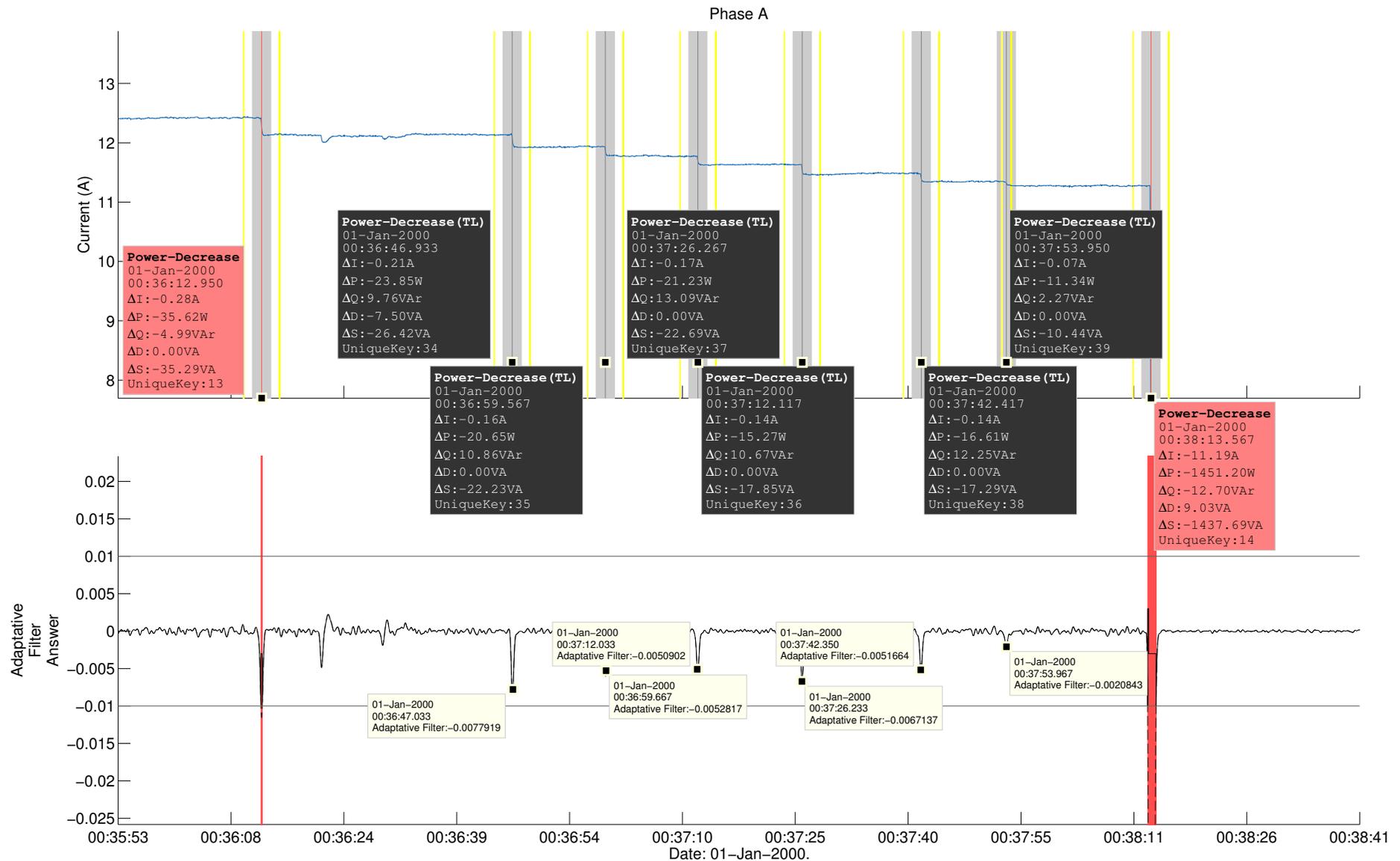


Figura 7.8: Exemplos de perdas de alvo no conjunto de dados *Empilhado4* para a configuração *ES 2*. Os eventos não detectados tem seus pontos de inflexão e seus valores em unidades da resposta do filtro exibidos. As etiquetas em preto indicam os valores de suas características, aonde é possível observar que o evento de maior potencia exibido que constitui perda de alvo apresenta um degrau em valor absoluto de cerca de 24 W.

representar uma configuração aplicável à realidade, como o caso do ES 3 que explorou uma brecha na fronteira do parâmetro n_{min} que permitia a utilização de valores muito elevados.

Além disso, também se empregou a utilização do Detector de Patamar Elaborado desenvolvido por Cañón [102]. Os valores dos parâmetros dessa configuração não foram alterados, de forma que seus valores estão exatamente como ajustado em seu trabalho. Os conjuntos de dados utilizados para o ajuste da técnica por esse autor não continham a atuação de equipamentos simultaneamente e, portanto, os resultados para os conjuntos *Temporizado*, *Empilhado4* e *Empilhado7* não apresentam a real capacidade de operação dessa técnica. Ainda, como foi dito na Seção 4.3.1, essa técnica é limitada apenas à eventos de transitório com acréscimo de consumo. Por isso, foi gerada uma nova tabela (Tabela 7.2) contendo os resultados das demais configurações anteriores apenas para esses eventos e adicionada uma nova coluna com os resultados para tal detector.

Conjunto de Dados		ES 1		Manual		CEPEL		ES 2		ES 3		Det.Elab.	
<i>Eventos acrés. consum.</i>		DET	FA	DET	FA	DET	FA	DET	FA	DET	FA	DET	FA
<i>NI00168</i>	Ocorr.	20	0	19	0	20	0	20	0	11	2	20	1
(20 eventos)	Taxa (%)	100,0	0,0	95,0	0,0	100,0	0,0	100,0	0,0	55,0	10,0	100,0	5,0
<i>NI00171</i>	Ocorr.	26	1	24	0	26	1	24	0	19	2	25	1
(26 eventos)	Taxa (%)	100,0	3,8	92,3	0,0	100,0	3,8	92,3	0,0	73,1	7,7	96,2	3,4
<i>NI00173</i>	Ocorr.	10	1	10	0	10	0	10	1	10	0	10	0
(10 eventos)	Taxa (%)	100,0	10,0	100,0	0,0	100,0	0,0	100,0	10,0	100,0	0,0	100,0	0,0
<i>NI00174</i>	Ocorr.	4	0	4	0	4	0	4	0	3	0	4	1
(4 eventos)	Taxa (%)	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	75,0	0,0	100,0	25,0
<i>NI00175</i>	Ocorr.	12	0	12	0	12	0	12	0	12	0	12	1
(12 eventos)	Taxa (%)	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	8,3
<i>NI00177</i>	Ocorr.	12	0	12	0	12	0	12	0	12	0	12	1
(12 eventos)	Taxa (%)	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	0,0	100,0	8,3
<i>Temp. Gab. 1</i>	Ocorr.	72	29	73	12	71	144	73	21	71	7	71	62
(74 eventos)	Taxa (%)	97,3	39,0	98,6	16,2	95,9	194,6	98,6	28,4	95,9	9,5	95,9	83,8
<i>Temp. Gab. 2</i>	Ocorr.	72	29	85	0	102	116	91	3	77	1	99	34
(105 eventos)	Taxa (%)	97,3	39,2	81,0	0,0	97,1	110,5	86,7	2,9	73,3	0,9	94,3	32,4
<i>Empilhado4</i>	Ocorr.	23	1	18	0	28	8	24	0	22	0	31	1
(40 eventos)	Taxa (%)	57,5	2,5	45,0	0,0	70,0	20,0	60,0	0,0	55,0	0,0	77,5	2,5
<i>Empilhado7</i>	Ocorr.	22	0	22	0	21	16	22	0	18	2	23	8
(24 eventos)	Taxa (%)	91,7	0,0	91,7	0,0	87,5	66,7	91,7	0,0	75,0	8,3	95,8	33,3

Tabela 7.2: Taxa de detecção de eventos de transitório e falso alarme para os três ajustes automáticos, os dois ajustes manuais e o Detector de Patamar Elaborado. Apenas os eventos com acréscimo de consumo foram considerados devido à limitação do Detector de Patamar Elaborado em detectar apenas esse tipo de eventos. As abreviaturas DET e FA referem-se a taxa de detecção e falso alarme, respectivamente.

Observa-se que o Detector de Patamar Elaborado, mesmo não ajustado para os conjuntos de dados analisados, comportou-se razoavelmente bem nos mesmos. Essa técnica obteve boa taxa de detecção, inclusive obtendo a taxa mais alta no conjunto *Empilhado4* e não detectando apenas 6 eventos no conjunto *Temporizado Gabarito 2*. A taxa de falso alarme no último conjunto citado foi de 34 eventos, próxima àquela obtida pelo ES 1, indicando o potencial dessa técnica. Porém, o conjunto *Empilhado7*

contém uma maior taxa de falsos alarmes para essa técnica, que ocorrem durante a operação da televisão CRT. Além disso, há a presença de alguns falsos alarmes nos conjuntos *NI00**** que não ocorrem na método com o filtro de Gaussiana. Esses falsos alarmes foram levantados e identificados pela operação das chaves do ventilador e secador de cabelo, que são do tipo “interrompa e então faça” (do inglês *break then make*), aonde há a interrupção do consumo do equipamento quando alterando seu estado de operação. Um exemplo de distúrbio que causa falso alarme de um acionamento na operação dessa chave pode ser observado na Figura 6.10, onde o segundo desacionamento nesse conjunto causa um evento de acionamento pelo Detector de Patamar Elaborado. Talvez seja possível trabalhar este método de forma a eliminar esse problema, que aliado com o ajuste de seus parâmetros, pode se tornar uma técnica bastante adequada para a detecção de eventos. O ajuste de parâmetros pode ser feito de maneira automática, utilizando uma metodologia similar àquela empregada neste trabalho.

Finalmente, como foi dito no Capítulo 4, é necessário indicar a capacidade de reconstrução de energia do NILM, porém, o mesmo não faria sentido nas simulações em laboratório, que não apresentam a operação dos aparelhos nos tempos que normalmente ocorrem nas residências. Indo além, os dados de residência disponíveis (ver Figura 5.2) não tinham a informação necessária para a construção do gabarito. Assim, as taxas aqui disponibilizadas não revelam a real capacidade de reconstrução de energia nas residências brasileiras, mas servem como um estudo inicial nesse sentido. Espera-se nos estudos posteriores empregar a análise em dados representativos, de forma a trazer as taxas já calculadas em termos de energia e capacidade de desagregação de consumo do NILM.

Capítulo 8

Conclusão

Este trabalho trouxe um extenso levantamento bibliográfico, desde um detalhamento das aplicações do NILM para a obtenção de uma melhor EE, às diversas técnicas aplicadas por outros autores em busca da obtenção de uma estimativa do consumo desagregado. A primeira parte do levantamento, em relação a aplicação do NILM para EE, detalhou um novo programa de EE proposto em países desenvolvidos como EUA e outros países na Europa Ocidental que permite a redução do consumo de energia elétrica no setor residencial através do retorno da informação (estimada) de consumo, aonde o NILM pode ter papel fundamental de tornar essa informação disponível com baixo custo. A intenção desses autores é tentativa da união da infraestrutura que será disponibilizada pelas redes elétricas inteligentes (*smart grid*) como sistema de aquisição de dados — os medidores inteligentes (*smart meters*). Além disso, o mesmo pode auxiliar na obtenção de uma informação com melhor qualidade e maior frequência para os estudos em EE realizados por entidades governamentais, aonde o mesmo pode auxiliar na informação obtida através das Pesquisa de Posse e Hábito de Eletrodomésticos (PPH) ou, até mesmo, futuramente substituí-la.

O levantamento das técnicas permitiu a realização de diversas considerações sobre os pontos envolvidos no projeto dessa tecnologia, aonde essa informação foi organizada e debatida na primeira seção Capítulo 4. Isso permitiu uma melhor compreensão do projeto como um todo, o que levou ao desenvolvimento de um ambiente para atender às especificações observadas. Dentre elas, destaca-se a necessidade de um ambiente único de análise para a evolução do projeto, que pode utilizar diversas técnicas e características diferentes durante o seu progresso, com o intuito de melhorar a capacidade de reconstrução de energia do NILM, que inclusive pode alterar a configuração de seu sistema de aquisição de dados e, portanto, precisa ser capaz de representar os dados e análises em uma única maneira. Esse ambiente permite analisar o comportamento do NILM antes que o mesmo seja colocado em operação.

A implementação desse ambiente adaptou a metodologia proposta pelo CE-

PEL/Eletrabras para detecção de eventos de transitório, passo necessário para a exploração da informação deixada por equipamentos durante a alteração de seu estado operativo e que foi atacado pelo grupo do CEPTEL/Eletrabras. A sua metodologia emprega um filtro de derivada de Gaussiana para a geração de candidatos, que ainda passam por outros dois cortes antes de serem aceitos como um evento de transitório: um para o valor mínimo de degrau de corrente que, caso seja demasiadamente pequeno, será removido por constituir um evento de ruído; e outro para candidatos próximos, adicionando um tempo morto de resposta do filtro antes que um novo candidato possa ser aceito.

Na versão da metodologia presente no ambiente de análise, adicionaram-se diversos avanços operacionais, alguns exemplos: leitura de redes elétricas com até três fases, transformação dos dados em um formato único para os dois medidores sendo utilizados; segmentação da informação em memória para permitir a leitura de dados longos; redução da necessidade de leitura em disco através do armazenamento das bordas dos arquivos em memória transitória; robustez indicando capacidade de recuperação de dados com falhas no medidor ou sua sobrecarga; aumento da velocidade da análise através da remoção de amostras irrelevantes na janela do filtro em convolução; melhor capacidade de interpretação gráfica dos dados e das análises; etc.

Porém, as maiores contribuições foram:

- Construção do gabarito: uma das dificuldades para o NILM é obter uma estimativa do consumo desagregado por aparelho a ser identificado pelas técnicas aplicadas no NILM. O ambiente tratou desse caso fornecendo uma interface gráfica para o usuário, que permitiu a criação da informação estimada a ser desagregada nos dados. Após completar o seu preenchimento para um determinado conjunto de dados, esse ambiente facilita a compreensão da informação contida no mesmo através de informações gráficas;
- Remoção de eventos por inconsistência: a mesma mostrou-se suficiente nos conjuntos de dados analisados para a remoção de candidatos a eventos de transitório antes removidos pela adição de tempo morto em resposta do detector pela metodologia do CEPTEL/Eletrabras, esses eventos causados por picos no consumo durante acionamento de equipamentos — como aqueles contendo motores. A inconsistência ocorre quando a resposta do filtro tem sinal diferente àquele presente no degrau (diferença entre pós e pré transitório) de corrente extraído durante a geração do candidato à evento;
- Capacidade de ajuste automático dos parâmetros: realizado através de uma adaptação do Algoritmo Genético de Estratégia Evolutiva inspirada nas

versões multiespécies de otimização multiobjetivo, porém para permitir uma dinâmica durante a evolução das espécies que possibilita separar maior capacidade computacional às espécies que melhor se adaptam para a solução do problema. O otimizador também apresenta outras capacidades:

- Determinação do ponto de operação: através do ajuste dos parâmetros livres para o cálculo da aptidão, é possível priorizar maiores taxas de detecção ou falso alarme, porém essa capacidade não pôde ser avaliada neste trabalho;
- Recuperar o processo caso ocorra interrupção não planejada: o algoritmo genético implementado é capaz de recuperar a evolução de um processo perdido, o que é especialmente importante durante a otimização de conjuntos com grande quantidade de dados.

Três ajustes utilizando o otimizador foram realizados:

- ES 1: antes de perseguir a aplicação do otimizador para os dados simulando cenários de aplicação prática, aplicou-se a otimização para os dados com condições mais simples — os conjuntos *NI00****. O otimizador apresentou menor falso alarme nos dados simulando cenários de aplicação prática. Atribuiu-se essa capacidade ao parâmetro γ_{rem} utilizado para o cálculo da aptidão dos indivíduos. Seu emprego permitiu explorar os limites do limiar de sensibilização para a geração de candidatos a eventos de transitório, o que reduziu a presença de falsos alarmes nos conjuntos representando cenários de aplicação reais, mesmo que eles não tenham sido vistos pelo otimizador;
- ES 2: como o objetivo do trabalho foi a extensão da técnica para condições mais próximas àquelas presentes nas redes elétricas residenciais — o ajuste de parâmetros realizado pelo CEPEL/Eletrobras foi realizado para dados com pouca presença de sinal ruído e operação de poucos (no máximo dois) equipamentos simultaneamente —, aplicou-se o ajuste automático para essas condições. A redução do falso alarme para o conjunto *Temporizado*, o conjunto com maior incidência de falsos alarmes, não correspondeu às expectativas, o que durante a investigação causou a descoberta de um distúrbio causado durante a operação da televisão LCD presente em tal conjunto. Ao considerar esse distúrbio como alvo, percebeu-se uma redução na taxa de detecção, o que mostra que nem todos eventos desse distúrbio são identificados pela metodologia. Assim, esse distúrbio ainda precisa ser melhor compreendido antes da aplicação do NILM;
- ES 3: para analisar a capacidade de generalização, ao invés de empregar os três conjuntos de dados simulando cenários de aplicação prática, realizou-se

a otimização com apenas dois desses conjuntos. Essa configuração permitiu identificar uma falha na fronteira determinada para o parâmetro responsável pelo tempo morto de resposta que permitia um limite superior de até 8,3 s aonde todos os candidatos seriam removidos caso um evento de transitório fosse confirmado. Apesar desse valor não ser justificável, para os dois conjuntos de dados analisados esse valor se constituía de uma vantagem por permitir a redução do falso alarme em detrimento de pouca perda na detecção, e o algoritmo genético explorou essa falha ao percebê-la. Assim, ainda que o ajuste seja realizado automaticamente, é necessário a realização de crítica pós-convergência para identificação de possíveis problemas como esse.

As taxas de detecção de falso alarme para o ajuste realizado pelo CEPEL/Eletrbras eram respectivamente de 100 % e 1,2 % para os dados *NI00****, enquanto esses valores foram de 90,2 % e 82,9 % para os dados com cenários de aplicação prática. A configuração ES 1 alterou os valores para os cenários de aplicação prática para 83,18 e 3,0 %, bem mais próximos a necessidade de aplicação do NILM, enquanto seus valores para os conjuntos de dados *NI00**** permaneceram praticamente constantes, em 98,8 % e 1,2 %. Os valores para os dados *NI00**** também permaneceram na mesma faixa para o ES 2, porém houve sua taxa para os dados com cenários de aplicação prática foram 80,1 % e 0,9 %, reduzindo o falso alarme em detrimento de uma menor taxa de detecção em comparação à configuração anterior. Portanto, além de todas facilidades e capacidades já referidas, o ambiente trouxe sistematização do processo de determinação dos parâmetros, que podem ser ajustados automaticamente para novas condições ainda não conhecidas presentes nas redes elétricas residenciais.

8.1 Trabalhos Futuros

O trabalho tratou apenas de uma das etapas necessárias para o desenvolvimento da tecnologia, o próximo passo é tratar do tema de discriminação e identificação do estado operativo dos aparelhos. Foram realizadas considerações sobre o tema das escolhas das características neste trabalho, que pode auxiliar próximos trabalhos em determinar quais delas serão utilizadas. Deseja-se aqui ressaltar a característica que se sobressaiu no levantamento bibliográfico, a estatística de uso. Apesar de não ser possível realizar a discriminação apenas utilizando a envoltória como característica única, uma vez que ela só é discriminante para mudanças de estados que causam acréscimo no consumo, a mesma mostrou-se discriminante nos estudos anteriores em colaboração com o CEPEL/Eletrbras e podem ser empregadas para auxiliar na discriminação. Propõe-se que a implementação seja realizada no ambiente de

análise pela questão de unificação e aproveitamento das ferramentas, além de toda a capacidade já oferecida por esse ambiente.

Outra questão importante é a utilização das medidas que levam em sua consideração a reconstrução em energia para a determinação da eficácia do NILM, bem como de conjuntos de dados reais. Também quanto a questão dos dados, é interessante a utilização dos conjuntos REDD e BLUED para possibilitar a técnica ser comparada em nível internacional. As medidas em termos de energia podem ser adaptadas também para o caso de detecção de eventos, que devem ser utilizadas com a evolução do trabalho.

A otimização automática de parâmetros também tem pontos que precisam ser trabalhados. Primeiro, a versão do ES implementada ainda não está completa, sendo necessário adicionar um método que permita a competição dos indivíduos somente após atingir um estágio de pré-evolução. Ainda, é necessário avaliar a capacidade do ES de determinar o ponto de operação através do ajuste dos parâmetros livres (γ_{det} e γ_{fa}) no cálculo da aptidão. Por fim, outros algoritmos podem ser analisados, como Inteligência de Enxames.

Quanto a técnica para detecção de eventos, podem ser aplicados outros métodos além do filtro de núcleo de Gaussiana. Mesmo sem alterar a ideia de convolução de um filtro, pode-se aplicar outra com uma função simétrica no sinal para a detecção de eventos. Além disso, durante este trabalho também se avaliou a capacidade da técnica proposta pelo estudo anterior realizado pela colaboração COPPE/UFRJ-CEPEL/Eletrabras, o Detector de Patamar Elaborado. Percebeu-se que o mesmo apresenta um potencial para ser estendido para os novos cenários, sendo necessário realizar o ajuste de seus parâmetros. Nesse caso, também será necessário a sua adaptação para não causar falsos alarmes durante a atuação de chaves *break then make*, presentes em alguns equipamentos como ventiladores e secadores de cabelo. Outra possibilidade seria a aplicação da Transformada *Wavelet*, que como observado durante o levantamento bibliográfico, tem aplicação bem-sucedida para a detecção de transitórios em redes elétricas. Ainda, o patamar não precisa ser determinado como um valor fixo, é possível estender essa metodologia para a utilização de patamares adaptativos, tendo um valor de sensibilização conforme o nível de dinâmica presente na rede elétrica.

Referências Bibliográficas

- [1] RIPPEL, R., RIPPEL, V. C. L., DE LIMA, J. F. “Percepções Genéricas sobre o crescimento populacional, a demanda por energia e os padrões de consumo dos recursos atuais”, set. 2006. Disponível em: <http://www.abep.nepo.unicamp.br/encontro2006/docspdf/ABEP2006_450.pdf>. Acessado em: 1 de fev. 2013.
- [2] IEA. *Key World Energy Statistics*. Paris, France, 2012. ISBN: 9789264086241. Disponível em: <<http://www.iea.org/publications/freepublications/publication/kwes.pdf>>. Acessado em: 26 de fev. 2013.
- [3] JATOBÁ, S. U. S., CIDADE, L. C. F., VARGAS, G. M. “Ecologismo, ambientalismo e ecologia política: diferentes visões da sustentabilidade e do território”, *Sociedade e Estado*, v. 24, pp. 47–87, 04 2009. ISSN: 0102-6992. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-69922009000100004&nrm=iso>. Acessado em: 29 de jan. de 2013.
- [4] EPE. *Avaliação da Eficiência Energética para os próximos 10 anos: 2012 - 2021*. Nota técnica, MME - Ministério de Minas e Energia, Rio de Janeiro, RJ, 2012. Disponível em: <http://www.epe.gov.br/mercado/Documents/S%C3%A9rie%20Estudos%20de%20Energia/20121221_1.pdf>. Acessado em: 6 de fev. 2013.
- [5] ONU. *Realizing the Potential of Energy Efficiency: Targets, Policies, and Measures for G8 Countries*. Relatório técnico, Expert Group on Energy Efficiency, Washington, D.C., USA, 2007. Disponível em: <http://www.globalproblems-globalsolutions-files.org/unf_website/PDF/realizing_potential_energy_efficiency.pdf>. Acessado em: 4 de fev. 2013.
- [6] DOS SANTOS, A. H. C. *Eficiência energética e a contribuição dos gases combustíveis: análise de caso das políticas de avaliação de edificações*. Tese

- de Mestrado, Universidade de São Paulo, São Paulo, SP, 2011. Disponível em: <<http://www.iee.usp.br/biblioteca/producao/2011/Teses/Disserta%E7%E3o%20Mestrado%20-%20Arthur%20Cursino.pdf>>. Acessado em: 1 de fev. 2013.
- [7] EPE. “Eficiência Energética”. Plano Nacional de Energia 2030, Rio de Janeiro, RJ, 2007. Disponível em: <http://www.epe.gov.br/PNE/20080512_11.pdf>. Acessado em: 7 de fev. 2013.
- [8] MME. *Plano Nacional de Eficiência Energética: Premissas e Diretrizes Básicas*. Brasília, DF, 2011. Disponível em: <www.mme.gov.br/mme/galerias/arquivos/PlanoNacEfiEnergetica.pdf>. Acessado em: 8 de fev. 2013.
- [9] IEA. *World Energy Outlook 2010*. Paris, France, 2010. ISBN: 9789264086241. Disponível em: <<http://www.iea.org/publications/freepublications/publication/weo2010.pdf>>. Acessado em: 26 de fev. 2013.
- [10] JANNUZZI, G. M. “Aumentando a Eficiência nos Usos Finais de Energia no Brasil”, 2002. Disponível em: <http://www.cgu.rei.unicamp.br/energia2020/papers/paper_Jannuzzi.pdf>. Acessado em: 23 de jan. 2013.
- [11] EPE. “Eficiência Energética”. Apresentado como parte do Seminário: O Plano Nacional de Energia 2030, 2006. Disponível em: <http://www.epe.gov.br/PNE/20070625_7.pdf>. Acessado em: 1 de fev. 2013.
- [12] GOLDEMBERG, J., LUCON, O. “Energia e meio ambiente no Brasil”, *ESTUDOS AVANÇADOS*, v. 21, n. 59, pp. 7, 2007. Disponível em: <<http://www.scielo.br/pdf/ea/v21n59/a02v2159.pdf>>. Acessado em: 27 de jan. de 2013.
- [13] EPE. *Projeção da demanda de energia elétrica para os próximos 10 anos: 2013 - 2022*. Relatório técnico, MME - Ministério de Minas e Energia, Rio de Janeiro, RJ, 2012. Disponível em: <http://www.epe.gov.br/mercado/Documents/S%C3%A9rie%20Estudos%20de%20Energia/20130117_1.pdf>. Acessado em: 6 de fev. 2013.
- [14] “Cartilha do Programa Brasileiro de Etiquetagem (PBE)”. Disponível em: <http://www2.inmetro.gov.br/pbe/pdf/folder_pbe.pdf>. Acessado em: 7 de fev. 2013.

- [15] “Etiqueta de eficiência energética do Programa Brasileiro de Etiquetagem (PBE)”. . Disponível em: <http://www2.inmetro.gov.br/pbe/a_etiqueta.php>. Acessado em: 7 de fev. 2013.
- [16] EPE. *Plano Decenal de Expansão de Energia 2020*. Relatório técnico, MME - Ministério de Minas e Energia, Rio de Janeiro, RJ, 2011. Disponível em: <http://www.epe.gov.br/PDEE/20120302_1.pdf>. Acessado em: 9 de fev. 2013.
- [17] MME - MINISTÉRIO DE MINAS E ENERGIA. *Balanço de Energia Útil - BEU 2005*. Brasília, DF, 2005.
- [18] EPE. *Balanço Energético Nacional 2012, Ano base 2011*. Relatório final, MME - Ministério de Minas e Energia, Rio de Janeiro, RJ, 2012. Disponível em: <<https://ben.epe.gov.br/BENRelatorioFinal2012.aspx>>. Acessado em: 1 de fev. 2013.
- [19] PROCEL/ELETROBRÁS. *Avaliação dos Resultados PROCEL 2005 - Ano Base 2005*. Classe residencial - relatório brasil, 2006. Disponível em: <<http://www.procelinfo.com.br/main.asp?View=%7B5A08CAF0-06D1-4FFE-B335-95D83F8DFB98%7D&Team=¶ms=itemID=%7BA1BA3112-38B0-46DB-81DF-38286C005960%7D;&UIPartUID=%7B05734935-6950-4E3F-A182-629352E9EB18%7D>>. Acessado em: 13 de fev. 2013.
- [20] “Pesquisa de Posse de Equipamentos e Hábitos de Consumo de Energia”. Disponível em: <<http://www.elektrobras.com/pci/main.asp?View=%7B05070313-120A-45FD-964D-5641D6083F80%7D>>. Acessado em: 13 de fev. 2013.
- [21] PROCEL/ELETROBRAS. *Resultados PROCEL 2012, Ano Base 2011*. Relatório completo, Rio de Janeiro, RJ, 2012. Disponível em: <<http://www.procelinfo.com.br/main.asp?View=%7B5A08CAF0-06D1-4FFE-B335-95D83F8DFB98%7D&Team=¶ms=itemID=%7BA1516CFD-2424-4E17-BA99-912E47D91F9D%7D;&UIPartUID=%7B05734935-6950-4E3F-A182-629352E9EB18%7D>>. Acessado em: 16 de mar. 2013.
- [22] EHRHARDT-MARTINEZ, K., DONNELLY, K. A., LAITNER, J. A. S. *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities*. Research report, ACEEE - American Council for an Energy-Efficient Economy,

Washington, D.C., USA, 2010. Disponível em: <<http://aceee.org/research-report/e105>>. Acessado em: 11 de mar. 2013.

- [23] EPRI. *Residential Electricity Use Feedback: A Research Synthesis and Economic Framework*. Final report 1016844, fev. 2009. Disponível em: <<http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001016844>>. Acessado em: 24 de mar. 2013.
- [24] SCHLEICH, J., KLOBASA, M., GOLZ, S. “Does smart metering reduce residential electricity demand?” In: *European Energy Market (EEM), 2012 9th International Conference on the*, pp. 1–4, 2012. doi: 10.1109/EEM.2012.6254779. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6254779>. Acessado em: 24 de mar. 2013.
- [25] QIU, Z., DECONINCK, G. “Smart Meter’s feedback and the potential for energy savings in household sector: A survey”. In: *Networking, Sensing and Control (ICNSC), 2011 IEEE International Conference on*, pp. 281–286, 2011. doi: 10.1109/ICNSC.2011.5874882. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5874882>. Acessado em: 24 de mar. 2013.
- [26] DARBY, S. *The effectiveness of feedback on energy consumption: a review for DEFRA of the literature on metering, billing and direct displays*. Relatório técnico, Environmental Change Institute, University of Oxford, 2006. Disponível em: <<http://www.eci.ox.ac.uk/research/energy/electric-metering.php>>. Acessado em: 24 de mar. 2013.
- [27] AYRES, I., RASEMAN, S., SHIH, A. *Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage*. Working Paper 15386, National Bureau of Economic Research, September 2009. Disponível em: <<http://www.nber.org/papers/w15386>>. Acessado em: 6 de mar. 2013.
- [28] DELMAS, M. A., FISCHLEIN, M., ASENSIO, O. I. “Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012”, *Energy Policy*, v. 61, n. 0, pp. 729 – 739, 2013. ISSN: 0301-4215. doi: <http://dx.doi.org/10.1016/j.enpol.2013.05.109>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0301421513004643>>.
- [29] BERGES, M. E., GOLDMAN, E., MATTHEWS, H. S., et al. “Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring”, *Journal of Industrial Ecology*, v. 14, n. 5, pp. 844–858, 2010.

ISSN: 1530-9290. doi: 10.1111/j.1530-9290.2010.00280.x. Disponível em: <<http://dx.doi.org/10.1111/j.1530-9290.2010.00280.x>>.

- [30] HE, H. A., GREENBERG, S., HUANG, E. M. “One size does not fit all: applying the transtheoretical model to energy feedback technology design”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 927–936, New York, NY, USA, 2010. ACM. ISBN: 978-1-60558-929-9. doi: 10.1145/1753326.1753464. Disponível em: <<http://doi.acm.org/10.1145/1753326.1753464>>. Acessado em: 6 de mar. 2013.
- [31] DARBY, S. “Making it obvious: Designing feedback into energy consumption”. In: *Proceedings of the 2nd International Conference on Energy Efficiency in Household Appliances and Lighting*, 2000.
- [32] CAIRES, L. E. *Aplicação de redes inteligentes nas instalações elétricas residenciais*. Tese de Mestrado, Universidade de São Paulo, São Paulo, SP, 2012. Disponível em: <http://www.iee.usp.br/biblioteca/producao/2012/Teses/Dissert_CairesCorrigida.pdf>. Acessado em: 1 de fev. 2013.
- [33] ANEEL. “REN502 - Regulamenta Sistemas de Medição de Energia Elétrica de Unidades Consumidoras do Grupo B”. 2012.
- [34] “Portal da Empresa Positive Energy”. Disponível em: <<http://opower.com/>>. Acessado em: 29 de mar. 2013.
- [35] “Portal da Empresa C3 Energy”. Disponível em: <<http://www.c3energy.com/>>. Acessado em: 29 de mar. 2013.
- [36] “Portal do medidor de energia TED”. Disponível em: <<http://www.theenergydetective.com/>>. Acessado em: 1 de abr. 2013.
- [37] “Portal do *display* de energia Wattson”. . Disponível em: <<http://www.diykyoto.com/uk/>>. Acessado em: 1 de abr. 2013.
- [38] “Portal do medidor de energia PowerCost Monitor”. Disponível em: <<http://www.bluelineinnovations.com/>>. Acessado em: 1 de abr. 2013.
- [39] “Portal do *display* de energia da Efergy”. Disponível em: <<http://www.efergy.com/index.php/default/>>. Acessado em: 1 de abr. 2013.
- [40] STÅHLBRÖST, A., LASSINANTTI, J., YLINIEMI, K., et al. “Energy Savings by User Interaction and Visualisation”. In: *Information*

Systems Research Seminar in Scandinavia, n. 35, 2012. Disponível em: <<http://pure.ltu.se/portal/en/publications/energy-savings-by-user-interaction-and-visualisation%28b20bf37d-9c86-4cfb-8091-c4d5ae7a97af%29.html>>. Acessado em: 6 de mar. 2013.

- [41] PETTERSSON, R. *It Depends: Principle and Guidelines*. 2012. ISBN: 978-91-85334-24-7.
- [42] LIIKKANEN, L. “Extreme-user approach and the design of energy feedback systems”. In: *International Conference on Energy Efficiency in Domestic Appliances and Lighting*, pp. 16–18, 2009.
- [43] ZEIFMAN, M., ROTH, K. “Nonintrusive appliance load monitoring: Review and outlook”, *Consumer Electronics, IEEE Transactions on*, v. 57, n. 1, pp. 76–84, 2011. ISSN: 0098-3063. doi: 10.1109/TCE.2011.5735484.
- [44] “Portal da Primeira Seminário Internacional em Monitoramento Não-Invasivo de Cargas Elétricas”. Disponível em: <<http://www.ices.cmu.edu/psii/nilm/>>. Acessado em: 3 de nov. 2013.
- [45] LIPOFF, S. “Home Automation – the Unrealized Promise”. In: *IEEE Consumer Electronics Society Newsletter*, pp. 13–14, 2010.
- [46] “Portal da empresa GetEmme”. Disponível em: <<http://www.getemme.com>>. Acessado em: 5 de nov. 2013.
- [47] “Portal da empresa Navetas”. Disponível em: <<http://www.navetas.com>>. Acessado em: 5 de nov. 2013.
- [48] HART, G. “Nonintrusive appliance load monitoring”, *Proceedings of the IEEE*, v. 80, n. 12, pp. 1870–1891, 1992. ISSN: 0018-9219. doi: 10.1109/5.192069.
- [49] BOULOUTAS, A., HART, G., SCHWARTZ, M. “Two extensions of the Viterbi algorithm”, *Information Theory, IEEE Transactions on*, v. 37, n. 2, pp. 430–436, 1991. ISSN: 0018-9448. doi: 10.1109/18.75270.
- [50] HART, G., BOULOUTAS, A. “Correcting dependent errors in sequences generated by finite-state processes”, *Information Theory, IEEE Transactions on*, v. 39, n. 4, pp. 1249–1260, 1993. ISSN: 0018-9448. doi: 10.1109/18.243442.
- [51] NORFORD, L. K., LEEB, S. B. “Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection

- algorithms”, *Energy and Buildings*, v. 24, n. 1, pp. 51 – 64, 1996. ISSN: 0378-7788. doi: [http://dx.doi.org/10.1016/0378-7788\(95\)00958-2](http://dx.doi.org/10.1016/0378-7788(95)00958-2). Disponível em: <http://www.sciencedirect.com/science/article/pii/S0378778895009582>.
- [52] LEEB, S., SHAW, S., KIRTLEY, J.L., J. “Transient event detection in spectral envelope estimates for nonintrusive load monitoring”, *Power Delivery, IEEE Transactions on*, v. 10, n. 3, pp. 1200–1210, 1995. ISSN: 0885-8977. doi: 10.1109/61.400897.
- [53] COLE, A., ALBICKI, A. “Data extraction for effective non-intrusive identification of residential power loads”. In: *Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings. IEEE*, v. 2, pp. 812–815 vol.2, 1998. doi: 10.1109/IMTC.1998.676838.
- [54] COLE, A., ALBICKI, A. “Algorithm for nonintrusive identification of residential appliances”. In: *Circuits and Systems, 1998. ISCAS '98. Proceedings of the 1998 IEEE International Symposium on*, v. 3, pp. 338–341 vol.3, 1998. doi: 10.1109/ISCAS.1998.704019.
- [55] POWERS, J., MARGOSSIAN, B., SMITH, B. “Using a rule-based algorithm to disaggregate end-use load profiles from premise-level data”, *Computer Applications in Power, IEEE*, v. 4, n. 2, pp. 42–47, 1991. ISSN: 0895-0156. doi: 10.1109/67.75875.
- [56] FARINACCIO, L., ZMEUREANU, R. “Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses”, *Energy and Buildings*, v. 30, n. 3, pp. 245 – 259, 1999. ISSN: 0378-7788. doi: [http://dx.doi.org/10.1016/S0378-7788\(99\)00007-9](http://dx.doi.org/10.1016/S0378-7788(99)00007-9). Disponível em: <http://www.sciencedirect.com/science/article/pii/S0378778899000079>.
- [57] MARCEAU, M. L. “Nonintrusive load disaggregation computer program to estimate the energy consumption of major end-uses in residential buildings”, 1999. Disponível em: <http://spectrum.library.concordia.ca/816/>.
- [58] BARANSKI, M., VOSS, J. “Nonintrusive appliance load monitoring based on an optical sensor”. In: *Power Tech Conference Proceedings, 2003 IEEE Bologna*, v. 4, pp. 8 pp. Vol.4–, 2003. doi: 10.1109/PTC.2003.1304732.
- [59] BARANSKI, M., VOSS, J. “Genetic algorithm for pattern detection in NIALM systems”. In: *Systems, Man and Cybernetics, 2004 IEEE International*

Conference on, v. 4, pp. 3462–3468 vol.4, 2004. doi: 10.1109/ICSMC.2004.1400878.

- [60] BARANSKI, M., VOSS, J. “Detecting Patterns of Appliances from Total Load Data Using a Dynamic Programming Approach”, *Data Mining, IEEE International Conference on*, v. 0, pp. 327–330, 2004. doi: <http://doi.ieeecomputersociety.org/10.1109/ICDM.2004.10003>.
- [61] MATTHEWS, H. S., SOIBELMAN, L., BERGES, M., et al. “Automatically Disaggregating the Total Electrical Load in Residential Buildings: a Profile of the Required Solution”, *Intelligent Computing in Engineering - ICE08*, p. 381, 2008.
- [62] LAUGHMAN, C., LEE, K., COX, R., et al. “Power signature analysis”, *Power and Energy Magazine, IEEE*, v. 1, n. 2, pp. 56–63, 2003. ISSN: 1540-7977. doi: 10.1109/MPAE.2003.1192027.
- [63] LEE, K. D., LEEB, S. B., NORFORD, L. K., et al. “Estimation of Variable-Speed-Drive Power Consumption From Harmonic Content”, *IEEE Transactions on Energy Conversion*, v. 20, pp. 566–574, 2005. doi: 10.1109/TEC.2005.852963.
- [64] WICHAKOOL, W., AVESTRUZ, A.-T., COX, R. W., et al. “Modeling and Estimating Current Harmonics of Variable Electronic Loads”, *IEEE Transactions on Power Electronics*, v. 24, pp. 2803–2811, 2009. doi: 10.1109/TPEL.2009.2029231.
- [65] SHAW, S., LEEB, S., NORFORD, L., et al. “Nonintrusive Load Monitoring and Diagnostics in Power Systems”, *Instrumentation and Measurement, IEEE Transactions on*, v. 57, n. 7, pp. 1445–1454, 2008. ISSN: 0018-9456. doi: 10.1109/TIM.2008.917179.
- [66] SRINIVASAN, D., NG, W. S., LIEW, A. “Neural-network-based signature recognition for harmonic source identification”, *Power Delivery, IEEE Transactions on*, v. 21, n. 1, pp. 398–405, 2006. ISSN: 0885-8977. doi: 10.1109/TPWRD.2005.852370.
- [67] PATEL, S. N., ROBERTSON, T., KIENTZ, J. A., et al. “At the flick of a switch: detecting and classifying unique electrical events on the residential power line”. In: *Proceedings of the 9th international conference on Ubiquitous computing, UbiComp '07*, pp. 271–288, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN: 978-3-540-74852-6. Disponível em: <http://dl.acm.org/citation.cfm?id=1771592.1771608>.

- [68] GUPTA, S., REYNOLDS, M. S., PATEL, S. N. “ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home”. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, pp. 139–148, New York, NY, USA, 2010. ACM. ISBN: 978-1-60558-843-8. doi: 10.1145/1864349.1864375. Disponível em: <<http://doi.acm.org/10.1145/1864349.1864375>>.
- [69] SULTANEM, F. “Using appliance signatures for monitoring residential loads at meter panel level”, *Power Delivery, IEEE Transactions on*, v. 6, n. 4, pp. 1380–1385, 1991. ISSN: 0885-8977. doi: 10.1109/61.97667.
- [70] CHAN, W., SO, A. T. P., LAI, L. “Harmonics load signature recognition by wavelets transforms”. In: *Electric Utility Deregulation and Restructuring and Power Technologies, 2000. Proceedings. DRPT 2000. International Conference on*, pp. 666–671, 2000. doi: 10.1109/DRPT.2000.855745.
- [71] LEE, W. K., FUNG, G. S. K., LAM, H. Y., et al. “Exploration on Load Signatures”, *International Conference on Electrical Engineering (ICEE)*, 2004.
- [72] LAM, H. Y., FUNG, G. S. K., LEE, W. K. “A Novel Method to Construct Taxonomy Electrical Appliances Based on Load Signatures”, *Consumer Electronics, IEEE Transactions on*, v. 53, n. 2, pp. 653–660, 2007. ISSN: 0098-3063. doi: 10.1109/TCE.2007.381742.
- [73] LIANG, J., NG, S. K. K., KENDALL, G., et al. “Load Signature Study - Part I: Basic Concept, Structure, and Methodology”, *Power Delivery, IEEE Transactions on*, v. 25, n. 2, pp. 551–560, 2010. ISSN: 0885-8977. doi: 10.1109/TPWRD.2009.2033799.
- [74] INAGAKI, S., EGAMI, T., SUZUKI, T., et al. “Nonintrusive appliance load monitoring based on integer programming”, *Electrical Engineering in Japan*, v. 174, n. 2, pp. 18–25, 2011. ISSN: 1520-6416. doi: 10.1002/eej.21040. Disponível em: <<http://dx.doi.org/10.1002/eej.21040>>.
- [75] BERGES, M., GOLDMAN, E., MATTHEWS, H. S., et al. “Training Load Monitoring Algorithms on Highly Sub-Metered Home Electricity Consumption Data”, *Tsinghua Science & Technology*, v. 13, Supplement 1, n. 0, pp. 406 – 411, 2008. ISSN: 1007-0214. doi: [http://dx.doi.org/10.1016/S1007-0214\(08\)70182-2](http://dx.doi.org/10.1016/S1007-0214(08)70182-2). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1007021408701822>>.

- [76] BERGES, M., GOLDMAN, E., MATTHEWS, H. S., et al. “Learning Systems for Electric Consumption of Buildings”, *ASCE International Workshop on Computing in Civil Engineering*, 2009.
- [77] LIANG, J., NG, S., KENDALL, G., et al. “Load Signature Study - Part II: Disaggregation framework, simulation and applications”. In: *Power and Energy Society General Meeting, 2010 IEEE*, pp. 1–1, 2010. doi: 10.1109/PES.2010.5589424.
- [78] ZEIFMAN, M., ROTH, K. “Non-Intrusive Appliance Load Monitoring (NIALM): Review and Outlook”. Apresentado na conferência: International Conference on Consumer Electronics, 2011. Disponível em: <http://cse.fraunhofer.org/Portals/55819/docs/nialm-ieee-2011.pdf>. Acessado em: 5 de nov. 2013.
- [79] BERGMAN, D., JIN, D., JUEN, J., et al. “Distributed non-intrusive load monitoring”. In: *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*, pp. 1–8, 2011. doi: 10.1109/ISGT.2011.5759180.
- [80] ZEIFMAN, M., ROTH, K. “Viterbi algorithm with sparse transitions (VAST) for nonintrusive load monitoring”. In: *Computational Intelligence Applications In Smart Grid (CIASG), 2011 IEEE Symposium on*, pp. 1–8, 2011. doi: 10.1109/CIASG.2011.5953328.
- [81] ZEIFMAN, M., AKERS, C., ROTH, K. “Nonintrusive appliance load monitoring (NIALM) for energy control in residential buildings”, *Sixth International Conference on Energy Efficiency in Domestic Appliances and Lighting, EEDAL*, 2011.
- [82] ZEIFMAN, M. “Disaggregation of home energy display data using probabilistic approach”, *Consumer Electronics, IEEE Transactions on*, v. 58, n. 1, pp. 23–31, 2012. ISSN: 0098-3063. doi: 10.1109/TCE.2012.6170051.
- [83] ZEIFMAN, M. “An approximate probabilistic approach for event-based disaggregation”, *1st International Workshop on Non-Intrusive Load Monitoring*, 2012.
- [84] ZEIFMAN, M., ROTH, K., STEFAN, J. “Automatic recognition of major end-uses in disaggregation of home energy display data”. In: *Consumer Electronics (ICCE), 2013 IEEE International Conference on*, pp. 104–105, 2013. doi: 10.1109/ICCE.2013.6486814.

- [85] EGARTER, D., SOBE, A., ELMENREICH, W. “Evolving Non-Intrusive Load Monitoring”. In: Esparcia-Alcázar, A. (Ed.), *Applications of Evolutionary Computation*, v. 7835, *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 182–191, 2013. ISBN: 978-3-642-37191-2. doi: 10.1007/978-3-642-37192-9_19. Disponível em: <http://dx.doi.org/10.1007/978-3-642-37192-9_19>. Acessado em: 3 de nov. 2013.
- [86] FROEHLICH, J., LARSON, E., GUPTA, S., et al. “Disaggregated End-Use Energy Sensing for the Smart Grid”, *Pervasive Computing, IEEE*, v. 10, n. 1, pp. 28–39, 2011. ISSN: 1536-1268. doi: 10.1109/MPRV.2010.74.
- [87] DO NASCIMENTO, D. M. *Identificação Não Invasiva de Cargas Elétricas Residenciais*. Tese de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, 2012.
- [88] RODRIGUES, R. D. P. *Reconhecimento de Cargas Elétricas Monofásicas Não-Lineares Através da Decomposição Wavelet e de Redes Neurais Artificiais*. Tese de Mestrado, UNIFEI, Itajubá, MG, 2009.
- [89] JAYASREE, T., DEVARAJ, D., SUKANESH, R. “Classification of Transients using Wavelet Based Entropy and Radial Basis Neural Networks”, 2009.
- [90] NAMBI, A. S. N., PAPAIOANNOU, T. G., CHAKRABORTY, D., et al. “Sustainable Energy Consumption Monitoring in Residential Settings”. In: *2nd IEEE INFOCOM Workshop on Communications and Control for Smart Energy Systems (CCSES 2013)*, 2013.
- [91] TOMEK, S., ANDRUSHEVICH, A., WALTHER, P., et al. “iHomeLab experience with NIALM for Smart Buildings”, *1st International Workshop on Non-Intrusive Load Monitoring*, 2012.
- [92] BERGES, M., SOIBELMAN, L., MATTHEWS, S. H. “Leveraging Data From Environmental Sensors to Enhance Electrical Load Disaggregation Algorithms”. In: *13th International Conference on Computing in Civil and Building and Engineering*, 2010.
- [93] CICHOCKI, A., ZDUNEK, R., PHAN, A. H., et al. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. Wiley. com, 2009.
- [94] ANDERSON, K., OCNEANU, A., BENITEZ, D., et al. “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, ago. 2012.

- [95] KOLTER, J. Z., JOHNSON, M. J. “REDD: A public data set for energy disaggregation research”. In: *SustKDD workshop on Data Mining Applications in Sustainability*, 2011.
- [96] “Portal do medidor de energia Watts Up”. . Disponível em: <<https://www.wattsupmeters.com/>>. Acessado em: 5 de nov. 2013.
- [97] FORNEY, G.D., J. “The viterbi algorithm”, *Proceedings of the IEEE*, v. 61, n. 3, pp. 268–278, 1973. ISSN: 0018-9219. doi: 10.1109/PROC.1973.9030.
- [98] FERNÁNDEZ, R. M. D. C., ROJAS, H. N. D. “An overview of wavelet transforms application in power systems”, *IEEE Transactions on Power Delivery*, pp. 24–28, 2002.
- [99] BASIR, O., YUAN, X. “Engine fault diagnosis based on multi-sensor information fusion using Dempster–Shafer evidence theory”, *Information Fusion*, v. 8, n. 4, pp. 379 – 386, 2007. ISSN: 1566-2535. doi: <http://dx.doi.org/10.1016/j.inffus.2005.07.003>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S156625350500076X>>.
- [100] AGUIAR, J. C. R. *Identificação Não-Invasiva de Cargas Elétricas Residenciais Utilizando Redes Neurais*. Tese de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, 1996.
- [101] DO PRADO, C. B. *Sistema de Monitoração de Carga Elétrica Residencial com Implementação para Diferentes Cenários de Aplicação*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2006.
- [102] CAÑÓN, A. D. O. *Sistema Neural de Monitoração Não Invasiva de Carga Elétrica Residencial*. Tese de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, 2009.
- [103] RIBEIRO, G. F., NAVARRO, V., DA SILVA, A. L., et al. *Metodologias para Detecção de Eventos*. Relatório técnico, CEPEL/ELETROBRÁS, Rio de Janeiro, RJ, 2011.
- [104] AKAGI, H., WATANABE, E. H., AREDES, M. *Instantaneous power theory and applications to power conditioning*, v. 31. Wiley. com, 2007.
- [105] “Portal do medidor Yokogawa utilizado”. Disponível em: <<http://tmi.yokogawa.com/us/products/digital-power-analyzers/digital-power-analyzers/model-cw240-clamp-on-power-meters/>>. Acessado em: 6 de nov. 2013.

- [106] OPPENHEIM, A. V., SCHAFER, R. W. *Discrete-time Signal Processing*. Upper Saddle River, NJ, USA, Prentice-Hall, Inc., 1989. ISBN: 0-13-216292-X.
- [107] EIBEN, A., SMITH, J. *Introduction to Evolutionary Computing*. Natural Computing Series. Springer, 2003. ISBN: 9783540401841.