Evaluating the Impact of Intermediate Facilities on Waste Collection Operations

Sebastian Ammon

A project report in partial fulfilment of the requirements for the degree

BACCALAREUS (INDUSTRIAL ENGINEERING)

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT, AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA

September 2017

Abstract

Given the importance of improving municipal solid waste management, the Optimisation Group from the University of Pretoria completed preliminary studies on data obtained from waste collection vehicle global position system traces from the City of Cape Town and found that some waste vehicles spend more time travelling between service locations and dump-sites as opposed to collecting waste. In order to improve these operations it was proposed that intermediate facilities be used to reduce the travel distance. The purpose of this project is to identify potential locations for intermediate waste facilities (transfer stations) by using a location model and then evaluating the impact of the intermediate waste facilities on waste vehicle collection operations.

This was achieved by splitting the project into three phases. Since some of the data that was required was not readily available, it had to be generated first. In phase one that data sets were generated. In phase two a variant of the location modelling formulation known as the multi-facility location problem was solved using the generalised Weiszfeld method for the multi-facility location problem. Up to 25 potential intermediate facility locations were identified and used for the remaining phase. In the next phase two models were built in order to quantitatively evaluate the impact of said facilities on waste collection operations. The first model was concerned with the travel distance and travel time for each collection vehicle associated to a single collection beat (service location). The second model was concerned with the waste allocation from the collection beats to either an intermediate facility or a landfill site and from the intermediate facilities to a landfill. The model was also used to conduct a cost benefit analysis.

It was concluded that as the number of intermediate facilities increased, the travel distance, travel time as well as the total transportation cost of the system decreased. Additionally, the location of the two existing intermediate facilities in the City of Cape Town were inefficient when compared to the same number of intermediate facilities located with the Weiszfeld algorithm. Furthermore, it was identified that the optimal number of facilities to locate, in the City of Cape Town, in order to reap the cost savings from improving the existing intermediate facility locations with respect to the capital cost of building new facilities, was three.

Contents

List of Figures

List of Tables

List of Algorithms

Acronyms

GIS Geographical Information System

GPS Global Positioning System

HLP Hub Location Problem

KNN K-Nearest Neighbour

LP Linear Programming

MFLP Multi-Facility Location Problem

MSW Municipal Solid Waste

VRP Vehicle Routing Problem

VRPIF Vehicle Routing Problem with Intermediate Facilities

Chapter 1

Introduction

1.1 Background

The importance of Municipal Solid Waste (MSW) management is not given much thought by the every day citizen who utilises the service. However, *"Solid waste management is the one thing just about every city government provides for its residents. While service levels, environmental impacts and costs vary dramatically, solid waste management is arguably the most important municipal service and serves as a prerequisite for other municipal action"* (Hoornweg and Bhada-Tata, 2012). Although MSW management might seem like a rather simple task it is highly complex consisting of various echelons, from municipal to governmental. Therefore, managing these activities must be done in an effective, efficient and sustainable manner.

Poorly managed waste operations can have a global impact on the environment due to greenhouse gas emissions arising from MSW. In fact, emissions from MSW is estimated to account for almost 5% of total global green house gas emissions (Hoornweg and Bhada-Tata, 2012). This of course includes emissions from waste collection operations and methane from landfills amongst others. Locally, uncollected solid waste contributes to flooding by blocking drainage, air pollution, and public health impacts by harbouring disease vectors (Hoornweg and Bhada-Tata, 2012). The importance of effective, efficient and sustainable MSW management does not only lie in reducing its impact on immediate and global environments, but also in reducing costs associated with it. As solid waste management is an important municipal service, the reduction in cost of MSW can help reduce financial pressures on local municipalities and the national government, while simultaneously maintaining adequate service delivery. The city of Cape Town for example budgeted R 128 million for the upgrade of solid waste facilities and R 80 million for the upgrade of drop-off facilities in the informal settlement areas for the year 2017/2018 alone (City of Cape Town, 2017).

The question might arise "Why is MSW management important now?". The answer lies in global as well as local MSW generation rates and urbanisation rates. According to Hoornweg and Bhada-Tata (2012) MSW "currently, world cities generate about 1.3 billion tonnes of solid waste per year. This volume is expected to increase to 2.2 billion tonnes by 2025." — generation levels are expected to effectively double by 2025. For the city of Cape Town, current MSW generation rates are expected to increase from 53 425 tonnes per day to 72 146 tons per day (City of Cape Town, 2017). Furthermore, MSW is growing even faster than the rate of urbanisation (Hoornweg and Bhada-Tata, 2012). This leads to problems, particularly in finding sites for new landfills or transfer stations, since locations which were previously unoccupied are now housing new occupants through urbanisation

(Bosompem et al., 2016).

Given the importance of improving MSW management, the Optimisation Group from the University of Pretoria completed preliminary studies on data obtained from waste vehicle global position system (GPS) traces and found that some waste vehicles spend most of their time travelling between service locations and dump-sites as opposed to actually collecting waste. It was reasoned that this is true in part because of the decentralised locations of landfills. The aim of this project was to identify potential locations for intermediate facilities (waste transfer stations) by using location models and then evaluating the impact of the transfer stations on waste vehicle operations. The City of Cape Town was used as the test case.

The implementation of transfer stations has the potential to improve the MSW by reducing the travelling time between service locations and dump-sites. Thus more time can be spent collecting waste and ultimately reducing the negative impact these inefficient operations have locally and globally.

1.2 Municipal Solid Waste Management and Intermediate Facilities

In this section the basic MSW management is explained and how intermediate facilities are linked in the process. This will give readers a better understanding of the proposed solution approach.

A waste management system consists of the following basic components (CSIR Building and Construction Technology, 2005):

- Waste Generation
- On-site Storage
- Collection
- Intermediate Facilities
- Incineration
- Recycling
- Disposal

How the components are interlinked is shown in the figure 1.1 in form of a typical waste cycle.

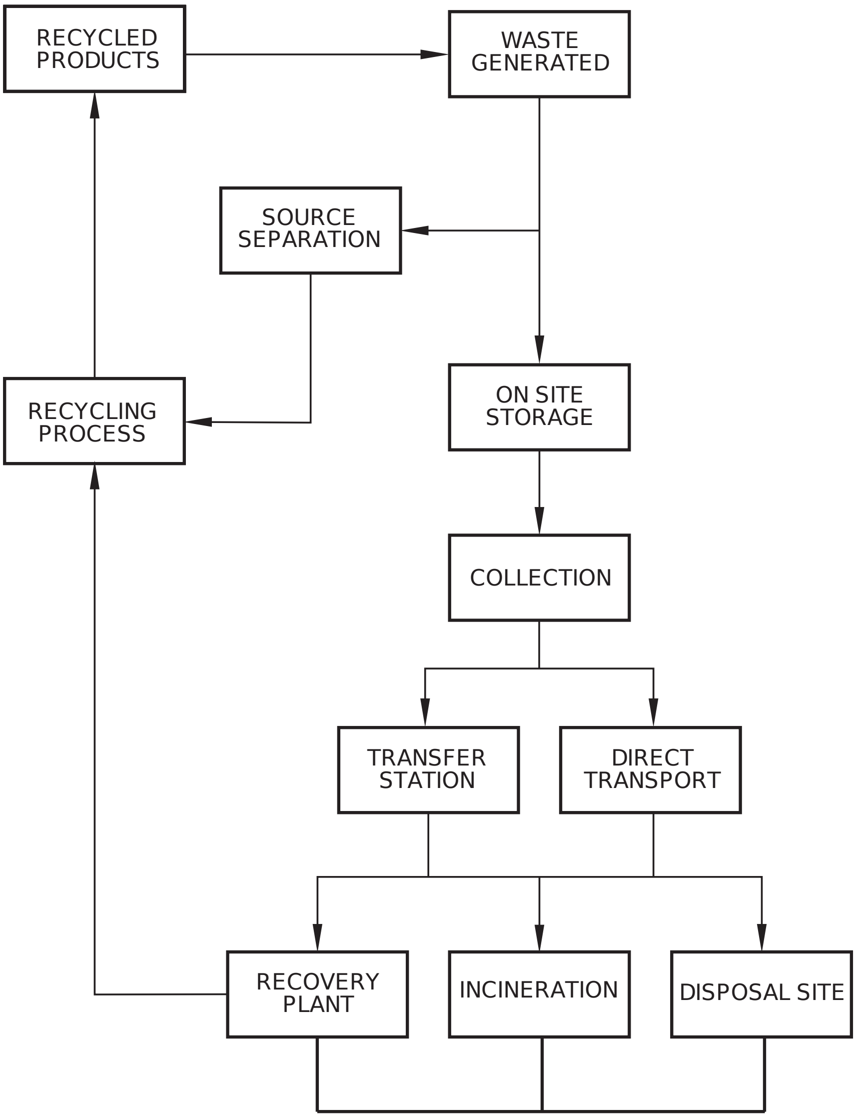
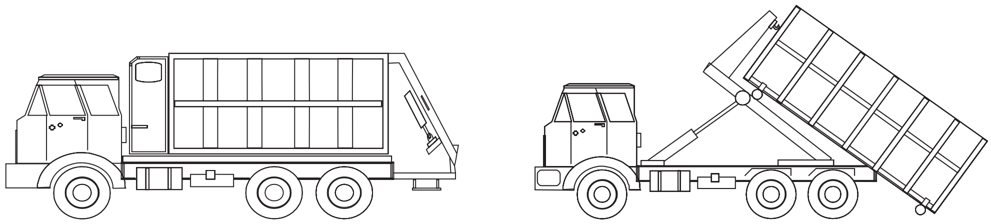


Figure 1.1: Typical waste cycle (CSIR Building and Construction Technology, 2005)

Waste is generated by people, for example at their homes (Waste Generation). The waste is then stored within their homes (On Site Storage) until garbage collection day where it is transferred to a 240L wheelie bin ready to be collected by garbage collection trucks (Collection). The garbage truck comes and collects the waste from the residences' homes by emptying the 240L bins in the back of the truck. The vehicle continues until it reaches its capacity. The driver either drops the waste off at the nearest facility (Recovery Plant, Incineration Plant or Disposal Site) through direct transport or at the nearest intermediate facility (transfer station). Waste at the transfer station is eventually collected by another vehicle which transfers it to the next facility (Recovery Plant, Incineration Plant or Disposal Site). The frequency of this operation depends on the municipalities' collection schedule and the volume of waste generated among other factors.

The reason intermediate facilities exist in the first place and that waste is not sent to other facilities directly is to make use of economies of scale. Where a typical rear-end loader refuse truck (Figure 1.2a) only has a capacity of $10m^3$ to $21m^3$, a roll-on roll-off vehicle (Figure 1.2b) has a capacity ranging from $18m^3$ to $30m^3$ (CSIR Building and Construction Technology, 2005).



(a) Rear-end loader refuse vehicle (b) Roll-on roll-off truck

Figure 1.2: Two types of vehicles used for waste collection operations (CSIR Building and Construction Technology, 2005)

The purpose is not only to reduce the transportation unit cost of collection vehicles, it also allows for quicker turn-around times and increased productivity. The need for intermediate facilities largely depends on the amount of waste generated, the collection system and the distance to disposal sites.

The City of Cape Town, which was used as the test case for the project, has five available disposable sites; three of which are landfills and the remainder are intermediate facilities (transfer stations). Their geographical locations can be seen in Figure 1.3.

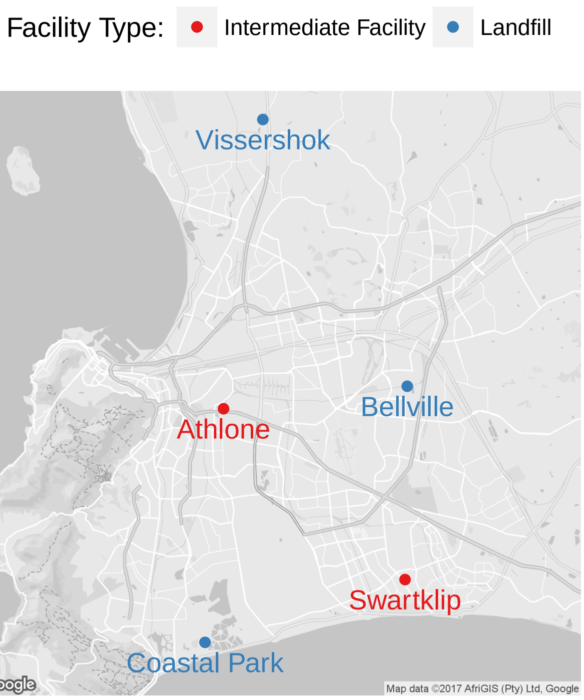


Figure 1.3: Disposable facilities

The collection schedule for the city is based on what are called collection beats. A waste collection beat depicts the area allocated to a single refuse compactor vehicle for the removal of waste from 240L bins at formal properties. Not all beats are serviced on the same day, the city has a collection schedule which is readily available online. The 723 collection beat areas can be seen in Figure 1.4.

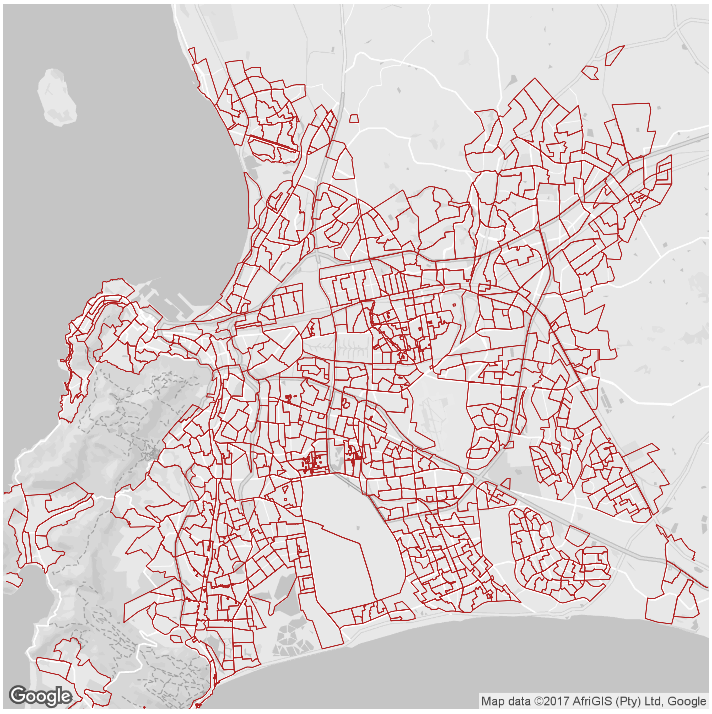


Figure 1.4: Beat outlines for the City of Cape Town

Intermediate facilities have the potential to improve the MSW collection operations for the City of Cape Town. By introducing intermediate facilities, the waste vehicles servicing a particular beat have the option to drop the waste off at a facility situated closer than the nearest landfill and may reduce the time required to travel between a drop-off facility and the collection beats.

The problem is then where and how many intermediate facilities should be located, so that the waste vehicles can better service the whole region.

1.3 Problem Statement

As MSW generation rates continue to increase, the importance of effective, efficient and sustainable MSW management systems becomes pertinent. It is therefore important to collect waste as quickly and efficiently as possible.

For the City of Cape Town it was identified that some waste collection vehicles spend more time travelling between collection beats (also known as service areas) and dump-sites than actually collecting waste. It is safe to say that such operations are neither effective, nor efficient. In light of the increasing rates of urbanisation and waste generation, although the collecting operations might not pose a serious problem now, in the future, increasing waste generation will put more pressure on the waste collection operations as well as local governments and municipalities as they try to uphold service delivery promises.

It was proposed that a possible solution to improve current waste collection operations was to introduce intermediate facilities. Intermediate facilities allow waste from several source to be consolidated and brought to a final destination at a later stage. This means that the vehicles which perform these operations need not drive to the final destination for each drop-off. Instead the vehicle can make shorter trips to a closer intermediate facility. This not only helps make use of economies of scale for transferring the waste to the final destinations, but also reduces turn around times for collection vehicles and increases productivity. However, waste from the sources must not necessarily be brought to the facility. In cases where for example a collection beat is situated closer to a destination location than to an intermediate facility, it might be more cost effective to make a direct transfer to the final destination. This solution approach can be seen in Figure 1.5. Nonetheless, this is the solution approach that was used.

Therefore, the problem this project attempts to solve is:

> *Where and how many intermediate facilities should be located to minimise the collection costs?*

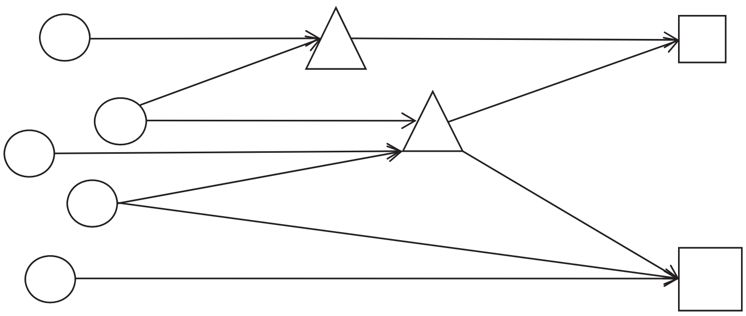


Figure 1.5: Solution approach

1.4 Research Design

In order to answer the research question, three mathematical models were required to gather quantitative information on the problem at hand and generate results in order to make educated recommendations and draw reasonable conclusions.

The first model that was required, is the *multi-facility location model*. It's purpose was to identify potential locations for intermediate facilities. It required the following inputs:

- number of intermediate facilities to be located;
- location of collection beats; and
- waste generated at each collection beat;

and returned as output:

- the location of the intermediate facilities.

After locations were identified, the *travel distance model* was built in order to calculate the distance travelled per collection vehicle for each collection beat. This distance was then divided by the average speed of a waste collection vehicle to get the average travel time per collection vehicle. The model took as inputs:

- the location of collection beats;
- the waste generated at each collection beat;
- the location of intermediate facilities;

- the location of vehicle depots; and
- the vehicle capacity;

and returned as output:

- the total distance travelled for each vehicle per.

Finally, to complete a cost benefit analysis the *waste allocation model* was used. This model is slightly different to the *travel distance model* in that it is concerned with the transportation cost of the entire system — the cost for transferring the waste from a collection beat to either an intermediate facility or landfill (first echelon), and the cost for transporting the compacted waste from the intermediate facility to a landfill (second echelon). The *travel distance model* on the other hand, is only concerned with the travel distance of the vehicle in the first echelon. The difference will become more clear in chapter 5.

The two most important models for the completion of the project were the *multi-facility location model* and the *waste allocation model*. The *travel distance model* was used to compare the travel time for collection vehicles with and without intermediate facilities.

1.5 Research Methodology

In this section a brief overview of the research methodology that was used is given.

From the Problem Statement it is clear that there were many factors that played a role in the success of the project. Furthermore, since the project was concerned with an aspect of MSW management, the project began to move away from a purely operational problem to a strategic and tactical one. To accommodate this the project was tackled in three parts. This allowed for strategic and tactical evaluation for each phase. Thus the three phases were defined as follows:

1. data set generation;
2. identifying potential locations for intermediate facilities; and
3. evaluating the impact of intermediate facilities on waste vehicle operations.

1.5.1 Identifying Potential Locations for Intermediate Facilities

In operations research the identification and selection of any facility is known as a *location modelling problem*. It is believed that the first location problem in literature is thanks to the mathematician Torricelli (1608–1647) (ReVelle and Eiselt, 2005). It is no surprise that location modelling has come a long way since then. Specifically many variants of the generic location modelling problem exist. These will be discussed in more detail in the Chapter 2.

The variant that was used to identify potential locations for the intermediate facilities was the *Multi-Facility Hub Location Problem (HLP)*. The algorithm that was used is the *Weiszfeld* algorithm which will be presented in the Chapters 2 and 4.

1.5.2 Evaluating the Impact of Intermediate Facilities on Waste Vehicle Operations

Once suitable intermediate facility locations were identified, the impact they had on the waste collection operations in terms of travel time were assessed. To do this a simple mathematical model was built in order to calculate the distance each vehicle had to travel to satisfy the demand at each collection beat which was then divided by the average speed of a waste collection vehicle. To make the results comparable to current operations, the model was also applied to current operations. However, the travel time alone was not sufficient to draw any conclusions about possible advantages or disadvantages of locating new intermediate facilities. Therefore the *waste allocation model*, which is a Linear Programming (LP) model, was built in order to conduct a cost benefit analysis.

Together these models were able to produce good results which aided in evaluating the impact of intermediate facilities on waste collection operations.

1.6 Document Structure

The remaining document is structured as follows: In the next chapter existing literature available for location modelling problems will be discussed – first in general and then more specifically to MSW management. Next, the approaches for how the required model input data were generated is discussed. Thereafter, the mathematical models are explained in detail followed by a presentation, discussion and analysis of the results. Finally, we conclude with our recommendations and make suggestions for further research.

Chapter 2

Literature Review

In this section existing literature will be presented and dissected to gain a clear picture of methodologies available to solve the problem at hand. Specifically, literature regarding location modelling for solid waste facilities, Hub Location Problem (HLP) models and impact evaluation models will be discussed chronologically.

2.1 Location Modelling for Solid Waste Facilities

Location modelling is concerned with "sitting facilities in some given space" (ReVelle and Eiselt, 2005). It includes the modelling, formulation and solution to a given problem. There are in fact many variants of the basic formulation, each dealing with a particular objective and unique structure. According to Eiselt and Marianov (2015) the two main approaches that exist when it comes to locating waste facilities such as landfills or intermediate facilities, are "the formulation of mathematical optimisation models, and the use of tools from the multi-criteria decision making toolkit". The mathematical models can be formulated comprehensively with very little involvement from the decision maker. However, solving such problems are often very complex and requires long computational times. On the other hand problems which are formulated as multi-criteria decision making models are much easier to solve, however they require much more involvement from the decision maker and a lot more data available before hand.

Since waste data is not readily available, and especially not for developing countries such as South Africa, the multi-criteria approach is not as applicable to the location modelling process in this case, as the mathematical optimisation model approach.

Many of the early contributions concerning mathematical optimisation models are almost exclusively single-objective mixed integer programming models with a cost minimisation objective function (Eiselt, 2007). However, since the nature of locating solid waste facilities is a strategic problem, it involves many stakeholders, thus multi-objective problems began to emerge. Examples include minimisation of risk and cost, or, minimisation of cost, the quantity that is landfilled, and the environmental impact.

Mitropoulos et al. (2009) present exact and heuristic approaches for the locational planning of an integrated solid waste management system. They solved a mixed integer program with the objective of minimising the total cost of the Municipal Solid Waste (MSW) management system. To be exact they solved a waste flow-allocation problem that simultaneously located the appropriate facilities (i.e. treatment plants, transfer stations and sanitary landfills). The waste generated at the sources can be sent either directly to treatment plants or through transfer stations. Similarly, landfills receive waste collected directly from the source, treatment plants, or compacted through transfer stations.

Eiselt (2007) evaluate the effectiveness of existing waste facilities with those obtained by a system of optimised locations.

Irrespective of the solution approach, the location results of a mathematical optimisation model are by no means the "be-all" solution (Eiselt, 2007). The locations can be used as a starting point, but later other constraints must be considered too.

In majority of the literature involving locating waste facilities, be it landfills or intermediate facilities, the optimal locations are selected from a set of potential locations. For this project however, the objective is to identify potential locations, thus we refer to other location modelling literature as a project basis. In particular the roots for many of the aforementioned problems lie with the so called HLP.

2.2 Hub Location Problems

Hub location problems were first introduced to the location literature by O'Kelly in 1986 (ReVelle and Eiselt, 2005). They play an important role in the transportation industry where hubs are required to "decrease the number of transportation links between origin and destination nodes" (Farahani et al., 2013). For this project, definitions are as follows: the origin nodes are the collection beats where the waste is generated and from where it is collected, hubs represent the intermediate facilities and the destination nodes are the landfills.

Within the realm of hub location problems, there exist three distinct solution domains: network, discrete and continuous (Farahani et al., 2013). The difference are as follows:

- network — the possible locations for the hub are all nodes in the network;
- discrete — the possible locations for the hub are a set of nodes in the network; and
- continuous — the possible locations for the hub are a plane or a sphere.

An example of a network model would be locating an ambulance along the entire road network, while an example of a discrete model would be locating a retail facility which may only be placed in areas which has been zoned for it.

One would be quite right to assume that locating an intermediate facility would also fall under the discrete solution space, however, often "discrete location models have gone through an additional preprocessing phase that has preselected candidate sites at which the facilities may be sited" (Eiselt and Marianov, 2015). This means that additional data of the network must available. If this is not the case, the best solution is to identify generally "good" areas for the intermediate facilities which would later be evaluated further. In particular, since more than one intermediate facility must be located, the variant of the HLP we are concerned with is the continuous Multi-Facility Location Problem (MFLP).

2.2.1 Solution Approaches

A variety of solution approaches exist for solving the continuous MFLP. Most notably there exist exact, heuristic and meta-heuristic algorithms. The reason for this is that as the size of the problem increases, the exact approach is no longer able to solve the problem in reasonable time. Heuristic and meta-heuristic approaches are the needed (Farahani et al., 2013).

One of the most well-known solution methods for solving the MFLP is the *Weiszfeld method* Iyigun and Ben-Israel (2010). In order to solve the continuous MFLP, which is $\mathcal{NP}$-Hard, the problem is relaxed by using probabilistic assignments. This means that a beat is assigned to an intermediate facility based on the probability that it is closer to that particular facility instead of another one. With each iteration the probability is re-calculated as the centre locations are updated too.

In order to answer our research question, the *Weiszfeld method* was used to solve the continuous MFLP by locating k intermediate facilities.

2.3 Evaluating the Impact of Intermediate Facilities

Seemingly no literature exists which is concerned with evaluating the impact of intermediate facilities on waste collection operations. However, one article dealing with a "satellite location analysis for a two-echelon vehicle routing problem" had the potential to lead the research approach in terms of evaluating the impact of intermediate facilities on waste collection operations.

Research done by Crainic et al. (2010) evaluated the impact of several parameters, directly linked to the system layout, on the total transportation cost of a distribution system. The parameters under investigation were the number of customers, the number of location hubs, the customer distribution, and the relationship between customer to hub and hub to destination node costs. Although this is not directly linked to waste collection operations, the work serves as a starting point in terms of which parameters we could also change in our models in order to evaluate the impact that intermediate facilities have on waste collection operations.

2.4 Conclusion

Existing literature was presented in this chapter. In particular work done by Eiselt (2007) and Mitropoulos et al. (2009) was used as the basis to evaluate the costs of the MSW management system with and without intermediate facilities. However, since both approaches required a set of pre-determined waste facility locations, we referred to literature from the MFLP in a continuous solution space in order to first find potential locations for the facilities. Furthermore, since the *Weiszfeld method* is one of the most well-known solution approaches to the continuous MFLP, we used it to solve for potential intermediate facility locations.

Finally, literature was found regarding the impact of several parameters on the total transportation cost of a distribution system. This was used to guide the decision in terms of which parameters to analyse when evaluating the impact that intermediate facilities have on the waste collection operations.

Chapter 3

Phase 1 — Data Generation

In this chapter the approach used to generate the input data for the models will be discussed.

A large problem concerning waste data is that it is generally inaccurate, outdated or incomplete (Hoornweg and Bhada-Tata, 2012). This is even more prominent in developing countries such as South Africa. This lack of information poses challenges on building accurate models, since the output of these models is only as accurate as their input.

The first step therefore, was to generate the required input data. For the City of Cape Town, two primary sources were available to our disposal and lay the foundation for all other data sets generated. These were data gathered by the Optimisation Group from the University of Pretoria and the "opendata" portal [1] from the City of Cape Town. From them we were able to extract the following datasets:

- collection beats;
- drop-off sites (i.e. landfill sites and existing intermediate facility locations); and
- waste vehicle Global Positioning System (GPS) data.

For the City of Cape Town, a collection beat represents the area allocated to a single collection vehicle for the removal of 240L bins at formal properties. These collection beats are readily available on the City of Cape Town's opendata portal and come packaged as so called "shapefiles". A "shapefile" is a collection of geometric shapes which make up complex polygons in order to represent any type of area. The collection beats for the a portion of the city can be seen in Figure 1.4.

The GPS traces from the waste vehicles were provided by the Optimisation Group from the University of Pretoria. Since the data was large in size, we only received the GPS data for a single collection day and assumed that it is representative of the general landscape of waste collection operations throughout Cape Town. With these datasets, we were able to continue with generating the remaining data required by the models.

Both models were built in *Python* version *3.5.2* using the available *numpy* and *SciPy* packages.

3.1 Waste Generation Rates

In order to identify potential location for intermediate facilities, the amount of waste that is generated at a collection beat is required, so that the intermediate facilities could be

[1] Available from http://web1.capetown.gov.za/web1/opendataportal/default

strategically placed in such a way to reduce the distance between collection beats and intermediate facilities for areas generating the most waste.

Waste generation rates are not readily available for the City of Cape Town. The only generation rate available stems from data by Hoornweg and Bhada-Tata (2012) in which it was estimated that South Africa generates approximately *2 kg per capita per day*. In reality this number varies depending on the income level of the area. For our purposes we assumed a constant rate throughout the city.

Still we were unable to say how much waste was generated at a collection beat. However, coupling the data available for the generation rate per capita per day (see above) with population data yielded promising results.

The City of Cape Town's population data for each ward was readily available online from the 2011 census. The City of Cape Town is divided into 115 wards with a total population of 3 740 026 people. A ward is a geopolitical subdivision of municipalities. Some of the wards can be seen in the Figure 3.1.

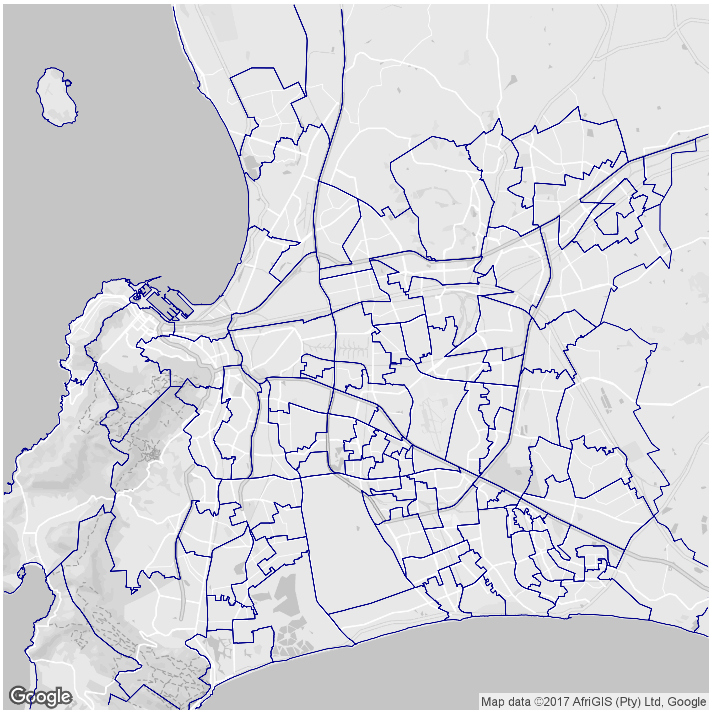


Figure 3.1: Ward outlines for the City of Cape Town

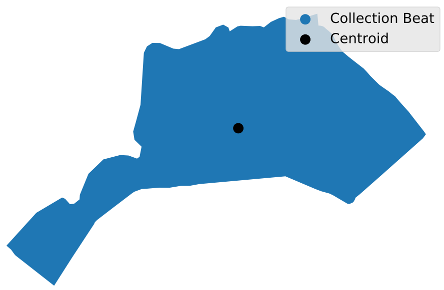
It was left to us to allocate each collection beat to a ward and then divide the total population of the ward by the number of collection beats associated to it to yield the average population per collection beat. This population was then multiplied by the approximated generation rate of *2 kg per capita per day* to estimate the amount of waste generated at each collection beat per day.

However, working with the collection beats and wards as shapefiles proved a difficult task. For example, allocating each collection beat to a ward by looking at whether or not the collection beat is contained within a ward yielded the results in table 3.1.

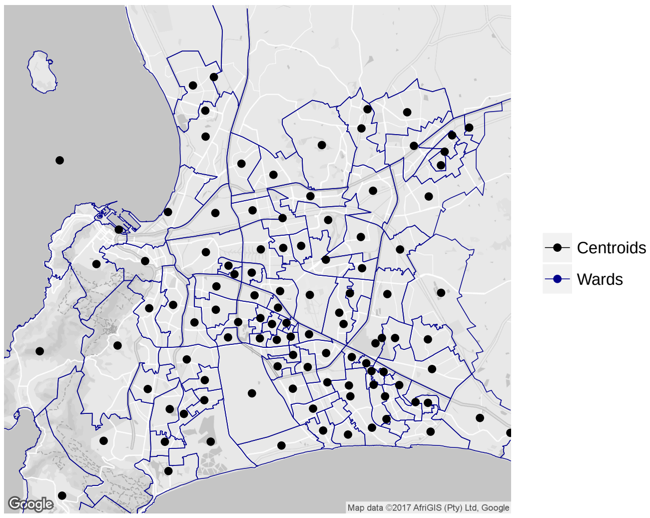
The reason that majority of the collection beats are contained within more than one ward is simply because the shapes are not perfectly aligned, and so it skews the results. For that reason the collection beats and wards were simplified to the centroid of the shape. Figure 3.2a illustrates this for a single ward.

Number of intersections	Count
2.0	278
1.0	186
3.0	162
4.0	69
5.0	20
6.0	6
8.0	1
7.0	1

Table 3.1: Number of intersection vs number of occurrences



(a) Simplified beat



(b) Simplified wards

Figure 3.2: Simplification of wards and collections beats

Thus we had a single GPS coordinate for each collection beat.

The model for estimating these generation rates per collection beat lay with one of the simplest machine learning algorithms called *K-Nearest Neighbour (KNN) algorithm*. The *KNN algorithm* finds the nearest neighbour for a point in data set **A** in data set **B**. In this case, the points in data set **A** represented the collection beats, while the points in data set **B** represented the wards. The results of the algorithm are shown in Figure 3.3.

(a) Unallocated collection beats (b) Allocated collection beats

Figure 3.3: Allocation of collection beats to wards

Once each collection beat was allocated to a ward, the population per ward (which was given by the 2011 census data) was divided over the total number of collection beats associated to it. This gave the average population per collection beat. This number was multiplied by the constant waste generation rate of *2 kg per capita per day*. An example of the results can be seen in Figure 3.2.

3.2 Depot Locations

In order to calculate the travel time using our *travel distance model* one had to know where the waste collection vehicles start and end their trips. As explained in Section 1.2, a waste collection vehicle would typically starts at a depot, collects waste at a collection beat, replenishes its capacity at an intermediate facility (if necessary), and returns to its depot.

These depot locations, however, were not simply found using our available sources, thus the idea was to estimate these locations using the collection vehicles' GPS traces.

Each waste collection vehicle is fitted with a GPS tracking unit which periodically sends out its current location and the position of its ignition key. A typical trace looks like this:

VEH_REG_NO	DATE	POS TIMESTAMP	POS_X	POS_Y	POS IGNITION
CA130563	2011-05-12	03:49:43.0000	18.681067	-33.9351	F

Table 3.3: Typical GPS trace

With this information it was possible to estimate the depot locations with reasonable confidence. The underlying idea was that if a waste collection vehicles' ignition position was off (POS_IGNITION == F) we can assume that the vehicle was at rest, for example at a depot. It could also be the case that the vehicle was for example stopping at the side of the road, or filling petrol etc. Therefore more data points were needed to improve the accuracy of our assumptions.

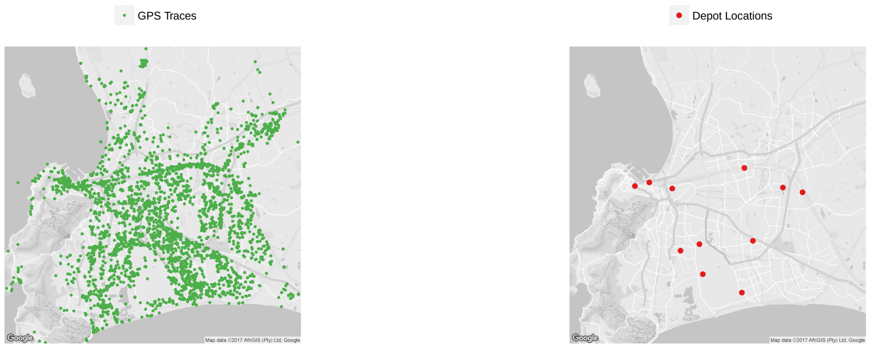
To do this we first had to extract the three most important fields: POS_X, POS_Y and POS_IGNITION from our dataset. In our case we extracted the traces of all vehicles for the waste collection operations on a Thursday over the period from 2011-05-05 to 2014-10-02 with more than 1 million rows. The more data available, the more accurate the

Table 3.2: Some results from allocating collection beats to wards

	XCoord	YCoord	COL DAY	COL DAY CLR	SW SRV AREA	OTSD	Ward	Waste
0	18.51285908	-33.82542082	Thursday	Purple	Atlantic	Departmental	106	4.124
1	18.50377584	-33.82303478	Thursday	Blue	Atlantic	Departmental	106	4.124
3	18.51384002	-33.83536041	Thursday	Blue	Atlantic	Departmental	83	6.82
5	18.49861219	-33.88850952	Tuesday	Orange	Atlantic	Departmental	101	8.884
6	18.50541732	-33.88444982	Tuesday	Orange	Atlantic	Departmental	52	9.656
13	18.53358006	-33.81500316	Trade Thursday	Blue	Atlantic	Departmental	58	6.722
15	18.47450195	-33.96197047	Wednesday	Purple	Atlantic	Departmental	65	4.77

algorithm output becomes. Next we filtered the data to include only entries for which POS_IGNITION == F. This left us with a total of 22 590 entries. The final step before feeding the data to the algorithm was to map the data on Google Earth in order to estimate the input parameters for the algorithm. Specifically we needed to estimate the sample density (minimum number of points per cluster), and the cluster distance. We estimated that a cluster distance of **100m** and a sample density of **300** points would be adequate.

The algorithm we used to identify the depot location was the *DBSCAN — density based clustering algorithm*. Its purpose is to discover clusters and noise from a spatial dataset (Ester et al., 1996). This was very relevant in our case as we were trying to identify the depot locations from a spatial dataset which included noise (e.g. the vehicle stopping at a non-depot location). The results of the *DBSCAN* algorithm are shown in Figure 3.4.



(a) Before DBSCAN implementation

(b) After DBSCAN implementation

Figure 3.4: Results of DBSCAN algorithm

The DBSCAN algorithm starts with an arbitrary point P in the dataset and returns all other points that fall within the circle of radius ϵ with centre P. If the total number of points including point P is greater than the number of points specified by min_points, then a cluster has been found. The algorithm then further loops through every point P' in the cluster, returns all other points that fall within the circle of ϵ with centre P'. If the total number of points is more than min_points then the cluster is updated to include the new cluster too. In case the number of points in the circle are less than min_points we ignore the point and continue with the next one. This explanation is represented in pseudocode in algorithm 1.

3.3 Conclusion

In this chapter we discussed how the datasets required for the project were acquired. The GPS traces for the waste collection vehicles were sourced from the Optimisation Group from the University of Pretoria, while the collection beats and drop-off sites (i.e. landfill sites and existing intermediate facility locations) were obtained from the opendata portal from the City of Cape Town. However, this information was not sufficient, we still required the amount of waste generated per collection beat in order to strategically locate the intermediate facilities. Hoornweg and Bhada-Tata (2012) estimated that South Africa generates approximately *2 kg per capita per day*. We took this insight and linked it with the population data per ward in order to estimate how much waste is generated per

Algorithm 1: Pseudocode for the DBSCAN clustering algorithm (Agrawal, 2013)

```
DBSCAN Algorithm
DBSCAN(dataset, epsilon, min_points):
    C = 0
    for each unvisited point P in dataset do
        mark P as visited
        sphere_points = regionQuery(P, epsilon)
        if sizeof(sphere_points) < min_points then
            ignore P
        else
            C = next cluster
            expandCluster(P, sphere_points, C, epsilon, min_points)

expandCluster(P, sphere_points, C, epsilon, min_points):
    add P to cluster C
    for each point P‘ in sphere_points do

    if P‘ is not visited then
        mark P‘ as visited
        sphere_points‘ = regionQuery(P‘, epsilon)
        if sizeof(sphere_points‘ ≤ min_points) then
            add sphere_points‘ to sphere_points
        if P‘ is not a member of any cluster then
            add P‘ to cluster C

regionQuery(P, epsilon):
    return all points within radius of size epsilon centred at P (including P)
```

ward. We then took this further and looked at which collection beats are contained within which ward and simply split the waste generated per ward evenly amongst the associated collection beats in order to estimate the average waste generated per collection beat.

Next we also generated the locations of the waste vehicle depots from the vehciles' GPS traces. Finally, we were left with following datasets:

- location of collection beats;
- waste generated per collection beat;
- location of drop-off sites (i.e. landfill sites and existing intermediate facility locations); and
- location of depots.

Next we present the models which made use of these datasets.

Chapter 4

Phase 2 — Multi-Facility Location Model

Once the datasets were generated, it was possible to find potential locations for intermediate facilities. In this chapter we present the Multi-Facility Location Problem (MFLP) model that was used to generate the intermediate facility locations. We also present our variants thereof in order to take into account the location of landfill sites as well as existing intermediate facilities. First, the *generalised Weiszfeld method for the multi-facility location problem* is presented, followed by an adaption to consider the locations of landfills and then considering the location of existing landfills. Finally, we conclude with choosing the best adaption for our problem.

Since the *Weiszfeld method* is a well-known solution approach to the MFLP, the *generalised Weiszfeld method for the multi-facility location problem* of Iyigun and Ben-Israel (2010) was used to find candidate locations for the intermediate facilities.

The model was coded in *Python* version *3.5.2*.

4.1 Model Formulation

Let $D = \{x_i : i \in \overline{1,N}\}$ be a set of N **data points** (customers) with given **weights** (demands) $\{w_i > 0 : i \in \overline{1,N}\}$ and $C = \{\boldsymbol{c}_k : k \in \overline{1,K}\}$ be a set of K many centre locations (intermediate facilities). Then given an integer $1 \leq K < N$, the MFLP is to locate K facilities and assign each customer to a facility by minimising the sum of weighted distances:

$$\min \sum_{k=1}^{K} \sum_{x_i \in \varrho_k} w_i d(\boldsymbol{x}_i, \boldsymbol{c}_k) \tag{4.1}$$

where $\boldsymbol{c}_k$ is the location of intermediate facilities and ϱ_k is the cluster of customers assigned to the k-th facility. In our case the centres represent the intermediate facilities and the customers represent the collection beats.

The *generalised Weisfeld method for multiple facilities* is given in Algorithm 2. The algorithm takes an iterative approach at finding optimal locations for K facilities. It alternates between assigning probabilities to the customers, that is the probability that a customer is associated to centre k, and updating the centres with respect to these probabilities and given weights per customer.

Algorithm 2: A generalised Weiszfeld algorithm for multiple facilities

Data: $D = \{x_i : i \in \overline{1,N}\}$ data points (location of customers),
$\{w_i : i \in \overline{1,N}\}$ weights,
K the number of facilities,
$\epsilon >$ stopping criterion

Initialisation: K arbitrary centres $\{\boldsymbol{c}_k : k \in \overline{1,K}\}$

Iteration:

Step 1: compute distances $\{d(\boldsymbol{x}, \boldsymbol{c}_k) : k \in \overline{1,K}\}$ for all $\boldsymbol{x} \in D$ using (4.2)

Step 2: compute probabilities $\{p_k(\boldsymbol{x}) : \boldsymbol{x} \in D,\ k \in \overline{1,K}\}$ using (4.3)

Step 3: update the centres $\{\boldsymbol{c}_k^+ = T_k(\boldsymbol{c}_k) : k \in \overline{1,K}\}$ using (4.4)

Step 4: if $\sum_{k=1}^{K} d(\boldsymbol{c}_k^+, \boldsymbol{c}_k) < \epsilon$ **stop**

return to step 1

The approach of the algorithm can be seen in Figure 4.1. There the inner working of the algorithm become visually clearer. K arbitrary centres are initiated (Initial Centres). The algorithm then computes the distances between centres and customers, and computes the probabilities that customer i is associated to centre k. Based on these values the previous centres are updated and the next iteration begins until a stopping criterion is reached (Final Centres)

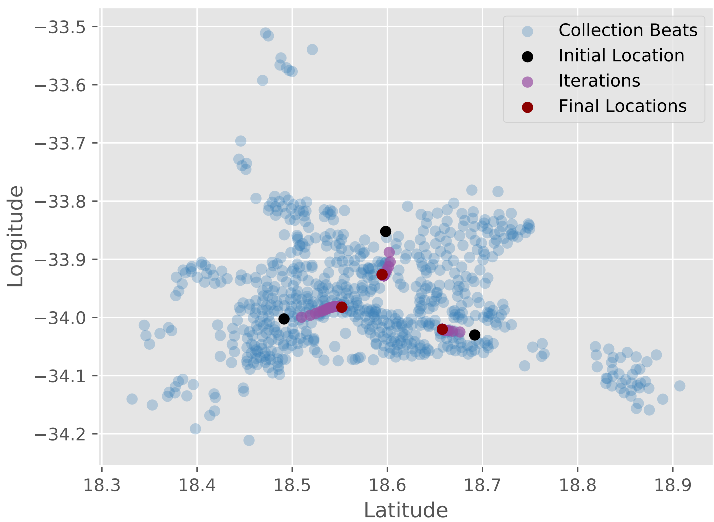


Figure 4.1: Weiszfeld iterations

Given two vectors $\boldsymbol{p} = (p_1, p_2)$ and $\boldsymbol{q} = (q_1, q_2)$, then their *euclidean* distance is calculated as:

$$d(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2} \tag{4.2}$$

Given the centres and distances $d(\boldsymbol{x}_i, \boldsymbol{c}_k)$, the probability that data point $\boldsymbol{x}_i$ is associated to centre $\boldsymbol{c}_k$ is given by:

$$p_k(\boldsymbol{x}_i) = \frac{\prod_{j \neq k} d(\boldsymbol{x}_i, \boldsymbol{c}_j)}{\sum_{l=1}^{K} \prod_{m \neq l} d(\boldsymbol{x}_i, \boldsymbol{c}_m)}, \quad k \in \overline{1, K} \tag{4.3}$$

To update the centres for $\boldsymbol{c}$ not in the set D (customer locations) the equation is given by:

$$T_k(\boldsymbol{c}) = \sum_{i=1}^{N} \left(\frac{\frac{p_k^2(\boldsymbol{x}_i) w_i}{\|\boldsymbol{x}_i - \boldsymbol{c}\|}}{\sum_{j=1}^{N} \frac{p_k^2(\boldsymbol{x}_j) w_j}{\|\boldsymbol{x}_j - \boldsymbol{c}\|}} \right) \boldsymbol{x}_i \tag{4.4}$$

It can be seen that Equation (4.4) is undefined if a centre $\boldsymbol{c}_k$ coincides with a data point. Iyigun and Ben-Israel (2010) propose an extension to deal with this problem. If a centre $\boldsymbol{c}_k$ does not coincide with one of the data points, then

$$\boldsymbol{R}_k(\boldsymbol{c}_k) = \sum_{i=1}^{N} \frac{p_k^2(\boldsymbol{x}_i) w_i}{\|\boldsymbol{x}_i - \boldsymbol{c}_k\|} (\boldsymbol{x}_i - \boldsymbol{c}_k) \tag{4.5}$$

Otherwise, if a centre $\boldsymbol{c}_k$ coincides with a data point $\boldsymbol{x}_j$, then $\boldsymbol{x}_j$ belongs with certainty to the kth cluster, and so

$$p(\boldsymbol{x}_j) = 1, \qquad p_m(\boldsymbol{x}_j) = 0, \quad \text{for all } m \neq k. \tag{4.6}$$

In this case, we define

$$\boldsymbol{R}_k(\boldsymbol{x}_j) = max\{\|\boldsymbol{R}_k^j\| - w_j, 0\} \frac{\boldsymbol{R}_k^j}{\|\boldsymbol{R}_k^j\|} \tag{4.7}$$

where,

$$\boldsymbol{R}_k^j = \sum_{i \neq j} \frac{p_k^2(\boldsymbol{x}_i) w_i}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|} (\boldsymbol{x}_i - \boldsymbol{x}_j) \tag{4.8}$$

Finally, we replace the updated centres equation (4.4) with (4.9):

$$T_k(\boldsymbol{c}) = \boldsymbol{c} + \boldsymbol{h}_k(\boldsymbol{c}) \boldsymbol{R}_k(\boldsymbol{c}) \tag{4.9}$$

$$\text{with} \quad \boldsymbol{h}_k(\boldsymbol{c}) = \frac{1}{\sum_{j=1}^{N} \frac{p_k^2(\boldsymbol{x}_j) w_j}{\|\boldsymbol{x}_j - \boldsymbol{c}\|}} \tag{4.10}$$

Essentially, what this extension does is, for a data point coinciding with a centre, it ignores that data point for the calculation of updating the centre. The next section discusses how the algorithm was adapted to locate intermediate facilities with respect to landfills too.

4.2 Locating With Respect to Landfills

From Section 1.2 it should be clear that a Municipal Solid Waste (MSW) management system is not only concerned with collecting waste from collection beats and bringing it to intermediate facilities, it us also concerned with hauling the the compacted waste from the intermediate facilities to the nearest landfill.

The *Weiszfeld* algorithm, however, only minimises the sum of weighted distances between the centre locations (intermediate facilities) and customer locations but ignores the location of the landfills, where the waste ultimately ends up at. The landfills should therefore be included in the calculation for locating the intermediate facilities as this would firstly consider the system as a whole, and secondly minimise the sum of weighted distances between the customer locations, centre locations and landfill locations.

A possible solution was to add the landfill locations to the dataset and assign them a dummy weight. It turned out, however, that the *Weiszfeld* algorithm was relatively insensitive towards the landfills at lower dummy weights. However, as soon as the dummy weight became too large relative to the actual weights of the customer locations, the centre locations would "snap" on top of the landfills. This can be seen in Figure 4.2.

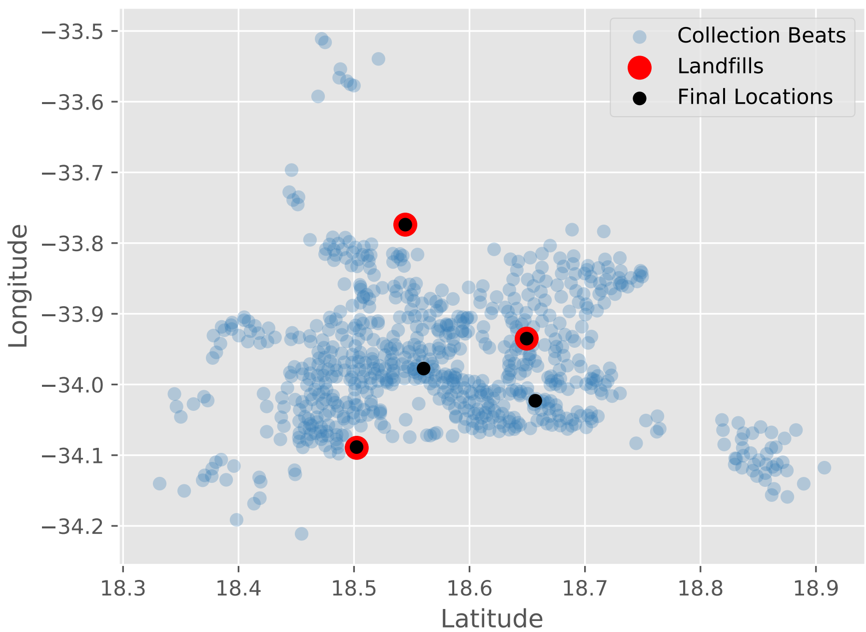


Figure 4.2: Locating With Respect to Landfills

This behaviour was to be expected. The reason is that the objective of the *Weiszfeld* algorithm is to locate K facilities and assign each customer to a facility by minimising the sum of weighted distances. As soon as the weight of a location becomes too large relative to the remaining weights of the data set the best solution becomes to locate the centre as close to the customer as possible, hence on top of it.

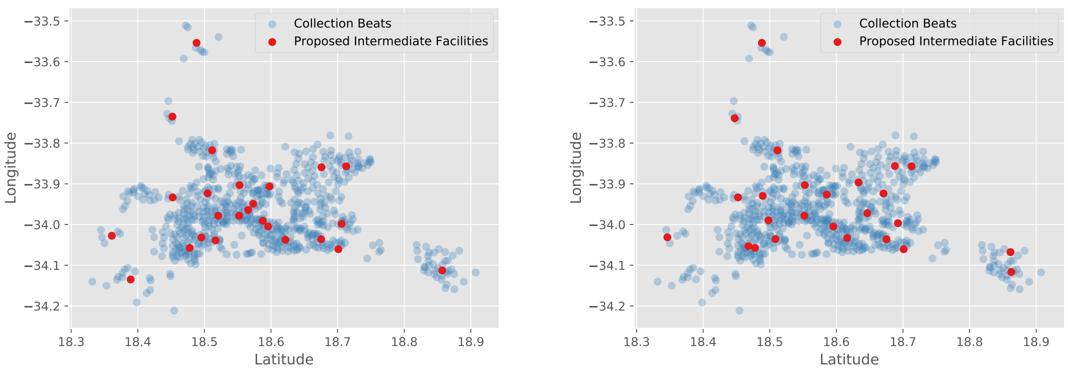
It is impractical to locate an intermediate facility next to a landfill let alone on top of it. Locating an intermediate facility too close to a landfill makes the benefit of an intermediate facility futile. It is therefore safe to say that this adaption of the *generalised Weiszfeld method for the multi-facility location problem* was not used further.

4.3 Locating With Respect to Existing Intermediate Facilities

Since the City of Cape Town already has two intermediate facilities, it may be more beneficial to find potential locations for new intermediate facilities with respect to existing facilities i.e. cover the areas that are missed by the current facility locations.

To achieve this, an additional step was added to the algorithm. The algorithm still calculated the distance between customer locations and centre locations as in step 1, but the distance between customer locations and existing intermediate facilities was also calculated. If the distance between the customer and the existing intermediate facilities was less than that of the customer and the new centre, the customer was removed from the dataset for that iteration. For each iteration the same process was executed.

The results of the normal locations versus the locations with this additional experimental step can be seen in Figure 4.3. From an initial inspection it can be seen that there are no prominent differences between the two data sets. The only visual difference is that the normal implementation seemingly located the facilities more even evenly among the collection beats. This can be seen in the South-Easterly corner of the data set. Where the normal implementation located only one facility, the adaption located two facilities in close proximity. Another example can be seen at the point (18.70, -33.85). In the adapted implementation the two facilities are located closer together than for the normal implementation. A last example can be seen in the South-Westerly corner. The one cluster of collection beats has access to an intermediate facility in the normal implementation, however has no access for the adapted implementation.



(a) Normal locations

(b) Experimental locations with existing Intermediate Facilities included

Figure 4.3: Locating With Respect to Existing Intermediate Facilities

More detailed differences were not studied as they did not help answer the research question. The question to be answered was:

> *Where and how many intermediate facilities should be located to minimise the collection costs?*

The aim was to locate **new** intermediate facilities and not to locate **more** facilities.

4.4 Conclusion

In this chapter we discussed the model that was used in Phase 2 of the research project's solution approach. More specifically we discussed the model that was used to solve the MFLP which was needed to identify potential locations for intermediate facilities. The MFLP was solved by applying the *generalised Weiszfeld method for multiple facilities*. We also discussed two experimental variants of the *generalised Weiszfeld method* which were believed potentially improve the location allocation process. It turned out, however, that locating the intermediate facilities with respect to landfill locations by assigning dummy weights to the landfill locations, located the facility on top of the landfill locations. This was impractical and therefore the adaption was not used.

The second variant was to locate the intermediate facilities with respect to the existing facilities. Also this approach was disregarded as the aim of the project was not to locate **more** facilities but rather to locate **new** facilities and see how they affect the collection operations. We do however, compare the location of the existing facilities with the location of new intermediate facilities in the Results and Discussion Chapter to evaluate how effective the location of the existing facilities is.

For the remainder of the project the standard implementation of the *generalised Weiszfeld method for multiple facilities* was used to generate the potential locations for intermediate facilities.

Chapter 5

Phase 3 — Impact Evaluation Models

Once the potential locations for intermediate facilities were identified, models were built to gain quantitative insight regarding the impact of intermediate facilities on waste collection operations. First the *travel distance mode* was built to help better understand the impact of the facilities in terms of travel distance and ultimately travel time. As such, the model is extremely simple and will be explained in the next section.

However, the travel distance/time alone was not enough to draw any real conclusions as it does not look at the Municipal Solid Waste (MSW) collection operations as a whole. It only calculates the travel distance for the waste vehicles per collection beats, for dropping the waste at an intermediate facility. It does not include the time for the vehicles transferring the compacted waste from the intermediate facility to the nearest landfill. Furthermore, it also does not allow for direct routes for collection beats to a landfill which might be closer than an intermediate facility. This is not a true reflection of reality.

Therefore, another model was built which was able to help the decision making process by taking more factors into account than the *travel distance model*. Since the *waste allocation model* is a Linear Programming (LP) model with an objective function for minimising cost, it was the perfect candidate to do a cost benefit analysis with.

The results of the models will be discussed in the next chapter.

5.1 Travel Distance Model

The starting point for evaluating the impact of intermediate facilities on waste collection operations was to look at the travel distance of the waste vehicles with and without intermediate facilities. In order to compare results from the Multi-Facility Location Problem (MFLP) model with current operations, the current situation was simplified by running it through the same model.

At a high level, the *travel distance model* looks at the total distance a waste collection vehicle would need to travel in order to satisfy the demand for its associated collection beat. The trip includes the distance from the starting depot to its collection beat, then any number of trips from the collection beat to the intermediate facility in order to satisfy the beat's demand, and finally a trip from the intermediate facility to the vehicle's starting depot.

The model was built on the following assumption:

1. all collection beats, intermediate facilities, depots and landfills are represented by a

single Global Positioning System (GPS) point;

2. no time restrictions are imposed on the vehicles — this means that the vehicle can make as many trips as necessary to fulfil the demand of its collection beat;

3. the distance between points is the haversine distance, and not the actual driving distance;

4. each collection beat is serviced by one and only one collection vehicle; and

5. no direct routes to a landfill are allowed — all waste must be collected and brought to an intermediate facility.

The model was coded in *Python* version *3.5.2* using the available *numpy* package in order to take advantage of *numpy*'s broadcasting capabilities. This dramatically increases computational speeds in comparison to the same function written in *pure Python*. The difference in computational speed between the haversine formula (5.1) written in *pure Python* and a vectorised function can be seen in table 5.1. A distance matrix was generated between 723 collection beats and 25 intermediate facilities using the *pure Python* and vectorised function of the formula.

	Number of function calls	Execution time
Pure Python	1 046 185	29.310 seconds
Vectorisation	1 451	0.129 seconds

Table 5.1: Execution time for calculating a haversine distance matrix using the pure python and vectorised implementation of the haversine formula

The computational speed increase is especially important since the haversine formula is called multiple times throughout the model. In fact the formula was also used in the linear program for the waste allocation model.

5.1.1 Model Formulation

Given two geographical coordinates with their latitude and longitude (in radians), the haversine distance between these points is given by:

$$d = 2r \arcsin \sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos\phi_1 \cdot \cos\phi_2 \cdot \sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)} \tag{5.1}$$

where,

- d: haversine distance between points;
- r: radius of the earth;
- ϕ_1, ϕ_2: latitude of point 1 and point 2, in radians; and
- λ_1, λ_2: longitude of points 1 and point 2, in radians.

The model is presented in pseudocode in Algorithm 3. Since each collection beat is serviced by a single collection vehicle, each collection beat only has one travel distance associated with it. We begin by calculating the haversine distance between a collection beat and each depot, in order to determine the closest depot for a collection beat and set it as the starting location for the vehicle. The distance as well as the associated depot is stored in a list. This is repeated for each collection beat and intermediate facility in order to determine the closest facility used to replenish the vehicle's capacity. Again, the distance and associated facility is stored. Next, the haversine distance between the associated closest depot and associated closest intermediate facility is calculated — this is the last trip the vehicle makes from the intermediate facility to end at the depot.

The total number of trips a vehicle must make between a service area and an intermediate facility was calculated using the following formula: $2\lceil \frac{w}{\alpha} \rceil - 1$. The demand for the collection beat is divided by the vehicle's capacity. The ceiling of the value is taken since no fraction trips can be made. This number is multiplied by two to include the return trips. However, since on the last trip, the vehicle does not return to the collection beat but rather returns to the depot, we reduce the trips by one. This process is repeated for all service areas.

Algorithm 3: Travel distances

Input : Set D of customers
Set W of demands for each customer
Set F of depot locations
Set C of intermediate facilities
The vehicle capacity as α

Output: Total travel distance per customer

```
h_1 ← ∞
h_2 ← ∞
for d ∈ D do
    for f ∈ F do
        h ← haversine(d, f)
        if h < h_1 then
            h_1 ← h
            f_closest ← f
    for c ∈ C do
        h ← haversine(d, c)
        if h < h_2 then
            h_2 ← h
            c_closest ← c
    h_3 ← haversine(f_closest, c_closest)
    trips ← 2⌈w/α⌉ − 1
    total distance ← h_1 + trips · h_2 + h_3

return total distance
```

5.2 Waste Allocation Model

It turned out that simply using the travel distance per vehicle was not enough to make an informed recommendation, as the previous model did not take into account the additional travelling that is required in order to bring the collected waste from intermediate facilities to landfills. Furthermore, it was also an unreasonable assumption that no direct routes for a vehicle to a landfill site was allowed. It is reasonable to assume that for some vehicles, if their collection beat is located directly next to a landfill, or if the distance to a landfill is less than that of the nearest intermediate facility, it might be more efficient to drop the waste off directly at the landfill.

Therefore, a LP model was built in order to tackle these shortfalls and to gain a more holistic view on the impact of intermediate facilities. The model was based on the work done by Mitropoulos et al. (2009) and Eiselt (2007) since both works were concerned with the collection operations as an integrated system, thus they looked at transporting the waste from collection beats to intermediate facilities and then further to landfills. Furthermore, they also allowed for direct routes to landfills. Finally, since both were LP models, their objective function was concerned with minimising total system costs. In our case we only looked at minimising the total system transportation cost.

The model was coded in *Python* version *3.5.2* using the optimisation package *PuLP* — a linear solver library.

5.2.1 Model Formulation

Similar to the model presented by Eiselt (2007), this model also consists of three layers; the collection beats (or customers) are the first layer, the intermediate facilities on the second layer, and the landfills on the third layer. Assuming that p intermediate facilities are to be chosen, then a binary variable o_j is one when intermediate facility j is to be opened and zero otherwise. Variable x_{ij} indicates the amount of waste transferred from collection beat i to intermediate facility j and variable y_{ik} indicates the amount of waste transferred from collection beat i to landfill k.

Since the travelling cost for waste vehicles is difficult to find, we avoid an exact transportation cost for the vehicles transporting the waste from intermediate facility but instead introduce a discount factor α. The justification for the discount factor is that unlike the collection vehicles that transport waste from the customer to intermediate facility or landfill, the vehicles which collect the waste from intermediate facilities have a much larger capacity. The model also assumes that the intermediate facilities are capacitated while the landfills are un-capacitated. Furthermore let M be an arbitrarily large number.

It is now possible to formulate the model.

Index sets

$$
\begin{aligned}
I &\triangleq \text{set of collection beat locations, where } j \in I \\
J &\triangleq \text{set of intermediate facility locations, where } j \in J \\
K &\triangleq \text{set of landfill locations, where } k \in K
\end{aligned}
$$

The decision variables are

$$
\begin{aligned}
x_{ij} &\triangleq \text{amount of waste carried from collection beat } i \text{ to intermediate facility } j \\
y_{ik} &\triangleq \text{amount of waste carried from collection beat } i \text{ to landfill } k \\
o_j &\triangleq \begin{cases} 1, & \text{if intermediate facility } j \text{ is opened} \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

Other model parameters are

k	$\triangleq$	number of intermediate facilities to be opened
cap	$\triangleq$	given maximum capacity per intermediate facility
c	$\triangleq$	given unit transportation cost for vehicles travelling in the first echelon
α	$\triangleq$	given discount factor for vehicles travelling in the second echelon
p_j	$\triangleq$	given waste flow at intermediate facility j
d_{ij}	$\triangleq$	given haversine distance between collection beat i and intermediate facility j
d_{ik}	$\triangleq$	given haversine distance between collection beat i and landfill k
d_j	$\triangleq$	given haversine distance between intermediate facility j and its nearest landfill
w_i	$\triangleq$	given waste generated at collection beat i

The objective is to minimise the total transportation cost given by:

$$\min z = \sum_{i \in I} \sum_{j \in J} x_{ij} d_{ij} c + \sum_{i \in I} \sum_{k \in K} y_{ik} d_{ik} c + \sum_{j \in J} d_j p_j c(1 - \alpha) \tag{5.2}$$

and subject to

$$\sum_{j \in J} x_{ij} + \sum_{k \in K} x_{ik} = w_i \qquad \forall i \in I \tag{5.3}$$

$$\sum_{i \in I} x_{ij} \leq M o_j \qquad \forall j \in J \tag{5.4}$$

$$\sum_{j \in J} o_j = k \qquad \forall j \in J \tag{5.5}$$

$$\sum_{i \in I} x_{ij} \leq p_j \qquad \forall j \in J \tag{5.6}$$

$$p_j \leq cap \qquad \forall j \in J \tag{5.7}$$

The object function (5.2) minimises the total system transportation cost. The first term is the total cost for transporting waste from collection beats to intermediate facilities. The second terms is the total cost for transporting waste from collection beats to landfills and the last term is the total discounted cost of transporting waste from intermediate facilities to landfills.

The first constraint, (5.3), ensures that the total flow from the collection beat is equal to the waste generated by the beat. Constraints (5.4) and (5.5) ensure that when waste is transferred to a facility it is opened and that the total number of facilities opened is equal to the specified number of facilities to be opened. Finally constraints (5.6) and (5.7) ensure that the total waste flow to the intermediate facility does not exceed its capacity.

5.3 Conclusion

The models that were used to solve the problem at hand were presented in this chapter. First the *travel distance model* was discussed. Its purpose was to calculate the total distance a waste vehicle would traverse in order for it to collect all the waste generated at its respective collection beat. However, the model had unrealistic simplifications as it did not take into consideration the entire collection operation, which included the waste to be transferred from intermediate facility to landfill site. Furthermore, the constraint imposed on the model were that the vehicles had to unload the waste at an intermediate facility and was not allowed to make direct trips to landfill sites. A more realistic model was developed in order to address these short-commings and to yield more accurate results in terms of the overall waste collection operations with respect to intermediate facilities. The model is a LP model with the objective of minimising the system's transport costs. In the next section we present the results obtained from the models for various parameters and scenarios.

Chapter 6

Results and Discussion

Once the models were built it was possible to run the models through various scenarios and evaluate their outcome. Since the City of Cape Town already has two intermediate facilities which are in use, another interesting question arose: *"could these intermediate facilities have been located more efficiently?"*. Remaining with the original problem statement of identifying potential locations for intermediate facilities and evaluating their impact on waste collection operations, the following questions were addressed:

- how does the system transportation costs change with respect to changing parameters;
- how does the waste allocation change with respect to changing parameters;
- assuming that no intermediate facilities exist for the City of Cape Town, how many intermediate facilities should be constructed and where should the City locate these facilities?

The remainder of the chapter is structured as follows. First the potential locations for the intermediate facilities are discussed, then the impact that the intermediate facilities have on waste vehicle travel time is analysed, followed by a cost benefit analysis and a summary of the results.

6.1 Intermediate Facility Locations

In Chapter 4, the various models and adaptions were discussed. It was decided that the best model to use was the standard implementation of the *generalised Weiszfeld method for the multi-facility location problem*.

The biggest question for this part of the project was *"how many facilities should be located?"*. As a test-case it was decided that up to a total of 25 facilities would be located. The reason was that locating more than 25 locations required more than 2 000 iterations per run which was relatively time consuming. Additionally with 25 intermediate facilities it was believed that the City of Cape Town would be covered sufficiently.

Using the *generalised Weisfeld method for multiple facilities*, Algorithm 2, 25 potential locations for the intermediate facilities were identified, shown in Figure 6.1. Their exact locations can be found in appendix A.

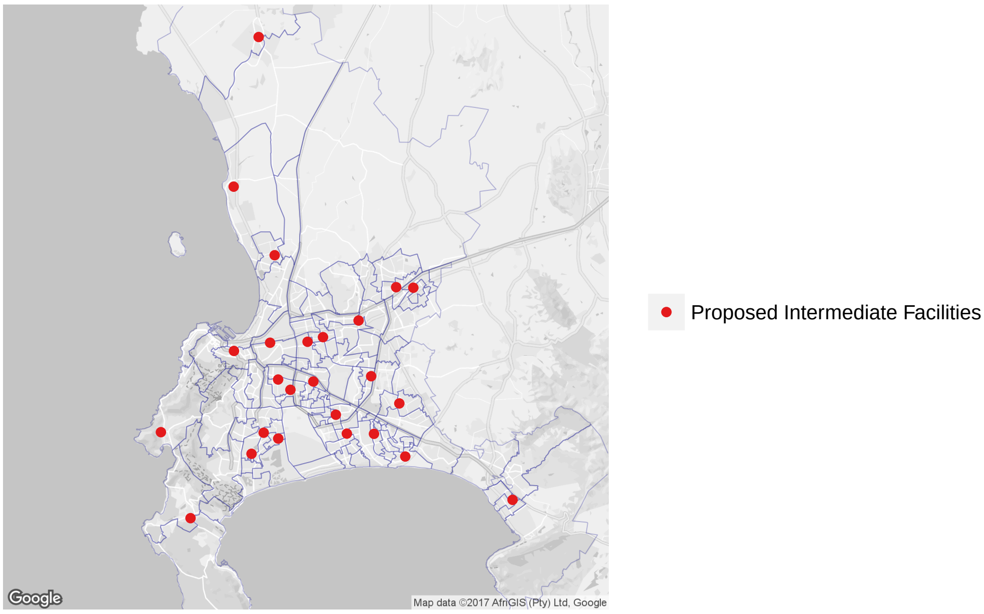


Figure 6.1: Potential locations for intermediate facilities

6.2 Travel Time

Before the travel time could be calculated for the vehicles, the travel distance had to be calculated. The data sets for the model were discussed in Section 1.4, the only additional parameter which was not mentioned is the vehicle capacity. It was assumed that the first echelon waste vehicles have an approximate capacity of **10 tons**.

Once the travel distance was calculated for each collection beat, it was possible to estimate a travel time for each collection vehicle per collection beat. The distance was divided by an average vehicle speed of **30 km/h** to yield the estimated time a waste collection vehicle spent travelling from the depot to a collection beat, between a collection beat and intermediate facility and finally back to the depot. The travel time did not include the time of collecting waste.

In order to see how the travel time changes with an increasing number of intermediate facilities, the travel time was calculated for the range of locating 1 to 25 intermediate facilities. Therefore, the *Weiszfeld method* was used to solve for one facility, for which the travel time was calculated, then for two facilities, then for three facilities and so forth.

The results for the average travel time for waste collection vehicles with various number of intermediate facilities located can be seen in Figure 6.2. As the number of intermediate facilities increase, the travel time decreases. This is because as more intermediate facilities are located around the city (see Figure 6.1) they cover a larger area, and so the facility becomes closer for more collection beats. This means that a waste collection vehicle no longer has to drive to a landfill further away, but can instead drop-off the waste at a nearby intermediate facility.

The figure also shows a dotted line at around 1.68 h. That was the case for when the three existing landfills were treated as landfills. In other words, no intermediate facilities were present. It can be seen that the dotted line intersects the graph when three intermediate facilities were located. This means that in terms of travel time, so long as waste collection vehicles are forced to visit an intermediate facility, not locating any

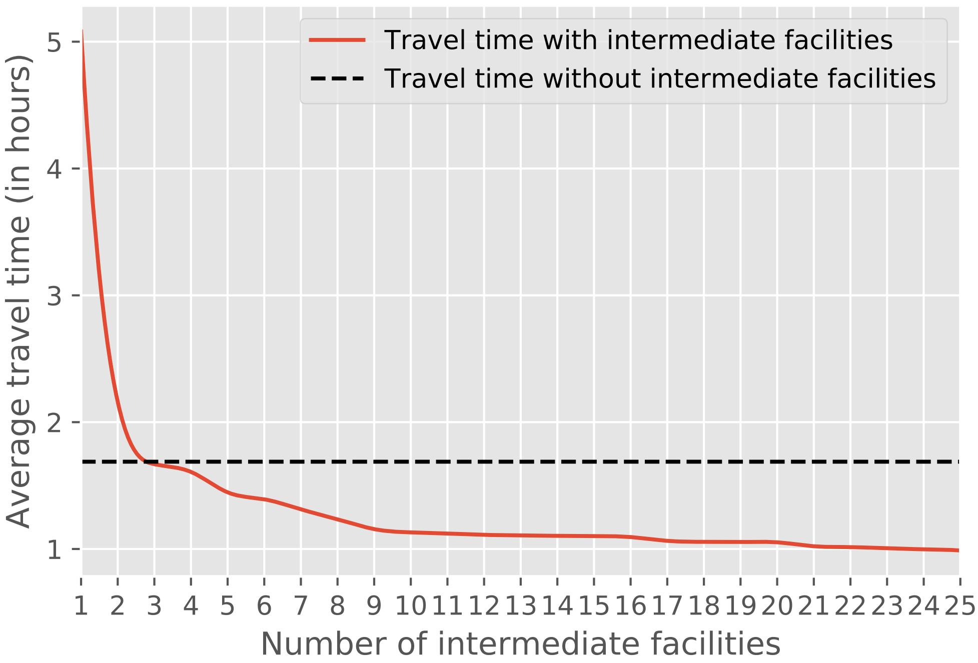


Figure 6.2: Travel time vs number of intermediate facilities

intermediate facilities when three landfills are available is the better option. However, as soon as more intermediate facilities than landfills are located, the model figure shows that the average travel time is less for the waste collection vehicles.

As mentioned previously, this was not a fair comparison, as the model did not take into the account the additional time that must be spent transferring the waste from intermediate facilities to the nearest landfill.

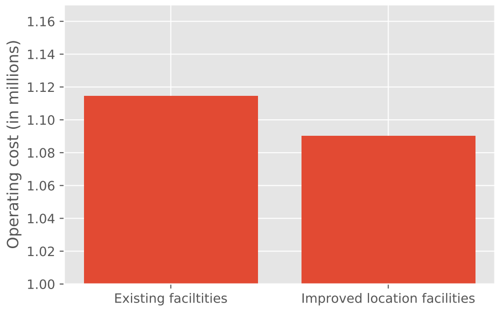
6.3 Cost Benefit Analysis

The results from the previous section were actually expected, especially since only the first echelon operations were considered. Without considering the second echelon operations, the travel time will decrease as more facilities are available because a larger area is serviced by the facility which is closer then the previous it was previously services by.

In this section a wholistic approach is applied by using the *waste allocation model*. The model was run for various input parameters to gain a better understanding of how intermediate facilities impact the waste collection operation system.

According to a feasibility report for Eskom regrading the "Municipal Solid Waste Diversion and Beneficiation Opportunities at Nelson Mandela Bay Metro Municipality", the unit travel cost of a waste collection vehicle is **R 13.73 per km per ton**. It was also assumed that the maximum capacity of an intermediate facility is **1000 tons** and that the discount factor for vehicles travelling the second echelon is 1/4 or **25%** of the first echelon costs.

The first question to be answered was, since the City of Cape Town already has two intermediate facilities, *"could these intermediate facilities have been located more efficiently?"*. To answer this question, the locations of the existing intermediate facilities and the potential locations according to the *Weiszfeld method* were each given as input for the *waste allocation model* and the results compared. Figure 6.3 shows the results. In the bar chart in Figure 6.3a it can be seen that had the intermediate facilities been located using the *Weiszfeld method* the transportation costs could have been reduced. The locations of the existing and new intermediate facilities are shown on the map in figure 6.3b.



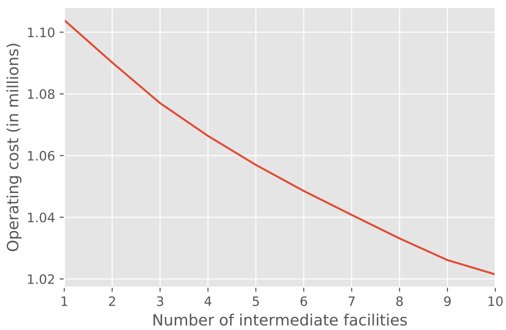
(a) System transportation cost

(b) Facility locations

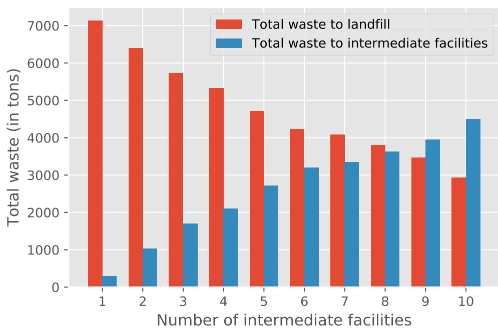
Figure 6.3: Existing vs new intermediate facilities

All this showed is that the locations of intermediate facilities with the *Weiszfeld method* could potentially have been more cost-effective than the location of the current facilities.

However, *how did these cost savings change as the number of intermediate facilities located increased?* Figure 6.4a shows that as the number intermediate facilities increased, the transportation costs decreased. This is expected for the same reason as explained in the previous section. Furthermore, since a discount factor is involved, it is beneficial to collect as much waste at the intermediate facilities and then transport it to the landfills at a later stage to make use of the economies of scale. This can be seen in figure 6.4b. As the number of intermediate facilities increased, more waste was allocated to the intermediate facilities as opposed to being sent directly to the landfills. This behaviour was also the case for when the discount factor and the capacity of the intermediate facilities were increased.



(a) Transportation cost vs number of intermediate facilities



(b) Facility locations

Figure 6.4: Affect of increasing number of facilities on transportation cost and waste allocation

Despite the fact that transport costs decreased as more intermediate facilities were located, *at what point do the cost benefits no longer outweigh the costs of building intermediate facilities?* To do this the annual cost savings per k intermediate facilities were compared with the capital cost required for building those k facilities to calculate a break even point (in years). This was done for a range from 1 to 10 intermediate facilities. The results were plotted on a graph in order to compare the scenarios. The project with the shortest break even period would be the most cost-effective scenario to implement considering the capital cost of opening that many facilities and the cost savings reaped from the scenario.

Assuming a capital cost per intermediate facility of **R 20 million**, the break even curve for up to 10 intermediate facilities is represented in figure 6.5.

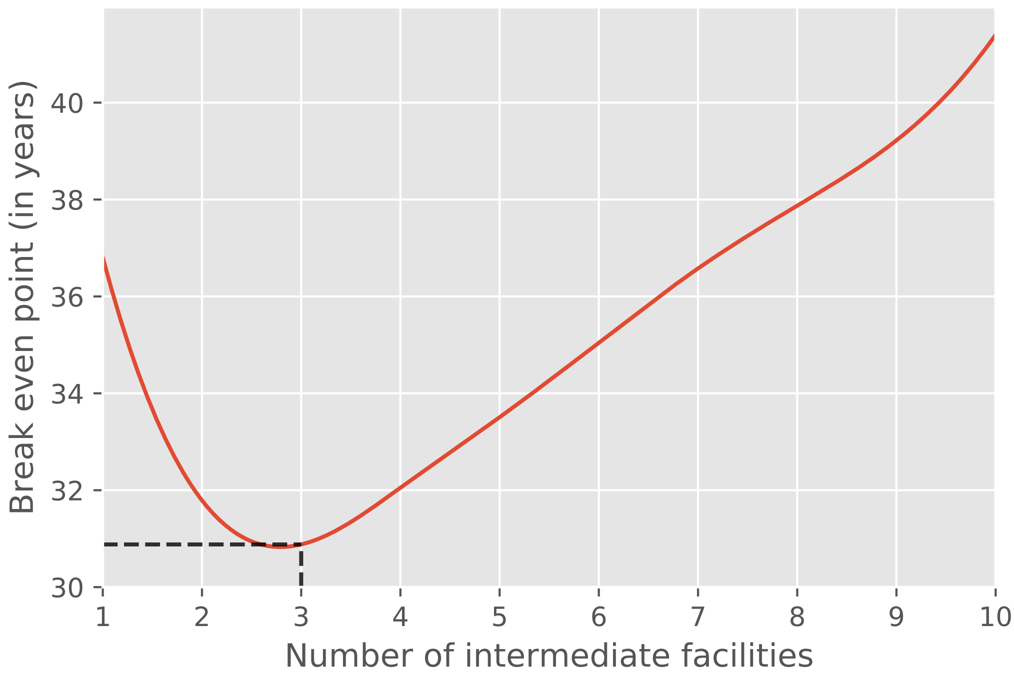


Figure 6.5: Break even curve for 10 intermediate facilities

The dashed line shows the shortest break even period and its associated number of intermediate facilities. With a break even period of 30.88 years, locating three intermediate facilities had the best cost benefit to capital cost ratio (i.e. break even period). The locations of the three intermediate facilities with respect to the existing facilities can be seen in figure 6.6.

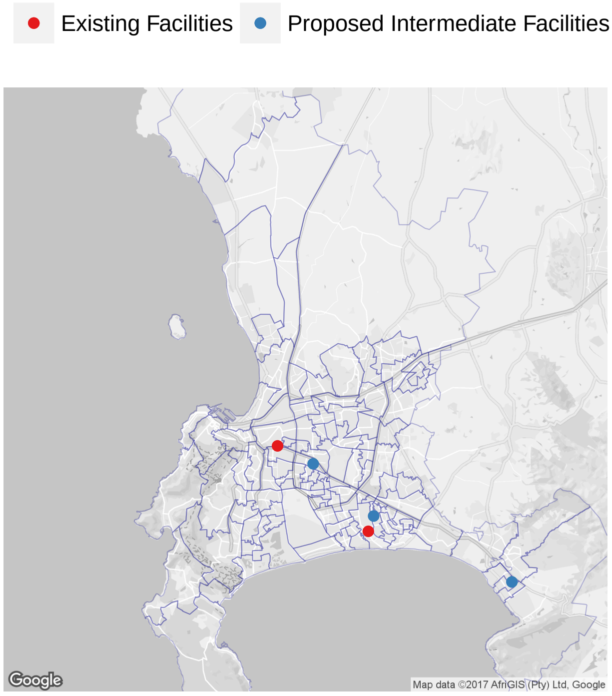


Figure 6.6: Proposed intermediate facilities vs existing locations

However, it is important to understand that this was just a test-case and that factors such as time-value of money, operating expenses of intermediate facilities, varying fuel costs and increases waste volumes were not considered. They would undoubtedly influence the results.

6.4 Discussion

Although the results showed that locating the intermediate facilities with the *Weiszfeld method* is superior to the location of the existing facilities, it is important to understand that these locations were based purely on Geographical Information System (GIS) data for collection beats and their estimated waste generation and no other information. Other factors may influence the results. It is therefore necessary to always confirm the potential location with a site visit in order to confirm the results of optimal sites Bosompem et al. (2016).

From the travel time analysis it was clear that the more number of intermediate facilities were located, the less the average travel time per vehicle was. In order to decrease the total travel time for waste vehicles in the first echelon, more intermediate facilities should be introduced. One can deduce that the converse is true for the waste vehicles in the second echelon as they must then visit more intermediate facilities to transfer the waste to the landfills.

However, the cost benefit analysis showed that even so, the total transportation cost for both echelons decreased with increasing number of intermediate facilities. Yet, the cost savings of locating more intermediate facilities with respect to the capital cost did not always increase. In fact, in this particular case, the best scenario was to locate three new intermediate facilities with a total capital cost of R 60 million at the locations specified in figure 6.6, and no longer use the existing intermediate facilities.

Chapter 7

Conclusion and Future Work

In this report the potential location for intermediate waste facilities and the impact thereof on waste collection operations were addressed.

In Chapter 1, the background concerning waste collection operations, the importance thereof and the problem statement were presented, among others. The Literature Review gave an insight to the existing literature available and what approaches had previously been taken to address similar problems. The approaches taken for generating the data sets were also presented as well as the models used to identify potential locations for intermediate facilities and to evaluate their impact.

Although the location of existing intermediate facilities fall short of the intermediate facility locations using the *Weiszfeld method*, it is important to understand that many assumption were made which may not actually justify this discrepancy. For example, although the capital costs were included in the *waste allocation model*, nowhere did we consider the cost of removing the existing facilities. Neither were other costs such as facility maintenance costs, un-loading costs or operating expenses for the facilities considered.

Furthermore, for the *Weiszfeld algorithm* its only two required input data sets were both simplified. The collection beats were reduced to single points, while the waste generation data was not verifiable and was generated by other data sets.

Additionally, the distances used throughout the project were also extremely simplified, either to the euclidean distance or the haversine distance. These distances also do not accurately represent the actual travel distance that a waste collection vehicle might take, notwithstanding that there might be cases (especially in Cape Town) where waste vehicles are not able to traverse as the roads are too narrow, or the there is only a one-way street.

Finally we also mentioned that other benefits from locating intermediate facilities had not yet been considered. For example, the added benefit of locating intermediate facilities is that since the intermediate facility is closer to other collection beats, the number of vehicles required in the first echelon could be dramatically reduced, since a vehicle is able to collect more waste in less time, since the travel distance between collection beat and landfill is drastically reduced. In this project these benefits were ignored as the assumption remained that one waste collection vehicle services one collection beat.

Therefore, the following could be done in the future in order to improve the accuracy and validity of the project:

1. increase the distance accuracy by using for example Google's Distance Matrix API;

2. apply a more rigorous process to identify potential locations for intermediate facilities;

3. possibly solve the project as a variant of the Vehicle Routing Problem (VRP) specifically as a Vehicle Routing Problem with Intermediate Facilities (VRPIF) in order to reap the additional benefits of the intermediate facilities;

4. use more accurate and relevant data, for example getting more accurate waste generation data could improve the accuracy of the project.

Even so, the project demonstrated novel approaches for generating additional data sets with limited available data. It also showed that intermediate facilities can be located using the *generalised Weiszfeld method for the multi-facility location problem* and how their impact can be evaluated using two models; a *travel distance model* and a Linear Programming (LP) *waste allocation model*.

This project can be used as the starting point for future research and has shown a clear path for how results can be refined for future iterations.

Bibliography

Agrawal, S. (2013). Machine Learning - DBSCAN. https://algorithmicthoughts.wordpress.com/2013/05/29/machine-learning-dbscan/. Online; accessed 3-September-2017.

Bosompem, C., Stemn, E., and Fei-Baffoe, B. (2016). Multi-criteria gis-based siting of transfer station for municipal solid waste: The case of kumasi metropolitan area, ghana. *Waste Management & Research*, 34(10):1054–1063.

City of Cape Town (2017). City of Cape Town annual budget 2017/18 - 2019/20. http://www.capetown.gov.za/Family%20and%20home/Meet-the-City/the-city-budget/The-Citys-Budget-2017-2018.

Crainic, T. G., Perboli, G., Mancini, S., and Tadei, R. (2010). Two-echelon vehicle routing problem: a satellite location analysis. *Procedia-Social and Behavioral Sciences*, 2(3):5944–5955.

CSIR Building and Construction Technology (2005). *Human Settlement Planning and Design*, volume 2. CSIR Building and Construction Technology.

Eiselt, H. A. (2007). Locating landfillsoptimization vs. reality. *European Journal of Operational Research*, 179(3):1040–1049.

Eiselt, H. A. and Marianov, V. (2015). Location modeling for municipal solid waste facilities. *Computers & Operations Research*, 62:305–315.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231.

Farahani, R. Z., Hekmatfar, M., Arabani, A. B., and Nikbakhsh, E. (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering*, 64(4):1096–1109.

Hoornweg, D. and Bhada-Tata, P. (2012). What a waste: a global review of solid waste management. *Urban development series knowledge papers*, 15:1–98.

Iyigun, C. and Ben-Israel, A. (2010). A generalized weiszfeld method for the multi-facility location problem. *Operations Research Letters*, 38(3):207–214.

Kahle, D. and Wickham, H. (2013). ggmap: Spatial visualization with ggplot2. *The R Journal*, 5(1):144–161.

Mitropoulos, P., Giannikos, I., and Mitropoulos, I. (2009). Exact and heuristic approaches for the locational planning of an integrated solid waste management system. *Operational Research*, 9(3):329–347.

ReVelle, C. S. and Eiselt, H. A. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165(1):1–19.

Appendix A

Weiszfeld Locations

Table A.1: Table showing the 25 potential intermediate facility locations

	XCoord	YCoord
0	18.712807110067782	-33.85711528004055
1	18.516465960006666	-33.96785460999634
2	18.6515972	-33.963851330000004
3	18.511482500041097	-33.81786668005036
4	18.567621847116392	-33.97029320213694
5	18.701078419988875	-34.06064018997564
6	18.856917610000004	-34.11279587
7	18.48818928	-33.55391653000001
8	18.477853264884782	-34.05734740341308
9	18.50480412038433	-33.92349569986156
10	18.655660900000004	-34.03332516999999
11	18.58163093000046	-33.916753720000024
12	18.452388860000003	-33.93340046
13	18.633334900000275	-33.89671335001809
14	18.49549166744577	-34.03196860424894
15	18.516638740000396	-34.039005569999965
16	18.34627769038505	-34.03130630995192
17	18.692484909999997	-33.99668034
18	18.38930131226647	-34.13483586383872
19	18.61627861564411	-34.03304298094496
20	18.559215537027686	-33.922332512988724
21	18.687923503042967	-33.85646828705292
22	18.534331900000307	-33.98023397000012
23	18.60030451542234	-34.010294744359456
24	18.45194658979699	-33.73502547018385

Appendix B

Project Sponsorship Form

Department of Industrial & Systems Engineering
Final Year Projects
Identification and Responsibility of Project Sponsors

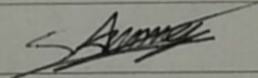
All Final Year Projects are published by the University of Pretoria on *UPSpace* and thus freely available on the Internet. These publications portray the quality of education at the University and have the potential of exposing sensitive company information. It is important that both students and company representatives or sponsors are aware of such implications.

Key responsibilities of Project Sponsors:

A project sponsor is the key contact person within the company. This person should thus be able to provide the best guidance to the student on the project. The sponsor is also very likely to gain from the success of the project. The project sponsor has the following important responsibilities:

1. Confirm his/her role as project sponsor, duly authorised by the company. Multiple sponsors can be appointed, but this is not advised. The duly completed form will considered as acceptance of sponsor role.
2. Review and approve the Project Proposal, ensuring that it clearly defines the problem to be investigated by the student and that the project aim, scope, deliverables and approach is acceptable from the company's perspective.
3. Review the Final Project Report (delivered during the second semester), ensuring that information is accurate and that the solution addresses the problems and/or design requirements of the defined project.
4. Acknowledges the intended publication of the Project Report on UP Space.
5. Ensures that any sensitive, confidential information or intellectual property of the company is not disclosed in the Final Project Report.

Project Sponsor Details:

Company:	Optimisation Group, Department of Industrial & Systems Engineering, University of Pretoria
Project Description:	Based on preliminary results from Project 023 the Optimisation Group found that waste vehicles spend most of their time traveling between service locations, dumpsites and depots, and very little time on actually collecting waste. This is partly due to the de-centralised locations of landfill. One way to improve the service utilisation of vehicles is to use centralised transfer stations. The purpose of this project is to identify and use location models to identify potential locations for transfer stations, and then evaluate the impact of the stations on waste vehicle operations.
Student Name:	Sebastian Ammon
Student number:	14053862
Student Signature:	
Sponsor Name:	**Dr Elias Willemse**
Designation:	Senior Lecturer

E-mail:	elias.willemse@up.ac.za
Tel No:	+27 12 420 3443
Cell No:	/
Fax No:	/
Sponsor Signature:	