Methods for Sensor Placement in Waste Management

Luís Manuel Martins Rodrigues

Thesis to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisors:
Prof. Tânia Rodrigues Pereira Ramos
Prof. Manuel Cabido Lopes

Examination Committee

Chairperson: Prof. José Rui De Matos Figueira
Supervisor: Prof. Tânia Rodrigues Pereira Ramos
Member of the Committee: Prof. Tânia Rute Xavier de Matos Pinto-Varela

April 2020

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Resumo

A presente dissertação, desenvolvida no âmbito do Mestrado em Engenharia e Gestão Industrial, pretende estudar uma integração economicamente sustentável das tecnologias de informação e comunicação nas operações de recolha de resíduos sólidos urbanos. Mais concretamente, é estudada a utilização de sensores de ultrassom que, quando instalados dentro dos contentores, são capazes de medir o seu nível de enchimento. Atualmente, as rotas de recolha são bastante ineficientes, pois são rotas fixas que não dependem dos níveis reais de resíduos. O uso desta tecnologia fornecerá informação que possibilita a introdução de rotas dinâmicas e mais eficientes, ao mesmo tempo que reduz o número de transbordos dos contentores.

Contudo, os custos de implementar esta tecnologia ainda são muito elevados. Nesta dissertação serão estudados critérios e estratégias que permitam identificar quais os contentores, de uma determinada população, em que a sua monitorização seja mais vantajosa. Assim, o principal objetivo desta dissertação é estudar a aplicação de diferentes métodos de seleção de um reduzido número de contentores para serem monitorizados, procurando-se reduzir os custos de investimento associados ao uso de sensores, ao mesmo tempo que se mantém os benefícios desta tecnologia.

Para tal, são analisadas as operações de recolha do material papel da ERSUC – uma empresa que atua na região litoral centro de Portugal – no município de Soure, donde se concluiu que monitorizar apenas uma fração dos contentores pode permitir reduções nos custos de 11% quando comparado com toda a população ser monitorizada e de 21% quando comparado com o cenário atual.

Palavras-Chave: Gestão de Resíduos, Recolha de Resíduos, Resíduos Sólidos Urbanos, Sensores, Rotas Dinâmicas, Monitorização Remota.

Abstract

The present dissertation, developed within the Industrial Engineering and Management Master's program, intends to study an economically sustainable incorporation of information and communication technologies on the collection operations of solid wastes. More precisely, it's studied the implementation of ultrasonic sensors that, when installed inside the containers, are capable of measuring its filling level. The current collection routes are quite inefficient, since they are fixed routes that don't depend on the real waste levels. The use of this technology will provide information that enables the introduction of dynamic and more efficient collection routes, while the number of overfull containers is also reduced.

However, the costs of implementing such technology are still very high. In this dissertation are studied criteria and strategies which may allow to identify what are the containers, of a certain population, in which their monitorization is more beneficial. Therefore, the main objective of this dissertation is to study the application of different methods for selecting a reduced number of containers to be monitored, in order to reduce the investment costs associated to the use of sensors, while maintaining the benefits of using this technology.

For this purpose, it is analysed the collection operations of ERSUC – company that operates in the central west coast of Portugal –, in the municipality of Soure, from where it was concluded that monitoring only a fraction of the containers may allow a cost reduction of 11% when compared to all containers being monitored and of 21% when compared to the current situation.

Keywords: Waste Management, Waste Collection, Municipal Solid Waste, Sensors, Dynamic Routing, Remote Monitoring.

Agradecimentos

Gostaria de agradecer aos meus orientadores, a Professora Tânia Ramos e o Professor Manuel Lopes, por toda a ajuda prestada ao longo da execução deste trabalho. Agradecer à Professora Tânia que esteve comigo desde o início e sempre mostrou imensa disponibilidade e nunca deixou de acreditar em mim. Agradecer ao Professor Manuel, que a partir do momento em que se juntou à tese, contribuiu muito com as suas intervenções perspicazes e conhecimento sobre o problema.

Agradeço também à minha família, especialmente aos meus pais, por todo o suporte não só ao longo desta tese, mas também do meu percurso académico.

Agradeço aos meus amigos que de uma maneira ou de outra me ajudaram ou com quem pude ir falando sobre a tese.

E, por fim, quero agradecer à minha namorada Marlene, que foi a pessoa que mais me apoiou durante estes tempos e sem a qual não sei se teria conseguido encontrar forças para terminar este trabalho.

Table of Contents

Resumo ii

Abstract iii

Agradecimentos iv

Table of Contents v

List of Figures viii

List of Tables x

List of Acronyms xi

1. Introduction 1
    1.1. Context and Motivation 1
    1.2. Dissertation’s Objectives 2
    1.3. Dissertation’s Methodology 3
    1.4. Dissertation’s Structure 3

2. The Waste Management Sector 5
    2.1. An overview of solid waste management 5
        2.1.1. A historical context of solid waste management 5
        2.1.2. Shift towards a sustainable solid waste management 6
    2.2. Municipal solid waste management 8
    2.3. Municipal solid waste management in Portugal 10
    2.4. Waste Collection 12
    2.5. Problem Characterization 14
        2.5.1. Integration of ICTs in Waste Collection 14
        2.5.2. Identification of the problem 16
    2.6. Chapter’s Conclusions 18

3. Literature Review 20
    3.1. Aspects and practices affecting waste collection 20
        3.1.1. Aspects regarding waste demand characteristics 20
        3.1.2. Spatial aspects regarding waste collection points 22
        3.1.3. Dynamic routing and scheduling practices to integrate with remote monitoring 23
        3.1.4. Other criteria – high priority bins 25
        3.1.5. Approaches for selecting containers to be monitored 26

3.1.5.1. Review of Correia's work 26

3.1.5.2. Oxfam's demonstration 27

3.2. Sensor placement problems 27

3.3. Routing problems 29

3.3.1. Travelling Salesman Problem 29

3.3.2. Vehicle Routing Problem 30

3.4. Chapter's Conclusions 31

4. Methods for sensor placement 33

4.1. Proposed Methods 33

4.1.1 Relevant criteria found in the literature 33

4.1.2. Method 1 – Waste demand rate of a container 34

4.1.3. Method 2 – Waste demand variability of a container 34

4.1.4. Method 3 – Remoteness of a container 35

4.1.5. Method 4 – Cost Reduction Algorithm 36

4.1.4.1. Degree of how much "intermediate" a container is 36

4.1.4.2. Description of the Cost Reduction Algorithm 37

4.1.4.3. Cost Reduction Algorithm – Variant with collection frequencies 40

4.1.5. Method 5 – Considering rate, variability and remoteness 41

4.2. Description of the validation methodology 42

4.2.1. Overview of the validation methodology 43

4.2.2. The dynamic collection policy 45

5. Study of a real-case scenario 48

5.1. Waste collection in Soure 48

5.1.1. Brief Presentation of ERSUC and Soure municipality 48

5.1.2. Current waste collection policy 50

5.2. Study of the individual demand 52

5.2.1. The gathering of each container's data 52

5.2.2. The treatment and analysis of the data 53

5.2.3. Statistical modelling of the waste demand 56

5.2.3.1. Selecting the probability distribution 56

5.2.3.2. The gamma distribution 57

5.2.3.3. Determining the parameters and modelling the waste demand 58

5.3. Study of the road network 59

6. Results...61

6.1. Values and assumptions used 61

6.2. Study of the baseline scenarios 63

6.2.1. Presentation of each scenario 63

6.2.2. Analysis of the baseline scenarios 64

6.3. Study of the partially monitoring scenario 66

6.3.1. Method 1 – Waste demand rate of a container 67

6.3.2. Method 2 – Waste demand variability of a container 69

6.3.3. Method 3 – Remoteness of a container 70

6.3.4. Method 4 – Cost Reduction Algorithm... 72

6.3.5. Method 5 – Considering rate, variability and remoteness 73

6.3.6. Overall conclusions and tendencies 74

6.4. Financial Analysis..75

7. Conclusions and Further Work 78

Bibliography 81

Appendix..89

Appendix A – Current routes 89

Appendix B – Waste demand: average and variability..91

Appendix C – Histogram vs PDF (2 examples more) 93

Appendix D – Threshold level and collection interval..94

Appendix E – The selection of containers 96

List of Figures

Figure 1: The Waste Hierarchy Principle pyramid (adapted from Recycle More, 2019). 7

Figure 2: Generic representation of the MSWM stages (adapted from Martínez et al., 2012). 9

Figure 3: Rates of preparation for reuse and recycling, in Portugal (adapted from APA, 2019d)........ 11

Figure 4: Evolution of separate collection, in Portugal. Data from PORDATA, 2019........................... 12

Figure 5: Waste collection routes with (b) and without (a) the ultrasonic sensor technology (adapted from Gonçalves, 2014). 16

Figure 6: Collection of only almost full containers (a) and of almost full and nearby containers with lower filling levels (b) (adapted from Nourinejad et al. 2018).. 24

Figure 7: Graph-based representation of a TSP with 10 nodes (adapted from Rivlin, 2019). 30

Figure 8: TSP vs VRP route patterns (source: Liu et al., 2014). 31

Figure 9: Selection of containers based on a single criterion.. 33

Figure 10: Representation of the criteria related to the waste demand. 34

Figure 11: Context for the criterion remoteness of container. 36

Figure 12: Representation of the criterion related to “intermediate” containers................................. 37

Figure 13: Flowchart of the cost reduction algorithm. 38

Figure 14: Graph of the small scenario example for the cost reduction algorithm.............................. 39

Figure 15: Three-dimensional classification of containers (adapted from Zhang et al., 2012). 42

Figure 16: Diagram of the validation methodology... 43

Figure 17: Flowchart of the dynamic collection policy algorithm. 46

Figure 18: Network of paper-carboard containers in Soure. 49

Figure 19: The paper-cardboard container (adapted from Vila Nova de Poiares, 2017)..................... 51

Figure 20: Negative and positive skewness of a probability distribution (adapted from ION Data Services, 2020).. 57

Figure 21: Probability density functions for different parameters k and θ (source: Wikipedia, 2020).. 58

Figure 22: PDF and Histogram (7 intervals with *w* = 4) of container C1. 59

Figure 23: Example of a small road network and its graph. 60

Figure 24: Graph of the Network of 98 Containers of Soure. 60

Figure 25: Scatter plot between costs and overflows of the baseline scenarios................................. 66

Figure 26: First 14 locations selected with method 1. 67

Figure 27: Overflows vs costs with a selection based on the waste demand average rates (method 1). 68

Figure 28: Overflows vs costs with a selection based on the waste demand variability (method 2). .. 69

Figure 29: Scatter plot between the coefficient of variation and the average rates of the containers. 70

Figure 30: Overflows vs costs with the two options based on remoteness method (method 3). 71

Figure 31: Overflows vs costs of a random selection... 71

Figure 32: First 14 locations selected with the Cost Reduction Algorithm (without considering frequencies)... 72

Figure 33: Overflows vs costs with a selection based on the cost reduction algorithm. 73

Figure 34: Results for the containers with high variability and high average rates. 74
Figure 35: Total costs, in 10 years, for a low overflow cost. 76
Figure 36: Total costs, in 10 years, for a medium overflow cost. 76

List of Tables

Table 1: Business models and costs practiced by sensor manufacturers. 18

Table 2: The values of the path costs for the different container locations. 39

Table 3: Small example with different frequencies of visit. 41

Table 4: The values of the cost reductions, considering frequency of visit. 41

Table 5: Steps of the dynamic collection policy algorithm. 46

Table 6: The collection of the 98 paper-cardboard containers, in 2017. 50

Table 7: Estimated distance, costs and collection times of routes A and B. 51

Table 8: Raw data of the waste-levels (%) measurements – example 54

Table 9: Estimated waste-levels (%) for all the days of the period – example 55

Table 10: Estimated daily waste demand rates, average and standard deviation (%) – example....... 55

Table 11: Daily waste demand rates - average, max and min values of the mean and coefficient of variation 55

Table 12: Frequency of each waste level at collection, in the 2019's studied period. 56

Table 13: Results of simulating the baseline scenarios. 64

List of Acronyms

APA – Portuguese Agency of Environment (*Agência Portuguesa do Ambiente*)

CVRP – Capacitated Vehicle Routing Problem

EPR – Extended Producer Reliability

EU – European Union

GDP – Gross Domestic Product

GIS - Geographic Information System

GPRS – General Packet Radio Service

GPS – Global Positioning System

ICT – Information and Communication Technology

IoT – Internet of Things

MSW – Municipal Solid Waste

MSWM – Municipal Solid Waste Management

PERSU – Portuguese Strategic Plan for Municipal Solid Waste (*Plano Estratégico dos Resíduos Sólidos Urbanos*)

PDF – Probability Density Function

PPP – Polluter Pays Principle

PVRP – Periodic Vehicle Routing Problem

SGRU – Municipal Solid Waste Management System (*Sistema de Gestão de Resíduos Urbanos*)

SVRP – Stochastic Vehicle Routing Problem

RFID – Radio-Frequency Identification

SC – Smart City

TSP – Travelling Salesman Problem

UK – United Kingdom

VRP – Vehicle Routing Problem

VRPTW – Vehicle Routing Problem with Time Windows

WCP – Waste Collection Problem

WEEE – Waste from Electric and Electronic Equipment

WHP – Waste Hierarchy Principle

WSN – Wireless Sensor Network

1. Introduction

1.1. Context and Motivation

Waste is part of our everyday lives. The production of waste as a side result of human activities is unavoidable, and long gone are the times when the waste produced by us could be recycled by Earth's natural processes. This situation affects directly the quality of life and health of the populations all over the world, as well the environment and, therefore, appears the need to manage these undesirable by-products, i.e., societies need to have waste management (The Environmental Literacy Council, 2015).

On top of that, not only the world's population is increasing, but it's also growing in terms of prosperity, urbanization and of technological developments. These factors are linked to an increasing consumption of goods, which means that, per capita, more natural resources are being used and more quantity of waste is being discarded (Kaza et al., 2018; Vergara & Tchobanoglous, 2012). Statistics according to Kaza et al. (2018) reveal that the world generates 2,01 billion tonnes of municipal solid waste per year, and that value is expected to grow to 3,40 billion tonnes as the year of 2050 arrives.

Municipal solid waste management (MSWM) has become incredibly complex and a centrepiece in the quest for building a more sustainable, modern, healthy and inclusive world. As a result, more than ever, is critical that cities and communities improve their solid waste management systems, being of utter importance to consider and to integrate waste management principles in the models of economic growth and innovation. It's fundamental that governments and entities find smarter and more sustainable ways to manage waste, promoting an efficient economic growth while minimizing the environmental impacts (Kaza et al., 2018; Vergara & Tchobanoglous, 2012).

One of the major problems of MSWM today is the lack of efficiency on the collection of solid wastes and associated transportation. Being the collection the costliest step in solid waste management, companies have been trying to find ways for tackling these inefficiencies with the integration of new information and communication technologies (ICTs). One example is the use of sensors that allow the measurement of the amount of waste existing inside a container in real time. By monitoring the real levels of waste, companies can develop models capable of responding to changes in the demand for collection and, consequently, they can optimize the collection routes and reduce the number of unnecessary pickups. Thus, with this technology, companies can have savings in fuel consumption and make a more proper use of their trucks, while mitigating the emissions of pollutant gases associated with the transportation (Anagnostopoulos et al., 2015; Gonçalves, 2014; Ion & Gheorghe, 2014; Longhi et al., 2012; Ramos et al., 2018).

The development of algorithms that permit to define dynamic collection routes, that change according to the needs of any given moment, it's a topic that has been receiving attention by the academic community, being already several the approaches proposed that confirm the ICTs' potential. Gonçalves (2014) suggested a technological system for the planning and management of truck collection routes and test it in a simulation environment, where it was possible to see the reduction on the number of routes performed, the total distance travelled, the number of containers visited, and the amount of hours worked by the employees. However, the author verified that the costs of installing the ultrasonic sensors in all containers were very high and wouldn't overcome the benefits of having them.

The present dissertation comes as a result of Gonçalves' work and his conclusions. Waste collection systems are complex ones that involve thousands of containers. Even if this type of technologies allows us to better understand these systems and to gain insights from the data collected, they are still very pricey, and it is necessary to develop financially more sustainable strategies for integrating them into the waste collection activities.

For that, it will be analysed in this work in which circumstances makes sense to have a sensor embedded in a container to measure its filling level, and a set of criteria, that takes into account the containers' characteristics and their relevance to the problem, will be identified. Therefore, several methods for selecting only a reduced number of containers to be monitored will be proposed. It will be studied how these methods allow reducing the investment costs associated with the use of the ultrasonic sensor technology (by monitoring only a reduced number of containers), while trying to maintain the benefits inherent to its use. Thus, a trade-off analysis between the cost and benefits of monitoring all the containers, or only the containers proposed by each method, will be done in order to comprehend the feasibility of the proposed solutions.

In order to perform this study, it will be studied the real-case scenario of the paper/cardboard collection operations, in the municipality of Soure. These operations are carried out by ERSUC, which is a waste collection company that operates in the central west coast of Portugal.

1.2. Dissertation's Objectives

The main objective of the present dissertation is to develop methods for selecting a reduced number of containers to be monitored, in a given collection area. With this, it is intended to evaluate and compare how the different methods effect the day-to-day collection operations, under a dynamic collection policy. This must be done by studying a real-case scenario, i.e., by analysing the individual characteristics of real containers.

It's also essential to perform a financial analysis, to understand if the proposed methods can help reducing the overall costs when compared to the current waste collection policy and to the scenario where all the containers are monitored. Finally, recommendations for the company will be devised.

In order to achieve these main objectives, the following secondary objectives are also contemplated:

- Contextualize the solid waste management sector;
- Present the potential benefits of integrating ICTs and, more specifically, the ultrasonic sensor technology on waste collection activities;
- Identify and characterize the problem of this work;
- Review of previous studies that address containers' characteristics that may be worth considering when installing sensors in a limited number of containers;
- Gather insights about traditional sensor placement problems;
- Introduce two routing problems, such as the TSP and the VRP, to better understand the context of optimizing waste collection routes, in order to gain knowledge that can be helpful when deciding how to strategically place the sensors;
- Identify the gap in the literature that supports the development of the present work;

1.3. Dissertation’s Methodology

In order to achieve the objectives mentioned, it was adopted the following methodology:

Contextualization of the solid waste management sector → The *modus operandi* was to start from the most-wide ranging topic and then start narrowing into the more specific subject matters that led to the identification of the problem under consideration. The basis for the investigation was both web research and scientific databases like Science Direct, Google Scholars and Research Gate.

Literature review → In this stage, it is performed a literature review on the aspects, practices and other topics worth considering when installing sensors in containers. The research was done by exploring the previously mentioned scientific databases. Some of the keywords used were: Waste Collection, Route Optimization, Sensor, Remote Monitoring, Dynamic Routing, Sensor Placement, Routing Problem.

Development of the sensor placement methods → In this stage, the methods for selecting the containers are developed, taking into account the ideas and criteria extracted from the literature review.

Development of a dynamic collection policy→ This stage consists on developing a dynamic collection policy that will be used to validate the methods proposed by simulating the day-to-day waste collection operations of each one.

Data collection → In this stage, real data from ERSUC regarding each individual container will be collected and analysed. These data are used to describe the current situation, but also to statistically model the waste demand, which is necessary for simulating the day-to-day collection operations.

Testing the proposed methods with the real-case scenario data → The methods for selecting the containers are applied to the real-case scenario (its data were used as inputs) and, subsequently, the several different selections are run in the simulation environment, in order to obtain the results.

1.4. Dissertation’s Structure

The present work respects the following structure:

1) **Introduction:** A brief explanation is given about the context and reasons that motivated the study of the problem in question. The objectives of the dissertation are outlined, as well the methodology to achieve them.
2) **The Waste Management Sector:** The solid waste management sector will be contextualized. It will be given a special focus to the waste collection activities and to the new technologies, such as ultrasonic sensors, that are appearing to change them. The last part of this chapter will expose and characterize the problem under consideration.
3) **Literature Review:** In this chapter, several studies related to the problem of this work will be covered and analysed. Examples are studies that address the use of sensors to monitor containers, dynamic routing practices, waste demand characteristics, sensor placement problems or node-routing problems such as the TSP and the VRP.
4) **Methods for Sensor Placement:** In this chapter, the several methods used for selecting a reduced subset of containers in a given collection area will be presented

and described. It will also be explained the validation methodology that will be used to evaluate the different methods in a simulation environment.

5) **Study of a real-case scenario:** The real-world scenario that will be used to validate the methods previously developed will be presented. The current practices undertaken will be studied and it will be explained how the individual data regarding each container were obtained and processed.
6) **Presentation of the results**: In this part of the dissertation, it is shown and analysed the results of the several proposed methods. A financial analysis that takes into account the investment costs necessary with each method is also carried out.
7) **Conclusions and Future Work:** The final conclusions of the present dissertation are outlined and recommendations concerning the future work are given.

2. The Waste Management Sector

This chapter has the objective of contextualizing the solid waste management sector and of explaining the problem that will be approached in this dissertation. Firstly, an overview of solid waste management will be given as a way to introduce the topic, with some background and its main principles being provided (section 2.1). From here, the focus is turned to the MSWM (section 2.2) and, in particular, its presence in Portugal (section 2.3). Following this, the current panorama regarding the waste collection will be looked in more detail (section 2.4), being acknowledged the immense importance of waste collection on the overall waste management system, and how and why it suffers from a poor usage of its resources. Next, in section 2.5, the problem under consideration is characterized. Here, the integration of ICTs and, more precisely, of ultrasonic sensors as a way to better the efficiency of waste collection is discussed. Given the fact these technologies are still a high-end solution for most waste collection companies, it is emphasized the work and progress that still needs to be done and how that led to the identification of the problem which inspired this work. Finally, in section 2.6, the main conclusions of the presented chapter will be addressed.

2.1. An overview of solid waste management

Nowadays, waste is considered to be a material that its owner intends to discard and solid waste management is the set of activities concerned with the production, handling, storage, collection, treatment, recycling and disposal of waste, as well with the improvement, planning and regulation of these mechanisms, in order to reduce the risk for social and environmental systems (Oxford Unity Press, 2013). Every consumption or production process that occurs originates waste as consequence, and to avoid pollution problems that may lead to public health or environmental issues, it is necessary to manage it, reduce it and control its impacts (Chang & Pires, 2015).

Waste management is strictly correlated with the evolution of a technological society and it has become one of the most serious problems of modern era. The benefits of mass production allowed for people to search for a better life and a higher standard of living, which history shows that goes hand in hand with an increase in the consumption of goods and generation of waste (Tchobanoglous & Kreith, 2002). Solid waste management plays a crucial role for decreasing the harmful impacts of this generated waste, so that it supports economic development and a better quality of life for societies.

2.1.1. A historical context of solid waste management

Before societies, waste generation was negligible. Only when humans started to gather in cities, waste started to pile (Wilson, 1976). At first, it was a problem of dirtiness and hygiene, and during centuries there were almost no changes on how waste was being managed which, with the growing degrees of urbanization, brought several public health problems (Schott, 2014).

When the 18th and 19th centuries arrived, the Industrial Revolution started to take place. If until Industrial Revolution almost all production processes were based on manual work, after it, steam powered machines were introduced into the manufacturing processes, increasing their production rates and efficiency and reducing the production costs like never seen before, kicking off an exponential rising trend in the world’s economy (Clark, 2014). This increase on production, inevitably, provoked an increase on waste generation, which until then remained relatively small (Marten, 2003).

In the 20th century, the environment had become a repository for solid waste (Barles, 2014). Public health turned a trending topic and, for solid waste management, the only concern was to take the waste out of cities, wherefore, as cities became gradually cleaner, the peripheries were getting dirtier (Schott, 2014).

Even though municipalities gradually started to provide waste services, due to several factors, the threat to the ecosystem was highly exacerbated. The improvements in public health had drastically reduced death rates in industrializing worlds and populations grew very rapidly (Marten, 2003; Wilson, 2007). However, not only societies were growing, but also their living standards were. As the new technological and mechanical developments were being introduced to the masses, their consumption habits were changing and requirements for improved shelter, comfort, clothing and food were taking place at the expense of the environment and the Earth’s resources (National Academy of Sciences et al., 1993). Added to this amount of waste produced by households there was also an excess production of waste due to industry and agriculture. Waste had been accepted as a necessary evil of development and quality of life (Barles, 2014).

Only with the arrival of the 1960s and 1970s, with the arising of environmental concerns in most developed countries, the harmful effects of consumption, industrialization and development were brought to discussion, and solid waste management came onto political agenda. Finally, it was recognized that Earth’s resources couldn’t continue to be used uncontrollably, in order to preserve the world for future generations (Barles, 2014; Wilson, 2007). The environmental awareness raised in these years, which described solid waste as the “third pollution” (just after water and air pollutions), and the introduction of the sustainable development concept would trigger several reforms on the MSWM, changing this sector into as we know it (Barles, 2014).

2.1.2. Shift towards a sustainable solid waste management

In the United States and across Europe, several legislations concerning solid waste, like the Solid Waste Disposal Act in 1965, the Recovery Act in 1976 or the European directive of July 1975, were proposed, discussed and approved by the political bodies. All these texts took under consideration the importance to reduce the waste created at production, for example by reducing materials used in packaging, and to re-use the waste collected, through more sustainable options like recycling or energy recovery (Barles, 2014).

The concept of waste management has become intrinsically connected with the concept of sustainable development. It began to be widely accepted that, as the economy grows and societies continue to develop, humans need to manage resources in a way that allows for future generations to also thrive. For that, is fundamental that solid waste management ensures the sustainability of the environment through proper waste collection, recycling and resources conservation (Chang & Pires, 2015).

In general, new principles and rules were adopted across developed countries which led to the emergence of a wide number of new concepts for waste management. Some of these concepts are the

Waste Hierarchy Principle (WHP), where the different ways of dealing with waste are ranked by their order of preference according to the environmental impact of each one, being prevention the most preferable option and the disposal of waste without any value extraction from the waste the least preferable (see Figure 1) (DEFRA, 2011; D. C. Wilson, 2007); the Polluter Pays Principle (PPP), that was introduced as a way of regularizing the management and pricing of waste and in a very simplified way can be described like the more a consumer pollutes, the higher the price it has to pay, and it is based on the thesis that the polluter will reduce pollution costs as costs surpass the benefits of polluting (Barles, 2014; European Environment Agency, 2012); or for example the Extended Producer Reliability (EPR), that constitutes an application of the PPP, but in a more extensive way, since the concept of polluter is not exclusively assigned to the final consumer using the product, but rather it is extended to the entities which participate in the different steps of the product chain and can take a decisive role in avoiding pollution through more ecological practices (e.g. eco-design) (Chang & Pires, 2015; Leitão, 2017).

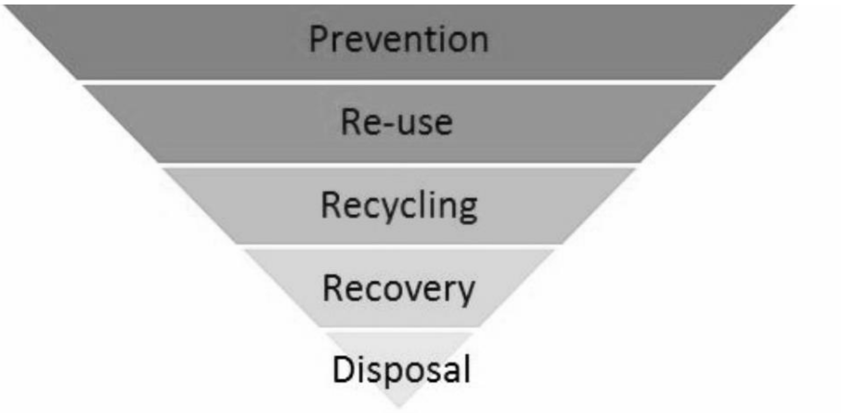


Figure 1: The Waste Hierarchy Principle pyramid (adapted from Recycle More, 2019).

These were important steps that became a reference, especially in developed countries, for achieving more integrated policies and for the establishing of statutory targets for environmental practices. They underlined that, in order to achieve a sustainable solution, the environmental, social, legal, political, institutional, economic and technical aspects must all be considered (Wilson, 2007). Waste management principles were particularly important to directly or indirectly force decision-makers, which lacked interest in environmental issues, to defend the adoption of good practices of waste management in the municipalities and in other solid waste management entities. One example is the case of Sociedade Ponto Verde, a non-profit organization, that due to the EPR policies, was formed by a group companies of the packaging sector that intended to avoid the high costs of operating individually and worked together to take responsibility for the collection, recycling and safe disposal of their products' resultant wastes (Leitão, 2017; Sociedade Ponto Verde, 2015).

2.2. Municipal solid waste management

According to its source, waste can be classified into MSW, which is the main focus of the present work, that regards the waste collected, generally by local government bodies, from households, schools, commercial establishments and from some small industrial operations. Commonly known as “trash” or “garbage”, it is the waste discarded by the “public” and it mainly consists on food wastes, garden and wood wastes, paper, plastics, as well some other inorganic wastes like glasses, metals and textiles (Regents of the University of Michigan, 2017; The Environmental Literacy Council, 2015). Besides MSW, waste can also be classified into medical waste; into industrial waste, which includes waste from, for example, manufacturing, construction, chemical or mining activities; or into “other wastes” that do not belong to any of the categories mentioned above (Chang & Pires, 2015).

MSW is growing in terms of quantity and complexity, which makes its sustainable development incredible challenging (Vergara & Tchobanoglous, 2012). The composition of MSW and its quantity are influenced by the income level of the families, their education (relates to their environmental concerns, for example), the spread of new products and technologies, and by cultural and geographic factors. At a local level, waste generation is also impacted by weather conditions, as well by seasonal festivities (Chang & Pires, 2015; Vergara & Tchobanoglous, 2012).

In a very generic way, and ignoring the backward flow of recovered and recycled materials, MSWM can be represented by the scheme of Figure 2. With the daily activities and consume of products, wastes are generated in the already mentioned sources, where they are handled (an initial separation is performed, and they are disposed into the containers) and next they are collected and transported by proper vehicles. After, and according to the type of waste and the type of collection (mixed or separate), waste will have different final destinations: they can go directly to the landfill where they are disposed, or they can pass firstly through treatment installations, where it is proceeded to their valorisation, being the recycling, material recovery, energic valorisation (e.g. incineration) and biologic treatment (e.g. composting and anaerobic digestion) some ways of possible treatment (Martínez et al., 2012).

Increasing population, rapid economic growth, development of urban areas, globalization and worldwide industrialization have made MSWM a universal problem. However, the challenges faced between nations can be significantly different, as we go from developed countries to developing countries and to lesser developed countries. The social, economic, technological and cultural conditions prevailing in developed countries are better suited for an easier, more efficient, more effective and broader employment of waste management strategies and policies, when in comparison to not developed countries, where these policies seem impractical and have often been resulting in less satisfactory results. Furthermore, the quantities of waste produced by a country are positively correlated to its economic development and, according to Kaza et al. (2018), even if developed countries only account for 16% of world’s population, they are responsible for 34% of the waste generated globally, a total of 683 million tonnes annually. Also, the composition of MSW also varies greatly from country to country and, usually, the waste materials in lesser developed and developing countries are rich in organic content which decreases the economic value of the recyclable materials. High-income countries have good financial investment on waste collection services (waste collection rates round up to 100% in these countries, whereas in lower income countries they are in between the 40-50%) and stakeholders are willing to cooperate in solutions for improving the delivery of these services (Kaza et al., 2018; Mmereki et al., 2016).

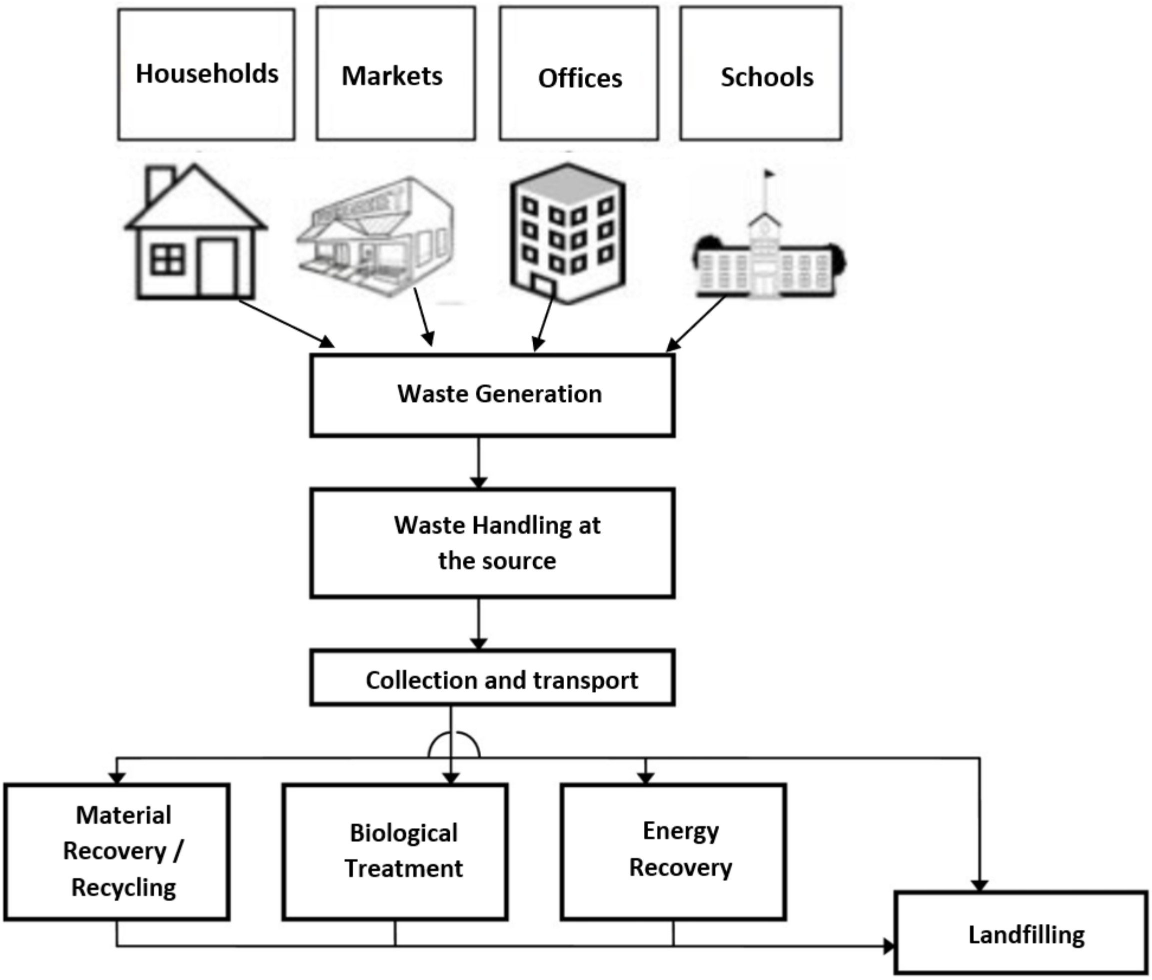


Figure 2: Generic representation of the MSWM stages (adapted from Martínez et al., 2012).

All these different aspects which continue to exist in developed countries and developing and lesser developed countries imply that solid waste management practices must differ among these countries. Vergara & Tchobanoglous (2012) claims that a greater diversity of successful MSWM systems is directly dependent on the recognition that the solutions must be local and contextual. Also, as stated by Kaza et al. (2018), it’s necessary to select the technological solutions taking into account the community and context where they will be applied, and often the best solution is neither the latest one or the most advanced technologically. Similarly, waste management practices also differ between urban and rural areas, or between industrial and residential areas (Pardini et al., 2019).

That being said, the management of MSW, across the world, and mostly in developed countries, is most commonly performed by municipalities in a decentralized manner or, when possible, these activities are performed on an intermunicipal scale to take advantage of economies of scale and reduce costs by saving on staffing needs, sharing investments and exchanging of know-how. Solid waste management programs are adapted to the local conditions such as financing, local norms and spatial layout of communities that often pay tariffs for these services, just like for water or electricity (Hoornweg & Bhada-Tata, 2012). MSWM can be seen as an opportunity for cities to create decent jobs, to raise awareness for a responsible consumption, standardize public services and to protect the natural environment (Fuss et al., 2018).

During the course of this work, when talking about solid waste management problems, challenges and solutions, the focus is on developed countries, capable of adopting the best practices and use state-of-the-art technologies, with special attention to the Portuguese situation.

2.3. Municipal solid waste management in Portugal

APA, the Portuguese Agency of Environment which is responsible for legislating the waste sector, defines MSW as any waste that comes from households, as well any other waste that, by its nature or composition, is identical to the waste that comes from households, with the particularity that large producers (daily production > 1100L) must be responsible for the management of their own waste (APA, 2019b)

Waste management, in Portugal, only became a priority concern in the 1990s. Until then, most waste (approximately 76%) was being disposed in more than 340 dumps around the country. Recycling was not very common and was mainly practiced in larger cities, for paper and glass only. During these years, several EU Directives guided the transformation of the Portuguese waste legislation, by defining management obligations for this sector. Thus, in 1997, the Portuguese Strategic Plan for MSWs (PERSU I) was approved. With this plan, appropriate treatment infrastructures were created, more adequate disposal was a main concern and all the operating dumps were shut down, the separate collection network for recycling was expanded and several multi-municipal and intermunicipal systems were created (Magrinho et al., 2006; Pássaro, 2003). Pássaro (2003) claims that, at the beginning of 2002, 100% of the MSW was being properly disposed and 70% of the Portuguese territory was being covered by separate collection programs for recycling.

As result of those first years of awakening to more sustainable practices in MSWM, the infrastructural and organizational situation in Portugal has advanced considerably, despite the increase on the total waste produced (Niza et al., 2014). The MSWM concerns addressed by PERSU I were maintained and updated with the implementation of PERSU II and PERSU 2020 plans which kept targeting the challenges and new requirements set by the EU, which have been becoming consistently more demanding with each new revision of the EU Directives (APA, 2019a; Niza et al., 2014). With the objective of decoupling the economic growth from the environmental impacts related to waste production, of promoting efficiency in the use of resources and of economically valuing recyclables, during the last 2 decades, Portugal has reduced significantly the waste that ends up on landfills, has increased the quantities of waste entering the facilities that prepare it for reuse or recycling (Figure 3) and, as can be seen in Figure 4, has expanded separate collection (APA, 2019d).

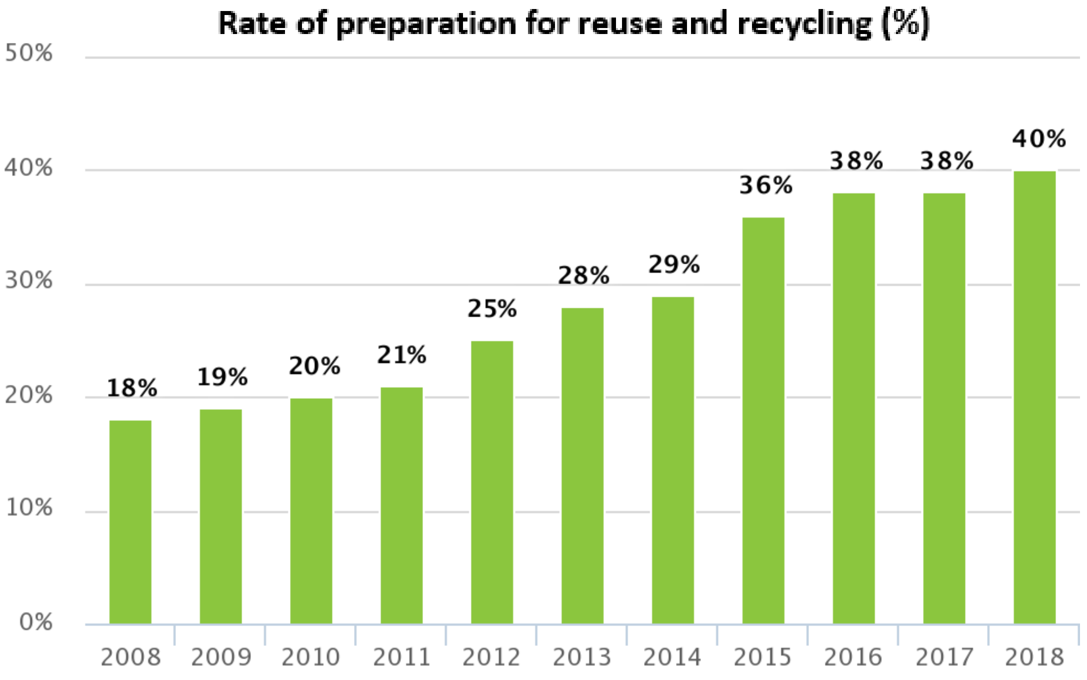


Figure 3: Rates of preparation for reuse and recycling, in Portugal (adapted from APA, 2019d)

Today, in Portugal, the MSW collection, transportation and disposal operations are the responsibility of 23 MSWM Systems (SGRUs in Portuguese terminology), which cover 100% of the Portuguese continental territory. Out of the 23 SGRUs, 12 are multi-municipal, meaning that the system involves several municipalities but, by decree-law, it is managed by a company of publicly-owned capital; and 11 are intermunicipal in which the several municipalities can delegate the management of the system to any company (APA, 2019c, 2019d; Sequeiros, 2012). These systems provide the human and logistical resources, as well the infrastructures and equipment necessary to ensure that the MSW produced in their areas is handled properly and delivered to the most appropriate destination (APA, 2019c, 2019d).

According to PORDATA (2019), in 2018, the average quantity of MSW collected, per capita, was of 507.8 kg, from which 103.5 kg came from a separate collection (glass, paper/cardboard and plastic/metal packaging), representing 20.4% of the total collection. The percentage of separate collection has been increasing during the course of the years (in 2001, this value only meant 4% of the total) which has led to a higher degree of complexity and higher costs associated to the collection and transportation activities, since separate collection requires new collection routes, more human resources, administration, investment in appropriate vehicles and containers and more fuel consumption (Teixeira et al., 2014). In order to ensure that all the population is covered by systematic, regulated and reliable waste collection and transportation, the MSWM systems could not escape to a substantial aggravation on the costs related to these services. Teixeira et al. (2014) concluded that, for the city of Porto, separate collection reveals a reduced overall performance as result of the operational, economic and environmental inefficiencies when in comparison to the mixed collection. According to the authors,

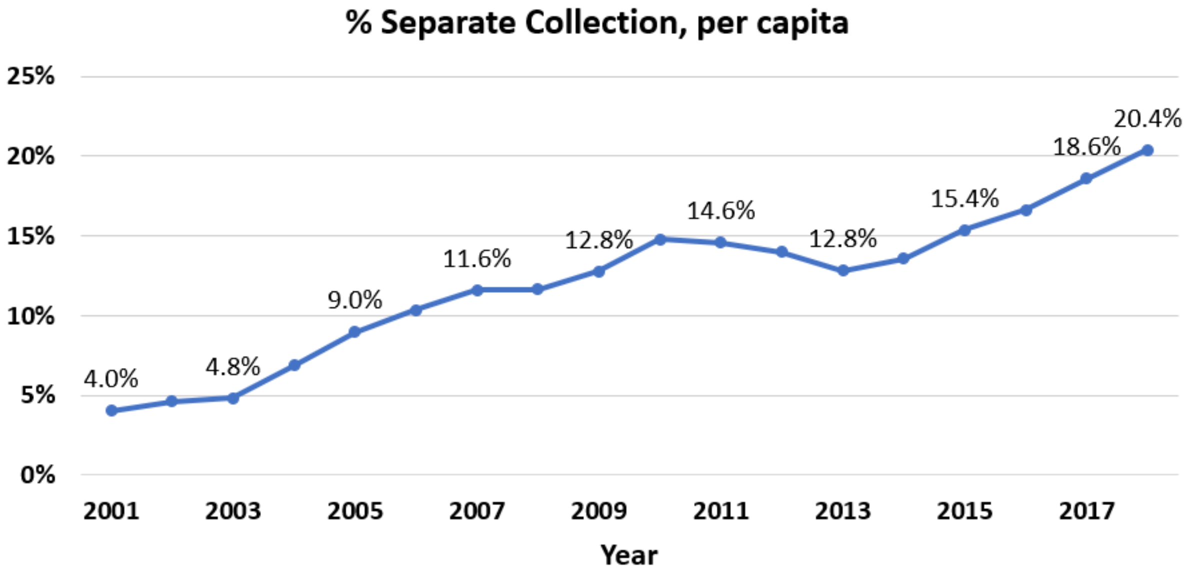


Figure 4: Evolution of separate collection, in Portugal. Data from PORDATA, 2019.

the fact that separate collection needs to be not only available but also accessible to the citizens despite the relatively low rate of separate collection causes these inefficiencies and leads to higher costs in this type of collection. Furthermore, focusing only on technological solutions such as the construction of a large number of infrastructures and equipment, which involved large investments, and an underestimation of policies that encourage the changing of attitudes and adoption of more sustainable social practices led to high and increasing costs of collection and transport (Niza et al., 2014). According to APA (2019d), the increase in separate collection is not proportional to the increase on the number of separate collection infrastructures and, therefore, instead of simply building new infrastructures, it is more and more urgent to upgrade the existing MSWM systems and their collection models.

2.4. Waste Collection

Waste collection is the face of MSWM, is where the waste generators (citizens) contact with the waste management system, and it is the process of removal of waste at source, preventing its accumulation where communities reside (Vergara & Tchobanoglous, 2012). Additionally, Christensen (2011) states that waste collection is the organizational interface that determines the success of such waste management system in accommodating the waste generated and avoiding uncontrolled dumping of waste. Without an effective waste collection, waste accumulates over time and problems like littering, overfull containers, contamination, odours and even flies may develop and become a public health and aesthetic problem to the community (Christensen, 2011; Pires et al., 2018).

The process of waste collection starts when the generated waste is thrown into appropriate receptacles, usually shared among the community, to be later picked up and emptied by collection vehicles. These vehicles are not only responsible for gathering the solid waste but they must also transport it to a processing facility, to a landfill disposal site or, more commonly, to a transfer station, where the collection vehicle is emptied and the waste is transferred to larger vehicles that will then transport these materials, usually through long distances, to the previously mentioned locations. (Christensen, 2011; Tchobanoglous & Kreith, 2002).

In high-income countries, waste collection operations have the most significant cost impact, usually ranging between 50 and 75% of the total costs of MSWM, as it involves massive labour, fuel consumption and maintenances (Christensen, 2011; Faccio et al., 2011; Kaza et al., 2018; Nguyen & Wilson, 2010; Pires et al., 2018; Tchobanoglous & Kreith, 2002). Teixeira et al. (2014), highlighted the extensive collection distances as one of the main threats to collection route efficiency. Furthermore, according to Faccio et al. (2011), the managerial complexity of waste collection is aggravated due to the difficult operational problems, the requirement of large investment costs (e.g. vehicle fleet) and the environmental problems associated such as emissions of pollutant gases, traffic congestions and noise. Johansson (2006) states that even if solid waste collection activities only account for a small part of the total freight transportations in a city (10-15%), due to the high number of stops and low average speed of this type of vehicles, the consequences on congestion, pollution and noise are higher than in comparison to other types of freight transportation.

It is well established among authors that waste collection is a critical and central part of any solid waste management program, but in practice this stage has often an underestimated role as Bilitewski et al. (2015) suggests. According to the authors, it is often neglected the fact that waste collection accounts for most of the total costs of MSWM, and thus any small improvement in this stage must be sought, considering that it would result in considerable savings in the overall system cost and in environmental benefits.

The collection efficiency is a tool for understanding the status of MSWM, and a proper bin collection system, an effective and efficient route planning and information about the collection schedule are prerequisites to achieve it (Jha et al., 2011; Mmereki et al., 2016). The structure of the collection system, the collection frequency; the type, number and location of the containers to be collected; the quantity collected per stop, material density, technologies used, route optimization efficiency, traffic and topography, the service level demanded, are all factors that influence the waste collection costs (Christensen, 2011; Pires et al., 2018; Teixeira et al., 2014). For Bilitewski et al. (2015), an efficient and optimal implementation and organization of waste collection must take into account the dimension of the collection area and its structural, economic and social settings, and legislative stipulations; the user demands; and the range of appropriate collection systems and technologies.

That being said, it is important to realize the fact that waste collection is subjected to such substantial costs is largely due to the lack of efficiency of its own operations. According to Faccio et al. (2011), optimizing the waste collection resources for an efficient management system is the most common struggle. Likewise, as Ramos et al. (2018) states, more than the intrinsic costs associated to transportation, waste collection companies suffer, each day, of an inefficient use of their resources, as their trucks repeatedly collect waste containers that are only partially full. This happens since, currently, waste collectors operate on the basis of static waste collection routes (“blind collection”), i.e., the routes which are performed are invariably the same ones, they are pre-defined routes that do not depend on the containers’ real filling levels. As the amount of MSW generated by the communities is highly variable and hard to predict, assumptions about the quantities of waste existing inside a bin are often incorrect. MSWM entities make use of forecasting software in order to try to guess those filling levels and, with this, traditionally, it’s defined when a specific route must occur and a fixed schedule for collection is obtained. This means that collections are being typically performed from each site on the same day and approximately at the same time each week, irrespective of how full the containers might be. Inevitably, container collection manoeuvres are often performed on the partially full or even empty containers, resulting in collection trucks that return to the collection center with loads well below their total capacity. These situations, when they occur, are obviously a source of wastage and of poor resources exploitation (Faccio et al., 2011; Gonçalves, 2014; Johansson, 2006; McLeod et al., 2013; Ramos et al., 2018).

Therefore, in order to improve the efficiency of the collection activities by optimizing collection routes and truck loads, reducing unnecessary pick-ups and adjust collection scheduling, it is needed to reduce the uncertainty associated to the bins' real filling levels. For this, the use of ICTs like ultrasonic sensors that signal, in real time, measurements about the level of waste existing inside the bin, is more and more seen as a possible solution to this problem. The integration of this technology enables the transition from a blind collection to a dynamic collection, where routes are taken according to the real needs of a given moment (Esmaeilian et al., 2018; Gonçalves, 2014; Ramos et al., 2018).

Finally, ensuring public health to citizens may be the main goal of waste collection, but this stage of MSWM encompasses today much more comprehensive duties. The ability of returning high quality recyclables, ready to be exploited, back to the industry is deeply enhanced by a separate collection that provides specific waste streams of source separated waste, which have better quality than mixed waste to be recycled and recovered (Pires et al., 2018; Vallero et al., 2019). However, as Pires et al. (2018) points out, and as seen in the previous subchapter, for the Portuguese situation, separate collection for different materials adds costs to the waste collection process and a trade-off between reduction of collection costs and higher amounts of waste for recycling and recovery has to be often balanced. The focus of this study will be on separate collection which, according to Teixeira et al. (2014), is particularly critical in areas of low population densities and of high dispersion (which constitutes great part of the Portuguese landscape). The authors highlight that separate collection is inefficient when in comparison to mixed collection, and as recycling targets aim higher, it is getting more and more stressed, leading to the necessity of having inefficient collection routes supressed. Furthermore, when containers fill-up quickly (which is the case of mixed collection) there is less need to monitor them since frequent fixed collections are fairly adequate (Straightsol, 2013). Hence, for separate collection, the integration of ICTs is particularly urgent and potentially more beneficial than for mixed collection.

2.5. Problem Characterization

2.5.1. Integration of ICTs in Waste Collection

By the year of 2050, circa 50 billion devices will be connected to the Internet (Pardini et al., 2019). The proliferation of the ICTs, not only in waste management but across many fields, has enabled a new era of Internet of Things (IoT) – which can be understood as the "interconnection via the Internet of computing devices embedded in everyday objects, enabling them to send and receive data" (Lexico.com, 2019) –, and of Wireless Sensor Networks (WSN), concepts which are intrinsically connected with the concept of Smart City (SC). Fundamentally, a SC tries to foster better quality of life for its citizens by meeting their social and economic needs in a sustainable way, through the intelligent use of ICTs on the several services, infrastructures and information and communication streams that a city provides (Anagnostopoulos et al., 2015; Goluboniv, 2018; Pardini et al., 2019). SCs are about automating routine functions, but also about monitoring, understanding, analysing, optimizing and planning the city itself and all the processes within; and where citizens play a central role in the decision processes (Lundin et al., 2017).

MSWM is not an exception of processes being optimized and automated with SC technologies. It starts to be more and more indispensable to upgrade waste management systems by integrating ICTs into the planning and collection of municipal waste, with all related processes susceptible of being benefited. The design of more efficient waste collection routes is a complex task, and waste collection companies, in order to automate their processes, improve their performance and reduce their costs must be aware of the new technologies that may allow them to better adapt to the always changing conditions, great number of variables, demanding objectives and constraints that municipal waste collection is subjected to. If with the beginning of the 2000s Geographic Information Systems (GIS) and optimization software have been used in order to produce a better scaling and planning of the waste collection services and thus reducing the distances travelled by collection vehicles, more recently, the development and implementation of ICTs (e.g. ultrasonic sensors, RFID systems, wired and wireless actuators, GPRS and GPS technologies) in these services, which enable real-time monitoring of routes and bins' waste levels, allowed companies to better their operational decision-making process, and those tasks can be performed in a more precise, cost-effective and dynamic way. With such waste collection based on this IoT approach, we're moving towards a smart waste collection paradigm, where route optimization has been the main motivation for further research and development of these smart waste collection systems (Lundin et al., 2017; Pires et al., 2018; Ramos et al., 2018; Vallero et al., 2019).

The waste management system includes numerous containers that are filled inconsistently, sometimes taking days to get full, other times weeks. MSW production is in itself hard to predict. Also, the variety and the irregularity of the discarded materials, and the seasonal changes that they might be exposed to constitute more difficulties to the forecasting process, which in some extent generates unnecessary costs to the municipal collection system (Pardini et al., 2019).

With the ultrasonic sensor technology, the uncertainty about the amount of real waste level in the bins is reduced. Although, having access to this data is not enough and the development of routing models capable of handling with this information is then being pursued. As this technology is increasingly being developed and implemented, static approaches to waste collection routing and scheduling are being replaced by real-time dynamic ones which fulfil the potentiality of this IoT technology. According to Bopardikar et al. (2010), vehicle routing problems are concerned with the planning of optimal vehicle routes when providing a service to a given set of customers. In scenarios with complete information about the customers, the route optimization can be done in a static way; whereas, contrastingly, dynamic vehicle routing is used in scenarios where not all the customer information is known beforehand, and thus routes are re-planned as new customer information becomes available. In the waste collection case, the always unknown information about the customers is their waste production. Furthermore, a dynamic model is running continuously, while in a static model we have an execution that is repeated in a given time interval (Anagnostopoulos et al., 2015; Bopardikar et al., 2010; Ramos et al., 2018).

The data collected by the sensors is sent to the cloud to be processed and analysed. This permits to a central system to "know" the real demand for waste collection/generation and, in function of that, decide which bin should be emptied and which should not, and thus define and schedule the routes in a more optimal manner (see Figure 5) (Anagnostopoulos et al., 2015; Faccio et al., 2011). Such system could also be highly responsive to rapid changes in the filling levels and rearrange the routes while collection trucks are performing them (Anagnostopoulos et al., 2015). Aside from sensors, in these systems, for each container, regardless of its type, must be considered its static GPS location and its volume capacity, and they often have embedded RFID tags for container tagging and to identify when its waste has been collected by a truck. For each moving truck, its dynamic GPS location must be considered, as well its volume capacity to collect waste from the containers (Anagnostopoulos et al., 2015).

With the optimization of the collection routes, it is almost always possible to reduce the distances travelled, the number of vehicles, the number of containers emptied, the time spent in collection, while maximizing the use of trucks' load capacity. This all results in a reduction of fuel consumption, operating costs and environmental impacts in face of an irregular demand (Faccio et al., 2011; Johansson, 2006; Pires et al., 2018). Besides the economic and environmental factors, municipalities may be motivated to integrate these technologies to control the quality of the service, eliminate health and safety hazards (dirty and overflowing bins), and increase the citizens' overall satisfaction with the service.

Figure 5: Waste collection routes with (b) and without (a) the ultrasonic sensor technology (adapted from Gonçalves, 2014).

2.5.2. Identification of the problem

Despite all the studies in recent years, the use of ultrasonic sensors in waste collection is still a topic where there is a lack of standards, methodologies and of best practices. Vallero et al. (2019) underlines that dynamic routing and scheduling approach using GPS and weight sensors, generally, is not an adopted option due to the complexities and sophistication that this approach still requires.

Most times, the needs and requirements to implement such advanced systems widely are not fully discussed in detail, which makes waste collection companies discouraged on adopting them. Just as Lundin et al. (2017) reported, after interviewing companies working in waste collection, even if there is a belief that the current state of ICT technologies is mature enough to be able to monitor trash bins, there's still a plethora of economic, managerial, societal and political concerns that usually impede them of making such investment, and there is still a need for a better understanding on how these technologies are used and how their utilization can be optimized.

Studies in the field have been focused on the operational point of view, on the cost reductions that might be achieved by having smarter waste collection routes but neglecting the investment costs which would be necessary to adopt these technologies. In Pires et al. (2018), it is denoted that with the development and application of ICTs, waste collection costs may, in overall, increase. Lundin et al. (2017) reveals that commercially available solutions on the market are often seen as very expensive and high-end solutions. According to the entities interviewed by the authors, most of them consider roughly 70€ to be an appropriate price to pay for each sensor, being clear that the greater obstacle to the widespread adoption of smart collection systems is the high capital investment cost.

In Gonçalves (2014) it was developed and tested a route optimization model, for the north and west regions of Lisbon, capable of reducing the number of routes performed, the total distance travelled, the number of containers visited, and the amount of hours worked by the employees. However, the author of this study came to the conclusion that the costs of installing the ultrasonic sensors in all the containers would not compensate for the cost reductions that they would allow. In the study, it was estimated how much the waste collection company (Valorsul) would have to pay for an Enevo system that would cover 928 containers of separate collection (glass, paper/cardboard and plastic/metal packaging). For this number of sensors and depending on the number of years of the contract considered, Valorsul would pay a different price for the rent of each sensor, on a monthly basis. For the years considered (3 to 7 years) the price per sensor would vary between 17€ and 13€ respectively, resulting in an annual cost between 189312€ and 144768€. Such high amount of investment required is obviously very intimidating for any waste collection company, even if it would result in considerable savings in the future. This was not the case for the study in question, where annual savings of 14650€ had no chance of standing against the best situation (144768€).

However, the business models practiced can vary from manufacturer to manufacturer. Furthermore, in Gonçalves (2014) it was not considered several types of costs such as installation costs, maintenance costs or even the cost of software needed to integrate all of the gathered data. In another study (Straightsol, 2014b), where Smartbin's sensors were used to monitor Oxfam's donations banks (textiles and books), Smartbin charged to Oxfam the purchase price of 162€ per sensor and it was considered an installation cost of 50€ per sensor. Additional running costs (due to communications) of 11€ per month were considered for each sensor and it was estimated a total maintenance cost of 54€ per month for the whole system. Additional labour costs for the set-up and planning of the project, as well for the ongoing management where also allocated to the project.

In order to better understand the current prices and practices, for the present work, two more companies were also contacted: Evox and a Slovakian start-up company which preferred to remain anonymous. Evox sells the sensors to the waste collection company but also charges an additional rent for each one. The purchase prices are between 140€ and 360€ and the rent costs 1€ per month. The installation costs are also held by the client and the estimation given is that they may vary between 5€ and 15€ per sensor. Besides this, Evox also charges the client for the software's license which adds a cost surrounding the 5000€, annually. The Slovakian-based company offers two main options: the customer can buy the sensor and pay a monthly subscription for the company's software; or it rents the sensors plus the software for a minimum of two years. In the first option, the customer can pay a purchase price of 142€ to 299€ for the sensor, depending on the type of sensor and the type of network, and a monthly subscription of 95.76€ for the software, for each sensor. In the second option the rent to pay for the sensors and software varies between 113€ and 238€, per sensor, monthly.

A recap of the business models depicted can be found in Table 1.

Table 1: Business models and costs practiced by sensor manufacturers.

	Enevo in (Gonçalves, 2014)	**SmartBin in (Straightsol, 2014b)**	**Evox**	**Slovakian company**
Business Model Practiced / Costs Considered	Monthly rent between 13€ and 17€ per sensor	Price: 162€/sensor Installation: 50€/sensor Monthly running costs: 11€/sensor Monthly maintenance of system: 54€	Price: between 140€ and 360€ per sensor Installation: between 5€ and 15€ per sensor Monthly rent: 1€/sensor Software's Licence: 5000€ annually	Model 1: Price: 142-299€/sensor Software's monthly subscription: 95.76€/sensor Model 2: Monthly rent between 113€ and 238€ per sensor

The sensor technology is still very expensive and requires high investments by the waste collection companies. However, the value of using a sensor does not come from the sensor itself but from how the information about the filling level is used. And the value of having more precise information varies among containers. If companies could detect what are the most important containers to be monitored, it's possible that by monitoring only part of the population of containers, most of the benefits could still be retained.

Failing to notice this may be a limitation that waste collection companies are incurring. That being said, addressing this limitation will be the focus of the present work. In this work it will be studied if this technology could be integrated in waste collection activities in a financially more sustainable way, by trying to understand in which circumstances it is more useful to have a container being monitored by an ultrasonic sensor. Containers' characteristics like location, historical data, type of population covered must be analysed and a set of criteria considering these characteristics must be identified.

So, in this dissertation, it will be analysed different methods for choosing a reduced number of containers to be monitored from a set of containers that belong to a particular collection area. Identifying the best way of reducing the investment costs associated with the implementation of ultrasonic sensors, while maintaining the benefits for route optimization and scheduling, can a be step forward in filling the gap that is found by waste collection companies when they consider investing in this technology.

2.6. Chapter's Conclusions

In this chapter, it was possible to have an extensive look on the solid waste management topic. This is today, more than ever, an incredible complex and critical issue. Since societies are becoming technologically more advanced and populations tend not only to grow, but to have life habits that result in massive waste production, it was recognized how important it is to reduce the harmful impacts of the discarded materials in order to safeguard the social and environmental systems of today and of future generations.

Even if MSWM is a global problem, it is also a very contextual and local one. It was noted how the requirements and challenges that occur, for example, between developed and developing countries, are totally different and that waste management practices must adapt to the conditions existing in each geographic location.

The evolution of the Portuguese situation during the last 20 years was also object of analysis. During this period, several sustainable practices were adopted that have directed the country to decrease the quantity of waste that is disposed irresponsibly and without any type of valorisation, and to increase the reuse and recycling of waste materials, where the expansion of a separate collection for these recyclables played a major role. However, these measures, even if important, often neglected the operational performance of such services, whose inefficiencies had led to a substantial aggravation of collection costs, being important now to upgrade the collection models.

Then, waste collection was specifically depicted. This phase is what sustains any waste management system, being where the majority of the MSWM costs are located. However, it was seen how the costs could be reduced, if waste collection's resources were used more efficiently. The main reason why resources are not better used is a consequence of waste collection companies not knowing the real filling levels of the containers to be collected, resulting in trucks collecting partially full or even empty containers.

It was described how the embedding of ultrasonic sensors in containers, which are capable of measuring the level of waste is, more and more, seen as a solution to reduce this uncertainty regarding the waste demand, and thus adapt routes to the real collection needs. With this, the distance travelled and the number of vehicles needed could, for example, be reduced, while the truck's load capacity is better exploited. Thus, waste collection companies can benefit from less fuel consumption and lower operating costs while reducing also the associated environmental impacts.

Nevertheless, when adopting this new type of solutions, waste collection companies still have valid concerns. They consider that the solutions that exist in the market are very expensive and is not clear to them if the benefits gained could compensate for the high investment costs that are required. It was important then to analyse the current business models and the magnitude of prices that some of ultrasonic sensor providers charge, from where it can be concluded that it seems realistic that there's still work to be done in order to integrate these technologies in waste collection operations in a financially more sustainable way.

In order to fill this gap, in this dissertation, it will be sought to understand the cost-benefit relation associated to the implementation of a sensor into a container. The value of using a sensor in a container will not be the same for each container and, in that way, it must be studied in which ones the sensor will be the most useful, i.e., in which ones the information they may provide, regarding waste demand levels, is more likely to be the most valuable for improving collection routes and scheduling. For that, containers' characteristics must be studied and, in the end, several methods for selecting a reduced number of containers to be monitored, must be analysed.

3. Literature Review

In this chapter it will be investigated what authors have written in the past that relates to some extent to the problem at hand and may help solving it. First, in section 3.1, works related to the use of sensors in waste collection or that represent similar problems are considered in order to understand what strategies and approaches make sense to explore and integrate when deploying sensors in waste containers and, in particular, when trying to restrict the use of sensors to a subset of the total population.

Sensor placement problems (section 3.2) from other subject areas will be reviewed in order to comprehend what can be incorporated from those type of problems, with some important distinctions being also discerned.

Before heading to the conclusions (section 3.4), in section 3.3, with the purpose of better grasping some concepts behind the optimization of waste collection routes that can be useful when developing the dissertation, the node-routing problems TSP and VRP – two central problems of operations research – will be briefly delved into.

3.1. Aspects and practices affecting waste collection

As described in the previous chapter, there's a whole range of factors that affect waste collection costs. With respect to the problem of choosing a reduced number of containers to be monitored in a certain collection area, it is certainly important to study the containers' characteristics. During the course of this subchapter, it will be reviewed works that identify some of those characteristics that are important to be addressed when planning waste collection activities such as the spatial location of the containers, the characteristics of the population and area it serves and the existing historical data regarding filling rates (the demand characteristics). Furthermore, some works which explore routing practices that, through the use of remote sensors on containers, might improve the performance of the solution to be implemented will be addressed. A work that creates a model for dynamic waste collection that has focus on social and public health concerns is also mentioned. Finally, it is given attention to works or situations that are directly comparable to the topic of this work. Important to refer that the selection of containers to be monitored will not take into consideration the possibility of improving the containers' placement, i.e., the existing network of containers will not be modified.

3.1.1. Aspects regarding waste demand characteristics

In order to address the uncertainties associated with MSWM and, more specifically, with waste generation rates, stochastic approaches are often adopted. In Zsigraiova et al. (2013) is emphasized how the waste collection problem (WCP) is a stochastic problem in its nature since the amount of MSW is highly variable and depends on several factors as the number of inhabitants per container, GDP per capita, lifestyle and season. It also considers that there is a generalized lack of reliable information regarding filling rates of individual containers. The authors criticize the fact that, in order to remedy this difficulty, many works used the same average fill-up rate value to model the generation of waste in the whole area under study. In their study, the authors applied a statistical analysis based on historical data (covering several years) of the waste collection system managed by Amarsul S.A. for glass waste, in

Barreiro. This allowed the authors to obtain for each individual container a specific fill-up rate value and, with this, more realistic estimations of the waste generation of each container were obtained, resulting in a more precise collection schedule and better ahead collection planning than the previous practiced fixed collection schedule. The historical data was provided in the form of daily average of fill-up increments (m^3/day) of each glass container in that area.

Mes et al. (2014) also treats the WCP as a stochastic problem, where future demand is known only in a probabilistic sense and is revealed over time through the use of sensors. The authors also consider important to learn with the historical fill-up levels in order to create more reliable predictions for the future. The “predictive function” is considered to have a normal distribution, to account for waste deposition’s variability. More studies such as Vonolfen et al. (2011), Faccio et al. (2011) or McLeod et al. (2013), for example, similarly consider the use of the normal distribution to represent the stochasticity of waste generation at each bin. In the latter, historical data was also used to specify the averages and standard deviations of the demand for each receptacle. Another option is to model waste arrival rate as a Poisson distribution (Markov et al., 2019; Omara et al., 2018).

The degree of waste demand’s variability affects the adoption of sensor technology. According to Mes (2012), the highest savings of the proposed dynamic policy and the highest potential benefits of using filling level sensors are attained in unstable environments where there is a high (unforeseen) demand fluctuation. In such cases, traditional waste collection’s trucks will empty too less in peak periods, and too much in periods with lower generation. Besides the existence of some seasonal fluctuations and some weekly and monthly patterns (e.g. more deposits on Mondays and less on Sundays; and slightly more at the beginning of the month), there’s a huge random variation from day to day and from week to week. The author also verified that the type of location in which the container is placed heavily influences the filling-up rates, being them higher in containers close to stores than those from households. Not only that, but the author also noticed huge differences in deposition between containers on the same location.

In the United Kingdom, the real-life logistics problem faced by Oxfam was studied under the Straightsol project. Oxfam is UK’s leading charity entity in servicing textile and book donation banks and faces essentially the same problem of waste collection companies in adjusting routes and schedules to the amounts of materials inside the receptacles. Several experiments and studies were conducted in order to study the installation of sensors in the donation banks to report their filling levels (Straightsol, 2015). In the related studies, McLeod et al. (2013) and McLeod et al. (2014), the authors also concluded that the best performance for dynamic optimization of collection routes and schedules through the use of sensors (relatively to the fixed collection) is obtained when the donation banks experience highly variable fill rates; and that when donation rates are fairly predictable, the need for remote monitoring at collection sites is less. This is in accordance to what Johansson (2006) concluded by comparing the impact of waste demand variability on the savings potential between a static and a dynamic scheduling and routing policy: even if a higher variability increases costs regardless of the policy, the impact is much more felt on the static one. Adding to this, if the donation banks, besides having high variability, also have relatively low fill rates (sporadic donations of large amounts), makes them even more suitable for remote monitoring (Straightsol, 2014a). Krikke et al. (2008) also concluded that the value of having accurate information about the quantity of material to be collected increases with decreasing collection frequencies (lower demands). Moreover, when collection points fill up more quickly, the less are the improvements in monitoring them, since frequent fixed rounds are capable of ensuring a relatively efficient collection (Straightsol, 2013).

Instead of looking into the historical data, when studying and predicting waste generation rates, and their composition and spatial variability across a certain area, it’s also possible to analyse demographic and economic variables such as population density, household sizes, income, size and type of the main commercial activities and even aspects as the presence or absence of gardens (Purcell & Magette, 2009). However, according to Niska & Serkkola (2018) the relationship between waste generation, consumer behaviour, products and environment is very complicated and often poorly understood. In their work, the authors analyse and model waste generation through the use of descriptive analytics and they are able to classify each site into a group having a similar waste generation pattern (i.e., each site has a type profile) by using a clustering algorithm, based on historical/monitored data. This data can also be merged with the demographic and economic data in order to better interpret and find influences behind waste generation.

3.1.2. Spatial aspects regarding waste collection points

MSWM systems are clearly of spatial nature and they depend on the facilities locations and containers distribution (Zsigraiova et al., 2013).

Besides considering that remote monitoring is more appropriate for receptacles with high variability, McLeod et al. (2013) and McLeod et al. (2014) also account that those that are located in more remote locations (more isolated) are preferred to be monitored since the risk of performing costly long-distance trips to partly filled receptacles needs to be avoided. This is a reason why savings can vary from region to region. As result of the density of population, networks of containers have often quite large concentrations in urban areas in opposition to being very dispersed in rural areas (Zsigraiova et al., 2013). The dispersion of collection points poses challenges to the MSWM and increases the cost per tonne of waste collected (APA, 2019d; Hoornweg & Bhada-Tata, 2012; Teixeira et al., 2014). According to Straightsol (2014a) and Straightsol (2014c), since the density of receptacles is higher in urban areas, the distances driven to collect each one are lower, resulting in fewer possibilities of reducing the total distance travelled due to dynamic routing based on the measured filling levels. Following this, relative remoteness of sites from the depot and from each other results in greater distances that provide a higher savings potential for a remote monitoring solution. It’s logical that, due to this, the benefits of deploying sensors in rural areas will be greater than in urban areas, meaning that the type of area contributes significantly to the success of the solution.

Furthermore, due to time constraints, for regions close to the depot, more receptacles can be included in one collection trip, as for regions far away from the depot, less receptacles can be collected per trip. The same happens with the average distance among receptacles: the longer the distance, the fewer the receptacles that can be collected and vice-versa. With fixed collection schemes, this risk is not taken into account and the collection frequencies to regions that have low quantities of material and are far from the depot can be set too high, resulting in higher collection costs (Liu et al., 2011).

If some containers of the network may be well isolated, others may be so close to each other or in a location where the collection truck will inevitably pass by as it performs the collection routes. These aspects may be considered into the problem of the presented work and into dynamic routing practices in general.

3.1.3. Dynamic routing and scheduling practices to integrate with remote monitoring

Dynamic scheduling is an inherently difficult problem where there's not a standard technique to be applied across all the instances of the problem (McLeod et al., 2013). When testing routing and scheduling policies that intend to ensure that containers are collected when they are close to full is common that authors define a threshold which when exceeded indicates that a specific bin must be collected. One example is Zsigraiova et al. (2013), where a threshold of 70% of the bin's capacity was established. Johansson (2006) developed further the understanding about the threshold value and concluded that it highly depends on the network density. It stated that policies that collect full containers are the most appropriate for dense networks, where distance between containers are smaller. For sparser networks, the threshold value decreases, meaning that it makes sense to collect containers that are *not entirely* full. Other works explore the use of two thresholds: one to indicate when the container *must be visited* in the next route and other to indicate that a container *is allowed to be visited* in the next route. For example, McLeod et al. (2013) considered values of 50% and of 75% of the container's capacity as the two thresholds; and values of 50% and 90% were used in Marchiori (2018). In the latter case, if in a day there is at least a bin that is above 90% full, a waste collection round is set up, and when this happens not only the critical bins (≥90%) are collected but also the potentially critical ones (≥50% and <90%).

A similar approach is employed in Mes et al. (2014) and in Stellingwerff (2011) where, after all "MustGo" containers have been scheduled, it is considered that collection vehicles might be able to still load more waste from other containers in order to improve the capacity utilization of these vehicles. These containers that might be planned into an existing route are defined as "MayGo" containers and are only added if they are more convenient to be emptied now than later. To decide if a "MayGo" should be added into a route it's estimated its insertion cost by calculating the ratio of the additional travelling and handling time the container requires and the volume of waste measured (it's convenient if the additional time it takes to visit the container is relatively small when compared to the quantity of waste collected). In this way, it would always be less convenient to insert a container into a route the more distant a container is located. To counterbalance this situation present ratios are compared with the smoothed average of past ratios in order to perceive relatively good opportunities. The authors proved that this approach improved routing flexibility, leading to better results; however, a limit for the number of "MayGo" containers to be collected should be established, especially for relatively low deposition rates, since too many emptyings and early collection of some container could lead to unnecessary costs.

Collecting waste from containers that are not close to full but that are relatively near to a full container seems an intuitive way of improving dynamic waste collection. Nourinejad et al. (2018) and Ramos et al. (2018) are two works that also state that collecting only containers that are near capacity (see Figure 6.a) is not necessarily the best solution, since it doesn't take advantage of visiting opportune containers that have lower filling levels in order to save future costs (see Figure 6.b). As shown in Ramos et al. (2018), an operational management approach that, by considering containers' filling levels, their location and a longer time horizon, is capable of not only deciding the best days to collect the containers, but also of recognizing what small additional distances are worthy to be travelled for less than full containers, is preferable when compared to an approach where routes are optimized for a specific day taking into account only the containers that have reached a specified minimum filling level. Similarly, in Nourinejad et al. (2018), the authors also propose an anticipative model that minimizes each day's costs while considering the consequences of that day's actions on future costs. Furthermore, after performing a sensitivity analysis, these authors concluded that for scenarios with lower production, the collection can be postponed and done at a higher filling level, where for a more intensive demand it's better to collect at a lower filling level. Likewise, Markov et al. (2019) advocates that "it may be cheaper to postpone a collection if the container in question fills up slowly. On the other hand, it may be worthwhile collecting an almost empty container today if we know that it will experience high demand in the coming days".

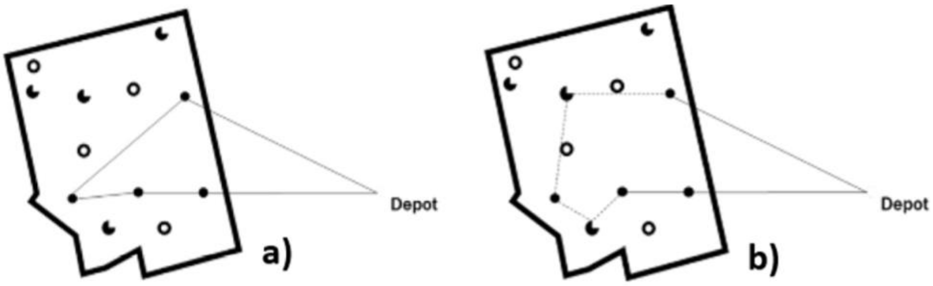


Figure 6: Collection of only almost full containers (a) and of almost full and nearby containers with lower filling levels (b) (adapted from Nourinejad et al. 2018).

This type of solutions can also be found in collection of other types of materials. In Krikke et al. (2008), for the collection of disassembled materials from end-of-life vehicles, the authors differentiate between "MUST" (filling level close to capacity) and "CAN" (significative filling level but not near capacity) orders, where the latter are only included in routes if they can be combined beneficially with "MUST" orders. For "CAN" orders to be included, they must be near "MUST" orders and can be inserted in routes at a low marginal cost. The authors came to the conclusion that this solution is more appropriate in collection networks with low waste demands. In Liu et al. (2011), for the collection of hazardous wastes, it was studied the possibility of: every time a campus has to be collected, neighbour campuses, if possible, would also be collected if their hazardous waste exceeds a pre-specified minimum quantity. This minimum quantity cannot be set too low or neighbouring campuses with low quantities would be collected too often, which is cost inefficient; but if it is set too high, some collection trips would not be as fully loaded as possible, which is cost inefficient, too. Moreover, in Elia et al. (2018), the authors study the problem of collecting Waste from Electric and Electronic Equipment (WEEE), which is highly unpredictable, with resource to IoT technologies to collect real time data on the waste levels. The authors studied the adoption of a "mixed" solution that combines features of dynamic and fixed scheduling policies. The "mixed" solution consisted on providing the collection service when at least one customer needed to be served, answering, in this way, to the necessity of the demand; and, when a service is started, all the other customers are served, independently of their filling level. This approach shown a reduction on the number of services provided and on distances travelled, and a higher utilization rate of the truck capacity was also obtained, which resulted in less costs when in comparison to exclusively fixed or dynamic policies.

As Mes (2012) puts it: “emptying a container that is far from full might still be efficient when a truck just passed this container”. This was felt by Oxfam’s drivers, who tested the implementation of dynamic routing based on remote monitoring, and highlighted that not collecting banks when they were passing by them on one day and then having to visit them on another, less convenient day, was one of the aspects that they disliked the most (Straightsol, 2014a). Omara et al. (2018) explored this idea by testing two heuristics where the collected bins would not only be the ones that triggered alarms (the ones that reached an upper threshold), but also the ones that are on the assigned route and have not triggered alarms. According to the authors, these approaches intend to reduce the cost by optimizing routes and reduce unnecessary trips. In one heuristic (“CUT” – only has upper threshold), bins in the assigned route are collected independently of their filling level, where in the other heuristic (“CULT” – has upper and lower threshold), bins in the assigned route are only collected after reaching a predetermined filling-level (a lower threshold), in an attempt to reduce unnecessary service to passed-by containers with a negligible amount of waste. Several scenarios were simulated. These two heuristics presented better results than the case where routes would be performed without collecting any intermediary bins that haven’t reached the upper threshold. Furthermore, between “CUT” and “CULT”, the former heuristic led to higher savings. It’s important to note that the authors intended to simulate an urban environment, where fill-up rates have higher values and is not common to have accumulative wastes in the bins for multiple consecutive days. Whereas, when they tested the heuristics for lower demand rates (more similar to rural areas), “CULT” shown a better performance, which is consensual to what was here discussed before.

All these practices were described since it’s important to better understand how (and in which situations) the information collected by using remote sensors can be well integrated with the truck collection routes before choosing in which containers to embed a sensor. Truck utilization can be improved by collecting waste from collection points that are nearby or in the way of the collection routes. It’s arguable that instead of looking at each container individually, it would make sense to look at them as a group/cluster of containers, grouping them to the other containers of the same region/location. For example, Marchiori (2018) suggests that selecting a number of representatives from each group with a relatively high number of close containers and then interpolate the demand accordingly for the other containers could be adopted to reduce the number of sensors used. However, this suggestion was not further developed by the author.

3.1.4. Other criteria – high priority bins

Anagnostopoulos et al. (2015) analysed a dynamic waste collection model based on the use of sensors to measure waste levels, which considered the existence of high priority wastes (in which the model is focused on). The authors considered that there exist areas where the waste deposited must be paid special attention to, since its presence can be particularly harmful to human lives and to their quality of life, and so there's a requirement for instant collection in these areas. The areas include areas where waste rejected is especially dangerous to humans such as hospitals or surroundings of factories; areas that are inhabited by or concentrate sensitive groups of people like schools or neighbourhoods where mainly elders reside; touristic areas or even the surroundings of gas stations (increased risk of fire). The authors characterize bins that exist in these areas as high priority bins and the dynamic model must adapt the routes rapidly after an alarm is triggered for this type of bins.

In a situation where a reduced number of bins has to be selected for waste monitoring, an method where sensors are embedded primarily into bins that impose the most risks for populations and is definitely a valid option due to the positive social impacts this method may have.

3.1.5. Approaches for selecting containers to be monitored

3.1.5.1. Review of Correia's work

To the best of the author's knowledge, in the literature, there's only one work that addressed the problem of monitoring a subset of containers through the use of sensors, in order to reduce the implementation costs (and still maintain the benefits of dynamic routing and scheduling), with the purpose of integrating this technology in a financially viable way. The referred work is Correia (2016) and it also follows up the conclusions of Gonçalves (2014) – it even considered the same region as Gonçalves (2014). In the work, the author proposes a methodology that revolves around grouping containers that have similar characteristics – forming "clusters" of containers – and where the installation of sensors is pretended to be carried out for a subset of containers of each individual cluster, producing a sample "more representative of the containers' diversity and specificity". For the clustering analysis, the proximity measures used were the number of inhabitants served by each container, the filling-up average of each container and the filling-up coefficient of variation of each container (quotient between the standard deviation and the average values).

After choosing a subset of containers for a cluster, data manually collected by waste collectors in a visual and approximate manner was used to simulate the sensors' measurements. These measurements "obtained" for the sample were then used to extrapolate the filling levels to the whole group of similar characteristics. After, the forecasted values were compared to the real ones and it was calculated the error associated to the difference between them. A trade-off between the costs of using the ICT technologies and the risks associated to the non-monitoring of all the containers was considered and several samples of different dimensions were studied in order to find the optimal percentage of containers to be monitored that balance the risk and investment.

Regarding the analysis that the author performed: first, it was assumed that, independently of the number of containers monitored, the benefits obtained in Gonçalves (2014) with the optimization of routes could be inherited into his work. Since it was also assumed that containers of the same cluster would be collected in the same day, this means that the routes would be totally different from the ones in Gonçalves (2014). Borrowing the results from Gonçalves (2014) is a faulty approach, since simulating the routes that would be originated by sensors' measurements (and have their costs calculated) is a core aspect to understand how well the proposed solution could in fact work. Also, from a topological point of view – and since the spatial aspects were not considered when creating each cluster –, the collection routes originated by collecting containers of the same cluster in the same day might themselves be incredibly inefficient and there's no way to know that in the proposed model.

Moreover, the author is using the sensors not to know *which containers are able to be collected* but to indicate *if a cluster of containers is ready to be collected*. In other words, no use is made of the *individual information* related to a container's waste filling level, being only considered in the model the *number* of containers that might be ready or not to be collected. This means that the model neither tries to tell if a specific non-monitored container is ready to be collected or not, nor does it enable to avoid containers that have an insufficient waste level to be collected. By collecting the entire cluster in the same day, the author is accepting the risk of collecting containers in that cluster that are not able to be collected and labels a cost to those cases. However, once again, that cost does not take into account the spatial relative position of that container and it is over-simplified by assuming an average value for all cases.

In sum, some of the assumptions and approaches adopted were fuzzy and represent an over-simplification of the problem in question; and the model is not taking full advantage of the information it has about the filling-levels.

3.1.5.2. Oxfam's demonstration

Only during the aforementioned Oxfam's demonstration, where sensors were installed in donation banks of textile and books, decision-makers were in a situation where they had to consider equipping a sample of the total population of containers. They had planned to equip all the donation banks in the demonstration area with sensors. However, some of the sensors were vandalised or damaged, despite being installed inside the containers. Due to this setback, the team revised the plan and had to choose in which banks they would install sensors. Thus, the distribution of the sensors was reorganized, now in a subset of the banks. From the different 58 bank sites where they intended to install sensors, only 21 ended up being monitored (36%). They preferred to install the sensors in the more remote sites, since they reasoned that these would lead to the greatest benefits through visiting them less often. The team considered that they would only schedule dynamically to where the sensor is working and, for the banks not equipped with a sensor they would use fixed collections. Even though the team was faced with this situation, the solution adopted was more of a "quick-fix" as they would always recommend to install sensors in all containers in order to collect the most benefits for the dynamic routing and scheduling (Straightsol, 2012, 2014a, 2014b). This means that they didn't consider exactly the problem that is being addressed in the present work, where is being studied the possibility that picking the right containers to monitor may be advantageous since it can reduce the implementation costs while still allowing to exploit the benefits of remote monitoring.

3.2. Sensor placement problems

There's a whole range of applications where sensors are becoming increasingly popular for monitoring phenomena, where environmental control for firefighting purposes, intruder surveillance on private property or marine ground erosion detection are some of the examples (Guerriero et al., 2011). Here, in these situations, choosing the right sensor locations is an essential task. Guestrin et al. (2005)

considers the optimization of sensor placement that increases the measurements' prediction accuracy and tested the proposed criteria for indoor temperature measurements and precipitation data. The optimization criterion chosen by the authors seeks to place the sensors in the locations that will most significantly reduce the uncertainty about the estimates in the rest of space. Other example is Elsersy et al. (2016), where sensor node optimization is used in structural health monitoring (controlling buildings' conditions). Here, both the node placement and the routing (communications between sensors) are taken into consideration conjointly, since the latter is highly dependent on the former. The best route for improving the performance is encountered in an iterative way. The initial set is updated if the sensor node that has the "least contribution" in the set can be replaced by one that contributes more positively; and the process repeats until no further improvement can be observed. According to the two works, since they are limited to installing a reduced number of sensors, they must carefully select where to place them. In other words, this problem can be formulated, very simply, as: finding the *k* best sensor locations out of a finite subset *V* containing *n* potential locations. A set $S = \{s_1, s_2, \dots, s_n\}$ must be found, where s_i is a binary indicator that is equal to one if location *i* has been selected and zero otherwise.

The sensor placement problem, as defined in Munagala (2009), is a stochastic optimization problem where the solution is built upfront based only on the probabilistic information about the inputs. For the *n* candidate locations, where sensors are possible to be placed, there's *k* ≪ *n* sensors. In each location correlated data is sensed, and the purpose is to place the sensors in such a way that the amount of information captured by the sensors regarding the physical phenomena being sensed is maximized. The aforementioned definition means that sensor placement problems (as a field of study) are fundamentally distinct from the problem being studied in the present work. In these types of problems, a sensor is used to capture information of a given area of interest that intersects with the area being monitored by other sensors, but in the present problem, a sensor only provides information about an individual container, meaning that each measurement depends on only one sensor. Moreover, whereas the goal in sensor placement problems is to maximize the information obtained from the sensors' measurements; in the present problem the main objective is to maximize the decision-making ability regarding which containers to collect, so that the best possible routes can be undertaken by the collection vehicles, i.e., it's related to the fact some containers may provide more useful information than others.

Another important distinction is the fact that the sensor placement problems described take advantage of the submodularity property which, intuitively, can be seen as a diminishing returns property where adding an element to a larger set gives a smaller gain than adding one to a smaller set. By exploring the submodularity property of a certain criterion, at each iteration the sensor that is added to the network will provide less benefits than the one added before it (Guestrin et al., 2005; Munagala, 2009; Summers & Lygeros, 2014). In the present problem such property does not exist, and it's not guaranteed that a subsequent sensor couldn't provide a higher benefit than the one that precede it.

Even though what is known to be a sensor placement problem not only does not include the problem here concerned, but is also categorically different from it, an analogue formulation can still be extracted: finding the *k* best containers to be monitored by sensors out of finite waste collection area *V* containing *n* containers. A set $S = \{s_1, s_2, \dots, s_n\}$ must be found, where s_i is a binary indicator that is equal to one if a container *i* is selected and zero otherwise. Considering Elsersy et al. (2016), the deployment of sensors in containers could encompass an estimate of how much that would allow to reduce the collection costs and deduce the containers that have a higher “contribution” for that matter.

3.3. Routing problems

Considering what was reviewed in subchapter 3.1, routing problems will be briefly explored here. The purpose of this is understand how these could be applied together with the knowledge gathered about the relevant spatial aspects and dynamic practices, in order to better decide where to strategically place the sensors.

A routing problem for a fleet of vehicles is concerned on finding efficient paths for transportation needs through a complex network and can be classified as a node-routing problem or as an arc-routing problem, depending on where the demand occurs in the network: it can occur in the nodes and hence the goal is to visit the locations, or it can occur in the arcs and then the goal is to visit the edges that connect the locations (Pearn et al., 1987). One application of an arc-routing problem is, for example, what is necessary for Google to solve when taking pictures of addresses to be fed into the Google Maps Street View technology, as it is needed to find the shortest route for each vehicle that crosses every street (the arcs) in an assigned region (Google Developers, 2019).

However, considering the problem here discussed, node-routing problems are the ones that make sense to investigate – in the context of optimizing waste collection routes, container sites and the depot are interpreted as nodes of the network.

Two of such problems, central to operations research, that might be helpful for developing methods for sensor placement, are the Travelling Salesman Problem (TSP) and the Vehicle Routing Problem (VRP). Typically, a huge number of solutions are possible to be obtained in this type of problems, and as a result exhaustive search is not a valid option, i.e., enumerating all the possible candidates for a solution, in order to find the (most suitable) solution (Vukmirović et al., 2019).

3.3.1. Travelling Salesman Problem

The TSP was inspired by the problem faced by a salesman that has to visit a certain number of customers that are located in different cities and wants to visit them by performing the shortest possible trip. Each city is visited exactly once and after visiting all of them, the salesman returns to its origin city (Jünger et al., 1995).. More abstractly and generically, it is the problem of, given a directed, edge-weighted graph, finding the cyclic path with minimal weight that visits every node in this graph exactly once (Hoos & Stützle, 2005). A graph representation of the TSP can be seen in Figure 7. The cities are represented by the nodes and the edges contain the distance between each city (the weights).

An instance of the TSP is called symmetric if, and only if, the weight between two nodes of the graph is the same in each direction, i.e., $w_{A,B} = w_{B,A}$ for any pair of different points A and B. If not symmetric, the instance is called asymmetric (Hoos & Stützle, 2005). One-way streets are an example of when distance between the same two points may be different depending on the direction considered, and thus optimizing street-level routes can be a practical problem instance that requires the use of an asymmetric TSP.

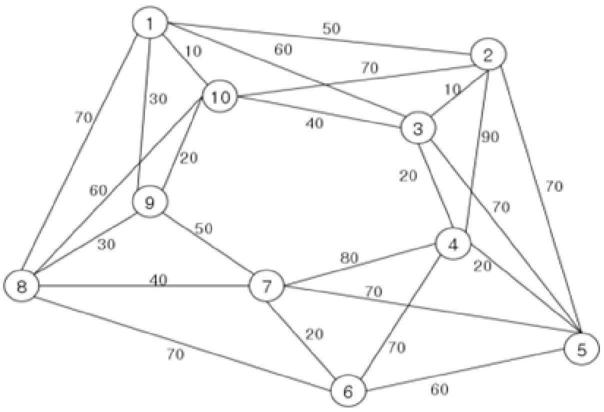


Figure 7: Graph-based representation of a TSP with 10 nodes (adapted from Rivlin, 2019).

For a TSP with *n* cities, an exhaustive search algorithm would have to enumerate ½ * (*n*-1)! possible routes and return the permutation with minimum cost (the product by ½ is to avoid overcounting solutions). This means that the time complexity of such algorithm would behave like Θ(n!): for 4 cities, the number of possible tours would be ½* 3! = 3; for 8 cities would be ½ * 7! = 2520; for 12 the number is 19958400; and for 100 cities this number equals 4.666311e+155, a more-than-astronomical number. Through Dynamic Programming solutions, the problem can be solved in $\Theta(2^{\mathrm{n}} * \mathrm{n}^2)$, which is much less than Θ(n!), but still represents an exponential running time. In fact, there's no polynomial time solution known for this problem (Cormen et al., 2009; Hoos & Stützle, 2005). Although, there are approximate algorithms that are capable of solving the problem like the Christofides algorithm, which guarantees an approximation ratio of at most 3/2, i.e., the worst possible ratio of the value of the answer obtained by the algorithm to the value of the optimal solution is of at most 3/2 (Christofides, 1976). These approximate algorithms work only if the problem instance satisfies the triangle inequality. Satisfying triangle inequality is a common restriction for the TSP, which means that weight between two points A and B is never of greater weight than the weight between A and B through other intermediate point C. More generally, a TSP instance satisfies the triangle inequality if, and only if, $w_{A,B} \leq w_{A,C} + w_{C,B}$ for any trios of different nodes A, B and C (Cormen et al., 2009; Hoos & Stützle, 2005).

The TSP, like many mathematical problems, can be formulated as integer programming problems, i.e., linear programming problems in which the variables must assume integer values.

3.3.2. Vehicle Routing Problem

The first time the VRP was formulated was in Dantzig & Ramser (1959), where it was named "The Truck Dispatching Problem". In the paper, the authors were concerned with finding the "optimum routing of a fleet of gasoline delivery trucks between a bulk terminal and a large number of service stations supplied by the terminal". With the shortest distance between any two points in the system being provided and a demand being specified for each of the stations, the trucks must be assigned to the stations in a way that demands are satisfied and the total travelled distance by the fleet is minimized. In the node-routing problem described by the authors, each carrier had a finite capacity *C* that could only supply a limited number of delivery points per trip.

The VRP can be seen as generalization of the TSP (Dantzig & Ramser, 1959). Whereas in the TSP one route serves all the orders, in the VRP more than one route is needed to serve all orders (see Figure 8) (Liu et al., 2014). Nowadays, the VRP is used when a set of routes for a fleet of vehicles that are based at one depot must be found for a number of geographically dispersed locations that have particular demands with the routes starting and ending at a depot (Díaz, 2006; Kumar & Panneerselvam, 2012).

Several variants of the VRP are formulated based on the nature of transported goods, the quality of service required and the characteristics of the customers and the vehicles. Some of its variants include the capacitated VRP (CVRP), where the vehicles have maximum loads; the VRP with time windows (VRPTW), where each location has to be visited within a predefined time window; the periodic VRP (PVRP) that requires visits to be done in specific days; or, for example, the stochastic VRP (SVRP) where some of the values (like customer demands or travel time) are random (Díaz, 2006; Kumar & Panneerselvam, 2012). According to Laporte (2010), the classic VRP consists of having *m* vehicles (which can be an input or an output parameter) of identical capacity *Q*, and the purpose is designing *m* vehicle routes of minimum total cost such that: 1) each route starts and ends at the depot; 2) each customer appears in exactly one route; 3) the total demand of a route does not exceed Q; and 4) the length of a route does not exceed a prespecified constant *L*.

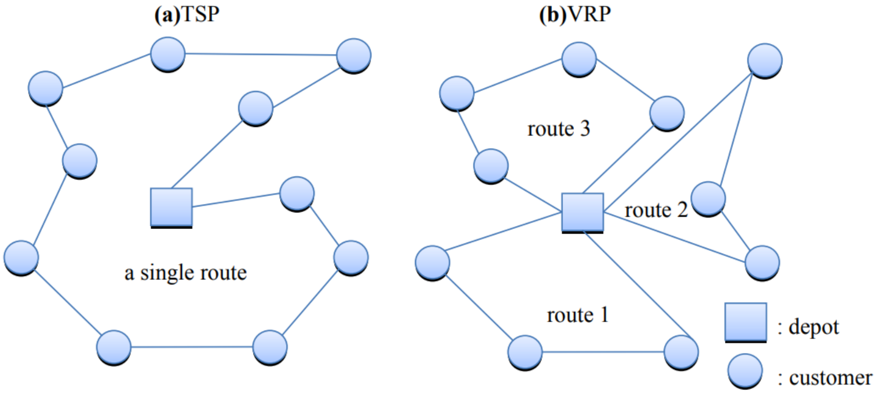


Figure 8: TSP vs VRP route patterns (source: Liu et al., 2014).

Similarly to the TSP, instances of this node-routing problem can be extremely difficult to solve in practice as the computational effort required increases exponentially with the problem size. Consequently, approximate solutions like some heuristic methods that give reasonably accurate solutions in limited time, are often preferable for real-life problem instances that involve large vehicle fleets and that affect greatly logistics and distribution strategies.

3.4. Chapter's Conclusions

In this chapter, the existing literature that relates to the problem in study was investigated. It was disclosed that this problem had only been directly addressed once in the existing literature, and some concerns related to the work's assumptions and methodology were brought up. Notwithstanding, in several instances was possible to obtain helpful knowledge from a broad range of works, which hinted at several aspects that may play an important role on the deployment of sensors in a reduced number of containers. These aspects were examined to better understand how to solve the problem in question.

First, the demand characteristics of waste were considered. It was acknowledged that waste demand at each bin must be dealt as having a probabilistic distribution and statistical analyses based on historical data may be used to improve solutions. Several works outlined how waste demands rates and variability affect remote monitoring's utility for the optimization of collection routes and schedules. The value of having accurate information about waste levels increases for collection areas that have low deposition rates (require lower collection frequencies) and that experience more unpredictable rates (have a higher variability).

It is also necessary to consider how containers' locations are spatially distributed. A higher dispersion of collection points results in relatively greater challenges and costs for waste collection companies. The use of volumetric sensors allows for a greater savings potential in situations where sites are, on average, more isolated from each other since there are more possibilities of reducing the total distance travelled due to the dynamic routing. With this, containers that are more isolated are also preferred to be monitored since unnecessary long-distance trips can then be avoided. This was the criteria adopted by Oxfam when some of the sensors they planned to use were damaged and decision-makers had to choose where to implement them in only 36% of the total number of sites.

It was carried out a review over several dynamic and scheduling practices that might improve the integration of remote monitoring technologies, especially in the context of the problem in study. It was highlighted how different authors agree that it can make sense to collect waste from bins, even if these are not near full capacity, in order to improve the collection vehicles' capacity utilization and save future costs. This is explained by the fact that it can be convenient to add partly filled-up containers to the collection process since sometimes their location is already included in (or is relatively close to) the assigned routes. Furthermore, as deposition rates increase, the better is to collect at lower filling levels and vice-versa.

The TSP and the VRP, were briefly explored to understand the complexity of optimizing routes in real-case scenarios. A notion was given about some concepts that may be helpful when developing the methods for selecting the containers to be monitored.

Besides considering only the financial benefits of using sensors, other approach is to consider the social benefits and use them preferentially to monitor high priority wastes, i.e., wastes that due to their composition or location are more likely to have adverse effects on communities.

Also, sensor placement problems from other fields of study were very briefly discussed. From this, it was possible formulate an analogue outlook for the problem here concerned.

Lastly, and even if it was possible to obtain insightful knowledge during the literature review, it is certainly correct to affirm that there is a literature gap related to the problem under study, which this dissertation intends to eliminate. There was not enough research focused on evaluating more economical ways of introducing remote sensors into the waste collection activities. It can be added value to the literature by studying how the different characteristics of the containers may make them more or less relevant to be monitored. This is important, since it may be what waste collection companies need to start investing in the remote sensors' technology, allowing them to upgrade their current waste collection systems and decrease operational costs.

4. Methods for sensor placement

In this chapter, it will be presented several methods for deciding which subset of a population of containers must be monitored with ultrasonic sensors. The methods proposed are based on some of the criteria that were identified during the literature review and which affect this decision. These criteria will be outlined and the different methods are described in detail in section 4.1. Next, in section 4.2., it will be explained how the methods will be evaluated and how the operational results will be simulated by adopting a dynamic collection policy.

4.1. Proposed Methods

4.1.1 Relevant criteria found in the literature

During the literature review it was possible to understand better in which circumstances may make more or less sense to invest in a sensor to be used in a specific container, with the purpose of maximizing the benefits of this technology. In this part of the work, it is important to clarify which criteria will be used and why they make sense to be incorporated. It was identified four main criteria: low waste demand rates, high waste demand variability, high remoteness, and container not being an intermediate one (i.e., a container not being in a location where the collection truck will inevitably pass by when going to other container locations). The different methods described in the following sections for selecting the containers will apply these criteria to the population of containers.

The first three methods will consist on a directly applying an individual criterion – waste demand rate, waste demand variability and remoteness. The fourth method is an algorithm that intends to indirectly apply some of the criteria – remoteness and the degree of how much intermediate a container is – and the concepts learned regarding routing problems. Finally, the fifth method, will explore the application of the three first criteria simultaneously.

For the direct application methods, the different containers must be ranked in each criterion, accordingly to their score in each one of these. From here an analysis of the relative score values can then be carried. In order to perform this analysis, it is necessary to define a measurement unit for evaluating each criterion. With the definition of the measurement units it is possible to specify how each container behaves in a given criterion. In this way, for each criterion, an ordered list with all the containers is obtained, where the order is given by starting in the container that scored higher in that criterion and ending with container that scored the lowest.

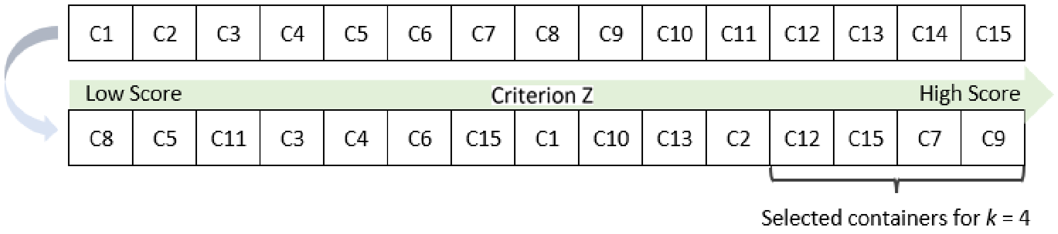


Figure 9: Selection of containers based on a single criterion.

The selection of the containers can then be ensued. In order to individually study the impact of the different criteria, different samples of different sizes can be evaluated. This means that a number k

must be established so that a sample of the *k* containers that scored higher, in each criterion, is selected. For example, as seen in Figure 9, the list of 15 containers is sorted according to a given criterion Z, and for a value of *k* = 4, sensors would be placed in containers C12, C15, C7 and C9.

4.1.2. Method 1 – Waste demand rate of a container

This criterion is related to the rate of deposition at which a container fills up. During the literature review it was concluded that it makes more sense to have a sensor in containers where the waste demand rate is lower (Krikke et al., 2008; Straightsol, 2013, 2014a). This means that the value of monitoring a specific container (the value of information) is inversely proportional to the waste demand rate of that container. In Figure 10.a, considering that the deposition rate of A is higher than B, and everything else equal, is preferable to monitor B than A – considering a fixed collection period for the containers without sensor, a sensor in B allows to reduce more the number of unnecessary kilometres, since the distance from the road junction to the container is saved more times.

The measurement unit that will be used to evaluate this criterion is very straightforward and is the sample mean of the waste demands rates of the container ($\bar{x}$) – the mean of the observed values $\{x_1, x_2, \dots, x_n\}$:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

Consequently, according to this criterion, a container will be higher ranked for being selected as its sample mean decreases.

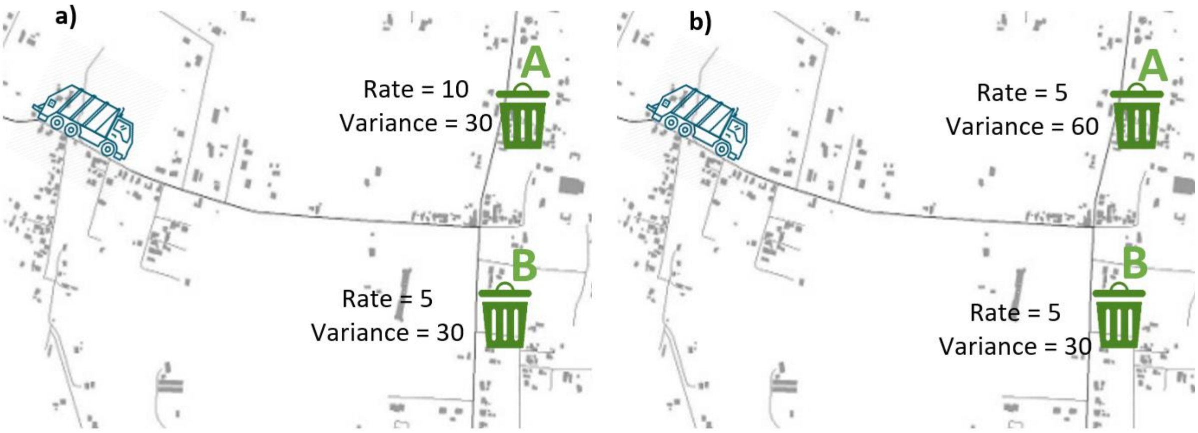


Figure 10: Representation of the criteria related to the waste demand.

4.1.3. Method 2 – Waste demand variability of a container

The waste demand variability is the criterion that is related to how the deposition rate varies in time. In the literature review it was concluded that when the deposition rate is uniform across time (lower variation of the deposition rate values), the need for a sensor is less, and when it is more floating (higher variation), the need for a sensor is higher (Johansson, 2006; McLeod et al., 2013, 2014; Mes, 2012). In other words, the value of information is proportional to the waste demand variability of a certain container. In Figure 10.b, assuming that the variability of A is higher than B, and everything else equal, is preferable to monitor A than B – we can predict more accurately the filling level of B without the need of a sensor.

In order to define a measurement unit for evaluating the variability of the waste demand, i.e., how far the waste demands rates observed for a container are spread out from the mean value, two direct tools such as the sample standard deviation (s) or its square, the sample variance (s^2), could be used:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{3}$$

However, in order to correctly compare the degree of variability amongst all the containers, these two measures cannot be used, since they are dependent on the value of the mean of each container. These two measures have always to be understood in the context of the mean of the data, so would be inappropriate for comparing, for example, two containers where the absolute values observed differ significantly. Therefore, a dimensionless measure like the sample coefficient of variation ($\widehat{C_V}$), that is independent of the unit in which the measurement took place and that can indicate the relative dispersion of the data, is required (Pang et al., 2005). It can be obtained by dividing the sample standard deviation to the sample mean:

$$\widehat{C_V} = \frac{s}{\bar{x}} \tag{4}$$

The sample coefficient of variation will then be used to assess the waste demand variability of a container. The higher this measure, the more a container is preferable to be selected.

Note that the terminology “sample” has been used, since the observed values constitute only a part of the population of all the waste demands that a container has experienced, i.e., the population is considered to be all the waste demands since a container was installed. Otherwise, the formula for calculating the standard deviation would be slightly different.

4.1.4. Method 3 – Remoteness of a container

The remoteness of a container is the criterion that is related to how much a container is isolated from the other containers of the collection network. It was seen in the literature review that the more remote a container is, the more logical is to have it monitored, in order to prevent the cost of long distances being travelled unnecessarily (Liu et al., 2011; McLeod et al., 2013, 2014; Straightsol, 2014a, 2014c). So, the value of information increases as the degree of remoteness of a container increases. From Figure 11, it can be understood how a sensor in F can lead to more costly trips being saved than if it was placed in container E, considering everything else equal and that the remaining network of containers is already behind the truck.

In order to evaluate how much a container is isolated from the rest of the group, several measurement units can be considered. Here, two different ways will be used – it may be interesting to try to understand if one leads to better results or not.

Furthermore, when working with distances, it will be used the distance by road (real distance). This implies that the distance from a point A to a point B can be different from the distance in the opposite direction ($d_{A,B} = d_{B,A}$ is not mandatory). So, both the distances from ($d_{A,B}$) and to ($d_{B,A}$) a container will be considered.

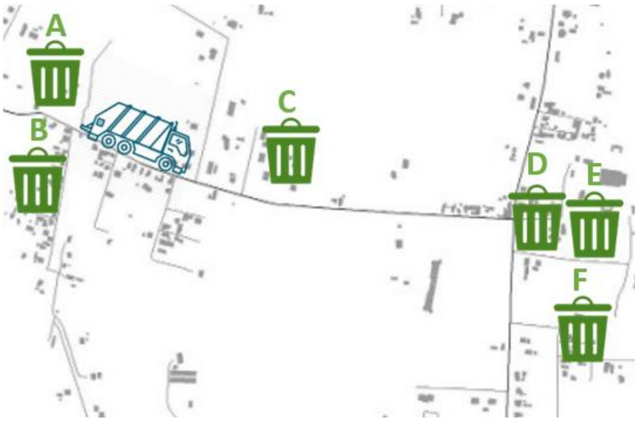


Figure 11: Context for the criterion remoteness of container.

A first way of analysing the remoteness can be by considering the distance to the closest container (option 1): the higher the distance to the closest container, the more remote a container is. Considering the locations $\{0, 1, \dots, \mathrm{n}\}$, where location 0 is the depot and *n* the is the number of containers, the remoteness of container *i* is evaluated in the following manner:

$$r_i^{\,1} = \min\left(\left\{\frac{d_{i,0} + d_{0,i}}{2}, \dots, \frac{d_{i,n} + d_{n,i}}{2}\right\} \backslash \{d_{i,i}\}\right) \qquad i = 1, \dots, n \qquad (5)$$

However, with this option, the remoteness of the container is not taking into account the remaining containers, because one container can, on average, be closer to the whole group and still have the higher distance to the closest container. As it can be seen in Figure 11, and considering that the remaining network of containers is already behind the truck, container C has the greater distance to the closest container, but, on average, is closer to the whole group than, for example, container F.

Therefore, it will also be studied a second option, where the remoteness of a container will be determined by calculating the average distance to all the other locations on the map (remaining containers + depot). In Vonolfen et al. (2011), the same approach is used to measure the remoteness of a site, but the authors didn't specify if they were using the Euclidean distance or the road distance. Considering road distances, option 2 will measure the remoteness of a container *i* in the following way:

$$r_i^{\,2} = \frac{1}{2n}\left(\sum_{j=0, j\neq i}^{n} d_{i,j} + \sum_{j=0, j\neq i}^{n} d_{j,i}\right) \qquad i = 1, \dots, n \qquad (6)$$

For both options, the higher the value of r_i, the more a container is preferable to be selected, according to its remoteness.

4.1.5. Method 4 – Cost Reduction Algorithm

4.1.4.1. Degree of how much "intermediate" a container is

Method 4 will indirectly apply some of the criteria studied. Besides considering the criterion remoteness of a container, this method will attempt to also take into account how much intermediate a container is – a priori, containers that are “more intermediate” will have less chances of being selected by this algorithm. Therefore, this criterion is here explained.

It was concluded that it can be advantageous for dynamic routing practices to collect the containers that are not near full capacity when the collection trucks will inevitably pass by them (Mes, 2012; Omara et al., 2018). Hence, in case such practice is carried, the value of information decreases for “very intermediate” containers, i.e., containers which are more likely to be on the paths to other containers. As can be seen in Figure 12, and if the waste demand behaves in a similar way for the different locations, monitoring A would be preferable than monitoring B, since it is likely that the collection vehicles will pass through B more often in order to reach C, D and E and thus the sensor wouldn’t be so relevant.

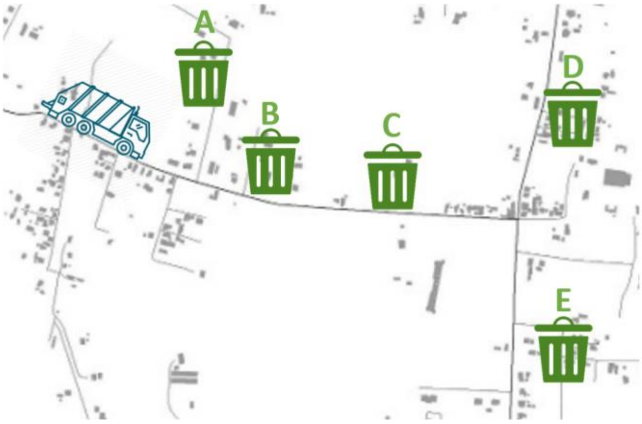


Figure 12: Representation of the criterion related to “intermediate” containers.

4.1.4.2. Description of the Cost Reduction Algorithm

In order to select the location of the sensors, it’s reasonable to make this decision based on the fact that embedding a sensor in different locations might lead to different cost reductions on the total path necessary to visit the locations, and thus it would be more desirable to select the location which provides the larger cost reduction. This algorithm follows the logic mentioned in subchapter 3.2 by estimating how much a sensor in a container would allow to reduce the collection costs – based on how it impacts the collection routes – and provides a ordered list that presents first the ones with a higher contribution.

Consider the following starting situation:

- a depot location *D* and *n* container locations (all without sensors)
- a cost matrix between all locations m_{ij}
- a fixed collection route for visiting locations without sensors, that costs C_{fixed}.

To know in which order the containers should be monitored, we can see, at each iteration *t*, which is the location *i* that when is “removed” – meaning that it turns to be a monitored container and is removed from the fixed collection route – will allow to obtain a collection route for the other (*n* – *t*) non-monitored locations, so that the difference $C_{fixed} - C_{n-t_i}$ is the maximum, for $i = 1, \dots, n+1-t$. In Figure 13 is presented the flowchart of the algorithm’s functioning.

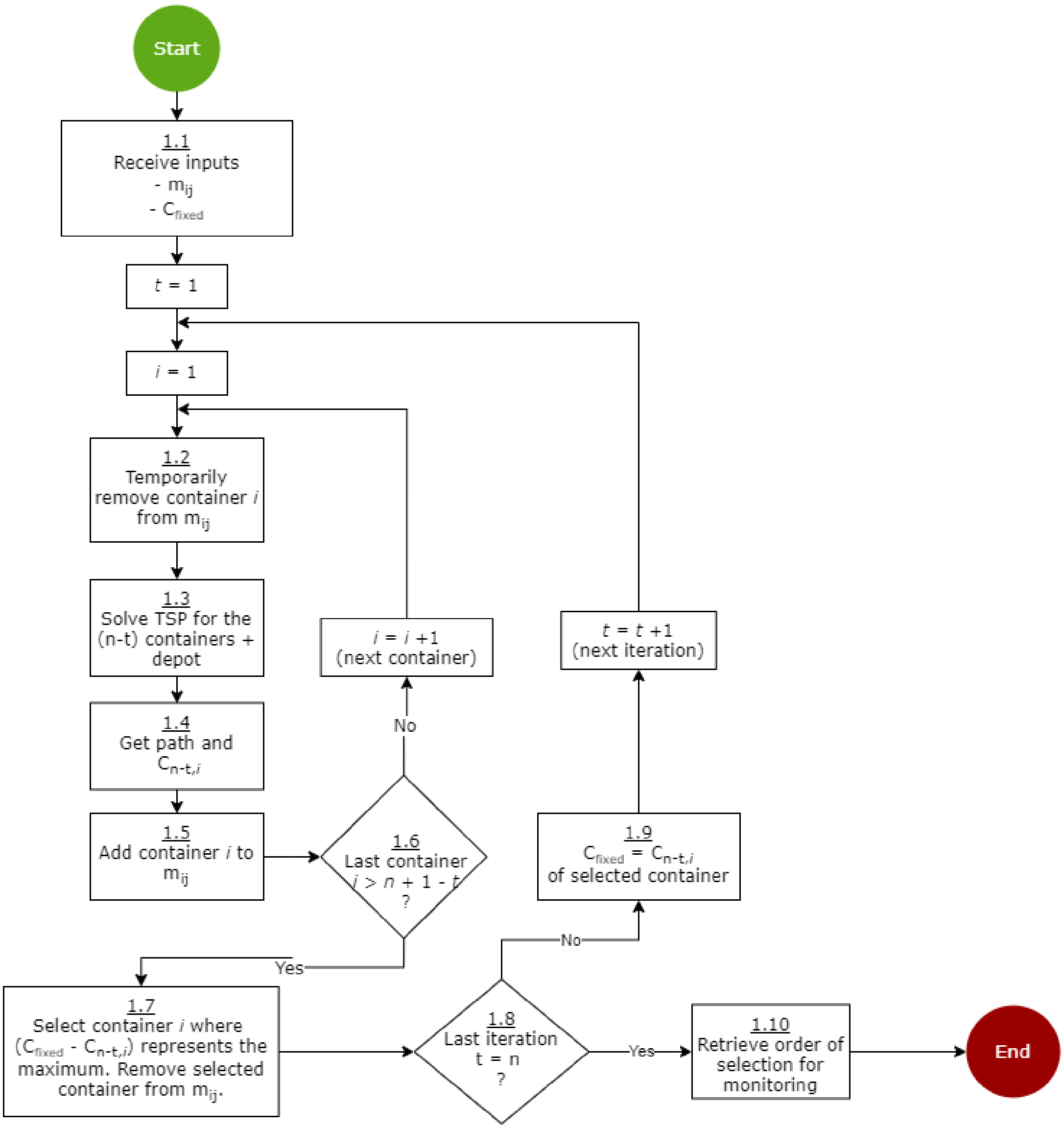


Figure 13: Flowchart of the cost reduction algorithm.

For example, in order to choose where to place the first sensor (*t* = 1), the location *i* that will be selected is the one that maximizes $(C_{fixed} - C_{n-1_i})$, with $i = 1, \dots, n$, meaning that it is the location that when is not included in the fixed collection route leads to it having the lowest cost (allows a higher cost reduction). It can be interpreted as the distance that is possible to avoid when the sensor gives the information that the container does not need to be visited.

In Figure 14 is presented a simplified scenario, where the node 0 represents the depot and the other 5 nodes represent the container locations. The optimal total cost of visiting all containers and returning to the depot is 13.

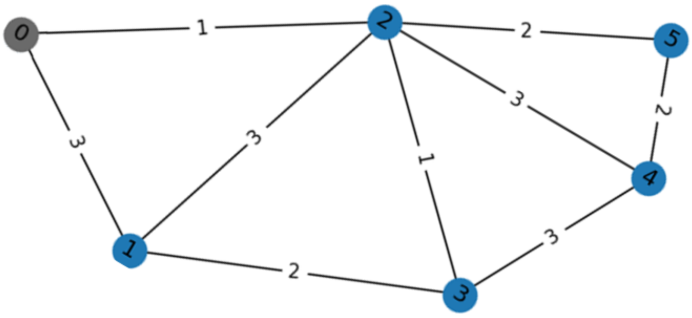


Figure 14: Graph of the small scenario example for the cost reduction algorithm.

With the purpose of determining in which order the sensors should be assigned on these container locations, the values of Table 2 would be considered. At iteration #1 (*t* =1), after solving the TSP five times (one for each location removal, *i* = 1,2,3,4,5) for four container locations plus depot, the algorithm selects location #1 (*i* =1), since it is the one that would allow the highest cost reduction. Next, *i* = 1 is removed and the algorithm continues with the remaining four locations and solves the TSP four times for three locations plus depot. By considering the path cost without location #1, the location that would allow now a higher reduction is location #4. By this method, the order in which the containers would be monitored would be: [#1, #4, #5, #3, #2]. So, similarly to what was explained, in the ending of section 4.1.1., for the direct application methods, if we wanted to assign sensors to, for example, 40% of containers (i.e., a sample of size *k* = 2), the locations #1 and #4 would be the ones selected to be monitored. It can also be seen how the location #2 was last in the order, since in all iterations except the last one, it never led to a reduction of the path cost given its “intermediate” location.

Table 2: The values of the path costs for the different container locations.

Iteration	Container removed	Path cost	Cost reduction
Initial Solution	-	13	-
#1	#1	10	3
	#2	13	0
	#3	13	0
	#4	11	2
	#5	12	1
#2	#2	10	0
	#3	9	1
	#4	8	2
	#5	9	1
#3	#2	8	0
	#3	6	2
	#5	4	4
#4	#2	4	0
	#3	2	2
#5	#2	0	2

When in comparison to the measurement units previously described for measuring the remoteness of a container (subchapter 4.1.4.), the logic of this method is, a priori, more correct, since a container with the higher distance to the closest container or with the higher average distance to all the other locations of the network may not be the one that when turns to be monitored enables the greatest cost reductions in the overall distances travelled.

Of course, there are some problems with selecting the containers in this fashion. First, as explained before, in subchapter 3.2, this problem is not submodular, so there's no guarantee that there wouldn't be a better way of selecting the containers than by starting with the one that provides the largest reduction. Also, after selecting a few sensors, in a dynamic collection policy, there could be some sensors that inform that a collection is required whereas there will be others that inform that it's not. This implies that it would be more correct to compute the decrease in cost reduction by estimating the expected value of cost reduction by considering all the possibilities – however, this would be an intractable combinatorial problem.

Finally, it is important to note that the TSP is being solved by using Google's OR-tools library (Google Developers, 2020). In order to find a solution is being used the provided variant of the Christofides heuristic (Christofides, 1976), meaning that the problem is not being solved in an optimal manner, which would be desirable but also very time consuming: in the first iteration it is solved *n* instances of the TSP with (*n* – 1) container locations plus depot, in the second it is solved (*n* – 1) instances with (*n* – 2) containers plus depot, etc. Finding the sensor locations with this algorithm by solving the instances of the TSP in an optimal manner would easily become impracticable when considering a real collection network. This may lead to the locations selected being not exactly the same ones that would eventually be chosen if the TSP was solved to optimality, since there is a dependency on the path that was found by the solver. However, it is likely that most of the locations that allow a higher cost reduction in optimal path will also allow a higher cost reduction in a less than optimal path, which safeguards the relevance of the results.

4.1.4.3. Cost Reduction Algorithm – Variant with collection frequencies

Within the previous algorithm, it is not being considered how much a container "needs to be visited". Given a period of time, the different containers will need to be visited a different number of times depending on the waste demand rates and variability. Following this, the number of times that a trip to a container can be avoided, in a given period *T*, is conditional on the number of trips that are expected to happen to that container in the same period.

With this, a variation of the previous algorithm will also be tested. Considering that each container *i* should be visited with a specific collection interval from x_i to x_i days – it will be explained in section 4.2.1 how this collection interval will be obtained for each container – we can calculate a frequency of visit $f_i = \frac{T}{x_i}$, for *i* = 1,…, *n*. Being *F* the max$(\{f_1, \dots, f_n\})$, and by assuming a fixed collection schedule that would collect with a frequency equal to *F*, the number of saved travels, during period *T*, for each container *i* with a sensor, would be equal to *F* – f_i. So, the cost reduction obtained in the previous algorithm is here multiplied by F – f_i in order to determine the container locations that allow a higher total cost reduction in a period *T*.

Considering, once again, the example from Figure 14, but now imagine that each container needs to be visited with the following frequency, as seen in Table 3.

Table 3: Small example with different frequencies of visit.

Container	x_i (days)	f_i, for *T* = 30	*F* - f_i, for *T* = 30
#1	15	2	2.5 – 2 = 0.5
#2	31	0.97	2.5 – 0.97 = 1.53
#3	12	2.5 (Max)	2.5 – 2.5 = 0
#4	22	1.36	2.5 – 1.36 = 1.14
#5	26	1.15	2.5 – 1.15 = 1.35

In this case, the previous Table 2 would now turn into Table 4. With this variation, the ordered list would be: [#4, #5, #1, #2, #3]. In the first iteration, even if the distance saved per trip is higher for container #1 than for container #4, the latter is chosen since a higher number of trips can be avoided and hence it leads to a higher total cost reduction. Note that when there's only containers #2 and #3 left, it would indifferent to choose either of them. Considering again, a sample of size *k* = 2, the selected locations would be #4 and #5, which is different than the ones obtained when the frequency is not considered (#1 and #4) – it may not seem much different since it is a small example, but it may lead to relevant differences when tested in a larger population of containers.

Table 4: The values of the cost reductions, considering frequency of visit.

Iteration	Container removed	Path cost	Cost reduction	Total cost reduction [cost reduction × (F – f_i)]
Initial Solution	-	13	-	-
#1	#1	10	3	3 × 0.5 = 1.5
	#2	13	0	0 × 1.53 = 0
	#3	13	0	0 × 0 = 0
	#4	11	2	2 × 1.14 = 2.28
	#5	12	1	1 × 1.35 = 1.35
#2	#1	8	3	1.5
	#2	11	0	0
	#3	11	0	0
	#5	7	4	5.4
#3	#1	4	3	1.5
	#2	7	0	0
	#3	7	0	0
#4	#2	4	0	0
	#3	2	2	0
#5	#3	0	4	0

4.1.5. Method 5 – Considering rate, variability and remoteness

The last method will consider the classification of each container, according to their behaviour in all the criteria that were presented in methods 1, 2 and 3. This is depicted in Figure 15. There it can be seen that, a priori, containers that are classified into octant 1 (in green) are expected to be the most critical (i.e., low demand rate, high variability and remote location) and therefore it makes sense that they should be monitored; and the ones in octant 8 (in red) are expected to be the less critical (i.e., high demand rate, low variability and non-remote location), and so they do not need to be monitored.

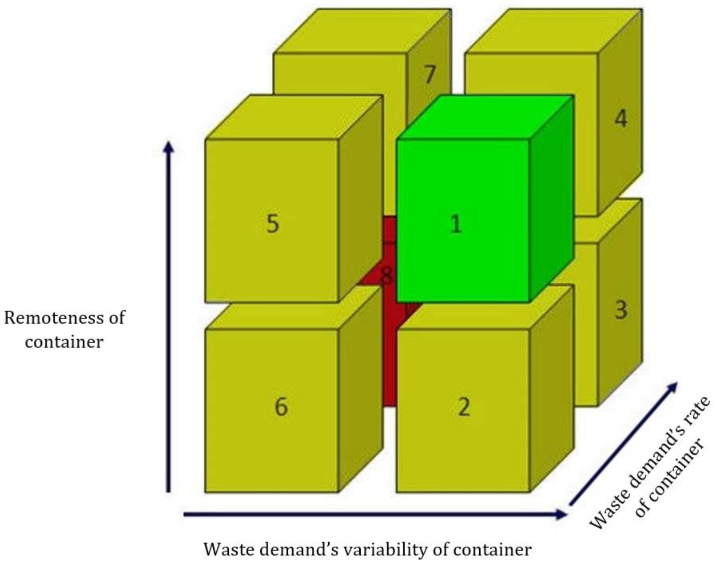


Figure 15: Three-dimensional classification of containers (adapted from Zhang et al., 2012).

By this method it is necessary define the limits that separate each octant, in each dimension. This can be done in an absolute manner by establishing a threshold value for each measurement unit of each criterion: For example, if l_{rate}, l_{var} and l_{rem} are the thresholds for the measurement units' values in the criteria waste demand rate, waste demand variability and remoteness, respectively; this means that a container C is considered to be in octant 1 if $\bar{x}_c \leq l_{rate}$, $\widehat{C_V}_c \geq l_{var}$ and $r_c \geq l_{rem}$. Or this can be done in a relative way by assigning a percentage value to each criterion (i.e., $x\%_{rate}, x\%_{var}$ and $x\%_{rem}$), and then a container C is considered to be in octant 1 if it's simultaneously in the $x\%_{rate}$ of containers that ranked higher in the criterion waste demand rate, in the $x\%_{var}$ of containers that ranked higher in the criterion waste demand variability, and in $x\%_{rem}$ of containers that ranked higher in the criterion remoteness.

After classifying each container in its respective octant, the selection of the containers to be monitored can take place. If selecting the containers of the 1st octant ought to be direct, for the ones in yellow (in Figure 15) further analysis is necessary. This means that when validating this method, several combinations of the octants to be selected must be evaluated. For example, we could first choose only octant 1 and then add the other octants individually in order to see which ones lead to the better results (e.g. octants 1 and 2; octants 1, 2, 4 and 5; etc.).

4.2. Description of the validation methodology

In order to validate the different methods for selecting the containers, it is necessary to test them in a simulation environment, where the day-to-day waste collection routes will have into account the filling levels of the monitored containers. So, firstly, it is given a broader and more generic description of what will be the validation methodology (4.2.1.). Afterwards, the dynamic collection policy adopted in this work, that intends to adapt the routes to the waste demand levels, will be described (4.2.2.): it will be provided the details on how the day-to-day collection activities will be simulated and tested.

The several methods for selecting the containers and the simulation environment where they will be tested were developed by programming in Python code.

4.2.1. Overview of the validation methodology

The diagram of the validation methodology is represented in Figure 16. There it can be seen the several steps that will be carried out in order to validate the several methods proposed.

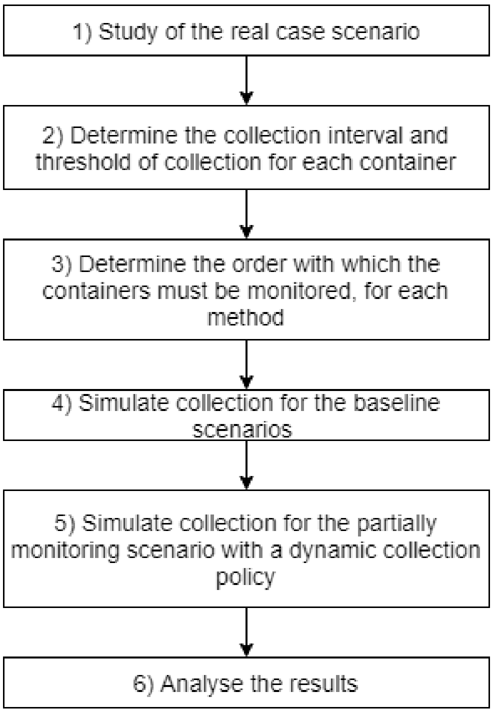


Figure 16: Diagram of the validation methodology.

1) **Study of the real case scenario:**

The validation methodology will be based on a real case. This step will be fully covered in chapter 5. A geographical area with fixed collection routes will be under study, where the author of this work will have access to historical data regarding the filling levels of the containers. These data will be used to determine the variance and average rate of each container's demand. Another sub-step involves the modelling of the waste demand into a continuous probability distribution, that will be necessary to represent the stochasticity of the simulations. Moreover, the real road network of the real case will be analysed, in order to obtain a matrix of costs and a graph that represent it, which are two inputs necessary for the dynamic collection policy – which is described in the next subchapter.

2) **Determine the collection interval and threshold of collection:**

In this step, for each container, it is calculated their specific collection interval and collection threshold level, based on a continuous probability distribution that will model the daily waste deposition, and on certain thresholds that intend to guarantee a decent level of service.

Threshold level: Being X_i a random variable that characterizes the daily deposition rate of a container *i,* and by considering the cumulative distribution function F_{Xi}: R → [0,1], there is a daily deposition rate threshold $(deposition_{limit_i})$ so that the probability of that container *i* fill-up in a day a quantity less than that deposition rate threshold is $\delta_{threshold}$. Meaning that we have:

$$\Pr(X_i \leq deposition_{limit_i}) = F_{Xi}(deposition_{limit}) = \delta_{threshold}$$
$$<=> \; deposition_{limit_i} = {F_{Xi}}^{-1}(\delta_{threshold}) = Q_{Xi}(\delta_{threshold}), \tag{7}$$

where Q_{Xi} represents the quantile function (or inverse cumulative distribution function). In this way, we can define a threshold level (ψ_i), so that when the container's filling level surpasses that value, the probability of that container not overflowing in the next day is less than $\delta_{threshold}$ (probability of overflowing is more than 1 – $\delta_{threshold}$). Then, ψ_i is given by:

$$\psi_i = 100 - deposition_{limit_i} \tag{8}$$

<u>Required Collection Interval</u>: Besides assuming a cumulative distribution function F_{Xi}: R → [0,1] of a random variable X_i that characterizes the daily deposition rate of a container *i,* it is also necessary to consider that the daily deposition rates are independent and identically distributed random variables, meaning that deposition rates of different days have the same characteristics and do not influence each other values (there is no memory of deposition rates generated in the past). That being said, if:

$$X_{i\ in\ any\ given\ day} \sim Continuous\ Probability\ distribution(\mu_i, {\sigma_i}^2) \tag{9}$$

Then we have, after b_i days:

$$X_{i\ after\ b_i\ days} \sim Continuous\ Probability\ distribution(b_i\mu_i, b_i{\sigma_i}^2) \tag{10}$$

and so, it is possible to calculate the number of days b_i that are necessary so that the probability of that container *i* being encountered with a filling level greater than its previous calculated threshold level (ψ_i) at day b_i is more than a certain value $\delta_{interval}$. Considering the survival function S_{Xi}, where $S_{Xi} = 1 - F_{Xi}$, the value of b_i can be calculated by discovering the largest integer value where the following condition is not verified:

$$\Pr\left(X_{i_{after\ b_i\ days}} > \psi_i\right) = S_{Xi_{after\ b_i\ days}}(\psi_i) \geq \delta_{interval} \tag{11}$$

The value of b_i that is calculated for each container is then their collection interval.

The calculated values will be displayed in the beginning of chapter 6.

3) <u>Determine the order with which the containers must be monitored:</u>

In this step, the methods described in subchapter 4.1 will be applied and obtained the ordered list of containers, that goes from the ones that must be the first to be monitored to the ones that can be monitored in last.

4) <u>Simulate day-to-day collection for the baseline scenarios:</u>

Here, it will be simulated the waste collection operations of several baseline scenarios that will be used as a basis of comparison for the partially monitoring scenario, such as: the current waste collection policy (blind collection with fixed collection routes) – baseline scenario 1; the case where all containers are being monitored (a dynamic collection policy with perfect information) – baseline scenario 4; a case where the current routes are maintained but their collection interval is increased – baseline scenario 2; and the case where the dynamic collection policy is applied but there is not a single sensor installed – baseline scenario 3. They will all be better depicted in the beginning of chapter 6.

5) **Simulate day-to-day collection for the partially monitoring scenario with a dynamic collection policy:**

Since with this work is intended to study if the benefits of a dynamic collection policy can be maintained, even if only a part of the population of containers is being monitored, the daily waste collection operations of the chosen real geographical area will be simulated as if sensors had been embedded in the selected containers and the routes are adapted to the monitored levels.

This study is performed for each method proposed and, in order to study the benefits for each method, the number of sensors to install will vary.

6) **Analyse the results:**

The costs, overflows and other relevant variables of the different scenarios are analysed and discussed.

4.2.2. The dynamic collection policy

To simulate the day-to-day waste collection routes that will result from monitoring only a pre-selection of locations it is needed to define a dynamic collection policy. It is proposed a simple stochastic dynamic collection policy where the performed routes are adapted to the waste demand by taking into account the filling levels of the monitored containers, and other aspects such as the historical data regarding the complete population of containers.

Therefore, with this dynamic collection policy:

- It is intended to simulate a period of collection operations, where part of the population of containers is being monitored and communicating the filling levels in real time;
- The decisions regarding when to collect must consider the knowledge we have about the filling levels of the monitored containers and must take into account the estimated needs for the containers without sensor. For these containers, it will be considered the number of days since the last visit and their specific collection intervals;
- The filling levels of the containers are determined by considering a stochastic function of the time, since last collection;
- Results such as the incurred costs and the number of overflows will be computed.

In Figure 17 is represented the flowchart of the algorithm behind this dynamic collection policy and in Table 5 the several steps are described (recommended to follow both at the same time).

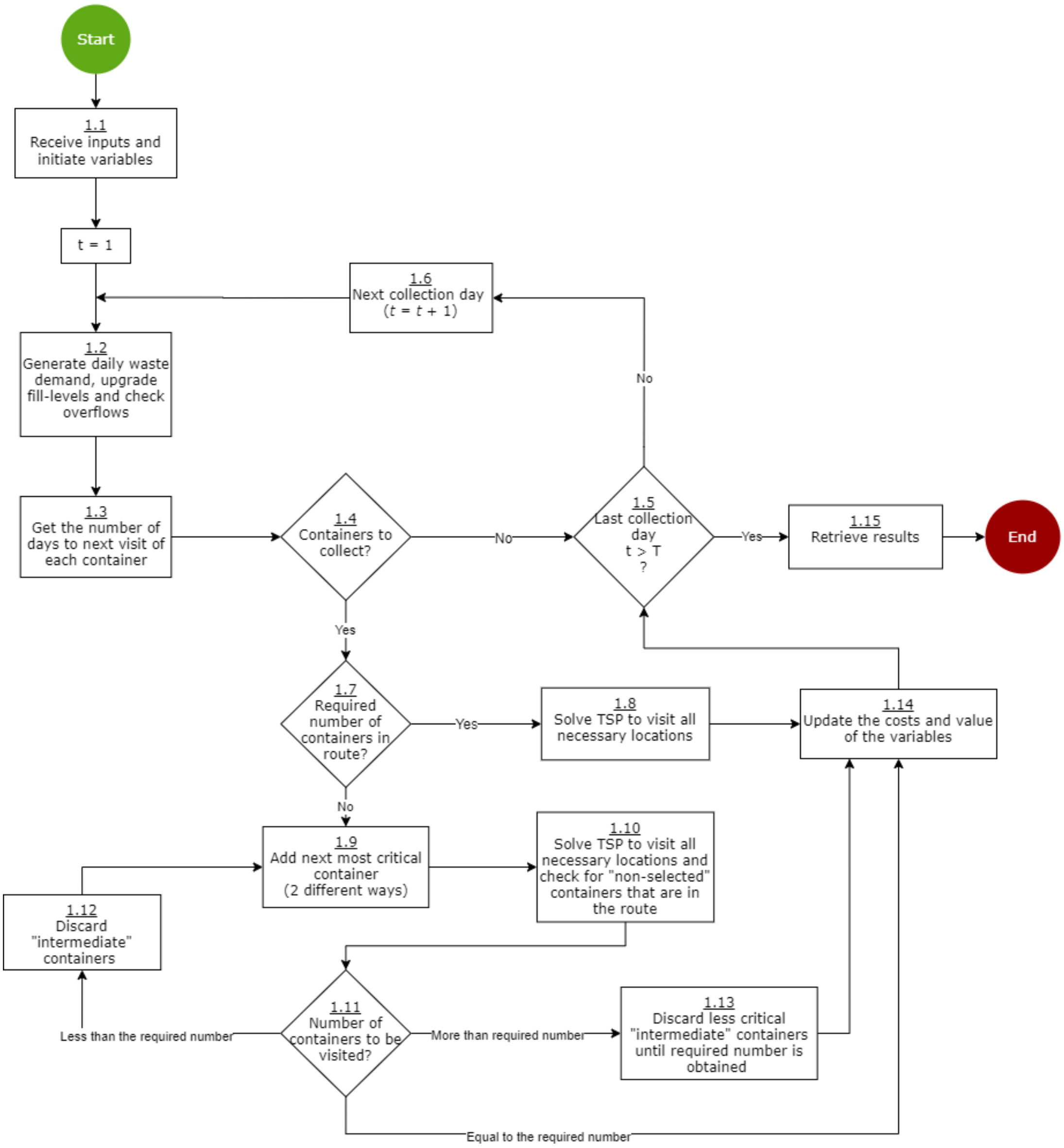


Figure 17: Flowchart of the dynamic collection policy algorithm.

Table 5: Steps of the dynamic collection policy algorithm.

Step	Description
1.1	The inputs are the matrix of costs between all locations (m_{ij}), the graph that represents the topology of the road network, the list of containers with sensor (S_L), the number of days of the simulation (T) and, for each container, the threshold level (ψ_i), the required collection interval (b_i), and the parameters that will allow to model the waste demand of each container into a continuous probability distribution. Each collection route will cover a required number of containers ($\text{number}_{\text{required}}$) that is adapted to the geographical area in question, which is also an input of the algorithm. The variables that will inform about the filling levels and count the cost, number of collections, number of trips, number of overflows, quantity collected and number of days since last visit are initiated. It is assumed that all the containers were collected in the previous day, so at the start of the simulation the filling level is 0 for all containers.
1.2	At first iteration we have $t = 1$. Based on the parameters of a continuous probability distribution, a daily deposition value is generated and added to the filling level. It is checked if there were any overflows – to note that it is assumed that the overflow quantity it is not collected, i.e., it is lost.

1.3	It is calculated for all containers in how many days each one of them must be collected (variable "$days\_To\_Next\_Visit$"): • For the containers *without* sensor: At each iteration, it is considered that a new day has passed since the last visit ("$last\_Visit$"). Since each container will have a required collection interval ("$days\_Collection\_Interval$") with which they must be visited, the number of days that have passed since the last collection is then subtracted to that collection interval, and thus, the next visit is calculated from there: $days\_To\_Next\_Visit = days\_Collection\_Interval - last\_Visit$ • For the containers *with* sensor: At each iteration, this decision is guided by the filling level of these containers. So, depending the measured filling-level ($current\_level$) and the historical data of those containers, it will be calculated a new time interval in which those containers need to be visited. This is similar to the way the required collection interval is calculated (see subchapter 4.2.1), but in this case, the current filling level is subtracted: $\Pr\left(X_{i_{after\ b_i\ days}} > \psi_i - current\_level\right) = S_{Xi_{after\ b_i\ days}} \geq \delta_{interval}$
1.4	Check if there is any container that has to be collected today (i.e., if $days\_To\_Next\_Visit = 0$).
1.5	Check if it is the last collection day, i.e., if $t > T$.
1.6	Go to next collection day ($t = t + 1$).
1.7	Here, it checked if the required number of containers to be collected ($\text{number}_{\text{required}}$) was reached.
1.8	The TSP instance is solved in order to visit all the necessary containers.
1.9	It is added the next most critical container. Two different ways of doing this will be tested: a) It is added the container that counts the minimum number of days until the next visit i.e., the min($days\_To\_Next\_Visit$). b) It is added the container where the ratio "days to next visit / required collection interval" has the lowest value, i.e., the min($days\_To\_Next\_Visit\ /\ days\_Collection\_Interval$).
1.10	The TSP instance is solved in order to simulate a visit to all the necessary containers. Here, it is also seen what are the locations that would be in the path, i.e., intermediate locations. If there exist these intermediate locations, they will be added temporarily to the list of locations to be visited.
1.11	Checks if the number of containers to be visited (including the intermediate containers) is less, equal or more the than required number of containers to be collected.
1.12	If the number is less, the intermediate containers are "forgotten" due to the fact they can stop being intermediate when the next most critical container is added, and a new route is determined.
1.13	If the number is higher than the required number, the intermediate containers that are the less critical (same evaluation as in step 1.9) are discarded, until the required number is obtained.
1.14	This step being reached means that there was a collection. So, the number of days since the last visit and the filling level of the collected containers are reset. The cost, number of collections and the total quantity collected are also updated.
1.15	Retrieves the number of collections, the number of trips, the number and quantity of overflows, quantity collected, and total costs incurred.

5. Study of a real-case scenario

The methods presented in the previous chapter will be validated on a real-case scenario. For that, the operations of ERSUC – Resíduos do Centro, S.A. in the municipality of Soure will be studied here. ERSUC is partner of this study and has shared several important data regarding their collection operations. ERSUC recommended to study the municipality of Soure since they consider the waste collectors from this municipality to be more reliable on registering the data values.

First, in section 5.1., the company ERSUC is briefly introduced, as well as part of the collection network that it operates in the municipality of Soure. Here, is also studied what are the current practices undertaken by ERSUC and an estimate of the costs of the current situation is given. In section 5.2., historical data, regarding part of the collection system is presented in order to study the demand characteristics of a population of containers. It is explained how these data were obtained, and how they are used to statistically model the waste demand's rate and variability of the containers.

Finally, in section 5.3., the geographic location of each container is also used in order to study the road network, i.e., the special facet of the problem.

5.1. Waste collection in Soure

5.1.1. Brief Presentation of ERSUC and Soure municipality

ERSUC – Resíduos do Centro, S.A. is a Portuguese waste management company responsible for exploring and managing a multi-municipal SGRU in the central west coast of Portugal, involving a total of 36 municipalities, covering 6700 km^2 (7.5% of the Portuguese continental area) and servicing a population of circa one million inhabitants. Nowadays, ERSUC possesses approximately 30 collection vehicles, has more than 13000 containers of separate collection, from where it collected 28513 tons of material, in 2018 (ERSUC, 2018a, 2018b, 2018c, 2018e). According to the 2018's financial statement report, ERSUC employs around 330 workers, holds assets superior to 87 million euros and has ended 2018 with a net result superior to 90000 euros (ERSUC, 2018a).

For the validation of the previously presented methods, it will be studied part of the waste collection system of ERSUC. In particular, it will be analysed the paper-cardboard waste collection of the municipality of Soure, which compromises 98 containers and a depot, and is represented in Figure 18. It can be seen that the containers are divided in two sub-groups that correspond to two different routes: one of 50 containers (let's denominate it group A) – blue markers in the figure –, and other of 48 containers (let's denominate it group B) – green markers. It's also possible to notice that the containers are spread out across, practically, the whole area of the municipality of Soure (the shaded area in the figure); and that the depot is considerably far from the area of collection (the black marker in the right top corner).

Figure 18: Network of paper-carboard containers in Soure.

As previously mentioned, in chapter 2, this work would focus in studying separate collection rather than studying mixed collection, since it was concluded that there is a greater urgency of updating the existing collection models in the former and is where the integration of ICTs is potentially more beneficial. This is mainly due to the low deposition rates that this type of collection suffers from. It was also noted that this is even truer for areas with low population density, such as rural areas, where the deposition rates are even lower, and the network of containers is more disperse.

The municipality of Soure fits in this description. Most of the area of Soure is predominantly rural. The population density of Soure is 72,9 people per km^2, a value below the national average of 111,5 people per km^2, and much lower than the values of urban centres like Lisbon where the population density is of 5066,4 people per km^2 (INE, 2012; PORDATA, 2019a).

Furthermore, by analysing the ERSUC's financial statement report of 2018, it is possible to see that, in 2018, 408.7 tons of separated waste were collected in the municipality of Soure, during separate collection. This value corresponds to an average of 22.6 kg being produced annually per inhabitant of Soure, which is less than the ERSUC's average for the 36 municipalities (30.9 kg per capita), which in its turn, is less than value for the Portuguese average (103.5 kg per capita, in 2018) (ERSUC, 2018a; PORDATA, 2019b).

In the case of the paper-cardboard waste stream, in 2018, these represented 29.4% of the separated waste collected by ERSUC in Soure, i.e., a total of 120.1 tons or an average of 6.62 kg per citizen of Soure (ERSUC, 2018a).

5.1.2. Current waste collection policy

According to ERSUC, separate collection is organized by circuits, for each material; and each circuit is performed with a pre-established frequency that is said to have in consideration the number of containers of the circuit and their waste production (ERSUC, 2018d).

Both groups A and B correspond to two different circuits, i.e., the collection of each group is not done in the same route nor in the same schedule. For the development of this work, ERSUC provided several data regarding the 98 containers of paper-carboard. In Table 6 some collection statistics of the year 2017 are presented for both groups.

Without sensors, in order to know the quantities of waste collected, a common approach is to weight the waste collection vehicles after the collection and with this obtain the collected quantity per route performed – the quantity related data from Table 6 was obtained by ERSUC in this fashion.

Table 6: The collection of the 98 paper-cardboard containers, in 2017.

	Group A	Group B
Number of containers	50	48
Number of collections	52	53
Average period between collections (days)	7.0	6.9
Maximum period between collection (days)	14	11
Minimum period between collection (days)	3	4
Average collection time	6h14m	6h25m
Average circuit distance (km)	138.6	174.5
Average collected quantity (kg)	1234.2	752.8
Average collected quantity per container (kg/container)	24.7	15.7

Even if the period between collections can vary within each group, on average, the containers are collected once per week, resulting in 52 and 53 collections per year. By viewing the locations of the containers of the two groups in Figure 18, it's reasonable to think that ERSUC would benefit from having only one route for the whole collection. However, in that scenario the maximum collection time would be exceed for the collection crew, which performs shifts of 6 hours and 40 minutes maximum – note that the average collection time is close to that value but does not exceed it. In this way, two routes are required due to this legal restriction. Other reason that could force the inclusion of two routes would be the collection trucks' capacity not being enough to collect all the containers in one route. According to the data of 2017, there were a few times that the ERSUC's truck almost reached full capacity (2200 kg), and only for the collection of the containers of group A. If there was not a restriction on the collection time, this factor could also imply the necessity of two routes.

It's also possible to see that the collection for group A is more efficient than for group B, since, on average, less kilometres are needed to be travelled to collect more quantities of material. This may be due to the fact that the containers of group A are less spread, are closer to the depot, and are mainly in areas with more population (e.g. centre of Soure). According to the previously identified criteria, an a priori assumption would be that the containers of group B are more appropriate for having a sensor.

Furthermore, the containers used by ERSUC are like the one seen in Figure 19. These are surface containers with a volume capacity of 2.5 m^3(Gonçalves, 2014)(Gonçalves, 2014)(Gonçalves, 2014). Values for the density of the paper-carboard material within a container were empirically obtained and provided by ERSUC: 30 $\mathrm{kg/m}^3$, in average. A very similar value was also adopted in Ramos et al. (2018) (29,5 $\mathrm{kg/m}^3$) – this was also provided by the waste collection company they partnered in their work, Valorsul. This means that the average collected quantity per container, for both groups, is far below from what a container could contain ($2.5\mathrm{m}^3 \ \times 30\frac{\mathrm{kg}}{\mathrm{m}^3} = 75\ \mathrm{kg}$).

Figure 19: The paper-cardboard container (adapted from Vila Nova de Poiares, 2017)

Both routes (A and B) can be found represented in Appendix A. The distance travelled for each route was obtained by using an app developed under the WSmart Route research project (WSmart Route, 2019). This app allows to determine the matrix of the road distances that exist amongst the containers – it is only necessary to provide, as an input, an Excel sheet with the GPS coordinates of each one –, and so, it is possible to obtain the distance of the performed routes. That being said, the route distance, the collection time and the cost per route were then estimated for each of these groups, as shown in Table 7.

Table 7: Estimated distance, costs and collection times of routes A and B.

	Group A	**Group B**
Total distance route (km)	133.7	179.6
Cost per route (€)	133.7	179.6
Collection Time	6h07m	6h33m

For the route costs was considered the information given by ERSUC: the current cost per travelled kilometre is estimated at 1 €/km and includes fuel consumption, the maintenance of the vehicle and the salary of the collection crew.

For calculating the collection time it was considered an average speed of 50 km/h when the vehicle moves towards or from the depot and it was considered an average speed of 30 km/h for group A and of 40 km/h for group B, when the vehicle moves between containers – the reason for this distinction is the fact that the containers of group A are closer to each other and are mainly located in populated areas (e.g. villages of Soure and Alfarelos) where the velocity needs to be reduced. An average of 3 minutes was considered for the collection manoeuvres. For example, for group A:

- The distance travelled from and to the depot is equal to $d_{0,1} + d_{50,0} = 62.5$ km, which is travelled at 50 km/h.
- The distance travelled amongst the containers is 133.7 – 62.5 = 71.2 km, which is travelled at 30 km/h.
- There are 50 containers.

So, we have:

$$\text{Collection time (A)} = \left[\left(\frac{62.5}{55} + \frac{71.2}{30}\right) \times 60 + 50 \times 3\right] / 60 \;\cong 6.12\text{h} = 6\text{h}07\text{m}$$

The estimated distances and collection times for the routes are slightly different in comparison with the average values taken from the data of 2017 in Table 6: the estimated time is a bit lower for route A and a bit higher for route B. For route A, this may lie in the fact that the path taken by the collection vehicles may not be exactly the one that matches the shortest distance between each container; or, in case of route B, it’s recorded that there were a few days when the vehicles of ERSUC didn’t collect all the containers of that group, which could lead to lower the average value. For the purpose of this work, it will be considered that for both routes the collection of ERSUC is always performed to all the containers and the path undertaken between two containers matches the shortest path between those two containers.

Under the circumstances, and considering that a year has exactly 52 weeks and that the collection happens one time per week, the yearly costs that ERSUC allocates to the separate collection of paper-cardboard in the municipality of Soure are: $(133.7 + 179.6) \times 52 = 16291.6$ €. Assuming the average values from Table 6 $(138.6 + 175.5) \times 52 = 16333.2$€, it can be considered that this estimation is fairly adequate.

5.2. Study of the individual demand

In this subchapter, it will be used data provided by ERSUC regarding the 98 containers of paper-carboard. This includes information about the individual demand that each container experienced during a period of almost four months. It will be explained how the data were obtained, and how it was subsequently investigated and handled in order to obtain the waste demand’s rate and variability for each container. In the last part of this subchapter, it is explained how the data will be modelled through a probability distribution in order to carry on with the analysis of the municipality of Soure.

5.2.1. The gathering of each container’s data

It is common that no data is gathered relatively to how much a specific container was filled at a particular collection (it was noted that a common approach is to only weight the vehicles after the collection). However, it was seen through the course of this work that having knowledge about the waste demand characteristics of each container can be very important if we want to study in which containers the remote monitoring can be more beneficial.

In order to overpass this problem, between April 1st and July 23rd of 2019, it was asked to the waste collectors of ERSUC to manual register the amount of waste inside a container at the moment just before its collection would take place; and to register also the amount inside the containers of the other materials that could be encountered in the same site, i.e., if, for example, the collection crew is collecting glass-waste containers, they would take a look into the containers of paper-carboard and of plastic-metal packaging wastes that would be on the same site, and register the waste level.

This method involves analysing the content of the container in a visual manner, and is, of course, subjective and can only be performed in an approximate way. Nevertheless, this does not constitute a problem, since it is not that important to precisely specify the amount of waste inside the container, but just to have a rough estimate of the waste level. Therefore, the worker registering the waste quantity would classify the container into the following values and intervals, constituting a Likert scale with six values:

- Container is empty (0%) → A value of 0% is assumed.
- Waste-level is between 0 and 25% → A value of 12.5% is assumed.
- Waste-level is between 25 and 50% → A value of 37.5% is assumed.
- Waste-level is between 50 and 75% → A value of 62.5% is assumed.
- Waste-level is between 75% and 100% → A value of 87.5% is assumed.
- Container is completely full (100%) → A value of 100% is assumed.

The same procedure was also adopted in Marchiori (2018), but in that work the data manually registered by the waste collectors was used to validate the measurements performed by the sensors, during a testing phase of four weeks. A linear correlation of 0.96 was found between the sensors data and the manually collected data, in that study.

By the end, the waste collectors of ERSUC, were responsible for registering a total of 3491 entries regarding the waste-levels of the 98 containers, during the period in analysis. The data collected were not without some inconsistencies, such as days where some containers didn't have an entry, but others did; or when a subsequent day would have a lower filling level than a previous day, with collection having not occurred in between. This incongruity could be due to human error (distraction or a "difficult-to-analyse" situation), or due to the fact that waste posteriorly thrown into the container could trample the materials already inside – which is relatively more likely to happen in the case of paper-cardboard wastes. Despite this, these situations represent a minority and workarounds were properly carried out when dealing with the data.

5.2.2. The treatment and analysis of the data

The purpose of studying these historical data is to find the values of the waste demand's rate and variability of each container. This is required because these aspects are object of analysis in two of the proposed methods, but also since it is necessary to simulate the results of applying the methods, and thus these parameters are used to attribute a continuous probability distribution to each container for modelling their waste generation. Therefore, the sample mean, the standard deviation (and variance) and the sample coefficient of variation of the waste demand will be found for each container, using the collected data.

The records made by the waste collectors are the raw data. These served to identify the waste level of a container, in percentage (%) of the total container, on a specific day. An elucidative example is presented in Table 8. Here, the records of a period of 8 days are presented for 12 containers. In this period, the containers were collected two times: in 01/04 and a week later, in 08/04. Not only the level of waste was annotated by the collection crew on these days (just before being collected), but also in a day between (04/04) by the team that was responsible for collecting the glass or plastic/metal packaging wastes on that site. It can also be seen that there is one entry missing for the waste level of C8 (in yellow), and that for container C4 it was registered a higher volume of waste for a previous day (37.5% in 04/04) than for a subsequent day (12.5% in 08/04) (in orange).

Table 8: Raw data of the waste-levels (%) measurements – example

	01/04/2019	02/04/2019	03/04/2019	04/04/2019	05/04/2019	06/04/2019	07/04/2019	08/04/2019
	Collection			Check				Collection
C1	12.5			12.5				12.5
C2	12.5			12.5				62.5
C3	12.5			12.5				37.5
C4	12.5			37.5				12.5
C5	37.5			12.5				37.5
C6	87.5			12.5				87.5
C7	37.5			62.5				62.5
C8	62.5							37.5
C9	12.5			12.5				62.5
C10	62.5			37.5				62.5
C11	37.5			12.5				37.5
C12	37.5			0.0				37.5

The raw data is then processed. The next, intermediary, step is to determine the containers' filling levels for all the days of the presented period. For this, two assumptions are made:

- The filling level grows linearly between two dates without entries, i.e., grows linearly for the days when there is no data input;
- The waste level is registered in the beginning of the day. So, the waste level of a container on the day *t* is the waste level of that container on the day *t* – 1 plus the waste deposited on the day *t* – 1:

$$WL(C)_{day\ t} = WL(C)_{day\ t-1} + \mathrm{deposition}(C)_{day\ t-1} \qquad (12)$$

Therefore, different filling-up rates are calculated for the two periods. In this example, the waste-level of 04/04 was used to fill the dates between 01/04 and 04/04, and the waste-level of 08/04 was used for the period between 04/04 and 08/04. Giving container C2 as example:

➔ <u>In the first period</u>: The level of 12.5% in 04/04 was deposited linearly in the three previous days. Thus, in these 3 days, the container filled-up a volume equal to $\frac{12.5}{3} \cong 4.2$, per day.

➔ <u>In the second period</u>: There's now four days of interval, and so, the incremental volume in this period is $\frac{62.5-12.5}{4} = 12.5$.

For the container C8, where there was an entry missing, it was considered the value of the day 08/04 to determine the remaining values for the period of 01/04 to 08/08 ($\frac{37.5}{7} \cong 5.4$). For the situation in C4 is assumed that the latest value to have been registered is the correct one (the one from day 08/04).

With this, the estimated waste-levels can be encountered in Table 9, and the estimated daily deposition rates can be found in Table 10, where the average and the standard deviation for the period of this example are also presented.

Table 9: Estimated waste-levels (%) for all the days of the period – example

	01/04/2019	02/04/2019	03/04/2019	04/04/2019	05/04/2019	06/04/2019	07/04/2019	08/04/2019
	Collection			Check				Collection
C1	12.5	4.2	8.3	12.5	12.5	12.5	12.5	12.5
C2	12.5	4.2	8.3	12.5	25.0	37.5	50.0	62.5
C3	12.5	4.2	8.3	12.5	18.8	25.0	31.3	37.5
C4	12.5	4.2	8.3	12.5	12.5	12.5	12.5	12.5
C5	37.5	4.2	8.3	12.5	18.8	25.0	31.3	37.5
C6	87.5	4.2	8.3	12.5	31.3	50.0	68.8	87.5
C7	37.5	20.8	41.7	62.5	62.5	62.5	62.5	62.5
C8	62.5	5.4	10.7	16.1	21.4	26.8	32.1	37.5
C9	12.5	4.2	8.3	12.5	25.0	37.5	50.0	62.5
C10	62.5	12.5	25.0	37.5	43.8	50.0	56.3	62.5
C11	37.5	4.2	8.3	12.5	18.8	25.0	31.3	37.5
C12	37.5	0.0	0.0	0.0	9.4	18.8	28.1	37.5

Table 10: Estimated daily waste demand rates, average and standard deviation (%) – example

	01/04/2019	02/04/2019	03/04/2019	04/04/2019	05/04/2019	06/04/2019	07/04/2019	08/04/2019		
	Collection			Check				Collection	Average	St. Dev.
C1	4.2	4.2	4.2	0.0	0.0	0.0	0.0	-	1.79	2.23
C2	4.2	4.2	4.2	12.5	12.5	12.5	12.5	-	8.93	4.45
C3	4.2	4.2	4.2	6.3	6.3	6.3	6.3	-	5.36	1.11
C4	4.2	4.2	4.2	0.0	0.0	0.0	0.0	-	1.79	2.23
C5	4.2	4.2	4.2	6.3	6.3	6.3	6.3	-	5.36	1.11
C6	4.2	4.2	4.2	18.8	18.8	18.8	18.8	-	12.50	7.80
C7	20.8	20.8	20.8	0.0	0.0	0.0	0.0	-	8.93	11.14
C8	5.4	5.4	5.4	5.4	5.4	5.4	5.4	-	5.36	0.00
C9	4.2	4.2	4.2	12.5	12.5	12.5	12.5	-	8.93	4.45
C10	12.5	12.5	12.5	6.3	6.3	6.3	6.3	-	8.93	3.34
C11	4.2	4.2	4.2	6.3	6.3	6.3	6.3	-	5.36	1.11
C12	0.0	0.0	0.0	9.4	9.4	9.4	9.4	-	5.36	5.01

By applying this procedure to the whole period of observation, is possible then to calculate what was the average of the daily waste demand rates, as well the measures related to the variability, that each container experienced. The values for the 98 containers can be found in Appendix B. In Table 11 it is presented the average value of the mean and of the coefficient of variation of each container, and the maximum and minimum values. It's also displayed the average values for groups A and B.

Table 11: Daily waste demand rates - average, max and min values of the mean and coefficient of variation

	Mean ($\bar{x}$)		C. Variation ($\widehat{C_V}$)
	in percentage	in weight	
Max	13.14 %	9.9 kg	2.36
Min	1.90 %	1.4 kg	0.57
Average	5.39 %	4 kg	1.33
Average Group A	5.62 %	4.2 kg	1.25
Average Group B	5.14 %	3.9 kg	1.42

Moreover, it is possible to note again, from the individual data collected in this period, that the containers of group B seem more appropriate for having a sensor. As shown in Table 11, these have a lower daily deposition and experience more variability.

Finally, in Table 12, it's also presented the number of times that the collection occurred for each waste level (intervals of 25%), in total and for both groups A and B. During this period, the collection was performed 17 times for each circuit, and a total of 1666 collection manoeuvres were applied to the containers. Half of the times, containers were collected at the lower interval (a waste level between 0 and 25%), and only one tenth were collected at the higher interval (between 75 and 100%). This reinforces the claim that sensors could help the separate collection of Soure to be more efficient. As expected, the containers of group B showcase worse results than the containers of group A.

Table 12: Frequency of each waste level at collection, in the 2019's studied period.

	Total		Group A		Group B	
% volume	**# collections**	**percentage**	**# collections**	**percentage**	**# collections**	**percentage**
[0;25]	851	51%	399	47%	452	55%
]25;50]	329	20%	182	21%	147	18%
]50;75]	301	18%	157	18%	144	18%
]75;100]	185	11%	112	13%	73	9%
Sum	1666	100%	850	100%	816	100%

5.2.3. Statistical modelling of the waste demand

5.2.3.1. Selecting the probability distribution

As explained before, in subchapter 4.2.2, it is intended to carry out simulations by using stochastic inputs in order to test and validate the methods proposed. To accomplish this, it is necessary to specify what is the continuous probability distribution of these stochastic inputs, i.e., the quantities of waste deposited daily at each container. Therefore, the purpose of this subchapter is to use the historical data previously studied to model the waste demand through a probability distribution.

First, it is necessary to understand that there will never be a probability distribution that represents exactly a set of data. There is, though, distributions that are accurate enough for the intended purposes of a given model (Law, 2011). Furthermore, it's important to notice that, in this case, trying to represent exactly the processed data could be a mistake, since the data themselves do not represent the reality – for example, a linear grow was assumed to happen between two consecutive registers (leading to a lot of values getting repeated), which is unlikely to be the case. Also, once again, the intent of these steps is only to provide a rough estimate of the real values.

With all this in mind, and since this is not the core of the current work, it will be not given too much depth on how to select the "correct" continuous probability distribution – for that, see Law (2011). Also, it will be given preference to the adoption of a more standard statistical methodology to the detriment of scrutinizing more complex (and possible more adequate) non-standard methodologies.

First thing to notice is that the waste generated daily at each container, cannot assume ever negative values and, by such reason, a normal distribution would fit very badly. Since the waste demand rates have, in average, values very close to zero, this aspect can't be neglected – otherwise negative values would be generated in the simulations with some frequency. Thus, the natural next choices are either the lognormal distribution or the gamma distribution, which assume only positive values and are capable of representing the positive skewness (i.e., asymmetric probability distribution around the mean with the “tail” to the right), so typical of these type of data sets (see Figure 20). The gamma distribution will be used instead of the lognormal, because it is easier to handle analytically (Linhart, 1965). Other option would be to consider the beta distribution, which could be used to limit the maximum value of waste generated daily. However, a maximum value will not be considered, and even if the collection crew didn’t register overfull containers (waste-level values > 100%), in the simulations that will not be considered impossible. Still, note that the question really being here discussed is if a container should be allowed to fill-up more than 100% in a single day. Due to the low deposition rates of most containers, that it will be extremely unlikely, and so this restriction will not be considered.

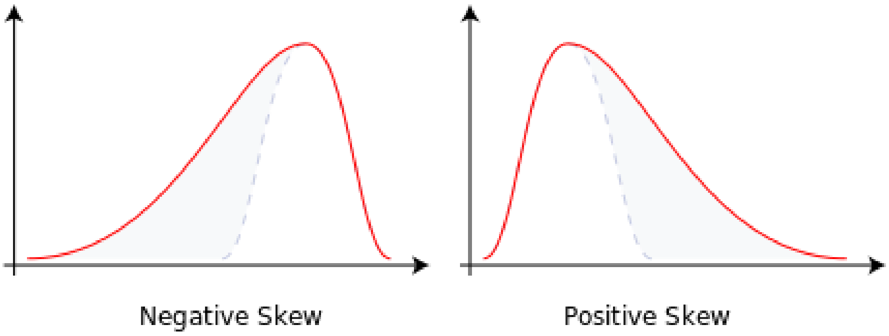


Figure 20: Negative and positive skewness of a probability distribution (adapted from ION Data Services, 2020).

One aspect that the standard gamma distribution does not suffice, is the fact that within the set of the collected data there are many zeros, which correspond to the days when no deposition happened. As said before, a gamma distribution only considers positive values, and so having zero values would immediately rule out the possibility of this distribution being suitable for modelling these data. However, this does not constitute a problem. Even though it is certainly the case that there are days when no deposition happens, the number of zeros may have been inflated due to the data registering procedure undertaken. Besides, it won’t affect the simulations results if extremely small values for the daily deposition are generated instead of the value zero, and thus the gamma distribution is still appropriate.

5.2.3.2. The gamma distribution

The gamma distribution is a positive continuous probability distribution. Three different parametrizations exist, but in this work is considered the one that makes use of a shape parameter k and a scale parameter θ. Both parameters must assume positive values. A gamma-distributed random variable X with shape k and scale θ is denoted:

$$X \sim Gamma(k, \theta) \tag{13}$$

In Figure 21 it is possible to see how the probability density function (PDF) varies with the values of the parameters k and θ. For values of $k \in]0;1]$, the curve of the PDF assumes a shape of an exponentially decreasing function; and for $k \in]1;+\infty[$, the function stops being monotonically decreasing, and starts resembling an “asymmetric bell-shaped” curve. As the $k \to +\infty$, the gamma distribution is identical to the normal distribution in shape.

The mean of the gamma distribution, μ, is:

$$\mu = k\theta \tag{14}$$

Whereas, the variance, σ^2, is calculated by:

$$\sigma^2 = k\theta^2 \tag{15}$$

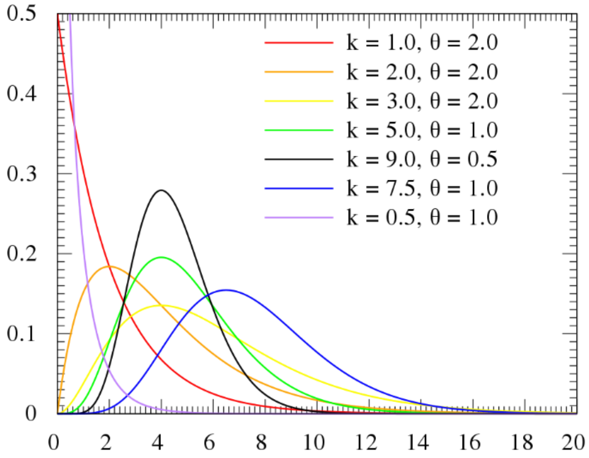


Figure 21: Probability density functions for different parameters k and θ (source: Wikipedia, 2020).

5.2.3.3. Determining the parameters and modelling the waste demand

Following (13), the deposition of waste in a certain container C, in a day *t*, is considered to be a gamma-distributed random variable, and is denoted:

$$\text{deposition}(C)_t \sim Gamma(k, \theta) \tag{16}$$

From equations (14) and (15), and by using the previously calculated values for the mean and variance of the waste demand rates (Appendix B), it is possible to determine the parameters k and θ, for modelling the gamma distribution of each container.

According to Law (2011), creating a histogram of the data, is one of the best ways for determining the shape of the underlying probability density function, since it is basically a graphical estimate of the density. However, according to the author, one problem of making a histogram consists in choosing the interval width *w*, and so it is recommended to use the smallest value interval width *w* that gives a reasonably “smooth” histogram. In Figure 22, for container C1 with k = 0.74 and θ = 6.39, it is presented its PDF, and its histogram that shows the number of times each set of values, for the daily deposition rate, occurred over the 112 days of the observation period. By analysing both the PDF and the histogram it is reasonable to confirm the fitness of the gamma distribution for modelling the collected data. Two more examples are provided in Appendix C. For all the three cases, the width *w* was determined automatically by the Microsoft Excel tool, since it appeared to provide the smoothest histogram.

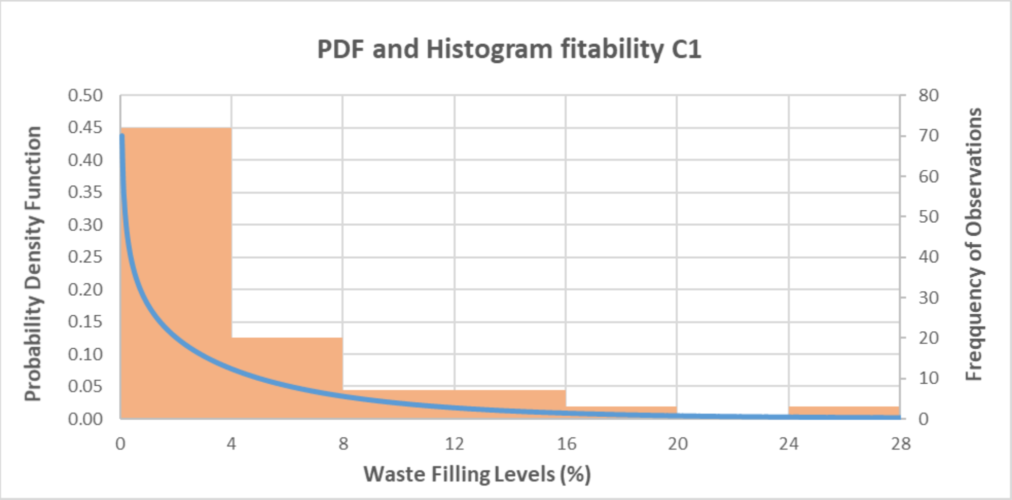


Figure 22: PDF and Histogram (7 intervals with *w* = 4) of container C1.

5.3. Study of the road network

In this part of the work it will be described how the spatial aspects of the network of containers, which are needed for the application of the methods and of the dynamic collection policy described in chapter 4 were handled. One of the main objectives of studying the road network is to capture the topology of the real-case scenario under analysis and, for this, a graph that represents the network of containers was created. As presented before, in subchapter 4.2.2., within the adopted dynamic collection policy, the TSP will be solved many times during the simulation of results, and it will be important to check if there are any containers that are “in the middle” of the encountered routes. This is the reason why this graph is particularly relevant since it constitutes an auxiliary tool that allows to know which of the containers are on the paths that go to other containers, i.e., which of them are intermediate.

The first step was to determine the distances between all containers. For this purpose, the app developed under the WSmart Route research project (WSmart Route, 2019) and the GPS coordinates (latitude and longitude) for each container were used.

Afterwards, the topology of the road network was analysed and the graph that represents it was created in Python. Henceforth, it is explained what was the adopted procedure for creating the graph, with a small example.

Imagine that the network includes only the containers (and a depot at yellow) presented in Figure 23.a. It’s easy to see that in order to go to some of the containers, some others will be inevitably passed by. This graph will be used in order to allow to know which of these containers are in-between.

Then, after the road network is presented, the steps for creating the graph are:

1) Visually analyse the road network;
2) Represent the positions of each location → The coordinates are given for each location taking into account their positions on the map, and with this the nodes are created.
3) Specify which locations are adjacent to each location → A vector is created with the adjacent locations, i.e., the locations with which there is a direct connection (a path that doesn’t pass by other locations). With this, the edges of the graph are obtained.

4) Indicate the distances between the adjacent locations → It is added a vector with the weights of the graph that, in this case, represent the road distance between adjacent containers.

5) Plot the graph.

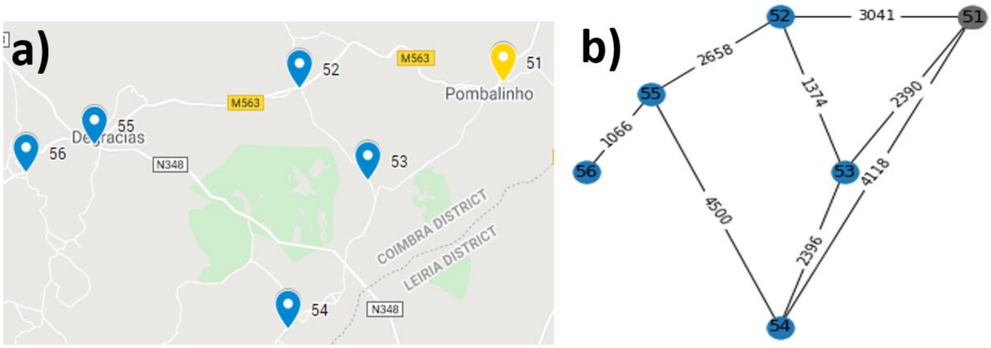


Figure 23: Example of a small road network and its graph.

The graph is saved in a variable and is plotted in Python. In the case of the example, the graph would look like the one represented in Figure 23.b (distances are in meters), and it is possible to see how it represents the topology of this road network example.

In Figure 24 it is shown the graph created for the real network of containers of Soure (see also Figure 18 for comparison). Even if the graph seems somehow disordered, note that what is important is not to visualize the graph but to call it and use it in the Python code. Also, notice that the edges have two weights because it is being considered the real distances and, therefore, the distances between two locations do not have to be symmetric.

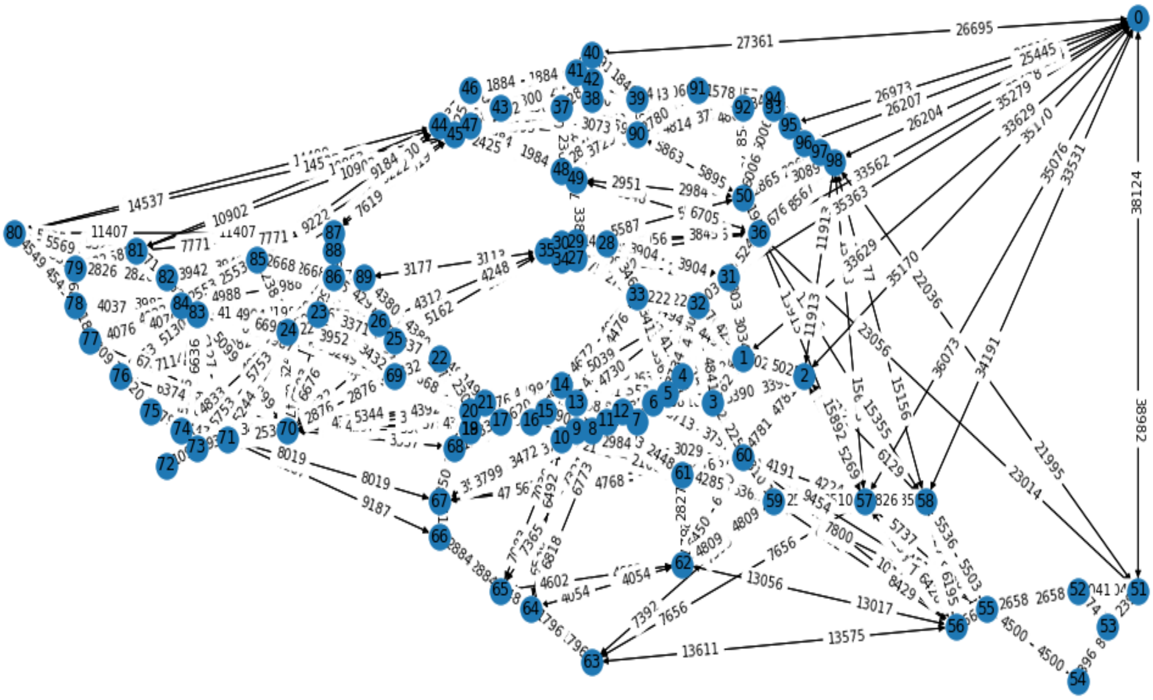


Figure 24: Graph of the Network of 98 Containers of Soure.

6. Results

In this chapter, we present the results of the application of the proposed methods. Firstly, it will be addressed the values and assumptions used for running the simulations (section 6.1). In section 6.2., the results for the baseline scenarios will be described. The partially monitoring scenario is presented in section 6.3, where the results for each method of selection are studied individually by performing a sensitivity analysis on the number of sensors to install. In section 6.4., a financial analysis will be carried out, where it is taken into account the investment costs necessary for adopting the remote sensor technology.

6.1. Values and assumptions used

Before presenting the results, it is important to evidence what were some of the values and assumptions that were adopted in the simulation of the results.

Values used for determining the threshold of collection and the collection interval:

In order to calculate the threshold of collection (ψ_i) and the collection interval (b_i), for each container, as it was depicted in subchapter 4.2.1., it was necessary to assume values for the parameters $\delta_{threshold}$ and $\delta_{interval}$.

It was considered that $\delta_{threshold} = 0.7$. This means that, for each container *i*, it was calculated the $deposition_{limit_i}$ that has a 70% chance of not being surpassed in a day. This value is then deducted to the value of 100% (full container), in order to obtain the threshold level ψ_i. The calculated values can be seen in Appendix D. To note that the lowest value is 83.9% for container #47, and that most of the threshold values are between 90% and 100%.

For the other parameter, it was considered that $\delta_{interval} = 0.8$, i.e., the number of days in which a container *i* requires a visit (b_i) will be determined by calculating the number of days necessary such that the probability of the filling-level of that container being higher than its threshold level (ψ_i) surpasses the limit of 80%. The values for all containers can also be found in Appendix D. The lowest values for each route are 8 days for container #47, in route A, and 9 days for container #65, in route B. However, these values range up to a collection interval of 56 days for container #90, with the most common values varying between 10 and 40 days.

This shows that in the geographical area of Soure, there's a lot of discrepancy amongst the behaviour of the different containers. Taking into account that ERSUC performs each route, in average, within an interval of 7 days, it can be concluded that the collection schedule seems adequate for preventing the most urgent containers from overfilling, while it goes excessively to the containers that do not require to be collected so frequently.

Note that the calculus also depends on the fact that it was chosen a gamma distribution to model the daily waste demands of the containers. So, for calculating the required collection interval (b_i), we have X_i which is a random variable with a gamma distribution (and parameters k_i and θ_i). That being said, considering (9), (14) and (15), we have:

$$X_{i\ in\ any\ given\ day} \sim Gamma(k_i, \theta_i) = Gamma\left(\frac{\mu_i^2}{\sigma_i^2}, \frac{\sigma_i^2}{\mu_i}\right) \tag{17}$$

and so (10) turns to be:

$$X_{i\ after\ b_i\ days} \sim Gamma\left(\frac{(b_i\mu_i)^2}{b_i\sigma_i^2}, \frac{b_i\sigma_i^2}{b_i\mu_i}\right) = Gamma\left(\frac{b_i\mu_i^2}{\sigma_i^2}, \frac{\sigma_i^2}{\mu_i}\right) = Gamma(b_ik_i, \theta_i) \tag{18}$$

Required number of containers to be collected:

Other restriction that was introduced in the dynamic collection policy was that each collection route would cover a required number of containers ($number_{required}$) that would be adapted to the geographical area under study. An important aspect is that the comparison between the real situation and the scenarios where it is simulated that the containers have a sensor must be the most “honest” possible, and thus the constraints that affect the collection in real life must not be forgotten. These constraints are the collection time and the collection vehicle’s capacity. Moreover, it was seen in subchapter 5.1.2. that, in the case of the paper-carboard collection in the municipality of Soure, the main reason why ERSUC needs two routes is due to the fact that, otherwise, the collection time would not be respected.

In order to consider this, and since the two existing routes involve the collection of 50 and of 48 containers, it was defined that in each collection it will be collected 49 containers by using the dynamic collection police. This will serve as a way of indirectly satisfy the collection time constraint.

To note that the author of this work is aware that this solution is not ideal, since forcing the collection vehicle to always collect 49 containers implies two limitations: i) there can be a few performed routes of 49 containers where the collection time would be surpassed; ii) there could be routes where it could make sense to collect less than 49 containers.

However, the main concern with this work was not to develop a very sophisticated dynamic collection policy, but only one that would allow to obtain results that can be considered a relevant approximation of what would happen in real life.

Overflow cost:

In the simulation environment, it was not considered a cost for the overflows. This means that in case there are overflows, the final collection costs will not be affected by the number of overflows. As it will be discussed later, it is very difficult to attribute a value to this parameter, and so different scenarios with different values for the overflow cost will be evaluated posteriori.

Number and period of the simulations:

The results were simulated for a year of collection, i.e., the simulations will run for 365 days.

Finally, it was decided to run/simulate each scenario 3 times, and so the results that will be presented will be the averages of the 3 simulations. This was done in order to reduce the chance of a simulation run being an outlier which could lead to a bias of the results.

Cost per kilometre:

As it was previously mentioned, the cost per kilometre travelled that ERSUC considers in their collection tours is of 1€/km and so this is the value assumed during the simulations. Since the value is 1€/km, from now on, values related to distances will be omitted and only the costs are presented (since they have the same numerical value).

6.2. Study of the baseline scenarios

Here, it will be presented the results of simulating the real case (baseline scenario 1), of simulating two scenarios where no sensors are used but the collection intervals were adapted (baseline scenarios 2 and 3) and, as well, of simulating the case where all containers are monitored (baseline scenario 4). It is of interest to analyse these scenarios, since they will be used as a basis of comparison for the partially monitoring scenario, which is carefully detailed in the section 6.3.

6.2.1. Presentation of each scenario

Baseline scenario 1 – Fixed Collection Routes with 7-Days of Interval (the real case):

As previously mentioned, the results of the current collection policy of ERSUC in Soure will be estimated, i.e., containers will be collected each seven days with the aforementioned fixed collection routes A and B. Even if the costs were already estimated in chapter 5.1.2., with the simulation it will be possible to get more insight into some other factors such as the number of overflows or the frequencies concerning the level at which a container is being collected, and thus perform a more legitimate comparison with the other simulations.

Baseline scenario 2 - Fixed Collection Routes with a Higher Time Interval:

With this scenario, it will be studied the results of maintaining the two fixed collection routes (route A and B) that ERSUC currently performs, but the interval between collections will be slightly increased, in order to match the containers which have to be visited with a higher frequency in each collection route. Following this, route A will be performed with a collection interval of 8 days (due to container #47, as explained in section 6.1), and route B will be performed from 9 to 9 days (due to container #65, as explained in section 6.1).

Baseline scenario 3 – Dynamic Collection Routes with Adapted Collection Intervals:

It will also be simulated the results of the scenario where the routes are adapted to the previously calculated collection intervals, specific of each container. So, for example, if container #48 has to be collected in every 8 days, and container #90 needs to be collected in each 56 days, after these are collected, $days\_To\_Next\_Visit$ will assume the value of 8 and 56, for the two containers, respectively. This scenario can be understood as the limit of trying to apply the dynamic collection policy presented but with no containers being monitored, i.e., the policy described in subchapter 4.2.2 will be implemented but with no real-time information regarding the filling levels. This can be particularly interesting, since the cases of the partially monitored populations can be interpreted as being situated between the case here described and the case where all containers are monitored.

The two different ways of choosing the next most critical container will also be studied here (see step 1.9 in subchapter 4.2.2.) – 1) minimum of $days\_To\_Next\_Visit$ or 2) minimum of $(days\_To\_Next\_Visit\ /\ days\_Collection\_Interval)$.

Baseline scenario 4 - All containers are monitored:

In this case, it is considered that all containers have a sensor embedded. This case is used to study how the waste collection operations would unfold under the dynamic collection policy by having access to the maximum possible information regarding the filling levels.

Also, the results for the two ways of choosing the next most critical container will be here studied.

6.2.2. Analysis of the baseline scenarios

The average results of the 3 simulations are presented in Table 13. It is presented the costs of the collection routes, the number of trips, the total number of days that containers were in overflow, the number of containers collected, and the fill levels at which the containers were collected, including an overfull level. Furthermore, a scatter plot between the costs incurred and the days in overflow, for each scenario, is also showed in Figure 25.

One first thing that is useful to notice is that the "Days in Overflow" column is different from the last column, since in the first we are counting all the days that a container was in overflow, whereas in the last one, it is only counted the encountered overflows at the days of collection. For example, if a container overflows in a certain day, but is only collected 5 days after, it means that this container was 5 days in overflow, but it will be counted only once for being encountered in overflow. If two containers were 5 days in overflow each, it will mean that the total number of days in overflow was 10.

Table 13: Results of simulating the baseline scenarios.

Costs	**Trips**	**Days in Overflow**	**Containers Collected**	**[0;25]**	**[25;50[**	**[50;75[**	**[75;100[**	**[100,+∞]**
Baseline scenario 1								
16293€	104	231 0.65%	5096	1956 38.4%	1866 36.6%	852 16.7%	301 5.9%	122 2.4%
Baseline scenario 2								
13201€	85	466 1.3%	4170	1136 27.2%	1584.3 38%	876.3 21%	364 8.7%	209.3 5%
Baseline scenario 3 Minimum of $days\_To\_Next\_Visit$								
8328€	54	2780.3 7.77%	2646	311.3 11.8%	570.7 21.6%	657.7 24.9%	544 20.6%	562.3 21.3%
Minimum of $(days\_To\_Next\_Visit\ /\ days\_Collection\_Interval)$								
9400€	60	1379.3 3.86%	2940	376.7 12.8%	733.7 25%	812.3 27.6%	584 19.9%	433.7 14.8%
Baseline scenario 4 Minimum of $days\_To\_Next\_Visit$								
14821€	97.3	55 0.15%	4769.3	1619.7 34%	1658.7 34.8%	1020.7 21.4%	415.3 8.7%	55 1.2%
Minimum of $(days\_To\_Next\_Visit\ /\ days\_Collection\_Interval)$								
14829€	94	36.3 0.1%	4606	1260.3 27.4%	1856.7 40.3%	1172.7 25.5%	280 6.1%	36.3 0.8%

By comparing the results of simulating the current waste collection policy (baseline scenario 1), to the results of Table 12, it may be possible to assume that the simulated waste demands slightly overestimate the waste demand relatively to the reality: in the simulations, the percentage of waste collected between 25% and 50% increased, while the percentage collected at the lower level decreased; and there is the occurrence of overflows. This may have to do with the fact of how the data regarding the waste levels were obtained, which certainly didn't allow to fully capture the behaviour of the waste demand. However, even if the overflows happen, they are considerably rare (2.4% of the collections). As for the days in overflow, which are in total 231, is pertinent to see that this value corresponds to 0.65% of the possible days that all containers could be in overflow $\left(\frac{231}{98 \times 365} = 0{,}65\%\right)$. Another way to put the results is by seeing that, in average, each container was 2.36 days in overflow $\left(\frac{231}{98} = 2.36 \text{ days}\right)$. In this way, it is easier to comprehend that the dimension that number represents is not very large. The existence of overflows will then be considered since, once again, it will allow that a fairer comparison can be ensued with the rest of the scenarios, since they will also have their overflows a bit overestimated.

Before comparing all the baseline scenarios, it's important to note that it seems to be a better option to select the next critical container, by considering the quotient between the days to next visit and the required collection interval $(days\_To\_Next\_Visit \ / \ days\_Collection\_Interval)$ than only considering the days to next visit $(days\_To\_Next\_Visit)$. In the scenario with sensors in all containers (baseline scenario 4), the costs among the two different approaches are practically the same, but with the former approach of selection there were less overflows (circa 20 less). In the scenario with the collection intervals adapted to each container (baseline scenario 3), the costs were better in the latter approach of selection (circa 1000€ less), but this meant that overflows were incredible high (and more than double) when in comparison to the former, which seems totally unreasonable. The fact that is better to select the next most critical containers by considering the relativeness to the required collection interval is that it allows containers that are at longer time without being collected of being chosen before containers that have being collected very recently but that have a short collection interval. For the scenario without sensors, the main advantage is that it anticipates collections of those that fill-up slower and are not being collected since long time (avoiding overflows); and for the scenario with sensors, it avoids collecting containers that were very recently collected but have shorter collection intervals, avoiding unnecessary collections, and focusing on emptying more filled up containers. This can be perceived by the fact that the percentage of collections at the lowest waste level decreased. Having said this, from now on, in the rest of the work, for each scenario, it will be only presented the approach that shown here the best results (in the partially monitoring scenario – section 6.3 –, the results were analogues, but this issue will not be addressed again, in order to save space).

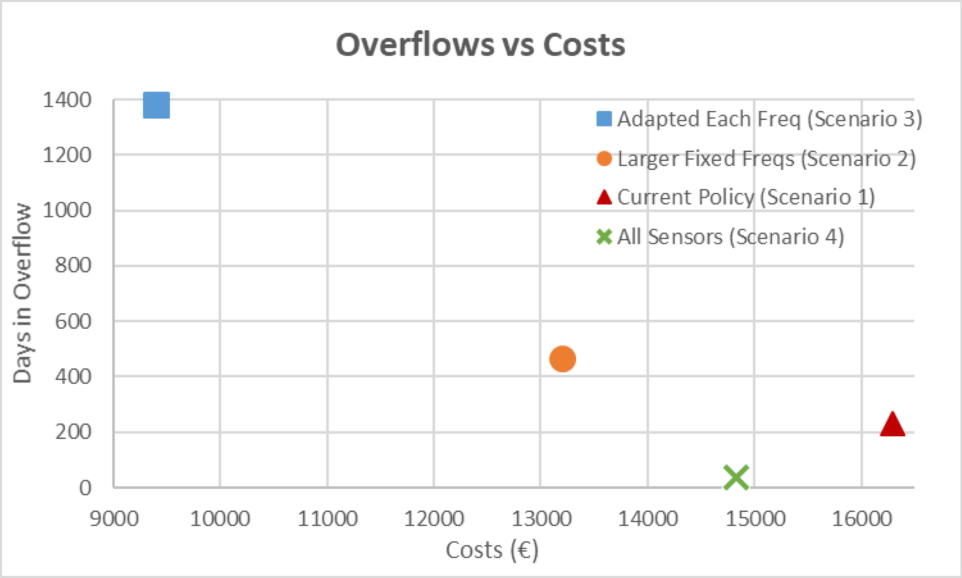


Figure 25: Scatter plot between costs and overflows of the baseline scenarios.

Looking to the different scenarios, one thing is practically inevitable: the higher the collection costs, the lesser the overflows and vice-versa, being the exception the case where all the containers are monitored, which manages to have fewer overflows and costs than the current collection policy. In the three scenarios where the routes do not depend on the measurements about the filling levels (scenarios 1, 2 and 3), it can be observed that by collecting the waste with a smaller average collection interval (number of trips is higher), the waste is collected at a lower level, but overflows are then also avoided. This is what happens with ERSUC: on one hand they are glad that they don’t have overflows, but on the other hand they feel they are visiting too many near-empty containers. However, the two things are interconnected. If by slightly increasing the collection interval of A and route B to 8 and 9 days (scenario 2), respectively, the costs decreased (circa 3000€ less, a reduction of 19%), it is inevitable that the number of days in overflow ends up increasing (it doubles). In scenario 3, where routes were adapted to the collection intervals of all containers, it’s shown a great cost reduction (42%), but due to a great number of overflows, and thus it is certainly a scenario that won’t be very desirable for the waste collection company (even if it would be considered a very small cost for each day in overflow). In this scenario, almost 15% of the performed collection manoeuvres would encounter an overfull container.

As for the scenario where all the containers have a sensor (scenario 4), we can say that the dynamic collection policy presented was successively validated, as it managed to reduce both costs (in 9%) and overflows (in 84%), relatively to the current collection policy. With this dynamic collection, there were overflows in only 0.1% of the possible days and they were collected in the very same day they overfilled (the number of days in overflow is equal to the number of containers encountered overfull). The cost results could eventually even be better, if the TSP was solved in an optimal manner.

6.3. Study of the partially monitoring scenario

In this section, it will be studied the results of dynamically collecting the waste, where only part of the population of containers makes use of the remote sensor technology. In the first 5 sections of this chapter, the results of applying each proposed method are presented. These were done by performing sensitivity analyses regarding the number of sensors to install. In section 6.3.6, the results are compared and discussed, and some overall conclusions and tendencies are outlined.

In the presented figures, it is shown again the results of the current policy (baseline scenario 1), of the scenario where all containers are monitored (baseline scenario 4) and of the scenario where the routes were adapted to the collection intervals of each container (baseline scenario 3), since the results will be in between these last two.

The ordered lists of all the sensor placement methods are presented in Appendix E. The methods where the map of the 14 first locations that were selected is not accompanying the text, they can be found also in the same appendix.

6.3.1. Method 1 – Waste demand rate of a container

The first 14 container locations that were selected by considering the waste demand rate of the containers, i.e., from the ones with lowest demand to the ones with highest demand, as described in section 4.1.1, are presented in Figure 26, with a red colour.

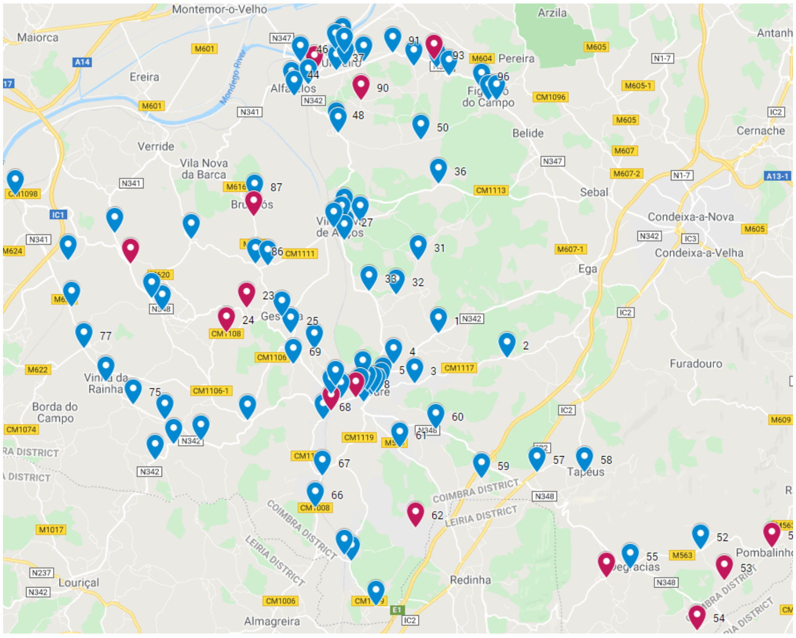


Figure 26: First 14 locations selected with method 1.

It was carried a sensitivity analysis on the number of sensors to install. It was chosen to study cases with a difference of 14 sensors each, by the mere reason that it is a divisor of 98 (the total number of containers). This resulted in 6 cases studied between 0 and 98 sensors (see Figure 27). The numbers at each bullet point correspond to the number of sensors to install (due to aesthetic reasons, for the values of “14”, “42” and “70” sensors it’s only represented the bullet point).

Furthermore, the results presented for all methods will be focused on comparing the operational/collection costs and number of days in overflow. In order to understand if the methods of selection have any meaningful impact, for all of them, the order of selection described in chapter 4 (in the charts: “selected order” at blue) will be compared to the results of the reverse order (in the charts: “inverted order” at orange).

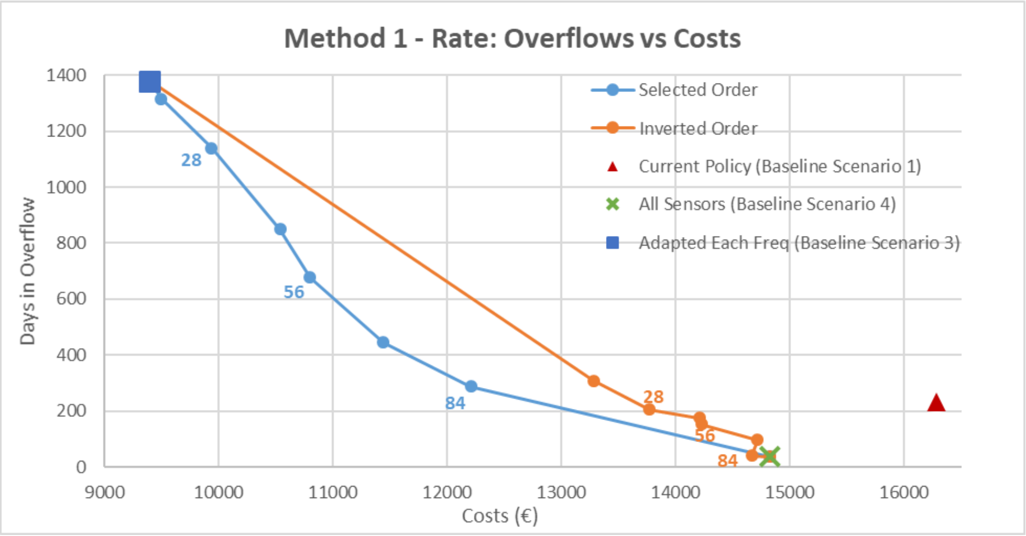


Figure 27: Overflows vs costs with a selection based on the waste demand average rates (method 1).

From Figure 27, it is possible to note that by starting to select the containers that are the slowest to fill-up (blue line), with a low number of sensors, there is almost no difference to the scenario where there are no sensors but the collection intervals were adapted (baseline scenario 3): costs are maintained at a very low value, but a very higher number of overflows exists. Whereas, when sensors are embedded first onto the containers that have the highest demand (orange line), there's a great reduction on the number of overflows but there's also a great increase on the collection costs. However, it can be seen that with only 28 sensors, both the costs and overflows presented are less than the costs and overflows of the real case (cost decreases around 16% for the same number of overflows). That situation never happens with the selected order – even with 84 sensors, the costs are very small, but the number of overflows is still higher than in the real case. Note also that the biggest differences in costs happened when the containers with the highest demand were added (last interval of the selected order and first interval of reverse order).

These results show that, in fact, with the dynamic collection policy adopted, this is a criterion that presents visible influences on the collection costs and number of overflows. However, contrary to what was before being considered, depending on the overflow cost, it may also make sense to start embedding the sensors on the containers with the highest demand.

This can be explained by the way the dynamic collection policy works. If we start placing sensors on the containers that fill-up very slowly, the routes will still be mainly imposed by the containers that fill-up quickly, since their required collection interval is shorter and will trigger the collection travels most of the times. Overflows are avoided in the containers that fill-up slowly, but when the containers that fill-up quickly occasionally experience even higher demands, the collection operations will not be able to respond, and overflows happen on those containers. However, if sensors are placed on the containers that fill-up quickly first, collection travels will be more tailored to the real necessities of the demand of these containers and will impede them from overflow. Costs will increase due to more trips being undertaken, but there's also the benefit of anticipating more collections of the non-monitored containers – that are added to fulfil the required number of containers that have to be collected – and so, the probability that these may overflow is also reduced.

6.3.2. Method 2 – Waste demand variability of a container

Figure 28 shows that by allocating the sensors based on the coefficient of variation's values of each container, the pattern across the several levels of monitorization seems to be very similar to what was obtained by considering the average rates (method 1).

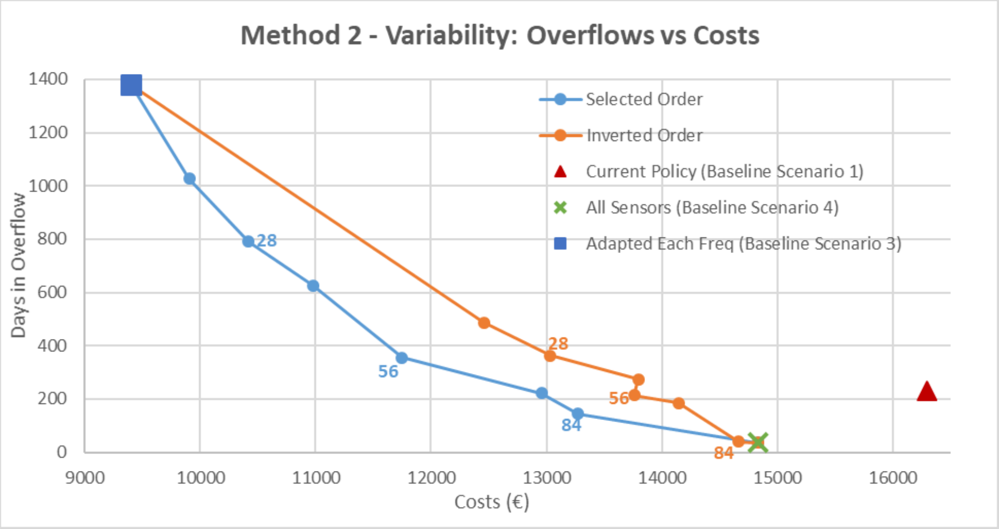


Figure 28: Overflows vs costs with a selection based on the waste demand variability (method 2).

At first sight, one could prematurely conclude that this is proof that the use of this criterion to select the containers has a meaningful impact on the costs and overflows. However, the results don't make sense and further analysis is required: it doesn't make sense that selecting first the containers with the lower coefficient of variation (inverted order) would provide a greater reduction on the overflows than selecting first the ones with a higher coefficient of variation (determined order).

It's important to note that it's being analysed a real case with real data, and thus when studying the effects of one variable, the other variables may be influencing the results of the simulations. In other words, the results may have been similar to the ones obtained by the average rate criterion, but due to the fact that the average rates may be influencing here the results. In order to better understand if this is the case, the coefficient of variation and the average rate value of all containers are presented in Figure 29.

By analysing the dispersion of the values of the graph, it can be concluded that, in fact, the results were mainly influenced by the average rates: by selecting first the ones with the higher coefficient of variation we are also selecting containers with low average rates; and by doing the opposite, i.e., by selecting first with the reverse order of the coefficient of variation, we end up rapidly selecting the ones with high average rates.

This is further supported by the fact that in the chart of Figure 28, relatively to the one of the average rates, the results of the selected order and of the inverted order seem closer to each other and are less of extremes. See, for example:

- The first interval (0 to 14), for the selected order, was very small in the first graph, but now is wider;
- The last interval (84 to 98), for the selected order, was very wide, but here is smaller;

- The first interval, for the reverse order, was very wide in the first graph, but now is smaller;
- The last interval (84 to 98), for the reverse order, was very small, but here is a bit larger.

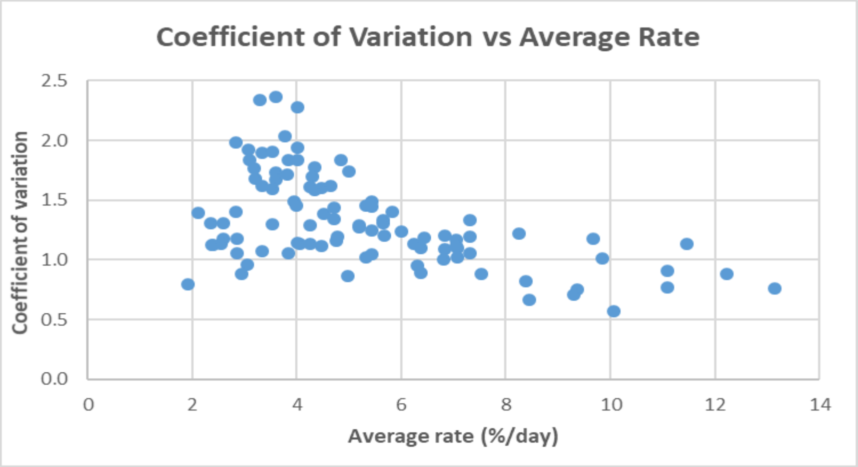


Figure 29: Scatter plot between the coefficient of variation and the average rates of the containers.

This is due to the fact that here it was first selected containers with low average rates – but not the lowest possible ones – with the “correct” order, and it was first selected containers with high average rates – but not the highest possible ones – with the reverse order.

No evidence is then present that, with the dynamic collection policy adopted, this criterion influences the collection costs and the number of overflows.

6.3.3. Method 3 – Remoteness of a container

Regarding the maps, that can be found in Appendix E, of the two options for selecting the containers based on the criterion remoteness, it is possible to see how they differ: the first option, since it is only concerned with the distance to the closest container, allows the selection of containers from more central positions (even it is a container along a road that is inevitably passed by), and the second option only allows the selection of containers that are more at the edges.

In Figure 30, it is displayed the simulation results of having chosen the containers by the two options used to evaluate remoteness. The curves of the two options used to evaluate remoteness are roughly similar. It’s possible to see that in the two cases, there is not much difference between the two curves (selected and inverted orders), and so it’s very difficult to discern any effect. In some cases, with a lower number of sensors, the selected order shows less costs but higher number of overflows, but then the opposite happens with a higher number of sensors, with the inverted order presenting less costs but higher number of overflows – see, for example, in the second graph how that happens with 14 and 28 sensors.

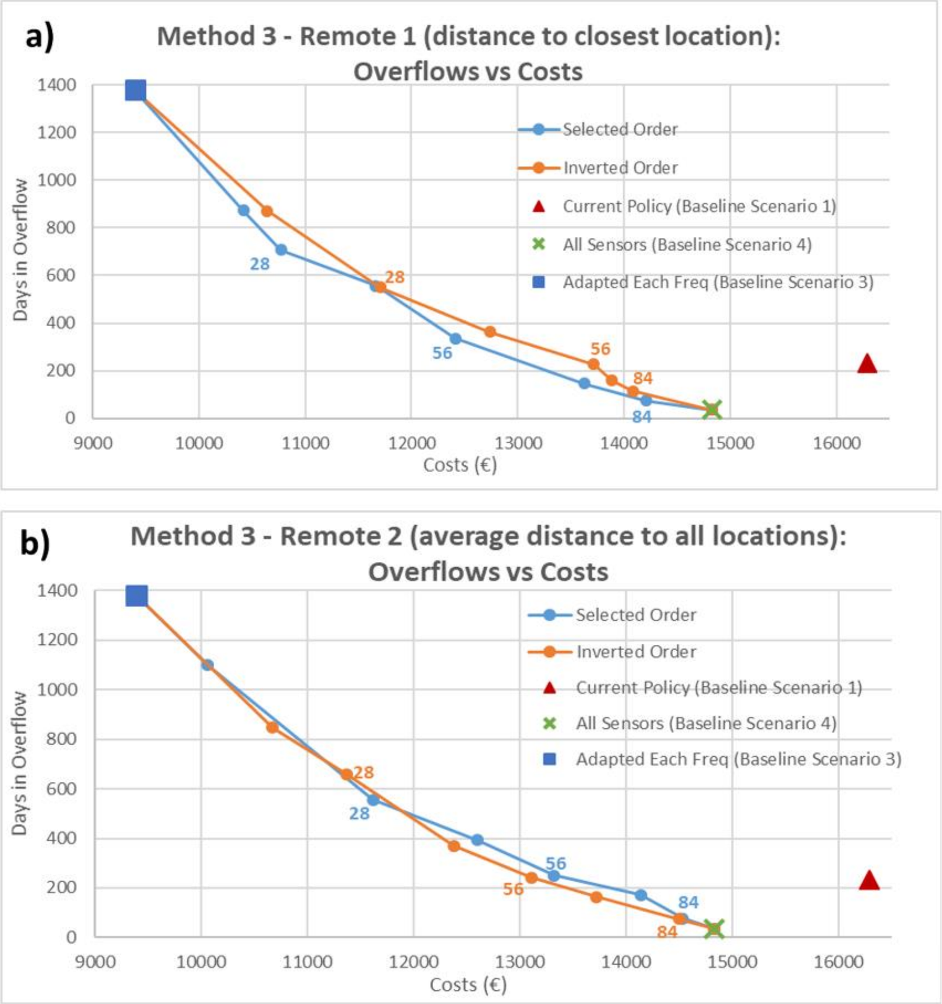


Figure 30: Overflows vs costs with the two options based on remoteness method (method 3).

Following this, it was decided to analyse a case where the sensors would be selected in a random fashion. The results are presented in Figure 31.

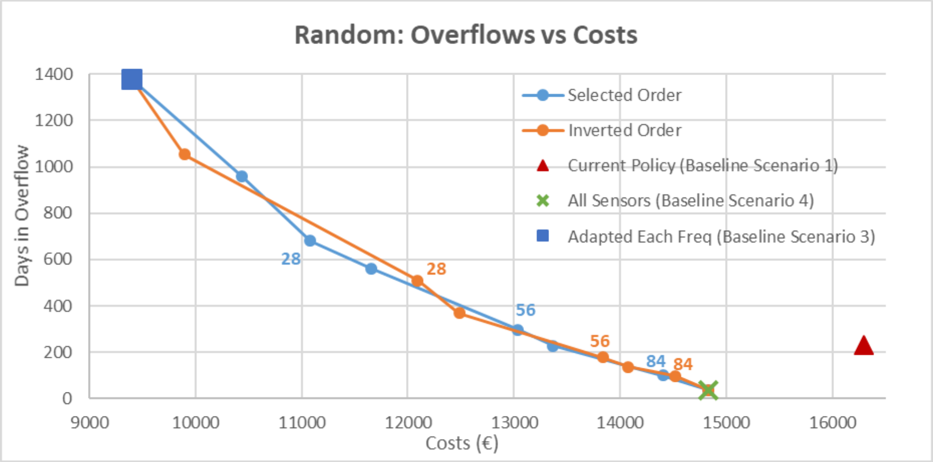


Figure 31: Overflows vs costs of a random selection.

As can be seen, with a random selection, there's also not much difference between the curves, and it is obtained a very similar pattern to the patterns obtained with the two options for evaluating the criterion remoteness. This seems to suggest that there is not enough evidence supporting that the results, by using the described dynamic policy, had been affected by having chosen the containers considering their remoteness, by the two presented options.

6.3.4. Method 4 – Cost Reduction Algorithm

Two versions of the Cost Reduction Algorithm had been proposed: with and without considering frequencies (section 4.1.4). Both versions were applied and the results were somehow similar, i.e., in each set of the 14, 28, 42, etc. first containers, maybe the order changed within the set, but the containers that would end up being selected were practically the same. For this reason, the results for the version that considers the frequencies will be skipped.

Therefore, Figure 32 presents the map for the 14 first chosen locations, considering the cost reduction algorithm (without considering frequencies). It can be seen that with this algorithm it was possible to successfully select locations that seem to allow a good reduction of the path costs – it selects isolated locations independently if they are on the boarder or in more central positions.

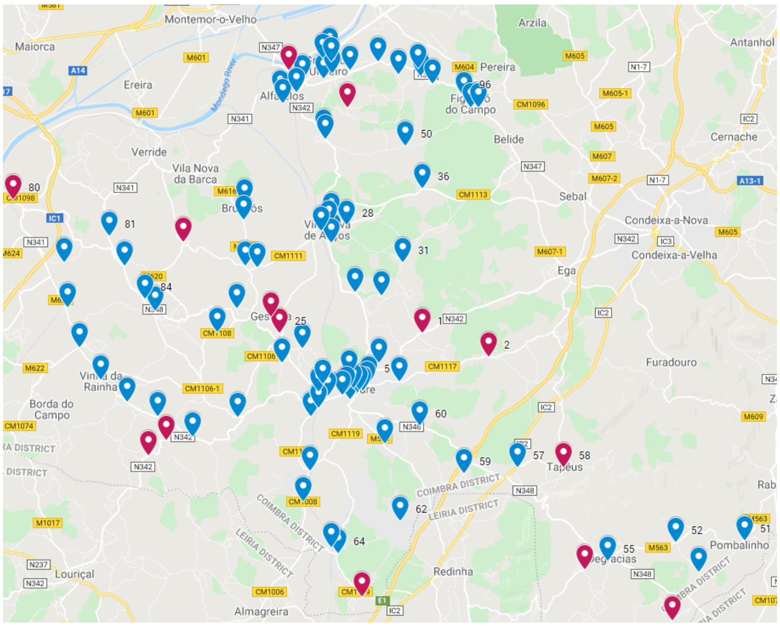


Figure 32: First 14 locations selected with the Cost Reduction Algorithm (without considering frequencies).

The results of simulating the daily waste collections are presented in Figure 33. Even if the evidence is not very strong, it seems that selecting containers by the order that the cost reduction algorithm provides is better than selecting containers with the reverse order. Note that with 28, 56, 70 and 84 sensors, by selecting with the “correct” order both costs and overflows are lower when in comparison to the inverted order, and with 14 and 42 sensors, the results are relatively close.

It’s still difficult to discern if these results were all due to the selection method, or if were also heavily influenced by the waste demand rates of the containers. However, it is considered that this method of selection shown a certain degree of impact.

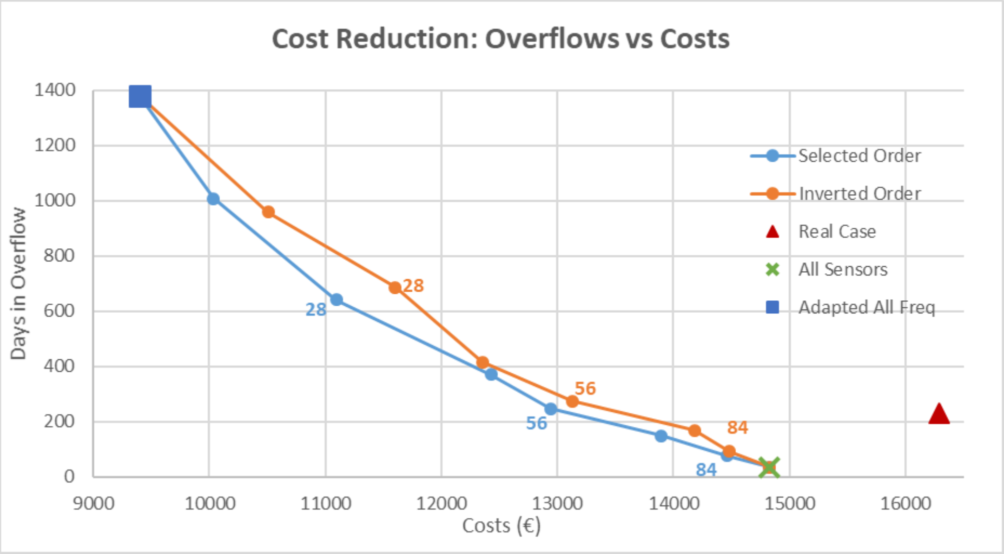


Figure 33: Overflows vs costs with a selection based on the cost reduction algorithm.

6.3.5. Method 5 – Considering rate, variability and remoteness

During the development of this work, it was concluded that it could make sense to evaluate the results of allocating the sensors in containers that were representative of three of the chosen criteria – waste demand rate, waste demand variability and remoteness. It is considered that this method may still make sense, however, it's important to have full disclosure with the fact that:

1) Only the criterion waste demand rate presented relevant results, and that went against the initial belief (it seems that starting by embedding sensors in containers with higher demand rates is also reasonable);
2) Since this is a real case, it's not guaranteed that it's possible to find elements with significative values across the different criteria. The ideal situation for performing this analysis would be, for example, by manipulating the several criteria, on the containers, in order to study what we sought to.

Therefore, as it was seen, there was a convergence between the low waste demand rates and high levels of variability, which influenced the results of when selecting by the variability criterion. So, the focus of this analysis will be more in studying the effects of intersecting containers with high variability with containers with a relatively high demand. Since the criterion of remoteness didn't present any relevance, it won't be considered in this analysis. So, the case that will be here studied is the one that corresponds to – as can be seen in Figure 15, in subchapter 4.1.5 – selecting containers from the octants 3 and 4 (simultaneously).

The group of containers that is simultaneously in the top 40% of containers with higher variability and of 40% containers with higher average rate, constitutes only a set of 4 containers: 49, 91, 68 and 29, which again shows how problematic it is to do this analysis with this real case. The results of simulating the collection routes are presented in Figure 34. It's also displayed the results of the top 4 containers with higher variability, of the top 4 containers with higher waste demand rate and a random selection of 4 containers. It's possible to see that the 4 containers selected shown better results than the random selection. However, this may be due to chance and may not happen when in comparison to another random group. It’s possible to see that the same pattern is repeated, and no further conclusion can be made regarding the use of the criterion variability: less overflows, but more costs relatively to the top 4 of variability, due to higher demand rates being present in this new selection; more overflows, but less costs relatively to the top 4 with the higher average rates, due to lower demand rates being presented in this new selection. This analysis should be done with a control group.

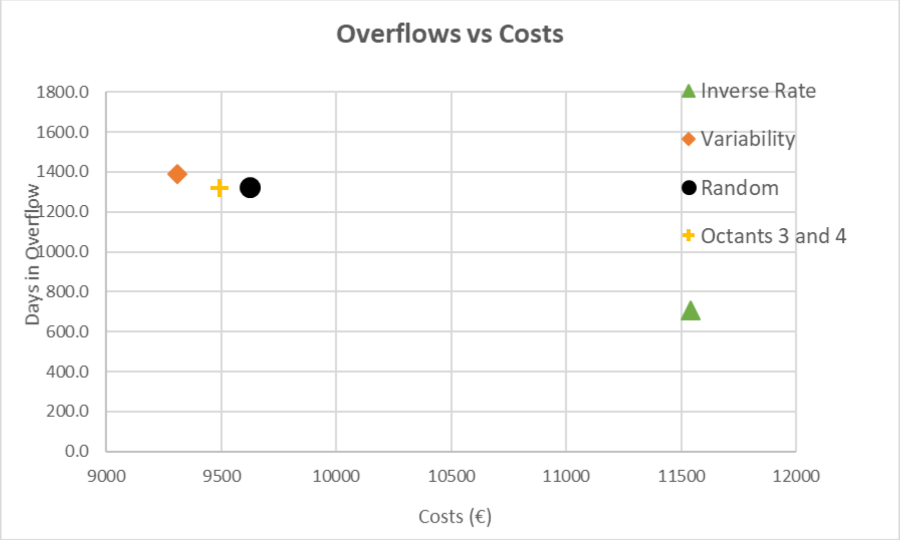


Figure 34: Results for the containers with high variability and high average rates.

6.3.6. Overall conclusions and tendencies

After analysing the several cases, it was possible to see that by adding sensors, more information about the filling levels is obtained, and so it’s possible to progressively define more efficient collection routes, and more capable of avoiding overflows. However, in order to decrease the number of overflows, the costs of collection end up increasing since more routes are triggered to avoid overfull containers.

It’s also possible to see that for a high number of sensors, the results start to become more uniform, and most of the methods present similar trends: first sensors added are more impactful (reduce more overflows but also increase more the costs), and the last sensors to be added seem to provide relatively less benefits (reduce less overflows but costs are also less increased).

The main exception to this, seems to be the case where sensors were added by the order of the lower waste demand rates, where the first sensors don’t present much impact but the last ones to be added do. And, correspondingly, the case where they were added in the opposite order (by the higher waste demand rates) was where this trend was most clear. This method of selection was undeniably the one that, within the dynamic policy presented, was more capable of modifying the simulation results.

Finally, it’s important to note that for most cases, having selected 70 or 84 sensors was enough to present less overflows and less costs simultaneously when in comparison to the current waste collection policy. With 56 sensors that was also possible some of the times, or the results were very close. When selecting by the lowest demand rates, obtaining both less overflows and less costs than the current waste collection policy was not feasible, whereas, by selecting by the highest waste demand first, with only 28 sensors that feat was made possible.

6.4. Financial Analysis

In order to study the financial feasibility of implementing sensors into a reduced population of containers, it's not enough to consider only the collection costs and the number of overflows. It's necessary to understand what are the investment costs of such technology, depending on the number of sensors and also what are the impacts of the overflows, meaning that is necessary to attribute a cost to each overflow.

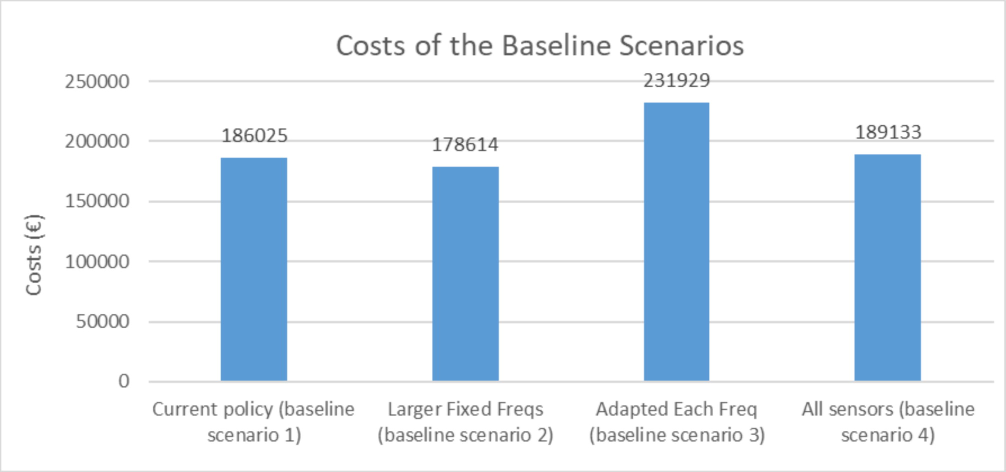
It must be pointed out that it is not easy to estimate a cost to the overflows. First, they can mostly be considered an intangible cost, i.e., a cost that is not quantifiable but that can impact the overall performance of a company. Waste collection companies are not comfortable in attributing a cost to each overflow, since an overflow does not have an immediate and direct economic impact. However, they often want to avoid overflows at all events, in order to ensure a good quality of service to the citizens and avoid complains. Therefore, in this financial analysis, it will be studied 2 different overflow cost schemes:

- A low overflow cost, where each day in overflow (per container) costs to the company 10€;
- A medium overflow cost, where the overflow cost is 35€ per day (per container);

Regarding the investment costs, it will be considered the use of sensors provided by the Evox, partner of this work, and it will be considered that the price is 250€ for each sensor acquired plus 10€ for each installation on a container (average of the range of prices given). It will also be paid a fee of 1€ per sensor, per month. It won't be considered any costs with software and the analysis will be performed for a time period with a length of 10 years, which corresponds to the battery life of the sensors.

Furthermore, this analysis will benchmark all scenarios (the four baseline scenarios and the partially monitoring scenario). Within the partially monitoring scenario, only the methods with the more relevant results will be analysed here - the selected and reverse order for the waste demand (method 1 low and high, respectively), and the selected order for the cost reduction algorithm (method 4).

The results of the costs during a period of 10 years, for the low overflow cost and for the medium overflow cost are presented in Figure 35 and Figure 36, respectively.



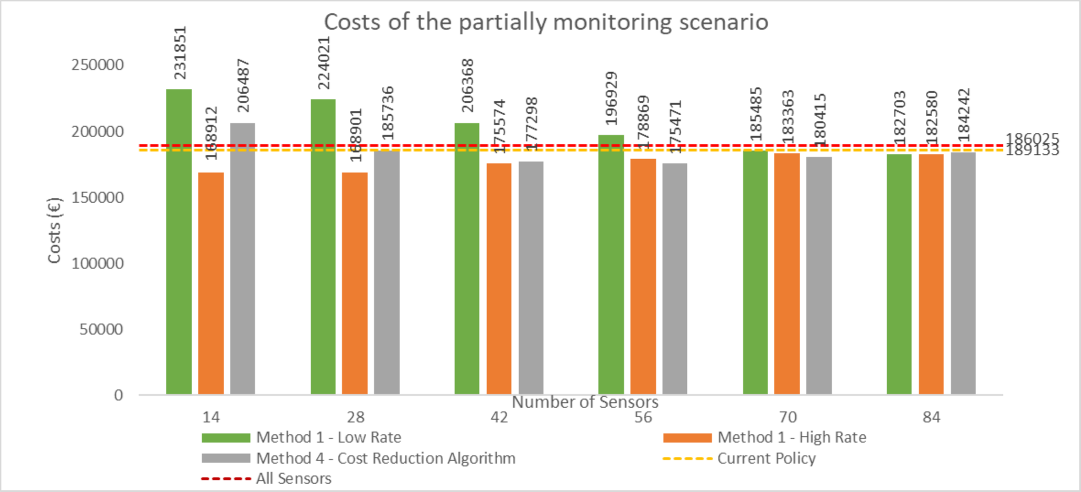


Figure 35: Total costs, in 10 years, for a low overflow cost.

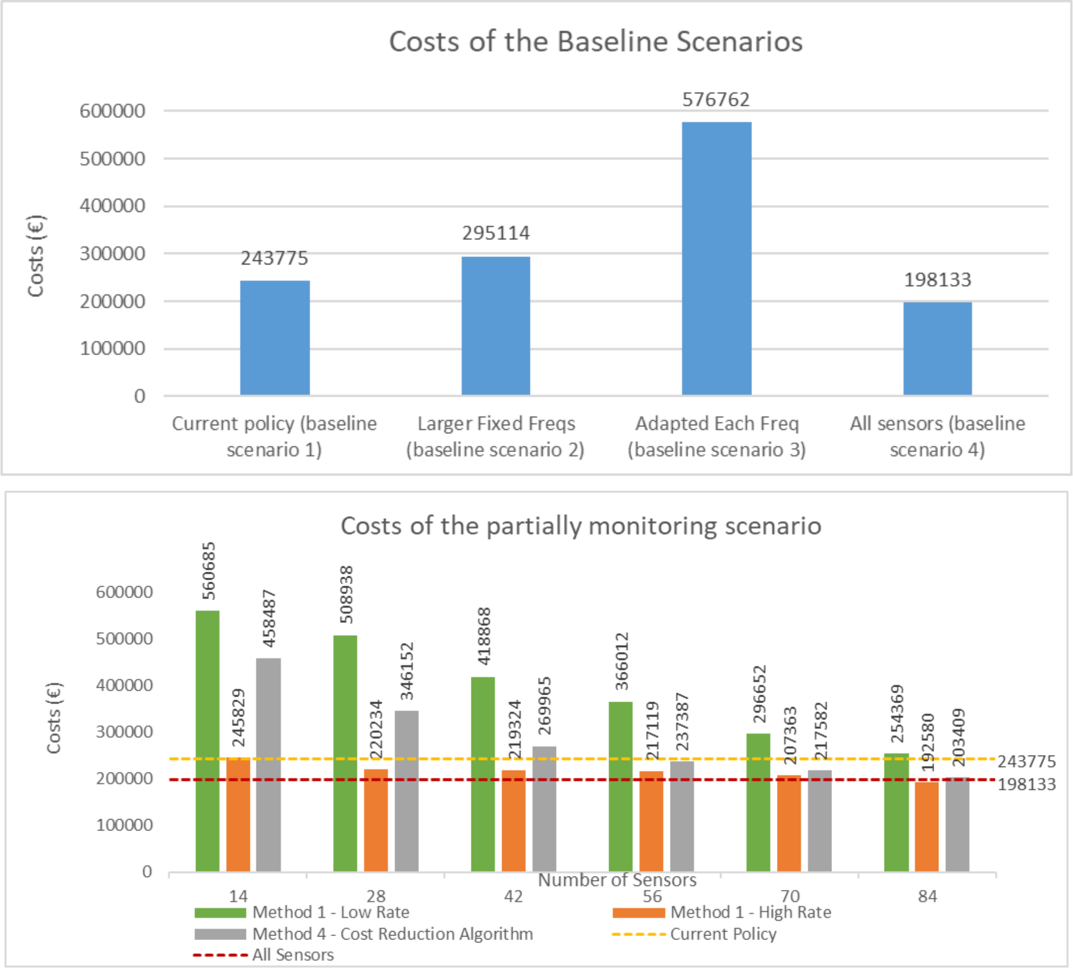


Figure 36: Total costs, in 10 years, for a medium overflow cost.

By analysing the two cost schemes for the overflow, it's possible to tell that the scenario where collections are adapted to the collection intervals of all containers (baseline scenario 3) represents the worst collection policy in both cost schemes, due to an extremely large number of overflows.

Regarding the other baseline scenarios, it can be noted that monitoring all containers (baseline scenario 4) is worse than keeping the current waste collection policy (baseline scenario 1) for a low overflow cost (2% more costs), but is the best baseline scenario for a medium cost (19% less costs). Augmenting the collection intervals by one day in route A, and by 2 days in route B (baseline scenario 2), is the best baseline scenario for a low overflow cost (less 4% costs comparing to the current situation), but shows worse results than the current policy for a medium cost (21% more costs) – is a scenario very impacted by the overflows.

Regarding the methods of the partially monitoring scenario, it's possible to see that there's a decrease in the costs where the sensors are added by the lower average rates (method 1-low) and by the path reduction algorithm (method 4), in both cost schemes, meaning that adding sensors brings more benefits than costs. With the second cost scheme this decrease is relatively higher, due to each overflow reduction being more impactful. To note that the cost reduction algorithm allowed, with 56 sensors and for a low overflow cost, a cost reduction of 7% when compared to case will full monitoring, and of 6% when compared to the current policy. For a medium overflow cost, with 84 sensors, it shown worse results than the full monitoring (costs increased 3%) but still better results than the current policy (costs decreased 17%).

When the sensors are added by the higher demand rates (method 1-high), there's no proportional reduction with the number of sensors in the first cost scheme, and there's a soft reduction in the second cost scheme. While with the other sensor selection methods, the benefits seem to grow proportionally with the number of used sensors, with this method, most of the benefits are obtained with the first sensors added, and then are practically held constant with the addition of new sensors, meaning that the incremental benefit diminishes. Considering the dynamic collection policy adopted, allocating sensors by this method, i.e., by partially monitoring the population of containers with method 1 – high rates, seems to be the wisest choice.

Depending on the available budget and on the overflow cost scheme considered, different allocations can be recommended to ERSUC. If the overflow cost is considered to be not very high (low cost), allocating a small number of sensors by this method can be the best solution for minimizing the costs (in the studied cases, 14 or 28 sensors present essentially the same results) – the total cost is reduced in 11% when comparing to installing 98 sensors, and is reduced in 9% when comparing to the current situation (it is recommended to install 14, since it requires a lower investment). Considering a medium overflow cost, with a low budget, sensors can be placed on 28 locations and most benefits are still attainable (it reduces 10% the costs relatively to the current situation but is 11% more costly than monitoring all containers), but with a sufficient budget, the best option is to allocate sensors in almost all containers by this method (in the studied cases, 84 sensors) – the reduction is of 3% when comparing to installing 98 sensors and of 21% when in comparison to the current policy.

There are also some other benefits that are not being taken into account. When a container is monitored, there's the benefit of having more accurate information that can be used to better model the historical data regarding that container, and thus it's possible to constantly adapt the dynamic collection policy to the encountered trends or changes in the overall demand.

7. Conclusions and Further Work

The objective behind this dissertation was to investigate the possibility of integrating ultrasonic sensors on the waste collection operations of solid waste in an economically more sustainable way by trying to reduce their investment costs while maintaining the benefits of monitorization. By analysing the existing literature, it was identified that four aspects could make a container more attractive to be monitored: a low waste demand rate, high waste demand variability, high remoteness, and not being an intermediate container. Several selection methods, that considered these criteria were then proposed, and a dynamic collection policy was also developed in order to simulate the results under the partially monitoring scenario of each method. This was an analysis that intended to fill a gap in the existing literature, since this had not been done before. It was analysed the real-case scenario of the paper/carboard collection of Soure, carried out by ERSUC, and used data regarding each individual container (a total of 98 containers), in order to simulate the results. Sensitivity analyses, regarding the number of sensors added were performed to comprehend the impacts on the results, of each method. From the analysis of this work, the following conclusions are outlined:

- There's always a trade-off between operational costs and overflows. For the scenarios that made use of the dynamic collection policy it was seen that with more sensors there's always a decrease on the number of overflows, but there's also an increase on the collection costs, since more trips are done to avoid overflows. For the scenarios with fixed collection routes, more frequent collections reduce the overflows, but increase the collection costs – is what happens in the current collection policy –, and with less frequent collections happens the opposite.
- The dynamic collection policy presented was validated, since by monitoring all the containers is possible to have both less operational costs (decrease of 9%) and less overflows (decrease of 84%) than with the current collection policy. The financial analysis shown that is better to adopt sensors in all containers than maintaining the current collection policy for a medium overflow cost (cost reduction of 19%). However, for a low cost this policy shows worse results (increase of 2% on the costs).
- If ERSUC does not consider each overflow to have a great impact (low overflow cost), the company can still have benefits (of less 4% costs) without using the sensor technology – by adopting the baseline scenario 2, i.e., augmenting routes A and B to a collection interval of 8 and 9 days, respectively.
- For the criteria remoteness (method 3) and variability (method 2), it was not perceptible any impact on the results. With the cost reduction algorithm (method 4) was detected a certain degree of impact since it was shown that using the order specified by the method would be better than using the reverse order. With 56 sensors, this method shown better results than the full monitoring scenario for a low overflow cost (cost reduction of 7%) but worse ones for a medium overflow cost (cost increase of 3%). When in comparison to the current situation, cost reductions of 6% and of 17% were obtained for a low and medium overflow cost, respectively.
- However, the method that considered the waste demand rate (method 1) was the one that considerably impacted the results. Contrary to the initial assumption, it is best, under the

dynamic collection policy adopted, to start by monitoring the containers that experience the highest demand rates, instead of the ones the experience the lowest ones. With this option, most of the benefits of the monitorization are obtained with a low number of sensors and, in the financial analysis, this way of selection presented the best results, for both cost schemes considered – considering a low overflow cost, the recommended number of sensors was 14 (compared to the scenario with sensors in all containers and to the current situation costs were reduced 11% and of 9%, respectively), and with a medium overflow cost, the recommended number of sensors was 84 (compared to the scenario with sensors in all containers and to the current situation costs were reduced 3% and 21%, respectively). This means that, in fact, it's possible to maintain the sensors' benefits by partially monitoring a collection area.

- Independently of the method, with 70 – 84 (sometimes 56) sensors it's possible to have better results than with the current waste collection policy (for example, by installing 70 sensors it's possible to reduce in 14-21% the operational costs while reducing overflows between 4-37%), and the first sensors seem to be the most impactful. The exception is when it's decided to start by the containers with the lowest demands.

It is also important to be aware of some difficulties and limitations that were encountered during the development of this work, and of some recommendations that can be useful for academics or waste collection companies interested in further exploring this topic:

- It's necessary to understand that the conclusions here drawn cannot be disassociated from the dynamic collection policy adopted. The author still believes that it makes sense to further study all the criteria identified in the literature review, but by using a different methodology, where it would be easier to detect the impact of each sensor embedded. It's probable that these criteria would work better in a scenario where routes are kept fixed to the containers that are not being monitored, and thus the sensor would only be used to decide if at the day of collection a container needs or not to be added to the fixed route. For example, note how under those circumstances, the cost reduction algorithm would be very reasonable to be considered – here, the effects of using that selection method are somehow "lost", due to future routes being dynamic and not the final fixed route that could be given by the algorithm. This proposal could also signify a smoother transition for the waste collection companies in adopting sensors, since they could see immediate results with little effort.
- Conversely, it's also reasonable to say that different dynamic collection policies that are more well-thought and better tailored to a partially monitoring scenario, could be considered in the future. Remember that, in the dynamic policy adopted, it was considered that a certain number of containers had to be collected in each route, in order to reflect collection time restrictions, and there were also no restrictions about the vehicles' capacity. In the future, a better option would be to consider a VRP with capacity and time restrictions.
- Other difficulty was when combining the different criteria (method 5). That part of the analysis wasn't very interesting, since it was being considered a real case and the presence (and intensity) of each criterion couldn't be modified. A more theoretical study regarding the criteria could benefit from manipulating the data and analyse the impacts on the results of each criterion.

- Other limitation was the fact that the number of overflows was a bit overestimated. Maybe this could be attenuated with the data being collected over a longer period of observations.
- A factor that has to also be taken into account is that the results can be greatly different if it is considered another price model for the sensors. Here, it was considered the average of the prices given by Evox for their sensors, but if it had been used the minimum (or the maximum) price given, the results would be different. Recall that, in the beginning of this thesis, it was presented the results of Gonçalves (2014), from where it was seen that, in terms of costs, using all sensors was way worse than maintain the current situation of the studied area. However, that was very impacted because it was considered a very high price for the sensors relatively to what was considered here, in this work – with a more reasonable price, monitoring all containers is not such an unreasonable option.
- One option that was not here explored (only mentioned) was the fact that containers could be grouped in spatial clusters, i.e., in a group of nearby containers, it could make sense to monitor only one (or a few) that are representative of the group (Marchiori, 2018).

Bibliography

Anagnostopoulos, T., Kolomvatsos, K., Anagnostopoulos, C., Zaslavsky, A., & Hadjiefthymiades, S. (2015). Assessing dynamic models for high priority waste collection in smart cities. *Journal of Systems and Software*, *110*, 178–192. https://doi.org/10.1016/j.jss.2015.08.049

APA. (2019a). Agência Portuguesa do Ambiente - Políticas > Resíduos > Planeamento em Resíduos > Plano Estratégico para os Resíduos Urbanos (PERSU). Retrieved May 2, 2019, from https://www.apambiente.pt/index.php?ref=16&subref=84&sub2ref=108&sub3ref=209

APA. (2019b). Agência Portuguesa do Ambiente > Políticas > Resíduos > Gestão de Resíduos Urbanos. Retrieved April 25, 2019, from https://www.apambiente.pt/index.php?ref=16&subref=84&sub2ref=933

APA. (2019c). Agência Portuguesa do Ambiente > Políticas > Resíduos > Gestão de Resíduos Urbanos > Sistemas de Gestão e Infraestruturas. Retrieved May 1, 2019, from https://www.apambiente.pt/index.php?ref=16&subref=84&sub2ref=933&sub3ref=934

APA. (2019d). Agência Portuguesa do Ambiente | Municipal waste production and management | Relatório do Estado do Ambiente. Retrieved April 25, 2019, from https://rea.apambiente.pt/content/municipal-waste-production-and-management?language=en

Barles, S. (2014). History of Waste Management and the Social and Cultural Representations of Waste. In M. Agnoletti & S. Neri Serneri (Eds.), *The Basic Environmental History* (pp. 199–226). Springer. https://doi.org/10.1007/978-3-319-09180-8_7

Bilitewski, B., Wagner, J., & Reichenbach, J. (2018). *Best Practice Municipal Waste Management*. Umweltbundesamt. Retrieved from https://www.umweltbundesamt.de/en/publikationen/best-practice-municipal-waste-management

Bopardikar, S. D., Smith, S. L., Bullo, F., & Hespanha, J. P. (2010). Dynamic vehicle routing for translating demands: Stability analysis and receding-horizon policies. *IEEE Transactions on Automatic Control*, *55*(11), 2554–2569. https://doi.org/10.1109/TAC.2010.2049278

Chang, N.-B., & Pires, A. (2015). *Sustainable Solid Waste Management*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119035848

Christensen, T. H. (2011). *Solid Waste Technology and Management, 2 Volume Set*.

Christofides, N. (1976). Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem. *Management Sciences Reseach Report*, *388*(February).

Clark, G. (2014). The Industrial Revolution. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 20, pp. 217–262). Elsevier B.V. https://doi.org/10.1016/B978-0-444-53538-2.00005-8

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms* (3rd ed.). Massachusetts Institute of Technology.

Correia, M. M. (2016). *Modelo de apoio à decisão para a utilização de TIC na otimização da recolha de resíduos recicláveis*. Retrieved from https://repositorio.iscte-iul.pt/handle/10071/14003

Dantzig, G. B., & Ramser, J. H. (1959). The Truck Dispatching Problem. *Management Science*, *6*(1), 80–91. https://doi.org/10.1287/mnsc.6.1.80

DEFRA. (2011). *Guidance on applying the Waste Hierarchy*. *Department for Environment, Food & Rural Affairs*.

Díaz, B. D. (2006). What is VRP? Retrieved from http://www.bernabe.dorronsoro.es/vrp/index.html?VRP-Intro.html

Elia, V., Gnoni, M. G., & Tornese, F. (2018). Improving logistic efficiency of WEEE collection through dynamic scheduling using simulation modeling. *Waste Management*, *72*, 78–86. https://doi.org/10.1016/j.wasman.2017.11.016

Elsersy, M., Elfouly, T. M., & Ahmed, M. H. (2016). Joint Optimal Placement, Routing, and Flow Assignment in Wireless Sensor Networks for Structural Health Monitoring. *IEEE Sensors Journal*, *16*(12), 5095–5106. https://doi.org/10.1109/JSEN.2016.2554462

ERSUC. (2018a). *ERSUC - Relatório de contas (2018).pdf*. Retrieved from http://ersuc.pt/media/12764/relatório-contas-2018-compactado.pdf

ERSUC. (2018b). MUNICÍPIOS | ERSUC. Retrieved January 23, 2020, from http://ersuc.pt/ersuc/municípios/

ERSUC. (2018c). Perfil|ERSUC. Retrieved December 19, 2019, from http://ersuc.pt/ersuc/perfil/

ERSUC. (2018d). RECOLHA SELETIVA | ERSUC. Retrieved January 23, 2020, from http://ersuc.pt/áreas-de-negócio/recolha-seletiva/#

ERSUC. (2018e). Triagem de Materiais Recicláveis | ERSUC. Retrieved January 23, 2020, from http://ersuc.pt/áreas-de-negócio/tratamento-e-valorização-de-resíduos/triagem-de-materiais-recicláveis/triagem-de-materiais-recicláveis/#

Esmaeilian, B., Wang, B., Lewis, K., Duarte, F., Ratti, C., & Behdad, S. (2018). The future of waste management in smart and sustainable cities: A review and concept paper. *Waste Management*, *81*, 177–195. https://doi.org/10.1016/j.wasman.2018.09.047

European Environment Agency. (2012). *Principles of EU Environmental Law*. Retrieved from http://ec.europa.eu/environment/legal/law/pdf/principles/2 Polluter Pays Principle_revised.pdf

Faccio, M., Persona, A., & Zanin, G. (2011). Waste collection multi objective model with real time traceability data. *Waste Management*, *31*(12), 2391–2405. https://doi.org/10.1016/j.wasman.2011.07.005

Fuss, M., Vasconcelos Barros, R. T., & Poganietz, W. R. (2018). Designing a framework for municipal solid waste management towards sustainability in emerging economy countries - An application to a case study in Belo Horizonte (Brazil). *Journal of Cleaner Production*, *178*, 655–664. https://doi.org/10.1016/j.jclepro.2018.01.051

Goluboniv, E. (2018). Improved Smart Waste Management for Smart City - Inovatink - Medium. Retrieved August 2, 2019, from https://medium.com/inovatink/improved-smart-waste-management-for-smart-city-7387a11f6204

Gonçalves, D. (2014). *Tecnologias de Informação e Comunicação para Otimização da Recolha de Resíduos Recicláveis - Projeto de Mestrado em Gestão*. Retrieved from https://repositorio.iscte-iul.pt/handle/10071/8890

Google Developers. (2019). OR-Tools | Google Developers. Retrieved December 17, 2019, from https://developers.google.com/optimization/routing

Google Developers. (2020). Routing Options | OR-Tools | Google Developers. Retrieved February 25, 2020, from https://developers.google.com/optimization/routing/routing_options

Guerriero, F., Violi, A., Natalizio, E., Loscri, V., & Costanzo, C. (2011). Modelling and solving optimal placement problems in wireless sensor networks. *Applied Mathematical Modelling*, *35*(1), 230–241. https://doi.org/10.1016/j.apm.2010.05.020

Guestrin, C., Krause, A., & Singh, A. P. (2005). Near-optimal sensor placements in gaussian processes. *ICML 2005 - Proceedings of the 22nd International Conference on Machine Learning*, *1*, 265–272. https://doi.org/10.1145/1102351.1102385

Hoornweg, D., & Bhada-Tata, P. (2012). What a waste: a global review of solid waste management. *World Bank, Washington DC*.

Hoos, H. H., & Stützle, T. (2005). Introduction. In *Stochastic Local Search - Foundation and Applications* (pp. 13–59). Elsevier Inc. https://doi.org/10.1016/b978-155860872-6/50018-4

INE. (2012). *Instituto Nacional de Estatistica, Censos 2011 - Resultados Definitivos - Região Centro*. Retrieved from https://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=ine_censos_publicacao_det&contexto=pu&PUBLICACOESpub_boui=156644135&PUBLICACOESmodo=2&selTab=tab1&pcensos=61969554

ION Data Services. (2020). Resolving Skewness in R. Retrieved February 3, 2020, from https://ionds.com/skewness-in-r/

Ion, I., & Gheorghe, F. F. (2014). The Innovator Role of Technologies in Waste Management towards the Sustainable Development. *Procedia Economics and Finance*, *8*, 420–428. https://doi.org/10.1016/s2212-5671(14)00109-9

Jha, A. K., Singh, S. K., Singh, G. P., & Gupta, P. K. (2011). Sustainable municipal solid waste management in low income group of cities: A review. *Tropical Ecology*, *52*(1), 123–121.

Johansson, O. M. (2006). The effect of dynamic scheduling and routing in a solid waste management system. *Waste Management*, *26*(8), 875–885. https://doi.org/10.1016/j.wasman.2005.09.004

Jünger, M., Reinelt, G., & Rinaldi, G. (1995). The traveling salesman problem. In M. O. Ball, M. T.L., C. L. Monma, & G. L. Nemhauser (Eds.), *Handbooks in Operations Research and Management Science* (Vol. 7, pp. 225–330). https://doi.org/10.1016/S0927-0507(05)80121-5

Kaza, S., Yao, L., Bhada-Tata, P., & Van Woerden, F. (2018). *What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050*. https://doi.org/10.1596/978-1-4648-1329-0

Krikke, H., le Blanc, I., van Krieken, M., & Fleuren, H. (2008). Low-frequency collection of materials disassembled from end-of-life vehicles. On the value of on-line monitoring in optimizing route planning. *International Journal of Production Economics*, *111*(2), 209–228. https://doi.org/10.1016/j.ijpe.2006.10.015

Kumar, S. N., & Panneerselvam, R. (2012). A Survey on the Vehicle Routing Problem and Its Variants. *Intelligent Information Management*, *04*(03), 66–74. https://doi.org/10.4236/iim.2012.43010

Laporte, G. (2010). The Traveling Salesman Problem, the Vehicle Routing Problem, and Their Impact on Combinatorial Optimization. *International Journal of Strategic Decision Sciences*, *1*(2), 82–92. https://doi.org/10.4018/jsds.2010040104

Law, A. M. (2011). A tutorial on how to select simulation input probability distributions. *Proceedings - Winter Simulation Conference*, *0*, 103–117. https://doi.org/10.1109/WSC.2016.7822083

Leitão, M. (2017). *Effects of Extended Producer Responsibility policies implementation - The case of Portuguese packaging waste system*. Instituto Superior Técnico. Retrieved from https://fenix.tecnico.ulisboa.pt/downloadFile/1970719973966881/Dissertacao - Manuel Leitao 76540 (Versao Finall).pdf

Lexico.com. (2019). Internet of things | Definition of Internet of things in English by Lexico Dictionaries. Retrieved August 2, 2019, from https://www.lexico.com/en/definition/internet_of_things

Linhart, H. (1965). Approximate Confidence Limits for the Coefficient of Variation of Gamma Distributions. *Biometrics*, *21*(3), 733–738.

Liu, K. H., Shih, S. Y., & Kao, J. J. (2011). Planning for hazardous campus waste collection. *Journal of Hazardous Materials*, *189*(1–2), 363–370. https://doi.org/10.1016/j.jhazmat.2011.02.046

Liu, W. Y., Lin, C. C., Chiu, C. R., Tsao, Y. S., & Wang, Q. (2014). Minimizing the carbon footprint for the time-dependent heterogeneous-fleet vehicle routing problem with alternative paths. *Sustainability (Switzerland)*, *6*(7), 4658–4684. https://doi.org/10.3390/su6074658

Longhi, S., Marzioni, D., Alidori, E., Di Buo, G., Prist, M., Grisostomi, M., & Pirro, M. (2012). Solid Waste Management Architecture Using Wireless Sensor Network Technology. In *2012 5th International Conference on New Technologies, Mobility and Security (NTMS)* (pp. 1–5). IEEE. https://doi.org/10.1109/NTMS.2012.6208764

Lundin, A. C., Ozkil, A. G., & Schuldt-Jensen, J. (2017). Smart Cities: A Case Study in Waste Monitoring and Management. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, 1392–1401. https://doi.org/10.24251/hicss.2017.167

Magrinho, A., Didelet, F., & Semiao, V. (2006). Municipal solid waste disposal in Portugal. *Waste Management*, *26*(12), 1477–1489. https://doi.org/10.1016/j.wasman.2006.03.009

Marchiori, M. (2018). The Smart Cheap City: Efficient Waste Management on a Budget. *Proceedings - 2017 IEEE 19th Intl Conference on High Performance Computing and Communications, HPCC 2017, 2017 IEEE 15th Intl Conference on Smart City, SmartCity 2017 and 2017 IEEE 3rd Intl Conference on Data Science and Systems, DSS 2017*, *2018-Janua*, 192–199. https://doi.org/10.1109/HPCC-SmartCity-DSS.2017.25

Markov, I., Bierlaire, M., Cordeau, J.-F., Maknoon, Y., & Varone, S. (2019). Waste Collection Inventory Routing with Non-stationary Stochastic Demands. *Computers & Operations Research*, *113*, 104798. https://doi.org/10.1016/j.cor.2019.104798

Marten, G. (2003). Human Ecology - Basic Concepts for Sustainable Development. *International Journal of Sustainability in Higher Education*, *4*(2), ijshe.2003.24904bae.004. https://doi.org/10.1108/ijshe.2003.24904bae.004

Martínez, T. G., Bräutigam, K. R., & Seifert, H. (2012). The potential of a sustainable municipal waste management system for Santiago de Chile, including energy production from waste. *Energy, Sustainability and Society*, *2*(1), 1–14. https://doi.org/10.1186/2192-0567-2-24

McLeod, F., Erdogan, G., Cherrett, T., Bektas, T., Davies, N., Shingleton, D., … Norgate, S. (2014). Improving collection efficiency through remote monitoring of charity assets. *Waste Management*, *34*(2), 273–280. https://doi.org/10.1016/j.wasman.2013.11.006

McLeod, F., Erdogan, G., Cherrett, T., Bektas, T., Davies, N., Speed, C., … Norgate, S. (2013). Dynamic collection scheduling using remote asset monitoring. *Transportation Research Record*, (2378), 65–72. https://doi.org/10.3141/2378-07

Mes, M. (2012). Using Simulation to Assess the Opportunities of Dynamic Waste Collection. In S. Bangsow (Ed.), *Use Cases of Discrete Event Simulation* (pp. 277–307). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28777-0_13

Mes, M., Schutten, M., & Rivera, A. P. (2014). Inventory routing for dynamic waste collection. *Waste Management*, *34*(9), 1564–1576. https://doi.org/10.1016/j.wasman.2014.05.011

Mmereki, D., Baldwin, A., & Li, B. (2016). A comparative analysis of solid waste management in developed, developing and lesser developed countries. *Environmental Technology Reviews*, *5*(1), 120–141. https://doi.org/10.1080/21622515.2016.1259357

Munagala, K. (2009). Lecture 1 : Sensor Placement and Submodularity Submodularity. Retrieved from https://users.cs.duke.edu/~kamesh/Lec_Fall09.pdf

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (1993). *Population Summit of the World's Scientific Academies*. Washington, D.C.: National Academies Press. https://doi.org/10.17226/9148

Nguyen, T. T. T., & Wilson, B. G. (2010). Fuel consumption estimation for kerbside municipal solid waste (MSW) collection activities. *Waste Management and Research*, *28*(4), 289–297. https://doi.org/10.1177/0734242X09337656

Niska, H., & Serkkola, A. (2018). Data analytics approach to create waste generation profiles for waste management and collection. *Waste Management*, *77*, 477–485. https://doi.org/10.1016/j.wasman.2018.04.033

Niza, S., Santos, E., Costa, I., Ribeiro, P., & Ferrão, P. (2014). Extended producer responsibility policy in Portugal: A strategy towards improving waste management performance. *Journal of Cleaner Production*, *64*, 277–287. https://doi.org/10.1016/j.jclepro.2013.07.037

Nourinejad, M., Malfara, N., & Roorda, M. J. (2018). Remote Assessment Sensor Routing: An Application for Waste Management. In E. Taniguchi & R. G. Thompson (Eds.), *City Logistics 1 - New Opportunities and Challenges* (1st ed., pp. 125–146). ISTE Ltd.

https://doi.org/10.1002/9781119425519.ch7

Omara, A., Gulen, D., Kantarci, B., & Oktug, S. F. (2018). Trajectory-assisted municipal agent mobility: A sensor-driven smart waste management system. *Journal of Sensor and Actuator Networks*, *7*(3). https://doi.org/10.3390/jsan7030029

Oxford Unity Press. (2013). A dictionary of environment and conservation. *Choice Reviews Online*, *45*(05), 45-2369-45–2369. https://doi.org/10.5860/choice.45-2369

Pang, W. K., Leung, P. K., Huang, W. K., & Liu, W. (2005). On interval estimation of the coefficient of variation for the three-parameter Weibull, lognormal and gamma distribution: A simulation-based approach. *European Journal of Operational Research*, *164*(2), 367–377. https://doi.org/10.1016/j.ejor.2003.04.005

Pardini, K., Rodrigues, J. J. P. C., Kozlov, S. A., Kumar, N., & Furtado, V. (2019). IoT-Based Solid Waste Management Solutions: A Survey. *Journal of Sensor and Actuator Networks*, *8*(1), 5. https://doi.org/10.3390/jsan8010005

Pássaro, D. A. (2003). Report: waste management in Portugal between 1996 and 2002. *Waste Management*, *23*(1), 97–99. https://doi.org/10.1016/s0956-053x(02)00142-3

Pearn, W. L., Assad, A., & Golden, B. L. (1987). Transforming arc routing into node routing problems. *Computers and Operations Research*, *14*(4), 285–288. https://doi.org/10.1016/0305-0548(87)90065-7

Pires, A., Martinho, G., Rodrigues, S., & Gomes, M. I. (2018). *Sustainable solid waste collection and management*. *Sustainable Solid Waste Collection and Management*. https://doi.org/10.1007/978-3-319-93200-2

PORDATA. (2019a). PORDATA - Densidade populacional. Retrieved January 23, 2020, from https://www.pordata.pt/Municipios/Densidade+populacional-452

PORDATA. (2019b). PORDATA - Resíduos urbanos recolhidos por habitante: total e selectivamente. Retrieved December 27, 2019, from https://www.pordata.pt/Portugal/Resíduos+urbanos+recolhidos+por+habitante+total+e+selectivamente-1229

Purcell, M., & Magette, W. L. (2009). Prediction of household and commercial BMW generation according to socio-economic and other factors for the Dublin region. *Waste Management*, *29*(4), 1237–1250. https://doi.org/10.1016/j.wasman.2008.10.011

Ramos, T. R. P., de Morais, C. S., & Barbosa-Póvoa, A. P. (2018, August 1). The smart waste collection routing problem: Alternative operational management approaches. *Expert Systems with Applications*. Pergamon. https://doi.org/10.1016/j.eswa.2018.03.001

Recycle More. (2019). The Waste Hierarchy. Retrieved April 22, 2019, from https://www.recycle-more.co.uk/business-zone/office-recycling-and-waste-management/the-waste-hierarchy

Regents of the University of Michigan. (2017). Municipal Solid Waste Factsheet | Center for Sustainable Systems. Retrieved April 6, 2019, from http://css.umich.edu/factsheets/municipal-solid-waste-factsheet

Rivlin, O. (2019). Reinforcement Learning for Combinatorial Optimization. Retrieved November 8, 2019, from https://towardsdatascience.com/reinforcement-learning-for-combinatorial-optimization-d1402e396e91

Schott, D. (2014). Urban Development and Environment. In M. Agnoletti & S. Neri Serneri (Eds.), *The Basic Environmental History* (pp. 171–198). Springer. https://doi.org/10.1007/978-3-319-09180-8_6

Sequeiros, B. A. H. L. de. (2012). *Aplicação de indicadores de desempenho nos serviços de resíduos urbanos*. UL. Instituto Superior Técnico. Retrieved from https://bibliotecas.utl.pt/cgi-bin/koha/opac-detail.pl?biblionumber=486723

Sociedade Ponto Verde. (2015). Sociedade Ponto Verde - Quem Somos. Retrieved April 22, 2019, from https://www.pontoverde.pt/quem_somos.php

Stellingwerff, A. (2011). *Dynamic waste collection - Assessing the usage of dynamic routing methodologies*. University of Twente. Retrieved from https://essay.utwente.nl/61129/

Straightsol. (2012). *Deliverable 4.1 - Monitoring of demonstration achievements - first period*. Retrieved from http://www.straightsol.eu/deliverables.htm

Straightsol. (2013). *Deliverable 2.1 - Urban freight and urban – interurban interfaces. Best practices, implications and future needs*. Retrieved from http://www.straightsol.eu/deliverables.htm

Straightsol. (2014a). *Deliverable 5.1 - Demonstration Assessments*. Retrieved from http://www.straightsol.eu/deliverables.htm

Straightsol. (2014b). *Deliverable 5.3 - Business Models for innovative and sustainable urban-interurban transport, Report*. Retrieved from http://www.straightsol.eu/deliverables.htm

Straightsol. (2014c). *Deliverable 5.3 Addendum 2 - Business concepts: critical design options for implementing urban logistics solutions*. Retrieved from http://www.straightsol.eu/deliverables.htm

Straightsol. (2015). Straightsol - Project Demonstrations. Retrieved September 30, 2019, from http://www.straightsol.eu/demonstration_C.htm

Summers, T. H., & Lygeros, J. (2014). *Optimal sensor and actuator placement in complex dynamical networks*. *IFAC Proceedings Volumes (IFAC-PapersOnline)* (Vol. 19). IFAC. https://doi.org/10.3182/20140824-6-za-1003.00226

Tchobanoglous, G., & Kreith, F. (2002). *Handbook of Solid Waste Management*. *Environmental Pollutants and Their Bioremediation Approaches*. McGraw-Hill Professional. https://doi.org/10.1036/0071356231

Teixeira, C. A., Russo, M., Matos, C., & Bentes, I. (2014). Evaluation of operational, economic, and environmental performance of mixed and selective collection of municipal solid waste: Porto case study. *Waste Management and Research*, *32*(12), 1210–1218. https://doi.org/10.1177/0734242X14554642

The Environmental Literacy Council. (2015). What is Waste? - The Environmental Literacy Council. Retrieved April 6, 2019, from https://enviroliteracy.org/environment-society/waste-management/what-is-waste/

Vallero, D. A., McLeod, F., & Cherrett, T. (2019). *Waste Collection*. *Waste* (2nd ed.). Elsevier Inc. https://doi.org/10.1016/B978-0-12-381475-3.10004-X

Vergara, S. E., & Tchobanoglous, G. (2012). *Municipal Solid Waste and the Environment: A Global Perspective*. *Ssrn*. https://doi.org/10.1146/annurev-environ-050511-122532

Vila Nova de Poiares. (2017). Vila Nova de Poiares - Município de Poiares e ERSUC expandem rede de ecopontos no Concelho. Retrieved January 27, 2020, from http://www.cm-vilanovadepoiares.pt/noticias2-2/1454-municipio-de-poiares-e-ersuc-expandem-rede-de-ecopontos-no-concelho

Vonolfen, S., Affenzeller, M., Beham, A., Wagner, S., & Lengauer, E. (2011). Simulation-based evolution of municipal glass-waste collection strategies utilizing electric trucks. *LINDI 2011 - 3rd IEEE International Symposium on Logistics and Industrial Informatics, Proceedings*, 177–182. https://doi.org/10.1109/LINDI.2011.6031142

Vukmirović, S., Čapko, Z., & Babić, A. (2019). Model of Using the Exhaustive Search Algorithm in Solving of Traveling Salesman Problem (TSP) on the Example of the Transport Network Optimization of Primorje-Gorski Kotar County (PGC). In D. Tipurić & D. Hruška (Eds.), *7th International OFEL Conference on Governance, Management and Entrepreneurship: Embracing Diversity in Organisations. April 5th - 6th, 2019, Dubrovnik, Croatia* (pp. 391–401). Zagreb: Governance Research and Development Centre. Retrieved from https://www.econstor.eu/handle/10419/196099

Wikipedia. (2020). Gamma distribution. Retrieved February 4, 2020, from https://en.wikipedia.org/wiki/Gamma_distribution

Wilson, D. C. (2007). Development drivers for waste management. *Waste Management & Research*, *25*(3), 198–207. https://doi.org/10.1177/0734242X07079149

Wilson, D. G. (1976). A brief history of solid-waste management. *International Journal of Environmental Studies*, *9*(2), 123–129. https://doi.org/10.1080/00207237608737618

WSmart Route. (2019). WSmart Route - Waste Collection Based on a Real Time Route Planning System. Retrieved February 7, 2020, from http://wsmartroute.tecnico.ulisboa.pt/

Zhang, T., Siebers, P. O., & Aickelin, U. (2012). A three-dimensional model of residential energy consumer archetypes for local energy policy design in the UK. *Energy Policy*, *47*, 102–110. https://doi.org/10.1016/j.enpol.2012.04.027

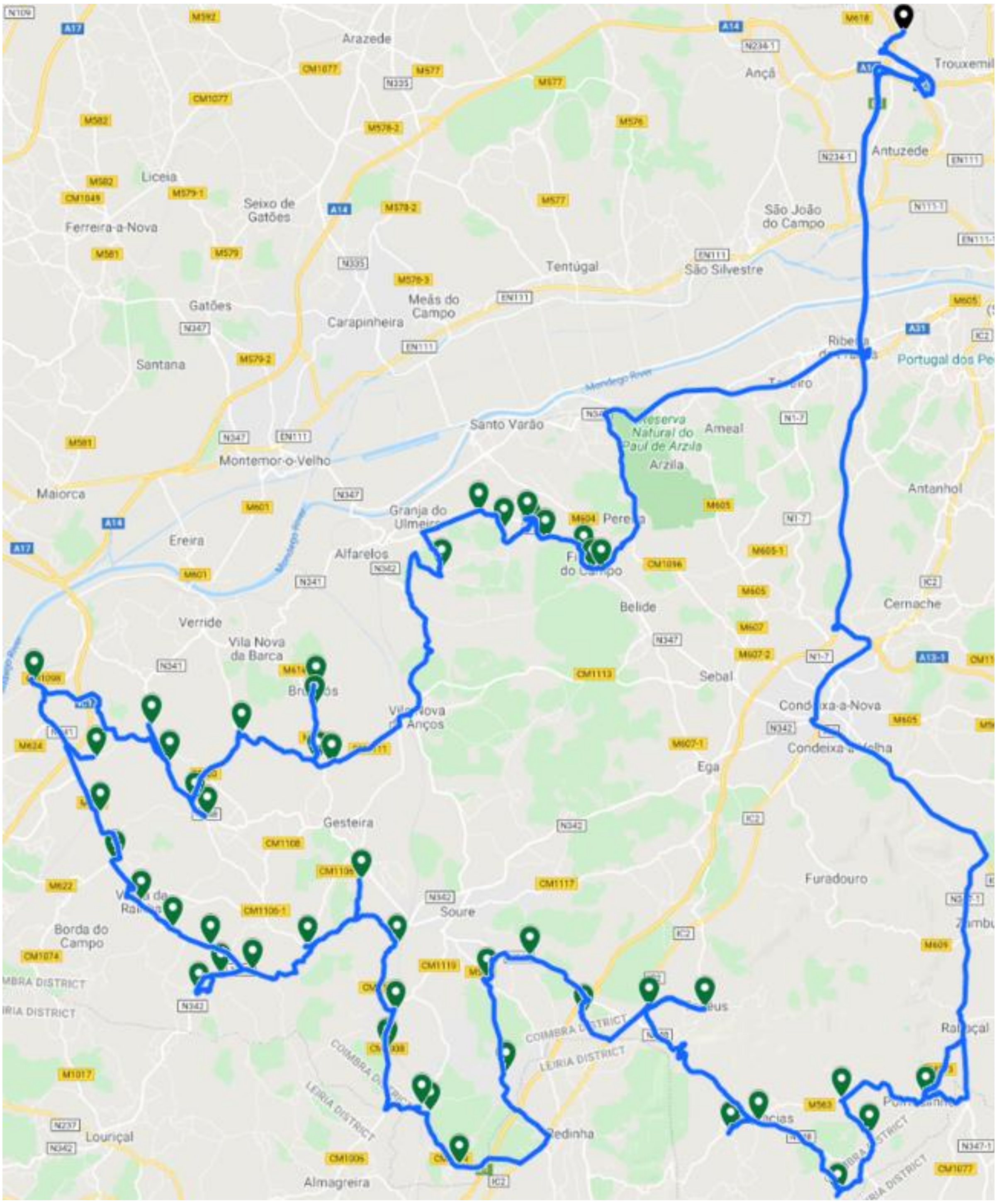
Zsigraiova, Z., Semiao, V., & Beijoco, F. (2013). Operation costs and pollutant emissions reduction by definition of new collection scheduling and optimization of MSW collection routes using GIS. The case study of Barreiro, Portugal. *Waste Management*, *33*(4), 793–806. https://doi.org/10.1016/j.wasman.2012.11.015

Appendix

Appendix A – Current routes

Route for group A:

DEPOT → C1 → C2 → … → C12 → C14 → C13 → … → C49 → C50 → DEPOT



Route for group B:

DEPOT → C51 → C52 → … → C97 → C98 → DEPOT

Appendix B – Waste demand: average and variability

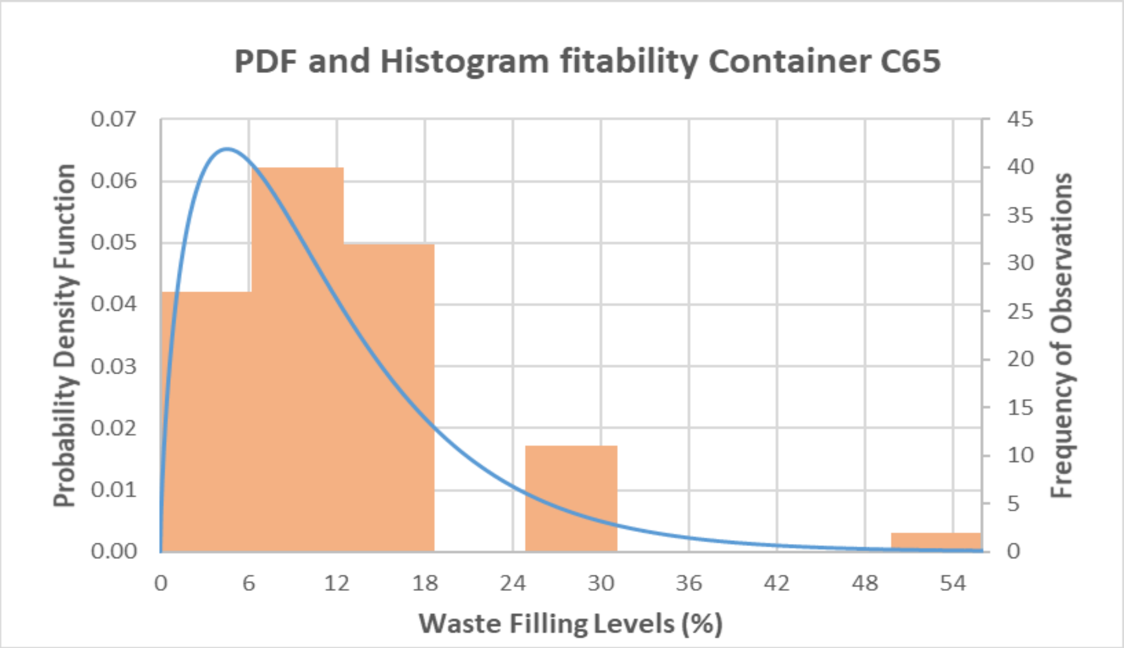
Container	Mean ($\bar{x}$) (%)	Sd. Dev. (s)	Variance (s^2)	C. Variation ($\widehat{C_V}$)
C1	4.75	5.50	30.30	1.16
C2	3.82	6.54	42.83	1.71
C3	7.06	8.27	68.32	1.17
C4	3.59	8.48	71.86	2.36
C5	4.98	4.33	18.72	0.87
C6	9.38	7.03	49.36	0.75
C7	9.84	9.99	99.88	1.02
C8	6.37	5.67	32.16	0.89
C9	4.05	4.60	21.13	1.13
C10	5.44	6.80	46.21	1.25
C11	3.59	6.22	38.64	1.73
C12	10.07	5.77	33.31	0.57
C13	8.45	5.64	31.78	0.67
C14	6.83	7.46	55.63	1.09
C15	5.32	5.42	29.34	1.02
C16	2.60	3.06	9.36	1.18
C17	6.31	5.98	35.80	0.95
C18	6.81	6.85	46.88	1.01
C19	2.85	3.00	8.98	1.05
C20	4.33	7.68	58.96	1.77
C21	3.34	5.40	29.18	1.62
C22	3.84	7.03	49.41	1.83
C23	3.06	2.94	8.64	0.96
C24	2.55	2.89	8.35	1.13
C25	4.64	7.52	56.60	1.62
C26	7.05	7.88	62.13	1.12
C27	3.84	4.05	16.38	1.05
C28	5.67	6.84	46.75	1.21
C29	5.32	7.75	60.12	1.46
C30	11.47	12.96	167.92	1.13
C31	8.38	6.86	47.03	0.82
C32	4.26	4.82	23.19	1.13
C33	3.99	5.81	33.77	1.46
C34	4.46	7.15	51.12	1.60
C35	4.84	8.89	79.06	1.84
C36	4.30	7.29	53.15	1.69
C37	3.19	5.38	28.94	1.68
C38	12.23	10.78	116.24	0.88
C39	4.79	5.72	32.70	1.19
C40	9.31	6.64	44.10	0.71
C41	3.19	5.63	31.75	1.77
C42	7.53	6.63	43.90	0.88

C43	2.93	2.60	6.74	0.88
C44	5.99	7.45	55.45	1.24
C45	6.25	7.10	50.35	1.14
C46	4.52	6.28	39.39	1.39
C47	13.14	10.02	100.43	0.76
C48	3.95	5.88	34.55	1.49
C49	5.82	8.17	66.77	1.40
C50	4.34	6.87	47.16	1.58
C51	2.10	2.93	8.58	1.39
C52	3.09	5.67	32.19	1.83
C53	2.60	3.40	11.59	1.31
C54	2.85	3.35	11.21	1.18
C55	3.34	3.58	12.79	1.07
C56	2.35	3.08	9.49	1.31
C57	3.54	4.59	21.07	1.30
C58	5.66	7.55	56.99	1.33
C59	7.31	9.77	95.38	1.34
C60	9.67	11.37	129.22	1.18
C61	5.19	6.61	43.67	1.27
C62	2.83	5.61	31.48	1.98
C63	3.07	5.88	34.62	1.92
C64	7.08	7.80	60.87	1.10
C65	11.08	8.51	72.36	0.77
C66	4.01	4.58	21.01	1.14
C67	7.08	7.23	52.33	1.02
C68	5.42	7.85	61.64	1.45
C69	4.48	5.01	25.08	1.12
C70	4.01	7.37	54.33	1.84
C71	4.25	6.83	46.70	1.61
C72	4.01	7.78	60.52	1.94
C73	3.54	6.73	45.25	1.90
C74	6.37	7.00	49.05	1.10
C75	3.77	7.67	58.85	2.03
C76	5.42	5.67	32.17	1.05
C77	4.72	6.32	39.96	1.34
C78	3.30	7.73	59.70	2.34
C79	3.33	6.31	39.81	1.89
C80	5.66	7.38	54.47	1.30
C81	8.25	10.05	101.07	1.22
C82	2.83	3.96	15.67	1.40
C83	7.31	8.70	75.76	1.19
C84	7.31	7.75	59.99	1.06
C85	4.72	6.77	45.90	1.44
C86	5.19	6.70	44.84	1.29
C87	3.61	6.03	36.42	1.67

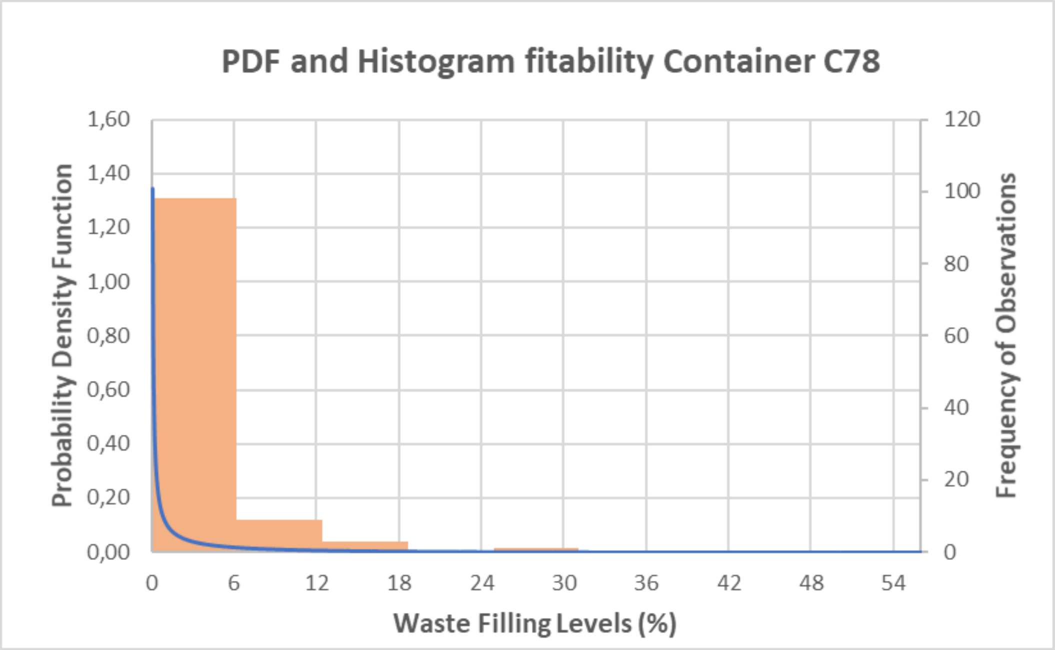
C88	2.40	2.71	7.34	1.13
C89	4.25	5.48	30.02	1.29
C90	1.90	1.51	2.27	0.79
C91	5.42	8.08	65.26	1.49
C92	5.00	8.71	75.90	1.74
C93	6.43	7.60	57.81	1.18
C94	2.38	3.41	11.66	1.43
C95	3.54	5.88	34.53	1.66
C96	4.01	8.90	79.16	2.22
C97	11.08	10.08	101.69	0.91
C98	6.84	8.26	68.18	1.21

Note: At green and red are marked the minimum and maximum values, respectively, for each column.

Appendix C – Histogram vs PDF (2 examples more)



(9 intervals of width 6)



(14 intervals of width 6)

Appendix D – Threshold level and collection interval

Container	Threshold level (ψ)	Collection Interval (n)
C1	94.5	24
C2	96.5	33
C3	91.8	17
C4	98.1	40
C5	93.9	22
C6	88.5	11
C7	88.2	11
C8	92.2	17
C9	95.2	28
C10	93.8	22
C11	96.7	35
C12	87.9	10
C13	89.7	12
C14	91.9	17
C15	93.6	21
C16	97	43
C17	92.4	18
C18	91.8	17
C19	96.6	39
C20	96.2	30
C21	96.7	37
C22	96.7	34
C23	96.3	36

C24	97	44
C25	95.5	27
C26	91.7	16
C27	95.4	29
C28	93.5	21
C29	94.4	23
C30	86.5	10
C31	89.8	13
C32	95	27
C33	95.8	30
C34	95.6	28
C35	95.9	28
C36	96	30
C37	97	39
C38	85.1	9
C39	94.5	24
C40	88.6	11
C41	97.2	40
C42	90.8	14
C43	96.4	37
C44	93.2	20
C45	92.7	19
C46	95.1	27
C47	83.9	8
C48	95.9	31
C49	93.7	21
C50	95.7	29
C51	97.7	55
C52	97.4	41
C53	97.1	44
C54	96.7	40
C55	96	34
C56	97.4	49
C57	96	33
C58	93.7	21
C59	91.9	17
C60	88.8	12
C61	94.1	23
C62	97.8	46
C63	97.5	42
C64	91.6	16
C65	86.5	9
C66	95.3	29
C67	91.5	16

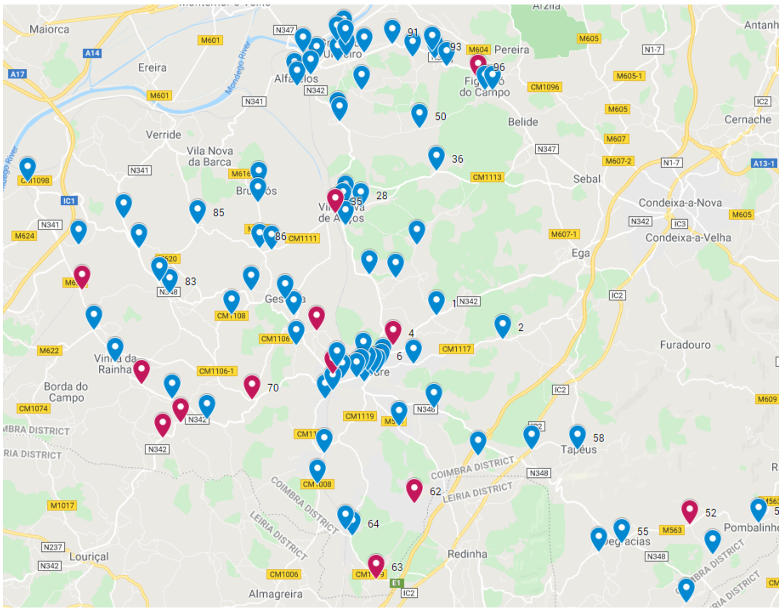
C68	94.3	23
C69	94.7	26
C70	96.6	33
C71	95.8	30
C72	96.8	33
C73	97.1	37
C74	92.5	18
C75	97.3	36
C76	93.5	21
C77	94.8	25
C78	98.2	42
C79	97.3	39
C80	93.7	21
C81	90.5	15
C82	96.9	42
C83	91.5	16
C84	91.3	16
C85	95	26
C86	94.2	23
C87	96.6	35
C88	97.2	47
C89	95.2	28
C90	97.7	56
C91	94.4	23
C92	95.5	26
C93	92.5	18
C94	97.5	49
C95	96.6	35
C96	97.5	35
C97	86.5	10
C98	92.1	17

Note: At green are marked the minimum values for the collection interval, for each route.

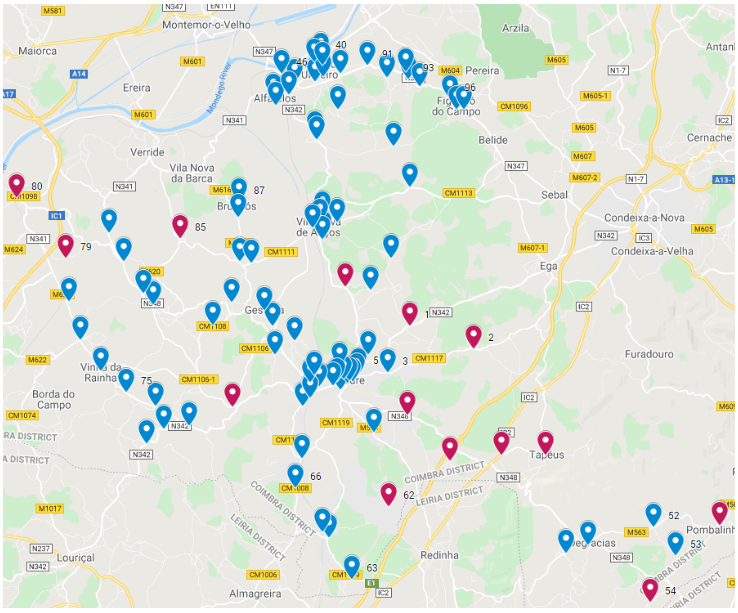
Appendix E – The selection of containers

Rate: [90, 51, 56, 94, 88, 24, 16, 53, 62, 82, 19, 54, 43, 23, 63, 52, 37, 41, 78, 79, 21, 55, 57, 73, 95, 4, 11, 87, 75, 2, 22, 27, 48, 33, 66, 70, 72, 96, 9, 71, 89, 32, 36, 20, 50, 34, 69, 46, 25, 77, 85, 1, 39, 35, 5, 92, 61, 86, 15, 29, 68, 76, 91, 10, 58, 80, 28, 49, 44, 45, 17, 8, 74, 93, 18, 14, 98, 26, 3, 64, 67, 59, 83, 84, 42, 81, 31, 13, 40, 6, 60, 7, 12, 65, 97, 30, 38, 47]

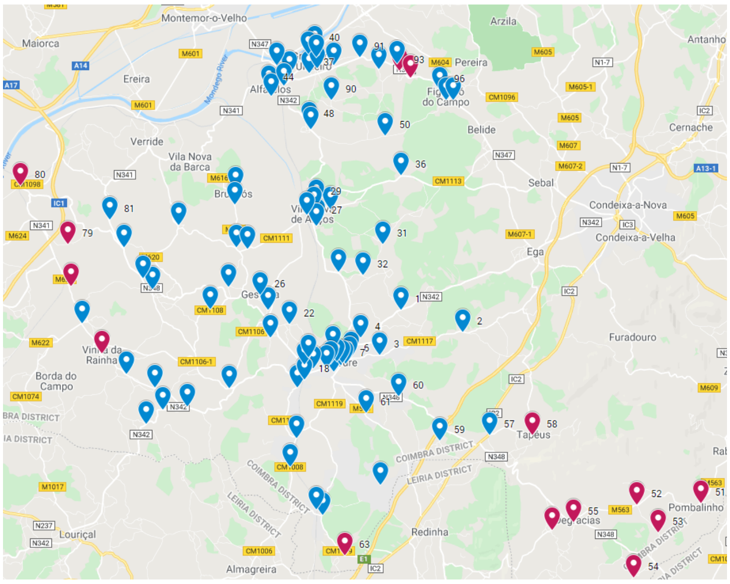
Variability: [4, 78, 96, 75, 62, 72, 63, 73, 79, 70, 35, 52, 22, 20, 41, 92, 11, 2, 36, 37, 87, 95, 25, 21, 71, 34, 50, 91, 48, 29, 33, 68, 85, 94, 49, 82, 51, 46, 77, 59, 58, 56, 53, 80, 57, 86, 89, 61, 10, 44, 81, 98, 28, 39, 83, 93, 16, 60, 54, 3, 1, 66, 45, 9, 24, 32, 30, 88, 26, 69, 64, 74, 14, 55, 84, 27, 19, 76, 67, 15, 7, 18, 23, 17, 97, 8, 43, 38, 42, 5, 31, 90, 65, 47, 6, 40, 13, 12]



Remoteness 1: [2, 80, 58, 62, 79, 85, 70, 57, 59, 54, 51, 1, 60, 33, 61, 31, 32, 36, 50, 78, 63, 77, 23, 24, 90, 52, 53, 22, 69, 76, 74, 75, 81, 82, 66, 67, 72, 3, 4, 55, 56, 39, 91, 68, 92, 71, 73, 95, 25, 26, 21, 28, 37, 38, 34, 14, 42, 20, 46, 83, 84, 43, 47, 17, 27, 87, 88, 64, 65, 96, 86, 89, 93, 94, 35, 10, 13, 44, 45, 5, 97, 98, 6, 40, 41, 29, 30, 15, 16, 48, 49, 7, 8, 9, 11, 12, 18, 19]

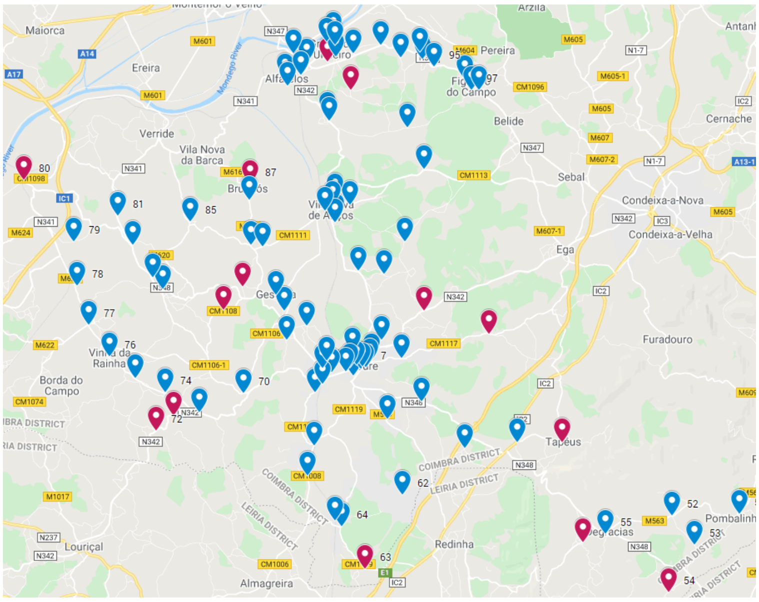


Remoteness 2: [51, 54, 53, 52, 56, 55, 80, 58, 63, 79, 78, 95, 93, 76, 94, 96, 77, 65, 92, 72, 98, 97, 64, 75, 81, 91, 57, 46, 40, 90, 42, 82, 73, 41, 39, 44, 62, 74, 59, 43, 45, 38, 37, 50, 47, 71, 66, 84, 83, 2, 36, 85, 87, 48, 67, 49, 88, 60, 61, 32, 31, 70, 24, 1, 23, 28, 68, 86, 89, 4, 3, 29, 33, 18, 19, 27, 14, 34, 30, 21, 35, 69, 22, 20, 17, 9, 25, 26, 13, 8, 15, 10, 12, 5, 16, 6, 7, 11]



Cost Reduction: [54, 2, 80, 63, 58, 90, 72, 85, 1, 56, 73, 25, 46, 26, 87, 88, 14, 59, 28, 81, 4, 23, 24, 83, 53, 5, 22, 94, 29, 57, 30, 27, 43, 13, 21, 44, 45, 47, 69, 20, 71, 40, 37, 12, 34, 9, 41, 79, 82, 78, 42, 38, 17, 16, 39, 8, 15, 10, 97, 98, 96, 7, 77, 84, 76, 75, 74, 70, 31, 32, 33, 86, 89, 3, 60, 61, 62, 35, 6, 18, 19, 68, 11, 36, 50, 48, 49, 51, 52, 55, 64, 65, 66, 67, 91, 92, 93, 95]

Cost Reduction considering frequencies: [54, 2, 80, 63, 90, 24, 23, 72, 58, 56, 1, 73, 37, 87, 88, 26, 25, 22, 4, 83, 69, 21, 53, 28, 14, 20, 59, 94, 81, 57, 29, 46, 43, 5, 44, 45, 27, 30, 40, 41, 42, 79, 82, 78, 71, 85, 17, 34, 13, 39, 16, 9, 8, 15, 10, 38, 12, 77, 84, 7, 97, 76, 75, 74, 70, 33, 32, 31, 98, 96, 86, 89, 3, 60, 61, 62, 35, 6, 18, 19, 68, 11, 36, 50, 47, 48, 49, 51, 52, 55, 64, 65, 66, 67, 91, 92, 93, 95]



Random Order Used: [74, 92, 63, 8, 82, 28, 37, 40, 35, 71, 19, 68, 97, 95, 27, 83, 1, 48, 96, 22, 12, 49, 87, 54, 6, 53, 90, 89, 43, 32, 70, 80, 73, 24, 58, 52, 75, 72, 64, 88, 29, 36, 98, 81, 69, 78, 44, 67, 38, 76, 11, 18, 51, 65, 86, 62, 13, 77, 9, 79, 2, 5, 56, 26, 7, 66, 33, 31, 57, 23, 59, 21, 3, 45, 20, 10, 30, 84, 60, 42, 47, 91, 50, 94, 17, 4, 34, 16, 85, 14, 39, 15, 61, 55, 46, 41, 93, 25]