

João Miguel Rodrigues de Oliveira

SIMULADOR DE CARGAS DO SECTOR RESIDENCIAL – AVALIAÇÃO DE IMPACTO DE MEDIDAS DE GESTÃO DA PROCURA

Mestrado Integrado em Engenharia Electrotécnica e de Computadores

Setembro 2012



UNIVERSIDADE DE COIMBRA



Departamento de Engenharia Electrotécnica e de Computadores
Mestrado Integrado em Engenharia Electrotécnica e de Computadores

SIMULADOR DE CARGAS DO SECTOR RESIDENCIAL
AValiação de Impacto de Medidas de Gestão da Procura

Aluno:

João Miguel Rodrigues de Oliveira

Presidente:

Professor Doutor Carlos Alberto Henggeler de Carvalho Antunes

Orientador:

Professor Doutor Álvaro Filipe Peixoto Cardoso de Oliveira Gomes

Vogais:

Professor Doutor Sérgio Manuel Ângelo da Cruz

Agradecimentos

Ao Professor Doutor Álvaro Filipe Gomes pelo total apoio, acompanhamento e disponibilidade prestada no decorrer da dissertação.

Aos meus pais e à minha irmã pelos incentivos constantes, pelo interesse e pela disponibilidade que sempre tiveram para me ouvir, aconselhar, apoiar e motivar.

À minha namorada pela compreensão, encorajamento, carinho e paciência.

A todos os meus amigos e colegas que, directa ou indirectamente, tornaram possível a concretização da dissertação.

A todos o meu Bem-haja!

Resumo

As questões com que as sociedades actuais se defrontam relacionadas com a problemática da preservação do meio ambiente e do desenvolvimento sustentado, impõe uma visão mais geral no processo de produção, gestão, distribuição e consumo do que é produzido e, no caso em apreço, da electricidade, que relativamente às outras mercadorias consumidas, não é viável armazenar em grandes quantidades, sendo esta mais uma razão para a sua gestão eficaz.

Actualmente impõe-se uma atitude diferente dos consumidores, passando de meros *clientes*, *consumidores*, para partes interessadas em todo o processo. Para isso é indispensável disponibilizar aos consumidores instrumentos eficazes e fáceis de utilizar que permitam uma efectiva adesão e participação nas estratégias de gestão da procura. Também eventuais entidades interessadas em promover este tipo de actividade terão toda a vantagem em apresentar medidas atractivas também para os consumidores, sendo, por isso, parte interessada nesses instrumentos.

Nesta dissertação é proposto um modelo de simulação de cargas residenciais (SIMCAR), tendo como objectivo principal a avaliação dos impactos das estratégias de gestão da procura, sendo capaz de lidar com a aleatoriedade e variabilidade habitualmente associadas à utilização dos serviços de energia no sector residencial.

Devido às diversas características dos diferentes tipos de cargas residenciais, é necessário utilizar estratégias diferentes para cada carga, quer relativamente à implementação de medidas de gestão, quer relativamente aos dados, e seu tratamento, que permitem caracterizar o comportamento das cargas. Assim pode:

- Recorrer-se a dois algoritmos de reconhecimento de padrões, sendo eles o *k-means* e de evolução diferencial, para caracterizar o funcionamento de algumas cargas ao longo do dia, com base em recolhas efectuadas com essas cargas em funcionamento;
- Recorrer-se a modelos fisicamente baseados para recriar diagramas de cargas termostáticas, onde todos os fenómenos físicos reais são tidos em conta, podendo dessa forma lidar com a aleatoriedade e variabilidade das cargas;
- Recorrer-se a (outras) fontes de dados existentes para a caracterização estatística da utilização das cargas ao longo do tempo.

Neste texto são apresentados estudos de caso para demonstrar as potencialidades do simulador, onde se podem visualizar e analisar os resultados da “aplicação” das técnicas de tratamento de dados para obtenção dos consumos das cargas, bem como verificar o impacto de algumas medidas de gestão da procura, tanto individualmente, como num grupo de

consumidores. É também efectuada uma análise comparativa entre o desempenho dos dois algoritmos de reconhecimento de padrões quando usados na obtenção de diagramas de cargas.

Palavras-chave:

Simulador de consumos residenciais, Diagramas de carga, Gestão da Procura, *Data Mining*, Evolução diferencial, *k-means*, *Clusters*, Modelo fisicamente baseado.

Abstract

The problems which the modern society faces nowadays related to the issue of preserving the environment and sustainable development, requires a broader view of the production, management, distribution and consumption process, and in this instance, the electricity that relative to the other goods consumed, cannot be stored. This fact is one more reason for its effective management.

Currently we need a different attitude of consumers. They need to be more than just simple clients, and be stakeholders in all the process. It is essential to provide the client with efficient tools and easy to use, allowing an effective participation in demand management strategies. Also, any entity interested in promoting this type of activity will have all the advantage in presenting measures also attractive to consumers, and therefore stakeholders in these instruments.

This thesis proposes a bottom-up approach for simulation of residential loads simulation, whose principal aim is to evaluate the impacts of management strategies, being able to handle the randomness and variability commonly associated with the use of energy services in the residential sector.

Due to the different characteristics of different types of residential loads, it's necessary to use different strategies for each load, either with regard to the implementation of management measures, or for the data and its treatment, which permit characterize the behavior of the loads. Therefore:

- Resorting to two algorithms for pattern recognition, namely the k-means and differential evolution to characterize the operation of some loads throughout the day, based on monitoring carried out on the loads in operation;
- Resorting to physically based models to recreate diagrams thermostatic loads, where all the real physical phenomena are taken into account, and can this deal with the variability and randomness loads;
- Resorting to (other) resources of data for statistical characterization of the use of the loads over time.

This paper presents case studies to demonstrate the capabilities of the simulator, where it is possible to view and analyze the results of the "application" of data processing techniques for obtaining the loads of consumption, as well as check the impact of some management measures

searching, both individually as a group of consumers. It also made a comparison between the performance of two algorithms for pattern recognition when used in obtaining diagrams loads.

Key words:

Residential load simulator, Load diagram, Demand Response, Data Mining, Differential Evolution, k-means, Cluster, Physically based model.

Índice

Agradecimentos	i
Resumo.....	iii
<i>Abstract</i>	v
Índice de Figuras.....	ix
Índice de Tabelas.....	xiii
Lista de Abreviaturas.....	xv
1. Introdução.....	1
1.1. Enquadramento Geral.....	1
1.2. Motivação.....	3
1.3. Estrutura.....	5
2. Gestão da Procura.....	6
2.1. Caracterização de cargas.....	9
3. Tratamento de dados.....	11
3.1. <i>Data Mining</i>	12
3.1.1. Determinação do número de <i>clusters</i>	15
3.2. Evolução Diferencial.....	16
3.2.1. Algoritmo Geral.....	16
3.2.2. Algoritmo de ED utilizado para <i>clusterização</i>	17
3.3. Teste de silhueta	19
3.4. Análise Comparativa entre <i>k-means</i> e ED.....	20
3.5. Conversor de resoluções temporais.....	21
4. SIMCAR – SIMulador de CArgas Residenciais.....	22
4.1. Carga Base.....	24
4.2. Máquina de lavar roupa, loiça e secar roupa.....	26
4.3. Ar Condicionado, <i>Inverter</i> e Termoacumulador.....	31
4.4. Interface	34
5. Caso de estudo.....	34
5.1. Algoritmos.....	35
5.2. Diagramas de carga individuais.....	39
5.3. Diagramas de carga finais.....	41
5.3.1. Apresentação de dois cenários distintos.....	41
5.3.2. Aplicação da estratégia de GP de reparametrização do termostato	44
5.4. Simulação do diagrama de carga de um consumidor e avaliação do impacto na factura da alteração de consumos	47

5.5. Obtenção de distribuições através dos algoritmos	49
6. Conclusão e Futuros Trabalhos	50
7. Bibliografia.....	52
Anexos	54
Anexo I – Aplicação do algoritmo <i>k-means</i>	54
Anexo II – Aplicação do <i>k-means</i> e Evolução Diferencial e silhuetas obtidas	55
Anexo III – Consumos de 14 dias de um consumidor.....	57
Anexo IV – Caracterização das Cargas.....	59
Anexo V – Parametrização do SIMCAR.....	61
Anexo VI – Determinação do valor K do <i>k-means</i>	64

Índice de Figuras

Figura 1: Evolução da produção de energia eléctrica até 2040 [1]	1
Figura 2: Produção de energia primária na UE [2].....	1
Figura 3: Procura de energia primária em Portugal [3]	2
Figura 4: Repartição do consumo de energia por sectores [4]	2
Figura 5: Objectivos de DSM [11] [15] [12].....	7
Figura 6: Sensibilidade / Protecção do consumidor com a variação do preço real do kWh [16].....	8
Figura 7: Classificação dos diferentes tipos de cargas de acordo com as estratégias de controlo que se podem aplicar [12].....	10
Figura 8: Representação dos processos aplicados a uma base de dados para retirar informação útil.	13
Figura 9: Comparação entre os algoritmos hierárquicos de aglomeração e de divisão [20] [23].....	14
Figura 10: Obtenção do número de <i>clusters</i> ideal para uma base de dados de 1.000 elementos aleatórios (esquerda) e uma base de dados de 10.000 elementos aleatórios.....	15
Figura 11: Representação dos cromossomas	18
Figura 12: Ilustração da população corrente (esquerda) e do cálculo de um vector-diferença (direita).....	18
Figura 13: Representação da população mutante (cima, esquerda), do processo de recombinação entre população mutante e população corrente (cima, direita), da população teste (baixo, esquerda) e da população corrente da 2ª Geração (baixo, direita).....	19
Figura 14: Variação da medida de silhueta e respectiva classificação da qualidade dos <i>clusters</i> [30]	19
Figura 15: Conversão do diagrama de um ciclo de uma MSR.	21
Figura 16: Dados de entrada do SIMCAR	23
Figura 17: Organização do ficheiro Excel de resultados do SIMCAR.....	23
Figura 18: Consumo de um consumidor de um dia, através de medições reais e através dos algoritmos de <i>clusterização</i>	24
Figura 19: Consumo médio de um consumidor ao longo de 14 dias, obtido através de medições reais e da aplicação dos algoritmos de <i>clusterização</i> a essas mesmas medições	25
Figura 20: Comparação do consumo médio real de CB (Preto) de um conjunto de consumidores, com o consumo médio obtido através da aplicação do <i>k-means</i> (Azul) e ED (Vermelho).....	25
Figura 21: Determinação do número de ciclos de funcionamento da MLR para Sábado, de acordo com a Tabela 6.....	27
Figura 22: Exemplificação da alocação de um ciclo.....	28
Figura 23: Distribuição de probabilidade inicial (Azul) e distribuição de probabilidade em que existem períodos que não é possível alocar ciclos de funcionamento (Laranja).	29
Figura 24: Cálculo do intervalo em que é possível alocar o ciclo. Laranja representa intervalos já ocupados, verde representa o <i>slot</i> livre e amarelo o intervalo em que o ciclo pode iniciar o funcionamento.	29
Figura 25: Calculadora de distribuição de probabilidades	30
Figura 26: Fluxograma simples para demonstração dos processos para a criação dos consumos pelo SIMCAR.....	31
Figura 27: Parametrização dos perfis de utilização dos serviços de energia.....	32
Figura 28: Tabela para a definição de uma estratégia de controlo personalizada pelo utilizador	32
Figura 29: Tabela existente nas folhas Excel do AC, <i>Inverter</i> e TR para a inserção da distribuição dos perfis de utilização para DU, Sábado e Domingo.....	33
Figura 30: Interface do SIMCAR.....	34

Figura 31: Determinação do número k de <i>clusters</i> para a situação de uma base de dados de medições de 10 diagramas de carga e de um de 365 diagramas.....	35
Figura 32: Consumo médio de CB de 50 consumidores (esquerda) e de CB de 500 consumidores (direita)	36
Figura 33: Resultados obtidos da aplicação do algoritmo de ED para diferente número de <i>clusters</i> e de cromossomas	36
Figura 34: Comparação da diferença média resultante dos diversos teste e tempos de execução.....	37
Figura 35: Consumo de carga base de 300 consumidores obtido através do algoritmo de Evolução Diferencial e do algoritmo de DM, <i>k-means</i>	37
Figura 36: Diferença entre os consumos de carga base dos 300 consumidores obtido pelos dois métodos	38
Figura 37: Distribuição de probabilidade do funcionamento das MLR obtido de [10].....	39
Figura 38: Comparação entre o consumo em funcionamento normal das máquinas de lavar roupa (azul) com a aplicação de diferentes técnicas de controlo de carga (outras cores).....	39
Figura 39: Consumo de 480 AC, com os mesmo parâmetros, apenas variando a temperatura definida do termostato para 25°C, 24,5°C e 24°C.....	40
Figura 40: Diagrama de Carga agregado de todos os consumidores para a Sexta-Feira.....	42
Figura 41: Diagrama de Carga agregado de todos os consumidores para o Domingo.....	42
Figura 42: Diagrama de carga agregado dos consumidores do segundo cenário para Sexta-Feira.....	44
Figura 43: Diagrama de carga agregado dos consumidores do segundo cenário para o Domingo	44
Figura 44: Comparação entre os consumos de Quarta-Feira para AC e Inverter parametrizados para 24°C (esquerda) e 25°C (direita).....	46
Figura 45: Comparação entre os consumos de Domingo para AC e Inverter parametrizados para 24°C (esquerda) e 25°C (direita).....	46
Figura 46: Diagrama de carga de um dia de um consumidor, sem a aplicação de nenhuma estratégia de GP.....	48
Figura 47: Diagrama de carga do mesmo consumidor da Figura 46, mas com a aplicação de estratégias de GP.....	48
Figura 48: Distribuição de probabilidade do funcionamento da MLR segundo [10] (laranja) e distribuição de probabilidade do funcionamento da máquina de lavar roupa quando aplicado o algoritmo <i>k-means</i> (azul).....	49
Figura 49: Distribuição de probabilidade do funcionamento da MLL segundo [10] (laranja) e distribuição de probabilidade do funcionamento da máquina de lavar roupa quando aplicado o algoritmo <i>k-means</i> (azul).....	49
Figura 50: Aplicação do algoritmo <i>k-means</i> , com $k=10$, a uma base de dados de 5000 elementos, gerada de forma aleatória através de uma distribuição uniforme.....	54
Figura 51: Aplicação do algoritmo <i>k-means</i> , com $k=10$, a uma base de dados de 5000 elementos, gerada através de uma distribuição normal com $\mu=0$ e $\sigma=1$	54
Figura 52: Resultados obtidos com o algoritmo de Evolução Diferencial com 5 <i>clusters</i> 5 cromossomas (esquerda) e a silhueta resultante (direita).....	55
Figura 53: Resultados obtidos com o algoritmo de Evolução Diferencial com 5 <i>clusters</i> 10 cromossomas (esquerda) e a silhueta resultante (direita)	55
Figura 54: Resultados obtidos com o algoritmo de Evolução Diferencial com 10 <i>clusters</i> 5 cromossomas (esquerda) e a silhueta resultante (direita).....	55
Figura 55: Resultados obtidos com o algoritmo de Evolução Diferencial com 10 <i>clusters</i> 10 cromossomas (esquerda) e a silhueta resultante (direita).....	56
Figura 56: Resultados obtidos com o algoritmo <i>k-means</i> com 5 <i>clusters</i> (esquerda) e a silhueta resultante (direita)	56

Figura 57: Resultados obtidos com o algoritmo <i>k-means</i> com 10 <i>clusters</i> (esquerda) e a silhueta resultante (direita).....	56
Figura 58: Resultados obtidos com o algoritmo <i>k-means</i> com 15 <i>clusters</i> (esquerda) e a silhueta resultante (direita).....	56
Figura 66: Diagramas de carga do consumidor monitorizado num período de 14 dias	58
Figura 59: Repartição dos consumos eléctricos no sector residencial [10]	59
Figura 60: Ciclo típico de funcionamento de uma MLR num programa de 30°C/40°C e consumo nacional das MLR [10]	59
Figura 61: Ciclo típico de funcionamento de uma MLL e consumo nacional das MLL [10]	59
Figura 62: Ciclo típico de funcionamento de uma MSR e consumo nacional das MSR [10]	60
Figura 63: Diferença entre AC convencional (esquerda) e <i>inverter</i> (direita) [18]	60
Figura 64: Funcionamento de um TR.....	60
Figura 65: Funcionamento de um frigorífico durante 6 horas [19]	60
Figura 67: Folha de parametrização “Geral”	61
Figura 68: Folha de parametrização “MLR”	61
Figura 69: Folha de parametrização “MSR”	61
Figura 70: Folha de parametrização “MLL”	62
Figura 71: Folha de parametrização “AC”	62
Figura 72: Folha de parametrização “TR”.....	62
Figura 73: Folha de parametrização “CriarControlo_AC” – Parte 1/2.	63
Figura 74: Folha de parametrização “CriarControlo_TR” – Parte 1/2.....	63
Figura 75: Parte 2/2 das folhas de cálculo para a criação dos perfis de utilização dos serviços de energia (“CriarControlo_AC”, “CriarControlo_Inv” e “CriarControlo_TR”).....	63
Figura 76: Determinação do número ideal de <i>clusters</i> a utilizar no para uma base de dados com 15 recolhas diária (esquerda), 30 recolhas (direita) e 185 recolhas (inferior).....	64

Índice de Tabelas

Tabela 1: Descrição dos objectivos de DSM.....	7
Tabela 2: Comparação dos resultados obtidos pelo teste da silhueta para os diferentes casos	20
Tabela 3: Diferença média entre os diferentes diagramas obtidos através da aplicação dos algoritmos ...	26
Tabela 4: Distribuição por programas dos ciclos de funcionamento obtidos para a MLR.....	26
Tabela 5: Distribuição dos ciclos de funcionamento das MLL.....	26
Tabela 6: Tabela explicativa de como é realizada a distribuição dos ciclos de funcionamento ao longo da semana (neste caso foram usadas cinco utilizações por semana).....	27
Tabela 7: <i>Centroids</i> de cada <i>cluster</i> obtidos pelo <i>k-means</i> e pelo algoritmo de ED para as 00:45h bem como a respectiva distribuição de probabilidade cumulativa.....	38
Tabela 8: <i>Centroids</i> de cada <i>cluster</i> obtidos pelo <i>k-means</i> e pelo algoritmo de ED para as 20:00h bem como a respectiva distribuição de probabilidade cumulativa.....	38
Tabela 9: Consumo de 480 AC, cujo termostato foi definido para 24°C, 24,5°C e 25°C.....	40
Tabela 10: Caracterização das máquinas simuladas no primeiro cenário	41
Tabela 11: Distribuição de probabilidades dos ciclos de funcionamento das máquinas utilizadas no primeiro cenário	41
Tabela 12: Distribuição dos perfis de utilização dos serviços de energia do TR, AC e <i>Inverter</i> para o primeiro cenário	42
Tabela 13: Caracterização das máquinas simuladas no segundo cenário.....	43
Tabela 14: Distribuição de probabilidades dos ciclos de funcionamento das máquinas utilizadas no segundo cenário.....	43
Tabela 15: Distribuição dos perfis de utilização dos serviços de energia do TR, AC e <i>Inverter</i> para o segundo cenário.....	43
Tabela 16: Distribuição dos ciclos da MLR, MLL e MSR ao longo da semana.....	45
Tabela 17: Distribuição dos perfis de utilização dos serviços de energia dos AC e <i>inverters</i> simulados ..	45
Tabela 18: Distribuição dos perfis de utilização dos serviços de energia dos TR.....	45
Tabela 19: Redução energética verificada com a reparametrização do AC e Inverter de 24°C para 25°C	47
Tabela 20: Diferença nos consumos verificados com e sem o desvio de consumos	48

Lista de Abreviaturas

Abreviatura	Descrição
SIMCAR	Simulador de Cargas Residenciais
UE	União Europeia
GP	Gestão da Procura
MFB	Modelo Fisicamente Baseado
DM	<i>Data Mining</i>
ED	Evolução Diferencial
SEE	Sistema de Energia Eléctrica
GEE	Gases com Efeito de Estufa
DSM	<i>Demand-side Management</i>
MLR	Máquina de Lavar Roupa
MLL	Máquina de Lavar Loiça
MSR	Máquina de Secar Roupa
CB	Carga Base
AC	Ar Condicionado
TR	Termoacumulador
DU	Dias Úteis

1. Introdução

1.1. Enquadramento Geral

As sociedades modernas caracterizam-se pela grande dependência que apresentam relativamente aos combustíveis fósseis. No entanto, quando se observam os impactos ambientais inerentes ao consumo destes combustíveis e as sucessivas crises energéticas, torna-se evidente a necessidade de desenvolver alternativas mais sustentáveis. Estas passam pela substituição das fontes primárias fósseis por fontes renováveis e pela criação de métodos de utilização da energia cada vez mais eficientes. Contudo, uma análise das tendências de evolução da utilização das fontes primárias usadas para a geração de electricidade nas sociedades actuais é suficiente para afirmar que será um percurso sinuoso até ser eliminada a dependência das fontes não renováveis por completo. Este facto é bem visível através dos gráficos da Figura 1, que demonstram que a produção de energia eléctrica através de fontes não renováveis continuará a aumentar.

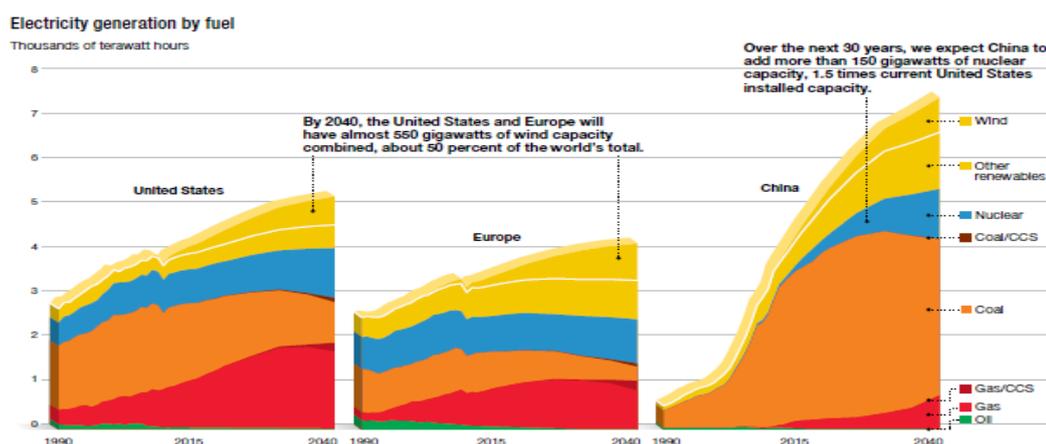


Figura 1: Evolução da produção de energia eléctrica até 2040 [1]

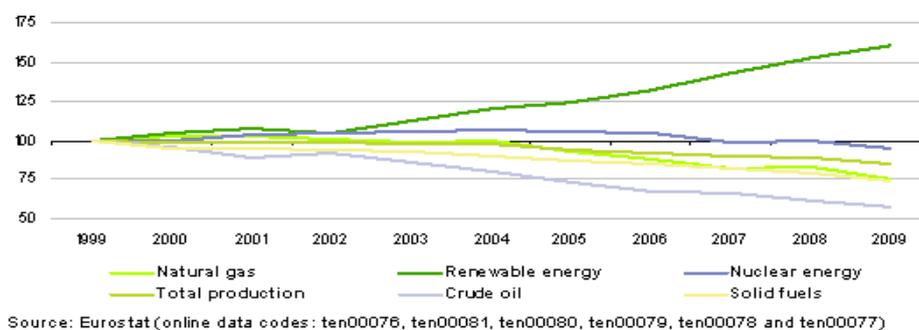


Figura 2: Produção de energia primária na UE [2]

Através da análise da Figura 2 é possível concluir que a União Europeia (EU) está cada vez mais dependente das importações de energia primária de modo a satisfazer a procura. Como consequência, o sistema energético europeu torna-se cada vez mais vulnerável às flutuações de preço praticadas nos mercados internacionais [2].

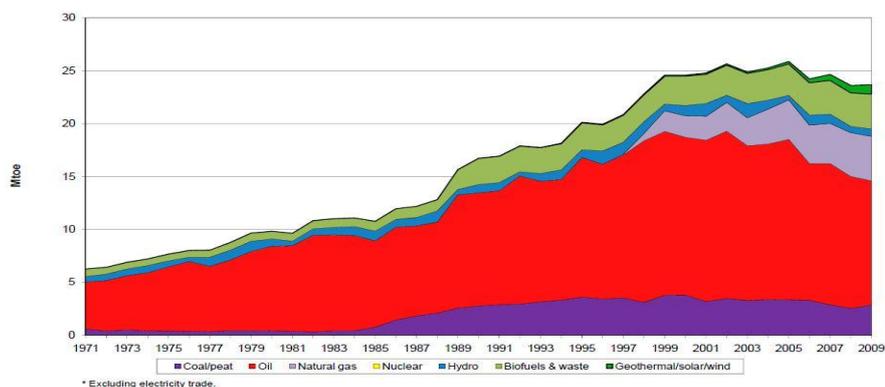


Figura 3: Procura de energia primária em Portugal [3]

Tal como se tem verificado a nível mundial, pode observar-se que em Portugal a procura de energia primária tem também vindo a crescer. Esta constatação comprova que, apesar do aumento do consumo de energias renováveis, estas ainda continuam a representar uma pequena parcela do consumo energético total [2].

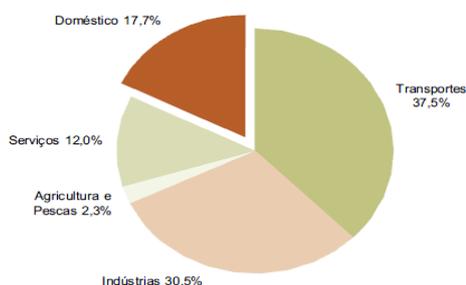


Figura 4: Repartição do consumo de energia por sectores [4]

Na Figura 4 é possível verificar que o sector dos transportes é o sector responsável pela maior parcela do consumo de energia a nível nacional, com cerca de 38%, seguido do industrial com quase 31%. Os sectores Residencial e de Serviços representam, em conjunto, quase 30% do consumo total.

A energia eléctrica apresenta características que a distinguem claramente de outros bens comercializados. Entre elas, podemos nomear o facto de não ser viável, por enquanto, o armazenamento de grandes quantidades, fazendo com que a produção tenha que ser contínua, de modo a satisfazer a procura. O sistema de energia eléctrica deve estar adaptado e ter capacidade para satisfazer integralmente a procura em qualquer instante, mesmo que a ponta de consumo ocorra apenas durante breves minutos no decorrer do ano. Como a procura energética é um reflexo do quotidiano das sociedades, a tendência será a sua variação ao longo do tempo, sendo continuamente necessário reagir de modo a satisfazê-la. Por outro lado, o mix de geração é diverso, quer em termos de tecnologias quer em combustíveis, fazendo com que o custo que advém da satisfação da procura seja oscilante. Para além dos diversos métodos que visam o aperfeiçoamento do modo como a energia eléctrica é utilizada por alguns equipamentos, a

eficiência na utilização dos recursos energéticos no seu conjunto pode ser melhorada através da chamada gestão da procura (GP). A GP contribui para a redução da ponta máxima ao desviar os consumos para períodos de menor procura ou através do controlo directo de cargas, com as consequentes reduções no *stress* respeitante ao aumento de capacidade do sistema, nas perdas e nos encargos para o consumidor, para além de não ser necessário o recurso a grupos geradores menos eficientes. Um dos sectores visados pela GP é o Residencial, que se torna teoricamente mais atractivo para aplicação destas medidas devido ao elevado número de cargas com potencial de actuação. No entanto, na prática, a implementação destas estratégias de controlo torna-se complexa devido ao facto deste sector se dividir em milhões de consumidores. Para além disso, é essencial garantir a não degradação da qualidade do serviço fornecido pelas cargas, aquando da aplicação das estratégias de controlo englobadas na GP.

A existência de simuladores de consumos residenciais ganha especial relevo no que respeita ao desenvolvimento das estratégias da GP, funcionando quer como instrumentos de auxílio na concepção das mesmas, quer como ferramentas de avaliação de impactos da sua implementação. A grande dificuldade da criação destes simuladores prende-se com a existência de um elevado grau de incerteza e variabilidade existente nos diagramas de carga de cada consumidor. Para além de cada consumidor apresentar uma procura diferente dos demais, o padrão e a amplitude de consumo pode também ser diferente de dia para dia, variando ao longo do ano. Constatase, então, que são muitos os factores que podem influenciar o padrão de consumo.

1.2. Motivação

Em Portugal têm também sido adoptadas algumas iniciativas de promoção do aumento da eficiência energética. Como exemplos, é possível citar a criação do Centro para a Conservação de Energia (CCE), em 1984, o qual foi reestruturado em 2000 e é actualmente conhecido como ADENE, o caso do Plano Nacional para o Desenvolvimento Económico e Social (PNDES), o Programa Nacional para as Alterações Climáticas (PNAC), o Programa E4 – Eficiência Energética e Energias Endógenas, o Programa Água Quente Solar para Portugal (AQSpP), o Programa para a Eficiência energética em Edifícios (P3E), a Estratégia Nacional de Desenvolvimento Sustentável (ENDS) e a Estratégia Nacional para a Energia.

O consumo de energia eléctrica de uma habitação e a variação da procura desta é fortemente influenciado pelo tempo de permanência dos residentes no domicílio e pela frequência de utilização de aparelhos consumidores de energia, muitas vezes com curta duração mas de elevada potência, como é o caso das chaleiras eléctricas [5]. Em [6] é apresentado um simulador do consumo de energia eléctrica para a iluminação. Neste modelo, a simulação é realizada com base em probabilidades relativas à duração de ocupação da habitação juntamente com os

diferentes níveis de luz natural ao longo do dia. A determinação dos consumos é, assim, realizada através do método *bottom-up*, no qual o consumo agregado é obtido a partir da soma dos consumos individuais. Para gerar os padrões de ocupação, é utilizada uma cadeia de Markov, em que é assumido que um individuo pode estar caracterizado de acordo com um de entre três estados (ausente, presente e activo, presente e inactivo) estando associada uma probabilidade à transição entre cada estado.

A estratégia *bottom-up* foi também utilizada por Capasso, Grattieri, Lamedica e Prudenzi [7] para a criação de um simulador dos consumos das residências italianas. Com base num inquérito sobre o uso do tempo, realizado pelo ISTAT (Instituto Nacional de Estatística Italiano), os autores desenvolveram uma função de distribuição de probabilidade respeitante à duração da presença de pessoas nas habitações. Através da intercepção da presença de pessoas em casa com os perfis de probabilidade de utilização dos electrodomésticos é possível calcular a função pretendida, a que os autores chamaram de recursos humanos, sendo esta a função de distribuição de probabilidade utilizada nas simulações.

O modelo de Gerbec, Gašperič, Šmon e Gubina [8] é um exemplo de determinação de padrões típicos de consumo com base em determinadas amostras. Os autores aplicaram um algoritmo hierárquico de *clusterização* aos dados obtidos de forma a obterem padrões típicos de consumo. Uma vez obtidos, cada consumidor não monitorizado é associado a um desses padrões.

O modelo [9] de Walker e Pokoski representou um grande avanço relativamente à introdução de funções de disponibilidade e tendência nos modelos de simulação de diagramas de consumo residencial. De modo a criar a função de disponibilidade, os autores basearam-se em estudos sobre densidades habitacionais e tempos de viagem para o trabalho e em dados de consumos extraídos em recolhas e inquéritos. A função de tendência tem como objectivo a simulação da rotina existente em certos consumos residenciais. Esta função é dividida em dois tipos, dependendo do tipo de actividade em estudo. No primeiro tipo estão englobadas as actividades que influenciam apenas um electrodoméstico (máquinas de lavar loiça ou roupa), enquanto o segundo tipo diz respeito às actividades que influenciam diferentes electrodomésticos, como é no caso da confecção das refeições, em que podem ser utilizados fornos, microondas, chaleiras, etc.

Destes modelos já existentes, o modelo [7] foi o que se demonstrou ser mais completo. No entanto, boa parte destes simuladores são ou dedicados, no sentido de reproduzirem o consumo de uma utilização final, ou não permitem a simulação de acções de controlo sobre as cargas. No sentido de analisar os impactos de acções de controlo sobre cargas e ajudar na concepção dessas acções é pois necessário construir uma ferramenta adequada.

Esta ferramenta deverá ser versátil, capaz de reconstruir diagramas de consumos residenciais realísticos, sem e com acções de controlo. Deve ainda ser capaz de lidar com as incertezas da

utilização dos serviços de energia e tirar partido de dados já existentes, embora muitas vezes em pequena quantidade e sem a aplicação de acções de controlo.

Na presente dissertação foi desenvolvido o SIMCAR (SIMulador de Cargas Residenciais). A grande vantagem deste simulador quando comparado com os já existentes prende-se com a sua flexibilidade relativamente à introdução de dados de alimentação do simulador. Através do SIMCAR é possível reproduzir diagramas de carga residenciais realísticos, possibilitando a sua utilização para qualquer objectivo, nomeadamente na avaliação do impacto das medidas de GP. As técnicas utilizadas para tratar os dados de entrada da simulação dos consumos são diversas, podendo recorrer-se à aplicação de algoritmos de *data mining* (DM) ou evolução diferencial (ED), utilização de dados de monitorizações já realizadas ou mesmo modelos fisicamente baseados (MFB). O SIMCAR simula o consumo individual de cada carga que, sendo estes agregados de forma a construir o diagrama de carga diário de cada consumidor.

As técnicas de DM e ED são utilizadas para identificar padrões de consumo nas bases de dados de monitorização de diversas cargas, sendo desta forma possível a obtenção da distribuição de probabilidade de consumo em cada intervalo de tempo no decorrer do dia. Estes algoritmos podem ainda ser aplicados para identificar a distribuição da probabilidade de algumas cargas iniciarem ciclos de funcionamento (tal como ocorre no caso da máquina de lavar roupa).

1.3. Estrutura

A dissertação está dividida em seis capítulos.

No primeiro capítulo é feito um enquadramento geral da situação energética actual, bem como apresentada a motivação para a realização deste simulador.

No segundo capítulo encontra-se a caracterização das cargas residenciais e é explicado em que consiste a gestão da procura.

O terceiro capítulo é referente ao tratamento de dados. Aqui é introduzido o conceito de *Data Mining* e de Evolução Diferencial, descrevendo os dois algoritmos utilizados. É também realizada uma análise comparativa entre ambos.

No quarto capítulo é descrita a forma de como o SIMCAR procede para a realização das simulações das diferentes cargas, bem como constrói os diagramas de carga.

O quinto capítulo consiste na apresentação de alguns casos de estudo, onde é realizada a comparação entre os dois algoritmos utilizados quando aplicados a um caso real de determinação de consumos de carga base, é apresentado o resultado de diversas simulações de cargas individuais, sendo possível visualizar impactos de medidas de gestão da procura em termos do consumo individual das cargas. Foram ainda simulados dois cenários com diversos consumidores, sendo apresentado o consumo desagregado por tipo de carga. Ainda é feita a

caracterização de um consumidor, sendo simulado o seu consumo de um dia, estando expostas possíveis poupanças que este poderá obter com recurso a medidas de gestão da procura. Finalmente é exibida a comparação da distribuição de probabilidade de as máquinas de lavar roupa e loiça iniciarem ciclos de funcionamento ao longo do dia obtida através da aplicação do algoritmo *k-means* com os dados obtidos do estudo da DGGE [10].

No sexto capítulo pode-se encontrar a conclusão do trabalho bem como algumas propostas de melhoramento futuras.

2. Gestão da Procura

O sistema de energia eléctrica (SEE) deve ser capaz de satisfazer sempre a procura. Devido à elevada dificuldade existente no armazenamento da energia eléctrica em grandes quantidades, torna-se necessário que esta seja produzida no exacto momento em que é necessária. Caso contrário, a integridade do sistema fica comprometida.

É necessário, no entanto, ter em conta que a necessidade de energia ao longo do dia é extremamente irregular. O SEE tem então de ser sobredimensionado de modo a conseguir satisfazer os picos existentes no consumo, mesmo que estes ocorram escassas vezes. Desta forma, verifica-se um subaproveitamento das infra-estruturas existentes [11].

Tendo em conta estas constatações é possível averiguar que os intervenientes do SEE são prejudicados pela existência de tal irregularidade. As entidades fornecedoras de serviços (produtores, transportadores e distribuidores) apresentam, então, custos superiores de manutenção e operação e maiores despesas no aumento da capacidade das infra-estruturas. Este acréscimo verificado nos custos de produção e transporte de energia eléctrica provoca também um aumento do custo do kWh pago pelo consumidor, ficando também este prejudicado. O meio-ambiente é outro dos lesados durante este processo. A necessidade da criação de reservas girantes com maior capacidade resulta num maior desperdício dos recursos naturais e contribui ainda para a emissão de gases com efeito de estufa (GEE).

Estes problemas podem ser atenuados através do recurso às estratégias de *Demand-Side Management* (DSM).

DSM consiste na alteração dos consumos de energia eléctrica no dia-a-dia em resposta a variações do preço da mesma ou em resposta a incentivos monetários fornecidos por parte das empresas prestadoras de serviços [12] [13].

Verifica-se a existência de duas principais estratégias de DSM: Eficiência Energética e Gestão da Procura (GP).

Nas estratégias de eficiência energética os consumidores são sensibilizados e incentivados para o uso de equipamentos mais eficientes, promovendo assim uma diminuição dos desperdícios energéticos e por conseguinte, da factura mensal do consumidor [14].

As estratégias de GP estão directamente relacionadas com a capacidade de resposta dos consumidores finais perante os estímulos provenientes da operadora de rede, alterando assim os seus consumos. Estes estímulos podem ser de origem diversa, desde a variação do preço do kWh, até à necessidade de aliviar a procura energética em determinadas situações.

De um modo geral, podem identificar-se seis objectivos de DSM, os quais se encontram apresentados na Figura 5 e descritas na Tabela 1.

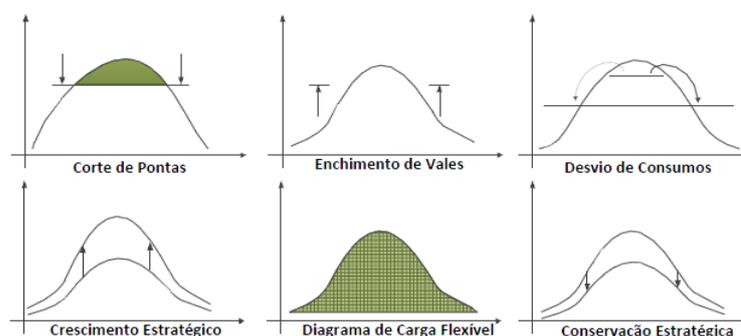


Figura 5: Objectivos de DSM [11] [15] [12]

Tabela 1: Descrição dos objectivos de DSM.

Corte de Pontas	Tem como objectivo a diminuição dos picos que surgem no diagrama de carga. O Controlo directo de carga é o método mais utilizado nesta estratégia.
Enchimento de Vales	O objectivo é o aumento do consumo nas horas em que este é menor.
Desvio de Consumos	O objectivo é desviar os consumos que ocorreriam nas horas de maior consumo (onde o preço da energia é mais elevado) para as horas em que a procura (e por sua vez, o preço) é mais reduzida.
Crescimento Estratégico	Tem como objectivo aumentar os consumos de energia, de uma forma inteligente e eficiente.
Diagrama de Carga Flexível	Pretende-se obter um diagrama onde é possível haver uma boa margem de manobra, sendo possível aumentar ou diminuir a procura sempre que tal seja necessário.
Conservação Estratégica	O consumidor é incentivado ao uso de equipamentos mais eficientes.

Os programas de incentivo à adopção de medidas de GP por parte dos consumidores são programas que se baseiam ou no preço ou na atribuição de incentivos.

Apesar do custo de abastecimento da energia eléctrica não se apresentar constante ao longo do dia, esta variação não é sentida pelos consumidores, uma vez que na maioria dos casos, as tarifas

são fixas durante um determinado período de tempo, geralmente três meses ou um ano. À medida que se vai adoptando outras tarifas que se aproximem dos preços em tempo real, o consumidor apercebe-se da variação do preço, tendo, desta forma, um estímulo para alterar os hábitos de consumo.



Figura 6: Sensibilidade / Protecção do consumidor com a variação do preço real do kWh [16]

O objectivo dos programas baseados nos preços é levar os consumidores a alterar os seus consumos, tirando partido da variação dos preços. Este tipo de programas são então totalmente dependentes do consumidor.

Dentro deste programa existem três tipos de planos [12]:

- **Preço Multi-Horário (TOU¹):** Um Tarifário com diferentes preços unitários, que variam de acordo com o período do dia. Este tarifário reflecte os custos médios de produção e distribuição de energia durante esses períodos de tempo. É o caso das tarifas bi-horárias, tri-horárias e tetra-horárias.
- **Preço em Tempo Real (RTP²):** Tarifário onde o preço da electricidade flutua num intervalo de tempo constante, tipicamente de hora a hora (podendo, no entanto, apresentar outros períodos de tempo, como de quinze em quinze minutos) sendo, deste modo, reflectidas as flutuações do preço da electricidade. Os consumidores são previamente informados destas alterações.
- **Preço de Pico Crítico (CPP³):** A estrutura base consiste numa estrutura multi-horária, diferindo no facto de se verificar uma variação do preço pago pelo consumidor quando o preço da energia eléctrica aumenta significativamente ou quando a integridade do sistema eléctrico se encontra em risco.

Nos programas baseados em incentivos os consumidores recebem incentivos monetários (traduzidos em prémios ou redução do valor do kWh) aquando da recepção do sinal, com o objectivo de os incitar a reduzir o consumo [11] [12]. Sendo assim, quando a integridade da rede

¹ TOU – Time-Of-Use

² RTP – Real Time Pricing

³ CPP – Critical Peak Pricing

eléctrica fica comprometida ou quando o valor da produção de energia eléctrica é muito elevado, são enviados sinais aos consumidores de forma a que estes reduzam os seus consumos. Caso os consumidores não o façam, sofrem determinadas penalizações [12].

Estes programas podem ser de quatro tipos [12] [15] [16]:

- **Controlo Directo de Carga:** Nestes programas o operador de rede tem controlo directo sobre algumas cargas dos consumidores, podendo, em determinadas ocasiões, ligá-las ou desliga-las, sendo o consumidor avisado num curto espaço de tempo. Estes programas aplicam-se geralmente a clientes residenciais ou de comércio em pequena escala.
- **Serviço Passível de Interrupção:** Os consumidores comprometem-se a reduzir os seus consumos, quando tal lhes for solicitado, para um valor pré-acordado. Caso ocorra incumprimento do estabelecido, o consumidor sofre uma penalização. Os consumidores que utilizam este tipo de programa são consumidores em grande escala, como ocorre no caso de grandes indústrias ou comércios.
- **Programas de Emergência:** São programas em que se oferecem incentivos aos consumidores para que reduzam os seus consumos em situações de emergência.
- **Programas de Licitação / Recuperação:** Estes programas apenas são aplicáveis a grandes consumidores, que são encorajados a comprar energia eléctrica no mercado grossista e a reduzir os consumos quando tal for necessário. Desta forma, a redução total disponível entra no processo de determinação do preço da energia eléctrica.

2.1. Caracterização de cargas

O conhecimento dos tipos de electrodomésticos e da sua contribuição no consumo total eléctrico de uma residência torna-se essencial para a construção de um simulador dos consumos residenciais e, conseqüentemente para o estudo da possibilidade da implementação de medidas de GP. Cada electrodoméstico deve ser caracterizado individualmente uma vez que os comportamentos diferem, podendo assim as potenciais medidas de gestão a aplicar a cada um deles também distinta. No caso do sector residencial, em Portugal, as cargas de frio alimentar (frigoríficos e combinados), de aquecimento ambiente, conjuntamente com o AQS eléctrico, congelador, máquinas de lavar e secar e arrefecimento ambiente representam cerca de 64% de consumo controlável [10], o que deixa antever um elevado potencial e flexibilidade no que respeita à aplicação de medidas de GP.

As cargas residenciais podem então ser classificadas como cargas controláveis ou não controláveis.

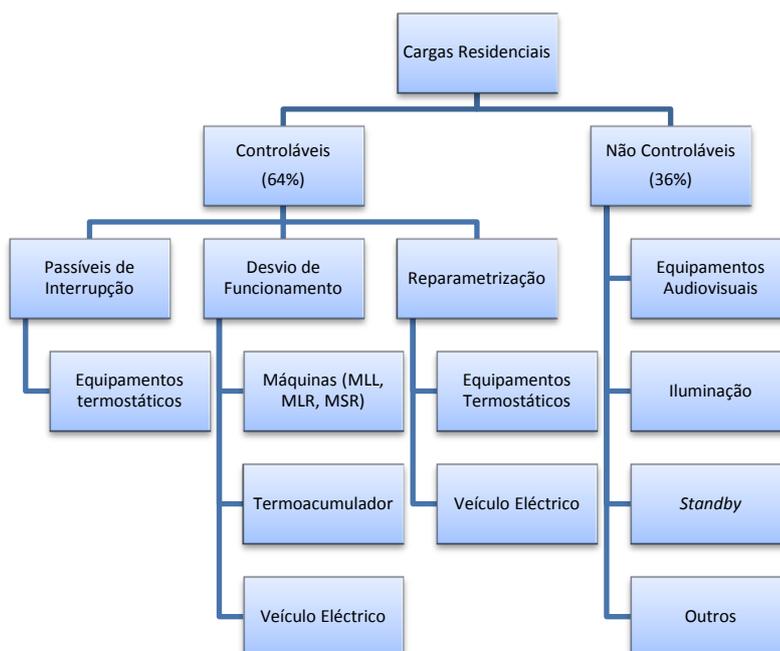


Figura 7: Classificação dos diferentes tipos de cargas de acordo com as estratégias de controlo que se podem aplicar [12]

As cargas controláveis são aquelas em que é possível a aplicação de medidas de GP. Como o seu funcionamento varia de carga para carga, as cargas controláveis podem ser divididas em três classes distintas: cargas passíveis de interrupção, desvio de funcionamento e reparametrização.

Cargas passíveis de interrupção são aquelas cujo funcionamento se pode interromper por breves instantes, sem que tal prejudique a sua performance. É o que ocorre no caso dos equipamentos termostáticos. Nas cargas termostáticas – climatização, frio alimentar, termoacumulador – o funcionamento tem por base um termostato: por exemplo, um frigorífico pode estar programado para manter a temperatura no seu interior a 2°C e, para evitar que este esteja sempre a ligar e desligar, existe uma banda morta. Se a banda morta for de 1°C, quando a temperatura no seu interior for de 2,5°C, o compressor é activado, arrefecendo a temperatura para 1,5°C. No entanto, se o SEE, por algum motivo alheio, estiver comprometido, a operadora pode enviar um sinal, impedindo que o frigorífico funcione por breves instantes. Uma vez que estes possuem um bom isolamento térmico, não haverá um acréscimo muito significativo da temperatura interior (mesmo que ocorra um aumento de 1°C ou 2°C, tal não se torna problemático para a conservação dos alimentos, pois a interrupção é por apenas breves instantes).

Numa tarifa multi-horária, o consumidor pode controlar estas cargas, de forma a obter poupanças na sua factura. Se ele controlar estas cargas para que no final de um determinado período do dia elas se desliguem (por exemplo, na última meia hora do período de cheias numa tarifa bi-horária), o consumo de energia, apesar de superior, será realizado apenas no período em que o preço do kWh é inferior.

Os desvios de funcionamento só podem ser aplicados a cargas cujo funcionamento possa ser desviado para um outro instante do dia. As cargas aptas para este tipo de controlo são as que apresentam capacidade de armazenamento de energia (termoacumuladores) ou as que possuem um funcionamento por ciclos [máquinas de lavar roupa (MLR), loiça (MLL) e secar roupa (MSR)].

Cargas que podem ser reparametrizadas, ou seja, aquelas em que é possível alterar a sua parametrização, têm como exemplos os equipamentos termostáticos e os veículos eléctricos. É o caso de um ar condicionado parametrizado para manter a temperatura no interior de uma sala a 25°C, quando o controlador altera a temperatura para 25,5°C, ninguém se irá aperceber da diferença. Contudo, a energia consumida por este será inferior.

Na categoria de cargas não controláveis estão inseridas as cargas que não se podem alterar no que respeita ao controlo do seu funcionamento, como é o caso do forno eléctrico, equipamentos audiovisuais, iluminação, *standby*, e outros equipamentos diversos (ferro de engomar, torradeiras, entre outros).

No âmbito desta dissertação, as cargas não controláveis são denominadas por carga base (CB). Nesta carga base estão ainda incluídos os equipamentos termostáticos de frio alimentar.

3. Tratamento de dados

O SIMCAR foi desenvolvido de modo a ser capaz de reproduzir os diagramas de carga de consumidores residenciais tendo em consideração, não só a variabilidade desses consumos, mas também os impactos da aplicação das estratégias de GP. Como já foi referido, no sector residencial existem diferentes cargas com características diversas que se prestam a diferentes acções de controlo. Daí advém a opção de utilização de uma abordagem *bottom-up* na construção desta ferramenta. Existe já bastante trabalho de modelação e simulação desenvolvido quer em aplicação a cargas individuais quer em termos de procura agregada [5] – [9], [17], mas como referido atrás, sem as características que acham interessantes e necessárias a uma ferramenta deste tipo. De modo a poder simular a procura das diferentes cargas, sempre que possível ou necessário, é utilizada, no SIMCAR, informação existente acerca do funcionamento das mesmas, a qual é obtida através do recurso a fontes de dados que podem ser distintas para diferentes cargas. Por exemplo, no caso das MLR, MSR e MLL, como o funcionamento é caracterizado pelos seus ciclos e programas, é necessário ter acesso a este tipo de dados e à sua distribuição ao longo do dia. Este último aspecto pode ser obtido de duas formas: através de relatórios ou estudos pré-existentes, ou através da aplicação de técnicas de *clusterização* utilizando evolução diferencial ou *data mining* a uma base de dados de consumos destas cargas. De referir que as

fontes de dados tanto podem ter sido construídas a partir de cargas semelhantes, permitindo, neste caso, a obtenção de regimes de funcionamento aproximados para as cargas a simular, como podem ser dados das cargas cujo funcionamento se pretende reproduzir numa ferramenta computacional. Desta forma será possível avaliar o impacto de diferentes acções de controlo.

Para além dos dados referidos atrás, no que respeita às cargas termostáticas, como é o caso do ar condicionado e termoacumuladores, é possível simular os consumos para determinadas situações através do recurso a MFB, que reproduzem os fenómenos físicos que ocorrem neste tipo de cargas. No caso do SIMCAR, foi utilizado o MFB apresentado em [18].

As técnicas de DM/ED foram inicialmente desenvolvidas com o objectivo de calcular a distribuição de probabilidade dos consumos da carga base para cada instante temporal no decorrer do dia. Para as cargas anteriores, que são tratadas individualmente, os consumos podem ser obtidos através de recolhas directas, de estudos de caracterização de cargas já existentes ([4] [10] [19]) ou através de MFB [18], enquanto que para a carga base tal já não se verifica. Apesar dos dados utilizados para a criação dos consumos serem obtidos através de recolhas, estes necessitam ainda de ser processados, de forma a ser possível adquirir conhecimento acerca da evolução da procura deste tipo de consumo. Para tal, são aplicados os algoritmos de DM e ED.

Uma vez criado o modelo destes algoritmos, é possível fazer uso dos mesmos para executar as restantes tarefas, descritas anteriormente para as outras cargas.

3.1. *Data Mining*

Data Mining (DM) pode ser definida como o processo de descoberta de padrões e tendências verificados em grandes quantidades de dados, não só através do uso de técnicas de reconhecimento de padrões como também de técnicas estatísticas e matemáticas. Desta forma é possível detectar informação extremamente útil em bases de dados extensas, que, numa primeira análise, aparentam não apresentar qualquer informação compreensível [20]. Na Figura 8 estão representadas os processos aplicados a uma base de dados por forma a retirar da mesma qualquer informação útil.

De um modo geral, pode referir-se que DM engloba três fases distintas: o armazenamento de dados, o tratamento (mineração) dos dados e a apresentação dos resultados [21].

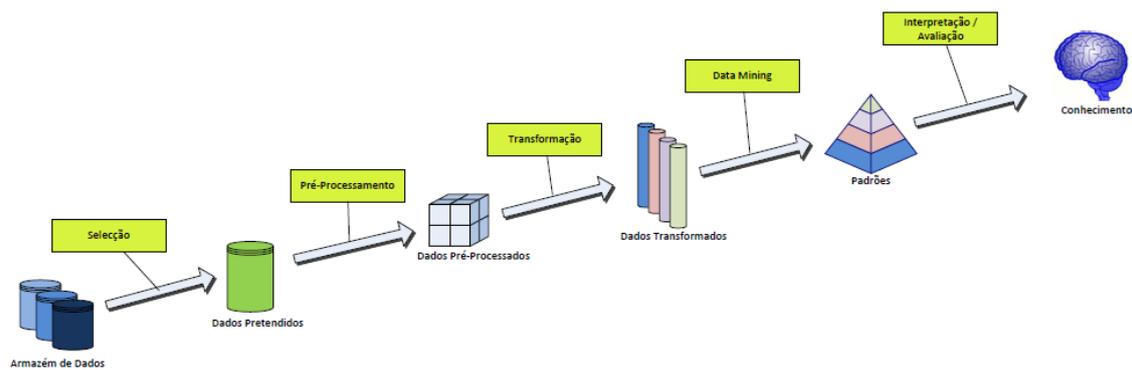


Figura 8: Representação dos processos aplicados a uma base de dados para retirar informação útil.

As técnicas de DM podem ser utilizadas para descrição, previsão, classificação, agrupamento (*clusterização*) e associação [20].

No entanto, antes da aplicação de qualquer uma destas técnicas é necessário proceder à limpeza dos dados. Quanto maior for a base de dados, maior será a probabilidade de existência de valores obsoletos, valores incompletos ou inconsistentes e a existência de *outliers*⁴ [20].

A técnica utilizada pelo SIMCAR foi a de *clusterização*, cujo objectivo é o de identificar conjuntos de elementos dentro de uma base de dados. Os elementos de cada conjunto identificado deverão ser homogéneos (ou similares) entre si e heterogéneos (ou dissimilares) dos restantes. Uma vez que foi possível obter conhecimento acerca da contribuição das diferentes cargas para o consumo residencial através de recolhas de dados, podem identificar-se, com recurso ao *clusterização*, características próprias de cada carga bem como adquirir conhecimento acerca do funcionamento das mesmas. Esta técnica revelou-se de especial importância para a caracterização dos consumos de carga base onde, com base em poucos dados, é possível reproduzir o comportamento deste tipo de carga.

A semelhança entre os elementos é calculada com base na distância entre eles, existindo diversas formas de calcular estas distâncias. A forma mais comum e mais simples é a distância Euclidiana [20] [22].

Outras distâncias a que se pode igualmente recorrer são a City-Block, Chebychev, Angular, Canberra, Mahalanobis, Hellinger, Bhattacharyya entre outras.

Nas situações em que os atributos que caracterizam os dados sejam de escalas e tipos muito diferentes entre si, é recomendável a que se proceda a uma normalização prévia destes dados. Os dois tipos de normalização mais utilizados são Normalização Min-Máx e *Z-Score Standardization*. [20]

⁴ *Outlier* é um valor que diverge acentuadamente dos restantes valores da amostra recolhida

Normalização Min-Máx:
$$X^* = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (1)$$

Z-Score Standardization:
$$X^* = \frac{X - \mu}{\sigma} \quad (2)$$

μ : Média do vector X

σ : Desvio Padrão

Após seleccionada a melhor forma de calcular a diferença entre os dados, procede-se à escolha do algoritmo de *clusterização*. Esta escolha depende do tipo de dados disponíveis e do objectivo pretendido. Os algoritmos de *clusterização* podem ser separados em dois tipos: hierárquicos e não hierárquicos ou de particionamento. [23] [24]

Os algoritmos hierárquicos podem, por sua vez, ser de aglomeração ou de divisão. Nos algoritmos de aglomeração, inicialmente todos os elementos formam um *cluster*. De seguida, os dois *clusters* mais similares são aglomerados num só. Este procedimento é repetido até ser obtido apenas um *cluster* que engloba todos os elementos. Nos de divisão, o processo é o inverso. Inicialmente todos os elementos estão aglomerados num só *cluster*, o elementos mais dissimilar dos restantes é separado e verifica-se a formação de um novo *cluster* [20] [23].

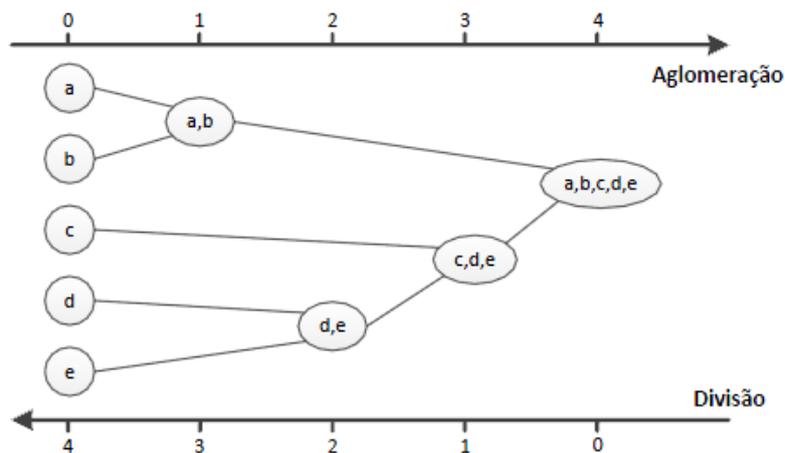


Figura 9: Comparação entre os algoritmos hierárquicos de aglomeração e de divisão [20] [23].

Por sua vez, os algoritmos de partição têm por base a divisão da base de dados em k *clusters*, sendo que cada um apresenta, no mínimo, um elemento e cada elemento pertence a um só *cluster*. Estes algoritmos criam uma partição inicial e iteram até um dos critérios de paragem se verificar.

No SIMCAR foi utilizado o algoritmo de partição *k-means*. Este algoritmo consiste no cálculo da diferença entre todos os pontos da amostra com os *centroids* actuais, atribuindo cada elemento ao *cluster* cujo *centroid* se encontra mais próximo. O *centroid* é, então, o valor médio de todos os pontos existentes no *cluster*, sendo ele que define o *cluster*. Tal como acontece com a grande maioria dos algoritmos de particionamento, neste algoritmo é necessário definir *a-priori* o número de *clusters*.

No Anexo I pode-se observar a aplicação do algoritmo *k-means* a uma base de dados de elementos distribuídos aleatoriamente de forma uniforme, e a outra base de dados de elementos distribuídos através de uma distribuição Normal. As desvantagens deste algoritmo baseiam-se no facto de ser necessário definir o número de *clusters* previamente, de ser sensível a *outliers* e da escolha dos *centroids* iniciais poder afectar os resultados finais, não sendo assim possível obter a solução óptima. No entanto, a sua fácil implementação e o rápido desempenho computacional e flexibilidade que este algoritmo apresenta, fazem dele o ideal para ser utilizado pelo SIMCAR. Uma vez que os dados foram pré-processados manualmente, a existência de *outliers* foi eliminada, não sendo assim o algoritmo afectado por estes.

3.1.1. Determinação do número de *clusters*

Tal como já foi referido anteriormente, uma das desvantagens dos algoritmos de DM de particionamento é a necessidade de se conhecer *a-priori* o número *k* de *clusters* a utilizar. Uma das formas de fazer esta identificação é calcular o erro quadrático médio para diferentes valores de *k*. O erro quadrático é inversamente proporcional ao número de *clusters*, ocorrendo o maior valor quando todos os elementos pertencerem apenas a um *cluster*, e tomando o menor valor quando existirem tantos *clusters* como elementos.

O erro quadrático é dado por:

$$E(C) = \sum_{j=1}^k \sum_{x_i \in C_j} |x_i - m_i|^2 \quad (3)$$

Onde *C* é o conjunto de *k clusters*, sendo definido como $C = \{C_1, \dots, C_k\}$, x_1, \dots, x_k elementos correspondentes a cada *cluster* e m_1, \dots, m_k os *centroids* de cada *cluster*.

Através da divisão do erro total pelo valor médio do erro, obtém-se o erro absoluto em percentagem. Este erro não deverá ser inferior a 1% [25].

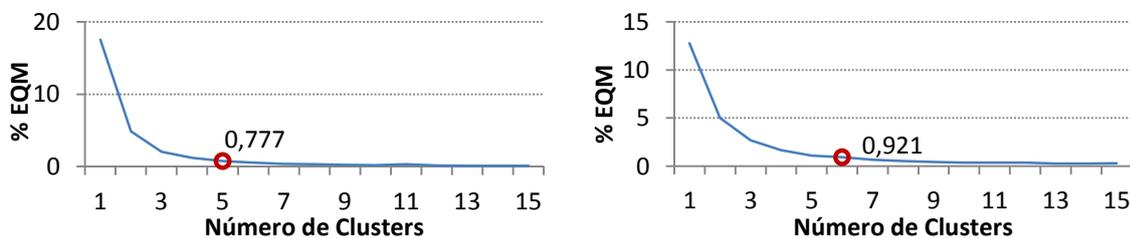


Figura 10: Obtenção do número de *clusters* ideal para uma base de dados de 1.000 elementos aleatórios (esquerda) e uma base de dados de 10.000 elementos aleatórios.

Na Figura 10 foi aplicado o método do erro quadrático para a obtenção do número *k* de *clusters*. Para o primeiro exemplo foi utilizada uma base de dados de 1.000 números aleatórios e no segundo foi utilizada uma base de dados de 10.000 números aleatórios.

3.2. Evolução Diferencial

Para além das técnicas de DM para tratamento dos dados foi também adoptado um método de *clusterização* baseado em evolução diferencial. Inicialmente proposto por Rainer Storn e Kenneth Price [26] com o intuito de otimizar funções não lineares com variáveis contínuas, este algoritmo pode também ser utilizado como ferramenta de identificação de agrupamentos de dados [27] [28].

Um algoritmo evolutivo tem por base a teoria da evolução Darwin. Segundo Darwin, através da selecção natural, os seres vivos de cada espécie mais aptos têm maiores probabilidades de sobreviverem, aumentando assim as probabilidades dos seus genes passarem para a próxima geração. É através da existência de mutações genéticas que se torna possível a biodiversidade entre indivíduos da mesma espécie, criando assim maiores competências para adaptação às frequentes alterações ambientais. Os algoritmos evolutivos inspiram-se então neste processo natural. As iterações destes algoritmos são analogias das gerações existentes na natureza. Na passagem de uma geração para outra os indivíduos podem sofrer recombinações ou mutações, replicando de certa forma o que se passa na natureza. Através do princípio de sobrevivência do mais apto (selecção natural), a(s) solução(ões) que maximiza(m) a função objectivo é seleccionada e transferida para a próxima geração. Estas iterações repetem-se até ser atingido um critério de paragem [26] [29].

3.2.1. Algoritmo Geral

Considerando um problema genérico de optimização não linear com variáveis contínuas, o algoritmo de evolução diferencial é então formulado da seguinte forma:

$$x^* = \arg \min f(x) \quad (4)$$

$$\text{Sujeito a: } \begin{cases} x \in \mathbb{R}^n \\ g_1(x) \leq 0 \\ \dots \\ g_m(x) \leq 0 \end{cases} \quad (5)$$

X_t é o conjunto de soluções candidatas ao problema.

$X_t = \{x_{t,i}\}$, $i = \{1, \dots, N\}$, em que i é o índice do indivíduo na população e t o índice da geração. $x_{t,i}$ é então um vector coluna de n linhas, em que n representa o número de variáveis do problema .

$$x_{t,i} = \begin{bmatrix} x_{t,i,1} \\ x_{t,i,2} \\ \dots \\ x_{t,i,n} \end{bmatrix} \quad (6)$$

O algoritmo de evolução diferencial utiliza vectores-diferença ($v_{t,i}$) criados a partir de vectores da própria população. Estes vectores-diferença são criados através da selecção aleatória de dois indivíduos e posterior cálculo da diferença entre eles, sendo essa diferença multiplicada por um factor de escala R e somada a um terceiro indivíduo seleccionado também aleatoriamente. Matematicamente verifica-se:

$$v_{t,i} = x_{t,r_1} + R(x_{t,r_2} - x_{t,r_3}), r_1, r_2 \text{ e } r_3 \in \{1, \dots, N\} \quad (7)$$

$v_{t,i}$ representa então a i -ésima solução mutante. Usando este procedimento obtém-se então a população mutante V_t .

Obtida a população mutante, os indivíduos mutantes podem ser recombinados com os indivíduos da população corrente, X_t . A probabilidade de ocorrência de recombinações CR (Crossover Rate) usualmente é um valor definido *a-priori*.

Sendo p um valor aleatório, com distribuição uniforme entre $[0,1]$, verifica-se:

$$u_{t,i,j} = \begin{cases} v_{t,i,j}, & \text{se } p \leq CR \vee j = \delta_i \\ x_{t,i,j}, & \text{caso contrário} \end{cases} \quad (8)$$

Em que δ_i é um valor aleatório $\in \{1, \dots, n\}$ que serve para garantir que irá existir pelo menos uma recombinação.

Finalmente é avaliado o valor da função-objectivo para cada solução teste $u_{t,i}$, comparando-o com o valor obtido para a solução corrente $x_{t,i}$. Desta forma é descartada a pior solução, passando a melhor para a próxima geração X_{t+1} .

3.2.2. Algoritmo de ED utilizado para *clusterização*

O algoritmo implementado é semelhante ao algoritmo geral de ED, com excepção dos valores da probabilidade de ocorrência de recombinações (CR) e do factor escala R .

A probabilidade de ocorrência de recombinações CR , em vez de ter um valor fixo, tem um valor dinâmico, sendo a variação deste valor dada por:

$$CR = (CR_{max} - CR_{min}) \times \frac{(MAXIT - iter)}{MAXIT} \quad (9)$$

Onde CR_{max} e CR_{min} são, respectivamente, os valores máximos e mínimos que CR pode tomar, $iter$ o valor da iteração actual e $MAXIT$ o número máximo de iterações. Desta forma, quando se está nas gerações iniciais, o valor de CR é relativamente superior, o que permite a

ocorrência de mais recombinações, aumentando assim a exploração da solução óptima do algoritmo. À medida que as gerações vão aumentando, como se caminha cada vez mais para a solução óptima, diminui-se a probabilidade de ocorrência de recombinações e, conseqüentemente, a obtenção de piores resultados [28].

O factor escala R também é um dos valores alterados, de forma a ser dinâmico. Este valor é calculado para cada iteração aleatoriamente através de uma distribuição uniforme, estando compreendido entre $[0.5,1]$. Desta forma o valor médio do factor escala é $0,75$, conseguindo-se assim que exista uma maior amplificação do vector-diferença, aumentando assim a diversidade da população [27] [28].

A constituição de cada cromossoma está representada na Figura 11, onde t corresponde ao número da geração, M ao número do cromossoma, k ao número do *cluster* e C é o valor real normalizado do *centroid* de cada *cluster*.

$x_{t,1}$	$C_{t,1,1}$	$C_{t,1,2}$...	$C_{t,1,k}$
$x_{t,2}$	$C_{t,2,1}$	$C_{t,2,2}$...	$C_{t,2,k}$
...
$x_{t,M}$	$C_{t,M,1}$	$C_{t,M,2}$...	$C_{t,M,k}$

Figura 11: Representação dos cromossomas

De forma a ilustrar melhor o funcionamento do algoritmo, estão representadas nas seguintes figuras as diferentes etapas deste.

Para criar os cromossomas iniciais que irão formar a primeira geração corrente (X_t), são criados N cromossomas, sendo de seguida definidos aleatoriamente, para cada cromossoma, k *centroids* (Figura 12, esquerda). Uma vez escolhidos os *centroids* iniciais, os elementos da base de dados são atribuídos ao *cluster* mais próximo. Esta atribuição é feita para todos os N cromossomas.

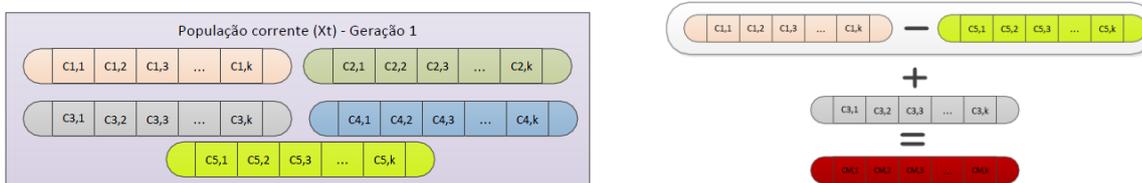


Figura 12: Ilustração da população corrente (esquerda) e do cálculo de um vector-diferença (direita)

Uma vez definida a população corrente da primeira geração, calcula-se o vector-diferença, sendo este cálculo feito como descrito em 3.2.1 e representado na Figura 12 (direita). Este processo é repetido para todos os cromossomas. O conjunto dos vectores que foram obtidos aplicando os vector-diferença forma a população mutante. Através de recombinação da

população corrente com a população mutante obtém-se a população teste. Estes processos estão representados na Figura 13.

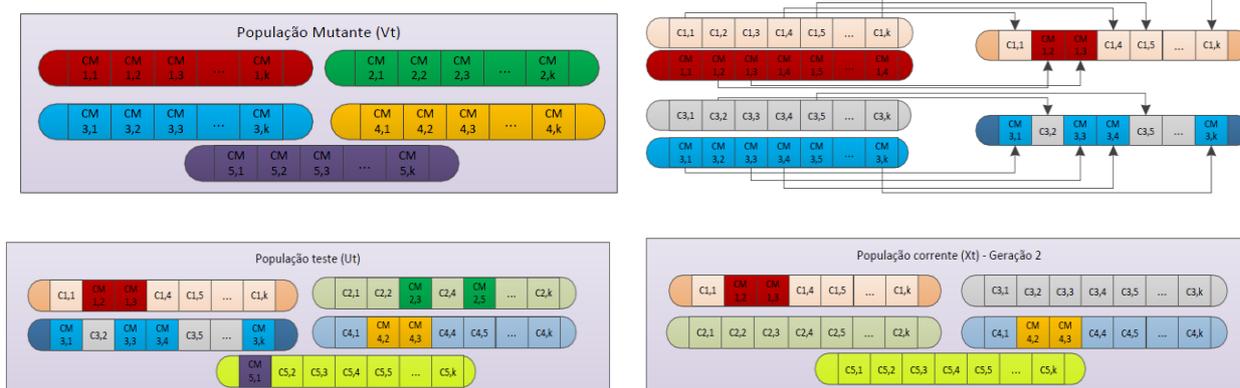


Figura 13: Representação da população mutante (cima, esquerda), do processo de recombinação entre população mutante e população corrente (cima, direita), da população teste (baixo, esquerda) e da população corrente da 2ª Geração (baixo, direita).

Uma vez obtida a população teste, os cromossomos são comparados com os cromossomos da população corrente, sendo os cromossomos que melhor optimizam a função objectivo que sobrevivem para a próxima geração. Este processo é repetido até uma condição de paragem se verificar ou o número máximo de iterações ser atingido.

3.3. Teste de silhueta

O teste da silhueta foi desenvolvido por Rousseeuw em 1987 [30], e serve para avaliar a qualidade da *clusterização* realizada. Neste teste é avaliada a coesão dos elementos de cada *cluster* bem como a separação entre *clusters*. A formula de cálculo da silhueta de cada elemento i é dada por

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}; s(i) \in [-1; 1] \quad (10)$$

Onde $a(i)$ representa a dissimilaridade média entre o elemento i e todos os elementos pertencentes ao mesmo *cluster* e $b(i)$ a dissimilaridade entre o elemento i e o *centroid* mais perto de i , com excepção do *centroid* do *cluster* a que i está atribuído.

Idealmente deve-se obter um valor de $a(i)$ próximo de zero (o que significa que o elemento i foi atribuído ao *cluster* certo) e um valor de $b(i)$ elevado (o que significa que o *centroid* do *cluster* mais próximo se encontra distante).

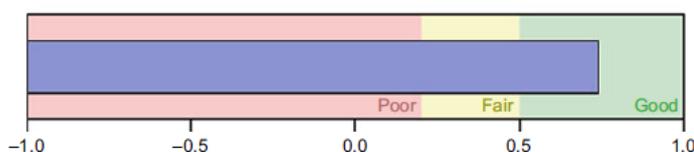


Figura 14: Variação da medida de silhueta e respectiva classificação da qualidade dos *clusters* [30]

Quanto mais próximo de 1 for este valor, melhor será a qualidade da *clusterização*. Na Figura 14 está apresentada a classificação da qualidade dos *clusters* obtidos de acordo com o valor da silhueta [30].

3.4. Análise Comparativa entre *k-means* e ED

Para comparar os dois algoritmos, foi gerada uma base de dados de 5.000 elementos normalizados entre 0 e 1, representativos dos diversos consumos para um dado intervalo de tempo, sendo de seguida ambos os algoritmos aplicados a esta base de dados, e através do teste da silhueta, foi possível comparar a qualidade dos *clusters* formados.

Nos testes realizados, foram também variados o número de *clusters* e de cromossomas nos algoritmos. Os resultados podem ser observados na Tabela 2, e as figuras encontram-se no Anexo II.

Tabela 2: Comparação dos resultados obtidos pelo teste da silhueta para os diferentes casos

Algoritmo	Média	Máximo	Mínimo	Tempo
ED 5 Clusters; 5 Cromossomas	0,72562	0,93075	-0,0222277	37,40 seg.
ED 5 Clusters; 10 Cromossomas	0,72562	0,93193	-0,026595	96,79 seg
ED 10 Clusters; 5 Cromossomas	0,70267	0,93134	-0,090848	29,6348 seg
ED 10 Clusters; 10 Cromossomas	0,69995	0,93041	-0,17043	61,35 seg
<i>k-means</i> 5 Clusters	0,72428	0,92992	-0,0142	2,52 seg
<i>k-means</i> 10 Clusters	0,70422	0,92983	-0,034969	4,51 seg
<i>k-means</i> 15 Clusters	0,698	0,92766	-0,057633	3,31 seg

Através da aplicação dos dois algoritmos, obtiveram-se resultados muito semelhantes, com excepção do tempo de execução, onde se pode verificar que o algoritmo *k-means* se demonstra ser mais eficiente. No algoritmo de evolução diferencial, verifica-se que quando o número de cromossomas aumenta, o tempo de execução deste também aumenta muito significativamente.

O algoritmo *k-means* possui um problema, que é a possibilidade de em determinados casos convergir para uma solução óptima local, uma vez que a solução final deste pode ser influenciada pelos valores iniciais dos *centroids*.

Na teoria, a vantagem do algoritmo de evolução diferencial quando utilizado como uma técnica de *clusterização* relativamente ao algoritmo *k-means*, é precisamente prevenir que o algoritmo reconheça uma solução óptima local como sendo a solução final. Ao existirem diversos cromossomas, a área de procura do algoritmo é maior. Contudo, numa situação em que se está a realizar uma *clusterização* unidimensional, o algoritmo *k-means* não demonstra ser

afectado por este problema. Devido à sua simplicidade, o tempo de execução deste último é muito inferior ao algoritmo de ED.

3.5. Conversor de resoluções temporais

De forma a tornar possível um utilizador que tenha em sua posse uma base de dados de diagramas de carga numa resolução temporal distinta da resolução temporal que se pretende utilizar na simulação, houve a necessidade de criar um conversor de resoluções temporais.

Quando o objectivo é passar de uma resolução temporal inferior para uma superior (isto é, passar de, por exemplo 15 em 15 minutos para uma resolução de minuto a minuto), o conversor expande de forma equitativa o consumo registado em cada intervalo de 15 minutos para intervalos de minuto a minuto. Se o objectivo for o inverso, supondo o exemplo referido, o conversor calcula a média do consumo em cada 15 minutos.

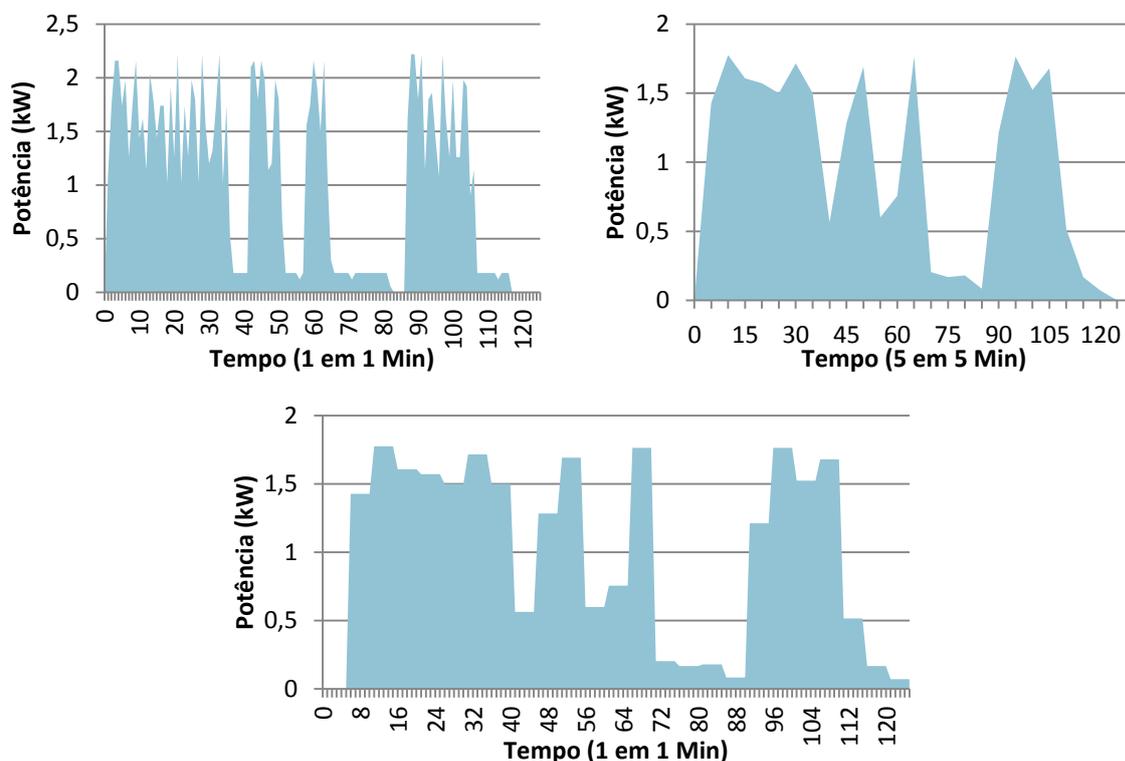


Figura 15: Conversão do diagrama de um ciclo de uma MSR.

Na Figura 15 está exemplificado um exemplo de conversões. Nesta situação, converteu-se um ciclo de uma MSR que se encontrava na resolução temporal unitária (diagrama superior esquerdo) e converteu-se para a resolução de cinco em cinco minutos (diagrama superior direito). De forma a demonstrar a situação inversa, reconverteu-se este último diagrama para a resolução temporal inicial (diagrama inferior).

4. SIMCAR – SIMulador de Cargas Residenciais

A construção do diagrama de carga residencial pelo SIMCAR é realizada através de uma estratégia de *bottom-up*, em que os consumos de cada carga são gerados individualmente e posteriormente agregados para obtenção do diagrama final. Sendo assim, para gerar diagramas de carga de diversas habitações, o SIMCAR realiza simulações individuais para cada tipo de carga, sendo, no final, agregados os consumos por tipo de carga. É dado a cada carga um tratamento diferenciado de acordo com o tipo de dados ou informação existentes e de acordo com as características de cada uma.

As cargas diferenciadas pelo simulador são todas as passíveis de serem sujeitas à aplicação de estratégias de GP, com excepção do frio alimentar. O SIMCAR tem a capacidade de simular os consumos em diferentes resoluções temporais, conforme o pretendido pelo utilizador.

Como foi referido, cada tipo de carga tem um comportamento e características distintas, pelo que a forma como a simulação é realizada também varia de carga para carga.

A Figura 16 representa os dados de entrada e dados de saída dos diferentes processos a para a simulação das diferentes cargas. No caso da CB, recorre-se à identificação de padrões em recolhas perviamente efectuadas, através da aplicação de um dos algoritmos de *clusterização*, *k-means* ou ED, obtendo-se desta forma a distribuição de probabilidade dos consumos para cada intervalo de integração⁵. Com esta distribuição de probabilidade é possível construir o diagrama de CB de cada consumidor.

Já para as cargas caracterizadas pelo seu funcionamento cíclico, como é o caso das MLL, MLR e MSR, a simulação é feita de forma distinta. Neste tipo de cargas, é necessário saber, em primeiro lugar, qual a distribuição de probabilidade das mesmas ao longo do dia bem como o número de vezes que funcionam em cada dia, para de seguida se poderem determinar qual(ais) os instantes em que elas operam. Esta distribuição de probabilidade pode ser obtida através de estudos ou relatórios já existentes, como é o caso de [10] e [19], ou através do recurso aos algoritmos de *clusterização*, os quais são aplicados a uma base de dados de medições reais. Uma vez determinados os instantes, é necessário saber quais os tipos de programas de funcionamento a utilizar, que são obtidos através de estudos de caracterização destas cargas. É ainda necessário fornecer uma base de dados de ciclos de funcionamento para os diferentes programas, para que estes sejam utilizados na construção dos diagramas de carga.

Por fim, no caso da simulação das cargas termostáticas é necessário possuir uma base de dados de diagramas diários. Esta base de dados pode ser construída através de recolhas ou

⁵ Intervalo de integração corresponde à resolução temporal das medições. Se estas foram feitas numa resolução de 5 em 5 minutos, um dia possui 288 intervalos de integração.

através de um MFB. Uma vez obtida a base de dados, pode-se utilizar os diagramas diários resultantes da simulação, ou então, no caso dos AC e TR, recorrer às técnicas de *clusterização* para criar a distribuição de probabilidade de os equipamentos se encontrarem em funcionamento.

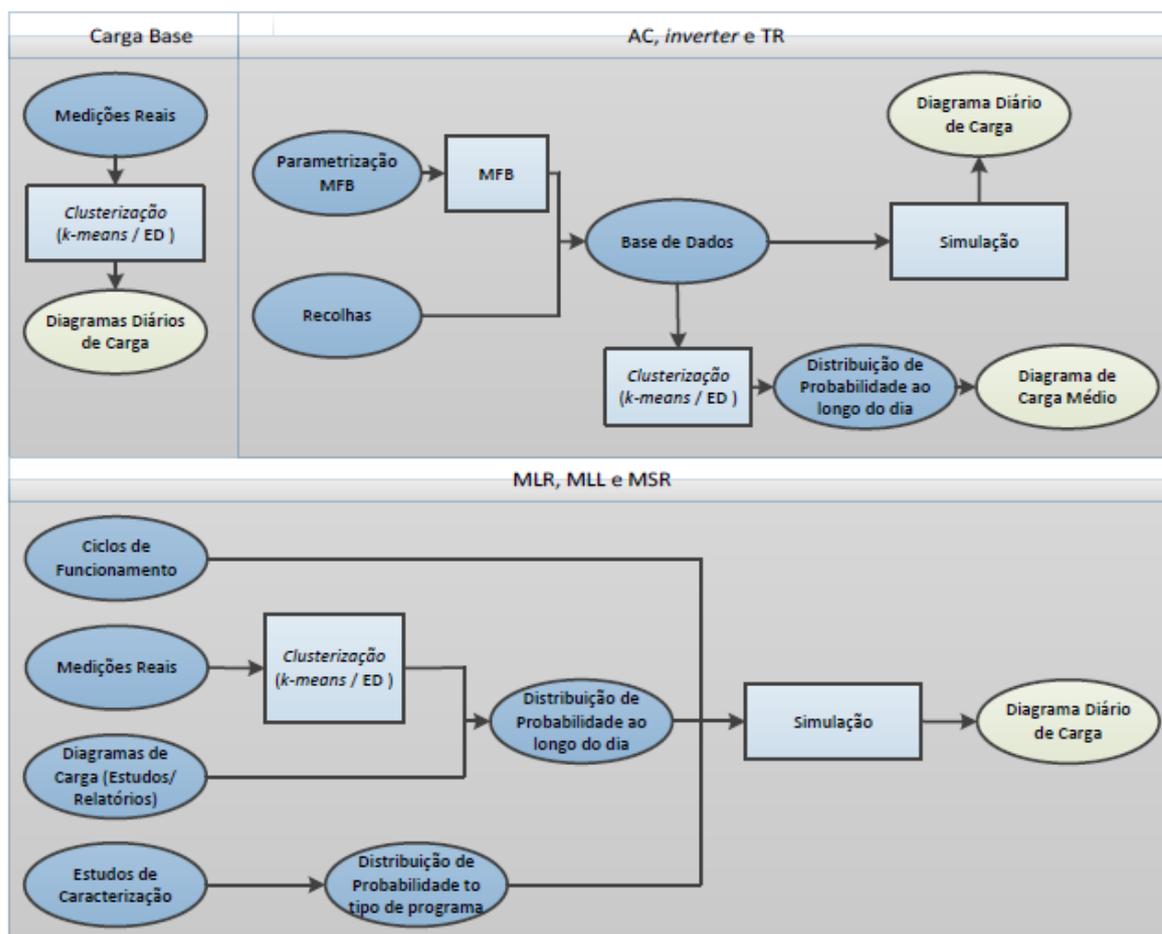


Figura 16: Dados de entrada do SIMCAR

Após executado, o SIMCAR fornece os consumos desagregados das cargas, sendo os resultados apresentados graficamente e guardados num ficheiro de Excel. Neste ficheiro, é possível ter acesso aos consumos de cada tipo de carga, os quais se encontram organizados pelas respectivas folhas de Excel (ver Figura 17), sendo cada coluna correspondente ao consumo total de cada tipo de carga por dia.



Figura 17: Organização do ficheiro Excel de resultados do SIMCAR

4.1. Carga Base

Tal como já foi abordado anteriormente, a carga base (CB) refere-se às cargas que não podem ser controladas (microondas, chaleiras eléctricas, fornos, equipamentos audiovisuais, etc.) e ao frio alimentar.

A simulação da CB consiste num diagrama diário de carga realístico, construído a partir de uma base de dados com diagramas diários recolhidos em campanhas de medição de consumos, na qual se faz a identificação da distribuição de probabilidade do consumo em cada intervalo de integração. Para isso é necessária a aplicação de uma técnica capaz de identificar padrões na base de dados. É nesta ocasião que se recorre a um dos dois algoritmos de *clusterização*: *k-means* ou *evolução diferencial*

Podemos então definir os seguintes passos:

Passo 1: Escolha de um dos dois algoritmos existentes e aplicá-lo numa base de dados de consumos de CB, obtendo a distribuição de probabilidade dos consumos ao longo do dia.

Passo 2: Após obtidas as distribuições de probabilidade, geração de números aleatórios para determinar a potência a ser consumida em cada instante.

Passo 3: Repetir os passos anteriores para o número de habitações pretendidas.

Passo 4: Agregar os consumos individuais para se obter o diagrama de carga global da CB.

A base de dados original é construída com recurso a recolhas realizadas. No nosso caso, usaram-se dados recolhidos pelo INESC, no âmbito de um projecto internacional [31]. Neste estudo foram recolhidos dados de diversos electrodomésticos, pertencentes a 66 consumidores, durante 15 dias. Dos 66 consumidores, em 22 foi possível, identificar de forma clara a contribuição dos consumos de CB, obtendo-se um total de 330 diagramas de carga diários.

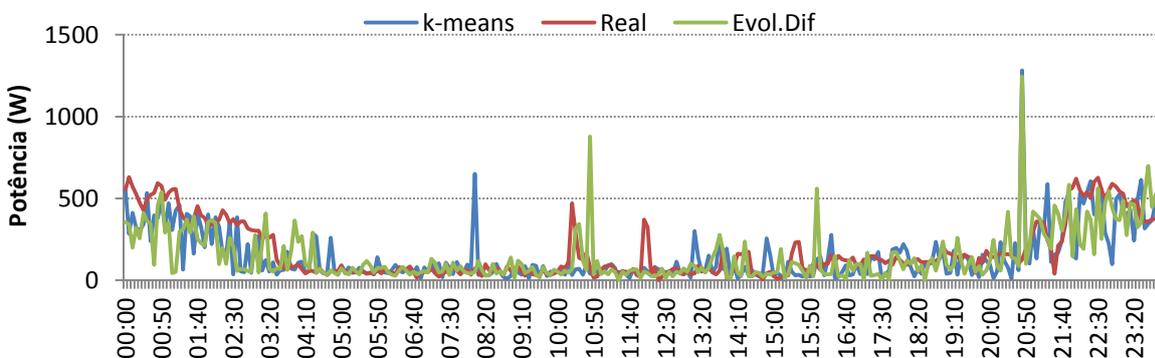


Figura 18: Consumo de um consumidor de um dia, através de medições reais e através dos algoritmos de *clusterização*

Na Figura 18 podemos observar o consumo obtido para um dia, através da aplicação dos algoritmos de *clusterização* a 14 dias monitorizados de um consumidor, podendo os diagramas destes 14 dias monitorizados ser observados no Anexo I/Anexo III.

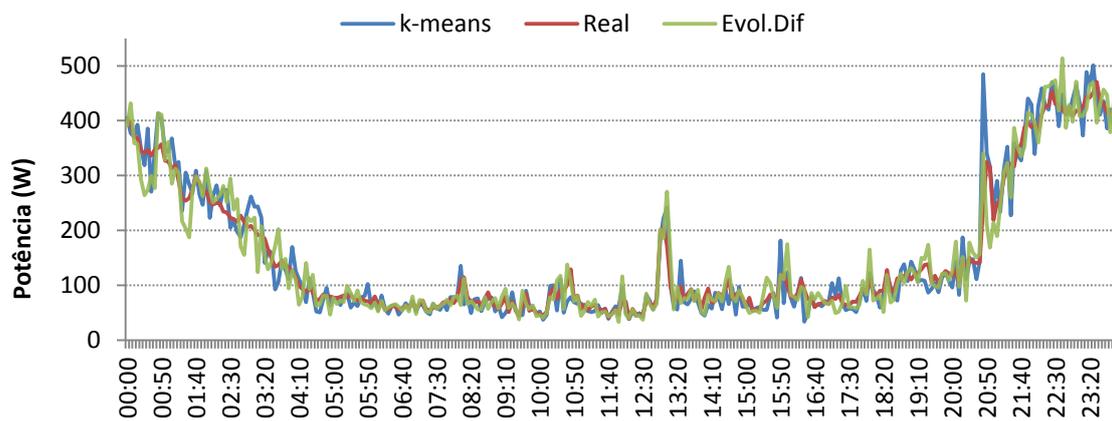


Figura 19: Consumo médio de um consumidor ao longo de 14 dias, obtido através de medições reais e da aplicação dos algoritmos de *clusterização* a essas mesmas medições

Foram ainda construídos 14 diagramas de carga diários deste consumidor com ambos os algoritmos. Na Figura 19 encontra-se representada a média dos 14 dias monitorizados deste consumidor, bem como a média dos diagramas simulados. Através da comparação destes diagramas, verifica-se que, apesar da existência de alguma variação, os consumos seguem a mesma tendência.

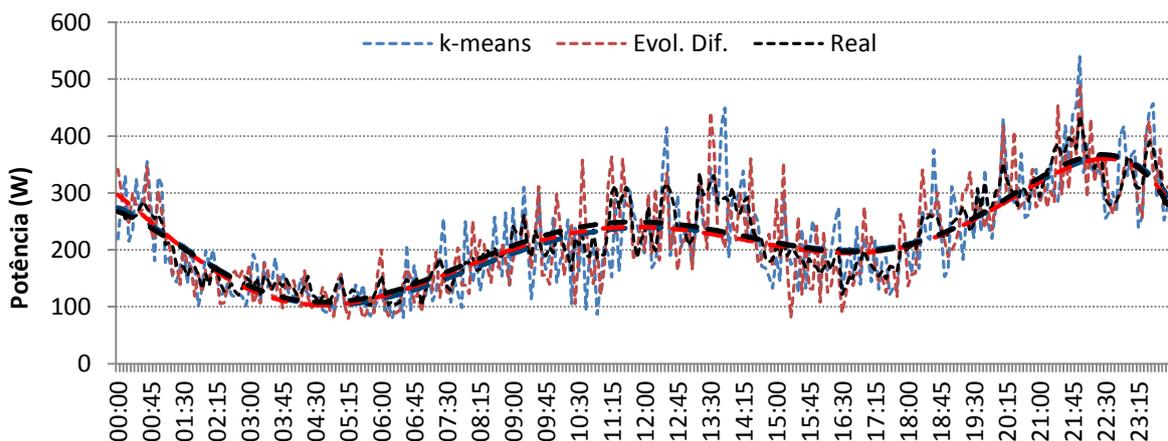


Figura 20: Comparação do consumo médio real de CB (Preto) de um conjunto de consumidores, com o consumo médio obtido através da aplicação do *k-means* (Azul) e ED (Vermelho)

Na Figura 20, para a construção do consumo médio real da CB, foram utilizados 51 diagramas de carga diários de diversos consumidores, os quais não estão incluídos na base de dados à qual foram aplicados os algoritmos. Os algoritmos foram aplicados a 185 diagramas de carga diários, obtendo-se assim as distribuições de probabilidade do consumo ao longo do dia. De forma a comparar os resultados obtidos pelos algoritmos com os diagramas reais, foram simulados 51 diagramas de carga e de seguida comparada a média dos mesmos. Mais uma vez, os consumos obtidos pelos algoritmos apresentam o mesmo comportamento que os dos diagramas reais.

A diferença média obtida entre os diagramas está apresentada na Tabela 3.

Tabela 3: Diferença média entre os diferentes diagramas obtidos através da aplicação dos algoritmos

	Real	<i>k-means</i>	Evolução Diferencial
Real		35,24	33,56
<i>k-means</i>	35,24		46,61
Evolução Diferencial	33,56	46,61	

4.2. Máquina de lavar roupa, loiça e secar roupa

Para construir os diagramas de consumo das MLR, MLL e MSR, é necessário que exista uma base de dados de ciclos de funcionamento para diferentes programas de lavagem. No caso apresentado, a base de dados da MLR foi construída com recurso a recolhas efectuadas em 47 equipamentos, de onde foi possível extrair 392 ciclos de funcionamento, que se encontram distribuídos da seguinte forma:

Tabela 4: Distribuição por programas dos ciclos de funcionamento obtidos para a MLR

Programa:	Frio	30°C/40°C	60°C	90°C
Número:	85	116	114	77

Já na MLL, o número de equipamentos monitorizados foi inferior, sendo de apenas 19 máquinas. No entanto, foi possível extrair 213 ciclos de funcionamento. Não foi possível nestes ciclos identificar de forma clara os programas de funcionamento. Contudo, o modelo de simulação do SIMCAR já se encontra preparado para a existência de quatro tipos de programas, também diferenciados pela temperatura de lavagem e classificados de acordo com o estudo de Stamminger [19]. A distribuição dos programas de lavagem está representada na Tabela 5.

Tabela 5: Distribuição dos ciclos de funcionamento das MLL

Programa:	35°C/45°C	50°C/55°C	65°C	70°C
Número:	13%	37%	36%	14%

Uma vez que não houve nenhuma monitorização de MSR, verificou-se a necessidade de criar alguns diagramas de carga para poderem ser utilizados pelo SIMCAR. Os diagramas criados seguem o exemplo de um ciclo típico de uma MSR que se encontra em [10], e pode ser observado no Anexo IV. Como o modelo de simulação também já se encontra operacional no SIMCAR, apenas é necessário actualizar a base de dados de ciclos da MSR, quando estes se encontrarem disponíveis.

O número de programas de funcionamento diferenciados para a MSR são também quatro, tendo por base o tempo de funcionamento dos mesmos. Os programas foram separados em 20 minutos, 40 minutos, 60 minutos e 80 minutos.

A metodologia de construção do diagrama de carga de uma máquina pode ser dividida em duas partes. Inicialmente é necessário determinar o número de vezes que a máquina entra em

funcionamento em cada dia da semana e em que altura(s) do dia. Só posteriormente é determinado qual(ais) o(s) programa(s) de funcionamento.

Para determinação do número de vezes em que a máquina funciona em cada dia é necessário que o utilizador indique o número médio de ciclos semanais e a distribuição ao longo da semana.

Através desta informação, é criada a seguinte tabela:

Tabela 6: Tabela explicativa de como é realizada a distribuição dos ciclos de funcionamento ao longo da semana (neste caso foram usadas cinco utilizações por semana).

	Segunda Feira	Terça Feira	Quarta Feira	Quinta Feira	Sexta Feira	Sábado	Domingo
I	10%	10%	10%	10%	10%	25%	25%
II	0,5	0,5	0,5	0,5	0,5	1,3	1,3
III	1	1	1	1	1	2	2
IV	2	2	2	2	2	3	3

I – O utilizador insere a percentagem de utilizações da máquina.

II – Número de ciclos por dia. Estes valores são calculados automaticamente através de $\text{Número de ciclos por semana} \times P(\text{dia da semana})$. Na Tabela 6, o número de ciclos utilizados por semana foi de cinco.

III – Número máximo de ciclos possíveis no dia. Este valor é obtido fazendo o arredondamento de **II** por excesso.

IV – Possibilidades de funcionamento. Por exemplo, no Sábado, em que a máquina funciona no máximo duas vezes, existem três possibilidades de funcionamento, sendo elas: “não funcionar”, “Funcionar uma vez” ou “Funcionar duas vez”. Neste exemplo, no máximo dos máximos, numa semana, a máquina poderá funcionar nove vezes ou pode nunca funcionar, mas a probabilidade de tal ocorrer é extremamente reduzido.

O cálculo do número de ciclos de funcionamento em cada dia da semana é feito através da geração de um número aleatório entre zero e o número máximo de ciclos nesse dia (**III**). Se o número aleatório for superior ao número de ciclos do dia (**II**) a máquina não inicia qualquer ciclo. Caso seja inferior, a máquina iniciará o número de ciclos correspondentes. A melhor forma de compreender é através da Figura 21.



Figura 21: Determinação do número de ciclos de funcionamento da MLR para Sábado, de acordo com a Tabela 6

Esta figura serve para exemplificar o caso de Sábado da Tabela 6. É então gerado um número aleatório real $rand \in [0,2]$.

$$\text{Se } \begin{cases} 0 \geq \text{rand} < 1, & \text{ocorre 1 ciclo de funcionamento} \\ 1 \geq \text{rand} < 1,3, & \text{ocorrem 2 ciclos de funcionamento} \\ 1,3 \geq \text{rand} < 2, & \text{n\~{a}o ocorre nenhum ciclo de funcionamento} \end{cases}$$

Para alocar os ciclos de funcionamento de cada máquina, em cada dia da semana, é necessário ter o cuidado para que a alocação de um ciclo não se sobreponha com um já existente, uma vez que isso seria fisicamente impossível. O algoritmo desenvolvido para a alocação dos ciclos de lavagem tem em conta esta restrição. Após determinado o número de ciclos de funcionamento em cada dia, o algoritmo divide o dia em *slots* livres. A resolução temporal dos *slots* é a mesma que está a ser utilizada pelo SIMCAR. Na primeira situação, em que ainda não foi alocado nenhum ciclo, existe apenas um *slot* livre, que é o dia completo. Depois do primeiro ciclo ser alocado, os *slots* livres são recalculadas.

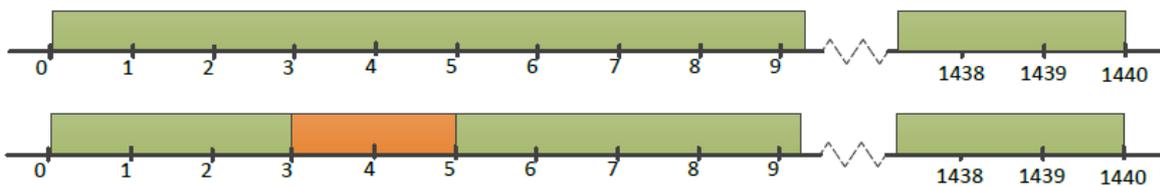


Figura 22: Exemplificação da alocação de um ciclo.

A Figura 22 apresenta uma pequena exemplificação deste processo de alocação de ciclos. Uma vez que se trata apenas de um exemplo, o ciclo alocado na figura tem a duração de apenas dois minutos.

A determinação do instante em que o ciclo é alocado é realizada tendo em conta a distribuição de probabilidade ao longo do dia da máquina entrar em funcionamento.

No decorrer da alocação dos diferentes ciclos da máquina ao longo do dia, muito provavelmente ocorrerão situações em que algumas horas do dia já possuem um ciclo de funcionamento, não sendo assim possível alocar mais nenhum nesse instante. Quando tal acontece, a probabilidade de a máquina iniciar nessa hora é 0%, sendo as restantes probabilidades recalculadas. Esta situação está representada na Figura 23. Neste caso, entre as 8 e as 9 horas e entre as 14 e 15 horas não há *slots* livres, logo a probabilidade nestes intervalos é 0%.

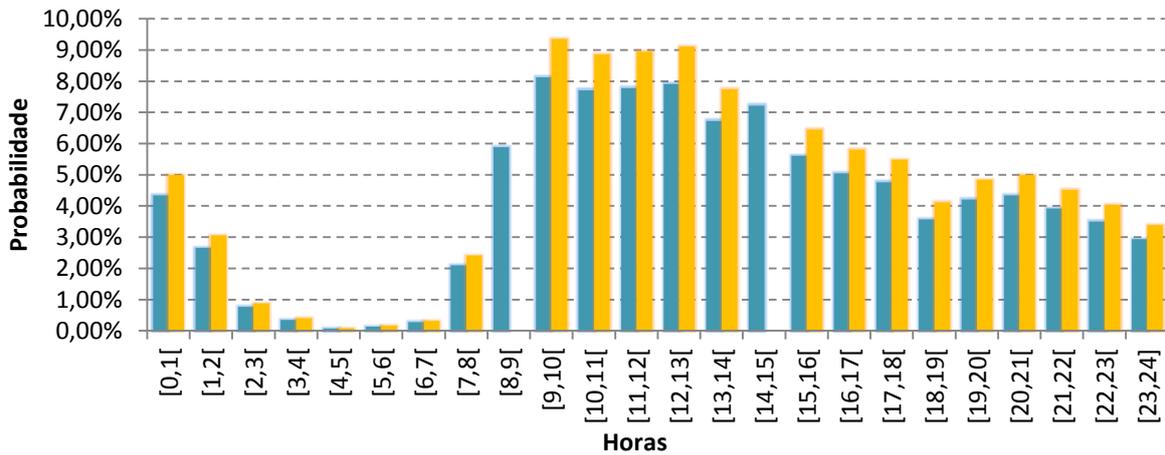


Figura 23: Distribuição de probabilidade inicial (Azul) e distribuição de probabilidade em que existem períodos que não é possível alocar ciclos de funcionamento (Laranja).

Uma vez que a distribuição de probabilidade de a máquina iniciar um ciclo de lavagem está definida como sendo de hora a hora, após determinada a hora de início, é necessário ainda determinar o instante (minuto) inicial. Para isso, dentro dos *slots* livres existentes nessa hora são identificados os instantes em que o ciclo pode ser alocado.

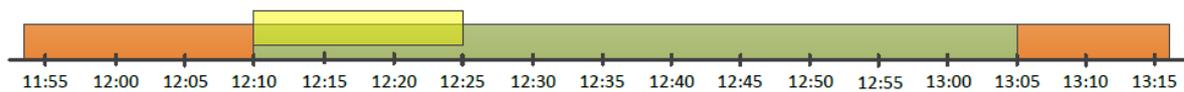


Figura 24: Cálculo do intervalo em que é possível alocar o ciclo. Laranja representa intervalos já ocupados, verde representa o *slot* livre e amarelo o intervalo em que o ciclo pode iniciar o funcionamento.

Por exemplo, se se pretender alocar um ciclo de 40 minutos no intervalo definido na Figura 24, verifica-se que este apenas pode iniciar o seu funcionamento no intervalo que se encontra a amarelo. Dentro deste intervalo, o cálculo do minuto exacto é feito através da geração de um número aleatório de distribuição uniforme entre as 12:10 e as 12:25.

Para as máquinas de lavar loiça e secar roupa, o processo de alocação de ciclos é idêntico.

Para saber a distribuição de probabilidade de as máquinas iniciarem os ciclos, o SIMCAR lê o ficheiro que se encontra em “Ficheiros\Input\Calc Prob (e a respectiva máquina)”.

Este ficheiro é constituído por duas folhas Excel, em que na primeira onde o SIMCAR vai ler os dados (ver Manual de Utilizador) e a segunda é onde foi concebida uma calculadora para construir a distribuição de probabilidade com base nos consumos horários das máquinas.

Para utilizar a calculadora, o utilizador apenas necessita de inserir os dados a verde-claro, sendo o resultado os valores que se encontram a laranja (ver Figura 25).

Taxa posse		90%		Nr Habitacoes		4.000.000		Pico Máximo Possível		2,1 kW	
t(h)	P(kW)	Nr Habitacoes		Potencia							
		3.600.000	100%	7.560.000	kW			Prob	Cumulativa	Prob 100%	Cum 100%
[0,1]	62.000		0,8201%	62.000	kW			0,8201%	0,8201%	4,34%	4,3448%
[1,2]	38.000		0,5026%	38.000	kW			0,5026%	1,3228%	2,66%	7,0077%
[2,3]	11.000		0,1455%	11.000	kW			0,1455%	1,4683%	0,77%	7,786%
[3,4]	5.000		0,0661%	5.000	kW			0,0661%	1,5344%	0,85%	8,1289%
[4,5]	1.000		0,0132%	1.000	kW			0,0132%	1,5476%	0,07%	8,1990%
[5,6]	2.000		0,0265%	2.000	kW			0,0265%	1,5741%	0,14%	8,3392%
[6,7]	4.000		0,0529%	4.000	kW			0,0529%	1,6270%	0,28%	8,6195%
[7,8]	30.000		0,3968%	30.000	kW			0,3968%	2,0238%	2,10%	10,7218%
[8,9]	84.000		1,1111%	84.000	kW			1,1111%	3,1349%	5,89%	16,6083%
[9,10]	116.000		1,5344%	116.000	kW			1,5344%	4,6693%	8,13%	24,7372%
[10,11]	110.000		1,4550%	110.000	kW			1,4550%	6,1243%	7,71%	32,4457%
[11,12]	111.000		1,4683%	111.000	kW			1,4683%	7,5926%	7,78%	40,2242%
[12,13]	113.000		1,4947%	113.000	kW			1,4947%	9,0873%	7,92%	48,1430%
[13,14]	96.000		1,2698%	96.000	kW			1,2698%	10,3571%	6,73%	54,8704%
[14,15]	103.000		1,3624%	103.000	kW			1,3624%	11,7196%	7,22%	62,0883%
[15,16]	80.000		1,0582%	80.000	kW			1,0582%	12,7778%	5,61%	67,6945%
[16,17]	72.000		0,9524%	72.000	kW			0,9524%	13,7302%	5,05%	72,7400%
[17,18]	68.000		0,8995%	68.000	kW			0,8995%	14,6296%	4,77%	77,5053%
[18,19]	51.000		0,6746%	51.000	kW			0,6746%	15,3042%	3,57%	81,0792%
[19,20]	60.000		0,7937%	60.000	kW			0,7937%	16,0979%	4,20%	85,2838%
[20,21]	62.000		0,8201%	62.000	kW			0,8201%	16,9180%	4,34%	89,6286%
[21,22]	56.000		0,7407%	56.000	kW			0,7407%	17,6587%	3,92%	93,5529%
[22,23]	50.000		0,6614%	50.000	kW			0,6614%	18,3201%	3,50%	97,0568%
[23,24]	42.000		0,5556%	42.000	kW			0,5556%	18,8757%	2,94%	100,0000%

Figura 25: Calculadora de distribuição de probabilidades

A determinação destas distribuições de probabilidade de as máquinas iniciar ciclos de funcionamento no decorrer do dia podem ainda ser calculadas através do uso dos algoritmos de DM e ED desenvolvidos. Esta aplicação torna-se extremamente útil para se poder criar as distribuições com base nos dados recolhidos. Desta forma, em vez de se utilizar uma distribuição generalista, que engloba todos os tipos de consumidores, é possível criar a distribuição de probabilidade de um determinado tipo de consumidores ou até mesmo para um só consumidor. Aliás, pode ser usada, em princípio, qualquer metodologia de agrupamento de consumidores. Por exemplo, pode ser calculada a distribuição para os consumidores alimentados a partir de um dado Posto de Transformação (PT), ou de uma dada cidade. Pode-se assim “gerar” conhecimento sobre o funcionamento destas máquinas, obtendo-se resultados mais realísticos do que se se utilizasse a distribuição nacional.

A escolha do tipo de programa de funcionamento é feita com base numa distribuição de probabilidade dos programas. Caso não seja definido nada em contrário nas definições da simulação destas cargas, esta distribuição é a que se encontra no estudo de Stamminger [19], sendo ela apresentada na Tabela 4 e Tabela 5 para a MLR e MLL respectivamente. Se o utilizador pretender, poderá ele mesmo fornecer esta distribuição, bastando para isso seleccionar na folha das definições destas cargas, em “Probabilidades do tipo de programa” a opção “Própria” e indicando na respectiva tabela a distribuição pretendida. O cenário mais interessante é quando essa distribuição é calculada com base em algum tipo de recolha de dados/informação junto dos consumidores.

Como já foi dito anteriormente, o método utilizado pelo simulador na representação de diversos diagramas de carga (*bottom-up*) consiste em simular cada carga individualmente, e através da agregação dos resultados de cada uma, criar o diagrama de carga final.

4.3. Ar Condicionado, *Inverter* e Termoacumulador

No caso dos sistemas de AC (Ar Condicionado Convencional e *Inverter*) e de aquecimento eléctrico de água, os diagramas de cargas são obtidos com recurso aos MFB ou a recolhas. Neste primeiro caso, em que se recorre ao MFB, a simulação dos fenómenos físicos permite a obtenção do padrão de consumos, sendo, para isso, necessário a caracterização dos padrões de utilização dos serviços de energia, das cargas e dos espaços climatizados, assim como de algumas variáveis ambientais (nomeadamente da temperatura). Uma vez que os MFB são muito exigentes no que respeita ao tempo de processamento, para que se possa utilizar o SIMCAR numa situação de simulação de um elevado número de cargas deste tipo ou numa situação de muitas simulações referentes a uma só carga mas com variação de algum dos parâmetros caracterizadores do seu funcionamento, resolveu criar-se, para cada tipo de consumidor, uma base de dados de diagramas de carga possíveis. Depois, e contrariamente ao que acontece com as cargas anteriormente descritas, o SIMCAR escolhe um dos diagramas diários existentes na base de dados. De referir que a opção por “criar” uma base de dados de diagramas de carga com base nos MFB é baseada na necessidade de lidar com a variabilidade do consumo de energia que a utilização do serviço de energia proporcionado por estas cargas acarreta. Assim, torna-se possível para um dado equipamento num dado cenário de espaço/temperaturas simular diversas utilizações do serviço, em que alguns dos parâmetros que caracterizam o funcionamento das cargas pode variar. A esta base de dados assim construída é também possível aplicar as técnicas de *clusterização*, por outro lado é possível aplicar as mesmas técnicas a uma base de dados construída com recolhas de dados reais.

Para que seja possível construir a base de dados é necessário criar previamente uma outra base de dados de perfis de utilização do serviço de energia.

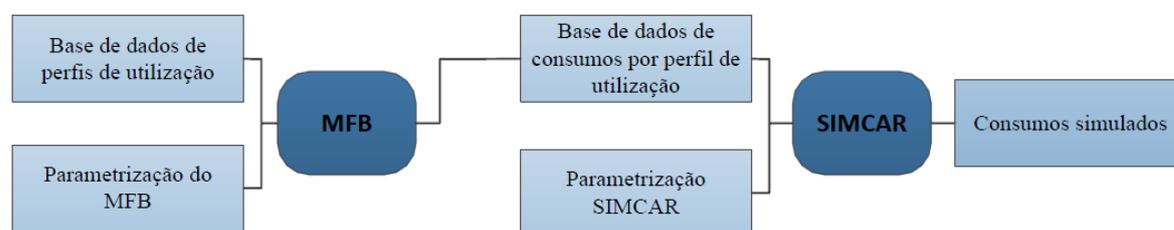


Figura 26: Fluxograma simples para demonstração dos processos para a criação dos consumos pelo SIMCAR

Cada pessoa é diferente das demais, e apresenta um comportamento específico, o que dificulta todo o processo de simulação. Contudo, é possível identificar e classificar diferentes tipos de perfis de utilização destas cargas. Uma pessoa que vai para o emprego de manhã e só regressa ao local de residência ao final do dia, vai apresentar um consumo diferente quando comparado com uma que se apresente desempregada ou reformada. Podemos então identificar três perfis principais:

No MFB, o utilizador pode definir diversos parâmetros, conforme o pretendido. No caso do ar condicionado e *inverter*, estes parâmetros podem estar relacionados com as dimensões do espaço em que está inserido, a insolação, a humidade, a temperatura interior e exterior, a potência do aparelho, fontes internas, temperatura pretendida, entre outros.

De forma a tornar a utilização do MFB mais prática, encontram-se já definidos 10 grupos de cargas com parâmetros distintos, permitindo também assim simular uma maior variedade de situações por simulação.

Para o termoacumulador, os parâmetros que podem ser definidos no MFB são a capacidade do termoacumulador, a sua potência, a temperatura inicial da água, temperatura da água da rede, os consumos ao longo do dia, etc.

Quanto à parametrização do SIMCAR no que respeita a estas cargas, o utilizador necessita de indicar o número de cargas a simular, qual a distribuição dos perfis de utilização e a percentagem de variação dos mesmos de carga para carga. O simulador permite que o utilizador altere as distribuições dos perfis de utilização entre dias úteis (DU), Sábados e Domingos.

Distribuição dos perfis de utilização						
1	2	3	4	5		
30%	60%	5%	3%	2%	100%	DU
30%	90%	95%	98%	100%		
20%	20%	50%	5%	5%	100%	Sábado
20%	40%	90%	95%	100%		
20%	10%	60%	5%	5%	100%	Domingo
20%	30%	90%	95%	100%		

Figura 29: Tabela existente nas folhas Excel do AC, *Inverter* e TR para a inserção da distribuição dos perfis de utilização para DU, Sábado e Domingo

É ainda possível utilizar os algoritmos de DM ou ED para realizar a simulação dos consumos destas cargas. Contudo, através destes algoritmos torna-se complexa a aplicação e avaliação dos impactos das estratégias de GP, uma vez que o consumo obtido será um consumo médio e não o real.

4.4. Interface

Com o objectivo de tornar o funcionamento do SIMCAR mais intuitivo para o utilizador foi também desenvolvida uma interface, que serve para executar as diferentes etapas do simulador. Esta interface está representada na Figura 30.

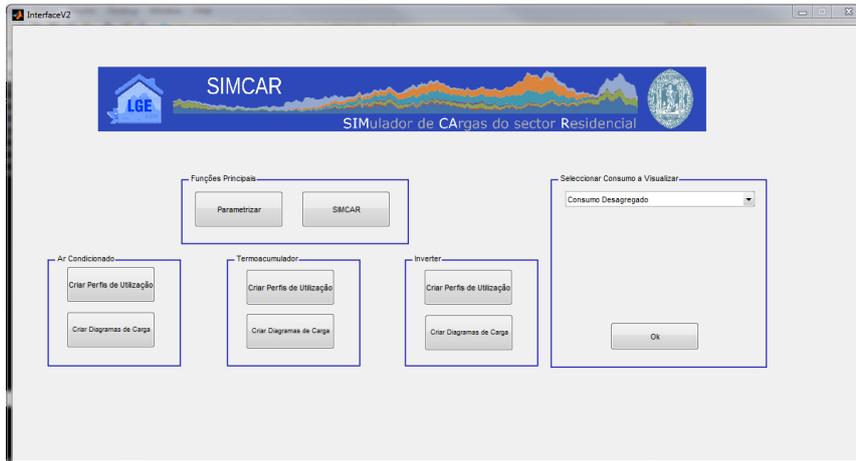


Figura 30: Interface do SIMCAR

Com esta interface o utilizador não necessita de escrever as funções na linha de comandos do *Matlab*. Após a conclusão da simulação pode-se visualizar os diferentes gráficos de consumos, podendo estes serem consumos agregados ou individuais. Para tal, basta seleccionar o consumo que pretende visualizar no lado direito da interface, sendo os consumos apresentados numa janela à parte, e guardados no ficheiro Excel “Resultados SIMCAR.xlsx”.

A parametrização do SIMCAR é feita toda ela em diversas folhas de cálculo do ficheiro Excel “Definicoes.xlsx”, podendo estas ser observadas no Anexo V.

5. Caso de estudo

De forma a demonstrar possíveis aplicações do SIMCAR, foram realizados alguns estudos de caso. Neste capítulo estão apresentados esses estudos, estando eles divididos em cinco partes. Numa primeira fase, é realizado o cálculo do número ideal de *clusters* para bases de dados de diferentes tamanhos contendo diagramas diários de CB e procede-se à aplicação do algoritmo de ED a alguns casos, de forma a se verificar diferenças na sua performance. É ainda feita uma comparação entre ambos os algoritmos de *clusterização* utilizados. Numa segunda fase, são simuladas individualmente algumas cargas, de modo a poderem observar-se detalhadamente os resultados individuais e compará-los quando são, ou não são aplicadas medidas de GP. Numa terceira fase são comparados os diagramas agregados de diversos consumidores. Procedeu-se

também à simulação o consumo de um consumidor e avaliada a poupança que este poderá obter na sua factura mensal se modifica os seus hábitos de consumo no que diz respeito à altura do dia em que coloca as suas máquinas em funcionamento. Finalmente numa última fase é demonstrado o resultado obtido com a aplicação do algoritmo *k-means* à base de dados de consumos das MLR monitorizadas, e comparado o resultado obtido com o do estudo da DGGE / IP-3E [10].

5.1. Algoritmos

Tal como foi referido, uma das desvantagens dos algoritmos de DM de particionamento é a necessidade de se introduzir *a-priori* o número k de *clusters* a utilizar. De forma a determinar qual o número ideal de *clusters* a utilizar, foi realizado o seguinte teste; aplicou-se o algoritmo *k-means* a bases de dados de diagramas de CB de contendo diferentes diagramas consumidores.

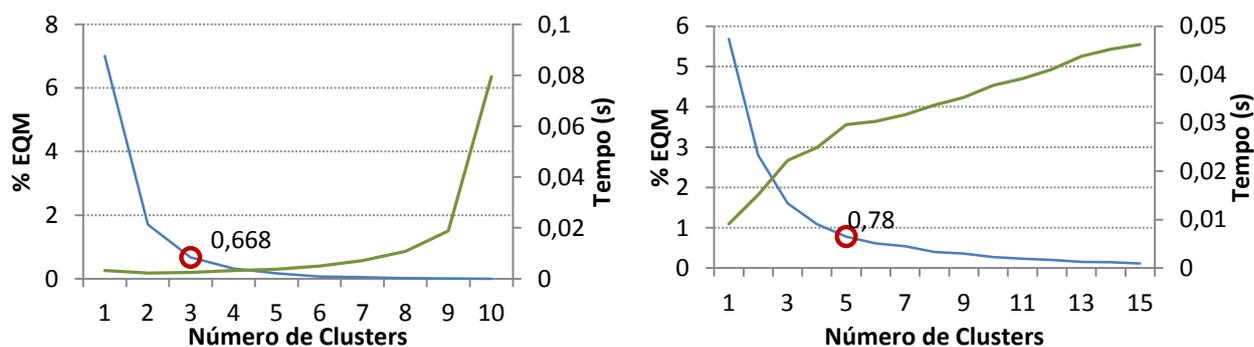


Figura 31: Determinação do número k de *clusters* para a situação de uma base de dados de medições de 10 diagramas de carga e de um de 365 diagramas

Na Figura 31 está a demonstração do número de *clusters* para o caso de apenas se ter o consumo de duas semanas de um consumidor (10 dias úteis) e para o caso de se ter uma base de dados de 365 diagramas (como por exemplo, medições ao longo de um ano de um consumidor). Neste teste foi utilizado o método do erro médio quadrático, sendo este calculado para cada intervalo do dia, e calculada a média do erro em cada intervalo, sendo no fim calculada a média do erro total para cada valor de k .

Nesta mesma figura, está também representada a evolução do tempo despendido no processo de *clusterização*. Este é o tempo médio por intervalo de tempo. Como seria expectável, quanto maior o número de elementos e de *clusters*, maior será o tempo despendido. No Anexo VI está realizado o mesmo teste para bases de dados de 15, 30 e 185 medições.

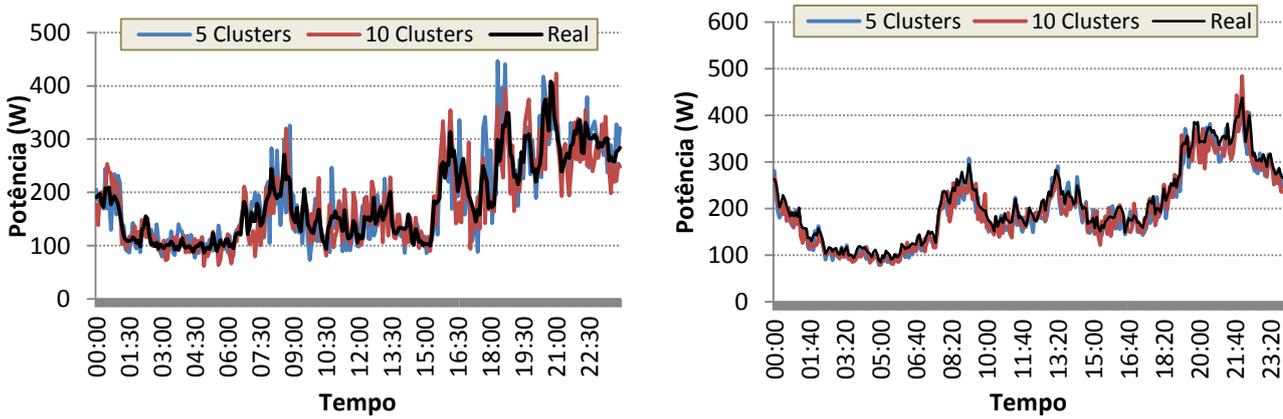


Figura 32: Consumo médio de CB de 50 consumidores (esquerda) e de CB de 500 consumidores (direita)

Devido à grande variabilidade existente neste tipo de carga, quanto menor for o número de consumidores a simular, maior será a diferença verificada quando comparados os consumos simulados com os reais. Na Figura 32 é possível observar esta variabilidade.

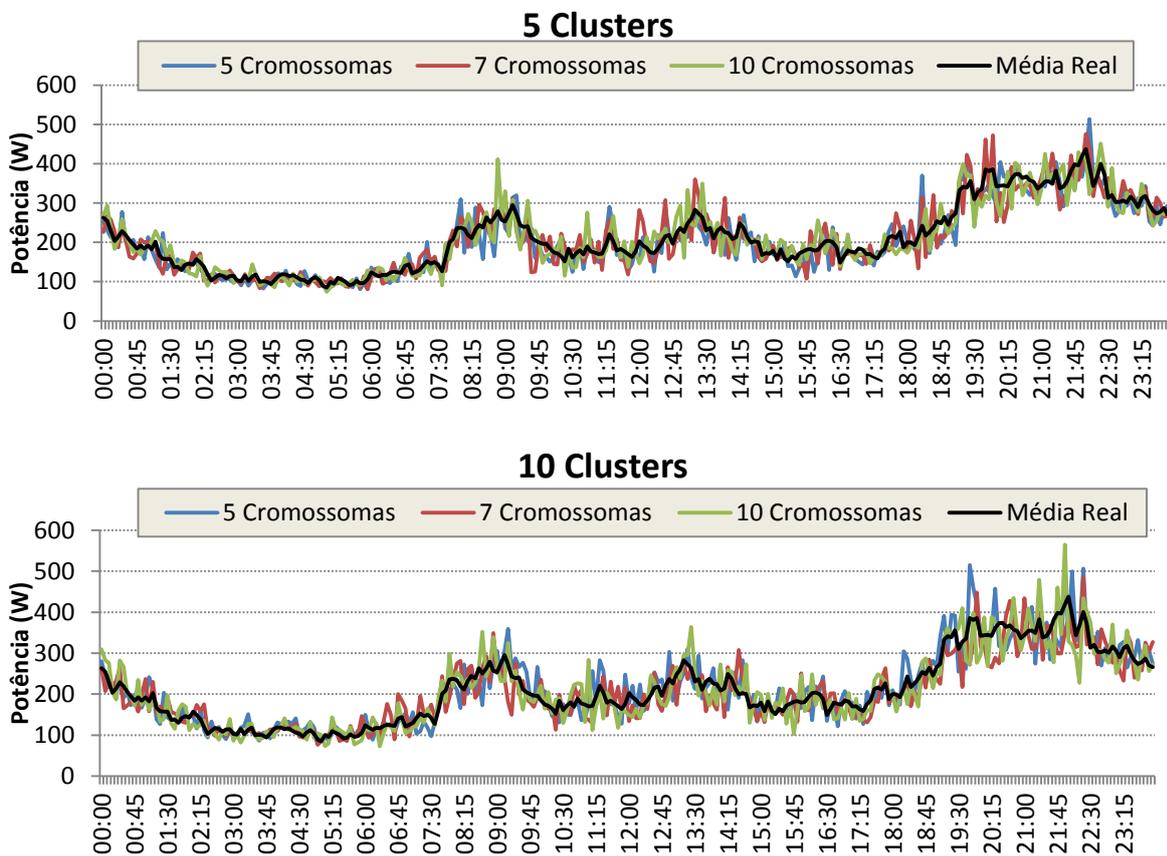


Figura 33: Resultados obtidos da aplicação do algoritmo de ED para diferente número de *clusters* e de cromossomas

Na Figura 33 estão os resultados obtidos através do algoritmo de ED, variando o número de cromossomas e o número de *clusters*. Apesar de não ser muito perceptível nas figuras, verifica-se que não há variações muito significativas entre o consumo simulado com o consumo real aquando a variação dos parâmetros.

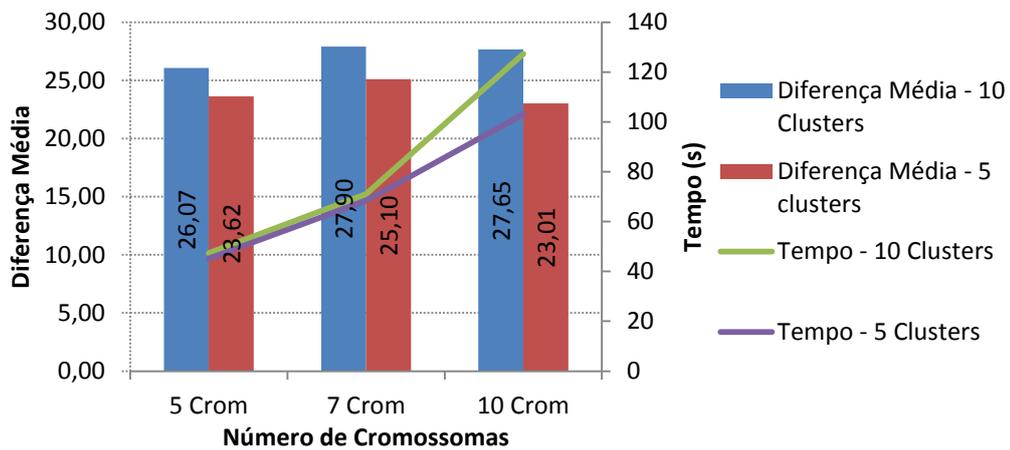


Figura 34: Comparação da diferença média resultante dos diversos teste e tempos de execução

De forma a tornar mais compreensível a diferença existente com a variação dos parâmetros do algoritmo de ED, está apresentada na Figura 34 a diferença média entre o consumo de carga base obtido por este algoritmo e o consumo real. Verifica-se que onde existe menor variação é quando se utilizaram cinco *clusters* e dez cromossomas. Contudo, o tempo necessário para o algoritmo executar foi de 103 segundos. O segundo melhor resultado obtido nesta comparação foi quando se utilizou cinco *clusters* e cinco cromossomas. Apesar de a diferença média dos consumos nesta situação ser um pouco superior à anterior, o tempo de execução foi o que demonstrou ser mais reduzido, sendo este de 45 segundos. O rácio entre performance e soluções obtidas é assim vantajoso.

De forma a comparar os dois algoritmos – *k-means* e evolução diferencial – está na Figura 35 representado o consumo médio de 300 consumidores obtido através destes. Como se pode verificar, a diferença existente entre eles é pequena, estando a percentagem de variação na Figura 36.

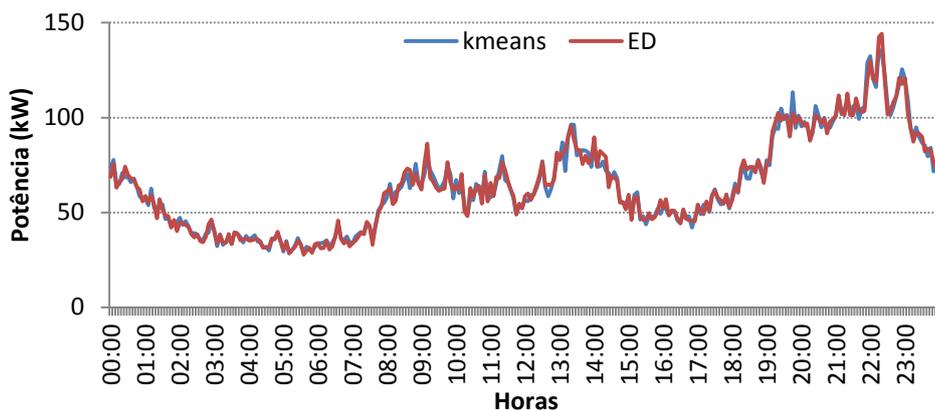


Figura 35: Consumo de carga base de 300 consumidores obtido através do algoritmo de Evolução Diferencial e do algoritmo de DM, *k-means*.

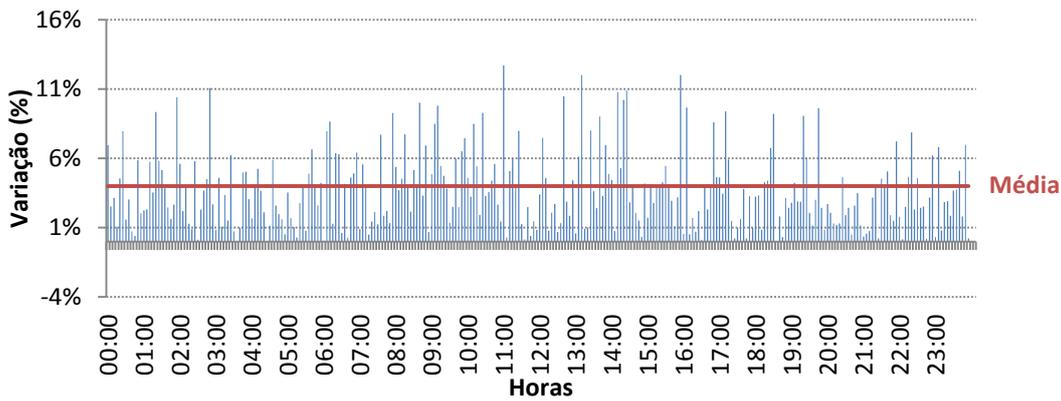


Figura 36: Diferença entre os consumos de carga base dos 300 consumidores obtido pelos dois métodos

Para esta simulação foram então utilizados dez *clusters* no algoritmo *k-means* e cinco cromossomas e cromossomas no algoritmo de ED. Nas Tabela 7 e Tabela 8 estão os centroides dos *clusters* obtidos no *k-means* e no melhor cromossoma obtido pelo algoritmo de ED, bem como a respectiva distribuição de probabilidade cumulativa de cada, para os instantes da 00:45h e das 20h. Apesar de os resultados serem semelhantes, o tempo de execução do *k-means* foi muito inferior.

Tabela 7: Centroids de cada *cluster* obtidos pelo *k-means* e pelo algoritmo de ED para as 00:45h bem como a respectiva distribuição de probabilidade cumulativa

Horas: 00:45											
Centroid	k-means	47,11	197,52	323,56	408,93	468,43	531,2	670,8	1040	1200	1453,6
	Evol Dif	67,778		275,036		547,552		1120		1453,6	
Prob.	k-means	0,5204	0,7953	0,8596	0,8947	0,9357	0,9532	0,9825	0,9883	0,9942	1
	Evol Dif	0,6374		0,8947		0,9825		0,9942		1	

Tabela 8: Centroids de cada *cluster* obtidos pelo *k-means* e pelo algoritmo de ED para as 20:00h bem como a respectiva distribuição de probabilidade cumulativa

Horas: 20:00											
Centroid	k-means	58,71	263,57	435,56	650,83	876,75	1143,31	1635,6	2323,2	2669,2	4480
	Evol Dif	160,827		800,741		1635,6		2553,867		4480	
Prob.	k-means	0,4773	0,7273	0,8409	0,8864	0,9318	0,9602	0,9773	0,9830	0,9943	1
	Evol Dif	0,8182		0,9602		0,9773		0,9943		1	

5.2. Diagramas de carga individuais

Para demonstrar as potencialidades das estratégias de GP de controlo directo de carga, foram simuladas três situações distintas. Nestas simulações foram utilizadas 3.000 MLR. A distribuição de probabilidade dos tipos programas de funcionamento é dada por 6,6% para programas de lavagem a frio, 54,3% para lavagem a 30°C/40°C, 32,3% para lavagem a 60°C e para os programas de lavagem a 90°C é de 6,8%. A distribuição de probabilidade do funcionamento das máquinas ao longo do dia utilizada na construção do consumo médio de funcionamento normal é a distribuição obtida através do estudo da DGGE / IP-3E, 2004 [10], e que se encontra representada na Figura 37.

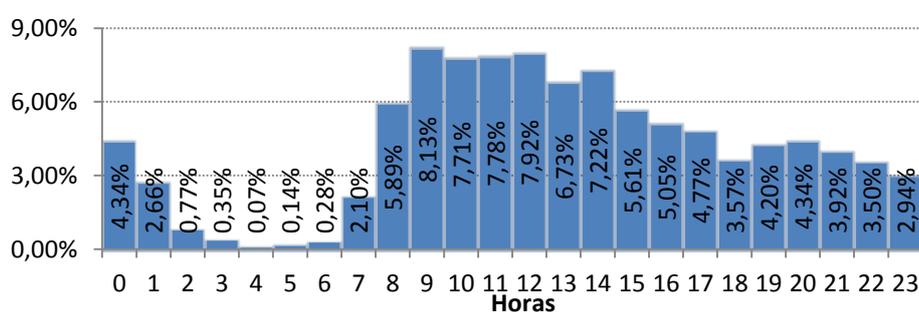


Figura 37: Distribuição de probabilidade do funcionamento das MLR obtido de [10]

Na Figura 38 podem ser observados os consumos das diferentes situações.

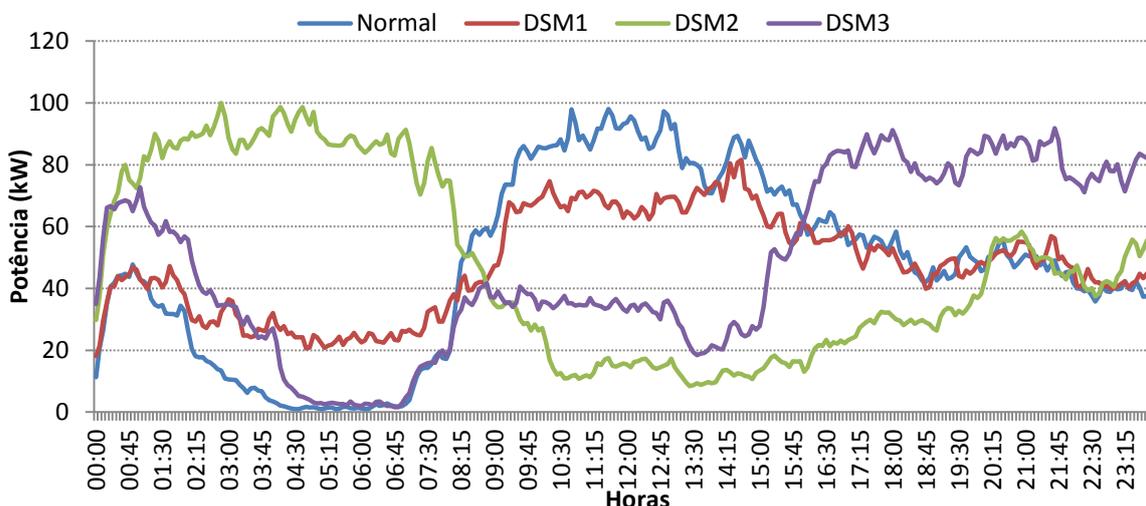


Figura 38: Comparação entre o consumo em funcionamento normal das máquinas de lavar roupa (azul) com a aplicação de diferentes técnicas de controlo de carga (outras cores)

Na primeira situação (DSM1), o controlo destas máquinas teve como objectivo o desvio dos consumos para o período de vazio, onde se verifica menor consumo, reduzindo-se desta forma a ponta que se verifica nas horas de cheias, onde existe maior procura. Nesta situação reduziu-se o pico em cerca de 13,351 kW, tal como se pode observar pela diferença do consumo da situação de funcionamento normal (azul) com o consumo registado com esta estratégia (vermelho).

Contudo, é necessário ter alguma atenção quando se realiza o desvio de consumos para as horas de vazio. O objectivo da simulação realizada na segunda situação (DSM2) foi precisamente demonstrar um dos cuidados a ter. Nesta simulação, ao serem transferidos os consumos para a madrugada, tal como se pode observar no gráfico, criou-se um novo pico, que poderá prejudicar o SEE. Na última situação (DSM3), o objectivo foi o de demonstrar a utilidade das estratégias de GP na ocorrência de um defeito no SEE, que neste caso ocorreu por volta das 08:15. Na ocorrência do defeito, a operadora envia um estímulo aos consumidores para que estes diminuam os seus consumos. Para este tipo de cargas, como não é possível interromper os ciclos de lavagem, as máquinas que já se encontram em funcionamento, continuarão em funcionamento. Apenas é impedido que novos ciclos sejam iniciados. Se um consumidor desrespeitar o sinal, sofrerá penalizações. Resolvido o problema, a operadora do SEE reenvia um novo sinal e os consumos são retomados.

Outra estratégia de GP existente consiste na reparametrização de cargas termostáticas. Com o intuito de demonstrar a poupança energética possível de obter através desta, foram simulados 480 AC, para os quais se procedeu à simulação dos consumos para o funcionamento como termostato definido para 24°C, 24,5°C e 25°C. Na Figura 39 pode-se observar graficamente o consumo dos 480 AC em estudo, estando na Tabela 9 apresentada a energia consumida por estes, sendo possível verificar as reduções ocorridas.

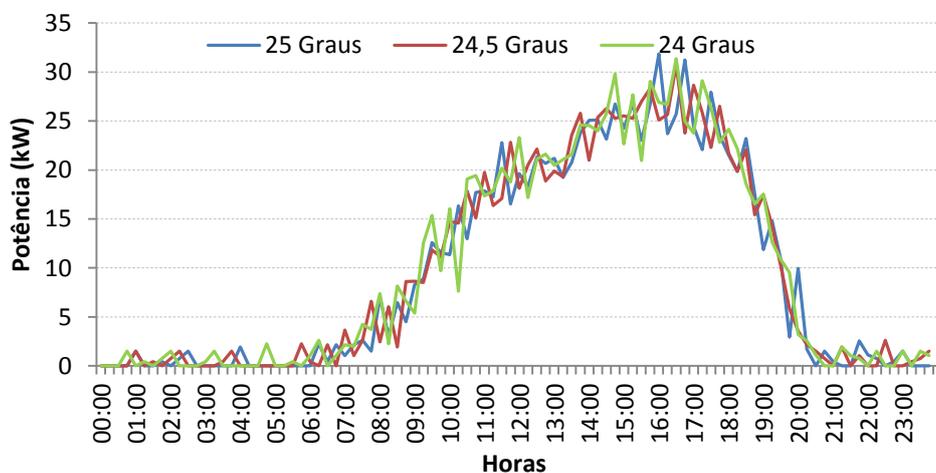


Figura 39: Consumo de 480 AC, com os mesmo parâmetros, apenas variando a temperatura definida do termostato para 25°C, 24,5°C e 24°C.

Tabela 9: Consumo de 480 AC, cujo termostato foi definido para 24°C, 24,5°C e 25°C

	24 Graus	24,5 Graus	25 Graus
Energia total consumida (kWh)	241,3	236,58	233,06
Poupança (kWh) no aumento de 24,5°C para 25°C			3,52
Poupança (kWh) no aumento de 24°C para 24,5°C		4,72	
Poupança (kWh) no aumento de 24°C para 25°C		8,24	

É assim possível verificar que com o aumento de apenas meio grau, obtiveram-se poupanças: 3,52 kWh quando se aumentou de 24,5°C para 25°C, 4,72 kWh quando se aumentou de 24°C para 24,5°C e finalmente, 8,24 kWh quando o aumento foi de 1°C, se parametrizar o termostato para 25°C.

Apesar de esta redução não ser muito significativa individualmente, em situações em que o SEE possa estar em risco de comprometer a sua integridade, a redução obtida poderá fazer a diferença, especialmente quanto mais elevado for o número de consumidores que adiram a estas medidas.

5.3. Diagramas de carga finais

Este capítulo encontra-se dividido em duas partes. Na primeira parte são apresentados dois cenários de simulação distintos e na segunda parte encontra-se a simulação de diversos de forma a avaliar o impacto da estratégia de GP de reparametrização de cargas termostáticas.

5.3.1. Apresentação de dois cenários distintos

Cenário 1

Num primeiro cenário foram simulados 1000 consumidores ao longo de uma semana. O número de máquinas a simular foi obtido através da utilizando a taxa de posse de [10], podendo o resultado ser observado na Tabela 10.

Tabela 10: Caracterização das máquinas simuladas no primeiro cenário

	Sexta	Domingo	Ciclos semanais	Número
MLR	10%	20%	4,9	900
MLL	10%	45%	4,2	700
MSR	10%	25%	4	400

Nesta tabela é também apresentado o número médio de ciclos semanais, bem como a distribuição dos ciclos de funcionamento para os dois dias da semana expostos na dissertação.

A distribuição dos programas de funcionamento foi a seguinte:

Tabela 11: Distribuição de probabilidades dos ciclos de funcionamento das máquinas utilizadas no primeiro cenário

Programa	Frio	30/40°C	60°C	90°C
MLR	10%	50%	20%	20%
Programa	35/45°C	50/55°C	65°C	70°C
MLL	12%	37%	36%	15%

Para o AC, *Inverter* e TR, as taxas de posse utilizadas foram de 30%, 35% e 45% respectivamente, sendo a distribuição dos perfis de utilização apresentados na Tabela 12.

Tabela 12: Distribuição dos perfis de utilização dos serviços de energia do TR, AC e *Inverter* para o primeiro cenário

	Distribuição dos perfis de utilização								
	Termoacumulador				AC e <i>Inverter</i>				
	1	2	3	4	1	2	3	4	5
DU	60%	15%	15%	10%	30%	45%	15%	5%	5%
	60%	75%	90%	100%	30%	75%	90%	95%	100%
Domingo	60%	30%	5%	5%	15%	30%	45%	5%	5%
	60%	90%	95%	100%	15%	45%	90%	95%	100%

Posto isto, os resultados obtidos podem ser visualizados nas Figura 40 e Figura 41. Tal como seria expectável, no Domingo o consumo verificado durante o dia é superior ao consumo verificado na Sexta-Feira. Isto deve-se em grande parte devido à presença de pessoas em casa durante o dia, que está directamente relacionado com o consumo de determinadas cargas, especialmente dos AC e *inverters*.

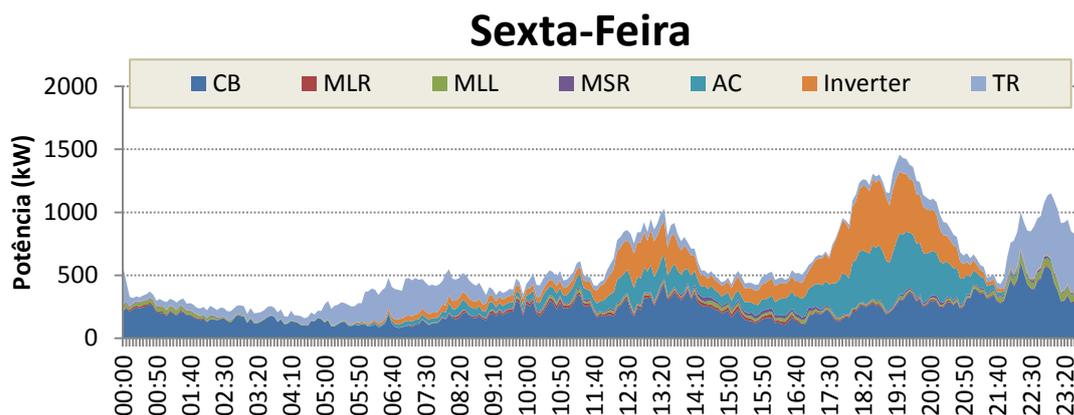


Figura 40: Diagrama de Carga agregado de todos os consumidores para a Sexta-Feira

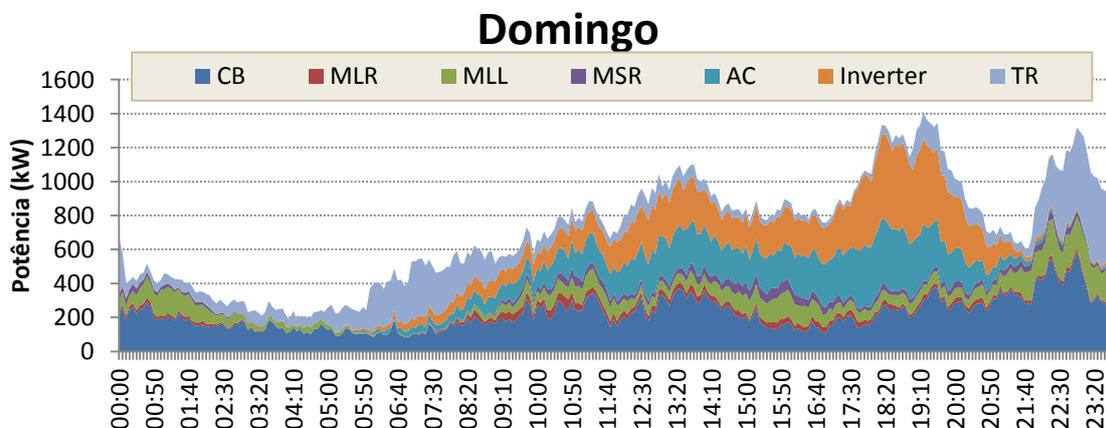


Figura 41: Diagrama de Carga agregado de todos os consumidores para o Domingo

No Domingo também se verifica a existência de consumos superiores no que diz respeito às máquinas de lavar e secar.

Cenário 2

Contrariamente ao primeiro cenário, em que se simulou um conjunto de consumidores que apresentam todos os tipos de consumos, neste segundo cenário pretende-se simular os consumos de um conjunto mais reduzido de consumidores, tendendo estes consumidores a se encontrarem nas respectivas habitações mais ao final da tarde. O número de consumidores simulados foi reduzido para 300. A taxa de posse das MLR, MLL e MSR são de 90%, 50% e 70% respectivamente.

Tabela 13: Caracterização das máquinas simuladas no segundo cenário

	Sexta	Domingo	Ciclos semanais	Número
MLR	15%	40%	5,5	270
MLL	20%	35%	5,5	210
MSR	10%	40%	2	150

Para este cenário foi também alterada a distribuição dos programas de funcionamento das máquinas, podendo estes ser observados na Tabela 14.

Tabela 14: Distribuição de probabilidades dos ciclos de funcionamento das máquinas utilizadas no segundo cenário

Programa	Frio	30/40°C	60°C	90°C
MLR	5%	25%	55%	15%
Programa	35/45°C	50/55°C	65°C	70°C
MLL	5%	55%	30%	10%

No caso do AC, *Inverter* e TR, as taxas de posse utilizadas foram de 15%, 12% e 17% respectivamente, sendo a distribuição dos perfis de utilização apresentados na Tabela 15.

Tabela 15: Distribuição dos perfis de utilização dos serviços de energia do TR, AC e *Inverter* para o segundo cenário

	Distribuição dos perfis de utilização								
	Termoacumulador				AC e <i>Inverter</i>				
	1	2	3	4	1	2	3	4	5
DU	55%	25%	10%	10%	10%	75%	5%	5%	5%
	55%	25%	10%	10%	10%	85%	90%	95%	100%
Domingo	50%	40%	5%	5%	10%	75%	5%	5%	5%
	50%	40%	5%	5%	10%	85%	90%	95%	100%

Nas Figura 42 e Figura 43 pode-se observar os diagramas de carga destes consumidores. Como se pode verificar, a grande maioria dos consumos ocorre ao final da tarde/início da noite.

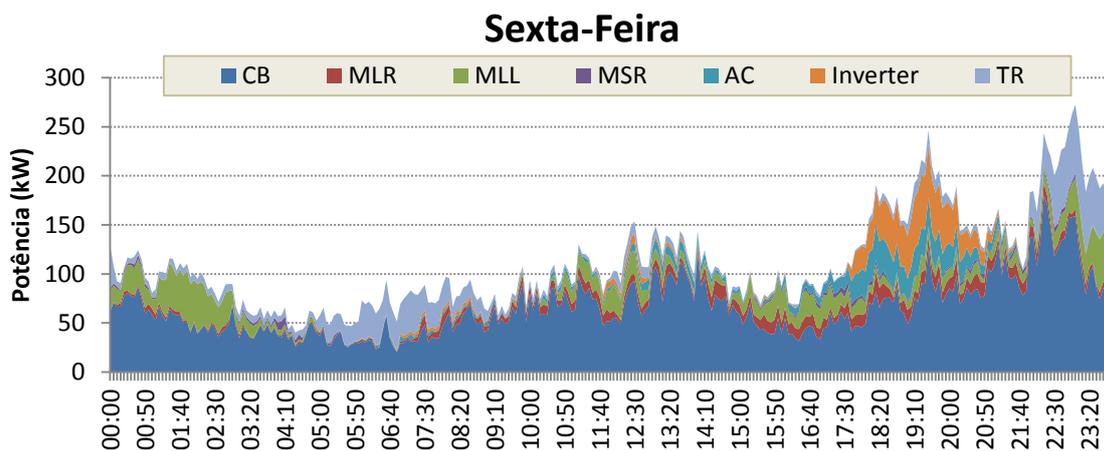


Figura 42: Diagrama de carga agregado dos consumidores do segundo cenário para Sexta-Feira

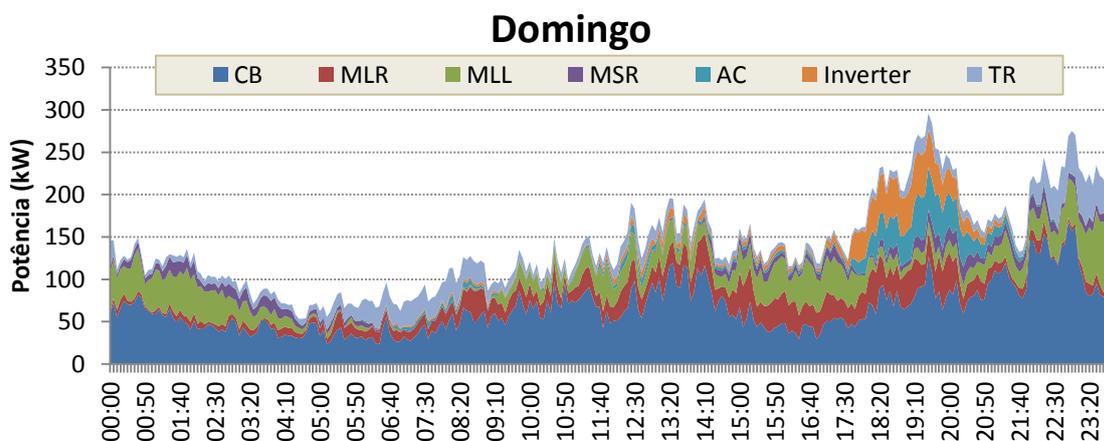


Figura 43: Diagrama de carga agregado dos consumidores do segundo cenário para o Domingo

É de notar que, apesar de os parâmetros das cargas controláveis terem sido definidos de forma a tornar possível a simulação deste tipo de consumidores, verifica-se à mesma um consumo mais elevado de carga base por volta da hora de almoço. Isto deve-se ao facto de a base de dados de consumos de carga base utilizada ser indiferenciada, contendo todos os tipos de consumidores.

No entanto, se o utilizador se encontrar na posse de consumos de carga base diferenciados por tipo de consumidor, é possível fazer a selecção dessa mesma base de dados na parametrização do SIMCAR.

5.3.2. Aplicação da estratégia de GP de reparametrização do termostato

Nas Figura 44 e Figura 45 foram simulados 500 consumidores residenciais. Neste grupo de consumidores existem 450 MLR, que realizam uma média de 4,9 ciclos por semana, sendo a distribuição de probabilidade destes ao longo da semana apresentado na Tabela 16. Relativamente aos programas de funcionamento das máquinas, a distribuição é dada por 6,6% para programas de lavagem a frio, 54,3% para lavagem a 30°C/40°C, 32,3% para lavagem a 60°C e para os programas de lavagem a 90°C é 6,8%.

Tabela 16: Distribuição dos ciclos da MLR, MLL e MSR ao longo da semana

	Segunda	Terça	Quarta	Quinta	Sexta	Sábado	Domingo
MLR	10%	10%	15%	10%	10%	20%	25%
MLL	10%	5%	15%	5%	20%	35%	10%
MSR	10%	10%	10%	10%	10%	25%	25%

O número de MLL simuladas neste cenário é 350, sendo o número médio de ciclos semanais de 4,2 com a distribuição apresentada na Tabela 16. A distribuição dos programas de funcionamento é a do estudo de Stamminger, que está apresentado na Tabela 5 [19].

Já nas MSR, sofreram simulação 200 máquinas, sendo o número médio de ciclos ao longo da semana de 4.

Nestas simulações foram utilizados 100 AC e 150 *inverters*. A distribuição dos perfis de utilização dos serviços de energia ao longo da semana é igual para ambos os tipos de carga, estando apresentado na Tabela 17.

Tabela 17: Distribuição dos perfis de utilização dos serviços de energia dos AC e *inverters* simulados

Distribuição dos perfis de utilização					
1	2	3	4	5	
15%	50%	20%	5%	10%	DU
15%	65%	85%	90%	100%	
15%	15%	50%	10%	10%	Sábado
15%	30%	80%	90%	100%	
15%	15%	60%	5%	5%	Domingo
15%	30%	90%	95%	100%	

Para a última carga, o termoacumulador, foram simulados 100 TR, cuja distribuição é dada pela Tabela 18.

Tabela 18: Distribuição dos perfis de utilização dos serviços de energia dos TR.

Distribuição dos perfis de utilização				
1	2	3	4	
60%	15%	15%	10%	DU
60%	75%	90%	100%	
60%	15%	15%	10%	Sábado
60%	75%	90%	100%	
60%	15%	15%	10%	Domingo
60%	75%	90%	100%	

O resultado obtido para a Quarta-Feira e para o Domingo estão apresentados na Figura 44 e Figura 45 respectivamente.

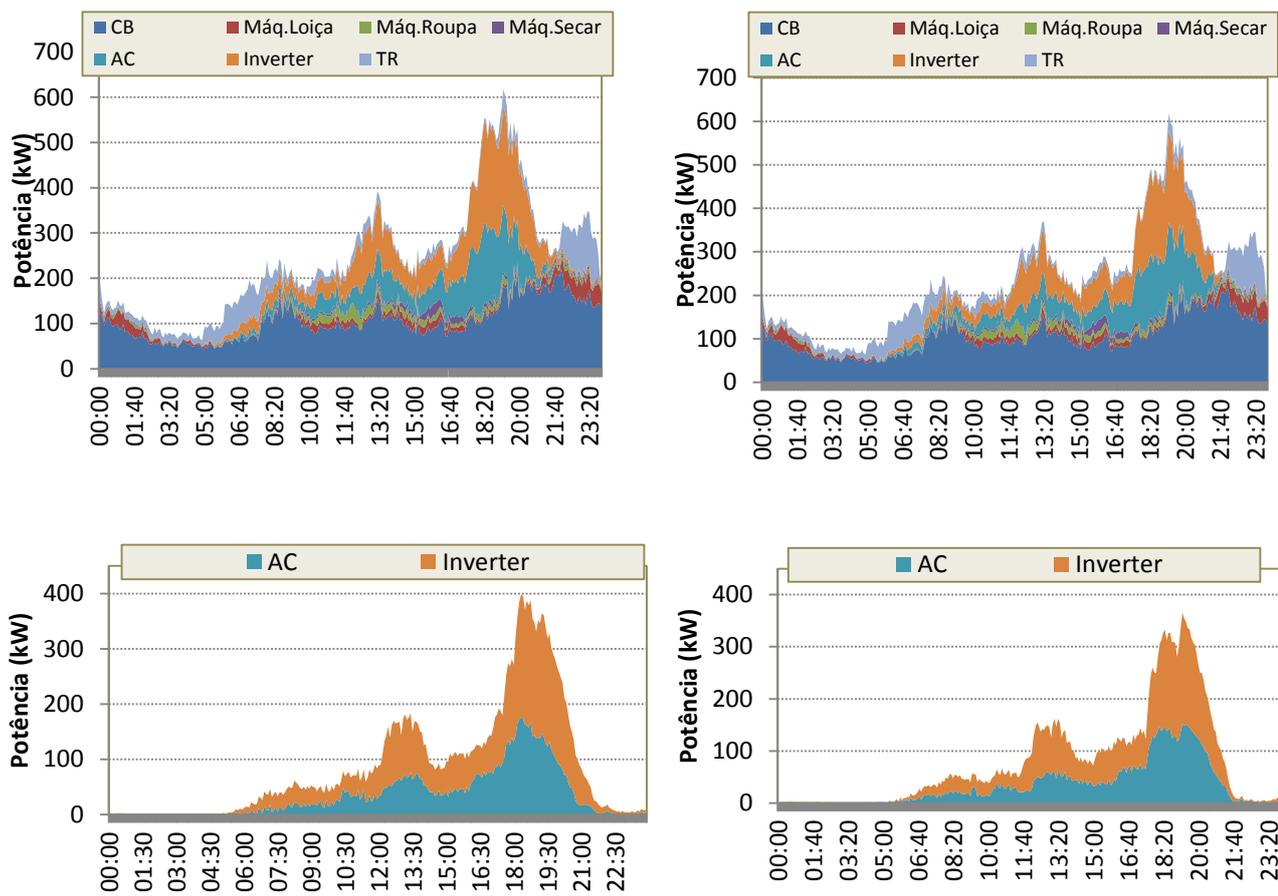


Figura 44: Comparação entre os consumos de Quarta-Feira para AC e Inverter parametrizados para 24°C (esquerda) e 25°C (direita)

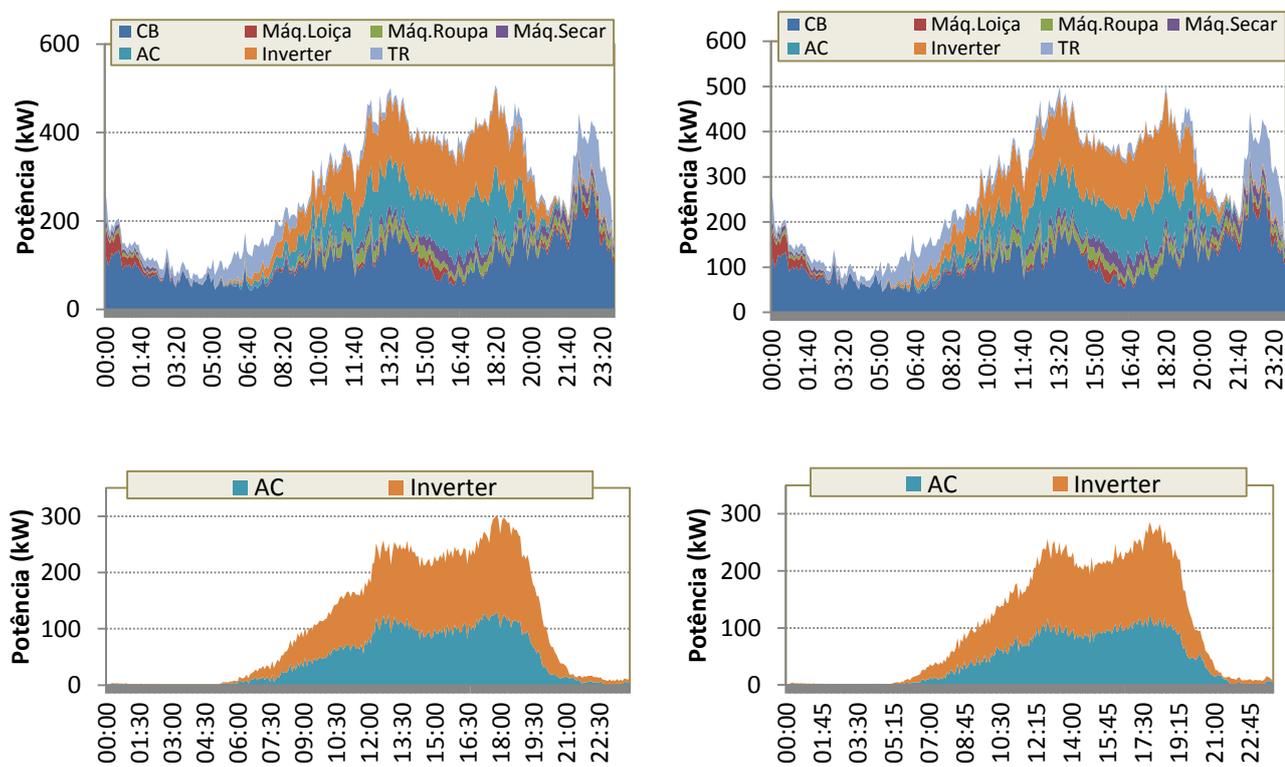


Figura 45: Comparação entre os consumos de Domingo para AC e Inverter parametrizados para 24°C (esquerda) e 25°C (direita)

As poupanças energéticas obtidas através desta reparametrização estão apresentadas na Tabela 19.

Tabela 19: Redução energética verificada com a reparametrização do AC e Inverter de 24°C para 25°C

<u>Energia Consumida</u>	Quarta-Feira		Domingo	
	AC	Inverter	AC	Inverter
Temperatura: 24	847,26 kWh	1.221,82 kWh	1.071,36 kWh	1.481,47 kWh
Temperatura: 25	830,76 kWh	1.076,15 kWh	1.041,87 kWh	1.434,05 kWh
Diferença	16,51 kWh	145,51 kWh	29,39 kWh	47,42 kWh
Total	162,02 kWh		76,81 kWh	

<u>Ponta máxima</u>	Quarta-Feira		Domingo	
	AC	Inverter	AC	Inverter
Temperatura: 24	175,798 kW	225,432 kW	128,824 kW	176,482 kW
Temperatura: 25	152,696 kW	212,385 kW	122,096 kW	166,289 kW
Diferença	23,100 kW	13,046 kW	6,728 kW	10,193 kW

Verifica-se assim que através desta medida de GP foi possível não só diminuir o consumo de energia em 162 kWh na Quarta-Feira e 76,81 kWh no Domingo, como foi também possível reduzir a ponta máxima destes dois tipos de cargas.

5.4. Simulação do diagrama de carga de um consumidor e avaliação do impacto na factura da alteração de consumos

Apesar de o SIMCAR não ter sido desenvolvido como ferramenta de calcular poupanças energéticas que os consumidores podem ter ao aderirem a determinadas estratégias de GP, foi feita uma pequena demonstração para um consumidor. Uma vez que a maioria dos consumidores apenas adoptaria as estratégias de GP se isto lhe trouxer benefícios directos, uma empresa prestadora de serviços, pode utilizar o simulador para representar um possível consumo do consumidor em causa, apresentando-lhe possíveis benefícios económicos que este poderá obter, para além dos incentivos monetários de algumas das medidas de GP.

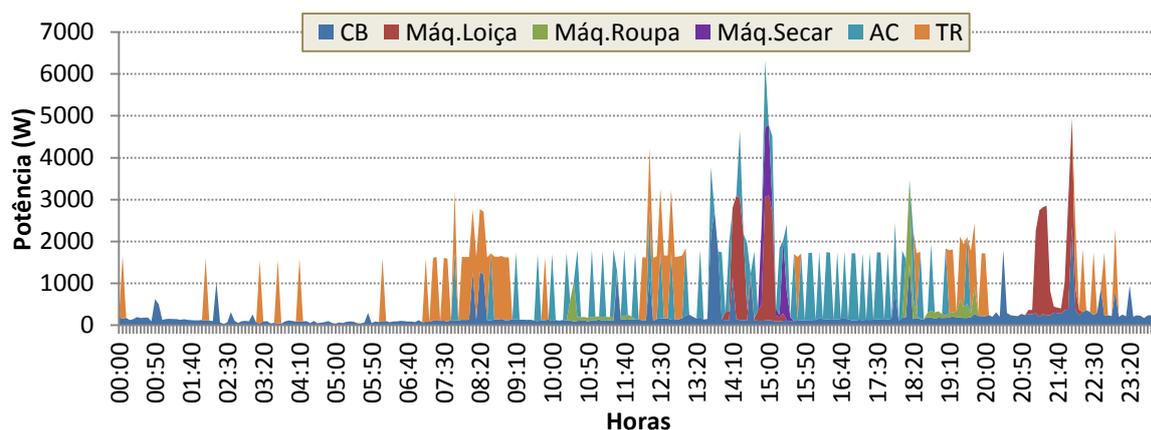


Figura 46: Diagrama de carga de um dia de um consumidor, sem a aplicação de nenhuma estratégia de GP

Para o consumidor em análise, consideraram-se o número médio de ciclos de funcionamento semanais da MLR, MLL e MSR como 4,9, 4 e 3 ciclos respectivamente, a potência do TR e como sendo de 1,5kW e 1,6kW a do AC. O consumo obtido para um determinado dia está apresentado na Figura 46.

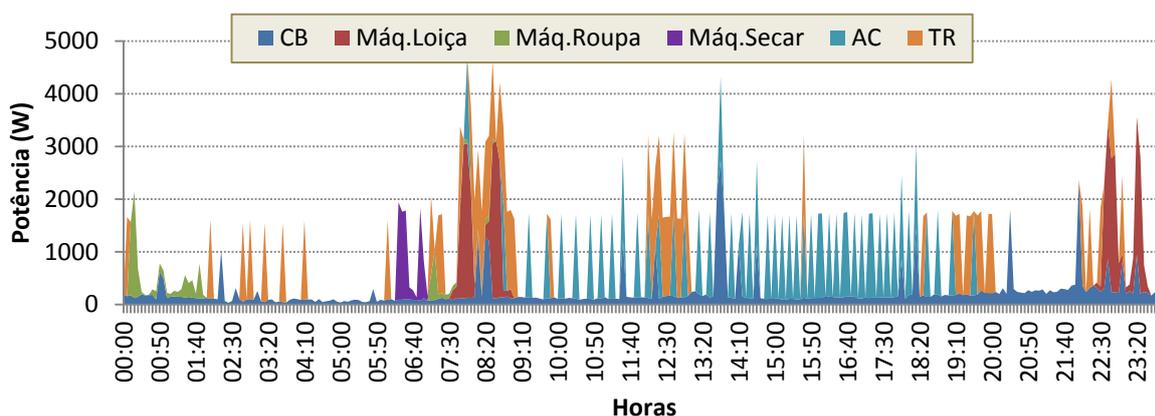


Figura 47: Diagrama de carga do mesmo consumidor da Figura 46, mas com a aplicação de estratégias de GP

Se o consumidor adoptar medidas de GP de controlo directo de cargas, bem como por um tarifário bi-horário (e utilizando o preço do kWh como 0,1551 €/kWh no período fora de vazio e 0,0833 €/kWh no período de vazio), numa situação em que a operadora de rede proceda ao controlo directo de cargas, sendo alguns dos consumos das máquinas desviados para horas de menor procura, é possível o consumidor obter poupanças na sua factura.

Tabela 20: Diferença nos consumos verificados com e sem o desvio de consumos

	Sem GP	Com GP	Gastos (Sem GP) (€)	Gastos (Com GP) (€)	Diferença
Consumo de energia no Período de Vazio (kWh)	5,146	8,744	0,4287	0,7284	0,2997
Consumo de energia no Período Fora de Vazio (kWh)	18,249	14,392	2,8304	2,2322	-0,5982
Total	23,395	23,136	3,2591	2,9606	-0,2985

Na situação da Figura 47, é possível poupar aproximadamente 0,30€ por dia. Extrapolando para o ano inteiro, supondo que esta poupança se mantém constante, o consumidor chega ao fim do ano com mais cerca de 110€. Caso seja possível este reduzir a potência contractada, poderá obter poupanças ainda mais significativas.

5.5. Obtenção de distribuições através dos algoritmos

Na Figura 48 pode-se comparar a distribuição de probabilidade de as MLR iniciarem um ciclo de lavagem de acordo com o estudo em DGGE / IP-3E, 2004 [10], com a distribuição obtida através da aplicação do algoritmo *k-means* à base de dados de consumos das MLR monitorizadas.

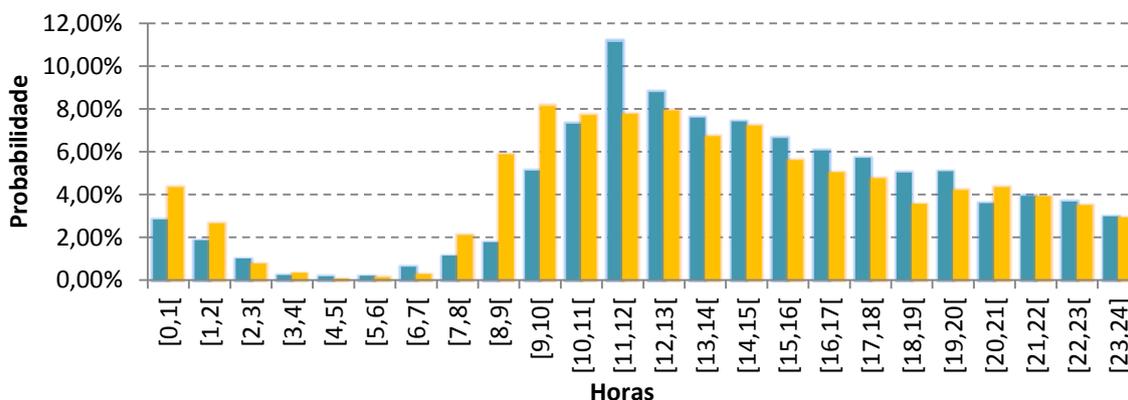


Figura 48: Distribuição de probabilidade do funcionamento da MLR segundo [10] (laranja) e distribuição de probabilidade do funcionamento da máquina de lavar roupa quando aplicado o algoritmo *k-means* (azul)

Pode-se observar que o resultado obtido para a MLR, quando se aplica o algoritmo aos consumos das 47 máquinas monitorizadas, é semelhante aos dados obtidos através do consumo nacional ([10]), o que significa que foi suficiente a quantidade das monitorizações das para abranger todos os tipos de perfis de consumidores. O mesmo se verifica para as MLL (ver Figura 49). Comparando os resultados obtidos através da aplicação do algoritmo aos consumos das 19 MLL monitorizadas com os dados obtidos através do consumo nacional ([10]) verifica-se que estes são semelhantes.

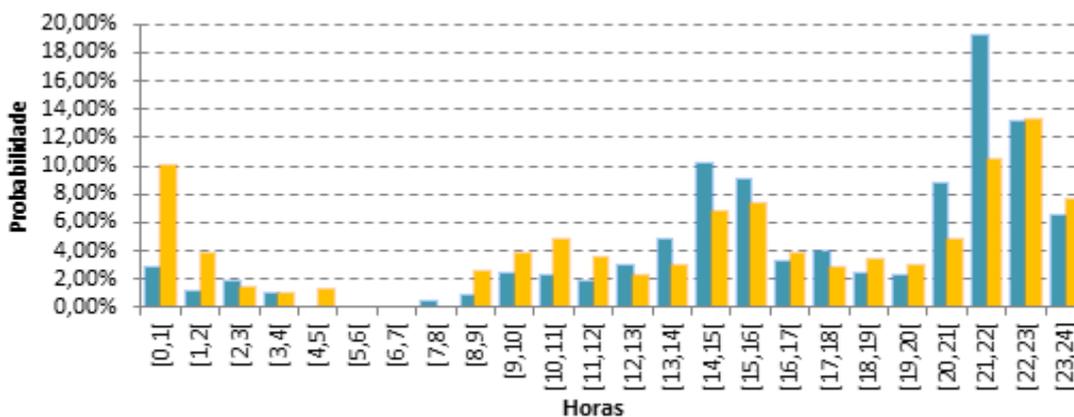


Figura 49: Distribuição de probabilidade do funcionamento da MLL segundo [10] (laranja) e distribuição de probabilidade do funcionamento da máquina de lavar roupa quando aplicado o algoritmo *k-means* (azul)

6. Conclusão e Futuros Trabalhos

O recurso directo aos diagramas de carga para a aplicação e avaliação das estratégias de GP torna-se uma tarefa bastante difícil, pelo que foi desenvolvido o SIMCAR com o intuito de recriar consumos residenciais realísticos.

A construção de diagramas de carga é feita através da simulação das cargas individualmente, sendo posteriormente aplicada uma abordagem *bottom-up* para construção do diagrama de carga final de um consumidor, sendo o processo repetido para todos os consumidores pretendidos. Um dos objectivos principais do SIMCAR ao gerar estes diagramas de carga, é tornar o estudo do impacto de medidas de GP mais prático e intuitivo, de modo a possibilitar uma adequada avaliação dos impactos que possam surgir aquando da aplicação das mesmas. Apesar de não ser essa a sua função principal, o SIMCAR pode também ser utilizado como ferramenta de demonstração para justificar ao consumidor final alguns investimentos que se propõe que a nível da aquisição de equipamentos de controlo, gestão de cargas e substituição de equipamentos mais velhos por outros mais eficientes.

O facto de a programação do SIMCAR ter sido realizado em blocos, demonstra-se ser uma faceta vantajosa deste simulador, uma vez que torna fácil para qualquer utilizador com alguns conhecimentos de programação em *Matlab*, testar novos algoritmos de simulação das diversas cargas, sem correr o risco de danificar o restante código. A sua grande flexibilidade no que toca à introdução de dados para alimentar o simulador e o modo funcionamento simples e intuitivo tornam o SIMCAR uma ferramenta prática e útil.

A metodologia utilizada para a construção de consumos de carga base demonstra ser capaz de, com base em poucos dados, gerar consumos de um qualquer número de consumidores. A incorporação de um MFB capaz de reproduzir consumos reais dos equipamentos termostáticos fornece ao utilizador a possibilidade de reprodução de diagramas de carga de situações mais específicas. Na alocação dos ciclos de funcionamento das MLR, MLL e MSR, através da utilização de *slots* ao longo do dia, é possível simular de forma realística o funcionamento de uma máquina com vários ciclos de funcionamento no mesmo dia. É também possível simular diferentes cenários, bastando para isso alterar a distribuição de probabilidade do funcionamento das mesmas.

Uma vez que alguns dos processos do SIMCAR podem demonstrar-se úteis em diversas situações alheias ao simulador, estes foram também desenvolvidos paralelamente, estando assim preparados para funcionarem de forma independente. É o caso do conversor de resoluções temporais e das diversas formas de cálculo da distribuição de probabilidades das diferentes

máquinas ao longo do dia – através da aplicação dos algoritmos de reconhecimento de padrões a uma base de dados ou utilização das calculadoras desenvolvidas em Excel.

Após a comparação dos dois algoritmos implementados, o algoritmo *k-means* demonstrou ser o melhor para esta situação, uma vez que, apesar de ambos apresentarem resultados semelhantes, o tempo de processamento deste foi muito superior.

Uma proposta de trabalho futuro pode relacionar-se com a aplicação de alterações ao SIMCAR, como é o caso da possibilidade de se definir a potência contractada dos consumidores, de forma a que, no processo de construção dos diagramas de carga de cada consumidor, esta limitação seja tida em conta. Também a eventual fusão do SIMCAR com um algoritmo de escalonamento de cargas, poderia trazer vantagens, uma vez que tornaria este simulador ainda mais versátil, no que toca à avaliação dos encargos mensais que cada consumidor tem.

Outra proposta para um trabalho futuro será a realização de uma desagregação mais exaustiva no que toca à carga base, tornando possível visualizar as contribuições de consumos de um maior leque de cargas.

Tendo em conta a grande diversidade e variabilidade existentes quando comparados os diagramas individuais e de carga base, a criação de métricas ou indicadores que tornem possível uma melhor comparação e validação destes, contribuiria para a robustez da aplicação.

Finalmente poder-se-á desenvolver um outro algoritmo de reconhecimento de padrões e compará-lo com os algoritmos utilizados, de modo a verificar qual o mais eficiente, tanto em termos de resultados obtidos como nos tempos de processamento. Com a implementação realizada, contrariamente ao que se esperava, o algoritmo de evolução diferencial não se demonstrou ser muito eficiente nesta situação, contudo aparenta possuir um grande potencial quando utilizado com a finalidade de *clusterização*. Sendo assim, uma investigação aprofundada nesta área poderá trazer resultados inovadores.

7. Bibliografia

- [1] ExxonMobil, 2012 The Outlook for Energy: A View to 2040, 2012.
- [2] “Eurostat,” [Online]. Available: <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>.
- [3] International Energy Association (IEA), 2012. [Online]. Available: www.iea.org.
- [4] INE / DGEG, Inquérito ao Consumo de Energia no Sector Doméstico 2010, Lisboa, 2011.
- [5] E. Carpaneto e G. Chicco, “Probabilistic Characterisation of the Aggregated Residential Load Patterns,” *IET Generation, Transmission & Distribution*, vol. 2, 2007.
- [6] J. Widén, A. Nilsson e E. Wäckelgård, “A combined Markov-chain and bottom-up approach to modelling of domestic lighting demand,” *Energy and Buildings*, vol. 41, pp. 1001-1012, 2009.
- [7] A. Capasso, W. Grattieri, R. Lamedica e A. Prudenzi, “A Bottom-Up Approach To Residential Load Modeling,” *IEEE Transactions on Power Systems*, vol. 9, 1994.
- [8] D. Gerbec, S. Gašperič, I. Šmon e F. Gubina, “An Approach to Customers Daily Load Profile Determination,” *IEEE*, vol. 1, pp. 587 - 591, 2002.
- [9] C. F. Walker e J. L. Pokoski, “Residential Load Shape Modeling Based on Customer Behavior,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, pp. 1703 - 1711, 1985.
- [10] DGGE / IP-3E, Eficiência energética em equipamentos e sistemas eléctricos no sector residencial, Lisboa, 2004.
- [11] A. T. Almeida e A. H. Rosenfeld, Demand Side Management and Electricity End-Use Efficiency, NASO ASI series, 1988, p. 99.
- [12] U.S. Department of Energy, “Benefits of Demand Response in electricity Markets and Recommendations for Achieving Them,” Fevereiro 2006.
- [13] Rocky Mountain Institute, “Demand Response: An Introduction,” Rocky Mountain Institute, Colorado, Abril 2006.
- [14] D. Violette, “Demand-Side Management (DSM): Future Role in Energy Markets,” em *Panel Discussion on Consumer Response to High Energy Prices*, Calgary, Alberta, 2007.
- [15] International Institute for Energy Conservation (IIEC), Demand Side Management, Best Practices Guidebook, Julho 2006.
- [16] The Brattle Group, “Dynamic Pricing: Past, Present and Future,” 2011.
- [17] A. Gomes, Simulação de Cargas Eléctricas no Contexto do Condicionamento da Procura de Energia

- Eléctrica, Coimbra: DEE-FCTUC, 1995.
- [18] J. Martinho, “Modelos de Cargas de Condicionamento Ambiente para Avaliação do Potencial de Demand Response,” DEEC-FCTUC, Coimbra, 2011.
- [19] R. Stamminger, *Synergy Potential of Smart Appliances*, 2008.
- [20] D. T. Larose, *Discovering Knowledge in Data*, New Jersey: Wiley, 2005.
- [21] L. S. Bora, “Data Mining and Ware Housing,” *IEEE*, 2011.
- [22] E. Mooi e M. Sarstedt, *A Concise Guide to Market Research*, Springer, 2011, p. Capítulo 9.
- [23] L. Kaufman e P. J. Rousseeuw, *Finding Groups in Data*, Wiley, 1990.
- [24] J. Oliveira, “Relatório de Progresso 1 - Data Mining,” Coimbra, 2012.
- [25] T. Chiu, D. Fang, J. Chen, Y. Wang e C. Jeris, “A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment”.
- [26] A. Gaspar-Cunha, R. Takahashi e C. H. Antunes, *Manual de Computação Evolutiva e Metaheurística*, Coimbra: Imprensa da Universidade de Coimbra/Editora da Universidade Federal de Minas Gerais, Junho 2012.
- [27] S. Das, A. Abraham e A. Konar, “Automatic Clustering Using an Improved Differential Evolution Algorithm,” *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS*, vol. 38, Janeiro 2008.
- [28] A. Abraham, S. Das e A. Konar, “Document Clustering Using Differential Evolution”.
- [29] W. Kwedlo, “A clustering method combining differential evolution with the K-means algorithm,” *Pattern Recognition Letters*, vol. 32, pp. 1613-1621, 2011.
- [30] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53-65, 1987.
- [31] Á. Gomes, H. Jorge, L. Neves, J. C. Gonçalves e A. M. Martins, “Rational Use of Energy Resources in Cities Through Demand-Side Management,” INESC.

Anexos

Anexo I - Aplicação do algoritmo *k-means*

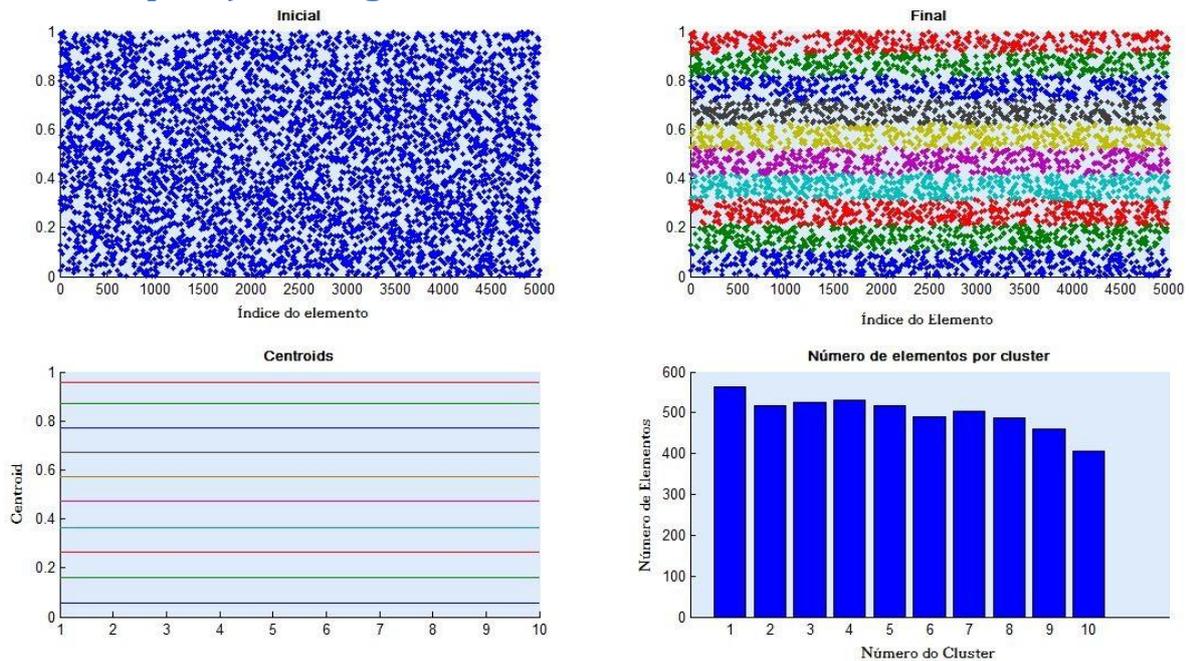


Figura 50: Aplicação do algoritmo *k-means*, com $k=10$, a uma base de dados de 5000 elementos, gerada de forma aleatória através de uma distribuição uniforme.

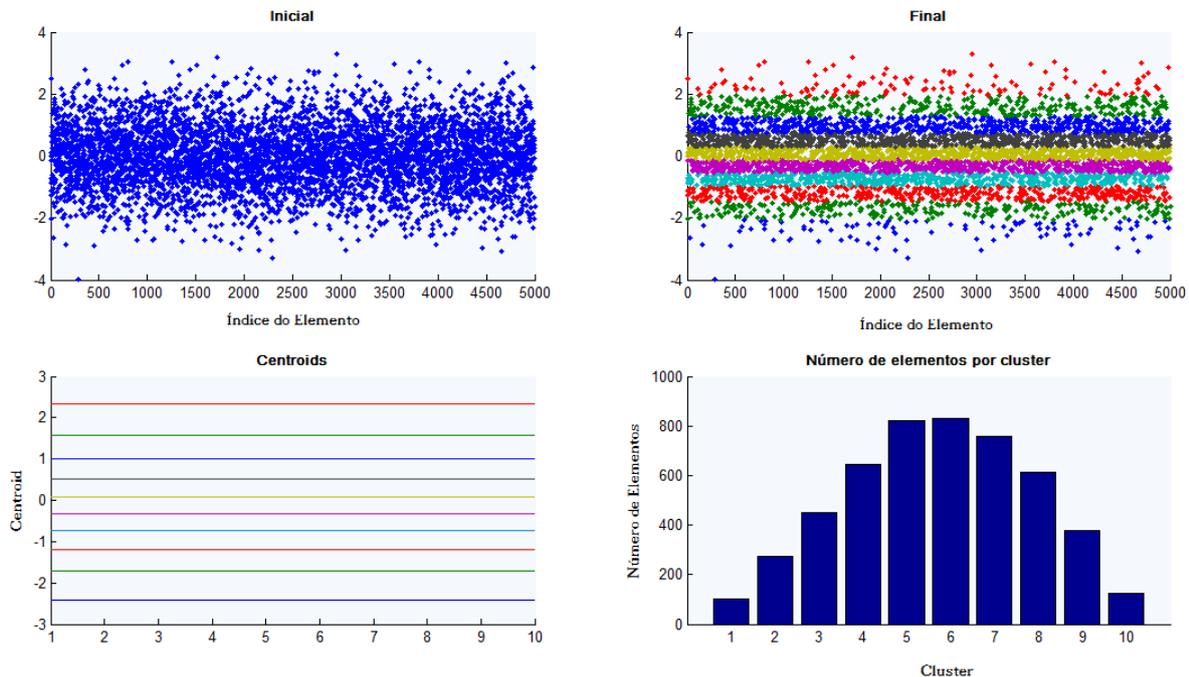


Figura 51: Aplicação do algoritmo *k-means*, com $k=10$, a uma base de dados de 5000 elementos, gerada através de uma distribuição normal com $\mu=0$ e $\sigma=1$

Os elementos em causa são unidimensionais, uma vez que a única variável existente é a potência. Na Figura 50 está exemplificada a aplicação do algoritmo a uma base de dados de 5000 elementos, gerada de forma aleatória através de uma distribuição uniforme e na Figura 51 a aplicação do algoritmo a uma outra base de dados contendo o mesmo número de elementos, mas gerados através de uma distribuição normal.

Anexo II – Aplicação do *k-means* e Evolução Diferencial e silhuetas obtidas

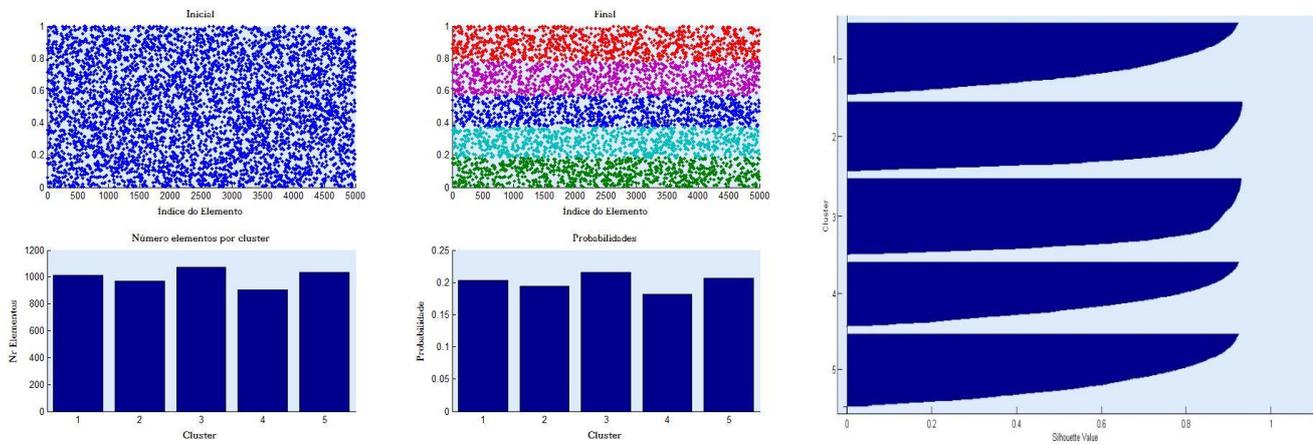


Figura 52: Resultados obtidos com o algoritmo de Evolução Diferencial com 5 clusters 5 cromossomas (esquerda) e a silhueta resultante (direita)

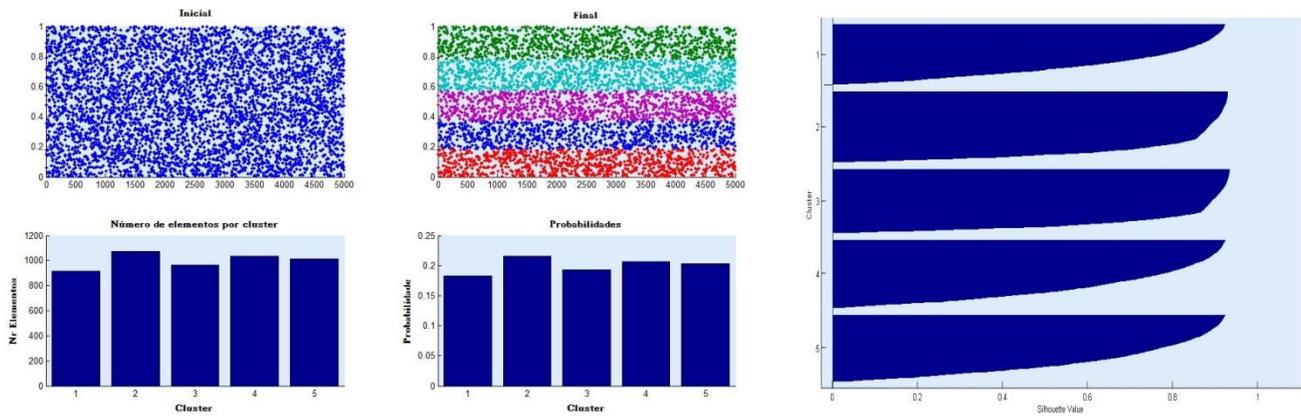


Figura 53: Resultados obtidos com o algoritmo de Evolução Diferencial com 5 clusters 10 cromossomas (esquerda) e a silhueta resultante (direita)

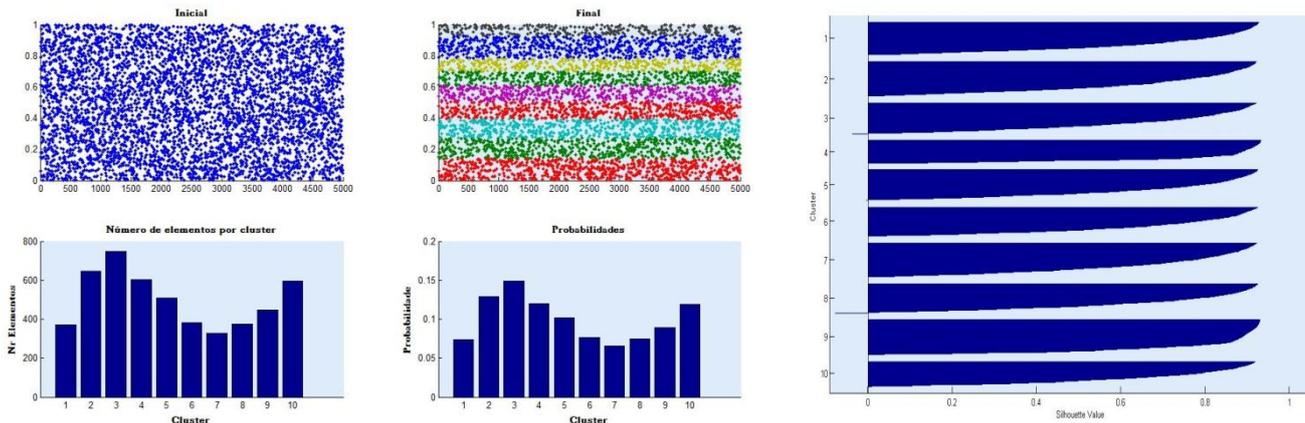


Figura 54: Resultados obtidos com o algoritmo de Evolução Diferencial com 10 clusters 5 cromossomas (esquerda) e a silhueta resultante (direita)

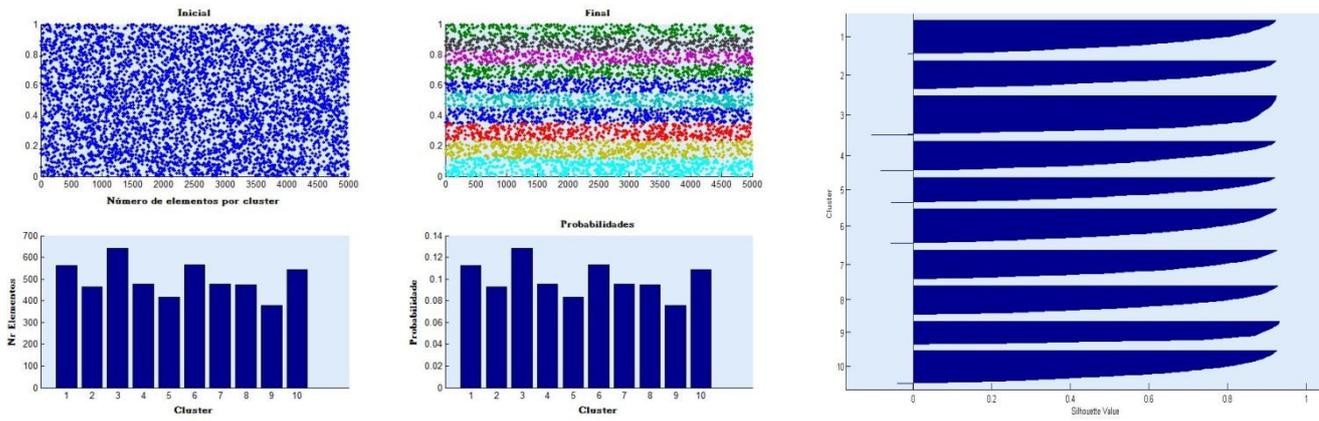


Figura 55: Resultados obtidos com o algoritmo de Evolução Diferencial com 10 *clusters* 10 cromossomas (esquerda) e a silhueta resultante (direita)

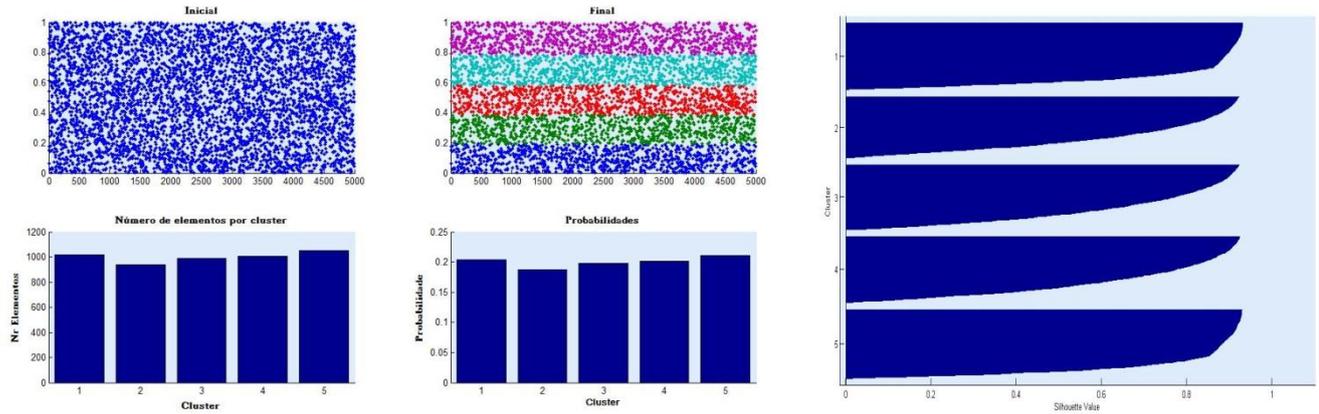


Figura 56: Resultados obtidos com o algoritmo *k-means* com 5 *clusters* (esquerda) e a silhueta resultante (direita)

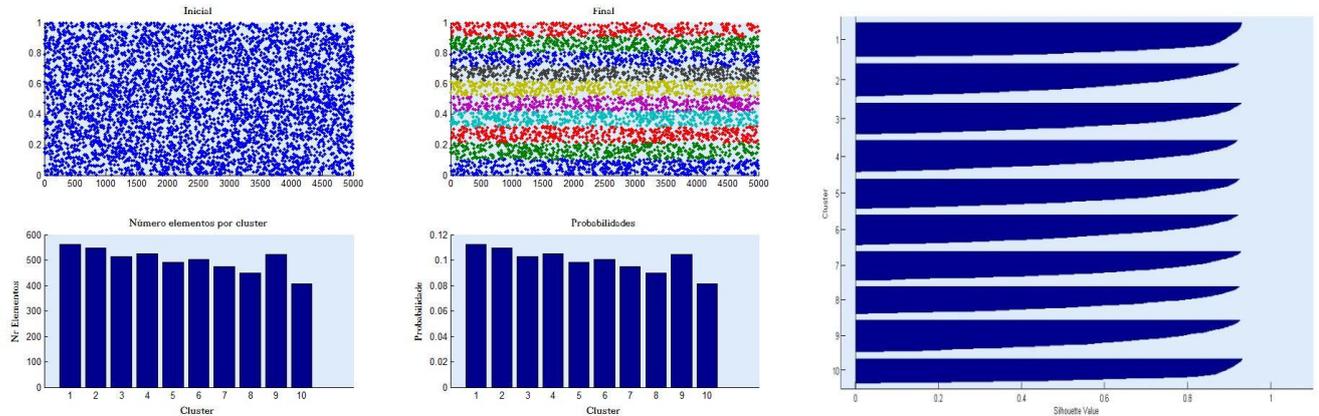


Figura 57: Resultados obtidos com o algoritmo *k-means* com 10 *clusters* (esquerda) e a silhueta resultante (direita)

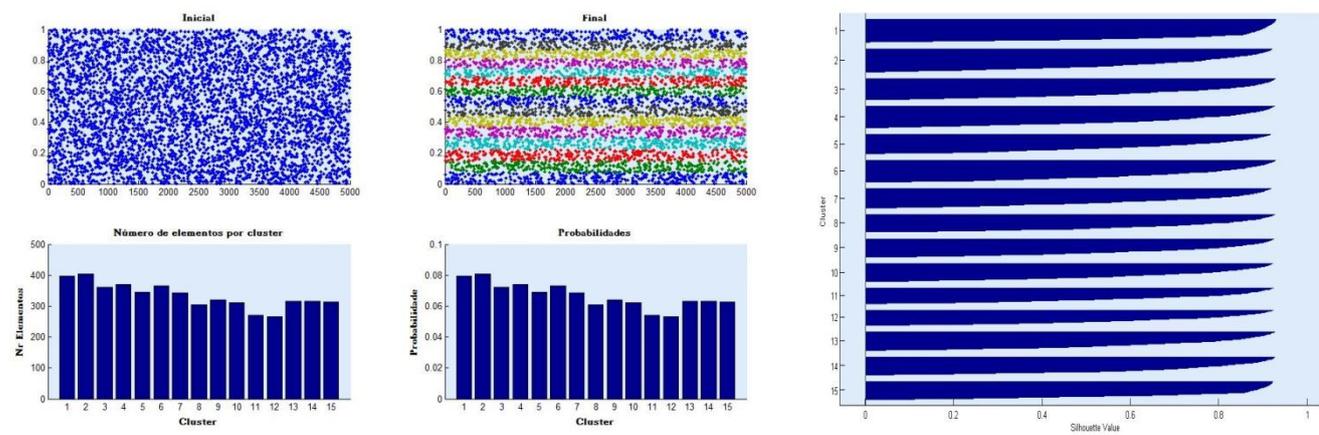
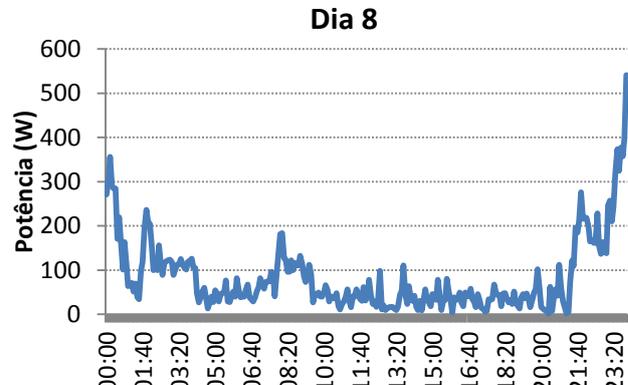
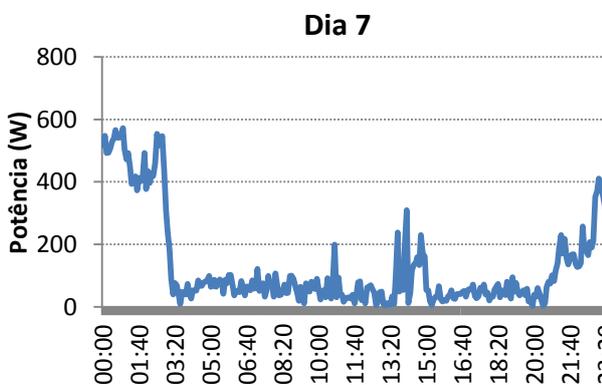
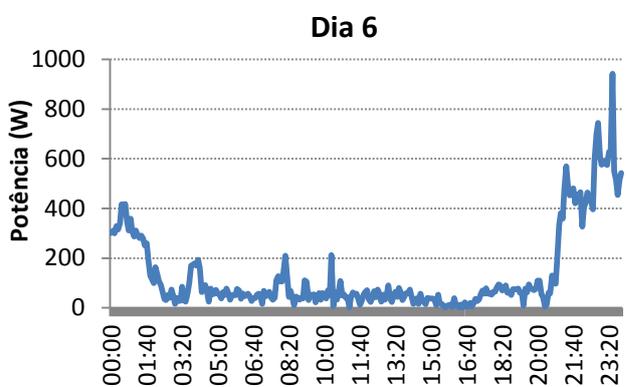
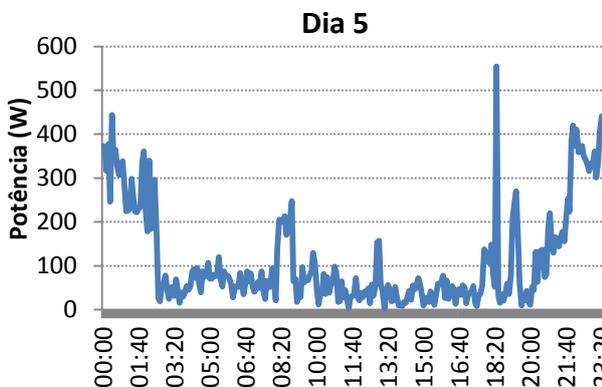
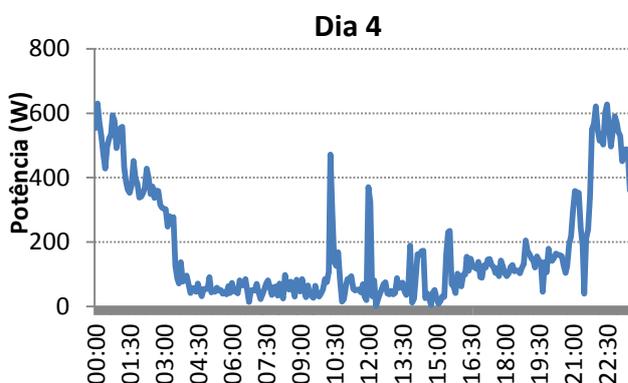
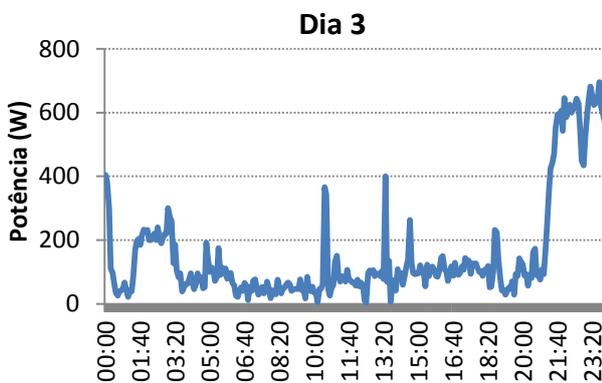
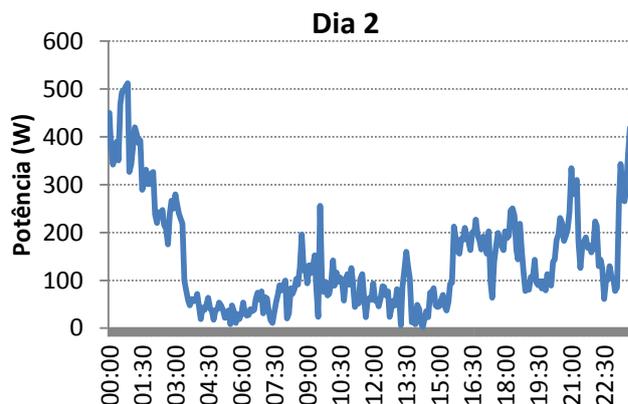
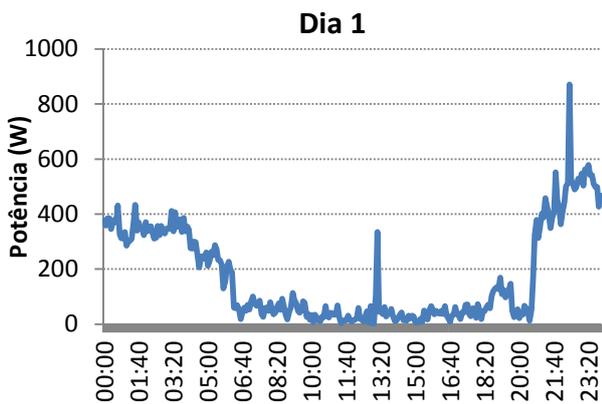


Figura 58: Resultados obtidos com o algoritmo *k-means* com 15 *clusters* (esquerda) e a silhueta resultante (direita)

Anexo III - Consumos de 14 dias de um consumidor



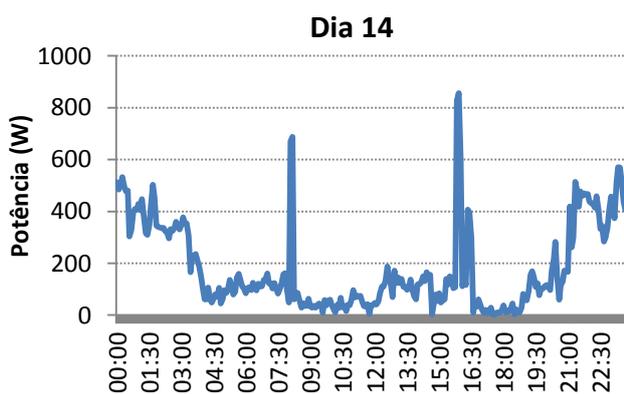
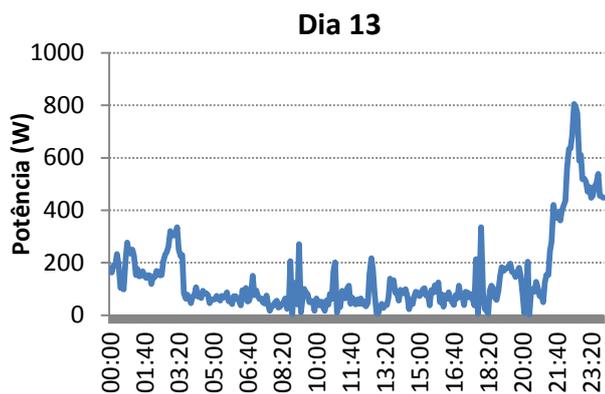
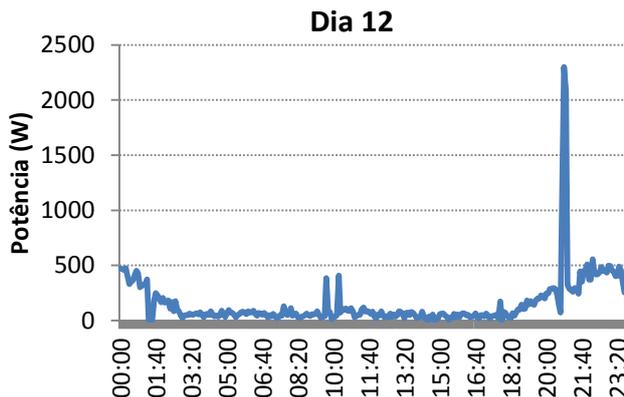
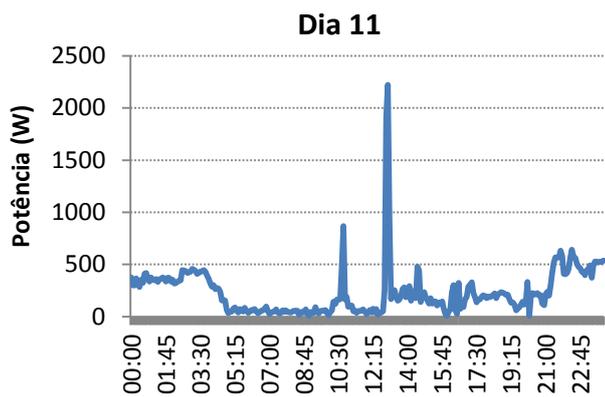
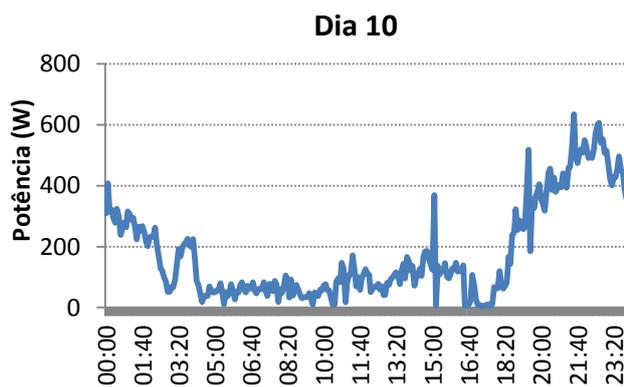
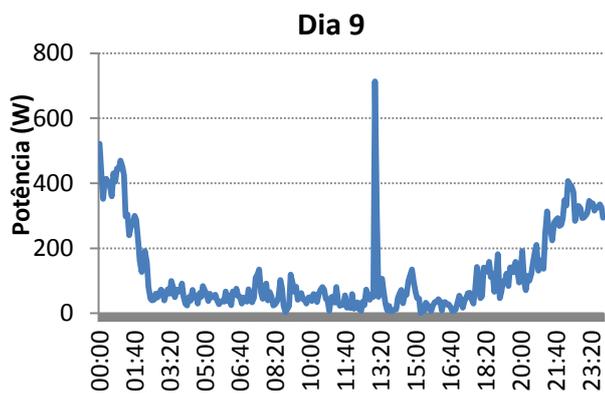


Figura 59: Diagramas de carga do consumidor monitorizado num período de 14 dias

Anexo IV - Caracterização das Cargas

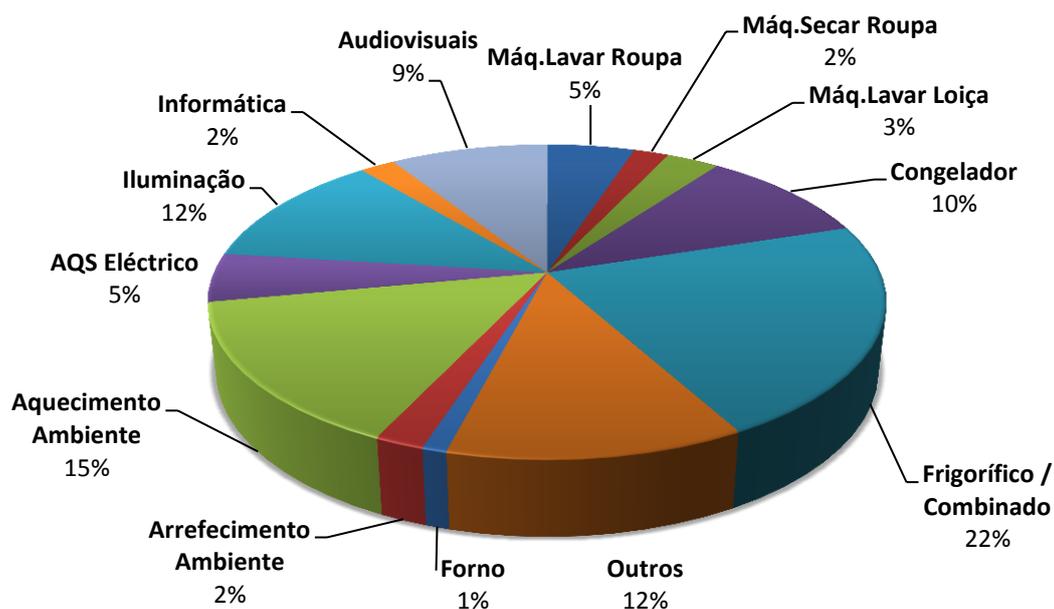


Figura 60: Repartição dos consumos eléctricos no sector residencial [10]

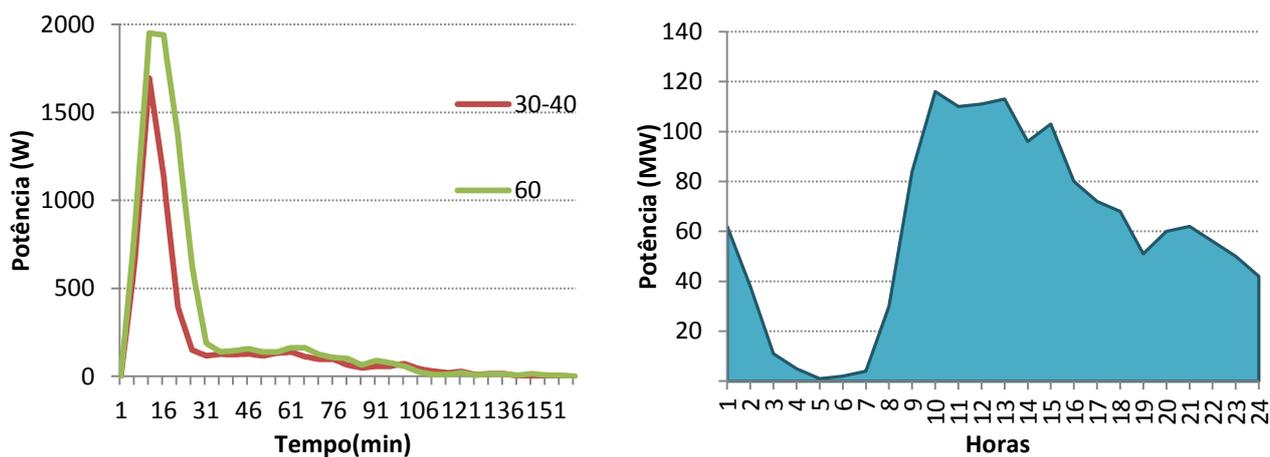


Figura 61: Ciclo típico de funcionamento de uma MLR num programa de 30°C/40°C e consumo nacional das MLR [10]

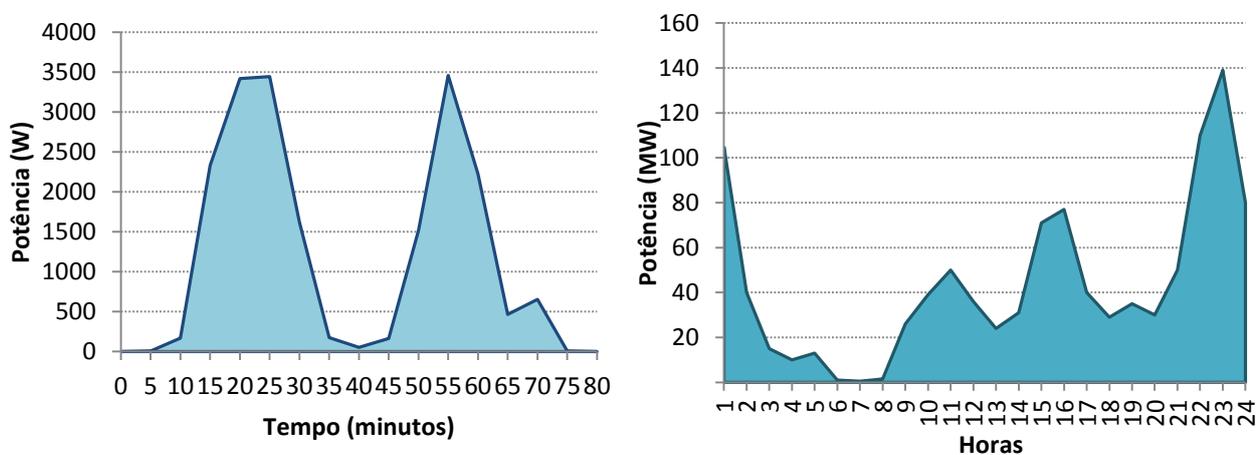


Figura 62: Ciclo típico de funcionamento de uma MLL e consumo nacional das MLL [10]

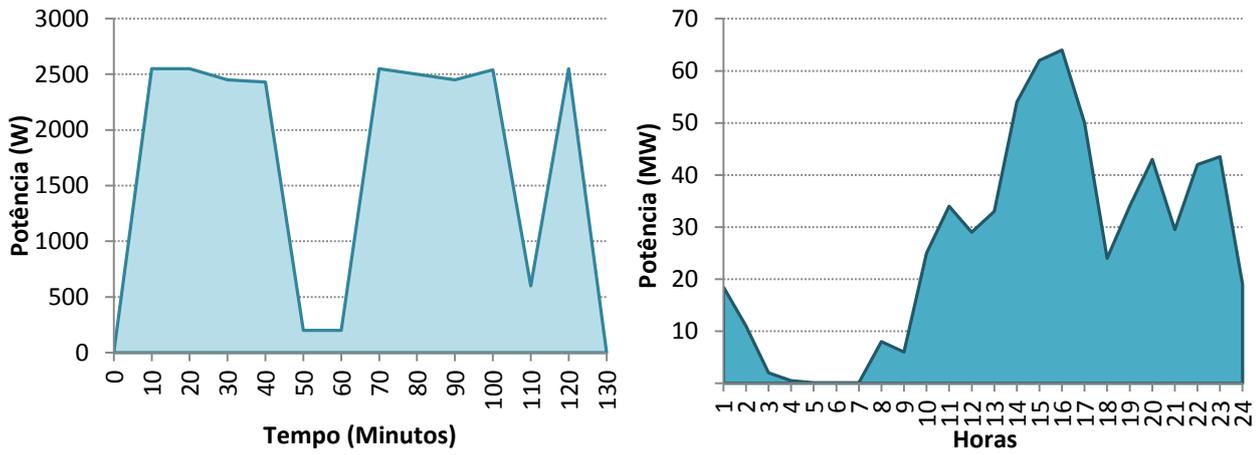


Figura 63: Ciclo típico de funcionamento de uma MSR e consumo nacional das MSR [10]

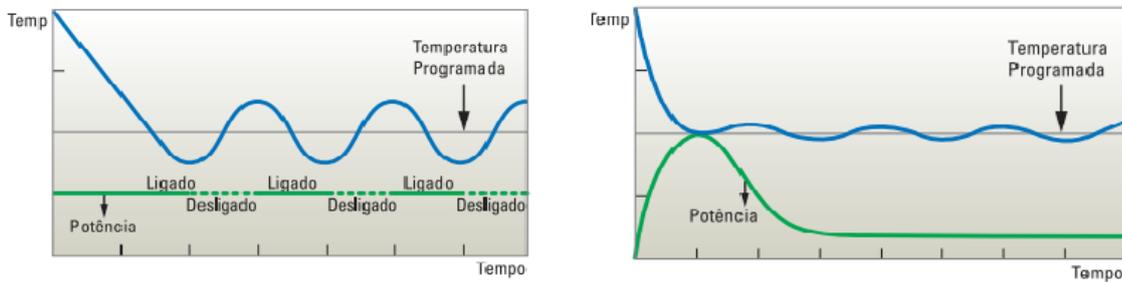


Figura 64: Diferença entre AC convencional (esquerda) e *inverter* (direita) [18]

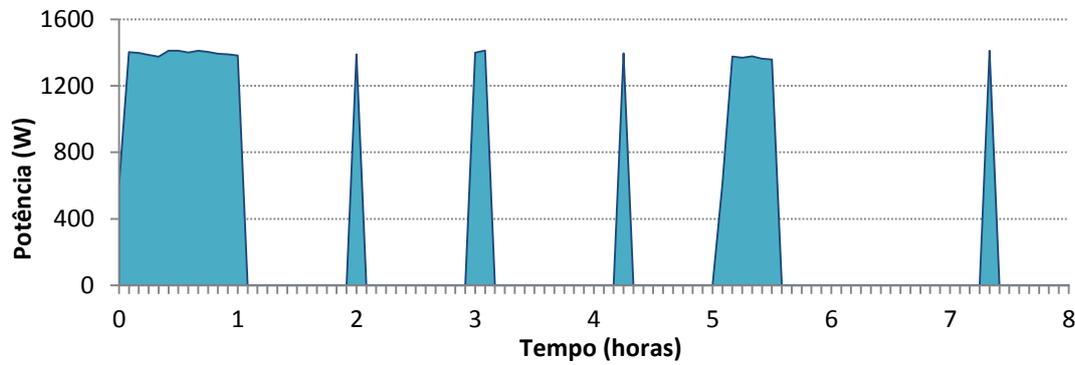


Figura 65: Funcionamento de um TR

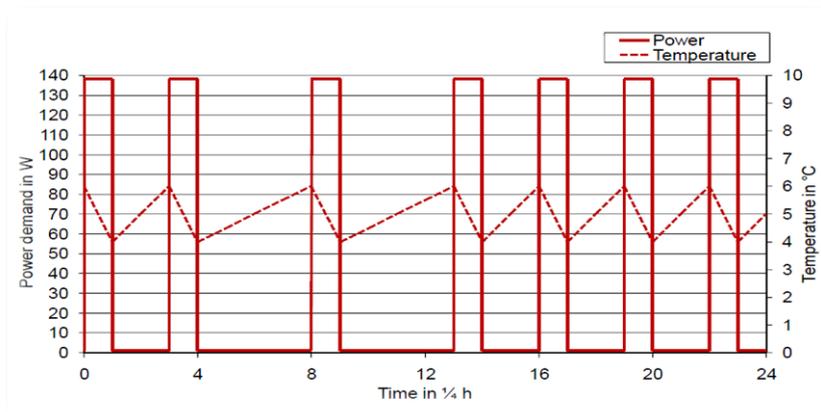


Figura 66: Funcionamento de um frigorífico durante 6 horas [19]

Anexo V – Parametrização do SIMCAR

Definições			
Dados a utilizar para a Carga Base			Default
Número de consumidores a simular:			300
Intervalo de Integração	5 min	Número de dias a simular	7
Algoritmo utilizado para a construção da carga base:		K-Means	
Localização onde guardar os resultados:			Default

Figura 67: Folha de parametrização “Geral”

Número de Máquinas de Lavar Roupa:	270	Taxa de Posse	90%	270	
Número médio de ciclos semanais:	5,5				20% 7
% de variação de número de ciclos semanais:	20%				-20% 4
% de variação de I:	0%				
% de variação dos programas de funcionamento:	0%				
Probabilidades do tipo de programa	Própria				
Valores	Default				

	Segunda Feira	Terça Feira	Quarta Feira	Quinta Feira	Sexta Feira	Sábado	Domingo	
I	5%	5%	5%	5%	15%	25%	40%	100%
II	0,3	0,3	0,3	0,3	0,8	1,4	2,2	
III	1	1	1	1	1	2	3	
IV	2	2	2	2	2	3	4	

Frio	30/40°C	60°C	90°C	
5%	25%	55%	15%	100%
5%	30%	85%	100%	

Tabela 2

Figura 68: Folha de parametrização “MLR”

Número de Máquinas de Secar Roupa:	150	Taxa de Posse	40%	120	
Número de ciclos semanais:	2				20% 2
% de Variacao de número de ciclos semanais:	20%				-20% 2
% de variação de I:	0				
% de variação dos programas de funcionamento:	0%				
Valores	Default				

	Segunda Feira	Terça Feira	Quarta Feira	Quinta Feira	Sexta Feira	Sábado	Domingo	
I	5%	5%	5%	5%	10%	30%	40%	100%
II	0,1	0,1	0,1	0,1	0,2	0,6	0,8	
III	1	1	1	1	1	1	1	
IV	2	2	2	2	2	2	2	

50 min	70 min	90 min	110 min	
12%	37%	36%	15%	100%
12%	49%	85%	100%	

Tabela 2

Figura 69: Folha de parametrização “MSR”

Número de Máquinas de Lavar Loíça	210	Taxa de Posse	70%	210	
Número de ciclos semanais:	5,5				10% 6
% de Variacao de número de ciclos semanais:	10%				-10% 5
% de variação de I:	0				
% de variação dos programas de funcionamento:	0				
Probabilidades do tipo de programa	Própria				
Valores	Default				

	Segunda Feira	Terça Feira	Quarta Feira	Quinta Feira	Sexta Feira	Sábado	Domingo	
I	5%	5%	5%	10%	20%	20%	35%	100%
II	0,3	0,3	0,3	0,6	1,1	1,1	1,9	
III	1	1	1	1	2	2	2	
IV	2	2	2	2	3	3	3	

35/45°C	50/55°C	65°C	70°C	
5%	55%	30%	10%	100%
5%	60%	90%	100%	

Tabela 2

Figura 70: Folha de parametrização “MLL”

Ar Condicionado a simular	
Número de Cargas	0
% Variação da Distribuição	0%
Valores a utilizar	Default

Criar Diagramas de Carga	
Perfil	5
Cargas	15

Distribuição dos perfis de utilização					
1	2	3	4	5	
10%	75%	5%	5%	5%	100%
10%	85%	90%	95%	100%	DU
25%	15%	40%	10%	10%	100%
25%	40%	80%	90%	100%	Sábado
10%	75%	5%	5%	5%	100%
10%	85%	90%	95%	100%	Domingo

Figura 71: Folha de parametrização “AC”.

A folha correspondente ao *Inverter* é semelhante à do Ar Condicionado.

Termoacumulador a simular	
Número de Cargas	0
% Variação da Distribuição	0%
Valores a utilizar	Default

Criar Diagramas de Carga	
Perfil	4
Cargas	15

Distribuição dos perfis de utilização				
1	2	3	4	
55%	25%	10%	10%	100%
55%	80%	90%	100%	DU
60%	15%	15%	10%	100%
60%	75%	90%	100%	Sábado
50%	40%	5%	5%	100%
50%	90%	95%	100%	Domingo

Figura 72: Folha de parametrização “TR”.

Anexo VI – Determinação do valor K do *k-means*

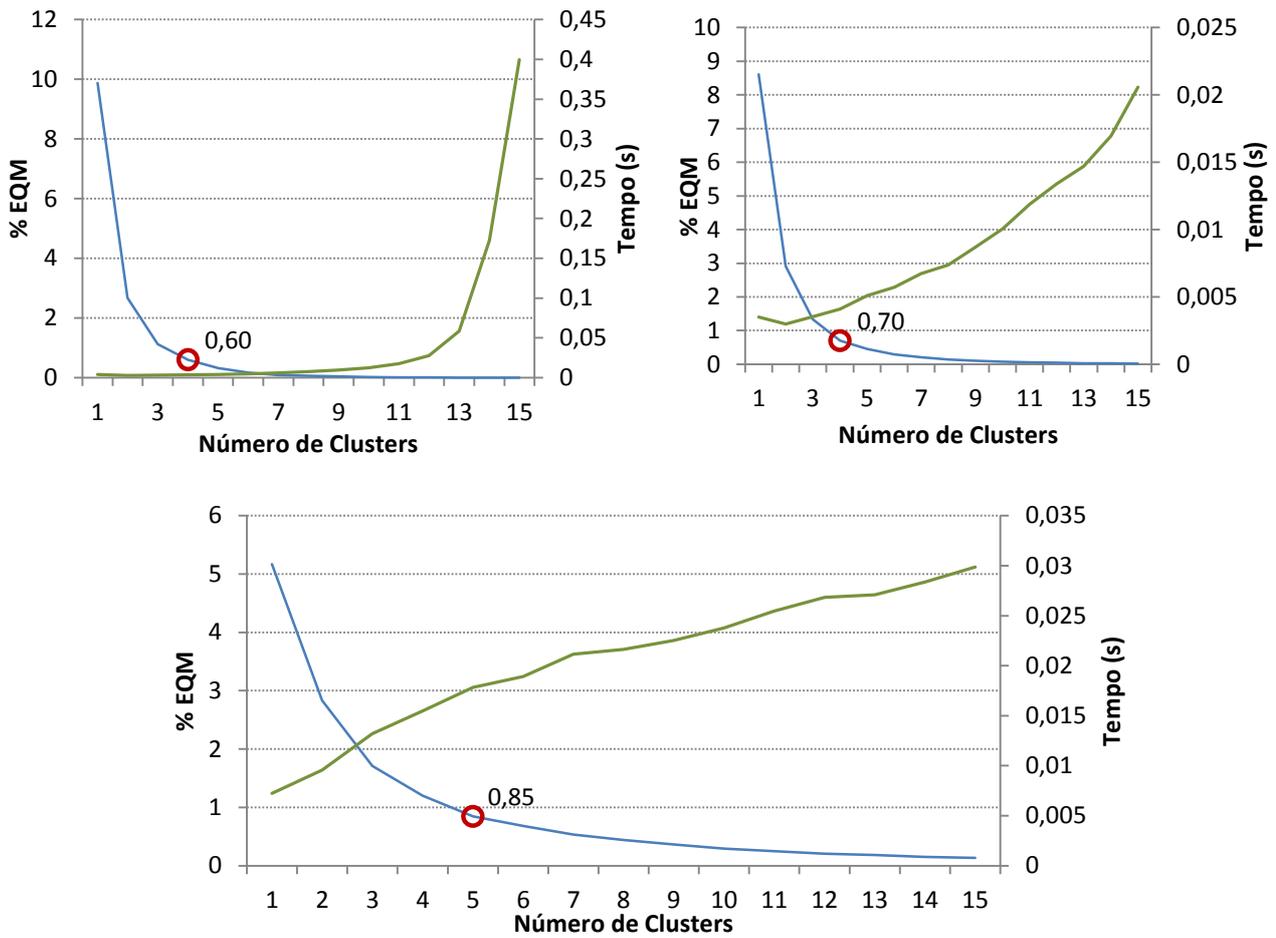


Figura 76: Determinação do número ideal de *clusters* a utilizar no para uma base de dados com 15 recolhas diária (esquerda), 30 recolhas (direita) e 185 recolhas (inferior)