

**Optimization of solid waste collections in Blantyre, Malawi**

Author: Nicolas Seemann-Ricard

Supervisor: Elizabeth Tilley

Tutor: Florian Dörfler

15 November 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**

Contents

[1 Introduction 5](#_Toc119395221)

[1.1 Background topic 1 (Solid Waste Management) 5](#_Toc119395222)

[1.2 Background topic 2 (SWM in Africa) 5](#_Toc119395223)

[1.3 Justification and Research Questions 5](#_Toc119395224)

[2 Data analysis 6](#_Toc119395225)

[2.1 Data 6](#_Toc119395226)

[2.1.1. 6](#_Toc119395233)

[2.1.2. Dump arrivals logs 7](#_Toc119395234)

[2.1.3. Skips filling data 7](#_Toc119395235)

[2.2 Methods 9](#_Toc119395236)

[2.1.4. Filling rates 9](#_Toc119395237)

[2.1.5. Dump logs 11](#_Toc119395238)

[2.3 Results 11](#_Toc119395239)

[2.1.6. Filling rates 11](#_Toc119395240)

[2.1.7. Dump logs 14](#_Toc119395241)

[2.1.8. 16](#_Toc119395242)

[3 Optimization model 20](#_Toc119395243)

[3.1 Methods 20](#_Toc119395244)

[3.2 Results 22](#_Toc119395245)

[4 Conclusions 22](#_Toc119395246)

[4.1 Capital expenditure 22](#_Toc119395247)

[4.2 Operational planning 22](#_Toc119395248)

[5 Recommendations for future work 22](#_Toc119395249)

[5.1 Skip-level data 22](#_Toc119395250)

[5.2 Current organizational and operational state 22](#_Toc119395251)

[6 References 23](#_Toc119395252)

# 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.

## Background topic 1 (Solid Waste Management)

## Background topic 2 (SWM in Africa)

Maybe now you dive into the global differences in collection

## 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?

# Data analysis

## Data

In order to formulate feasible and pertinent recommendations, parameters reflecting the situation need to be calculated.

The data analysis is based on three datasets. The first is a timeseries of specific skips’ levels over a certain period of time. 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 inorganic waste landfill 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.

Each skip is 7m3. All of the 53 skips are assumed to be mixed waste (organic and inorganic) except for the ones described in 2.1.2. The skips, the municipal dump (Mzedi dump), the truck storage facility and the compost facility are mapped in Figure 1.

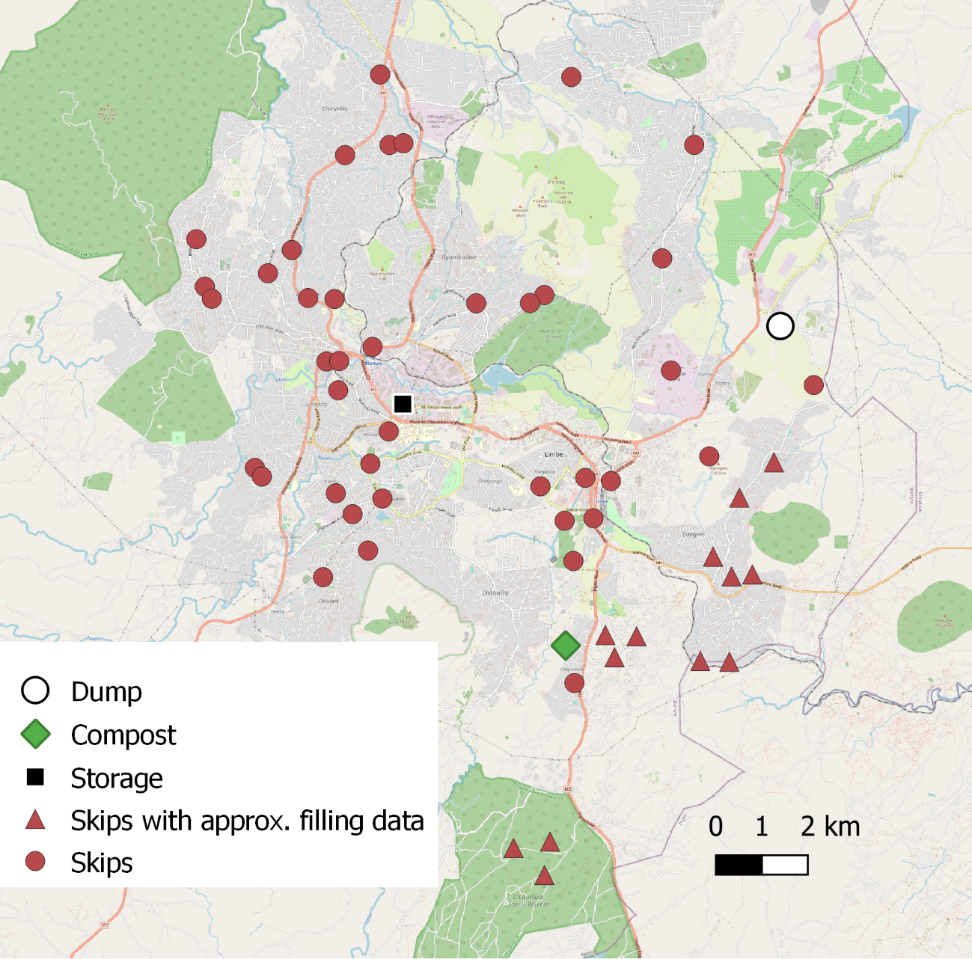


Figure 1 Locations of skips, dump, compost facility and truck storage in Blantyre

### Dump arrivals logs

This is a log lists arrivals at the Mzedi dump, along with the origin of the skip carried by each truck. The origins match exactly the geographical locations in Figure 1. The period of this series is 2020-12-05 to 2021-12-31.

### Skips filling data

Filling data for several skips is provided. Which skips from Figure 1 they exactly refer to is ambiguous, but the skips within the area which could match are annotated as “Skips with approx.. filling data”. The areas are *Bangwe*, *BCA*, *Chigumula* and *Naizi*.

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 are shown inFigure 2**-**4.

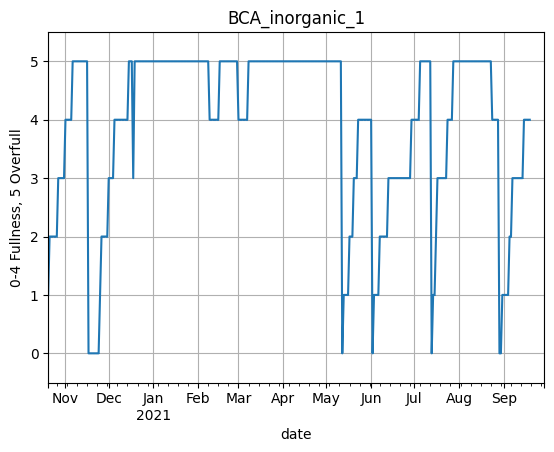


Figure 2 Skip filling data for “BCA inorganic 1”



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 and has to be dismissed. Finally, the data from *“Bangwe inorganic 1”* indicates that it fills up extremely fast, sometimes within a day.

## Methods

### Filling rates

Filling rates are extracted from the skip-level data. The goal is to identify a rising trend over several days, and to compute the filling rate as the slope of this ramp. It performs the following steps:

1. Removing spikes

Spikes occur in the dataset frequently. These are points where the level quickly drops or increases before going back to its original level. Those are assumed to be measurement errors. Figure 5 shows the action of the filter for “*Chigumula\_inorganic\_1*”. The limitations of this method is that quick filling might be dismissed as a spike, effectively putting a lower bound of 0.5 skip/day filling rate.

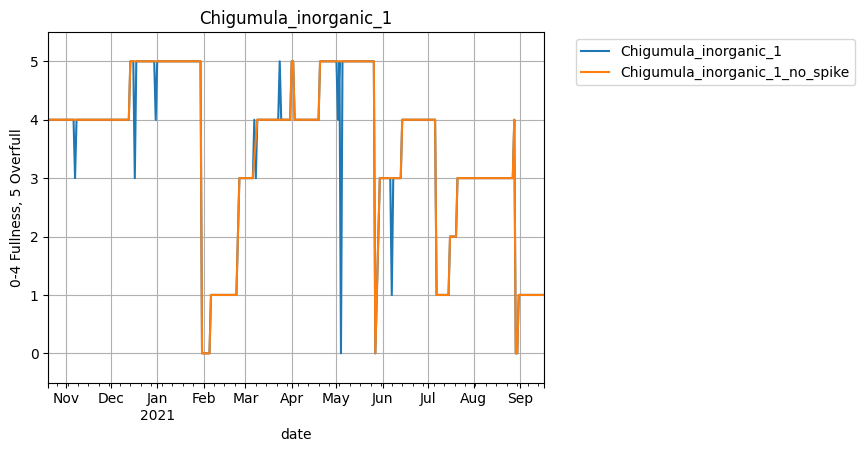


Figure 5 Action of filtering spikes

1. Detecting the top of ramps

The top of the ramp is 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 (e.g. a direct transition between 3 and 5), the top value will be the 5.

A backward pass detects the first date at which the top value appears. These are considered to be the time where the skip achieves the fullness level at which it is emptied, or when it reaches full capacity. Finding this date is particularly important when the skip stays overfull for long periods of time. Using the top ends to calculate the filling rate would bias it to a lower rate.

1. Eliminating downgrading

Downrating is clear in Figure 6 (a zoomed in version of Figure 2), on two occasions between February and April. These events 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 the 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 the current top end if the level rises.

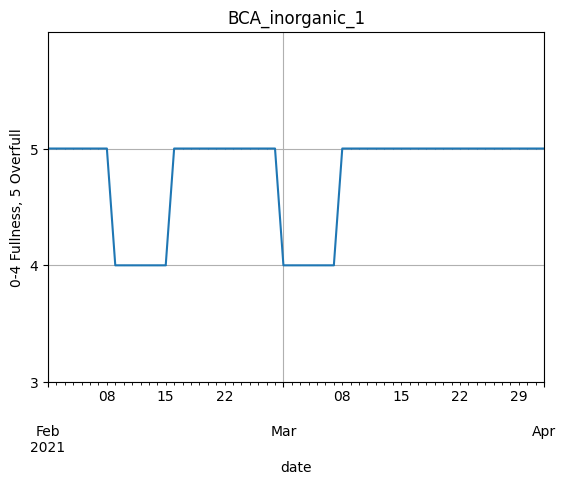


Figure 6 Zoomed in "BCA\_inorganic\_1" filling data

1. Detecting the bottom of ramps

Finally, the bottom values and dates are extracted 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.

### Dump logs

As previously mentioned, dump logs are useful in characterizing the current operation

## Results

### Filling rates

The result is shown in Figure 7 for BCA inorganic 1 and in Figure 8 for Bangwe inorganic 1. Clearly, the fit in the second case is less ideal. In this case, it is more accurate to not to remove spikes, leading to the ramps shown in Figure 9.



Figure 7 Results of ramps for “BCA inorganic 1”



Figure 8 Results of ramps for “Bangwe\_inorganic\_1” **with** spikes removed

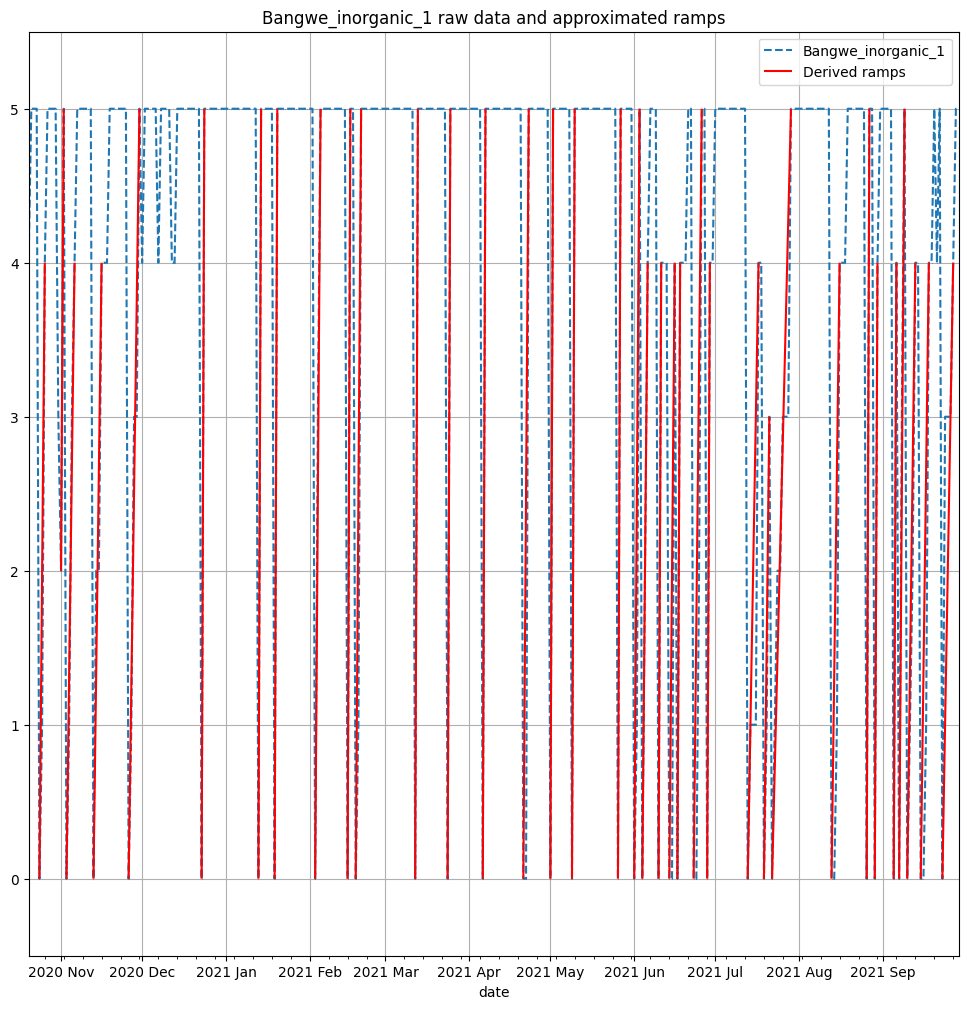


Figure 9 Results of ramps for “Bangwe\_inorganic\_1” **without** spikes removed

The average slopes of the ramps are calculated for each skip in the dataset for the case where spikes are removed and where they are not. This corresponds to the minimum and maximum average filling rates respectively in Table 1. For skips such as “*Bangwe\_inorganic\_1”*, there is 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 9, removing the spikes produces accurate ramps. The process is repeated for each skip based on the visual fit. In Table 1, the filling rates determined to be more accurate is highlighted.

Table 1 Aggregate data from ramps analysis. Highlighted filling rates are selected to be the most accurate based on visual fit of the ramps. # indicates the number of ramps fitted. Prop. overfull is the proportion of days where the level is 5. The period is the period over which levels were measured

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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 | - | - | - | - | - | - |

Figure 10 is a histogram of the filling rates. These samples indicate a possible log-normal distribution of filling rates. However, this is based on a small geographical area and only 12 sample skips.

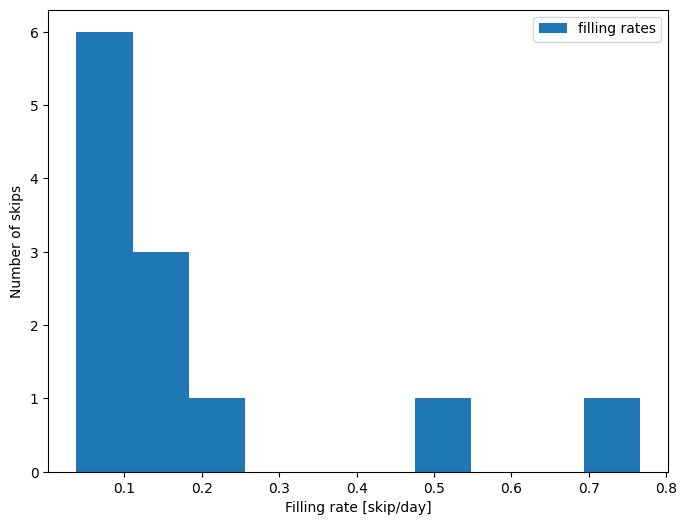


Figure 10 Histogram of extracted filling rates

### Dump logs

The sum of arrivals in each week is shown in Figure 11**.** 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 12.

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 13. 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 11 Sum of deliveries (arrivals) at Mzedi dump over the entire period of measurements



Figure 12 Sum of deliveries to Mzedi dump per weekday over the entire period of measurements

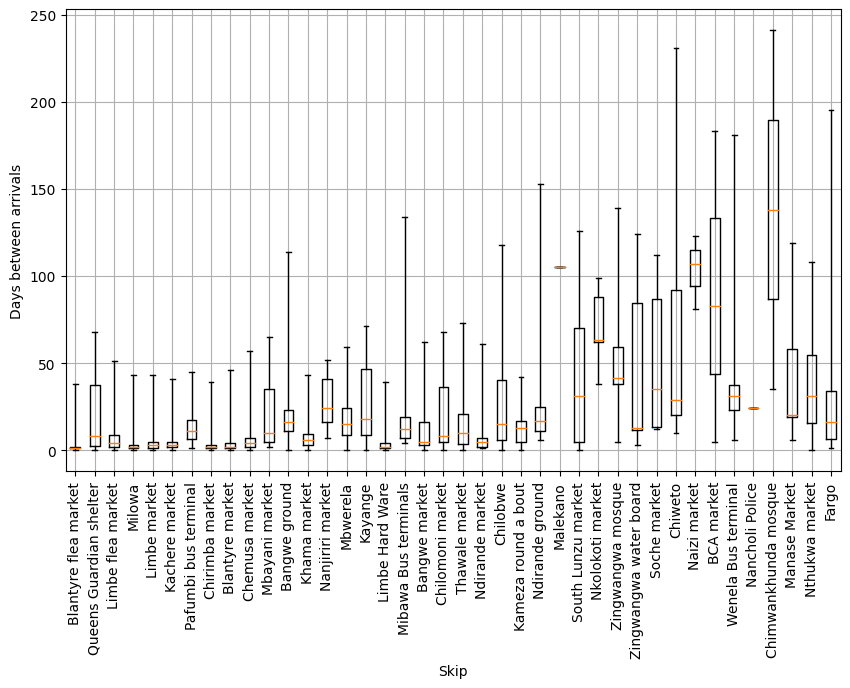


Figure 13 Box plots of number of days between arrivals for each skip. The boxes represent the range between the quantiles Q1 (0.25) and Q3 (0.75). The whiskers extend to the minimum and maximum gaps.

### Extrapolating filling rates

As previously mentioned, 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 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. The organic skips are not considered since they should not go to the dump.

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\_inorganic\_1  Bangwe\_inorganic\_2 | 23  16 | Bangwe ground  Bangwe market | 14  31 |
| BCA | BCA\_inorganic\_1  BCA\_inorganic\_2 | 6  1 | BCA market | 4 |
| Naizi | Naizi\_inorganic\_1 | 15 | Naizi market | 4 |
| Chigumula | Chigumula\_inorganic\_1 | 4 | none |  |

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 2.1.2) 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 14 contradicts this notion, however. It shows many skip emptying events do not match with arrivals at the dump. In Figure 15, 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%.

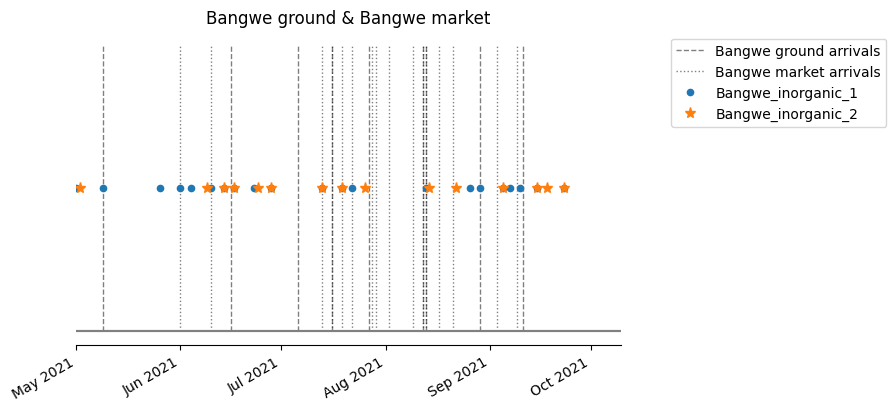


Figure 14 Timeline of Bangwe arrivals at Mzedi dump and emptying events of select skips

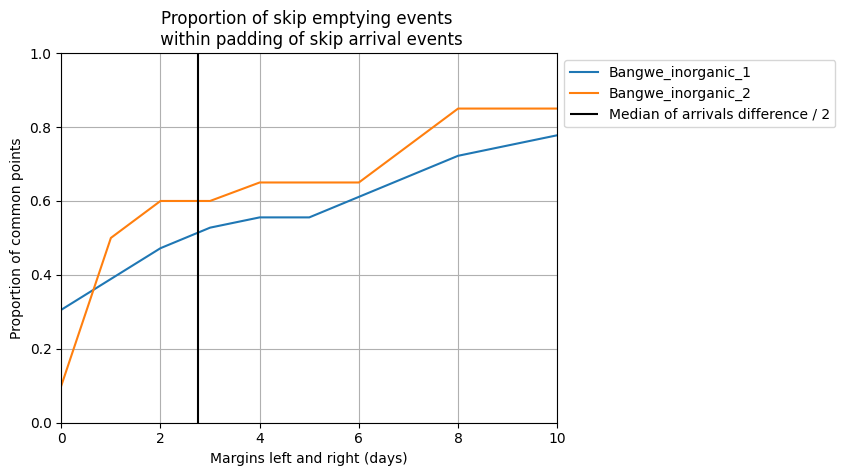


Figure 15 Proportion of skip emptying events matching arrivals at the dump

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 |

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 16. Interestingly, the reciprocal of the median emptying frequency exhibits a similar behaviour, as seen in

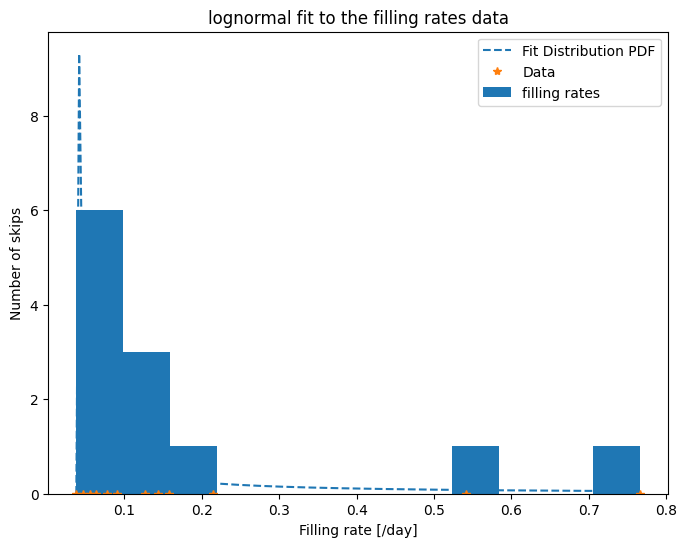


Figure 16 Lognormal fit of the filling rates skip level data

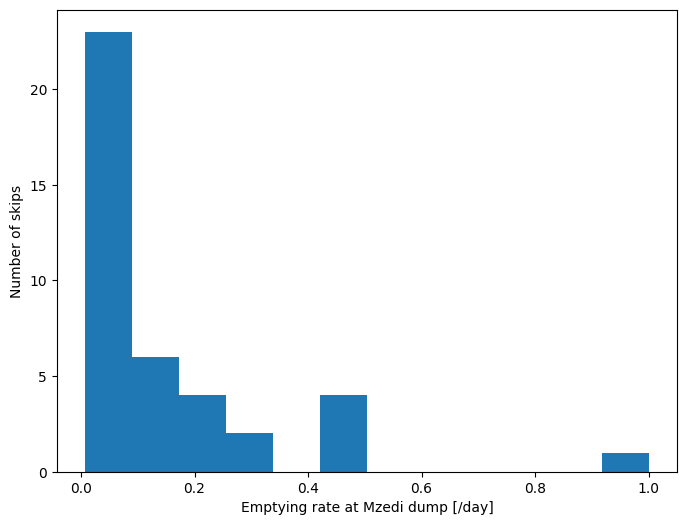


Figure 17 Histogram of emptying rates by skip arrivals at Mzedi dump

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.

# Optimization model

The problem is first formulated as a series of features, assumptions, objective and constraints.

## Methods

* Features:
  + **28 days planning period**
  + Morning and evening collection
    - The two useful cases are:
      * Afternoon collection on Saturday and morning collection on Monday
      * Multiple collections per day
    - Heuristically assign sets of schedules to skips and add skips
      * Assume that a double collection is always done on the same day for both skips in an area
    - Both cases are specific
  + **No emptying on Sunday (and Saturday)**
* 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)
  + Cost constraints
  + **Daily operation constraints (max 8h of travelling 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 18 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 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 still 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.

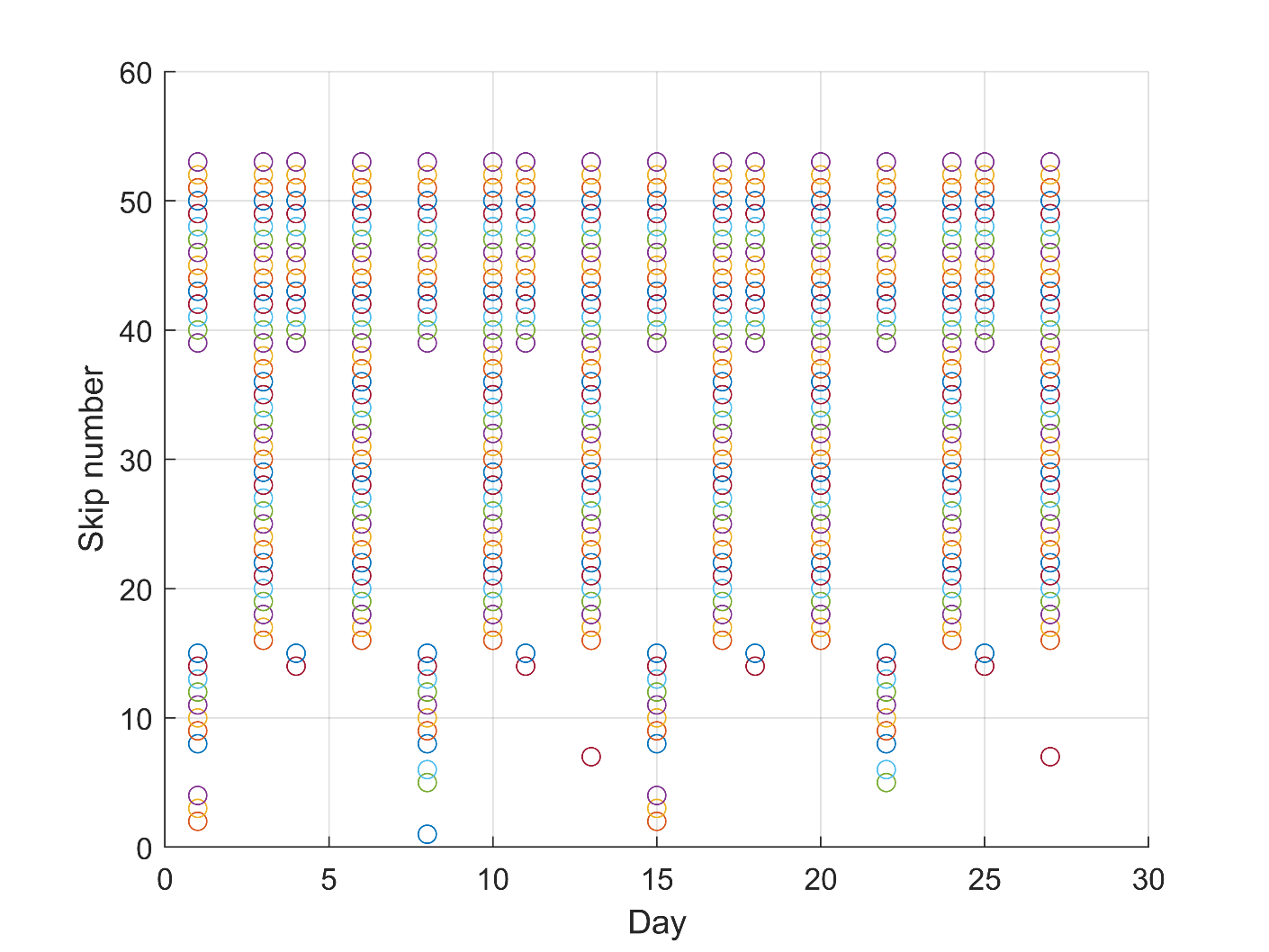


Figure 18 Schedule diagram (first version, more work will be done on operational planning representation)

## Results

# Conclusions

## Capital expenditure

## Operational planning

# Recommendations for future work

## 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.

## Current organizational and operational state

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

# References