

MODELO DE APOIO À DECISÃO PARA A UTILIZAÇÃO DE TIC NA OTIMIZAÇÃO DA RECOLHA DE RESÍDUOS RECICLÁ- VEIS

Dissertação submetida como requisito parcial para obtenção do grau
de Mestre em Gestão

Manuel Mendes Correia

Orientadora:

Prof. Doutora Tânia Rodrigues Pereira Ramos, Prof. Auxiliar,
Departamento de Engenharia e Gestão, Instituto Superior Técnico, Universidade de Lisboa

setembro de 2016

Modelo de apoio à decisão para a utilização de TIC na otimização
da recolha de resíduos recicláveis

RESUMO

A gestão de resíduos vive atualmente um paradigma de otimização dos seus processos, onde as tecnologias de informação e comunicação (TIC) têm um papel cada vez mais preponderante. No entanto, os custos associados à utilização das TIC nos processos de recolha de resíduos são ainda muito elevados, logo é necessário elaborar estratégias para uma integração mais sustentável dessas tecnologias nesta indústria.

No presente trabalho propõe-se um modelo de apoio à decisão de monitorização em tempo real das quantidades de resíduos recicláveis depositadas nos ecopontos tendo em vista a otimização das suas rotas de recolha. Dado o custo elevado de monitorizar todos os ecopontos com recurso às TIC, o modelo desenvolvido baseia-se na análise de agrupamentos de ecopontos com características semelhantes, efetuando a monitorização de uma amostra de cada *cluster* em detrimento da monitorização de todos os ecopontos. A incerteza associada à extrapolação dos dados das amostras para o total dos *clusters* leva à necessidade de se realizar uma análise de risco com recurso à simulação de Monte Carlo para poder avaliar eficazmente os benefícios desta solução. O modelo desenvolvido foi aplicado e validado com os dados cedidos por uma empresa real.

PALAVRAS-CHAVE

Gestão de resíduos recicláveis, TIC, análise de *clusters*, análise de risco, modelo de apoio à decisão.

ABSTRACT

Waste management addresses nowadays a paradigm of optimizing its processes in order to follow up efficiently the growth of recyclable waste production using at the same time information and communication technologies (ICT). The costs associated to the implementation of ICT in the processes of waste collection, even if enabling an optimization of the processes, are high, hence it is necessary to develop strategies to a sustainable integration of those technologies in that industry.

In the present work it is proposed a decision support model regarding the monitoring of recyclable waste containers with ICT aiming to the optimization of the collection routes. The model is based on the grouping of similar containers into clusters, and then monitoring a sample of each cluster, instead of monitoring all containers. The uncertainty associated to the extrapolation of data from the samples to the clusters implies that a risk analysis (with Monte Carlo simulation) must be carried out in order to effectively assess the benefits of the decision support model when applied to a real case.

KEYWORDS

Recyclable waste management, ICT, cluster analysis, risk analysis, decision support model.

“Pão suado, sabor redobrado.”

Ditado popular

AGRADECIMENTOS

Agradeço à minha mulher, Sofia, que sempre me apoiou nesta aventura de tirar mais um mestrado, agora em horário pós-laboral. A paciência pelas noites passadas no INDEG e os fins de semana de sol passados em casa a estudar ou a desenvolver esta tese. É a ela que dedico este trabalho.

Agradeço aos meus pais por, mais uma vez, me terem apoiado, de várias formas, nestas nova etapa dos meus estudos e da minha vida.

Agradeço à Professora Tânia Ramos que gentilmente aceitou orientar este trabalho e propôs o tema que aqui é desenvolvido. Agradeço adicionalmente o facto de ter aceitado continuar a orientar este trabalho mesmo após se ter mudado para o IST.

Agradeço aos meus amigos e colegas Roberto, Zé Miguel e ao melhor grupo de trabalho do Mestrado Executivo em Gestão de Projetos no INDEG (Bárbara, Lúcia, Maria João, Mónica, Nuno e Soraia).

Agradeço aos Professores Leandro Pereira e Carlos Jerónimo, que contribuíram para a minha formação em Gestão de Projetos, facto que me levou até ao INDEG e conseqüentemente ao ISCTE.

Agradeço a todos os meus colegas de trabalho da Rui Prata Ribeiro, Lda. por me terem facilitado os horários para a realização da tese. Um obrigado especial ao João e ao Rui.

Agradeço ao Sr. Paulo Rodrigues da Lipor pela amabilidade das informações prestadas sobre os ecopontos inteligentes em Vila do Conde.

Agradeço por fim à Valorsul por ter fornecido os dados necessários à aplicação do modelo desenvolvido.

ÍNDICE

CAPÍTULO 1. INTRODUÇÃO	1
1.1. ENQUADRAMENTO GERAL	1
1.2. OBJETIVOS DA DISSERTAÇÃO	3
1.3. ORGANIZAÇÃO DA DISSERTAÇÃO	3
CAPÍTULO 2. REVISÃO DE LITERATURA	5
2.1. GESTÃO DE RESÍDUOS RECICLÁVEIS	5
2.2. MONITORIZAÇÃO DO ENCHIMENTO DE CONTENTORES	7
2.2.1. <i>Identificação por radiofrequência (RFID)</i>	7
2.2.2. <i>Sensor de imagem</i>	8
2.2.3. <i>Sensor volumétrico</i>	9
2.2.4. <i>Sensor ótico</i>	10
2.2.5. <i>Ecoponto inteligente</i>	11
2.3. MODELOS DE APOIO À DECISÃO	12
2.3.1. <i>Modelos com incerteza pura</i>	13
2.3.2. <i>Modelos probabilísticos</i>	14
2.3.3. <i>Gestão do risco</i>	16
2.4. ANÁLISE DE <i>CLUSTERS</i>	18
2.4.1. <i>Métodos de análise de clusters</i>	19
2.4.2. <i>Medidas de proximidade</i>	20
2.4.3. <i>Agregação em clusters</i>	22

CAPÍTULO 3. METODOLOGIA	25
CAPÍTULO 4. MODELO DE APOIO À DECISÃO PARA A UTILIZAÇÃO DE TIC NA OTIMIZAÇÃO DA RECOLHA DE RESÍDUOS	27
4.1. ANÁLISE DE <i>CLUSTERS</i>	28
4.1.1. <i>Medidas de proximidade</i>	29
4.1.2. <i>Agregação em clusters</i>	30
4.2. ANÁLISE DE DADOS	31
4.2.1. <i>Técnicas de amostragem</i>	31
4.2.2. <i>Estimação de valores</i>	32
4.3. ANÁLISE DE RISCO	34
4.3.1. <i>Simulação de Monte Carlo</i>	35
CAPÍTULO 5. APLICAÇÃO DO MODELO A UM CASO DE ESTUDO REAL	39
5.1. ENQUADRAMENTO.....	39
5.1.1. <i>Situação atual</i>	41
5.1.2. <i>Dados iniciais</i>	42
5.1.3. <i>Monitorização de ecopontos</i>	43
5.2. ANÁLISE DE <i>CLUSTERS</i>	44
5.2.1. <i>Método gráfico</i>	44
5.2.2. <i>Método hierárquico</i>	45
5.2.3. <i>Método das k-médias</i>	46
5.2.4. <i>Caracterização dos clusters</i>	47
5.3. ANÁLISE DE DADOS	49

5.4.	ANÁLISE DE RISCO	54
5.4.1.	<i>Método hierárquico</i>	56
5.4.2.	<i>Método das k-médias</i>	58
5.4.3.	<i>Comparação das soluções</i>	60
5.5.	ANÁLISE CUSTO/BENEFÍCIO	60
CAPÍTULO 6.	CONCLUSÕES	63
6.1.	CONCLUSÕES	63
6.2.	LIMITAÇÕES DO MODELO E DO CASO DE ESTUDO	65
6.3.	PERSPETIVAS DE TRABALHO FUTURO.....	65
BIBLIOGRAFIA		67
ANEXOS		71
ANEXO A1 – CUSTO MENSAL DA SOLUÇÃO		71
ANEXO A2 – ANÁLISE DE SENSIBILIDADE		73
ANEXO A3 – SIMULAÇÃO DE MONTE CARLO		75

ÍNDICE DE FIGURAS

Figura 1.1 – Evolução dos resíduos reciclados recolhidos em Portugal entre 1998 e 2015. (Sociedade Ponto Verde, 2016).....	1
Figura 2.1 – Dimensão da gestão de resíduos. Adaptado de APA (2014).	5
Figura 2.2 – Ecoponto sobrelotado. Adaptado de Davide (2012).....	6
Figura 2.3 – Ilha ecológica em Cascais. Adaptado de Davide (2012)	6
Figura 2.4 – Esquema de funcionamento de um sistema RFID. Adaptado de Lelah et al. (2010)	7
Figura 2.5 – Utilização de sensores de imagem para monitorização de enchimento de contentores: (a) imagem real; (b) imagem binária. Adaptado de Arebey, Hannan e Basri (2011).	8
Figura 2.6 – Sistema de monitorização remota do enchimento de contentores com sensores de ultrassom. Adaptado de Lelah et al. (2012).	9
Figura 2.7 – Sensor de capacidade: (a) elétrodos coplanares; e (b) elétrodos frente a frente. Adaptado de Reverter; Gasulla e Pallas-Areny (2003).	10
Figura 2.8 – Sensor com infravermelhos. Adaptado de Gonçalves (2014).....	11
Figura 2.9 – Esquema da relação entre a incerteza e risco na tomada de decisões.....	13
Figura 2.10 – Esquema explicativo das probabilidades em função das ocorrências para os modelos com incerteza pura, probabilísticos e determinísticos.	13
Figura 2.11 – Processos da gestão de risco segundo o PMBoK.	16
Figura 2.12 – Exemplo de matriz de impacto.	17
Figura 2.13 - Etapas da análise de clusters adaptado de (Jain et al., 1999).	19
Figura 3.1 – Metodologia seguida.....	25

Figura 3.2 – Esquema representativo da curva do investimento e do erro em função da percentagem de ecopontos com sensor.	26
Figura 3.3 – Esquema representativa da curva do custo da solução (investimento + erro) em função da percentagem de ecopontos com sensor.....	26
Figura 4.1 – Representação esquemática do modelo de apoio à decisão desenvolvido.....	27
Figura 4.2 – Esquema explicativo de uma amostra aleatória recolhida de um universo de ecopontos.....	28
Figura 4.3 – Esquema explicativo de várias a amostras aleatórias recolhidas de um universo de ecopontos dividido em clusters.	28
Figura 4.4 – Janela do software SPSS para a análise de clusters pelo método hierárquico.	30
Figura 4.5 – Esquema da relação entre população e amostra.....	31
Figura 4.6 – Representação esquemática da seleção de amostras no modelo proposto.....	32
Figura 4.7 – Categorias de resultados do enchimento de ecopontos.....	33
Figura 4.8 – Esquema do ajuste da estimativa de uma amostra pequena à realidade.	33
Figura 4.9 – Esquema do ajuste da estimativa de uma amostra maior à realidade.	33
Figura 4.10 – Esquema explicativo do risco de recolher um ecoponto não apto.	34
Figura 4.11 – Esquema explicativo do risco de recolher um ecoponto sobrelotado.....	34
Figura 4.12 – Esquema da integração do risco no custo da solução.	36
Figura 4.13 – Esquema da frequência cumulativa em função do custo da solução.	36
Figura 5.1 – Área de intervenção da Valorsul com zona estudada a vermelho.	40
Figura 5.2 – Configuração típica de ecopontos.....	40
Figura 5.3 – Exemplo do cálculo do enchimento médio diário de um ecoponto.	42
Figura 5.4 – Exemplo do cálculo da média, desvio-padrão e coeficiente de variação do enchimento médio diário.....	42

Figura 5.5 – Sensor de ultrassons Enevo ONe (Enevo, 2016).	43
Figura 5.6 – Representação gráfica dos valores em função das medidas selecionadas.	45
Figura 5.7 – Representação gráfica dos valores com uma simulação do número de clusters..	45
Figura 5.8 – Representação gráfica de clusters com o critério do vizinho mais próximo.	45
Figura 5.9 – Representação gráfica de clusters com o critério do vizinho mais afastado.....	45
Figura 5.10 – Representação gráfica de clusters com o critério das médias.....	46
Figura 5.11 – Representação gráfica de clusters com o critério dos centroides.....	46
Figura 5.12 – Representação gráfica de clusters com o critério de Ward.....	47
Figura 5.13 – Representação gráfica de clusters com o método das k-médias.....	47
Figura 5.14 – Gráfico com comparação de número de ecopontos por cluster pelo método hierárquico.....	48
Figura 5.15 – Gráfico com comparação da média do enchimento de ecopontos por cluster pelo método hierárquico.....	48
Figura 5.16 – Gráfico com comparação do coef. de variação do enchimento de ecopontos por cluster pelo método hierárquico	48
Figura 5.17 – Gráfico com comparação da população por ecopontos por cluster pelo método hierárquico.....	48
Figura 5.18 – Gráfico com comparação de número de ecopontos por cluster pelo método das k-médias	49
Figura 5.19 – Gráfico com comparação da média do enchimento de ecopontos por cluster pelo método das k-médias.....	49
Figura 5.20 – Gráfico com comparação do coef. de variação do enchimento de ecopontos por cluster pelo método das k-médias	49

Figura 5.21 – Gráfico com comparação da população por ecopontos por cluster pelo método das k-médias.....	49
Figura 5.22 – Gráfico com o erro e o investimento mensal no sistema para proporções de monitorização de 1% a 100% no cluster 1 do método hierárquico.....	52
Figura 5.23 – Gráfico com o custo mensal do sistema para proporções de monitorização de 1% a 100% no cluster 1 do método hierárquico.....	52
Figura 5.24 – Curva S que relaciona o custo da solução com a confiança para o cluster 1 do método hierárquico.....	56
Figura 5.25 – Comparação das curvas com o custo da solução em função da confiança para o método hierárquico e para o método das k-médias.....	60
Figura A.1 – Custo mensal da solução Cluster 1 Método hierárquico.	71
Figura A.2 – Custo mensal da solução Cluster 2 Método hierárquico.	71
Figura A.3 – Custo mensal da solução Cluster 3 Método hierárquico.	71
Figura A.4 – Custo mensal da solução Cluster 4 Método hierárquico.	71
Figura A.5 – Custo mensal da solução Cluster5 Método hierárquico.	71
Figura A.6 – Custo mensal da solução Cluster 6 Método hierárquico.	71
Figura A.7 – Custo mensal da solução Cluster 1 Método das k-médias.	72
Figura A.8 – Custo mensal da solução Cluster 3 Método das k-médias.	72
Figura A.9 – Custo mensal da solução Cluster 4 Método das k-médias.	72
Figura A.10 – Custo mensal da solução Cluster 5 Método das k-médias.	72
Figura A.11 – Custo mensal da solução Cluster 6 Método das k-médias.	72
Figura A.12 – Custo mensal da solução Cluster 7 Método das k-médias.	72
Figura A.13 – Simulação de Monte Carlo para o Cluster 1 Método hierárquico.	75
Figura A.14 – Simulação de Monte Carlo para o Cluster 2 Método hierárquico.	75

Figura A.15 – Simulação de Monte Carlo para o Cluster 3 Método hierárquico.	75
Figura A.16 – Simulação de Monte Carlo para o Cluster 4 Método hierárquico.	75
Figura A.17 – Simulação de Monte Carlo para o Cluster 5 Método hierárquico.	75
Figura A.18 – Simulação de Monte Carlo para o Cluster 6 Método hierárquico.	75
Figura A.19 – Simulação de Monte Carlo para o Cluster 1 Método das k-médias.	76
Figura A.20 – Simulação de Monte Carlo para o Cluster 3 Método das k-médias.	76
Figura A.21 – Simulação de Monte Carlo para o Cluster 4 Método das k-médias	76
Figura A.22 – Simulação de Monte Carlo para o Cluster 5 Método das k-médias	76
Figura A.23 – Simulação de Monte Carlo para o Cluster 6 Método das k-médias	76
Figura A.24 – Simulação de Monte Carlo para o Cluster 7 Método das k-médias	76

ÍNDICE DE TABELAS

Tabela 2.1 – Propriedades das medidas de proximidade adaptado de Reis (2001).	21
Tabela 2.2 – Vantagens e desvantagens dos critérios de agregação para o método hierárquico (adaptado de Reis (2001)).....	23
Tabela 5.1 – Dados de rotas de recolha da Valorsul entre janeiro e setembro de 2013.	41
Tabela 5.2 – Dados de enchimento médio e percentagem de ecopontos aptos para recolha. ..	41
Tabela 5.3 – Resultados da otimização de rotas realizadas por Gonçalves (2014).....	43
Tabela 5.4 – Valores de monitorização de 5% e 75% ecopontos do cluster 1 com o método hierárquico.....	50
Tabela 5.5 – Valores da estimativa das amostras com 5% e 75% ecopontos monitorizados comparados com os valores reais para o cluster 1 com o método hierárquico. (Legenda: NA - Não apto; A - Apto; S - Sobrelotado).....	51
Tabela 5.6 – Cálculo dos valores de custos mensais para as amostras de 5% e 75% dos ecopontos monitorizados.....	52
Tabela 5.7 – Valores de custo mensal e número de ecopontos monitorizados para cada cluster para o método hierárquico.....	53
Tabela 5.8 – Valores de custo mensal e número de ecopontos monitorizados para cada cluster para o método das k-médias	53
Tabela 5.9 – Valores de custo mensal e número de ecopontos monitorizados para cada método de análise de clusters.	54
Tabela 5.10 – Coeficiente de variação de cada cluster do método hierárquico.	56
Tabela 5.11 – Custo da solução para diferentes níveis de confiança para o método hierárquico	57
Tabela 5.12 – Coeficiente de variação de cada cluster do método hierárquico.	58

Tabela 5.13 – Custo da solução para diferentes níveis de confiança para o método das k-médias.....	58
Tabela 5.14 – Análise de custo/benefício anual das várias soluções existentes.	61
Tabela A.1 – Análise de sensibilidade para os clusters do Método hierárquico	73
Tabela A.2 – Análise de sensibilidade para os clusters do Método das k-médias	74

CAPÍTULO 1. INTRODUÇÃO

1.1. ENQUADRAMENTO GERAL

De acordo com a Sociedade Ponto Verde (2016) 61% dos resíduos recicláveis foram recolhidos e tratados em 2015. Na Figura 1.1 apresenta-se um gráfico com a evolução entre 1998 e 2015 da quantidade de resíduos reciclados recolhidos em Portugal.

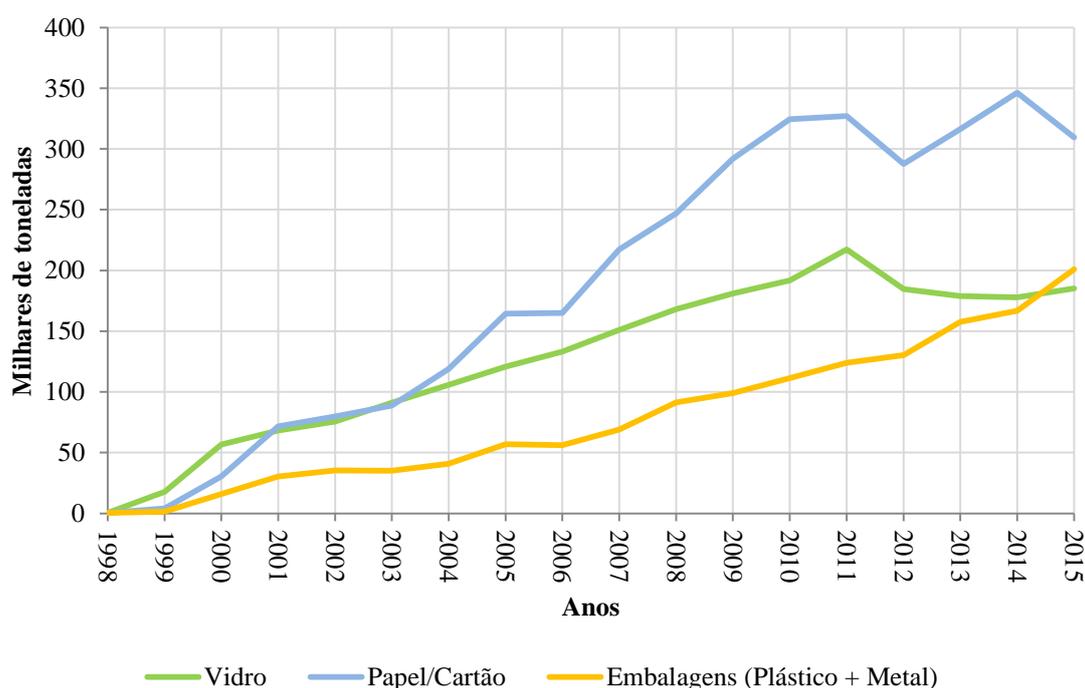


Figura 1.1 – Evolução dos resíduos reciclados recolhidos em Portugal entre 1998 e 2015. (Sociedade Ponto Verde, 2016).

Da observação do gráfico é possível depreender que a recolha de resíduos recicláveis em Portugal tem vindo a crescer de ano para ano, havendo, contudo, ainda espaço para evoluir. Porém, sem uma recolha eficaz, pode transformar-se num problema de escala ainda superior.

Num estudo realizado por Gonçalves (2014), foi identificado que na empresa de gestão de resíduos **Valorsul** – que opera essencialmente na zona oeste de Portugal – a grande maioria dos contentores de resíduos recicláveis – os ecopontos – eram recolhidos quando ainda não estavam aptos para a recolha (a Valorsul considera uma taxa de enchimento mínima de 0,75/1,00 para considerar o ecoponto apto). Isto significa que são percorridos muitos quilóme-

tros para recolher ecopontos com uma baixa taxa de enchimento, o que leva ao consumo de recursos de forma ineficiente.

O mesmo autor verificou que a introdução de **Tecnologias de Informação e Comunicação (TIC)** nos ecopontos, nomeadamente um sensor de ultrassons, para medir o seu enchimento em tempo real, poderia otimizar as rotas de recolha, evitando que ecopontos não aptos fossem recolhidos. Porém, os custos de implementação de um sensor por ecoponto revelaram-se inviáveis.

Assim, é necessário identificar formas de realizar a monitorização do enchimento dos ecopontos com um método menos oneroso, podendo a solução passar pela monitorização de alguns ecopontos em detrimento de outros, extrapolando os resultados de uma amostra para o universo total dos ecopontos a recolher. Essa extrapolação acarreta, naturalmente, riscos. *Risco de recolher um ecoponto que não esteja apto* (com uma taxa de enchimento inferior a um parâmetro δ), o que implica um acréscimo de quilómetros percorridos sem recolher o peso esperado, e portanto, uma ineficiência na operação; *risco de recolher um ecoponto que esteja sobrelotado* (com uma taxa de enchimento superior a 100%), o que implica a não recolha de outros ecopontos aptos por falta de capacidade do veículo que recolheu mais quantidade do que o previsto, para além de implicar um risco social e de saúde e higiene pública grave pela sobrelotação dos ecopontos; *risco de não recolher um ecoponto que esteja apto*, o que terá como consequência a sobrelotação do mesmo e as implicações já mencionadas para o caso da recolha de um ecoponto sobrelotado.

Na escolha do tema para a presente tese foram ponderados vários fatores, incluindo a motivação que os resultados deste projeto podem implicar.

A **sustentabilidade ambiental** é um fator fundamental de motivação para este trabalho – numa macroperspetiva – e que nos dias de hoje é um assunto que é altamente debatido. Os recursos finitos do nosso Planeta fazem com que tenham de existir mecanismos de reaproveitamento que garantam a sustentabilidade das sociedades; o ambiente não é exceção. A reciclagem de resíduos é um processo que começa na consciência de cada indivíduo, mas que apenas tem resultados práticos se houver uma estrutura organizada de recolha e tratamento. Nesse sentido, esta tese pretende dar o seu contributo para o melhoramento da forma como a recolha de resíduos recicláveis é realizada.

Outro fator de motivação é a identificação de uma solução com **adequação à realidade das empresas**. O desenvolvimento de soluções e de modelos de apoio à decisão deve ter em conta a realidade das empresas, desde a sua capacidade de investir num novo sistema, assim como a integração dessas mesmas soluções com os princípios orientadores da cada organização. As soluções e modelos devem procurar uma aplicabilidade real nas empresas, pois só assim se poderá medir o seu sucesso.

1.2. OBJETIVOS DA DISSERTAÇÃO

Tendo em conta a problemática tratada neste trabalho e enquadrando-a com a realidade das empresas, a presente dissertação tem como principais objetivos:

- Desenvolver um modelo que permita apoiar a decisão de utilização de TIC na otimização de resíduos recicláveis. Este modelo deve permitir decidir qual a dimensão da amostra de ecopontos que devem ser monitorizados através de TIC, tendo em conta o *trade-off* entre o custo de utilização das TIC e os riscos associados à não monitorização do universo de ecopontos.
- Aplicação do modelo de apoio à decisão ao caso de estudo da Valorsul, analisando o custo/benefício de vários cenários, incluindo a situação atual (sem monitorização) e a solução identificada por Gonçalves (2014) (monitorização de todos os ecopontos).

1.3. ORGANIZAÇÃO DA DISSERTAÇÃO

O tema desenvolvido nesta dissertação foi dividido em seis capítulos. O presente capítulo tem carácter introdutório, iniciando-se com um breve enquadramento geral do âmbito deste trabalho no domínio da gestão de resíduos, com uma descrição da problemática de investigação, seguida de uma breve apresentação dos objetivos do trabalho e uma descrição da organização do documento.

No segundo capítulo fez-se uma revisão bibliográfica sobre o tema, com particular enfoque na gestão de resíduos recicláveis e na forma como esta se enquadra no panorama socioeconómico atual e quais os meios existentes para a monitorização do enchimento de contentores. Fez-se ainda um levantamento de métodos de apoio à decisão, terminando com uma abordagem à

análise de *clusters*, por forma a identificar vários métodos que permitam auxiliar a construção do modelo que se pretende desenvolver com esta tese.

No terceiro capítulo é explicada a metodologia de abordagem ao problema. No quarto capítulo é detalhado o modelo desenvolvido para resolver a problemática identificada da monitorização de ecopontos com recurso a TIC. O modelo recorre a uma análise de *clusters* para agrupar os ecopontos, a uma análise de dados para caracterizar as amostras e a uma análise de risco para ter em conta as imprevisibilidades da operação.

No quinto capítulo são apresentados os resultados da aplicação do modelo apresentado no capítulo anterior a um caso de estudo real, o da Valorsul, realizando uma análise de custo/benefício por forma a comparar as soluções obtidas pelo modelo de apoio à decisão proposto com outros cenários já estudados.

Por fim, no sexto capítulo são tecidas algumas conclusões relativamente ao modelo de apoio à decisão desenvolvido bem como da sua aplicação ao caso de estudo. São ainda feitas algumas considerações relativamente a perspetivas de trabalho futuro.

CAPÍTULO 2. REVISÃO DE LITERATURA

No presente capítulo procurou-se fazer uma recolha da informação existente sobre os assuntos relacionados com a temática abordada neste trabalho. Entendeu-se que a revisão de literatura deveria partir do geral – a gestão de resíduos recicláveis – até ao particular, ou seja, as tecnologias de informação e comunicação existentes para a monitorização, assim como os modelos de apoio à decisão e a análise de *clusters*.

2.1. GESTÃO DE RESÍDUOS RECICLÁVEIS

A **gestão de resíduos** é, de acordo com a Diretiva Europeia n.º 2006/12/CE, o conjunto de atividades de recolha, transporte, valorização e eliminação de resíduos, incluindo o controlo destas operações e dos locais de descarga. Por sua vez, em Portugal, o Decreto-Lei n.º178/2006 acrescenta a essa diretiva as atividades de armazenagem, triagem, tratamento e as operações de descontaminação de solos dos locais de descarga.

De acordo com a Associação Portuguesa do Ambiente (2014), para além das atividades acima descritas deverão ainda ser consideradas na gestão de resíduos a prevenção, a regulamentação, a legislação e políticas ambientais (Figura 2.1).



Figura 2.1 – Dimensão da gestão de resíduos. Adaptado de APA (2014).

Desta forma é possível identificar os elementos que podem ser objeto de otimização na gestão de resíduos (recicláveis). Todos estes elementos são indissociáveis, logo a manipulação de um fator pode ter consequências nas restantes.

Num estudo levado a cabo na Turquia, foram observadas alterações significativas na geração de resíduos dependendo da densidade populacional, tipo de consumo dos utilizadores, sazonalidade e fatores demográficos (Berkun *et al.*, 2011).

Por outro lado, num estudo realizado por Shan-shan e Lo (2004), verificou-se que existe uma correlação entre a reciclagem e a consciência ambiental dos utilizadores (*i.e.* da população) com algumas variáveis demográficas e sociais (idade, nível de educação, existência de boas práticas por parte de personalidades, etc.). Segundo esses autores é um fator a ter em consideração na análise de tendências e deve ser alvo de estudo e de medição.

Segundo Davide (2012) alguns dos fatores dissuasores da (maior) adesão da população à reciclagem é a imagem negativa deixada passar pela utilização abusiva de ecopontos (Figura 2.2), que para além de esteticamente negativa, emana cheiros e é potenciadora de vandalismo (*e.g.* incêndios). O mesmo autor cita Luís Marau (Departamento Técnico e I&D da EMAC – Empresa de Ambiente de Cascais) que refere que as ilhas ecológicas – ecopontos enterrados – solucionam grande parte desses problemas e minimizam essa imagem negativa associada aos ecopontos sobrelotados (Figura 2.3).



Figura 2.2 – Ecoponto sobrelotado. Adaptado de Davide (2012)



Figura 2.3 – Ilha ecológica em Cascais. Adaptado de Davide (2012)

No que diz respeito mais concretamente ao processo de recolha, segundo Arebey *et al.* (2011) os principais problemas são os seguintes:

- Falta de informação sobre a área abrangida e tempo das rotas de recolha;

- Carência de um sistema capaz de monitorizar em tempo real veículos e contentores recolhidos;
- Ausência de estimativas do enchimento dos contentores a recolher;
- Falta de resposta rápida para casos urgentes como avaria/acidente de um veículo de recolha ou atraso na rota.
- Falta de resposta rápida para reclamações de utilizadores devido a resíduos não recolhidos.

O presente trabalho debruça-se essencialmente sobre a problemática da ausência de **estimativas sobre o enchimento dos ecopontos** a recolher.

2.2. MONITORIZAÇÃO DO ENCHIMENTO DE CONTENTORES

Na literatura existente são exploradas diversas **tecnologias de informação e comunicação (TIC)** que permitem monitorizar o enchimento de contentores de resíduos tentando dessa forma colmatar a ausência de estimativas, identificadas no capítulo anterior.

2.2.1. IDENTIFICAÇÃO POR RADIOFREQUÊNCIA (RFID)

O sistema de identificação por radiofrequência (RFID da sigla inglesa *Radio Frequency Identification*) foi um sistema desenvolvido durante a II Guerra Mundial que permite identificar objetos sem que seja necessário contacto, utilizando ondas de rádio (Lelah *et al.*,2010). A atribuição de uma etiqueta ao objeto (*i.e.* contentor) permite saber várias informações sobre o mesmo através da leitura do mesmo com equipamento próprio instalado nos camiões de recolha (ver Figura 2.4).

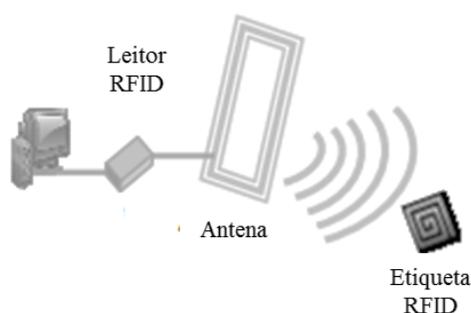


Figura 2.4 – Esquema de funcionamento de um sistema RFID. Adaptado de Lelah *et al.* (2010)

Uma das grandes vantagens desta tecnologia é o facto de não necessitar de uma fonte de energia para funcionar. Contudo, as etiquetas RFID permitem apenas guardar informação e transmiti-la através de um leitor, não fazendo a monitorização do enchimento em si (Hannan *et al.*, 2009).

2.2.2. SENSOR DE IMAGEM

Arebey, Hannan e Basri (2011) realizaram um estudo recorrendo a câmaras de vídeo que recolhiam imagens de contentores de lixo doméstico. As câmaras foram estrategicamente posicionadas de forma a abrangerem os 3m² circundantes do contentor, uma vez que o enchimento pode exceder a capacidade do contentor (Rovetta *et al.*, 2009).

As imagens recolhidas para vários enchimentos são recolhidas e tratadas, através da transformação da imagem, em primeiro lugar para uma escala de cinzas e depois para uma escala binária preto/branco (Figura 2.5). Essa segunda imagem permite que seja feita uma contagem de píxeis brancos e pretos e consoante a proporção de uns e outros é possível aferir o enchimento do contentor.



Figura 2.5 – Utilização de sensores de imagem para monitorização de enchimento de contentores: (a) imagem real; (b) imagem binária. Adaptado de Arebey, Hannan e Basri (2011).

Segundo os autores do estudo, o sistema mostrou-se eficaz na monitorização do enchimento de contentores, bem como na transmissão dessa informação em tempo real através de um sistema passivo de RFID e GPS.

A classificação dos resíduos através das imagens pode ser realizada pela aplicação da transformada de Hough – a técnica matemática que deteta formas geométricas em imagens digitais – que assim permite saber o que se encontra dentro do contentor (Hannan *et al.* 2014).

2.2.3. SENSOR VOLUMÉTRICO

O estudo realizado por Lelah *et al.* (2012) utilizou sensores de ultrassons que permitiam a medição do volume no interior de contentores de vidro através do efeito de Doppler – medição da diferença entre a onda enviada e a onda recebida. Essa tecnologia, associada a um módulo rádio que transmita via GSM para uma antena, permitia fazer chegar ao servidor central que calculava as rotas de recolha de recolha do contentor (Figura 2.6).

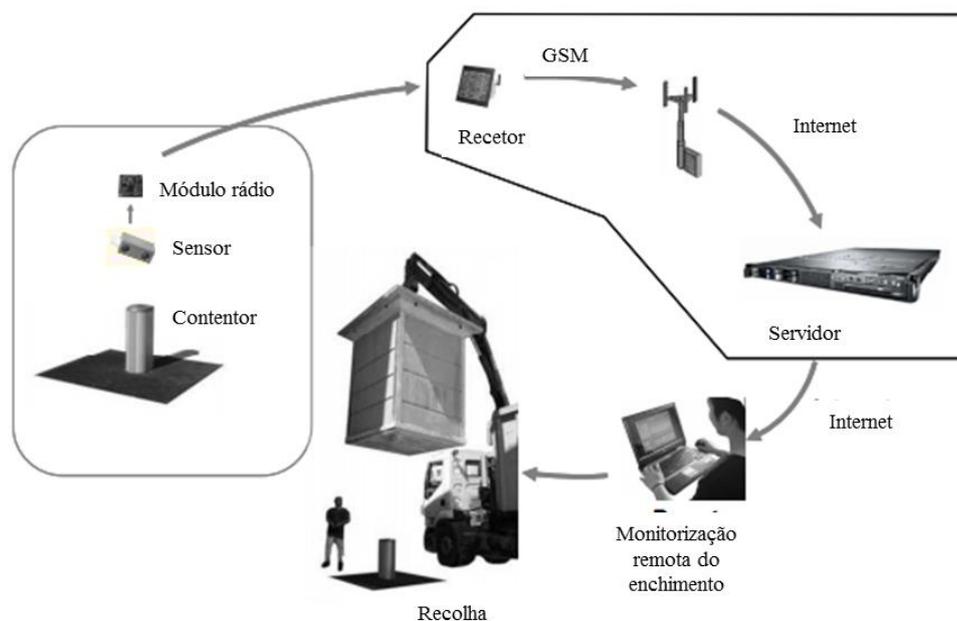


Figura 2.6 – Sistema de monitorização remota do enchimento de contentores com sensores de ultrassom. Adaptado de Lelah *et al.* (2012).

Esta tecnologia permite ter um conhecimento em tempo real do enchimento do contentor (Vincentini *et al.*, 2009). Contudo, os autores do estudo destacam como melhoramento a integração do sensor no fabrico do contentor para evitar operações de manutenção e substituição do sensor.

Outra forma de medir o volume é a utilização de sensores de capacidade, estudados por Reverter *et al.* (2003) com a instalação de um sistema de elérodos num contentor de papel. Os elérodos eram instalados de forma a medir a ultrapassagem de um limite máximo. Foram

estudadas duas soluções, uma com os elétrodos coplanares e outra com os elétrodos frente a frente (Figura 2.7).

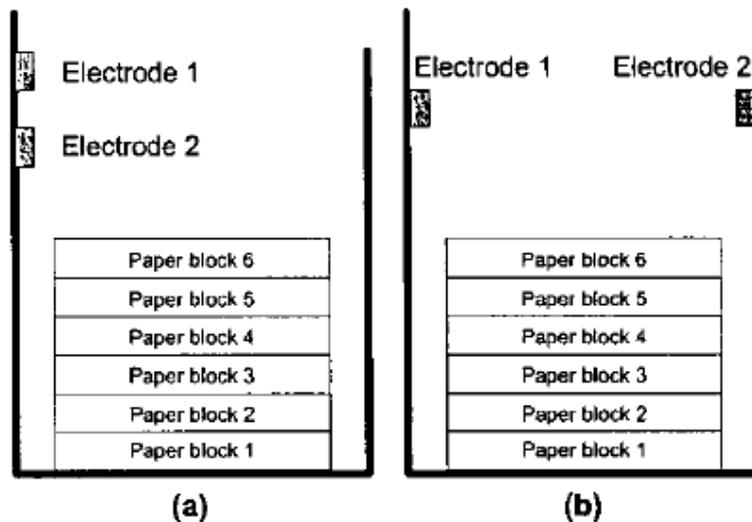


Figura 2.7 – Sensor de capacidade: (a) elétrodos coplanares; e (b) elétrodos frente a frente. Adaptado de Reverter; Gasulla e Pallas-Areny (2003).

Os resultados obtidos pelos autores foram satisfatórios, porém a qualidade dos mesmos pode ser afetada pela presença de elementos metálicos nos resíduos que entram em conflito com os elétrodos quando coplanares. As condições atmosféricas, em particular a humidade, afetam igualmente o funcionamento desta tecnologia.

2.2.4. SENSOR ÓTICO

Johansson (2006) utilizou sensores com infravermelhos para monitorizar o enchimento de contentores. Cada contentor tinha instalado quatro díodos de sensores que quando interrompidos (por resíduos depositados) indicam o enchimento do contentor (Figura 2.8).

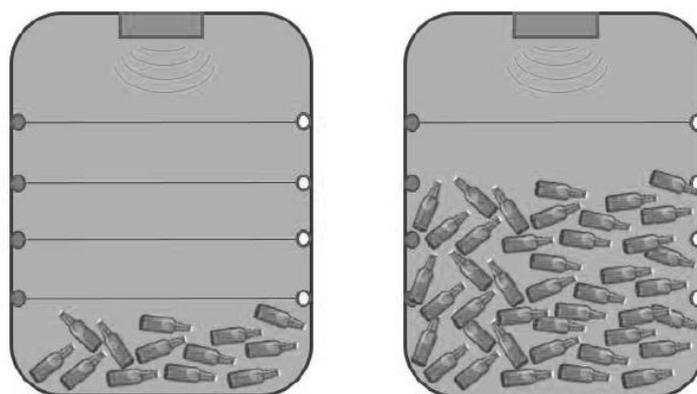


Figura 2.8 – Sensor com infravermelhos. Adaptado de Gonçalves (2014).

A cada hora era avaliado o enchimento do contentor; ao ser acionado o terceiro sensor, era enviada uma mensagem via GSM que alertava o operador. A interrupção do quarto, e último, sensor enviava nova mensagem de alerta. Como medida de controlo de qualidade foram utilizadas nesse estudo penalidades para os operadores que deixassem passar mais de 24h ou 48h entre alarmes em dias úteis ou fim de semana, respetivamente, antes de procederem à recolha.

Apesar das vantagens, a tecnologia de infravermelhos apresenta a desvantagem de não permitir fazer uma leitura contínua da informação uma vez que a avaliação é feita, neste caso, em períodos horários.

2.2.5. ECOPONTO INTELIGENTE

Foi inaugurado em Vila do Conde no dia 3 de fevereiro de 2016 um sistema de ecoponto inteligente que mede o enchimento dos contentores para além de caracterizar os resíduos produzidos pelo utilizador através de um sistema de identificação. (Lipor, 2016).

Segundo Paulo Rodrigues da Lipor, o utilizador introduz os resíduos num compartimento com 100 L (1 unidade) que quando fechado liberta os resíduos para dentro do contentor; após completar um determinado número de unidades – isto é, de utilizações – parte-se do princípio que o contentor está apto para recolha. Uma vez que cada tipo de resíduo tem densidades diferentes, a unidade de enchimento é retificada de acordo com o histórico (70% para embalagens, 40% para vidro e 60% para papel), sendo que esses valores são atualizados.

A informação recolhida é transmitida para um portal onde o utilizador, pela visualização dos resíduos que produziu, acumula pontos com vista a prémios. A iniciativa tem como objetivo

sensibilizar a população para a separação de resíduos, assim como monitorizar o tipo de resíduos produzidos por cada utilizador para a otimizar e personalizar o tipo de campanhas de sensibilização a realizar junto dos utilizadores.

2.3. MODELOS DE APOIO À DECISÃO

Tendo em vista um objetivo, as ações que podem ser tomadas para o alcançar dependem de uma relação entre as alternativas que se podem escolher e os resultados que elas produzem. A tomada de decisão consiste na escolha da alternativa. A **tomada de decisão** constitui um meio para alcançar objetivos e tomar uma decisão consiste em fazer uma escolha, que se traduz numa ação de alocação de recursos.

A tomada de decisões pode ser influenciada pela incerteza e complexidade do problema que se pretende resolver. Por um lado, a **complexidade** está associada às diferentes variáveis e critérios de um problema, bem como as relações entre eles, que crescem exponencialmente (Goodwin e Wright, 1991). De acordo com o mesmo autor, um método utilizado para diminuir a complexidade de um problema é “dividir e conquistar”, que consiste em partir o problema em vários problemas mais simples de resolver, cuja soma das soluções permite resolver o problema inicial. No capítulo 2.4 é abordada a análise de *clusters* que se enquadra nessa abordagem.

Por outro lado, a **incerteza** associada aos resultados de uma decisão deve ser tida em conta uma vez que nem sempre o desfecho de um problema é determinístico (conhecimento total das probabilidades e resultados), podendo existir uma probabilidade de ocorrência, e consequentemente o risco associado (Figura 2.9).

Os modelos existentes podem ser divididos em três tipos:

- Decisão com incerteza pura;
- Decisão com risco (ou modelos probabilísticos);
- Decisão com certeza absoluta (ou modelos determinísticos).

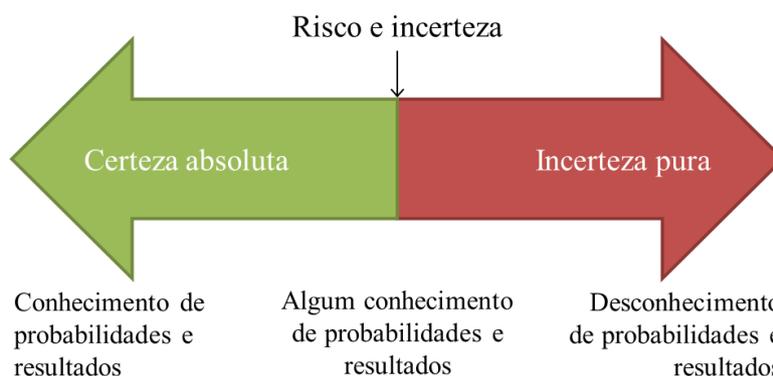


Figura 2.9 – Esquema da relação entre a incerteza e risco na tomada de decisões.

Nos capítulos seguintes serão abordados cada um desses tipos de modelos que podem ser resumidos com a Figura 2.10. Nos modelos com incerteza pura, como existe desconhecimento total dos resultados, a probabilidade é teoricamente igual para todas as alternativas. Nos modelos probabilísticos, a probabilidade é conhecida logo podem ser ponderadas as várias alternativas tendo em conta esses valores. Por fim, os modelos determinísticos não dependem das probabilidades, mas apenas da qualidade da alternativa (Howson e Urbach, 2005). Este último não será abordado, uma vez que se trata de uma situação limite sem escolha efetiva.

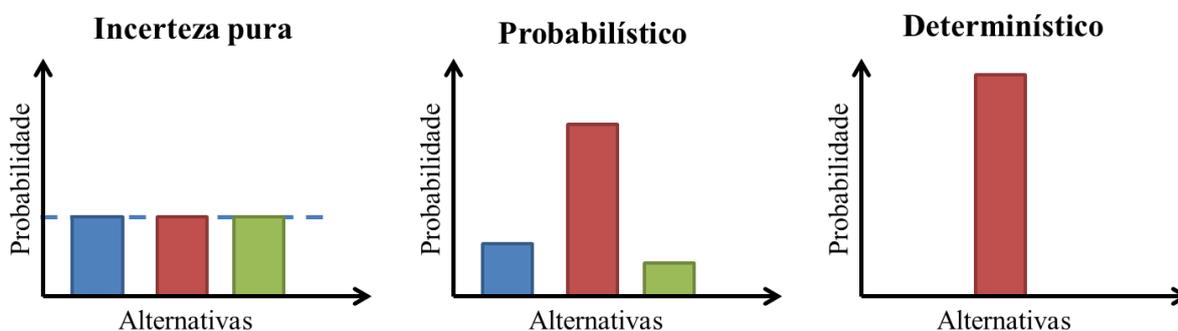


Figura 2.10 – Esquema explicativo das probabilidades em função das ocorrências para os modelos com incerteza pura, probabilísticos e determinísticos.

2.3.1. MODELOS COM INCERTEZA PURA

Nos modelos de **decisão com incerteza pura**, o decisor atua sem conhecer os possíveis resultados da sua escolha. Nestes casos, a decisão depende do tipo de personalidade do decisor (Bizwas, 1997; Ossadnik *et al*, 2013):

- Pessimista/conservativo (*Maximin*): escolhe o melhor pior resultado;
- Otimista/agressivo (*Maximax*): escolhe o “melhor melhor” resultado;

- Coeficiente de otimismo (Índice de Hurwicz): escolhe uma ponderação entre o melhor e o pior resultado.
- Minimizar o arrependimento (Oportunidade perdida de Savage): escolhe a alternativa que provoca o mínimo arrependimento (por não ter escolhido outra alternativa).

Este tipo de modelo é vantajoso quando existe um grande desconhecimento, pois permite ultrapassar a contingência de o decisor não ter informação sobre o problema. Contudo, o facto de as probabilidades não serem conhecidas faz com que exista um potencial risco elevado na decisão. Este tipo de modelo é usualmente mais aplicável a decisões de cariz pessoal e não profissional; neste último caso existe uma necessidade de ter um pouco de informação para suportar uma decisão (Goodwin e Wright, 1991).

2.3.2. MODELOS PROBABILÍSTICOS

Os **modelos probabilísticos** integram o risco na tomada de decisão. Neste tipo de modelos, por um lado não existe desconhecimento sobre as probabilidades de ocorrência de cada alternativa, mas por outro existe ainda alguma falta de informação que não permite uma escolha direta (Wang *et al*, 2012; Broekhuizen *et al*, 2015).

O **valor esperado** é um método probabilístico que permite ter em consideração o resultado e a probabilidade de ele ocorrer. A decisão é feita através da seleção da alternativa com o valor esperado mais elevado. (Botezatu, 2016) O valor esperado pode ser calculado com a seguinte expressão,

$$VE_i = \sum_j r_{ij}p_j \quad (2.1)$$

onde r_{ij} é o valor do resultado da alternativa i , tendo em conta o estado da natureza j e p_j é a probabilidade de acontecimento do estado da natureza j (Ragsdale, 2015).

A **função de utilidade** é um método que permite relacionar o resultado com a utilidade que este tem para o decisor. Ou seja, um mesmo resultado pode ser melhor ou pior dependendo do perfil de risco do decisor (Wellman e Doyle, 1992). A função utilidade, para um decisor avesso ao risco pode ser dada pela seguinte equação,

$$U(x) = 1 - e^{-x/R} \quad (2.2)$$

onde x é o resultado e R um parâmetro representativo da tolerância ao risco do decisor (Ragsdale, 2015).

Existem outros métodos que utilizam as probabilidades para simulações de base estatística, como é o exemplo da **simulação de Monte Carlo**. Essa é utilizada para determinar resultados através de parâmetros estocásticos totalmente aleatórios ou pseudoaleatórios caso sejam gerados por computador (Rubinstein e Kroese, 2011).

Assumindo como entrada no modelo uma dada distribuição de probabilidade, são levadas a cabo um grande número de simulações com o objetivo de determinar um histograma de frequências de resultados separados e independentes (Miguel, 2013).

Algumas das distribuições de probabilidade que podem servir de entrada para a simulação de Monte Carlo são enumeradas a seguir (Carmo, 2014):

- **Distribuição normal:** curva de probabilidades simétrica com o valor mais provável no centro (média). É descrita pelo valor da média e do desvio padrão e é dada pela seguinte fórmula,

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}, -\infty < x < +\infty, \sigma > 0 \quad (2.3)$$

onde x é a abcissa, μ é a média e σ o desvio padrão.

- **Distribuição log-normal:** curva cujo logaritmo é a distribuição normal. Caracteriza-se por apresentar apenas valores positivos e é dada pela seguinte fórmula,

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)}, 0 < x < +\infty, \sigma > 0 \quad (2.4)$$

- **Distribuição de Poisson:** é uma distribuição discreta que representa a probabilidade de um número de ocorrências suceder num dado intervalo de tempo/espaço. É dada pela seguinte fórmula,

$$f(k, \lambda) = \frac{e^{-\lambda} \lambda^k}{k!} \quad (2.5)$$

onde, k é o número de ocorrências e λ é o número de ocorrências por unidade de tempo/espço.

- **Distribuição beta:** caracteriza-se por estar compreendida entre 0 e 1, e é parametrizada por dois valores positivos α e β .
- **Distribuição uniforme:** todos os valores têm igual probabilidade de ocorrência.

As vantagens do método de Monte Carlo passam pela sua capacidade de integrar a incerteza nos resultados de uma dada decisão, permitindo visualizar os resultados em função de um intervalo de confiança (Kwak, 2007). Por outro lado, é um método que exige um processamento muito pesado de dados o que dificulta a sua aplicação (Loizou e French, 2012).

2.3.3. GESTÃO DO RISCO

Por risco entende-se como um acontecimento/evento incerto, estando associada uma probabilidade de ocorrência e um impacto esperado caso o evento ocorra. Os modelos de apoio à decisão estão intrinsecamente ligados à gestão de risco, uma vez que estes permitem integrar a probabilidade e o impacto (Sadgrove, 2015). A gestão do risco é composta por seis processos: a) plano de gestão do risco; b) identificação de riscos; c) análise qualificativa de risco; d) análise quantitativa de riscos; e) planeamento de respostas; e f) controlo dos riscos (PMI, 2013).

O primeiro dos seis processos é transversal a todos eles, pelo que o foco poderá ser dado aos restantes cinco processos (Figura 2.11).

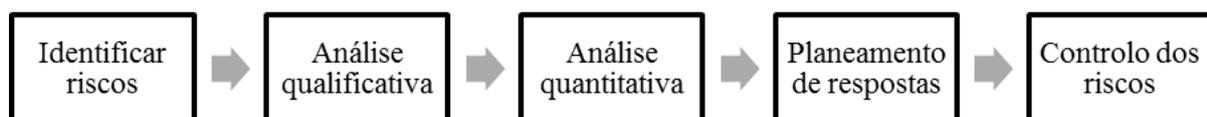


Figura 2.11 – Processos da gestão de risco segundo o PMBoK.

Na identificação de riscos pretende-se listar exhaustivamente os riscos que podem comprometer uma determinada situação ou decisão. Essa recolha de informação pode ser feita através de experiências passadas, dos conhecimentos de um especialista, entre outras fontes. Os riscos

identificados devem depois ser analisados qualitativamente por forma a averiguar a quais deve ser dado maior destaque por implicarem maiores impactos (Miguel, 2012). A classificação do risco pode ser feita através de uma relação entre a probabilidade de ocorrência e o impacto que tem; essa representação é normalmente feita sob a forma de uma matriz de impactos (Figura 2.12). Apesar de ser uma forma aceite pela generalidade da comunidade científica, as matrizes de impacto apresentam algumas desvantagens tal como o facto de não integrar a frequência de ocorrência (Cox, 2008).

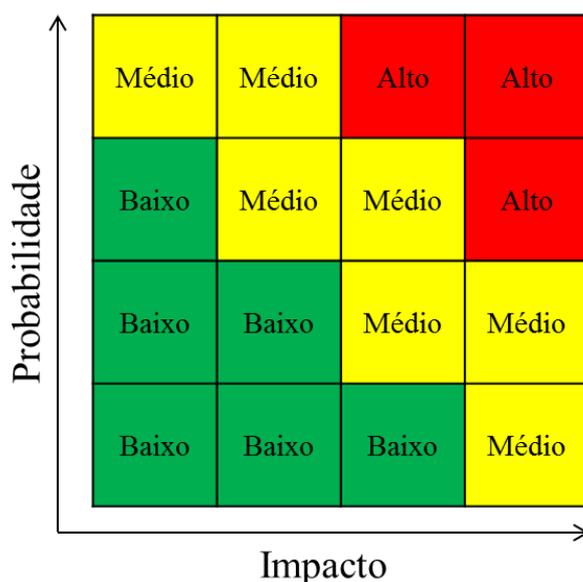


Figura 2.12 – Exemplo de matriz de impacto.

Existem várias estratégias para gerir o risco: a) eliminar; b) mitigar; c) transferir e d) aceitar. A primeira foca-se na eliminação da causa do evento evitando que este possa ocorrer. A mitigação passa por suavizar o impacto do evento. A transferência imputa a um terceiro a responsabilidade de lidar com o impacto do evento. A aceitação implica que se assuma a totalidade da responsabilidade do impacto. (PMI, 2009).

Após aplicar as estratégias mencionadas, se não houver alteração da classificação de risco “alto”, deverá ser realizada uma análise quantitativa. Essa permite que se tenha uma noção do impacto sob forma numérica, facilitando o processo de planeamento de respostas aos riscos.

A simulação de Monte Carlo permite avaliar qual o impacto do risco dependendo da probabilidade de ocorrência. Desta forma, pode ser feita uma reserva de contingência que tenha em conta os possíveis resultados de um risco assim como a probabilidade de ocorrência dos

mesmos, diminuindo o impacto da incerteza da solução. A reserva de contingência tem como principal objetivo fazer face aos *known-unknown*, ou seja, é uma folga (que pode ser de orçamento ou de tempo) para gerir os riscos (*unknown*) que foram identificados e estimados (*known*).

Por fim, durante a execução deverá ser feito o controlo dos riscos identificados por forma a avaliar se existem alterações quanto às suas classificações, alterações de pressupostos ou aparecimento de novos riscos.

2.4. ANÁLISE DE CLUSTERS

Na presença de uma grande quantidade de dados, existe a necessidade de os organizar, com o objetivo de resumir e facilitar a sua compreensão. A análise multivariada aplicada apresenta várias formas de organizar e reduzir a quantidade de informação, auxiliando na resolução de problemas. Esta permite que se estabeleça uma relação direta com as suas variáveis ou que se agregue os indivíduos em grupos com características comuns – análise de *clusters*. (Reis, 2001).

A **análise de *clusters*** pode ser dividida em cinco etapas (Jain e Dubes, 1988):

- Representação da amostra, podendo já focar as características mais relevantes dos indivíduos dessa amostra e/ou apresentar novas características resultantes de transformações de características iniciais dos indivíduos;
- Definição de medidas de proximidade que mais bem se adequem às características da amostra;
- Agrupamento de indivíduos em *clusters*;
- Classificação de cada *cluster*;
- Validação de resultados.

Os passos acima descritos estão representados na Figura 2.13.

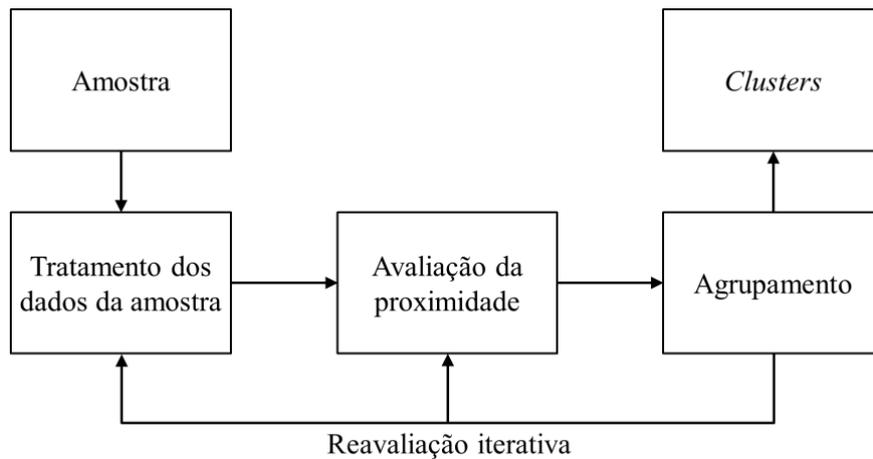


Figura 2.13 - Etapas da análise de clusters adaptado de (Jain et al., 1999).

Uma dada amostra, X , pode ser representada por matriz com n indivíduos e p características, onde x_{ij} representa o indivíduo i com a característica j .

$$X = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{bmatrix} \quad (2.6)$$

Para uniformizar o valor das variáveis, anulando os efeitos de escala e da unidade de medição, pode transforma-se em novas variáveis, com média nula e desvio-padrão unitário (i.e. com valores entre -1 e 1) através da seguinte expressão,

$$x' = \frac{(x - \mu)}{\sigma} \quad (2.7)$$

onde x' é o valor estandardizado de x , μ é a média e σ é o desvio-padrão.

2.4.1. MÉTODOS DE ANÁLISE DE *CLUSTERS*

Existem diversos métodos de análise de *clusters* na literatura, contudo serão focados apenas aqueles que foram considerados pelo autor como sendo adequados à problemática em discussão – os métodos hierárquicos e os métodos partitivos.

Os **métodos hierárquicos** podem dividir-se em dois subgêneros: **aglomerativos** ou **divisivos**. Os aglomerativos iniciam-se com n *clusters*, contendo apenas um indivíduo cada. Tal como o nome indica, esses *clusters* “unitários” podem ser aglomerados através da matriz de distâncias/semelhanças, criando *clusters* com vários indivíduos próximos/semelhantes, até chegar, no limite, a um *cluster* com todos os n indivíduos. Esta técnica tem a particularidade de o somatório dos *clusters* do nível superior corresponder ao *cluster* de nível imediatamente abaixo. Por outro lado, os divisivos partem de um *cluster* com n indivíduos, dividindo-o em vários, até no limite existirem n *clusters* com um indivíduo apenas. Este método, por outro lado, tem a particularidade de o somatório dos *clusters* do nível inferior corresponder ao *cluster* de nível imediatamente acima. Em termos gerais as suas características são:

- Partem de uma amostra ou de uma matriz de proximidade;
- Uma vez que um indivíduo seja integrado num *cluster* não mais o abandona;
- O número de *clusters* não é conhecido no início da análise;
- São úteis para objetos e variáveis.

Os **métodos partitivos**, ao contrário dos métodos hierárquicos, não se baseia na matriz de distâncias/semelhanças entre indivíduos, mas parte diretamente dos dados iniciais. É feita uma partição inicial dos dados num número predefinido de *clusters*. Para cada *cluster* é calculado o centroide, transferindo a cada iteração indivíduos de uns *clusters* para os outros por forma a encontrar a menor distância ao centroide do *cluster*. Quando a transferência de indivíduos estabilizar, a análise termina. Em suma, os métodos partitivos:

- Efetuam os cálculos sobre a matriz da amostra;
- Aplicam-se apenas a objetos (excluindo variáveis);
- Os grupos devem satisfazer os critérios de isolamento externo e coesão interna;
- O número de *clusters* é fixado no início da análise;
- Um indivíduo pode mudar de *cluster* durante a análise.

2.4.2. MEDIDAS DE PROXIMIDADE

A amostra com os valores estandardizados X' permite que se construa uma outra matriz – a **matriz de proximidade** – que representa uma medida de proximidade entre indivíduos. O

valor d_{kl} representa uma medida de proximidade entre o indivíduo k e o l , que variam entre 1 e n .

$$D = \begin{bmatrix} 0 & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \dots & 0 \end{bmatrix} \quad (2.8)$$

A medida de proximidade não tem de ser uma distância; pode também representar uma semelhança ou dissemelhança entre as características dos indivíduos analisados (Everitt *et al.*, 2011). Na Tabela 2.1 apresentam-se as propriedades das medidas de proximidade.

Tabela 2.1 – Propriedades das medidas de proximidade adaptado de Reis (2001).

Propriedades	Observações
$d_{ij} \geq 0$	Diferenciabilidade de idênticos
$d_{ii} = 0$	Indiferenciabilidade de idênticos
$d_{ij} = d_{ji}$	Simetria
$d_{ij} \leq d_{ik} + d_{ki}$	Desigualdade triangular

As medidas de proximidade podem ser classificadas em quatro categorias:

- Coeficientes de correlação;
- Medidas de distância;
- Coeficientes de associação;
- Medidas de semelhança probabilística.

Apesar de cada medida apresentar as suas vantagens e inconvenientes, as mais utilizadas são as duas primeiras, pelo que se fará apenas uma breve descrição desses dois tipos de medidas.

Os **coeficientes de correlação** quantificam a correlação entre dois indivíduos i e j variando o valor entre -1 e 1, sendo que 0 significa que não existe correlação. Para um conjunto p de variáveis, a correlação entre dois indivíduos pode ser dada pela seguinte expressão,

$$r_{ij} = \frac{\sum_{v=1}^p (X_{iv} - \bar{X}_i)(X_{jv} - \bar{X}_j)}{\sqrt{\sum_{v=1}^p (X_{iv} - \bar{X}_i)^2 \sum_{v=1}^p (X_{jv} - \bar{X}_j)^2}} \quad (2.9)$$

onde:

X_{iv} é o valor da variável v para o indivíduo i ($v=1, \dots, p$);

X_{jv} é o valor da variável v para o indivíduo j ($v=1, \dots, p$);

\bar{X}_i é a média de todas as variáveis para o indivíduo i ;

\bar{X}_j é a média de todas as variáveis para o indivíduo j .

Esta medida destaca-se pelo facto de não ser influenciada pelas unidades de medição ou ordens de grandeza das variáveis, pois é imposta uma standardização do valor. Contudo, uma vez que a correlação é estabelecida com base na associação linear entre variáveis, os agrupamentos podem sofrer enviesamentos. Para além disso, uma vez que é tida em conta a média de todas as variáveis de um indivíduo, pode acontecer que alguma das variáveis não tenha significado claro.

Por outro lado, para as **medidas de distância** existem várias técnicas que podem ser usadas, entre as quais a distância euclidiana, o quadrado da distância euclidiana, a distância absoluta, a distância de Minkowski, entre outras. Na técnica hierárquica aglomerativa, a matriz de proximidade/distância terá de ser recalculada a cada agrupamento uma vez que a distância passará a ser medida entre grupos diferentes a cada nível que se avança.

Este tipo de medida tem alguns inconvenientes ao nível das diferenças resultantes de diferentes ordens de grande ou de unidades de medição. Contudo, pode ser realizada uma standardização de acordo com a fórmula (2.7), apesar de não ser totalmente eficaz (Reis, 2001).

2.4.3. AGREGAÇÃO EM *CLUSTERS*

Após a definição de medidas de proximidade para a análise, é necessário proceder à agregação dos dados. O princípio utilizado é a maximização da proximidade dos indivíduos dentro do mesmo *cluster* e a minimização da proximidade aos indivíduos de outros *clusters*. É neces-

sário nesta fase que o utilizador defina o número k de *clusters* que pretende construir para que se possa proceder à agregação de indivíduos.

Existem vários critérios de agregação dos quais se destacam os seguintes: vizinho mais próximo, vizinho mais afastado, média dos grupos, centroides e critério de Ward. Estes critérios são válidos para o método hierárquico. Nos métodos partitivos, uma vez que não existe matriz de proximidade, o processo é diferente; o mais comum para este tipo de método é as k -médias.

Na Tabela 2.2 são apresentadas as principais vantagens e desvantagens de cada critério apresentado para o método hierárquico.

Tabela 2.2 – Vantagens e desvantagens dos critérios de agregação para o método hierárquico (adaptado de Reis (2001)).

Critério	Vantagens	Desvantagens
Critério do vizinho mais próximo	<ul style="list-style-type: none"> • Deteta grupos de forma muito variada • São necessários apenas dois objetos para determinar a distância entre <i>clusters</i>. • Deteta valores discrepantes. • Insensível a transformações monótonas da matriz de distâncias. • Não é afetado pela existência de relações nos dados iniciais. 	<ul style="list-style-type: none"> • Não tem capacidade para isolar grupos cuja separação não é nítida. • Não robusto (<i>i.e.</i> adição de dados pode alterar o resultado). • Tendência para se agrupar a um grupo já definido do que para formar o núcleo de um novo grupo. • Tendência para encontrar grupos esféricos.
Critério do vizinho mais afastado	<ul style="list-style-type: none"> • Tendência para encontrar <i>clusters</i> compactos. 	<ul style="list-style-type: none"> • Nem sempre apresentam um elevado grau de concordância com a estrutura inicial dos dados. • Tendência para encontrar grupos esféricos.
Critério da média dos grupos	<ul style="list-style-type: none"> • Evita valores extremos. • Toma em consideração toda a informação dos grupos. 	
Critério dos centroides		<ul style="list-style-type: none"> • Se dois grupos forem muito diferentes em termos de dimensão, o centroide do

		novo agrupamento estará mais próximo daquele que for maior e as características do grupo menor tenderão a perder-se.
Critério de Ward	<ul style="list-style-type: none">• Tendência para criar grupos com tamanho semelhante.	<ul style="list-style-type: none">• Tendência para criar grupos com tamanho semelhante.• Encontrar soluções que podem ser ordenadas a partir dos perfis relativamente às variáveis iniciais.

No que diz respeito aos métodos partitivos, destaca-se a técnica das k-médias. Esta técnica tem como características o facto de o representante do *cluster* ser um indivíduo do próprio grupo – o mesoide – ou seja o indivíduo mais central do *cluster*. Esse indivíduo serve de referência para alterar o agrupamento de cada indivíduo no cálculo iterativo característico dos métodos partitivos (Reis, 2001).

CAPÍTULO 3. METODOLOGIA

Um dos objetivos desta tese é desenvolver um modelo que permita apoiar a decisão de utilização de TIC na otimização de resíduos recicláveis. Este modelo deve permitir apoiar a decisão sobre qual a dimensão da amostra de ecopontos que devem ser monitorizados através de TIC, tendo em conta a relação entre o custo do investimento nas TIC e o erro associado à extrapolação dos resultados de uma amostra de ecopontos para o universo de ecopontos.

Na revisão de literatura realizada não foi identificado nenhum modelo que servisse o propósito do presente trabalho. Assim, o autor optou por desenvolver um modelo de apoio à decisão que combinasse várias análises e técnicas existentes.

Pretende-se assim identificar a dimensão ótima da amostra de ecopontos que devem ser monitorizados, visto que a monitorização do número total de ecopontos é economicamente inviável (Gonçalves, 2014). O caminho que o autor optou por percorrer para chegar a esse fim encontra-se esquematizado na Figura 3.1.

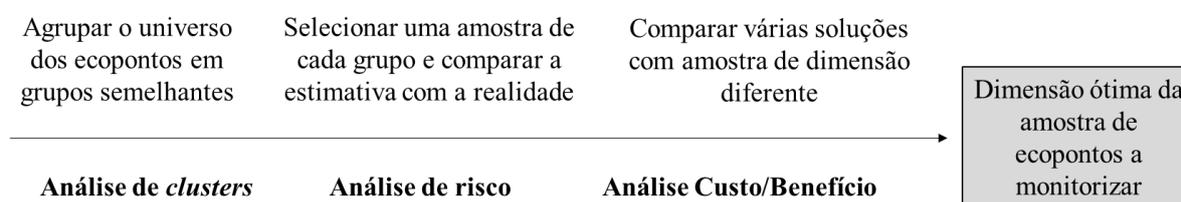


Figura 3.1 – Metodologia seguida

Por forma a abordar a alta complexidade do problema, onde existem milhares de ecopontos, cada um com as suas características, optou-se por agrupar os ecopontos em grupos semelhantes, para depois selecionar uma amostra de cada grupo e por fim extrapolar os valores de enchimento para todo o grupo calculando o erro associado à diferença entre a estimativa da amostra e a realidade (Figura 3.2). Para diferentes dimensões de amostra, o erro e investimento variam, logo a comparação de soluções permite identificar a solução ótima (Figura 3.3). Sobre essa solução é executada uma análise de risco, considerando alguns fatores associados ao enchimento de ecopontos, que permitem desenhar uma reserva de contingência para gerir a incerteza dos resultados obtidos na solução ótima.

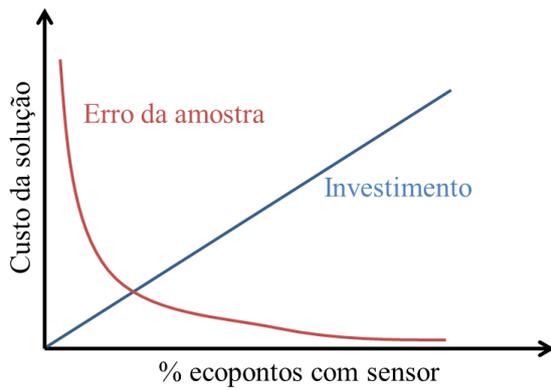


Figura 3.2 – Esquema representativo da curva do investimento e do erro em função da percentagem de ecopontos com sensor.

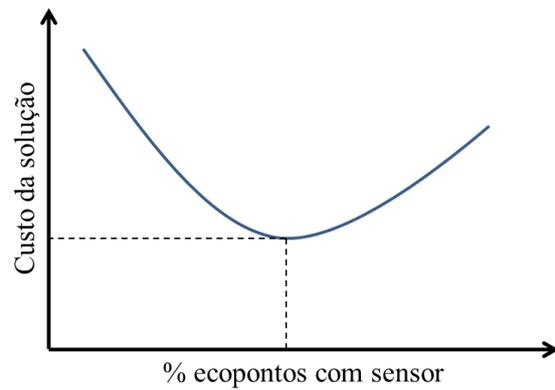


Figura 3.3 – Esquema representativo da curva do custo da solução (investimento + erro) em função da percentagem de ecopontos com sensor.

Por fim, uma análise de custo/benefício permite comparar a solução ótima determinada pelo modelo com a solução atual e com a solução em que todos os ecopontos são monitorizados.

CAPÍTULO 4. MODELO DE APOIO À DECISÃO PARA A UTILIZAÇÃO DE TIC NA OTIMIZAÇÃO DA RECOLHA DE RESÍDUOS

No presente capítulo pretende fazer-se a apresentação de um modelo de apoio à decisão para a problemática decorrente do trabalho desenvolvido por Gonçalves (2014), onde se verificou que a monitorização de ecopontos poderia ser benéfica para a otimização das rotas de recolha; contudo, a monitorização de todos os ecopontos mostrou-se inviável, pelo que o presente modelo pretende identificar o rácio de ecopontos a monitorizar que permite equilibrar o risco e o investimento.

Na Figura 4.1 está esquematizado o modelo que se desenvolveu para abordar o problema proposto.

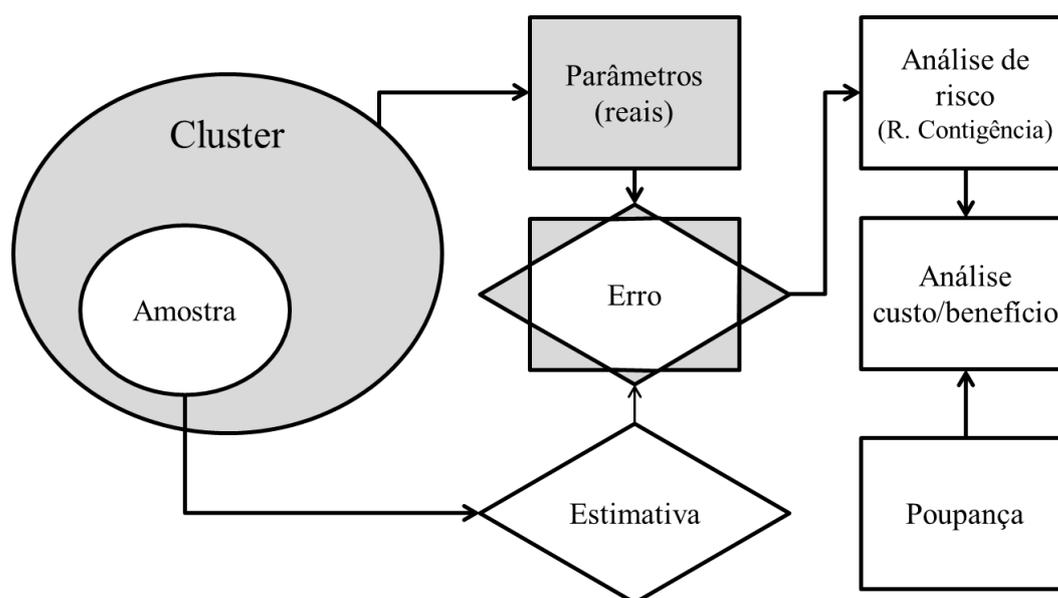


Figura 4.1 – Representação esquemática do modelo de apoio à decisão desenvolvido.

Uma vez que existe a necessidade de analisar uma grande quantidade de dados que são inter-relacionáveis, optou-se por agrupar os ecopontos recorrendo à técnica de **análise de clusters**.

Depois de criar e validar os *clusters*, é definida uma amostra de ecopontos por *cluster* onde são instalados os sensores através da **análise dos dados**. Os dados da amostra permitem reali-

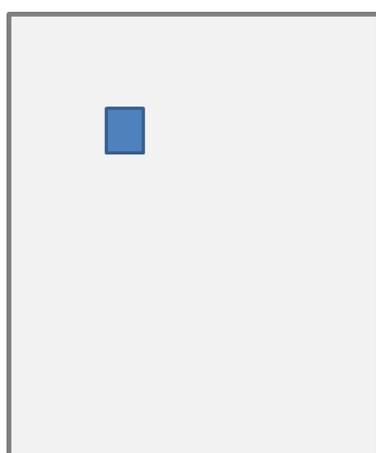
zar uma estimativa que caracteriza a totalidade do *cluster* que quando comparados com a realidade implicam a existência de um erro.

O erro tem um custo associado, pelo que é realizada uma **análise de risco**, uma vez que existe alguma incerteza associada aos ecopontos cujos dados de enchimento não são recolhidos pelos sensores. Esta análise permite definir uma reserva de contingência.

Por fim, realizar-se-á uma **análise de custo/benefício** avaliando as vantagens económicas da adoção de uma solução face aos benefícios expectáveis.

4.1. ANÁLISE DE CLUSTERS

No presente modelo optou-se por realizar uma análise de *clusters* por forma a dividir um problema inicial com um elevado grau de complexidade em vários pequenos “subuniversos” (os *clusters*). Definindo as medidas de proximidade entre ecopontos e agregando-os em *clusters*, é possível que uma amostra de dimensão semelhante a uma amostra retirada aleatoriamente do universo de ecopontos (Figura 4.2) seja mais representativa da diversidade e especificidade dos ecopontos (Figura 4.3).



Universo de ecopontos

Figura 4.2 – Esquema explicativo de uma amostra aleatória recolhida de um universo de ecopontos.



Universo de ecopontos

Figura 4.3 – Esquema explicativo de várias amostras aleatórias recolhidas de um universo de ecopontos dividido em clusters.

4.1.1. MEDIDAS DE PROXIMIDADE

Para permitir o agrupamento de ecopontos em *clusters*, foram identificadas as seguintes potenciais medidas de proximidade para cada ecoponto e para cada tipo de resíduos (embalagens, papel e vidro):

- **Geolocalização:** através das coordenadas M e P (latitude e longitude).
- **População por ecoponto:** os ecopontos foram agrupados por freguesia, cruzando essa informação com a população residente na freguesia. Apesar de ser uma medida com um grau de precisão relativo, pode contudo caracterizar a população por ecoponto (número de habitantes por ecoponto).
- **Enchimento diário médio (média e desvio-padrão):** o enchimento diário médio de ecopontos foi obtido através das medições de enchimento recolhidas (em volume), dividindo o enchimento pelo número de dias em que o ecoponto esteve sem ser recolhido. Foi calculada a média e o desvio-padrão para cada ecoponto deste enchimento diário médio. O coeficiente de variação de Pearson (obtido pela divisão do desvio-padrão pela média) permite determinar se existem grandes variações do enchimento diário médio.
- **Sazonalidade:** as variações no enchimento médio diário que todos os anos se verificam na mesma altura do ano (fator apontado por Berkun *et al.* (2011) como sendo relevante para a caracterização da produção de resíduos).

A geolocalização, apesar de se apresentar como uma medida de proximidade evidente para a agregação de *clusters* foi excluída, uma vez que poderia enviesar a análise ao agrupar ecopontos em função da sua localização em detrimento das restantes medidas de proximidade, nomeadamente o seu padrão de enchimento. Uma vez que o objetivo é o de agrupar os ecopontos com comportamento de enchimento semelhante, para mais facilmente obter uma amostra representativa, a geolocalização iria limitar os *clusters* à distância (em quilómetros) entre ecopontos, não sendo isso que se pretende.

Assim, as medidas de proximidade a considerar são:

- Número de habitantes servidos pelo ecoponto;
- Média de enchimento do ecoponto;

- Coeficiente de variação de enchimento do ecoponto;
- Sazonalidade (semanas/meses de maior utilização).

4.1.2. AGREGAÇÃO EM *CLUSTERS*

Após a definição de medidas de proximidade para a análise, é necessário proceder à agregação dos dados. Nesse sentido é realizada uma análise de *clusters* através do *software* SPSS versão 23 (IBM, 2015). Este *software* permite realizar a análise de *clusters* de três formas: a) método hierárquico; b) método das k-médias; e c) método de 2 etapas.

Para a análise presente neste modelo foram apenas considerados os dois primeiros, uma vez que o último é uma combinação dos dois primeiros, não permitindo identificar claramente as fronteiras entre cada método. Para o cálculo da medida de proximidade, o *software* permite escolher de entre várias formas de a calcular. Optou-se por considerar o quadrado da distância euclidiana.

Na Figura 4.4 está representada uma janela do SPSS com as informações necessárias para a análise.

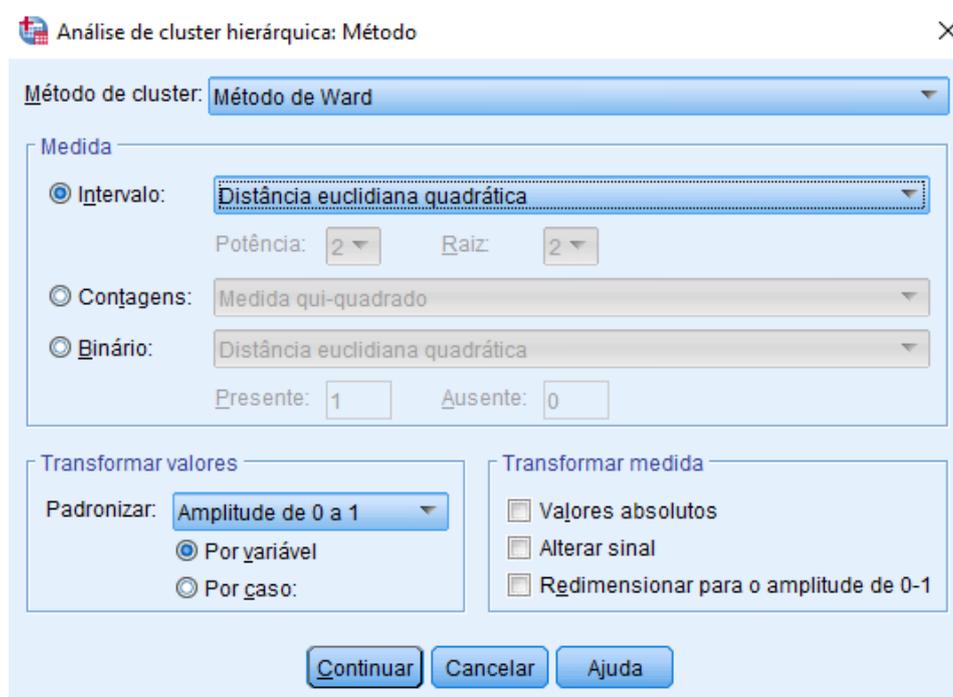


Figura 4.4 – Janela do software SPSS para a análise de clusters pelo método hierárquico.

4.2. ANÁLISE DE DADOS

A análise de dados parte da aquisição de dados a partir de uma dada população. A população é um conjunto de unidades estatísticas (neste caso ecopontos). O estudo da população pode ser realizado através de uma **amostra** dessa população. Os parâmetros de uma população podem ser aferidos de uma amostra através de estatísticas (Figura 4.5) que, apesar de algum erro implícito pode de certa forma ser útil para estimar parâmetros da população.

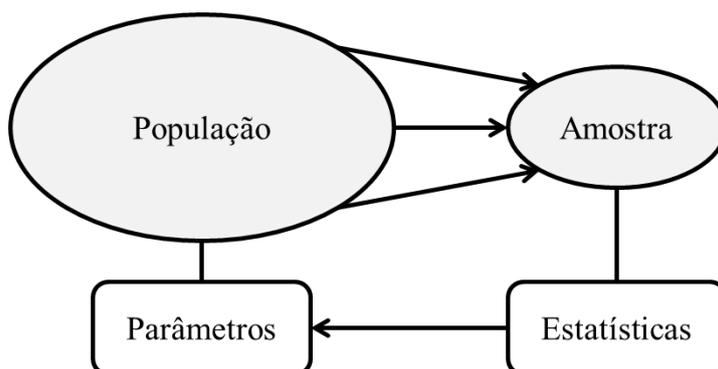


Figura 4.5 – Esquema da relação entre população e amostra.

Para que as estatísticas de uma amostra tenham uma estimativa útil é necessário garantir que a amostra é representativa da população.

4.2.1. TÉCNICAS DE AMOSTRAGEM

As técnicas de amostragem podem ser sumariamente divididas em aleatórias e não aleatórias. Na presente dissertação serão apenas apresentadas técnicas de amostragem aleatórias, por forma a evitar enviesamentos dos resultados (Murteira, 2007).

As **amostras aleatórias simples** conduzem a resultados onde as várias amostras com n elementos tenham a mesma probabilidade de serem selecionadas para uma população com N elementos. De notar que a probabilidade igual diz respeito à amostra e não somente aos indivíduos, isto porque existem outras técnicas onde cada indivíduo tem a mesma probabilidade de pertencer à amostra.

Por outro lado, as **amostras sistemáticas** podem ser recolhidas a partir de uma população de dimensão N , cujos dados estejam ordenados, escolhendo-se aleatoriamente um dos primeiros

k elementos, onde k é a parte inteira de N/n . São então selecionados os k -ésimos elementos da população a partir daí por forma a completar a amostra com n elementos.

As **amostras estratificadas** têm por base a divisão da população em estratos. Esses estratos, que podem ser mais ou menos homogêneos, consoante o critério escolhido pelo investigador, permitem que sejam selecionados aleatoriamente indivíduos de forma a que cada estrato seja proporcionalmente representado na amostra.

Cada técnica apresenta vantagens e desvantagens, que dependem dos parâmetros de cada população, pelo que não existe uma técnica melhor que outra.

No presente modelo é utilizada uma amostra aleatória sistemática, pois dessa forma garante-se que cada amostra retirada do *cluster*, tenha a amostra uma dimensão de 5 ecopontos ou 500 ecopontos, representa as singularidades do enchimento do grupo. Os ecopontos são ordenados por ordem crescente de taxa de enchimento, permitindo desta forma selecionar aqueles que farão parte da amostra (Figura 4.6).

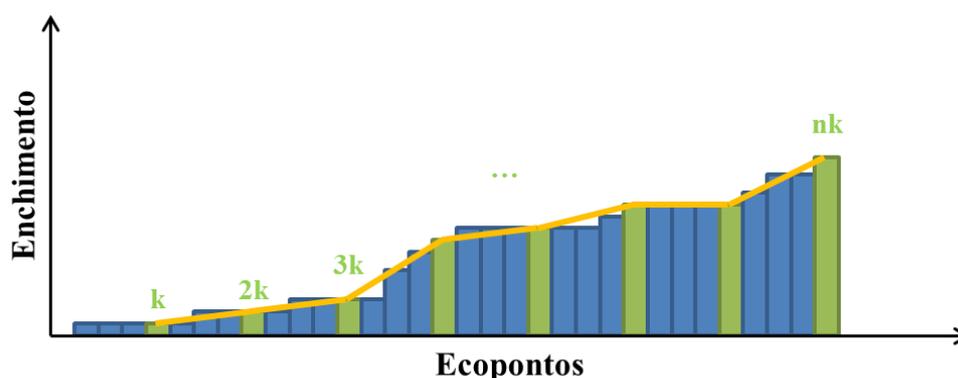


Figura 4.6 – Representação esquemática da seleção de amostras no modelo proposto.

4.2.2. ESTIMAÇÃO DE VALORES

Os valores recolhidos pela amostra são extrapolados para o *cluster* e comparados com os valores reais. Na Figura 4.7 estão representados as três categorias de resultados possíveis para o enchimento dos ecopontos: a) não aptos (enchimento abaixo de δ); b) aptos (enchimento entre δ e 100%); e c) sobrelotados (enchimento superior a 100%). O valor δ deverá ser definido pela empresa, uma vez que é o parâmetro que define a apetência do ecoponto para ser recolhido.

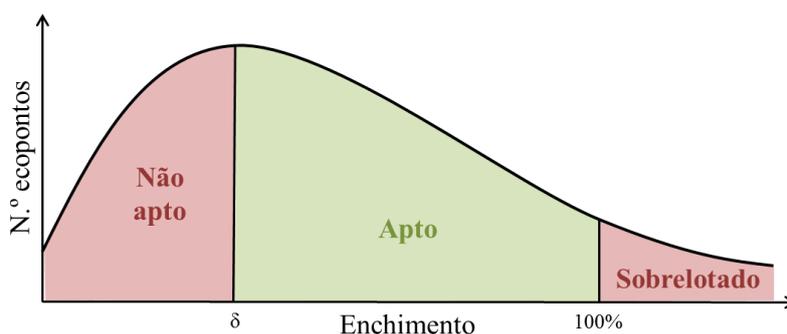


Figura 4.7 – Categorias de resultados do enchimento de ecopontos.

O objetivo desta estimativa é minimizar a diferença de ecopontos não aptos e sobrelotados no ato da recolha, uma vez que são esses casos que tornam a operação ineficiente. Assim, compara-se o valor da estimativa com o dos parâmetros (real) do *cluster*. Dependendo da dimensão da amostra, esta ajusta a estimativa melhor ou pior à realidade (Figura 3.2), ou seja, quanto maior a dimensão da amostra, mais próxima será a estimativa da realidade (Figura 4.9).

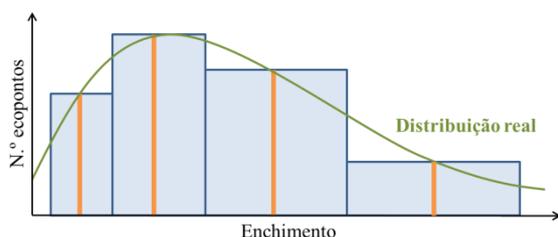


Figura 4.8 – Esquema do ajuste da estimativa de uma amostra pequena à realidade.

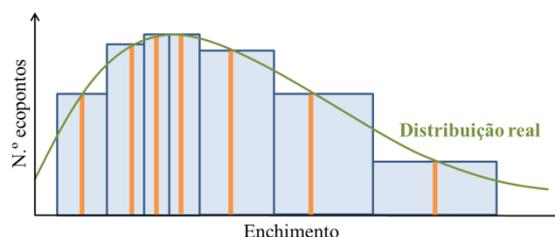


Figura 4.9 – Esquema do ajuste da estimativa de uma amostra maior à realidade.

A diferença/erro entre a estimativa da amostra e os valores reais, seja esse erro negativo ou positivo, tem um custo; no caso de recolha de um ecoponto não apto é o custo do desvio necessário fazer (em quilómetros) para recolher um ecoponto que o não deveria ser (Figura 4.10); no caso de recolha de um ecoponto sobrelotado é o custo de encher o camião de recolha com mais resíduos do que o planeado, podendo isso significar que outros ecopontos (aptos) não serão recolhidos pelo facto do camião já estar na sua máxima capacidade (Figura 4.11). A não recolha de um ecoponto apto numa determinada rota implica que noutra rota tenha de ser feito um desvio para o recolher.

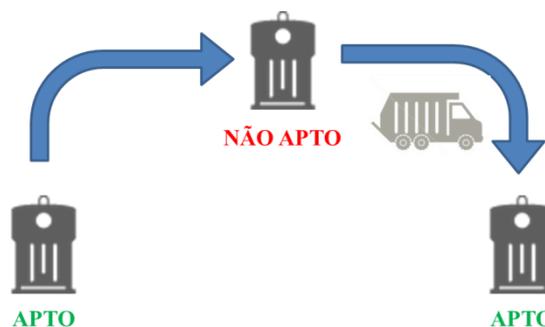


Figura 4.10 – Esquema explicativo do risco de recolher um ecoponto não apto.

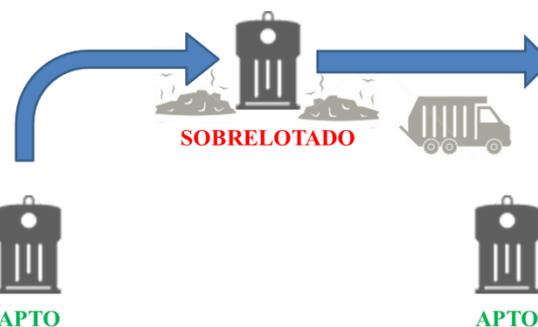


Figura 4.11 – Esquema explicativo do risco de recolher um ecoponto sobrelotado.

Esse erro pode ser contabilizado pela distância percorrida (ida e volta) do desvio necessário para recolher o ecoponto não apto/sobrelotado. Para essa distância pode ser considerada a distância média entre ecopontos. Tal como verificado na Figura 4.8 e Figura 4.9, o risco da estimativa estar mais bem ajustada à realidade é crescente com o número de ecopontos monitorizados; logo, quanto mais ecopontos monitorizados, menor o risco (que no limite é nulo, caso todos os ecopontos sejam monitorizados).

Por outro lado, o investimento em sensores é crescente (Figura 3.2). A soma das duas curvas permite a determinação de um valor ótimo, que equilibra o investimento necessário e o risco associado (Figura 3.3).

Uma vez que o valor da distância média entre ecopontos pode ter uma relevância significativa nas conclusões a tirar da análise, é necessário realizar uma análise de sensibilidade para avaliar até que ponto o resultado é dependente da distância considerada.

4.3. ANÁLISE DE RISCO

Tal como referido no capítulo 2.3.3, a gestão de risco passa por: a) identificar os riscos; b) classificar os riscos (qualitativa e quantitativamente); e c) controlar e monitorizar os riscos.

No que toca à identificação dos riscos, é feita uma listagem não exaustiva de possíveis riscos associados a implementação de sensores para a medição do enchimento dos ecopontos:

- Variação do enchimento;
- Alteração dos índices de adesão da população à reciclagem;
- Avaria/falha na comunicação do sensor com a central;

- Eventos extraordinários (*e.g.* festividades locais).

No presente modelo são considerados os dois primeiros fatores de risco identificados, uma vez que são considerados riscos que necessitam de uma resposta mais efetiva por afetarem diretamente o enchimento dos ecopontos, variável central para a análise dos dados. A quantificação do risco pretende definir a reserva de contingência que o decisor terá de ter disponível. A consideração de uma reserva de contingência faz parte da estratégia de aceitação ativa do risco, onde, uma vez que não foi possível eliminar o risco, se aceita que esse possa ocorrer, havendo para essa eventualidade uma reserva para fazer face ao risco.

4.3.1. SIMULAÇÃO DE MONTE CARLO

Por forma a realizar a análise de risco foi levada a cabo uma simulação de Monte Carlo para o enchimento dos ecopontos. Essa simulação consiste em recolher os dados de uma série de n observações, avaliando a frequência com as observações se repetem, representando-se em forma de histograma. Para simplificar a análise, foi assumida como base para o histograma a distribuição normal por permitir uma adaptação adequada à distribuição real.

Desta forma, é possível perceber de que forma poderá a solução ótima, calculada a partir do enchimento dos ecopontos, ser afetada pela imprevisibilidade. Assim, é integrado no custo do erro, identificado para a solução ótima (diferença entre a estimativa e o real), o risco desse erro poder ser diferente do previsto. Esse valor pode ser superior ou inferior, uma vez que esse erro foi calculado tendo por base um valor médio. Somando o valor do investimento nos sensores ao valor do erro contemplado o risco é possível determinar o custo da solução (Figura 4.12).

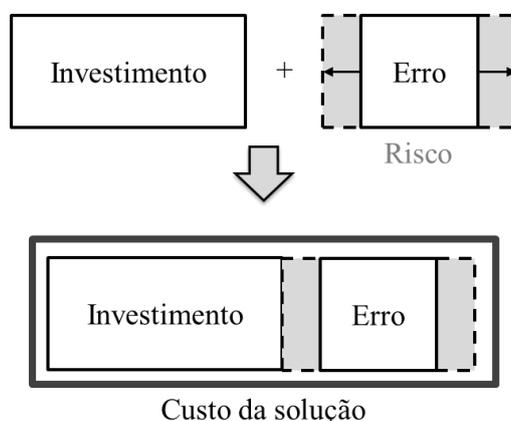


Figura 4.12 – Esquema da integração do risco no custo da solução.

O valor do erro sem risco é suficiente para fazer face a 50% das ocorrências de enchimento (i.e. variações do enchimento que são iguais ou inferiores ao erro calculado). Porém, existe uma probabilidade de 50% do enchimento poder ser superior ao erro, levando a que a reserva de contingência definida não seja suficiente. Assim, devem ser identificados os valores do custo da solução para probabilidades superiores a pelo menos 80%, valor mínimo aceitável, até 99% (Figura 4.13).

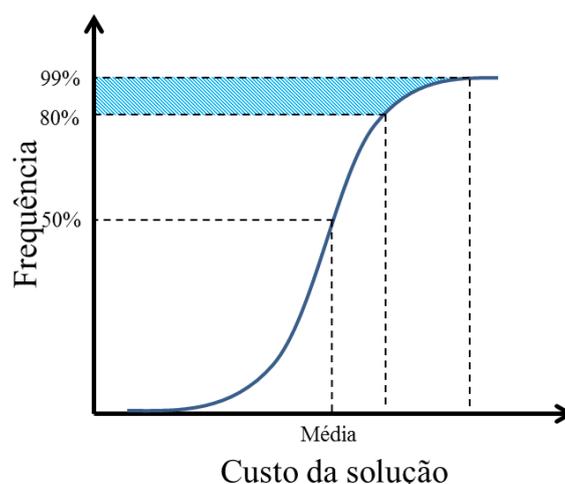


Figura 4.13 – Esquema da frequência cumulativa em função do custo da solução.

Associada à probabilidade está a confiança na solução, pelo que a confiança é equivalente à probabilidade. Quanto maior é a probabilidade da ocorrência estar dentro do limite definido, maior é a confiança, e conseqüentemente, menor o risco. A construção da curva é realizada através de uma função normal cumulativa, tendo por base a média (valor do custo do erro) e de um desvio padrão. Realça-se o facto de o investimento não ser considerado nesta análise de risco, uma vez que esse é constante.

Logo, é necessário definir a média e o desvio padrão da distribuição. A média considerada é a média do enchimento médio diário de cada *cluster*.

O desvio-padrão tem em consideração o coeficiente de variação do enchimento médio diário do *cluster* ($S_{Coef} = Coef \times média$) mas também o potencial da população aderir à reciclagem, calculado através da seguinte equação de desvio-padrão,

$$S_{Adesão} = \sqrt{\frac{1}{N-1} \sum (X - \bar{X})^2} \quad (4.1)$$

onde X é a adesão da população (enchimento/hab.dia.ecoponto), \bar{X} a média da adesão na população total e N o número de ecopontos no *cluster*. Os dois valores, do coeficiente de variação e de adesão da população, foram combinados através da equação de soma de desvios padrão,

$$S = \sqrt{S_{Coef}^2 + S_{Adesão}^2} \quad (4.2)$$

CAPÍTULO 5. APLICAÇÃO DO MODELO A UM CASO DE ESTUDO REAL

5.1. ENQUADRAMENTO

O caso de estudo foca-se na gestão de resíduos recicláveis da Valorsul – Valorização e Tratamento de Resíduos Sólidos das Regiões de Lisboa e do Oeste, S.A. – responsável pelo tratamento e valorização de resíduos urbanos em 19 municípios da Grande Lisboa e da Região Oeste. Apesar de a área de intervenção desta empresa representar apenas cerca de 4% da superfície nacional, são valorizados mais de 20% dos resíduos produzidos em Portugal.

A **missão** da Valorsul é *implementar e gerir um sistema integrado, tecnicamente avançado, ambientalmente correto e economicamente sustentável, para tratamento e valorização dos resíduos urbanos de 19 municípios das regiões de Lisboa e Oeste de Portugal* (Valorsul, 2016).

A **visão** passa por *tratar e valorizar resíduos (...) só possível através de um moderno Sistema de Gestão Integrada de resíduos urbanos, adequado à sua evolução quantitativa e qualitativa. A Valorsul é responsável pela conceção e concretização deste Sistema de Gestão Integrada que representa uma otimização das opções ambientais, sociais, económicas, técnicas e institucionais, ao mesmo tempo que prevê uma aposta na prevenção e aplicação de diferentes opções de tratamento e valorização de acordo com os diferentes tipos de resíduos recolhidos* (Valorsul, 2016).

Os **valores** da empresa são:

- *Ambição – ser uma empresa de excelência, sustentável e socialmente responsável;*
- *Eficiência – potenciar valor aos nossos acionistas, clientes, colaboradores e população;*
- *Transparência – evidenciar a transparência na nossa relação com as comunidades envolventes;*
- *Ética – cumprir os mais elevados padrões de ética em todas as nossas atividades.*
- *Competência – garantir elevados padrões de desempenho profissional em todos os setores da empresa, motivados pela excelência de serviço;*

- *Inovação – garantir e inovar em todos os setores na procura das melhores soluções.* (Valorsul, 2016).

A presente dissertação utiliza dados fornecidos pela Valorsul para 14 concelhos (ficam de fora Lisboa, Odivelas, Amadora, Loures e Vila Franca de Xira pois a recolha seletiva nestes concelhos fica a cargo dos próprios municípios) com um total de 6702 ecopontos (Figura 5.1). Os agrupamentos de ecopontos normalmente contam com três contentores (um para cada tipo de resíduo: vidro, papel e embalagens) (Figura 5.2), existindo contudo alguns agrupamentos com apenas um ecoponto ou com vários do mesmo tipo.



Figura 5.1 – Área de intervenção da Valorsul com zona estudada a vermelho.



Figura 5.2 – Configuração típica de ecopontos.

As rotas de recolha são efetuadas para apenas um tipo de resíduos de cada vez. A periodicidade das mesmas é calculada através do histórico de recolhas, tentando dessa forma prever picos de utilização (por exemplo na época balnear existe um reforço de recolhas em zonas costeiras e de veraneio).

A aplicação do modelo desenvolvido ao caso da Valorsul pretende contribuir para a otimização da definição de rotas para melhorar a **eficiência** e **ambição** da empresa, apostando na **inovação** com soluções tecnológicas com grande impacto no **sistema de gestão integrada**. É desta forma que os objetivos do presente trabalho pretendem alinhar-se com a missão, visão e valores da Valorsul.

5.1.1. SITUAÇÃO ATUAL

Os dados fornecidos pela Valorsul, para as rotas realizadas entre janeiro e setembro de 2013, são apresentados na Tabela 5.1.

Tabela 5.1 – Dados de rotas de recolha da Valorsul entre janeiro e setembro de 2013.

Tipo de resíduo	Distância percorrida [km]	Quantidade transportada [kg]	Tempo gasto [h]	Velocidade média [km/h]
Embalagens	129.909	1.922.225	6.683	19,44
Papel/Cartão	51.723	3.353.580	5.928	8,72
Vidro	118.051	2.425.040	2.354	50,15

Tendo em conta que a Valorsul atribui um custo total de operações de 1€/km (incluindo combustível, mão de obra e desgaste do material), o custo anual estimado para cada tipo de resíduo é de 173.212€, 68.964€ e 157.401€, para embalagens, papel/cartão e vidro, respetivamente.

Na Tabela 5.2 é possível verificar que existe uma grande percentagem de ecopontos recolhidos que não estavam aptos para a recolha, isto é, com uma taxa de enchimento abaixo de 0,75 (este valor foi definido pela Valorsul).

Tabela 5.2 – Dados de enchimento médio e percentagem de ecopontos aptos para recolha.

Tipo de resíduo	Enchimento médio	Ecopontos aptos para recolha
Embalagens	0,597	51%
Papel/Cartão	0,496	39%
Vidro	0,220	9%

Verifica-se desta forma que existe de facto uma oportunidade para otimizar as rotas de recolha. De notar que os valores de enchimento apresentados dizem respeito a dados de janeiro a setembro de 2013. Logo, a representatividade da amostra é relativa, uma vez que não se possuem mais dados que permitam uma análise mais completa.

5.1.2. DADOS INICIAIS

Foram disponibilizados pela Valorsul dados relativos à recolha de resíduos recicláveis (papel, embalagens e vidro) durante os três primeiros trimestres de 2013, assim como a localização dos contentores, as rotas realizadas e a frota automóvel disponível.

Os dados de enchimento foram recolhidos por observação e registo do motorista no momento da recolha durante as rotas observadas. O enchimento é quantificado com um número entre 0 e 1 (de 0,25 em 0,25), em que 0 significa que o ecoponto se encontrava vazio e 1 que o ecoponto se encontrava cheio. Esses dados constituem a base da análise a realizar.

A partir desses dados foi estimado enchimento médio diário para cada ecoponto (Figura 5.3). Desta forma foram determinados os enchimentos médios diários para cada ecoponto entre cada recolha, possibilitado assim determinar o valor médio e o desvio padrão do enchimento médio diário, assim como o coeficiente de variação (quociente entre desvio-padrão e média) (Figura 5.4).

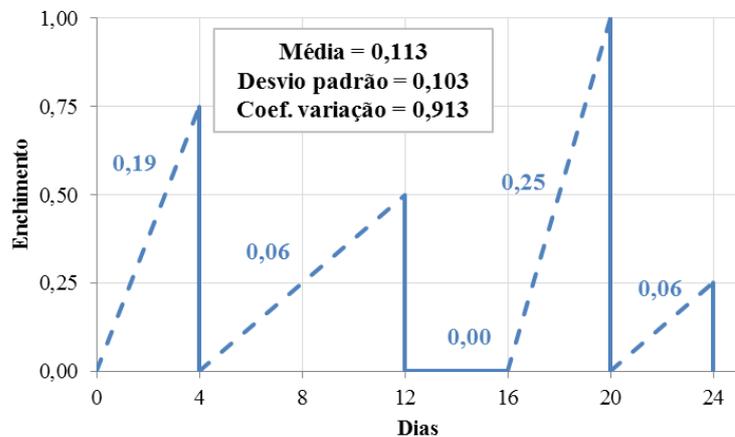
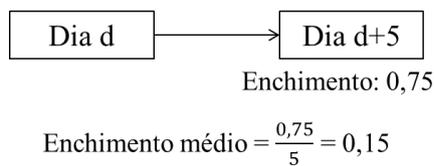


Figura 5.3 – Exemplo do cálculo do enchimento médio diário de um ecoponto.

Figura 5.4 – Exemplo do cálculo da média, desvio-padrão e coeficiente de variação do enchimento médio diário..

Estes dados permitem caracterizar o enchimento de cada ecoponto; essa é informação necessária para introduzir no modelo.

5.1.3. MONITORIZAÇÃO DE ECOPONTOS

No estudo realizado por Gonçalves (2014), uma das tecnologias apresentadas para a monitorização do enchimento dos ecopontos é o *Enevo ONE* (Figura 5.5), um sensor de ultrassons, sem fios, com autonomia de 10 anos.



Figura 5.5 – Sensor de ultrassons Enevo ONE (Enevo, 2016).

A escolha de tecnologia recaiu sobre os ultrassons pelo facto de revelar vantagens ao nível da instalação, funcionamento e manutenção.

Os resultados obtidos no estudo acima mencionado são apresentados na Tabela 5.3.

Tabela 5.3 – Resultados da otimização de rotas realizadas por Gonçalves (2014).

Tipo de Resíduo	Número de Rotas Realizadas	Ecopontos Recolhidos	Carga Recolhida	Distância Percorrida
Embalagens	-10%	-64%	+36%	-27%
Papel/Cartão	-14%	-60%	+65%	-28%
Vidro	-25%	-84%	-32%	-48%

O sistema tinha um custo mensal associado a cada sensor de 17€ (contrato de 3 anos) ou 15€ (contrato de 5 anos) à data do estudo. Atualmente, o custo foi revisto para 10,25€ (contrato de 3 anos) ou 8,33€ (contrato de 5 anos).

A poupança anual adquirida pela monitorização de ecopontos com esse sistema é superior a 140 mil euros, implicando, contudo, um custo anual de 1,35 milhões de euros (à data do estu-

do) ou 790 mil euros com os valores revistos. Tal como identificado anteriormente, e aqui novamente exemplificado, o investimento mostra-se inviável para a Valorsul.

Reforça-se desta forma a tese de que a monitorização de ecopontos deverá passar por uma solução em que apenas alguns ecopontos são monitorizados em detrimento de outros.

5.2. ANÁLISE DE *CLUSTERS*

Para efeitos de aplicação do modelo proposto ao caso de estudo foi apenas considerado o resíduo embalagens. Considerando as medidas de proximidade definidas em 4.1.1 foram analisados os dados para a criação de *clusters*:

- Número de habitantes servidos pelo ecoponto;
- Média de enchimento do ecoponto;
- Coeficiente de variação de enchimento do ecoponto;
- Sazonalidade (semanas/meses de maior utilização).

Note-se, contudo, que a não existência de dados em quantidade suficiente para tirar conclusões relevantes relativamente à sazonalidade implicam que esta medida tenha de ser excluída para este caso de estudo concreto. Apenas estão disponíveis para análise dados relativos a 9 meses, não sendo possível tirar conclusões sobre a sazonalidade. Desta forma, a estimativa para o número de *clusters* pode ser dada pelo método gráfico, pois existe apenas três medidas de proximidade que permitem a representação dos dados em forma de gráfico.

5.2.1. MÉTODO GRÁFICO

Por forma a identificar o número de *clusters* numa primeira iteração, na Figura 5.6 pode observar-se a representação gráfica dos dados em função das medidas de proximidade selecionadas no modelo. Foram identificados alguns ecopontos com valores extremos no que diz respeito às medidas de proximidade, tendo esses dados sido retirados da análise a fim de permitir uma agregação mais homogênea de indivíduos em *clusters*. Os ecopontos retirados da análise ficam automaticamente selecionados para serem equipados com sensor. Foram identificados, de forma visual oito possíveis *clusters* (Figura 5.7).

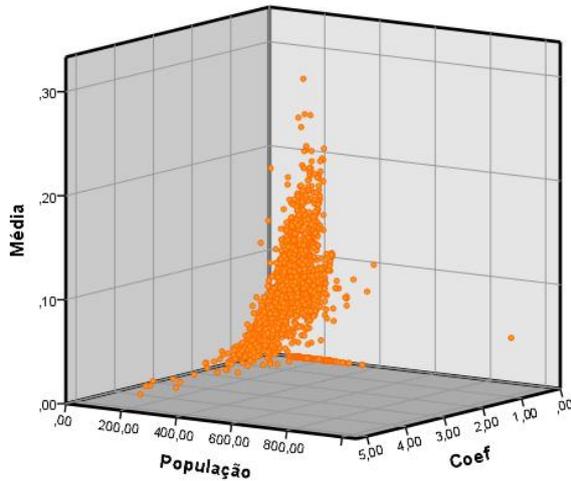


Figura 5.6 – Representação gráfica dos valores em função das medidas selecionadas.

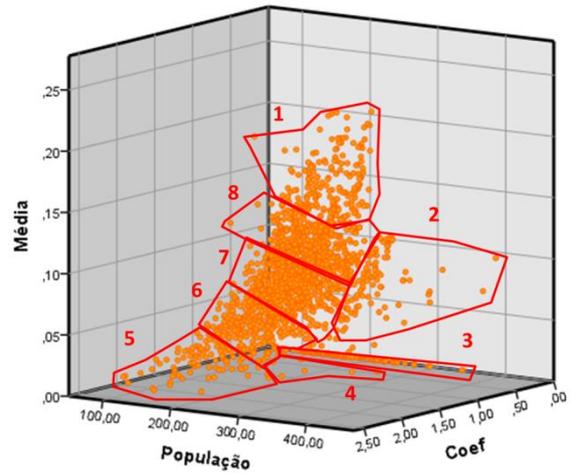


Figura 5.7 – Representação gráfica dos valores com uma simulação do número de clusters.

5.2.2. MÉTODO HIERÁRQUICO

Com o método hierárquico foram utilizadas alguns critérios diferentes por forma a verificar possíveis variações: a) vizinho mais próximo (Figura 5.8); b) vizinho mais longe (Figura 5.9); c) médias (Figura 5.10); d) centroides (Figura 5.11); e e) Ward (Figura 5.12). Com base na agregação gráfica dos dados foram definidos oito *clusters* para uma primeira iteração.

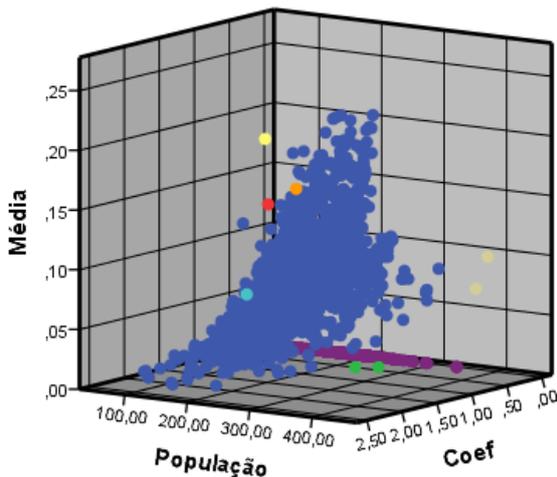


Figura 5.8 – Representação gráfica de clusters com o critério do vizinho mais próximo.

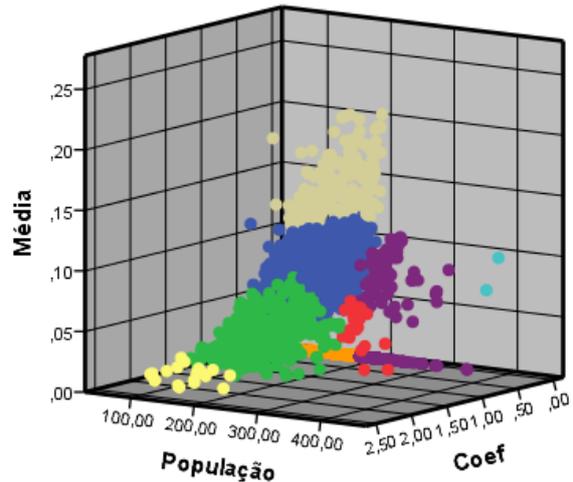


Figura 5.9 – Representação gráfica de clusters com o critério do vizinho mais afastado.

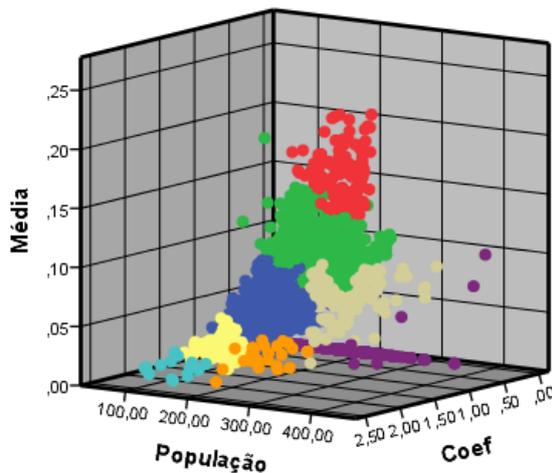


Figura 5.10 – Representação gráfica de clusters com o critério das médias.

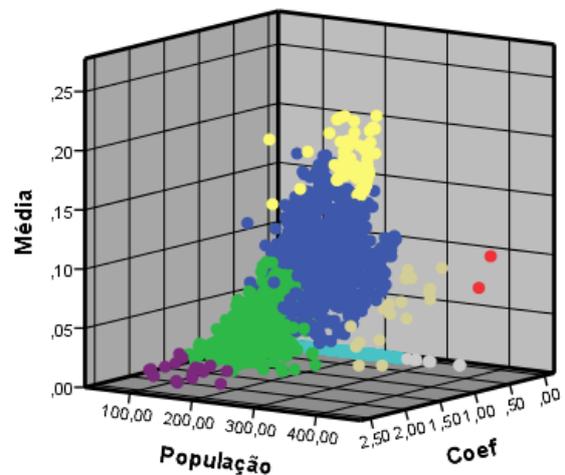


Figura 5.11 – Representação gráfica de clusters com o critério dos centroides.

Da observação dos gráficos é possível verificar que existe uma clara diferença entre os critérios. Enquanto que o critério do **vizinho mais próximo** agrupa a maioria dos ecopontos num só *cluster*, fazendo com que os restantes sete *clusters* tenham um reduzido número de indivíduos, verifica-se que os restantes critérios geram *clusters* aparentemente mais equilibrados em dimensão e posicionamento espacial. O critério do **vizinho mais afastado** é aquele que aparenta possuir *clusters* mais compactos, tal como referido na literatura, estando os elementos mais dispersos agrupados em *clusters* diferentes. O critério das **médias** formou *clusters* que aparentam ter em conta toda a informação dos dados. Os *clusters* formados com o critério dos centroides gerou dois *clusters* com tamanho superior tendo esses criado uma maior tendência para neles agrupar mais elementos. O critério de **Ward**, como espectável, produziu *clusters* com tamanho semelhante.

5.2.3. MÉTODO DAS K-MÉDIAS

À semelhança do método hierárquico, o método das k-médias considerou oito *clusters*. O cálculo foi limitado a 30 iterações tendo estas sido suficientes para alcançar o critério de convergência definido (convergência para zero). Na Figura 5.13 está representada a distribuição de *clusters* com o método das k-médias. Tal como no método hierárquico, a medida de distância selecionada foi o quadrado da distância euclidiana.

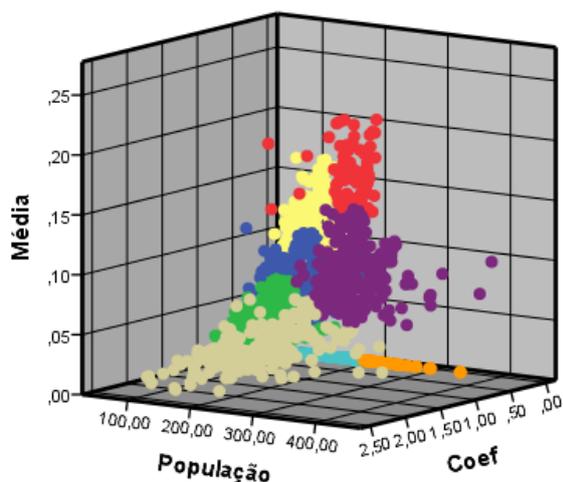


Figura 5.12 – Representação gráfica de clusters com o critério de Ward.

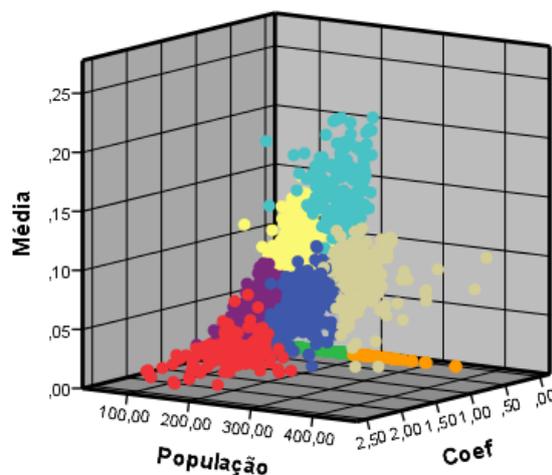


Figura 5.13 – Representação gráfica de clusters com o método das k-médias.

O facto de os métodos partitivos (entre os quais está as k-médias) não considerarem a matriz de semelhanças entre os valores permite que haja uma maior liberdade ao nível dos cálculos para agrupar os dados diretamente do original. Esse facto é facilmente identificado uma vez que os *clusters* definidos por este método se aproximam bastante daqueles definidos graficamente de forma visual e empírica.

5.2.4. CARACTERIZAÇÃO DOS *CLUSTERS*

Para o presente estudo foram seleccionados os dois métodos – hierárquico e k-médias. Apesar de não existir um critério superior a outro, no método hierárquico foi escolhido o critério de Ward por apresentar *clusters* mais uniformes em dimensão e por se aproximarem da divisão alcançada com o método das k-médias (facilitando a comparação).

Na Figura 5.14 está representado o número de ecopontos por *cluster* e na Figura 5.15 a média do enchimento de ecopontos, demonstrando assim a heterogeneidade da população agrupada em diferentes *clusters*.

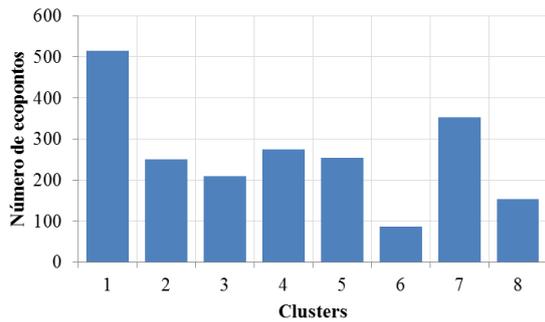


Figura 5.14 – Gráfico com comparação de número de ecopontos por cluster pelo método hierárquico.

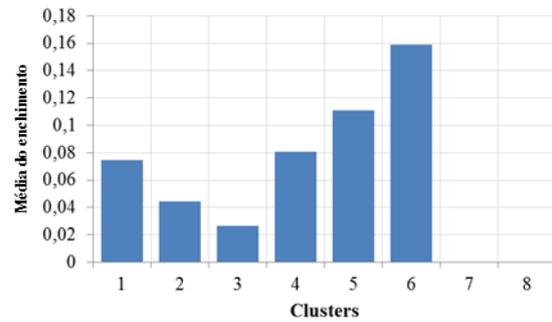


Figura 5.15 – Gráfico com comparação da média do enchimento de ecopontos por cluster pelo método hierárquico

Na Figura 5.16 está representado o coeficiente de variação do enchimento dos ecopontos. A população por ecoponto é representada na Figura 5.17, demonstrando que é praticamente constante para todos os ecopontos, revelando poder ser uma medida de proximidade que não tenha uma influência tão predominante quanto esperado.

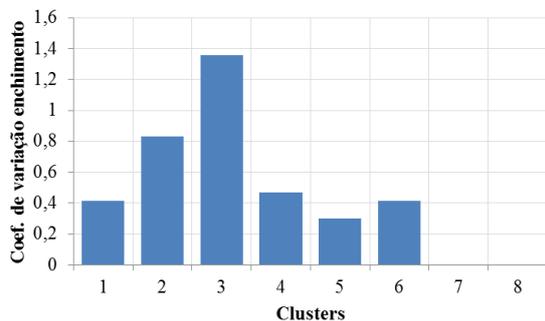


Figura 5.16 – Gráfico com comparação do coef. de variação do enchimento de ecopontos por cluster pelo método hierárquico

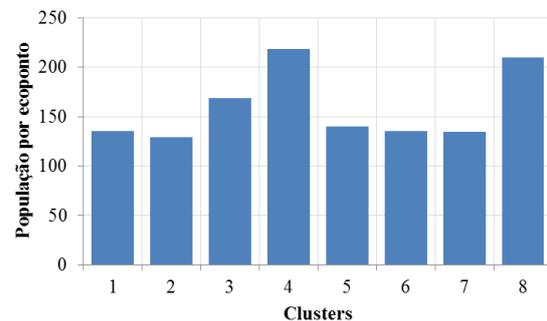


Figura 5.17 – Gráfico com comparação da população por ecopontos por cluster pelo método hierárquico

Verifica-se à partida que os *clusters* 7 e 8, que têm valor zero, serão excluídos da atribuição de sensores, uma vez que não registam qualquer atividade ao longo de seis meses. É importante ressaltar que os dados disponíveis levam a esta conclusão, mas que seria necessária uma análise com recurso a mais dados e mais informação para poder chegar a uma decisão final que poderia até passar pela retirada desses ecopontos.

Na Figura 5.18 está representado o número de ecopontos por *cluster* e na Figura 5.19 a média do enchimento de ecopontos, demonstrando assim a variedade da população agrupada em diferentes *clusters* criados pelo método das k-médias.

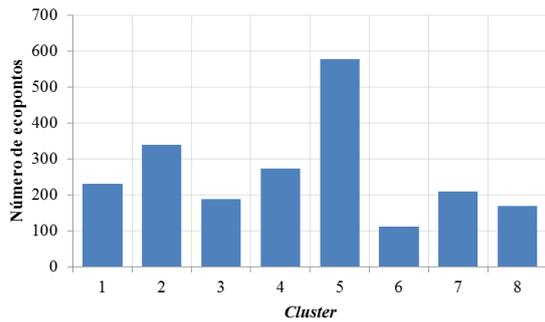


Figura 5.18 – Gráfico com comparação de número de ecopontos por cluster pelo método das k-médias .

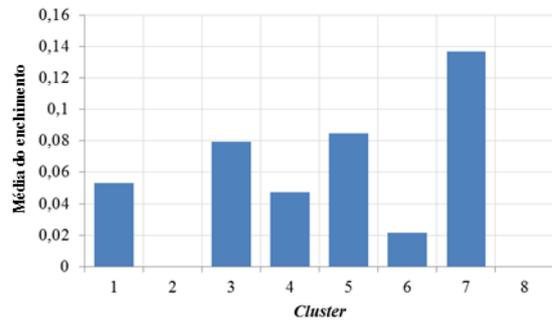


Figura 5.19 – Gráfico com comparação da média do enchimento de ecopontos por cluster pelo método das k-médias

Na Figura 5.20 está representado o coeficiente de variação do enchimento dos ecopontos. A população por ecoponto é representada na Figura 5.21, demonstrando pouca variação entre ecopontos, reforçando a tese de poder ser uma medida de proximidade que não tenha uma influência tão predominante quanto esperado, tal como já verificado para o método hierárquico.

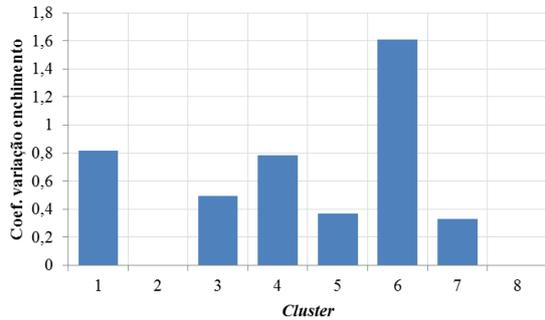


Figura 5.20 – Gráfico com comparação do coef. de variação do enchimento de ecopontos por cluster pelo método das k-médias

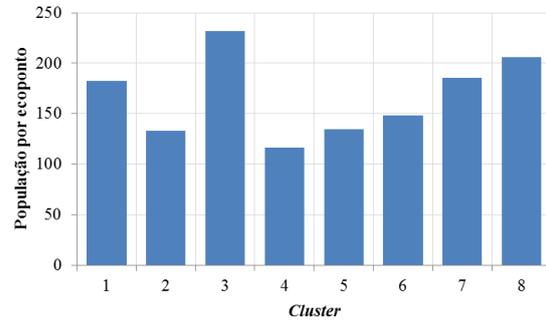


Figura 5.21 – Gráfico com comparação da população por ecopontos por cluster pelo método das k-médias

5.3. ANÁLISE DE DADOS

Com o objetivo de identificar a proporção ideal de sensores por *cluster*, realizou-se uma análise de cenários em que várias proporções foram testadas. Na análise realizada não foram feitas as simulações de rotas, tendo-se assumido como pressuposto que os benefícios adquiridos com a otimização de rotas realizada por Gonçalves (2014) são constantes para todos os

cenários que consideram ecopontos monitorizados. Assume-se que a variação dos benefícios recolhidos possa estar representada no risco de cada solução.

Cada *cluster* foi caracterizado com o número de ecopontos nele incluído, assim como a média de enchimento, o coeficiente de variação e a população por ecoponto. O valor da média estará na base do cálculo do dia da recolha – número de dias necessários para que a média do enchimento seja igual a δ , que no caso da Valorsul é 0,75.

Para a análise foram testadas várias percentagens de ecopontos monitorizados (de 1% a 99%) onde se avalia a capacidade da amostra de ecopontos monitorizados extrapolar os resultados para a totalidade do *cluster*. Utilizou-se uma amostra aleatória seletiva de ecopontos (ver capítulo 4.2.1).

Com base na média dos valores das amostras, foi calculado o dia em que deveria ser feita a recolha (ou seja, quando atinge um enchimento de 0,75). Foi considerado o mesmo dia de recolha para todos os ecopontos do *cluster* em análise, por forma a não tornar o processo ainda mais complexo. Fixado o dia, são identificados os *k-ésimos* ecopontos, de uma lista ordenada crescente do enchimento médio diário dos ecopontos do *cluster*, para formar a amostra.

Na Tabela 5.4 estão representados, a título de exemplo, os valores obtidos para uma monitorização de 5% e 75% ecopontos no *cluster* 1. Esse *cluster* caracteriza-se por ter 515 ecopontos, ou seja 25 e 386 ecopontos para uma amostra de 5% e 75%, respetivamente.

Tabela 5.4 – Valores de monitorização de 5% e 75% ecopontos do cluster 1 com o método hierárquico

Ecopontos monitorizados	Média do enchimento médio diário	Dia da recolha	Amostra		
			Não aptos	Aptos	Sobrelotados
5% (25)	0,07523	10	12 (48%)	12 (48%)	1 (4%)
75% (386)	0,07436	11	134 (34%)	222 (58%)	30 (8%)

De acordo com a amostra de 25 ecopontos monitorizados, ao fim de 10 dias ($0,75/0,07523=9,97$ dias), 12 ecopontos encontravam-se aptos, 12 não aptos e 1 sobrelotado. Por outro lado, a amostra de 386 ecopontos monitorizados, ao fim de 11 dias ($0,75/0,07436=10,08$ dias), 222 ecopontos encontravam-se aptos, 134 não aptos e 30 sobrelotados.

Os valores obtidos com as amostras foram extrapolados para a totalidade dos 515 ecopontos do *cluster* e comparados com os valores reais no dia da recolha respetivo, ou seja, ao 10º dia e ao 11º dia, para as amostras de 5% e 75% respetivamente. Na Tabela 5.5 estão representados os valores da estimativa da amostra, os valores reais e a diferença/erro entre eles.

Tabela 5.5 – Valores da estimativa das amostras com 5% e 75% ecopontos monitorizados comparados com os valores reais para o cluster 1 com o método hierárquico. (Legenda: NA - Não apto; A - Apto; S - Sobrelotado).

Ecopontos monitorizados	Estimativa			Real			Diferença/Erro (Real-Estimativa)		
	NA	A	S	NA	A	S	NA	A	S
5%	247	247	21	264	241	10	17	-6	-11
75%	178	296	40	178	296	41	0	0	1

É possível verificar que para a amostra com 5% de ecopontos a diferença entre a estimativa e a realidade é significativa, enquanto que para a amostra com 75% de ecopontos a diferença é praticamente nula. Assim constata-se que as expectativas iniciais de que quanto menor for o número de ecopontos monitorizados, maior será o erro dos resultados de enchimento relativamente à realidade.

Esse erro pode ser transformado em custo, de acordo com o explicado no capítulo 4.2.2. Assim, a distância média entre ecopontos foi calculada com base nos quilómetros percorridos em cada rota da Valorsul e a quantidade de ecopontos existente. O valor da distância média entre ecopontos é de 2,235 km, logo o desvio na rota (ida e volta) tem um comprimento de 4,470 km ($2 \times 2,235$ km), o que equivale a 4,47€/desvio para o caso de recolher um ecoponto não apto. Uma vez que para ecopontos sobrelotados o custo depende da quantidade de resíduos extra, foi adotado o mesmo valor de 4,47€, uma vez que podem ditar a realização de uma nova rota de recolha.

Na Tabela 5.6 está exemplificado o cálculo dos custos para as amostras com 5% e 75% dos ecopontos monitorizados.

Tabela 5.6 – Cálculo dos valores de custos mensais para as amostras de 5% e 75% dos ecopontos monitorizados.

Amostra	Ecopontos monitorizados	Investimento (10,25€/ecop.)	Erro na estimativa (NA+S)	Custo do erro (4,47€/ecop.)	Erros / mês	Total
5%	25	256,25 €	(17+11) = 28	125,16€	30/10 dias	631,73€
75%	386	3.956,50 €	(0+1) = 1	4,47€	30/11 dias	3.968,69€

Na Figura 5.22 estão representados os custos do erro da amostra e do investimento em função da proporção de ecopontos monitorizados no *cluster* 1 do método hierárquico. Na Figura 5.23 está representada a soma das duas curvas que permite, identificar o valor ótimo que equilibra o erro e o investimento. Os resultados para os restantes *clusters* são apresentados no Anexo A1.

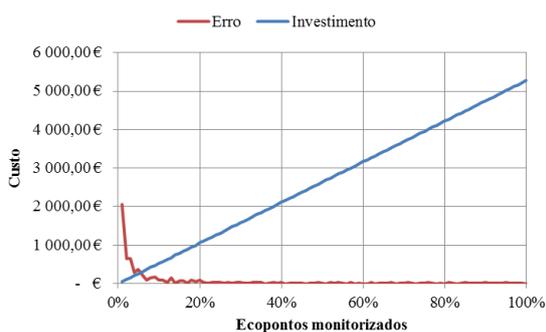


Figura 5.22 – Gráfico com o erro e o investimento mensal no sistema para proporções de monitorização de 1% a 100% no cluster 1 do método hierárquico.

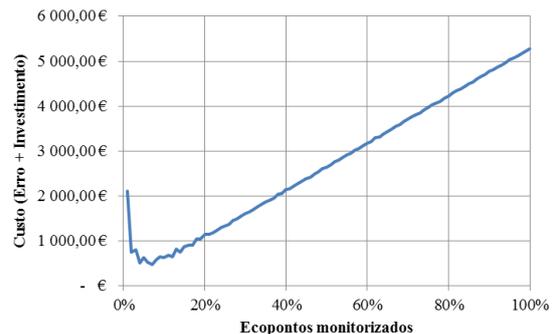


Figura 5.23 – Gráfico com o custo mensal do sistema para proporções de monitorização de 1% a 100% no cluster 1 do método hierárquico.

Da observação dos gráficos é possível verificar que existe um ponto mínimo que equilibra o erro e o investimento da melhor forma (amostra com 7% e custo mensal de 467,94 €). O investimento tem uma progressão linear com derivada igual a 10,25€, enquanto que o erro estabiliza para rácios de monitorização a partir de aproximadamente os 10%. O valor da distância entre ecopontos, utilizado como métrica para a quantificação do erro foi testado através de uma análise de sensibilidade, tendo-se alterados os valores de distância por forma a identificar se haveria alterações na solução obtida. Para o *cluster* 1 do método hierárquico o valor de 7% mantém-se para distâncias do dobro e do triplo daquela considerada; para soluções onde a distância é metade ou um terço dessa distância observou-se a alteração da solução para 4%,

contudo, de uma forma conservativa, foi assumido que a solução está validada. No Anexo A2 são apresentados os valores para os restantes *clusters* no que diz respeito a análise de sensibilidade.

Na Tabela 5.7 é feito um resumo dos valores associados a cada *cluster* para o método hierárquico para o valor mínimo de custo identificado.

Tabela 5.7 – Valores de custo mensal e número de ecopontos monitorizados para cada cluster para o método hierárquico.

<i>Cluster</i>	1	2	3	4	5	6	7	8	Total
Ecopontos monitorizados	36	17	10	19	10	11	0	0	103
	7%	7%	5%	7%	4%	13%	0%	0%	4,9%
Custo mensal [€]	467,94	230,27	157,93	221,96	180,24	167,17	0,00	0,00	1.425,51 €

Aos 103 ecopontos identificados é necessário somar os 14 que foram retirados na análise de *clusters* por ser demasiado discrepantes dos restantes. O custo mensal total obtido pelo método hierárquico é **1.569,01€**.

Na Tabela 5.8 é feito um resumo dos valores associados a cada *cluster* para o método das k-médias para o valor mínimo de custo identificado.

Tabela 5.8 – Valores de custo mensal e número de ecopontos monitorizados para cada cluster para o método das k-médias .

<i>Cluster</i>	1	2	3	4	5	6	7	8	Total
Ecopontos monitorizados	6	0	24	13	28	8	16	0	95
	3%	0%	13%	5%	5%	8%	8%	0%	4,5%
Custo mensal [€]	129,52	0,00	273,21	226,78	347,47	106,01	209,35	0,00	1.292,34

À semelhança do que foi feito para o método hierárquico, também para os 95 ecopontos identificados com as k-médias é necessário somar os 14 que foram retirados na análise de *clusters* por ser demasiado discrepantes dos restantes. O custo mensal total obtido pelo método hierárquico é **1.435,84 €**.

Ambos os métodos produziram soluções equivalentes e da mesma ordem de grandeza. Enquanto que através do método hierárquico se obteve um custo mensal de 1.425,51 € o método das k-médias obteve 1.292,34 € (-10%). Se se retirar a esses valores a parcela relativa ao investimento em sensores, isolando desta forma a parcela do erro, obteve-se 369,76 € para o método hierárquico e 318,59 € para o método das k-médias, 26% e 25% dos custos mensais das soluções, respetivamente.

Na Tabela 5.9 estão resumidos os valores mais significativos das soluções alcançadas através de cada um dos métodos.

Tabela 5.9 – Valores de custo mensal e número de ecopontos monitorizados para cada método de análise de clusters.

Método	Custo mensal da solução	Custo mensal do erro	Ecopontos monitorizados	Custo mensal do erro/ecoponto
Hierárquico	1.425,51€	369,76€	103	3,59€
K-médias	1,292,34€	318,59€	95	3,35€

É possível novamente constatar que os valores são da mesma ordem de grandeza, não se verificando alterações significativas entre ambos. Porém, há que destacar que o método das k-médias é aquele que menor custo do erro por ecoponto apresenta, ou seja, o custo do erro combinado com o número total de ecopontos monitorizados é inferior (3,35€) ao verificado para o método hierárquico (3,59€).

5.4. ANÁLISE DE RISCO

Na descrição do modelo de apoio à decisão, no capítulo 4.3, foi identificada a necessidade de fazer um levantamento dos riscos existentes, quantificá-los e gerir aqueles que não forem passíveis de serem eliminados. Neste caso, foi apenas considerado para a análise o risco do erro calculado devido à diferença entre a estimativa e a realidade (risco associado a variações na taxa de enchimento e alterações dos índices de adesão da população à reciclagem). Assim, a análise de risco resume-se à determinação da reserva de contingência necessária para a aceitação ativa do risco associado ao custo determinado para a solução.

O custo de investimento é um valor fixo associado à amostra identificada. Contudo, o erro entre a estimativa da amostra e a realidade, que dá origem a um custo do erro (relacionado com os quilómetros a mais percorridos e o custo que essa distância tem para a Valorsul), permitem definir uma reserva de contingência, à disposição do gestor do sistema, para fazer face às recolhas adicionais que terá de realizar. Apenas o custo do erro foi considerado para a simulação de Monte Carlo, uma vez que é aí que reside a incerteza. Assim, foi utilizado o método de Monte Carlo, que tem por base a simulação de uma mesma solução várias vezes, recolhendo num histograma a frequência com que cada solução ocorre. Por razões de simplificação do modelo, foi adotada uma distribuição normal para a simulação de Monte Carlo. A análise foi feita com recurso ao MS Excel 2007, tendo assumido simulações infinitas, logo a curva resultante da análise representa uma soma cumulativa dos valores da distribuição normal.

Quanto maior for a probabilidade de uma solução ocorrer, maior é a confiança que existe nessa solução, e conseqüentemente, menor é o risco. Logo, para o cálculo do custo da solução para os vários níveis de confiança foi utilizado o custo do erro como média e como desvio padrão foi utilizado o coeficiente de variação, obtido com a fórmula (4.2) multiplicado pela média; uma vez que o coeficiente de variação é o quociente entre o desvio padrão e a média, quando multiplicado pela média, transforma-se em desvio padrão.

Na Figura 5.24 está representada a título de exemplo a curva para todos os níveis de confiança entre 0% e 100% para o *cluster* 1 do método hierárquico. Os gráficos para os restantes *clusters* do método hierárquico, bem como para o método das k-médias encontram-se representados no Anexo A3.

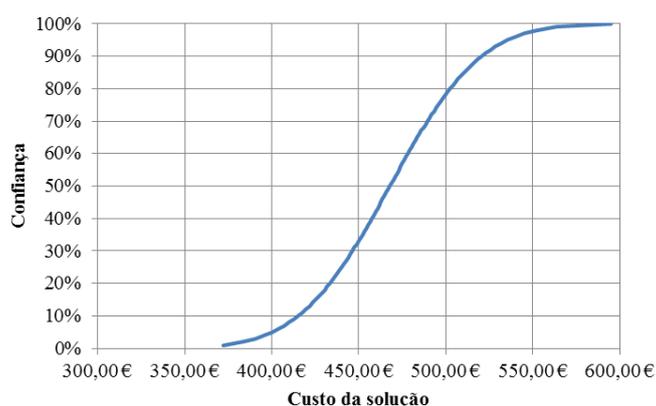


Figura 5.24 – Curva S que relaciona o custo da solução com a confiança para o cluster 1 do método hierárquico.

Da observação da curva é possível verificar que para níveis baixos de confiança (até sensivelmente aos 20%) o custo da solução aumenta mais rapidamente, estabilizando para valores intermédios (entre os 20% e os 80%), voltando a ter um comportamento em que o custo da solução aumenta mais rapidamente, não havendo um acompanhamento proporcional do aumento do nível de confiança. A curva tem uma maior ou menor amplitude consoante o valor do desvio-padrão; assim, para valores de desvio padrão menores, a amplitude é menor, fazendo com que haja uma maior proximidade dos valores do custo da solução para os diferentes níveis de confiança; em oposição, para valores de desvio padrão maiores, a amplitude é maior, fazendo com que haja uma menor proximidade dos valores do custo da solução para os diferentes níveis de confiança.

5.4.1. MÉTODO HIERÁRQUICO

Na Tabela 5.10 são apresentados os coeficientes de variação para cada *cluster* do método hierárquico, exceto para os *clusters* 7 e 8, cujo valor de erro é zero, pois são ecopontos sem enchimento.

Tabela 5.10 – Coeficiente de variação de cada cluster do método hierárquico.

<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>	<i>Cluster 4</i>	<i>Cluster 5</i>	<i>Cluster 6</i>
0,415	0,833	1,357	0,468	0,301	0,364

Na Tabela 5.11 são apresentados os custos de solução associados a vários níveis de confiança para cada *cluster* obtido com o método hierárquico. Os níveis de confiança selecionados foram 50%, 80%, 90%, 95% e 99%. O nível de confiança de 50% foi selecionado por se apresentar como sendo neutro, na medida em que o valor apresentado não tem qualquer tipo de influência do desvio padrão. Os valores que se seguem, começando no 80% até ao 99%, são considerados valores aceitáveis do ponto de vista da confiança na solução, uma vez que integram no seu resultado uma componente suficientemente considerável de incerteza/risco que permite obter uma reserva de contingência mais robusta e cautelosa. Com os valores a aproximarem-se dos 100% a cautela é crescente, e consequentemente a probabilidade de se exceder a reserva de contingência é decrescente.

Tabela 5.11 – Custo da solução para diferentes níveis de confiança para o método hierárquico

Confiança	50%	80%	90%	95%	99%
<i>Cluster 1</i>	467,94 €	502,50 €	520,57 €	535,48 €	563,47 €
<i>Cluster 2</i>	230,27 €	269,54 €	290,07 €	307,02 €	338,83 €
<i>Cluster 3</i>	157,93 €	221,23 €	254,32 €	281,64 €	332,90 €
<i>Cluster 4</i>	221,96 €	232,68 €	238,28 €	242,90 €	251,58 €
<i>Cluster 5</i>	180,24 €	199,93 €	210,23 €	218,73 €	234,68 €
<i>Cluster 6</i>	167,17 €	183,84 €	192,55 €	199,75 €	213,25 €
<i>Cluster 7</i>	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
<i>Cluster 8</i>	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
Total	1 425,51 €	1 609,72 €	1 706,02 €	1 785,52 €	1 934,71 €

A diferença entre o custo da solução para uma confiança de 50% e 99% tem uma relação direta com o coeficiente de variação, tal como se esperava. O *cluster 3*, que era o que apresentava

o maior coeficiente de variação, é também aquele que apresenta a maior diferença relativa entre as soluções para 50% e 99% de confiança, sendo a segunda 111% superior à primeira. Por outro lado, o *cluster* 5, que é aquele que apresenta o menor coeficiente de variação não é aquele que apresenta uma menor diferença relativa entre as soluções para 50% e 99% de confiança (+30%), uma vez que a solução identificada para este *cluster* tem uma amostra de apenas 4% do *cluster*, tendo valores de investimento e custo do erro de uma ordem de grandeza semelhantes, logo a variação do valor do erro (e consequentemente do custo da solução) é mais variável.

No que diz respeito ao valor total da soma dos *clusters*, a diferença relativa entre as soluções para 50% e 99% de confiança é de +36%.

5.4.2. MÉTODO DAS K-MÉDIAS

Na Tabela 5.10 são apresentados os coeficientes de variação para cada *cluster* do método k-médias, exceto para os *clusters* 2 e 8, cujo valor de erro é zero, pois são ecopontos sem enchimento.

Tabela 5.12 – Coeficiente de variação de cada cluster do método hierárquico.

<i>Cluster 1</i>	<i>Cluster 3</i>	<i>Cluster 4</i>	<i>Cluster 5</i>	<i>Cluster 6</i>	<i>Cluster 7</i>
0,819	0,490	0,785	0,366	1,608	0,322

À semelhança do ocorrido para o coeficiente de variação calculado para o método hierárquico, também para o método das k-médias não se verifica qualquer influência de uma possível maior adesão da população no valor do coeficiente de variação.

Na Tabela 5.13 são apresentados os custos de solução associados a vários níveis de confiança para cada *cluster* obtido com o método das k-médias. À semelhança do método hierárquico foram selecionados os níveis de confiança 50%, 80%, 90%, 95% e 99%.

Tabela 5.13 – Custo da solução para diferentes níveis de confiança para o método das k-médias

Modelo de apoio à decisão para a utilização de TIC na otimização da recolha de resíduos recicláveis

Confiança	50%	80%	90%	95%	99%
<i>Cluster 1</i>	129,52 €	153,28 €	165,70 €	175,96 €	195,20 €
<i>Cluster 2</i>	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
<i>Cluster 3</i>	273,21 €	284,48 €	290,37 €	295,23 €	304,35 €
<i>Cluster 4</i>	226,78 €	259,45 €	276,53 €	290,63 €	317,08 €
<i>Cluster 5</i>	347,47 €	366,09 €	375,83 €	383,87 €	398,95 €
<i>Cluster 6</i>	106,01 €	138,50 €	155,48 €	169,51 €	195,82 €
<i>Cluster 7</i>	209,35 €	222,02 €	228,64 €	234,11 €	244,37 €
<i>Cluster 8</i>	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
Total	1 292,34 €	1 423,82 €	1 492,54 €	1 549,30 €	1 655,76 €

À semelhança do já verificado para o método hierárquico, verifica-se que a diferença entre o custo da solução para uma confiança de 50% e 99% tem uma relação direta com o coeficiente de variação. O *cluster 6*, que era o que apresentava o maior coeficiente de variação, é também aquele que apresenta a maior diferença relativa entre as soluções para 50% e 99% de confiança, sendo a segunda 85% superior à primeira. Por outro lado, o *cluster 5*, que é aquele que apresenta o menor coeficiente de variação não é aquele que apresenta uma menor diferença relativa entre as soluções para 50% e 99% de confiança (+15%), uma vez que a solução identificada para este *cluster* tem uma amostra de apenas 5% do *cluster*, tendo valores de investimento e custo do erro de uma ordem de grandeza semelhantes, logo a variação do valor do erro (e consequentemente do custo da solução) é mais variável.

No que diz respeito ao valor total da soma dos *clusters*, a diferença relativa entre as soluções para 50% e 99% de confiança é de +28%.

5.4.3. COMPARAÇÃO DAS SOLUÇÕES

Numa comparação entre os valores obtidos para a reserva de contingência com ambos os métodos, é possível verificar que o método das k-médias apresenta valores relativamente mais baixos. Na Figura 5.25 estão representadas as curvas, para ambos os métodos em função do custo da solução e da confiança.

À semelhança do que já tinha sido verificado após a quantificação do custo do erro, constata-se que a tendência se mantém, mesmo para níveis de confiança mais elevados. Verifica-se que a diferença entre solução até aumenta para níveis de confiança mais elevados. Esta diferença pode ser justificada por uma maior homogeneidade dos *clusters* no método das k-médias, agrupando num *cluster* os ecopontos com enchimentos mais inconstantes, ficando os outros com enchimentos mais previsíveis, logo, menos incertos.

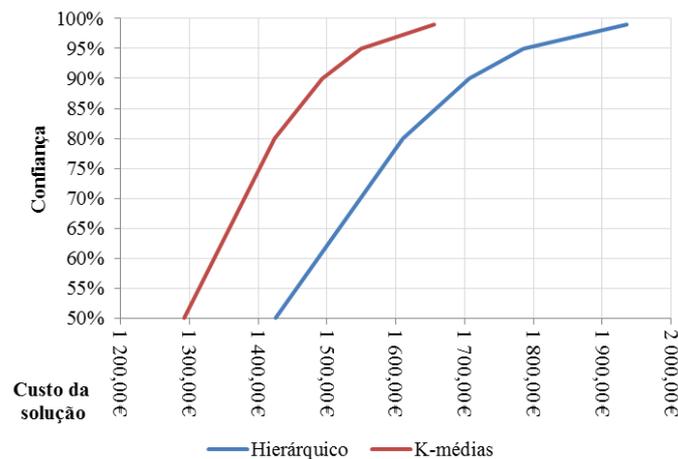


Figura 5.25 – Comparação das curvas com o custo da solução em função da confiança para o método hierárquico e para o método das k-médias.

5.5. ANÁLISE CUSTO/BENEFÍCIO

Para a realização da comparação entre as soluções determinadas através do modelo proposto e as soluções já conhecidas previamente (a situação atual e a solução proposta por Gonçalves (2014)) foram determinados os valores relativos a operação, o investimento necessário na nova tecnologia e a reserva de contingência associada ao risco. Na Tabela 5.14 são apresentados os acima mencionados valores para cada uma das soluções. No que diz respeito às reser-

Modelo de apoio à decisão para a utilização de TIC na otimização
da recolha de resíduos recicláveis

vas de contingência são apresentados os valores para 99% de confiança, uma vez que é o valor mais gravoso.

Tabela 5.14 – Análise de custo/benefício anual das várias soluções existentes.

Método	Custos operacionais	Investimento em sensores	Reserva de contingência*	Total
Cenário atual	173.212,00€,	0,00€	0,00€	173.212,00€,
Proposta Gonçalves (2014)	126.444,76€	259.407,00€	0,00€	385.851,76€
Método hierárquico		12.669,00€	10 547,52 €	149 661,28 €
Método das k-médias		11.685,00€	8 184,17 €	146 313,93 €

* Valores das reservas de contingência para a solução com 99% de confiança.

Da análise dos valores totais, que são uma soma dos custos operacionais associados à realização das rotas de recolha, do investimento em sensores e da reserva de contingência, é possível verificar que a solução alcançada através do modelo proposto – para qualquer um dos métodos utilizados – é a mais competitiva por apresentar um valor mais reduzido.

Face ao cenário atual, a solução proposta apresenta uma poupança anual estimada de aproximadamente 25k euros – uma redução de 14%. Essa diminuição só não obtém valores mais elevados devido ao facto dos custos operacionais ainda representarem grande parte do total (cerca de 90%). Por outro lado, o modelo permitiu identificar soluções que apresentam valores substancialmente menores face à solução identificada por Gonçalves, obtendo-se poupanças na implementação da solução (investimento + reserva de contingência) de 95%. É ainda de referir que a reserva de contingência pode ter valores inferiores, uma vez que pode ser considerado um nível de confiança inferior a 99%; todavia, essa decisão depende do perfil de risco do decisor. Uma vez que o valor da reserva de contingência não necessita ser totalmente utilizado, é uma questão exclusivamente da gestão a alocação da verba para esta reserva.

De realçar que a poupança adquirida para as embalagens é possivelmente a mais pequena dos três resíduos, pois foi aquela que apresentou o valor mais reduzido de otimização (Tabela 5.3). Assim, pode perspetivar-se que para o papel e para o vidro se venham a observar valores de redução de custo ainda mais relevantes.

CAPÍTULO 6. CONCLUSÕES

6.1. CONCLUSÕES

A falta de eficiência nas rotas de recolha de resíduos recicláveis é fator crítico para as empresas de gestão de resíduos. Com uma tendência crescente da quantidade de resíduos recicláveis produzidos, são desperdiçados recursos e tempo para recolher ecopontos que se encontram com níveis de enchimento insuficientes; e a tendência é que esse desaproveitamento de recurso continue caso não se adotem medidas de melhoria no sistema. A monitorização do enchimento de ecopontos com TIC é um primeiro passo para a otimização das rotas, contudo revelou-se demasiado dispendioso segundo o estudo de Gonçalves (2014).

Foi realizada uma revisão de literatura onde se constatou, de acordo com diversos autores, que existem uma série de fatores que tornam a gestão de resíduos, em particular os resíduos recicláveis, pouco eficiente, nomeadamente o forma como se processa a recolha dos resíduos. Foram também analisados alguns estudos com a aplicação de diversas TIC para a monitorização do enchimento de contentores, avaliando os seus prós e contras. Para melhor identificar o modelo de apoio à decisão a aplicar na presente tese foram ainda revistos alguns estudos de vários autores sobre modelo de apoio à decisão, com particular destaque para os modelos com incerteza, e a sua integração com os princípios de gestão de risco. Por fim, foram ainda pesquisados os fundamentos da análise de *clusters* bem como as várias técnicas que podem ser utilizadas para o agrupamento de dados.

O presente trabalho propôs um modelo de apoio à tomada de decisão para tornar a solução de utilização das TIC economicamente viável ao conceber um sistema de monitorização parcial dos ecopontos. Agrupando os ecopontos mais semelhantes em clusters, tendo em conta parâmetros relacionados com o enchimento, entre outros, é possível obter grupos mais homogêneos. A partir de cada *cluster* foram selecionadas amostras que permitiam caracterizar, com um certo grau de erro, a totalidade do *cluster*. O erro da amostra combinado com o investimento necessário em TIC permitiu identificar qual a dimensão ótima da amostra para cada *cluster*. Contudo, uma vez que o enchimento de ecopontos não é constante, foram considerados alguns fatores que podem introduzir alguma incerteza, tal como as flutuações de enchi-

mento ou a adesão da população à reciclagem. Essa incerteza foi quantificada através de uma análise de risco com base na simulação de Monte Carlo.

O modelo foi aplicado a um caso de estudo, por forma a validar a sua adequação para a realidade de uma empresa – a Valorsul. Os dados fornecidos permitiram aplicar o modelo de apoio à decisão, sendo que as conclusões mais específicas de cada quadrante do modelo são abordadas de seguida.

A análise de *clusters* permitiu dividir o universo inicial de 2109 ecopontos, que devido à sua dimensão e heterogeneidade era muito complexo, em *clusters* de dimensão inferior (entre 87 e 577 ecopontos), o que permitiu a análise independente de cada um deles. Foram definidas como medidas de proximidade: (i) a média de enchimento; (ii) o coeficiente de variação de enchimento; e (iii) os habitantes/ecoponto. Dos vários métodos utilizados para a análise de *clusters* foram selecionados dois: método hierárquico com o critério de Ward e método das k-médias. Foi possível identificar que os *clusters* gerados por ambas as análises agruparam ecopontos com valores de enchimento e coeficientes de variação semelhantes em *clusters* diferentes; contudo, constatou-se que a população por ecoponto não teve uma influência relevante na definição dos grupos de ecopontos.

A seleção das amostras em cada *cluster* permitiu que fossem realizadas estimativas que previam o real estado de cada cluster. O erro resultante entre a comparação da estimativa produzida pela amostra e a realidade foi quantificado tendo em conta o número de quilómetros a realizar a mais (designado por custo do erro). Ao somar o custo do erro com o custo de investimento em sensores de monitorização de enchimento, para diversas diferentes dimensões de amostras, foi possível identificar os valores ótimos que melhor equilibravam o custo do erro e o investimento necessário. Foram assim apurados os custos mensais para a dimensão ótima da amostra (dizer qual a dimensão ótima – dizer quantos ecopontos devem ser monitorizados) para cada um dos métodos de *clusters* utilizados: 1.425,51€ e 1.292,34€ para o método hierárquico e das k-médias, respetivamente. Apesar de serem valores da mesma ordem de grandeza, foi possível verificar que o método das k-médias permitiu alcançar um resultado mais vantajoso ao nível do investimento necessário.

Os valores obtidos na análise de dados foram submetidos a uma análise de risco com o objetivo de definir uma reserva de contingência que permitisse fazer face ao erro existente entre a

estimativa e a realidade. Com uma versão simplificada da simulação de Monte Carlo foi possível determinar para diferentes níveis de confiança, o valor que deveria ser considerado mensalmente para a reserva de contingência tendo as soluções atingido (para 99% de confiança) 1.934,71 € e 1.655,76 € para o método hierárquico e das k-médias, respetivamente. As soluções apresentaram custos finais anuais de 149.661,28 € e 146.313,93 € para o método hierárquico e das k-médias, respetivamente.

As soluções obtidas com o modelo desenvolvido, quando comparadas com o cenário atual, em que nenhum ecoponto é monitorizado (173.212,00€) e com o cenário proposto por Gonçalves (2014) em que todos os ecopontos eram monitorizados (385.851,76€), permitiram constatar que apesar da existência de uma reserva de contingência, o valor total das soluções desenvolvidas é significativamente inferior aos cenários com que são comparados. A monitorização de 5% de ecopontos permite uma poupança de 14% face à situação atual e de 61% face ao cenário estudado por Gonçalves (2014). O modelo de apoio à decisão ficou assim testado para a Valorsul e comprovou-se a sua aplicabilidade e adequabilidade.,

6.2. LIMITAÇÕES DO MODELO E DO CASO DE ESTUDO

Apesar dos ganhos obtidos na aplicação do modelo de apoio à decisão para o caso de estudo da Valorsul, o modelo não deve ser considerado como totalmente válido, uma vez que seria necessário aplicá-lo a mais casos, por forma a permitir tirar conclusões mais robustas.

Por outro lado, os dados fornecidos pela Valorsul foram insuficientes para obter resultados mais conclusivos do modelo. Caso tivessem sido facultados valores para um maior período de tempo (como por exemplo 2 ou 3 anos) seria possível verificar tendências sazonais, evoluções anuais, entre outros. Esses valores permitiriam que o modelo fosse testado a nível superior de complexidade, o que seria obviamente benéfico.

6.3. PERSPETIVAS DE TRABALHO FUTURO

O modelo proposto pelo presente trabalho foi desenvolvido tendo como base a informação prestada pela Valorsul. Os dados facultados, apesar de úteis, revelaram-se insuficientes para uma mais completa análise do problema. Assim, um dos principais motivos de trabalho futuro no seguimento da presente dissertação seria a aplicação do modelo de apoio à decisão consi-

derando dados completos, para os vários tipos de resíduos e durante um maior período de tempo, preferencialmente alguns anos, por forma a avaliar fenómenos como a sazonalidade.

Por outro lado, o facto de os custos operacionais terem sido extrapolados dos resultados alcançados por Gonçalves (2014), em que todos os ecopontos eram monitorizados, pode enviesar as conclusões relativamente aos benefícios da solução proposta. Nesse sentido, poderia ser relevante realizar simulações das rotas de recolha para vários cenários de monitorização dos ecopontos.

Por fim, a elaboração de um teste-piloto real, na região oeste ou outra, seria uma forma adequada de validação do modelo proposto, podendo contribuir para a otimização e refinamento do modelo de apoio à decisão para a utilização de TIC na otimização da recolha de resíduos recicláveis.

BIBLIOGRAFIA

AGÊNCIA PORTUGUESA DO AMBIENTE (2014) *Plano Nacional de Gestão de Resíduos 2014-2020*.

AREBEY, M., HANNAN, M. A., e BASRI, H. (2011) *Integrated technologies for solid waste bin monitoring system*, Environment Monitoring Assessment, vol. 177, 399–408.

BERKUN, M., ARAS, E. e ANILAN T. (2011) *Solid waste management practices in Turkey*, The Journal of Material Cycles and Waste Management, vol. 13, pp. 305-313.

BISWAS, T. (1997), *Decision Making Under Uncertainty*, St. Martin's Press.

BOTEZATU, M. (2016), *Insight into project risk management*, Journal of Information Systems & Operations Management, pp 1-14.

BROEKHUIZEN, H; GROOTHUIS-OUDSHOORN, C; VAN TIL, J.; HUMMEL, M.; e IJZERMAN, M. (2015), *A Review and Classification of Approaches for Dealing with Uncertainty in Multi-Criteria Decision Analysis for Healthcare Decisions*, Pharmacoeconomics, vol. 33, n.º 5, pp. 445-45.

CARMO, J. (2014), *Aplicação do método Monte Carlo na estimativa das necessidades nominais de energia útil para aquecimento e arrefecimento de um parque residencial*, Tese de Mestrado, Lisboa, Faculdade de Ciências UNL.

COX, T. (2008), *What's Wrong with Risk Matrices?*, Risk Analysis, vol. 28, n.º 2, pp. 497-512.

DAVIDE, P (2012) *Sistema de Gestão Para a Recolha de Material Reciclável*, Tese de Mestrado, Lisboa, Faculdade de Ciências UNL.

DIÁRIO DA REPÚBLICA (2006) Decreto-Lei n.º 178/2006, Lisboa,

GONÇALVES, D. (2014) *Tecnologias de Informação e Comunicação para Otimização da Recolha de Resíduos Recicláveis*. Tese de Mestrado, Lisboa: ISCTE Business School.

GOODWIN, P., WRIGHT, G. (1991), *Decision Analysis for Management Judgment*, John Wiley & Sons

GUIMARÃES, R. e CABRAL, J. (1997), *Estatística*, McGraw-Hill de Portugal, Lisboa

HANNAN, M., AREBEY, M., BEGUM, R., e BASRI, H. (2009), *Radio frequency identification (RFID) and communication technologies for solid waste bin and truck monitoring system*. Waste Management, vol. 31, pp. 2406–2413.

HANNAN, M, ZAILA, W. , AREBEY, M., BEGUM, R. e BASRI, H. (2014), *Feature extraction using Hough transform for solid waste bin level detection and classification*, Environment Monitoring Assessment, vol. 186, pp. 5381–5391.

HOWSON C., e URBACH, P. (2005), *Scientific Reasoning: The Bayesian Approach*, 3ª Edição, Open Court Publishing.

INSTITUTO NACIONAL DE ESTATÍSTICA, *Censos de 2011*, http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos_quadros, visitado a 08 de novembro de 2015 às 19h00.

JOHANSSON, O. M. (2006) *The effect of dynamic scheduling and routing in a solid waste management system*. Waste Management, vol. 26, pp. 875–885.

JORNAL OFICIAL DA UNIÃO EUROPEIA (2006) *Diretiva 2006/12/CE do Parlamento Europeu e do Conselho de 5 de abril de 2006*, Bruxelas.

KWAK, Y. e INGALL, L. (2007), *Exploring Monte Carlo Simulation Applications for Project Management*. Risk Management, vol. 9, n.º 1, pp.: 44-57.

LELAH, A., MATHIEUX, F., BRISSAUD, D., e GIMENO, V. (2010) *SensCity: a new project opening the way for sustainable services in the city based on a mutualized M2M infrastructure*. CIRP IPS2 Conference 2010, pp. 99-106. Linköping, Suécia.

LELAH, A., MATHIEUX, F., BRISSAUD, D. e VINCENT L. (2012), *Collaborative network with SMEs providing a backbone for urban PSS: a model and initial sustainability analysis*. Production Planning and Control, Taylor & Francis, vol. 23, pp. 299 - 314.

LIPOR, <http://www.lipor.pt/pt/noticias/municipio-de-vila-do-conde-inaugura-1-equipamento-inteligente-do-pais-para-a-gestao-de-residuos/>, visitado em 08 de fevereiro de 2016 às 15h00.

LOIZOU, P. e FRENCH, N. (2012), *Risk and uncertainty in development: A critical evaluation of using the Monte Carlo simulation method as a decision tool in real estate development projects*, Journal of Property Investment & Finance, vol. 30 n.º 2, pp. 198-210

MIGUEL, A. (2013), *Gestão Moderna de Projectos*, FCA – Editora de Informática, 7ª Edição

MURTEIRA, B (2007), *Introdução à Estatística*, Mcgraw-Hill Portugal, 2.ª Edição

OSSADNIK, W.; WILMSMANN, D. e NIEMANN, B. (2013), *Experimental evidence on case-based decision theory*, Theory and Decision, vol.75 n.º 2, pp. 211-232.

PROJECT MANAGEMENT INSTITUTE (2009), *Practice standard for Project Risk Management*, PMI, 1ª Edição.

PROJECT MANAGEMENT INSTITUTE (2013), *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*, PMI, 5ª Edição.

RAGSDALE, C. (2015), *Spreadsheet Modeling & Decision Analysis – A practical introduction to business analytics*, Cengage Learning, 7ª Edição.

REIS, E. (2001), *Estatística multivariada aplicada*, Edições Sílabo, 2ª Edição, Lisboa.

REVERTER, F.; GASULLA, M. e PALLAS-ARENY, R. (2003), *Capacitive level sensing for solid-waste collection*, Sensors Proceedings of IEEE, vol. 1, pp. 7-11.

ROVETTA, A. *et al.* (2009) *Early detection and evaluation of waste through sensorized containers for a collection monitoring application*. Waste Management, vol. 29, 2939–2949.

RUBINSTEIN, R. e KROESE, D. (2011), *Simulation and the Monte Carlo method*, John Wiley & Sons.

SADGROVE, K. (2015), *The Complete Guide to Business Risk Management*, Routledge, 3ª Edição, Nova Iorque.

SHAN-SHAN, C e LO, C. (2004) *Waste Management in Guangdong Cities: The Waste Management Literacy and Waste Reduction Preferences of Domestic Waste Generators*, Environmental Management vol. 33, n.º 5, pp. 692–711.

SIMCHI-LEVY, D.; KAMINSKY, P. e SIMCHI-LEVY, E. (2008), *Designing and managing the supply chain: Concepts, strategies and case studies*, McGraw-Hill International Editions.

SOCIEDADE PONTO VERDE, <http://pontoverde.pt/numeros.php>, visitado em 01 de junho de 2016 às 22h00.

VALORSUL, <http://www.valorsul.pt/>, visitado em 13 de fevereiro de 2016 às 22h00.

VICENTINI, F., GIUSTI, A., ROVETTA, A., FAN, X., HEB, Q., ZHU, M., ET AL. (2009). *Sensorized waste collection container for content estimation and collection optimization*. Waste Management, vol. 29, 1467–1472.

WANG, J.; BYRNES, J.; VALTORTA, M.; e HUHNS, M. (2012), *On the combination of logical and probabilistic models for information analysis*, Applied Intelligence, vol. 36, n.º 2, pp. 472-497.

WELLMAN, M. e DOYLE, J. (1992), *Modular Utility Representation for Decision-Theoretic Planning*, AIPS-92, pp. 236-242.

ANEXOS

ANEXO A1 – CUSTO MENSAL DA SOLUÇÃO

As figuras A.1 a A.6 representam os gráficos de custo mensal da solução de cada *cluster* para o método hierárquico utilizando o critério de Ward.

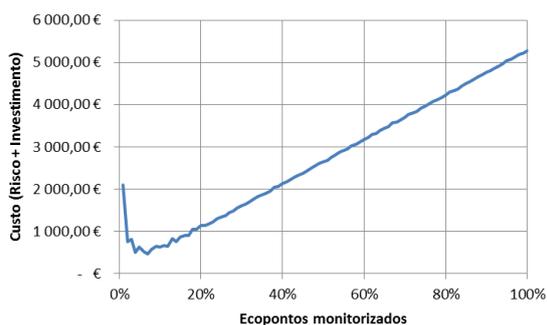


Figura A.1 – Custo mensal da solução Cluster 1 Método hierárquico.

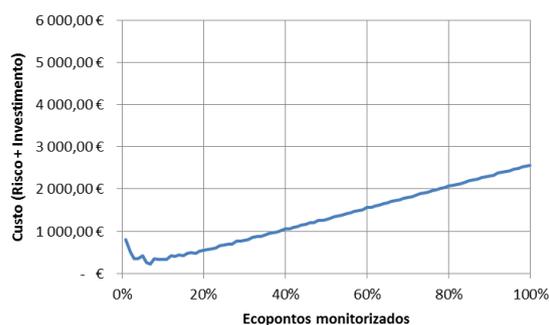


Figura A.2 – Custo mensal da solução Cluster 2 Método hierárquico.

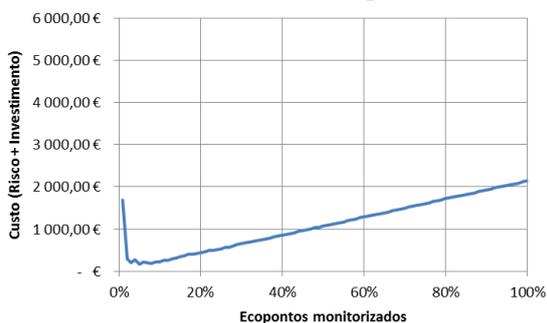


Figura A.3 – Custo mensal da solução Cluster 3 Método hierárquico.

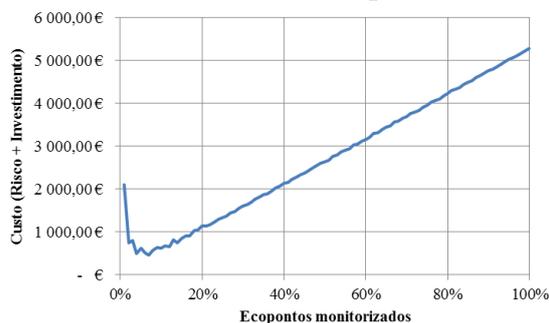


Figura A.4 – Custo mensal da solução Cluster 4 Método hierárquico.

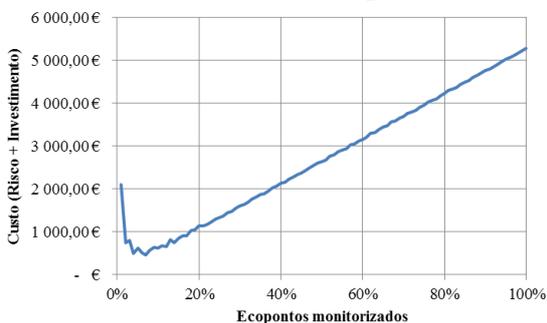


Figura A.5 – Custo mensal da solução Cluster 5 Método hierárquico.

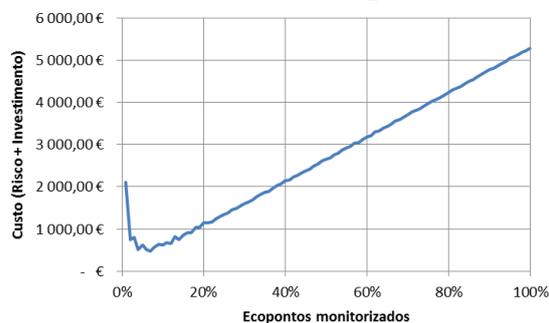


Figura A.5 – Custo mensal da solução Cluster 5 Método hierárquico.

As figuras A.7 a A.12 representam os gráficos de custo mensal da solução de cada *cluster* para o método k-médias.

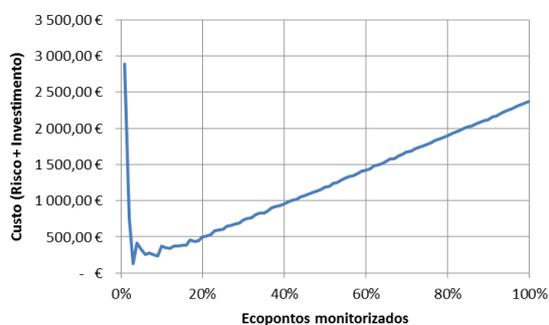


Figura A.7 – Custo mensal da solução Cluster 1 Método das k-médias.

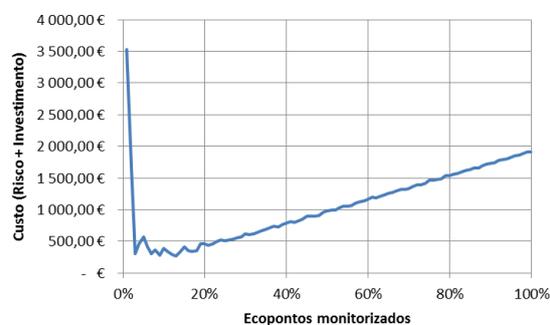


Figura A.8 – Custo mensal da solução Cluster 3 Método das k-médias.

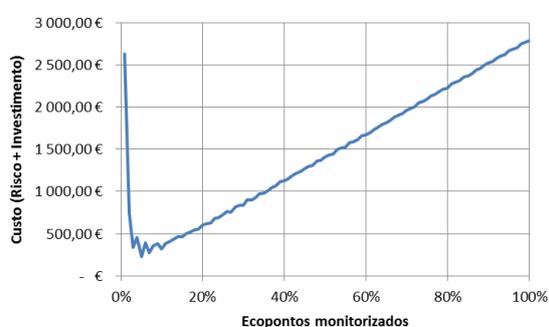


Figura A.9 – Custo mensal da solução Cluster 4 Método das k-médias.

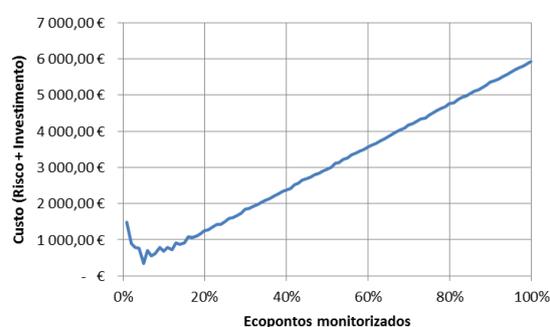


Figura A.10 – Custo mensal da solução Cluster 5 Método das k-médias.

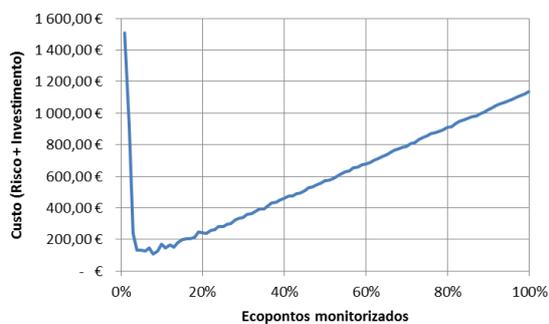


Figura A.11 – Custo mensal da solução Cluster 6 Método das k-médias.

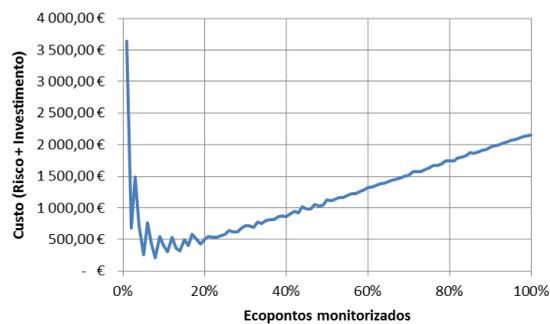


Figura A.12 – Custo mensal da solução Cluster 7 Método das k-médias.

ANEXO A2– ANÁLISE DE SENSIBILIDADE

A tabela A.1 apresenta os valores relativos à análise de sensibilidade de cada *cluster* para o método hierárquico utilizando o critério de Ward para vários valores de distância entre ecopontos. Essa distância foi definida através de um fator multiplicado pela distância considerada.

Tabela A.1 – Análise de sensibilidade para os clusters do Método hierárquico

Solução de monitorização obtida face à população total do <i>cluster</i>	Fator de distância				
	1/3	1/2	1	2	3
<i>Cluster 1</i>	4%	4%	7%	7%	7%
<i>Cluster 2</i>	3%	7%	7%	7%	7%
<i>Cluster 3</i>	3%	5%	5%	5%	5%
<i>Cluster 4</i>	4%	4%	7%	7%	7%
<i>Cluster 5</i>	4%	4%	7%	7%	7%
<i>Cluster 6</i>	4%	4%	7%	7%	7%
<i>Cluster 7</i>	0%	0%	0%	0%	0%
<i>Cluster 8</i>	0%	0%	0%	0%	0%

A tabela A.2 apresenta os valores relativos à análise de sensibilidade de cada *cluster* para o método k-médias para vários valores de distância entre ecopontos. Essa distância foi definida através de um fator multiplicado pela distância considerada.

Tabela A.2 – Análise de sensibilidade para os clusters do Método das k-médias

Solução de monitorização obtida face à população total do <i>cluster</i>	Fator de distância				
	1/3	1/2	1	2	3
<i>Cluster 1</i>	3%	3%	3%	3%	3%
<i>Cluster 2</i>	0%	0%	0%	0%	0%
<i>Cluster 3</i>	3%	3%	13%	13%	13%
<i>Cluster 4</i>	5%	5%	5%	5%	10%
<i>Cluster 5</i>	5%	5%	5%	5%	5%
<i>Cluster 6</i>	4%	4%	8%	8%	8%
<i>Cluster 7</i>	5%	5%	8%	8%	8%
<i>Cluster 8</i>	0%	0%	0%	0%	0%

ANEXO A3 – SIMULAÇÃO DE MONTE CARLO

As figuras A.13 a A.18 representam da simulação de Monte Carlo, apresentando o custo da solução em função da confiança, para cada *cluster* para o método hierárquico utilizando o critério de Ward.

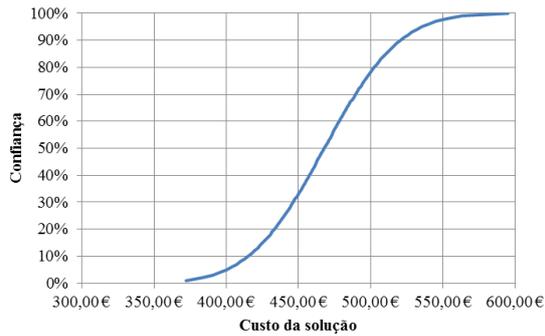


Figura A.13 – Simulação de Monte Carlo para o Cluster 1 Método hierárquico.

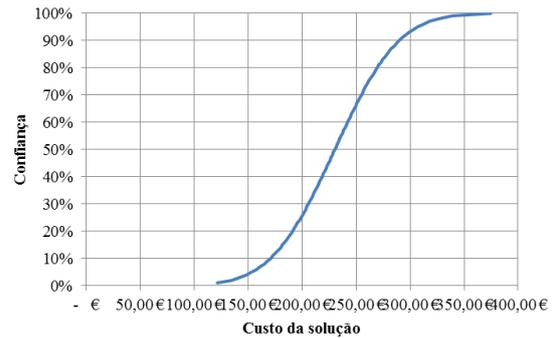


Figura A.14 – Simulação de Monte Carlo para o Cluster 2 Método hierárquico.

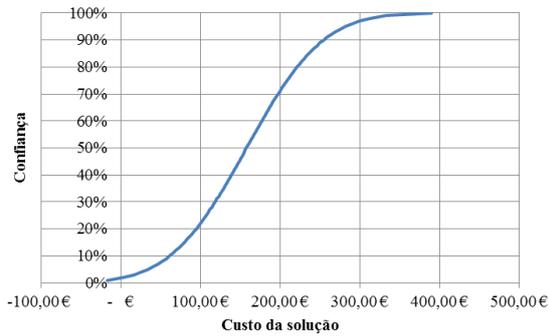


Figura A.15 – Simulação de Monte Carlo para o Cluster 3 Método hierárquico.

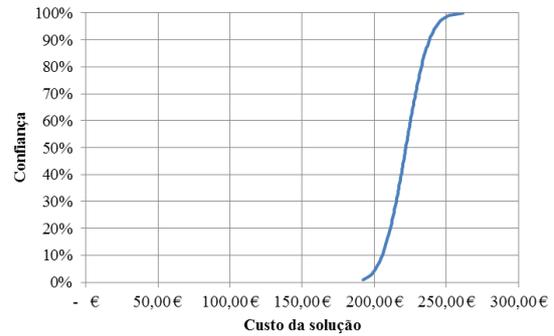


Figura A.16 – Simulação de Monte Carlo para o Cluster 4 Método hierárquico.

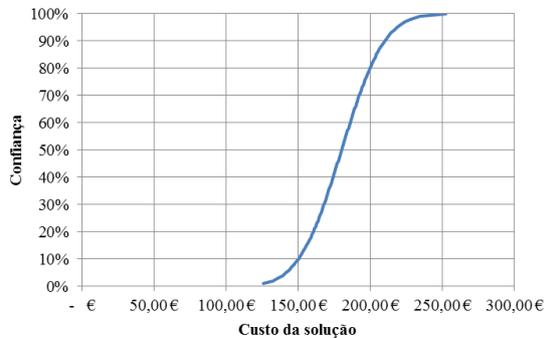


Figura A.17 – Simulação de Monte Carlo para o Cluster 5 Método hierárquico.

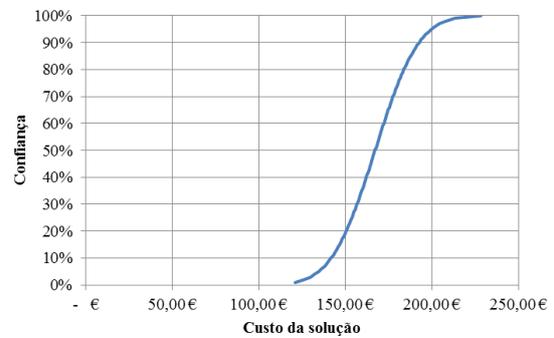


Figura A.18 – Simulação de Monte Carlo para o Cluster 6 Método hierárquico.

As figuras A.19 a A.24 representam da simulação de Monte Carlo, apresentando o custo da solução em função da confiança, para cada *cluster* para o método das *k*-médias..

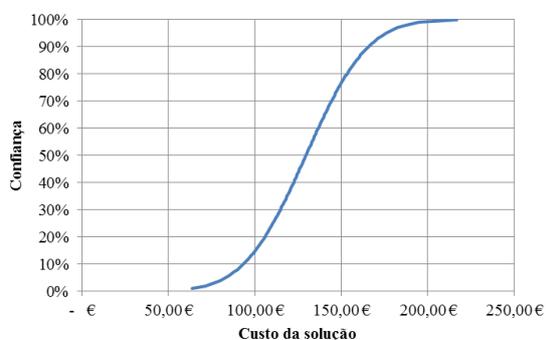


Figura A.19 – Simulação de Monte Carlo para o Cluster 1 Método das k-médias.

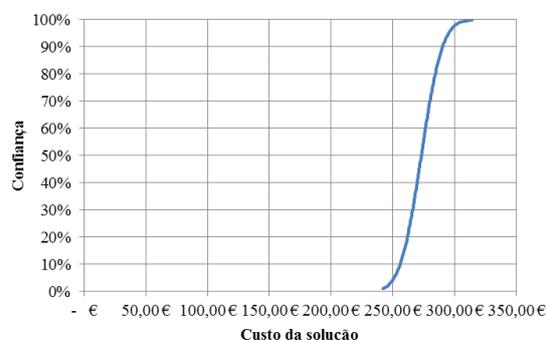


Figura A.20 – Simulação de Monte Carlo para o Cluster 3 Método das k-médias.

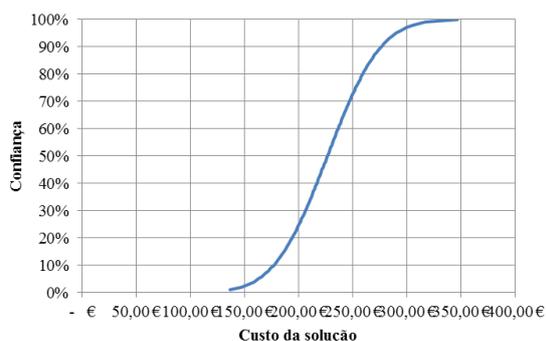


Figura A.21 – Simulação de Monte Carlo para o Cluster 4 Método das k-médias.

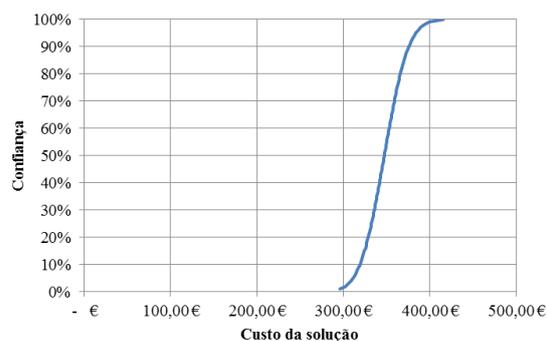


Figura A.22 – Simulação de Monte Carlo para o Cluster 5 Método das k-médias.

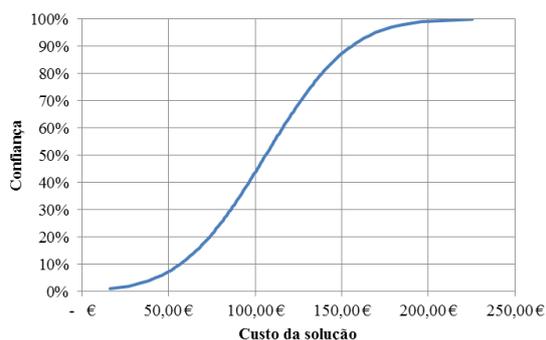


Figura A.23 – Simulação de Monte Carlo para o Cluster 6 Método das k-médias.

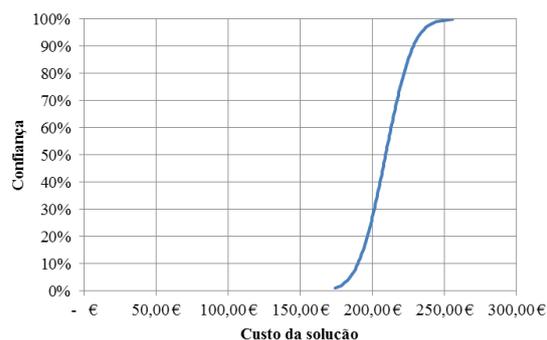


Figura A.24 – Simulação de Monte Carlo para o Cluster 7 Método das k-médias.