



Optimization of solid waste collections in Blantyre, Malawi

Author: Nicolas Seemann-Ricard

Supervisor: Elizabeth Tilley

Tutor: Florian Dörfler

21 October 2022

Abstract

Your abstract should succinctly summarise the research gap, the methods you employed, your results, conclusions, and recommendations. Don't use acronyms if possible and keep the language as general as possible. Keep the abstract to a maximum of 500 words. The abstract stays on its own page.

Table of Contents

1	Introduction	4
1.1	Background topic 1 (Solid Waste Management)	4
1.2	Background topic 2 (SWM in Africa)	4
1.3	Justification and Research Questions	4
2	Data analysis	5
2.1	Set of skip locations	5
2.2	Skips filling data	6
2.3	Dump arrivals logs	13
2.4	Estimating filling rates	Error! Bookmark not defined.
3	Results and Discussion	Error! Bookmark not defined.
3.1	Discussion	Error! Bookmark not defined.
4	Conclusions and Recommendations	Error! Bookmark not defined.
5	References	21

1 Introduction

The project seeks to minimize the costs of operation of the municipal solid waste management service in Blantyre Malawi, while limiting overflowing of the communal skips.

1.1 Background topic 1 (Solid Waste Management)

1.2 Background topic 2 (SWM in Africa)

Maybe now you dive into the global differences in collection

1.3 Justification and Research Questions

1. How can we model with limited data?
2. How many trucks are needed to service all the skips? What would be the mileage and cost of servicing all skips without overflow?
3. What would be the optimal routing schedule?

2 Data analysis

In order to formulate feasible and pertinent recommendations, parameters reflecting the situation need to be calculated. Specifically, the location of skips and the rate at which skips are filling, to know the frequency at which they need to be emptied (services). The more robust set, as seen in this section, is a timeseries of specific skips filling and emptying. However, the scope of it is quite narrow, with only 12 useful filling rates extracted, all in a small area of Blantyre. A second dataset gives the arrivals at Mzedi dump, the main waste ***word*** in Blantyre. Though it covers all the skips studied, it presents strong limitations. Namely, arrivals at the dump do not reflect the speed at which the skips fill up. Indeed, the individual filling data show that some skips go a long time without being emptied, overflowing and presenting a public health risk.

2.1 Set of skip locations

A set of skips, taken to be all the community skips to be serviced in Blantyre, along with GPS coordinates is given. Each skip is 7m^3 . Some data points have the same name, when several skips are in the same area (e.g., three skips in the same market). It is unknown if more than one skip can be represented by one coordinate point. Furthermore, there is no information about the intended nature of the waste (organic or inorganic). The locations of the municipal dump (Mzedi dump), the truck storage facility and the compost facility are also given. The locations are mapped in Figure 1.

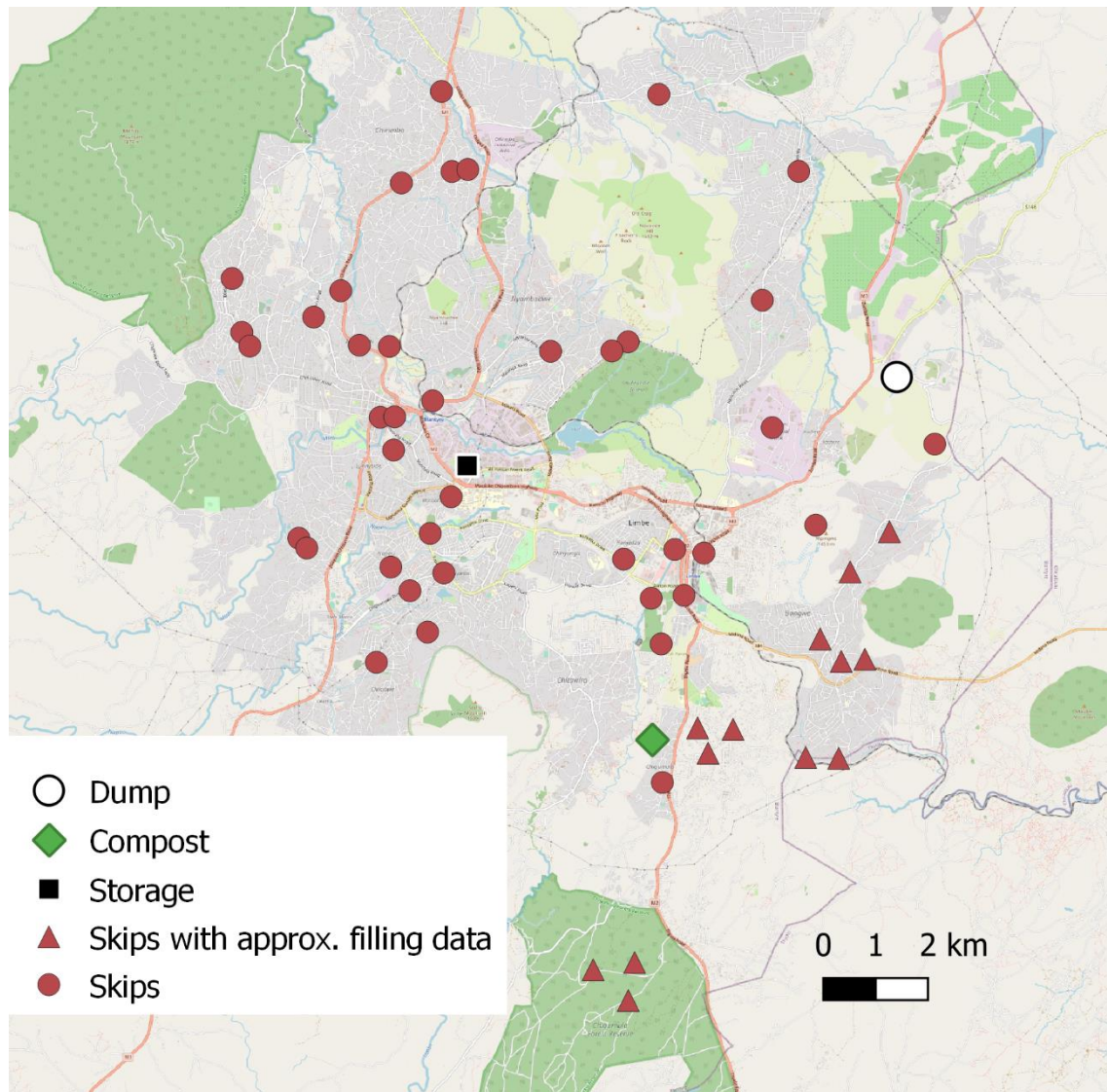


Figure 1 Locations of skips, dump, compost facility and truck storage in Blantyre

2.2 Skips filling data

Filling data for several skips is provided. The time series are named in the following fashion: “area+type of waste+number”. The names of the areas do not always match the ones from the set of skip locations, in addition to the ambiguity of several skips present in a certain area. The type of waste also complicates the analysis, as some are marked as organic, and other inorganic.

The areas covered are Bangwe, BCA, Chigumula and Naizi.

While efforts have been made to separate organic and inorganic waste at the skip level, those have, to my knowledge, not been successful. The composting facility does not seem to be in use.

Over a certain period (depending on the skip), a measurement on a scale from 1-5 was taken visually (generally) every day at those skips. A score between 0 and 4 indicate the estimated fullness of the

skip, while a 5 means the skip was overflowing. Three of the 14 provided sets these are shown in Figure 2-4.

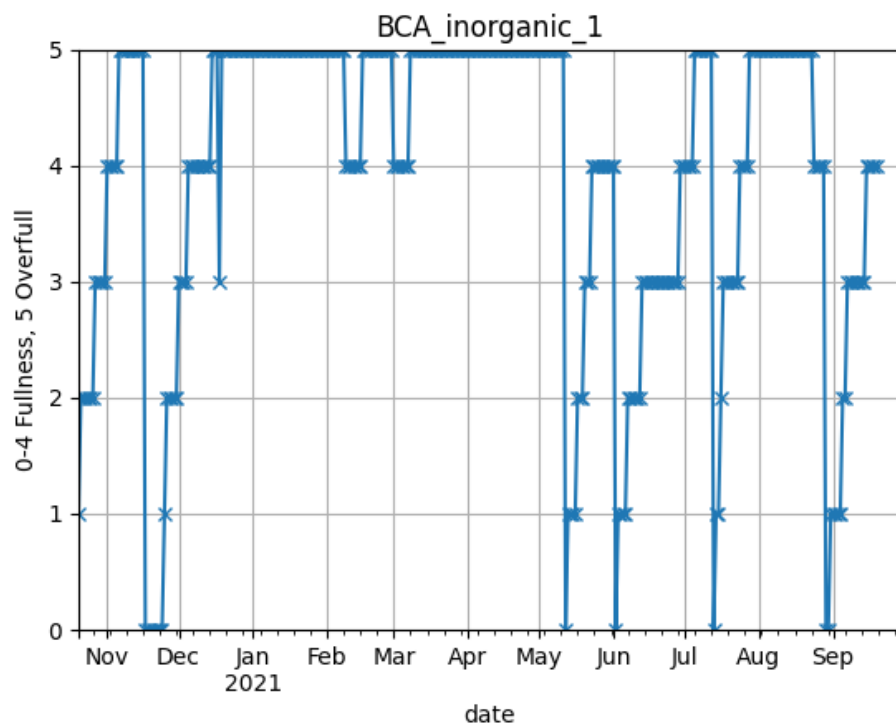


Figure 2 Skip filling data for BCA inorganic 1

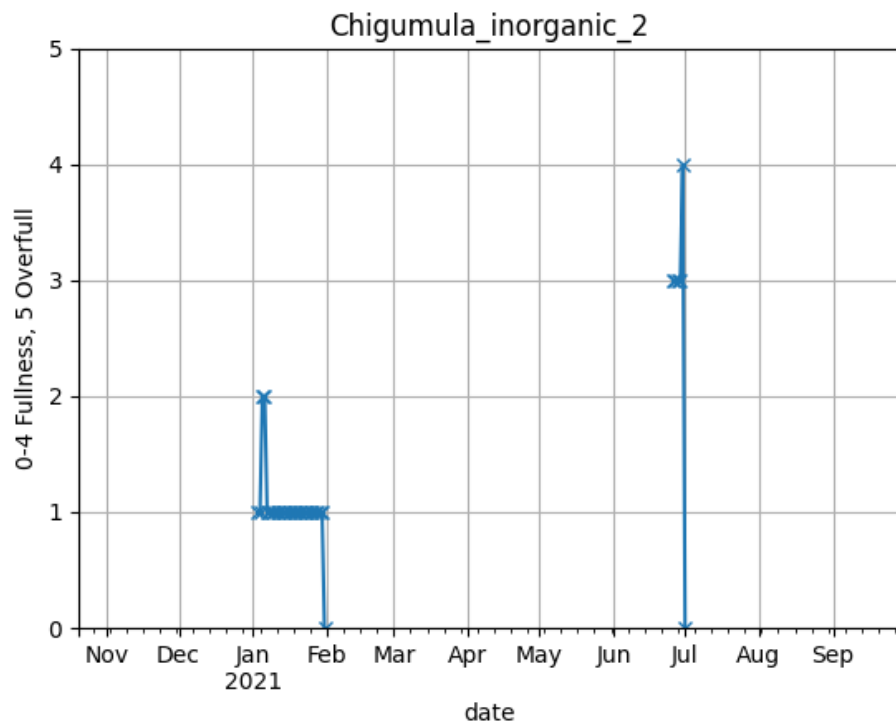


Figure 3 Skip filling data for Chigumula inorganic 2

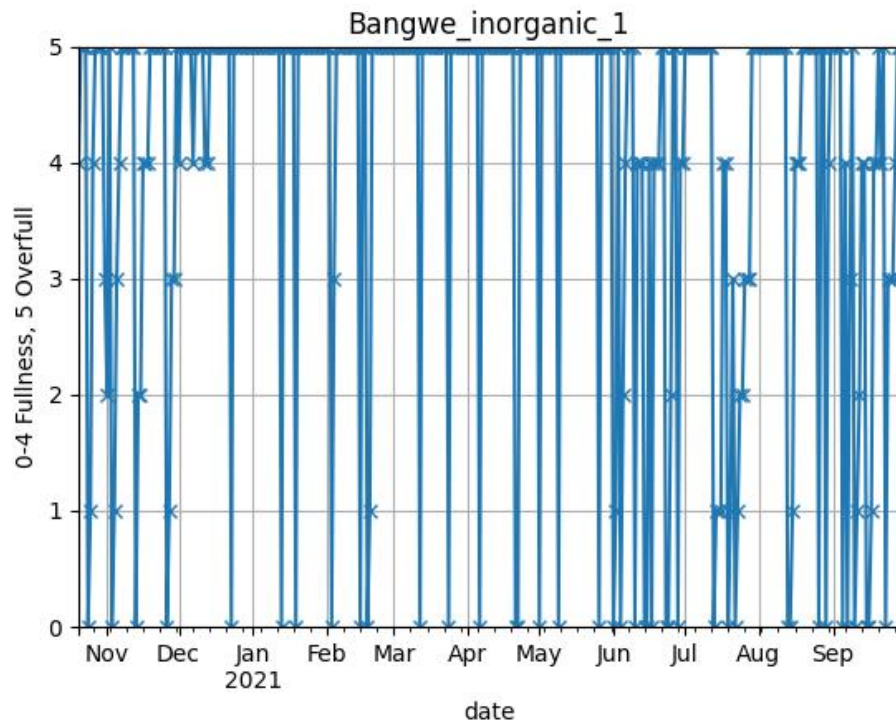


Figure 4 Skip filling data for Bangwe inorganic 1

BCA Inorganic 1 gives a good example of a profile from which rising trends (ramps) are discernible and allow to easily calculate a filling rate. This is useful in estimating the frequency at which the skip needs to be serviced. Chigumula Inorganic 2, however, simply does not have enough data points. Finally, Bangwe inorganic 1 seems to be quite active. On many occasions the fullness dips for only one day only to be full the next day. This seems to indicate that it fills up quickly.

An algorithm is developed to extract the rising ramps, in order to estimate the filling rate of the skips in the dataset. It performs the following steps:

1) Removing spikes

Spikes occur in the dataset frequently. An example is Figure 2, between December and January. After a relatively slow ramp up, the level drops to a level 3 then back to level 5. There are also positive spikes which appear unlikely. Those are taken as measurement errors. Spikes are therefore defined as a single point where the value before and after are the same. Although it might remove some legitimately fast filling events, such as ones seen in Figure 2, the erroneous spikes in most timeseries have a stronger bias on the ramps estimations.

2) Detecting the top of ramps

The top of the ramp is defined here as the end point of a rising ramp fitted to the time series. These may occur at any level. They are determined by a forward and a backward pass through the series. The forward pass identifies top “ends”, when the next value of fullness is smaller than the current one (which characterises the skip being emptied). A top end is also identified when the current value is 4 and the next is a 5 (or overfull). The reasoning behind this is the preference for ramps to finish with a 4. This level is defined as full, which is more precise than the “overfull” denomination. Still, in the absence of an intermediate 4 in a transition e.g., between 3 and 5, the top value will be the 5.

From the top “ends”, the top “beginning” values are extracted. They are given to be the first appearance of the top value in a plateau (a series of the same level). They are the points considered to be the time where the skip achieves the fullness level at which it is emptied, or when it reaches full capacity. This shift is particularly important when long periods during which the skip is not emptied occur. Using the top ends to calculate the filling rate would bias it to a lower rate.

3) Eliminating derating ramps

Derating is clear in Figure 2 on two occasions between February and April. This might occur for several reasons, such as:

- Trash being burned to eliminate the overflowing waste.
- Waste being cleared by another party.
- An error in data collection or ambiguity in the measurement scale.

In the backward pass described in 2), the top values at the end of derating periods are removed by adding a condition that deletes top ends if the value at the beginning of the 4 plateau is 5.

4) Detecting the bottom of ramps

Finally, the bottom values and dates are extracted by simply iterating backward from each top beginning, until the previous value is larger than the current one, at which point it is assumed the skip has just been emptied.

The pseudo code for each block is shown in **APPENDIX**.

The result is shown in Figure 5 for BCA inorganic 1 and in Figure 6 for Bangwe inorganic 1. Clearly, the fit in the second case is less ideal. In Figure 7, the removal of spikes is effectively displayed.

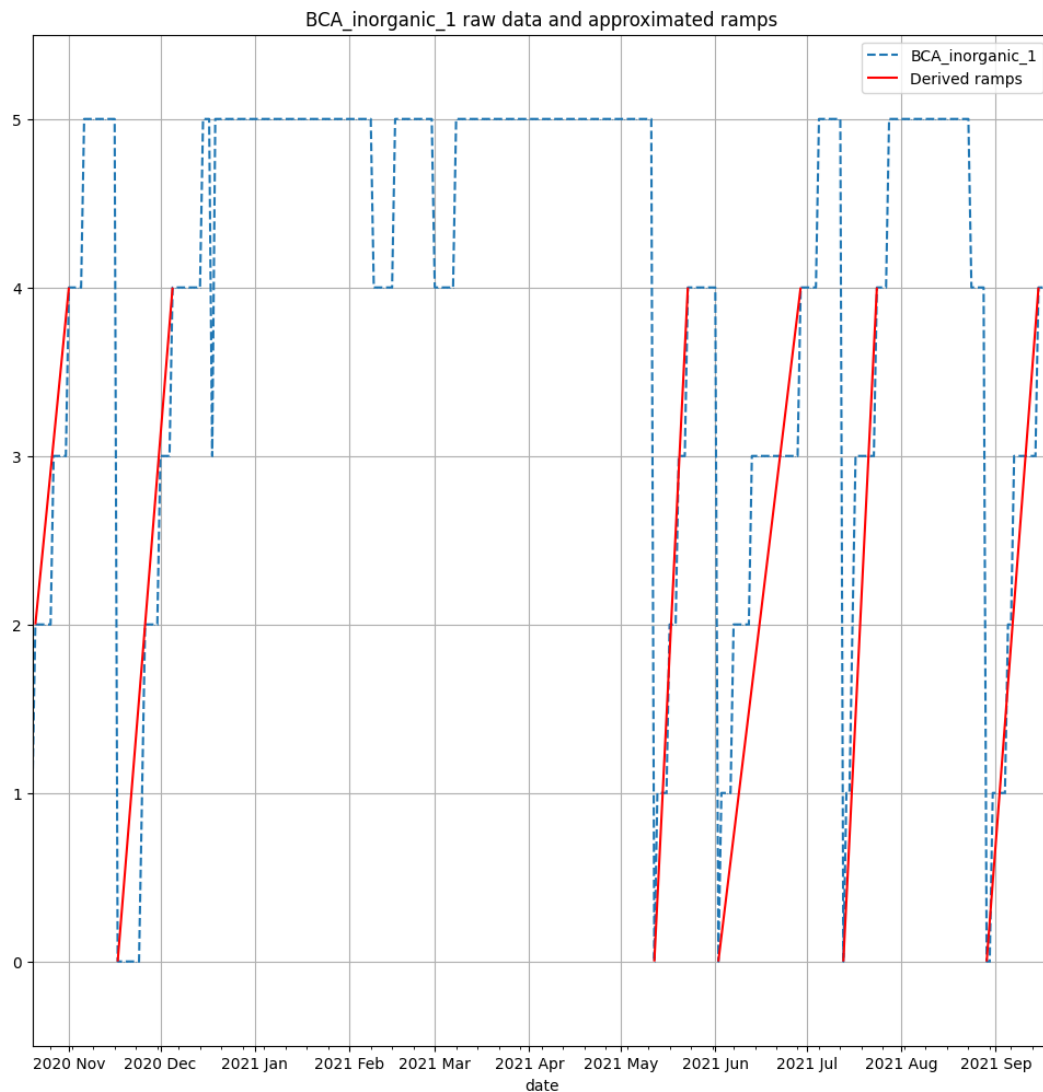


Figure 5 Results of ramps for BCA inorganic 1 (spikes removed)

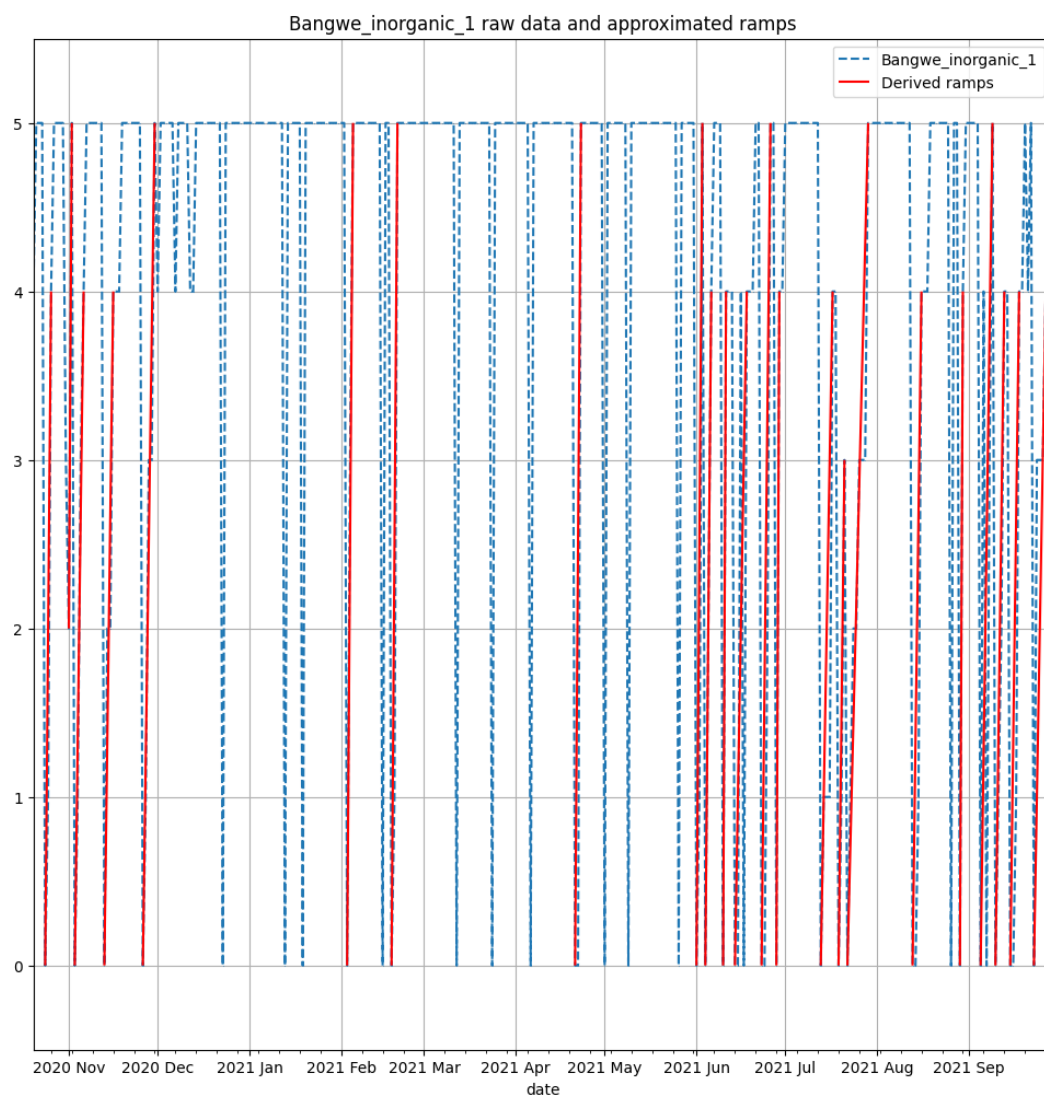


Figure 6 Results of ramps for Bangwe inorganic 1 (spikes removed)

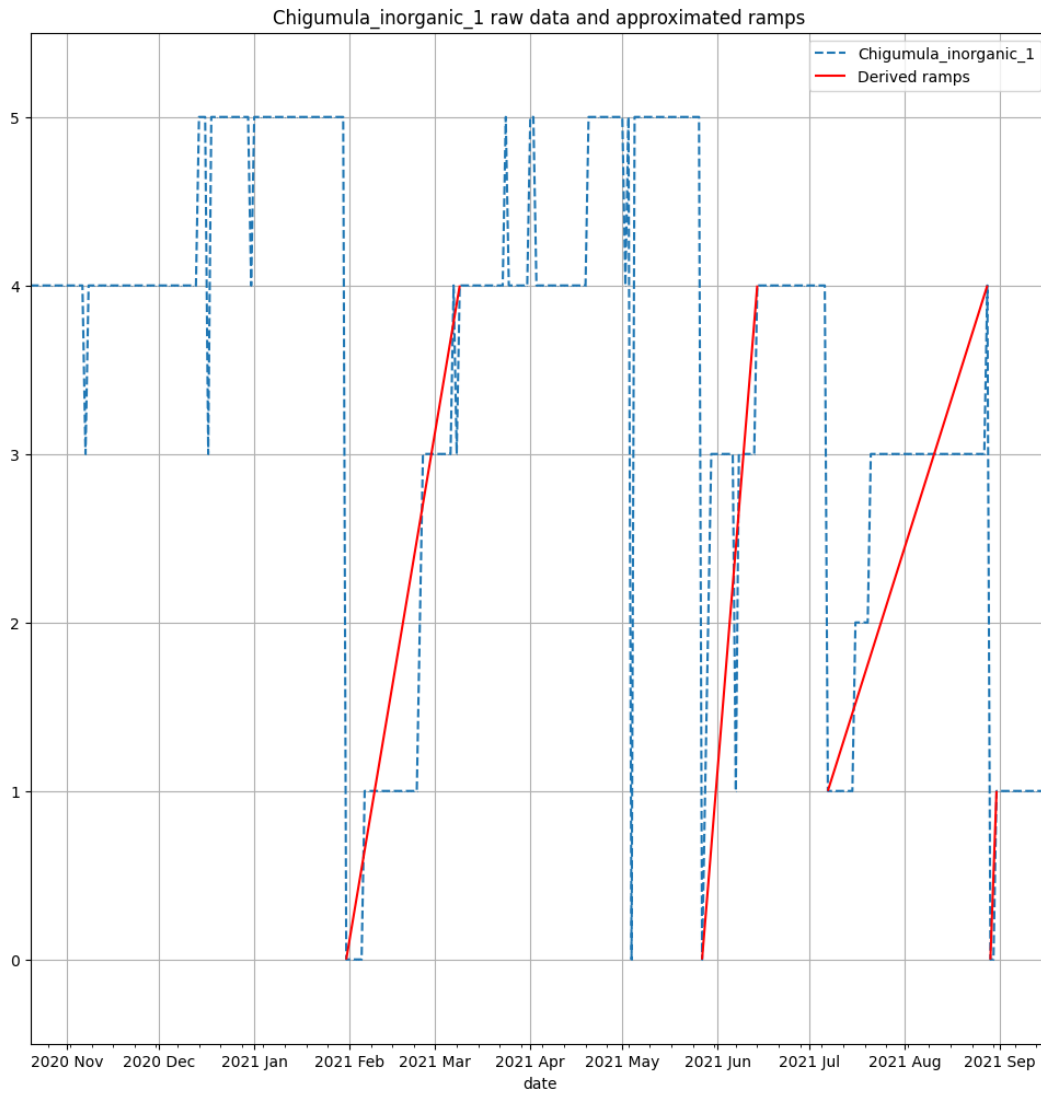


Figure 7 Results of ramps for Chigumula inorganic 1 (spikes removed)

The aggregate results are shown in Table 1. The minimum and maximum values indicate if spikes were removed (min) or not (max). For skips such as Bangwe inorganic 1, it turns out to make a large difference, as spikes and quick filling of the skip are confounded. Chigumula inorganic 1 also has a large difference between the minimum and maximum filling rates. However, in this case and as shown in Figure 7, the ramps fitting with spikes removed is correct, and including the spikes incurs significant noise. In Table 1, one of the two filling rates is highlighted as chosen qualitatively based on the visual fit. Having recourse to this is attributed to the low granularity of data as well as unexplained events that result in spikes.

Table 1 Aggregate data from ramps analysis

Skip	Avg filling rate min	Avg filling rate max	# min	# max	Prop. overfull	period (days)

Bangwe_Organic_1	0.215	0.215	24	26	0.258	341
Bangwe_Organic_2						
Bangwe_inorganic_1	0.49	0.766	23	36	0.725	345
Bangwe_inorganic_2	0.4	0.541	16	20	0.737	265
BCA_Organic_1	0.091	0.105	11	13	0.386	324
BCA_Organic_2	0.077	0.077	2	2	0	54
BCA_inorganic_1	0.063	0.126	6	7	0.536	335
BCA_inorganic_2	0.038	0.047	1	2	0.627	178
Naizi_Organic_1	0.046	0.046	8	8	0.21	324
Naizi_Organic_2						
Naizi_inorganic_1	0.144	0.161	15	17	0.293	342
Naizi_inorganic_2						
Chigumula_Organic_1	0.158	0.212	21	24	0.183	327
Chigumula_Organic_2	0.127	0.142	13	19	0.106	223
Chigumula_inorganic_1	0.056	0.308	4	9	0.256	333
Chigumula_inorganic_2						

2.3 Dump arrivals logs

A separate time series is used. It lists arrivals at the Mzedi dump, along with the origin of the skip carried by each truck. The origins match exactly the set of skip locations, but once again, not the filling data. The period of this series is 2020-12-05 to 2021-12-31. The sum of arrivals in each week during this period is shown in Figure 8. A sizeable gap is noticeable for almost the whole of February 2021. This is reflected in the skips filling data, where many skips were overflowing and not emptied during this period. The reason for this gap is unknown but assumed here to be the service simply not operating. Intra-weekly, the arrivals are relatively homogeneous, except for Sunday, as illustrated in Figure 9.

Of particular interest is the number of days between arrivals at the dump for each skip. The analysis of these gaps is shown in Figure 10. A noticeable characteristic is the variability of the time gaps, as illustrated by the size of the boxes and whiskers. For areas that are serviced often, such as Blantyre Flea Market, Limbe, Ndirande and Chirimba (where the median gap is at maximum 3), the outliers are frequent, indicating gaps where the skips are likely overflowing.

Importantly, the arrivals at the dump provide useful insights into the current operation of the municipal solid waste management system in Blantyre. It does not, however, allow by itself, to infer the filling rate of bins. As seen in Table 1, at least some skips spend a considerable amount of time overfull. As such, the time between collections is dependent on other factors as well.

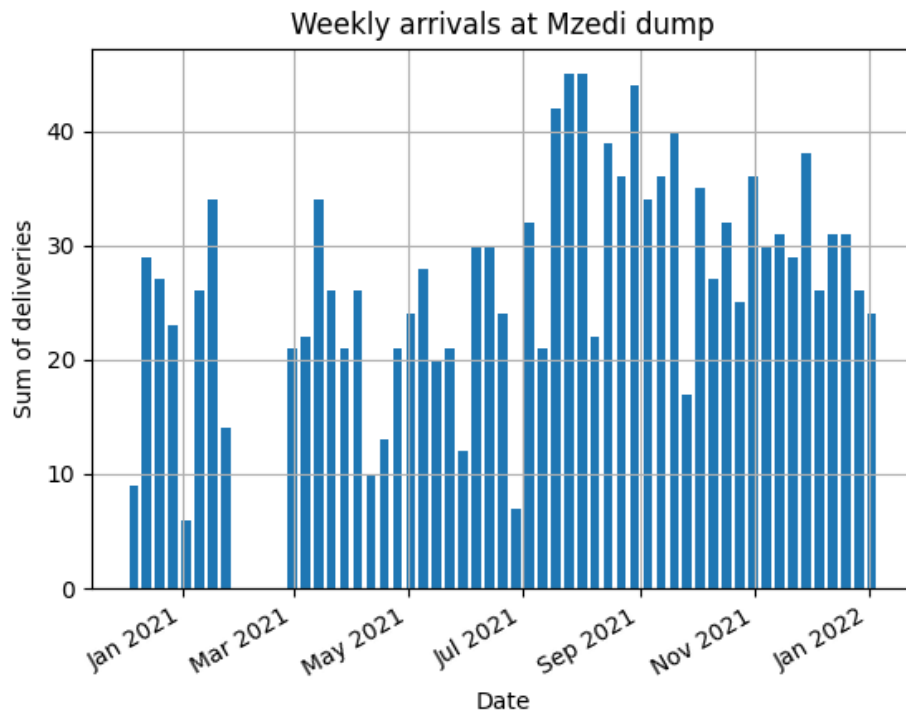


Figure 8 Sum of deliveries (arrivals) at Mzedi dump over the entire period of measurements

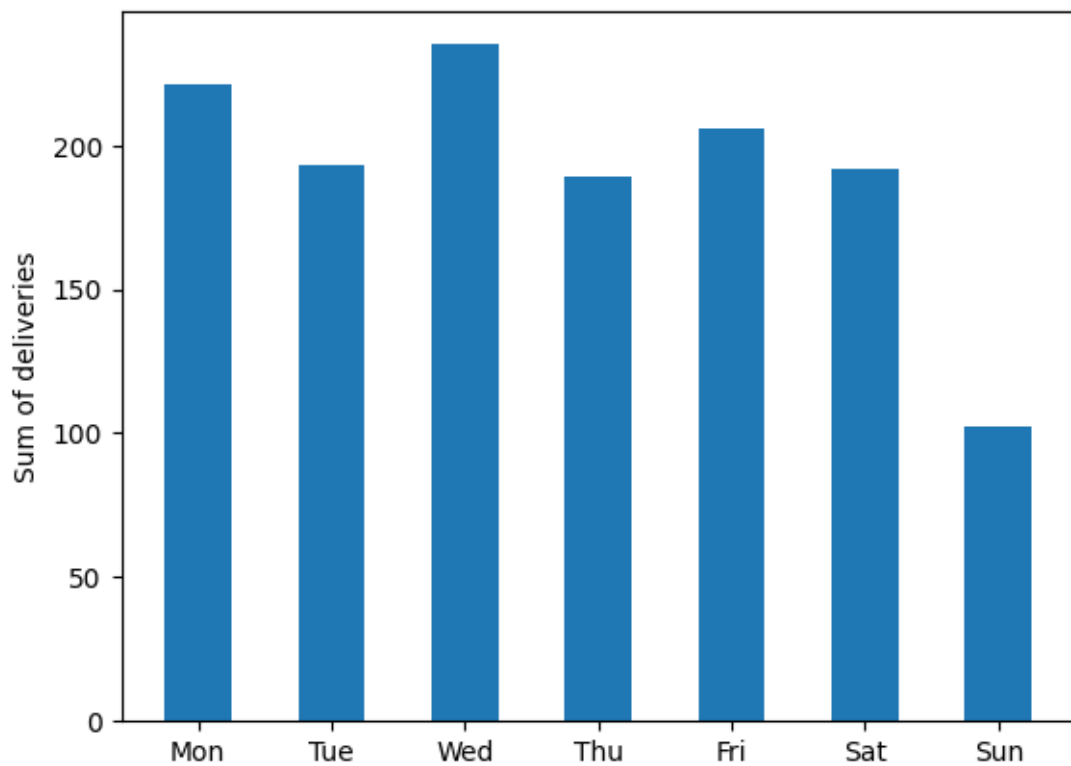


Figure 9 Sum of deliveries to Mzedi dump per weekday over the entire period of measurements

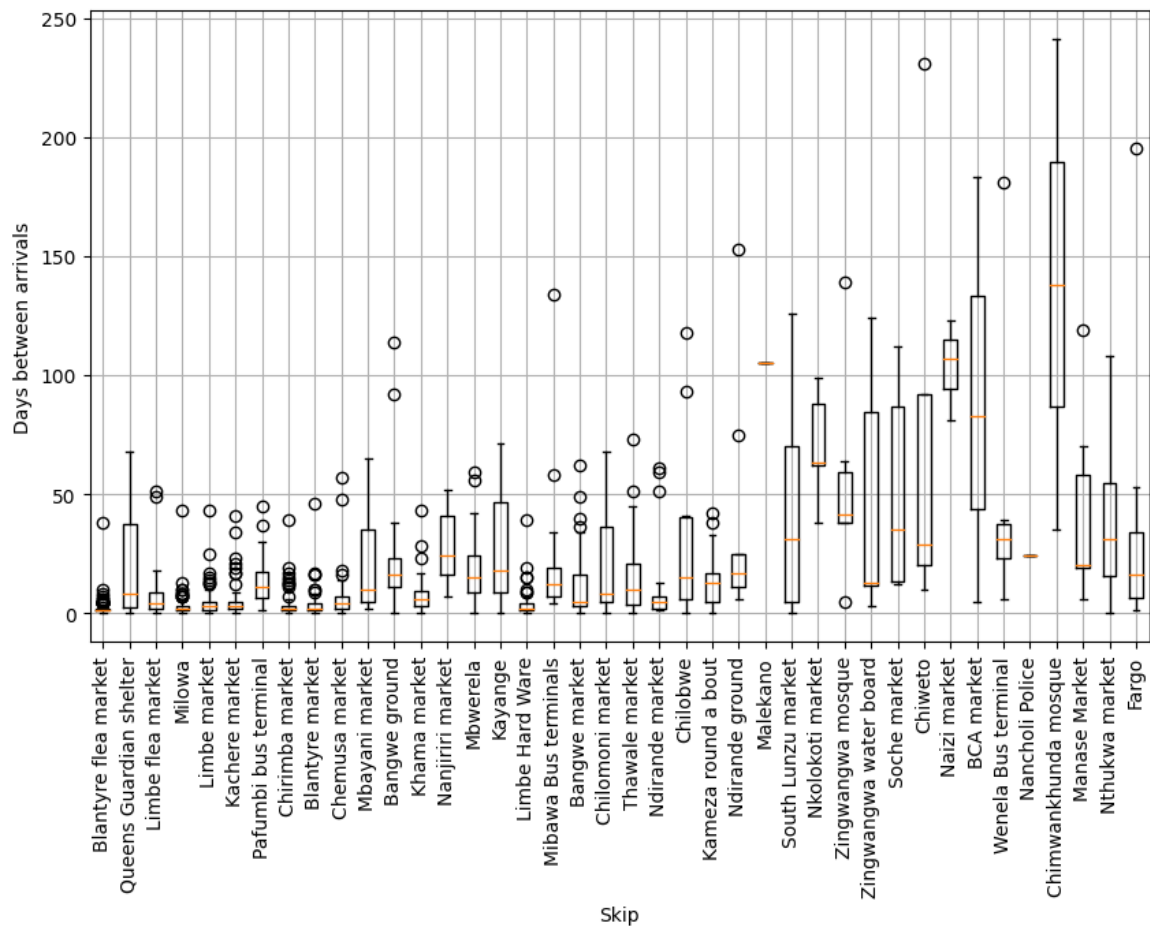


Figure 10 Box plots of number of days between arrivals for each skip

2.4 Relationship between the datasets

The first dataset, though useful in estimating the filling rates, is limited to twelve skips. Those skips are also concentrated in one area of Blantyre, as seen in Figure 1. This makes it difficult to extrapolate to other areas in the city. It is attempted to use characteristics from the two datasets to get filling rates estimates for each skip. Since the skips in the filling dataset and the areas from the dump logs do not match, they are aggregated as shown in Table 2. Chigumula, though in the original list of skips (see 2.1), does not appear in the dump arrival logs. This despite the fact that skip filling data indicates the inorganic skip was emptied at least 3 times in the period, and the organic skips were emptied dozens of times.

Table 2 Aggregation of skips and skip areas and number of data points

Aggregate area	Skips from filling data	# ramps	Area from dump logs	# arrivals
Bangwe	Bangwe_Organic_1,	24	Bangwe ground Bangwe market	13 30
	Bangwe_inorganic_1	23		
	Bangwe_inorganic_2	16		
BCA	BCA_Organic_1	11	BCA market	3
	BCA_Organic_2	2		

	BCA_inorganic_1	6		
	BCA_inorganic_2	1		
Naizi	Naizi_Organic_1	8	Naizi market	3
	Naizi_inorganic_1	15		
Chigumula	Chigumula_Organic_1	21	none	
	Chigumula_Organic_2	13		
	Chigumula_inorganic_1	4		

This points to a larger problem of mismatch in the data. Bangwe has the most data points in both data sets, so it is used to illustrate the issue. Since it is known that not all Bangwe ground and Bangwe market skips are included in the skips filling data, the emptying events (here we use the bottoms described in 2.2) should all fit within the arrival events described in the Mzedi dump logs. This holds even when quick ramps are wrongly identified as spikes, since they are simply disregarded. Figure 11 contradicts this notion, however. It shows many skip emptying events do not match with arrivals at the dump. In Figure 12, the proportion of emptying events matching arrival events is broken down for the three skips.

To account for errors in calculating the beginning of the ramps, a “padding” is introduced, where an emptying event is said to match if it is within a period of time defined as a margin. A margin of 1 indicates a period of 3 days, one before and one after the emptying event. The vertical line is the median of time differences between Mzedi arrivals (both from Bangwe ground and market) divided by 2, to account for the margin being on both sides. At this point, most skip emptying events should match. However, for all three skips, they are only around 50%.

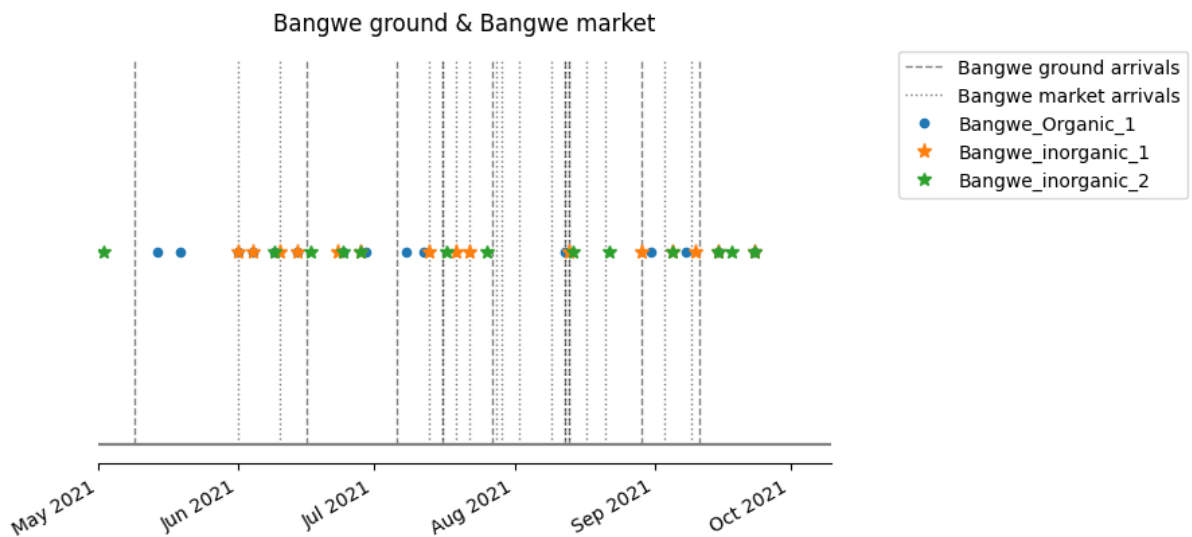


Figure 11 Timeline of Bangwe arrivals at Mzedi dump and emptying events of select skips

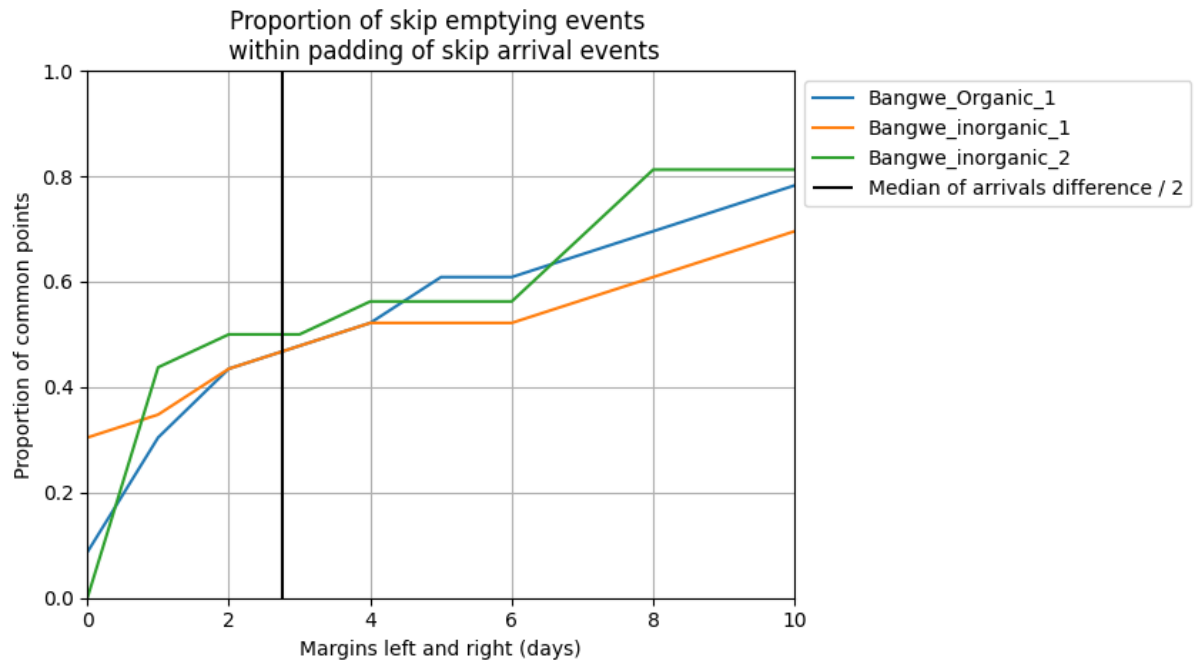


Figure 12 Proportion of skip emptying events matching arrivals at the dump

Possible reasons for this mismatch are:

- Inaccuracy of the ramp fitting method
- Measurement ambiguity in the skip filling data
- Issues with data logging at Mzedi dump, such as not all arrivals being logged

In Table 3, the aggregated data shows the difficulty to extract a filling rate from the dump logs. The weighted average ramps filling rate is weighted according to the number of ramps of each skip. The weighted average of median gaps is weighted (only for Bangwe), based on the number of arrivals from each sub-area described in the logs.

Table 3 Aggregated results for filling rate and gaps between arrivals

Aggregate area	Weighted average of ramp filling rates	# ramps	Proportion of time overfull	Weighted average of median gaps between arrivals	# arrivals
Bangwe	0.3261	63	0.5571	10	43
BCA	0.0786	20	0.4673	83	3
Naizi	0.1102	23	0.2523	107	3

2.5 Generating artificial filling data

Though the histogram of the filling rates indicate a possible lognormal distribution, the data is too scarce to do a fit, as shown in Figure 13. Interestingly, the reciprocal of the median emptying frequency exhibits a similar behaviour, as seen in

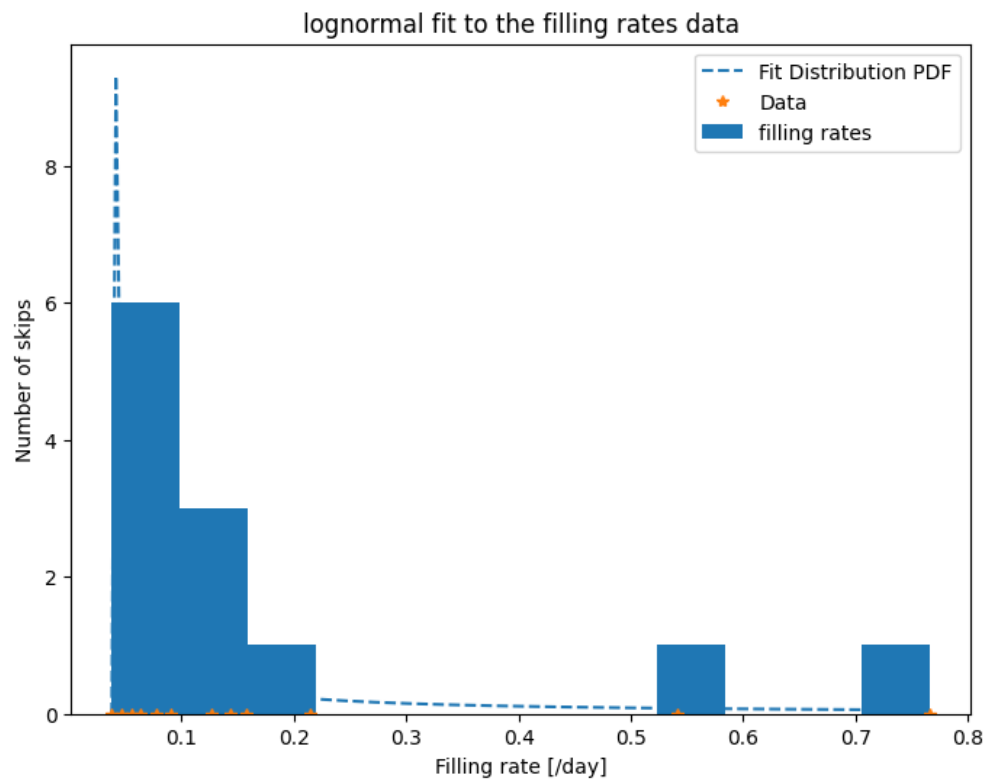


Figure 13 Lognormal fit of the filling rates skip level data

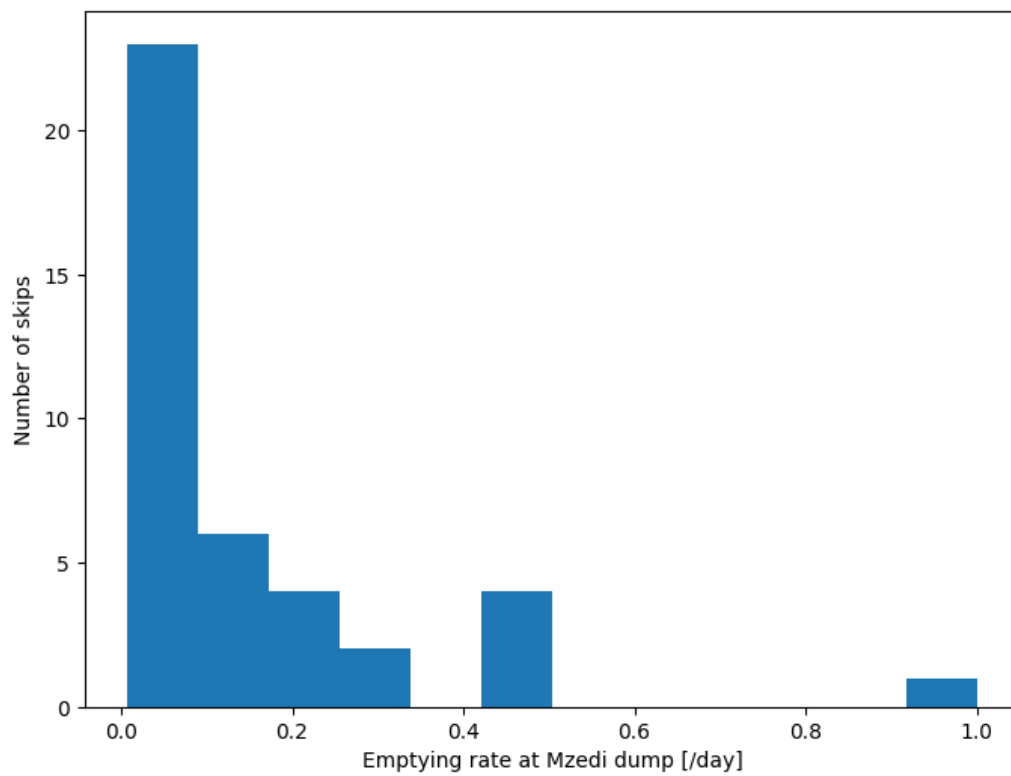
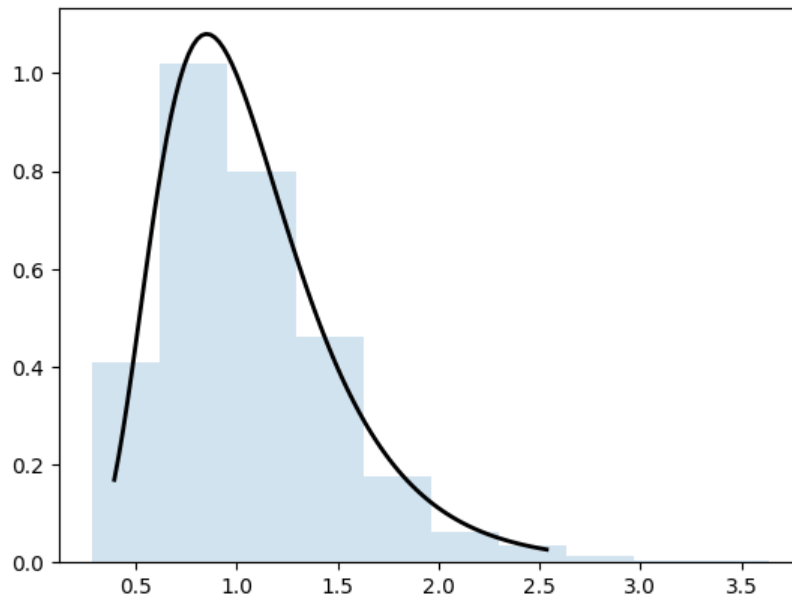


Figure 14 Histogram of emptying rates by skip arrivals at Mzedi dump

Given the incompleteness of the available data, an artificial dataset of filling rates is created. It is based on the filling rates, with an added random element.

Other possibility: artificial log-normal distribution with a specific number of bins at an upper value bound and maximum probability at system median.



3 Optimization model

3.1 Problem formulation

- Features:
 - **28 days planning period**
 - Morning and evening collection (e.g. an afternoon collection on Saturday and a morning collection on Monday)
 - **No emptying on Sunday**
- Assumptions:
 - **Constant filling rate over the planning period**
 - Constant speed (i.e. time of travel is directly proportional to distance)
- Objective: Minimize cost of capital, labour and of operation in addition to overflowing costs:
 - Cost of capital:
 - Number of vehicles
 - Number of extra skips
 - **Cost of labour: based on number of days operating**
 - **Operational costs: based on distance travelled**
 - Overflowing costs: based on amount overflowed
- Constraints:

- **Visits following a regular pattern**
- **Emptying schedule preventing overflowing** (with some overflow allowable subject to high costs)
- Daily operation constraints (2x4hour blocks per day)

Currently, the optimization algorithm is implemented in MATLAB, and solved by Gurobi through YALMIP. The points in bold in the previous list have been implemented for now. It is to note that the mathematical formulation of the problem involve many equality and inequality constraints that do not show in the above logistical problem.

Figure 15 gives the schedule representation of a first optimization formulation. The implication of the daily cost is obvious in the fact that many skips are scheduled on the same days. This is also caused by the lack of time constraints for working days, which would severely limit this. Additionally, the filling rates in this case are randomly generated in a range between 0-0.5. The upper limit of 0.5 is because the model does not offer any flexibility around overflowing (which should have an associated cost), and the collections are assumed to be at the same time every day (instead of morning/afternoon collections).

Additionally, the problem assumes one vehicle only, and that all collections are done by bringing an empty bin and exchanging it (meaning as many empty bins as vehicles). Since there are multiple trucks, the formulation will have to be adapted.

Finally, this is formulated as a one stage optimization problem, focusing on operation. However, the number of trucks, additional skips will become decision variables and the problem will become a multi-stage problem. This may increase complexity drastically and require dynamic programming.

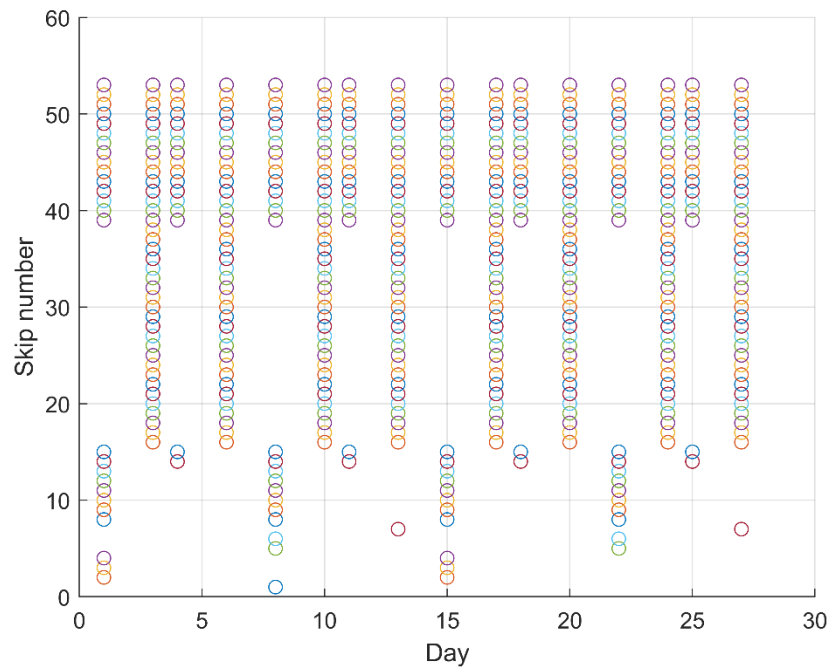


Figure 15 Schedule diagram (first version, more work will be done on operational planning representation)

4 Results

4.1 Capital

4.2 Operational planning

5 Recommendations for future work

5.1 Skip-level data

- More complete data, with a more robust scale, two measurements per day and a more spatially and categorically representative set of skips.

5.2 Current organizational and operational state

- Current operational protocols
- More specific constraints
- Accurate costs analysis
- Measured time of travel between skips
- Stakeholder analysis

6 References

