Digital Transport for Africa: Advocating for Maputo's Transit Digitalization

Final Report

Final Report v1

Date: February 2023

Produced by:

Transport for Cairo
مواصلة للقاهرة

For:

World Resources Institute

Digital Transport for Africa

Title picture: Ponto Final Cross Bus Stop, Chapas Project

Table of Contents

List of Figures

List of Tables

Glossary

Acronym	Term
GTFS	General Transit Feed Specification
AFD	French Development Agency
WRI	World Resource Institute
TfC	Transport for Cairo
OMT	Observatório da Mobilidade
AMT	Agência Metropolitana de Transporte De Maputo
REM	Rede Estrutural da Área Metropolitana
DT4A	Digital Transport for Africa
MMA	Maputo Metropolitan Area

OSM	OpenStreetMap
GIS	Geographic Information Systems
PT	Public Transport
BRT	Bus Rapid Transit
GPS	Global Positioning System
GIS2GTFS	Geographic Information Systems to General Transit Feed Specification
FR	Field Research

Executive Summary

This report explains the technical support provided by Transport for Cairo to Mobility and Transport Observatory in Maputo City as part of the Digital Transport for Africa (DT4A) Innovation Challenge. The project is led by the World Resource Institute (WRI). The report provides background information on the current situation of paratransit and transit data in Maputo, an assessment of OMT's existing data and mapping methodology, and the support activities done.

TfC's assessment of existing data shared by OMT concluded that there is room for improving the mapping methodology used at OMT to produce both spatial and temporal data. This can enable more analyses than the existing data could. Moreover, OMT, having an upcoming mapping project, expressed strong interest in knowledge transfer of TfC's mapping experience.

Two main activities were conducted:

1. Demonstration of a spatial analysis using OMT's existing paratransit data to engage the local transport authority and raise interest in paratransit mapping.
2. Capacity Building for OMT and its new mapping team to implement the improved data collection method.

Activity (1) showed that almost 50% of existing Chapas routes overlap geographically with proposed BRT corridors in a significant way, deeming them as competitive and liable for relocation. The analysis was based on available OSM data from previous mapping efforts and wasn't meant to influence actual BRT planning rather than advocating the power of digital transport data and it's potential.

Activity (2) comprised multiple sub-activities:

a. Producing technical documentation of the improved methodology and step-by-step guides for mappers and data processing personnel on the open tools used for implementation.
b. Developing a graphical software tool that implements the complex data transformations done by TfC. This was needed since no current member of OMT had coding experience.
c. A 2-day virtual training organized with OMT. The training went over all the conceptual and practical aspects of the improved methodology. Participants from OMT and other local stakeholders attended the training.

The support activities have provided OMT with better tools to map the paratransit network and produce insights from the data. With the Maputo Metropolitan Area (MMA) going through the process of professionalizing its transit network and planning a BRT system, there is strong interest from local stakeholders to ensure a successful transition, by better understanding of the existing paratransit network through data.

OMT will put the outcomes of the capacity building and associated tools to implement its upcoming mapping project in the Matola suburb in MMA.

1 Introduction

1.1 Stakeholders

- **French Development Agency (AFD):** is a public financial institution that implements the policy defined by the French Government. It works to promote sustainable development and fight poverty. AFD Group finances and supports development projects in; energy, urban development, education, climate, agriculture, health, water and sanitation, digital technologies, biodiversity and sport. AFD is one of the main sponsors of the DT4A intiative.
- **World Resource Institute (WRI):** is a global research non-profit organization that works on tackling multiple global challenges including mobility in cities. WRI-Africa is currently leading the DT4A project implementation and have hired Transport for Cairo as technical support consultant
- **Transport for Cairo (TfC):** TfC provides data, tools, and research to improve urban mobility in emerging cities, primarily in Africa. They are the technical support consultant responsible of the project implementation. Namely to undertake mapping and going beyond mapping projects in Alexandria, Egypt; Kumasi, Ghana and Maputo, Mozambique
- **Mobility and Transport Observatory (OMT):** The Mobility and Transport Observatory (OMT) is a monitoring and research association in the various sectors of transport with special focus on the field of Urban Mobility. It's an initiative of the local Think-and Do Tank WAZA. OMT creates knowledge in the form of research papers and digital data. They have already mapped Maputo city's Chapas network and look to map more cities in the Maputo Metropolitan Area (MMA)
- **Agência Metropolitana de Transporte De Maputo (AMT):** Maputo Metropolitan Transport Agency was created by the government of Mozambique in 2017 as a public institution with administrative and financial autonomy, aiming to plan and manage the integrated transport system in the municipalities of Maputo, Matola and Boane, and Marracuene in Maputo province.

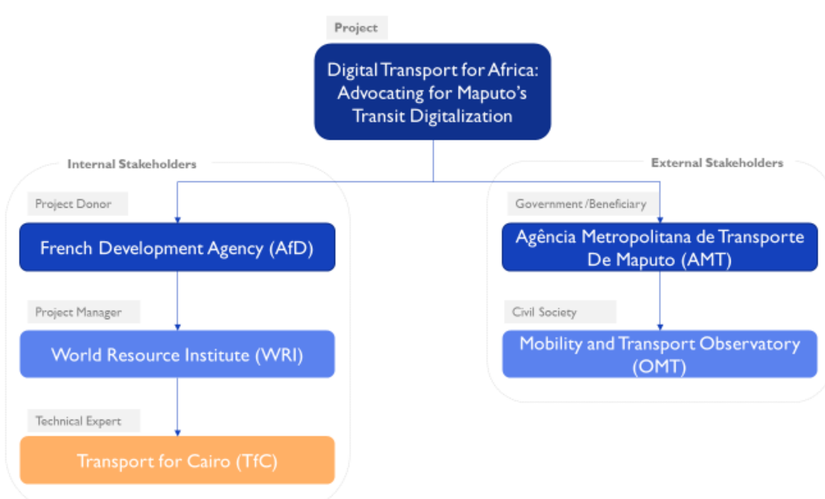


Figure 1: Project Stakeholders

1.2 Background & Objectives

Acting on a comprehensive urban transport master plan done by JICA in 2015, the city of Maputo has been taking steps to mitigating existing urban transport issues regarding road infrastructure, traffic management, traffic safety, and public transport.

As with other cities in Africa and the global south, Maputo is aiming to transition from an ungoverned low-capacity paratransit network to mass transit modes and professionalization of existing services. This shift is currently coordinated by the Maputo's Metropolitan Transport Agency (AMT) which oversees and coordinates the implementation of the Metropolitan Structural Network (REM), a network of buses operated by 5 public companies, 10 cooperatives, and more recently three private companies in 2020[1].

The REM network covers 7 corridors with 94 routes and is looking to expand to include all MMA's corridors. The world bank announced last year a 250 million USD grant to establish Maputo's first BRT system, accelerating the shift from existing paratransit services to mass transit modes[2].

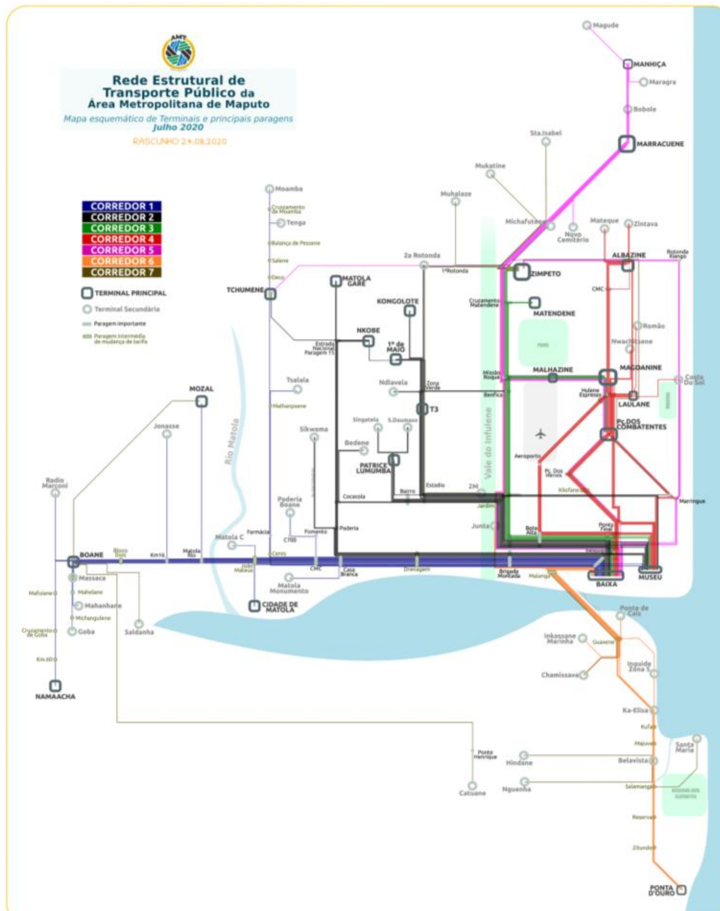


Figure 2: Metropolitan Structural Network: AMT (2020)

This rapid change prompts the need for comprehensive and accurate data to inform decision makers working on situational analysis. This is to accommodate the needs of passengers (demand) and current operators (supply) as well as other stakeholders whose livelihoods will be affected by the professionalization of the paratransit industry in the city.

One important aspect of this analysis is the operational data for the services currently running in the MMA, especially those who are by default "invisible" to authorities. This is where OMT's transport mapping initiative becomes an important tool in filling this knowledge gap.

[1] Romero, J., Mazzolini, A., Machanguana, C., Matos, A., Cavoli, C., Oviedo, D., Macamo, G., (2022) "Perfil Urbano de Maputo. Mobility, accessibility and land use in the Maputo Metropolitan Area". T-SUM project. UCL

[2] https://www.worldbank.org/en/news/press-release/2022/08/24/world-bank-supports-urban-transport-in-the-maputo-metropolitan-area

Thus, the objectives of the assignment were defined as follows:

1. Assessment of **existing digital data**: on public transport in Maputo
2. Assessment of **local capacity**: for mapping & data processing
3. Formulation of **recommendations**: to leverage local mapping capacity
4. Transfer **knowledge** based on recommendations: through documentation of learning materials & conducting training sessions with local mappers

2 Current Situation

2.1 Transport in Maputo

There are 3 main public bus transport modes in Maputo:

1. **REM buses**: operating on 7 corridors with both varying sizes of passenger buses
2. **Chapas**: the main paratransit mode operating with 14-seater buses and with the highest share of daily trips
3. **Myloves**: Pick-up trucks that load passengers and merchandise in areas with low road quality

(a) Chapa

(b) Mylove

(c) REM Buses

Figure 3: Public bus modes in Maputo, source: O País

In terms of modal split, as per 2012 estimates from the JICA 2015 masterplan, Chapas had the highest share at 32.9% followed by what is classified as “large buses” at 9.2%. Assuming the latter refers to the REM network, it is safe to assume that since then, the percentage of REM might have increased due to increasing fleet size and AMT influence.

2.2 OMT’s Digital Mapping & Available Data

Amongst their activities to promote sustainable urban mobility, the OMT team have undertaken the task of consolidating standardized digital data for public transport in the MMA, namely for chapas, formalized buses, and rail. For the chapas, field surveys are necessary to actualize data collection since there is no secondary sources such as sensor or physical data available.

OMT has completed a field research exercise in 2017 to digitize the chapas services in the city of Maputo. This leaves the districts of Matola, Boane, and Marracuene. OMT plans to create a comprehensive database and GTFS feed for the entirety of MMA.

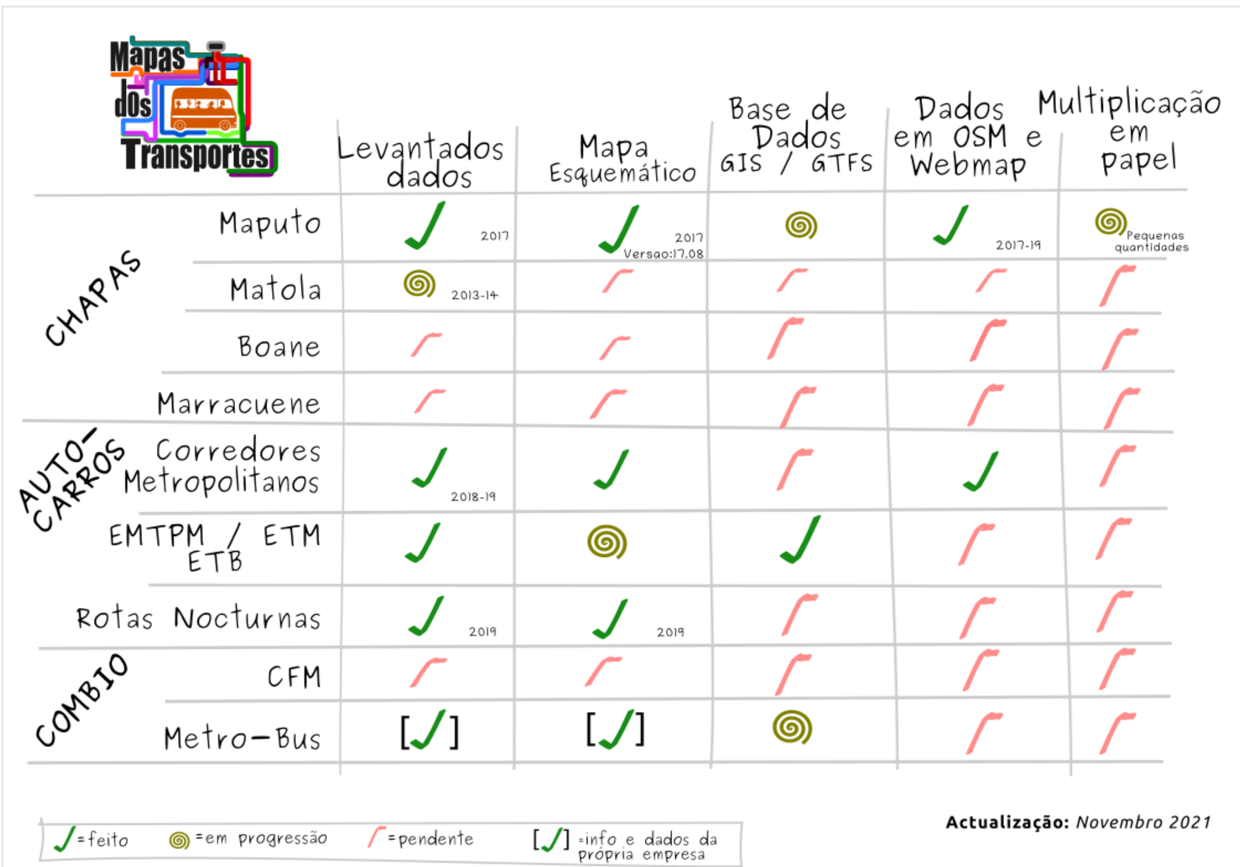


Figure 4: OMT MMA Mapping Progress as of November 2021

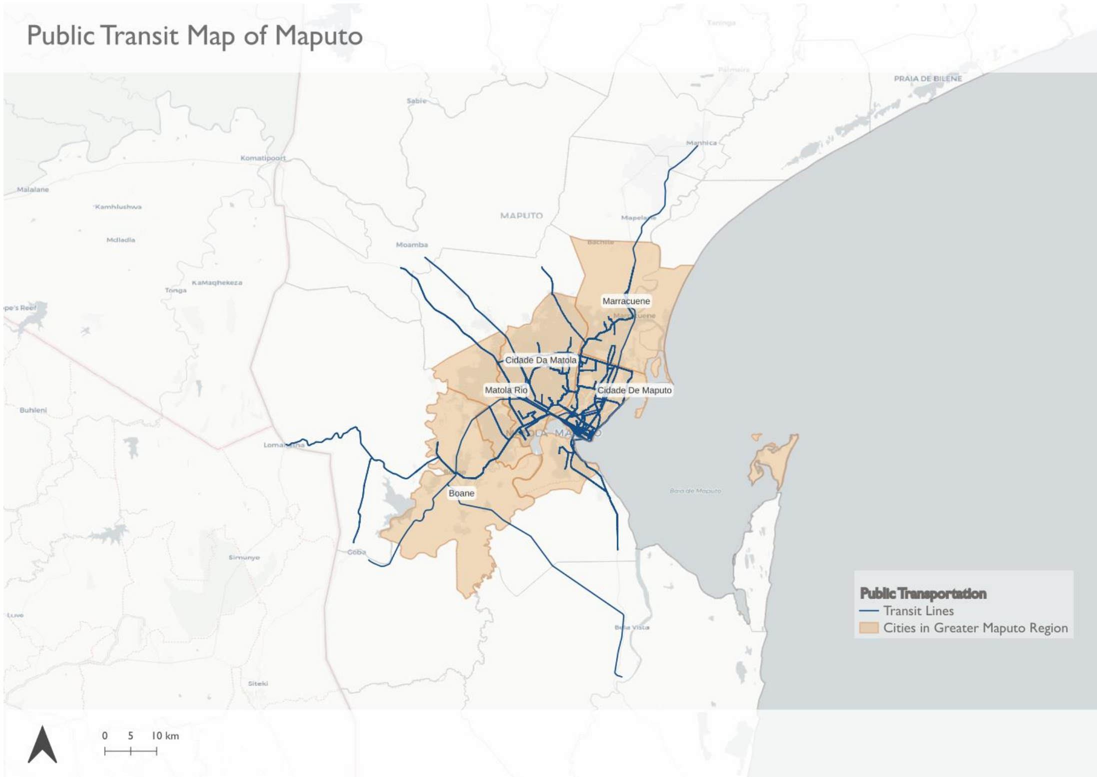


Figure 5: Transit lines data previously collected by OMT

2.3 OMT's Upcoming Project: Mapping Matola

OMT's upcoming mapping project, taking place mid-2023, covers Matola. Matola is the largest suburb of the MMA. With a population reaching 1,032,197 inhabitants[3], it is one of the most populated cities in Mozambique. Mapping of Matola is a continuation of OMT's mapping efforts aiming to cover the whole of MMA.

OMT are working closely with AMT to accomplish the mapping, presenting a unique opportunity for having the public sector onboard with the mapping process and later on using the data outputs in decision making, especially in planning for BRT implementation.

Given the convenient timing of this mapping project, TfC team decided to design the local capacity building and knowledge transfer activities closely with OMT as a training for mappers and data processors who are going to be working on the Matola mapping project. Thus, the mapping of Matola is expected to be the first real-world application of the set of recommendations provided by TfC.

3 Assessment

To design an effective intervention, TfC team needed to assess the existing local capacity for mapping. The assessment included existing data, processes, software tools used, and human resource capacity.

[3] Census, 2017, https://www.citypopulation.de/en/mozambique/admin/

3.1 Data Assessment Framework: Maturity Model

To properly assess data quality and comprehensiveness for different cities and in different contexts, TfC developed the "Data Maturity" model. The model is the framework used to:

1. Act as a benchmark for existing dataset outputs of previous mapping projects
2. Act as a reference for data outputs throughout the project's activities, for consistency

We define the data maturity model as following three pillars of data availability:

1. Geographic Data
2. Temporal Data
3. System Adequacy Data

OMT shared the details of their mapping process in previous projects, as well as the raw data collected and GTFS created from it. The data was assessed against the Data Maturity Model.

Table 1: TfC's Data Maturity Model

Code	Levels (Sub-categories)	Metadata	In OMT Data?
Geographic Mapping			
1.1 **BASIC**	Trip GPS-trace	- Route Number / Name - Trip Head-sign	Yes
	Trip Stops (i.e. **where** did the bus stop during the trip)	- Stop Location	Yes
1.2 **INTRM.**	Trip GPS-trace with basic temporal data	- Bus Type & Capacity - Route Description - Trip Fares - Trip Fare Variability 1.0 (YES/NO)	Yes
	Trip Stops (i.e. **when** did the bus stop during the trip)	- Stop Name - Stop Type (Planned or Actual or Both) - Trip Duration (between stops) - Dwell Time (at stops)	Yes
1.3 **ADV**	Multiple Trip GPS-Traces for same route with detailed temporal data	- Route Deviation (Alternative Routes due to congestion, etc.) (YES/No)	No

Temporal Mapping			
2.1 **BASIC**	Baseline Headway	- Frequency of service (Headway per route) - General Operating Time	No
2.2 **INTRM.**	Headway Changes	- Operating Time Variations (Seasonal Schedules, …etc.) - Availability of Passenger Information System 2.0 (in Stops)	No
2.3 **ADV**		- Frequency Variation per stop per Vehicle (x Vehicles "same route number or name" / stop / hour)	No
System Adequacy			
3.1 **BASIC**	End of Trip Survey	- Availability of Safety Equipment and measures (Door closing, Standing Bar … etc.) - Availability of Passenger Information System 1.0 (in Vehicles) - Emergency Preparedness (Vehicle) - Wheelchair Accessibility	No
3.2 **INTRM.**		- Crowding 1.0 (General Score on the trip level) - Use of Renewable or Clean Energy - Comfort Measures (AC availability, comfortable seating … etc.) - Stop Amenities (Availability and Quality) - Emergency Preparedness (in Stops)	No
3.3 **ADV**	General	- Crowding 2.0 (Per Stop to measure passenger flow in and out of the vehicle) - Safety and Walkability of Stop's catchment area (Built Environment Infrastructure Data) - Trip Fare Variability (Value by distance) - Energy Efficiency - Noise Pollution Level	No

A comprehensive data assessment exercise concluded OMT's data was based on intermediate geographic mapping efforts, while little to no surveys were done to capture temporal and system adequacy data. This was a good starting point to shift focus onto temporal mapping and the optimization of existing geographic mapping processes.

3.2 Software Tools

This section describes the software tools used by OMT for mapping and provide details on their strength points and limitations.

3.2.1 OSM Tracker

OSM Tracker is a free and open-source android app designed to record GPX tracks and mark points of interest along that track. Its data model (GPX tracks with waypoints) is made to match the format which can be directly uploaded to OpenStreetMap, hence the name of the app.

The app supports the following features:

- Works offline.
- Sharing collected traces can be through:
    - uploading to OSM
    - exporting to GPX files
    - uploading GPX files to an online drive.
- The setting for GPS logging interval or distance, to balance between accuracy and battery efficiency can be changed.
- During the track you can take a photo, record a voice note, mark a point of interest with an OSM-defined or custom class (Bus stop), or take a note.

Given the tight integration between the app and OSM, it has some limitations and points of friction when applied to collecting transit data within the Data Maturity model in Table 1: TfC's Data Maturity Model. They can be summarized as follows:

- GPX Tracks have, by default, are exported with a date-time stamp as their name, which can later be changed.
- No way to define custom fields to input structured data in waypoints, even though the GPX specification supports custom field[4]s. OSMTracker supports "text notes" which provides a free-form text input.

3.2.2 TransitWand

OMT staff informed TfC team that they have been given access to "TransitWand", a transit mapping software tool, by Transitec, a French consultancy who are working with OMT on a SUMP for the

[4] https://www.topografix.com/GPX/1/1/#type_extensionsType

MMA[5].OMT did not have the technical capacity to use "TransitWand" at the time of the assignment, however.

TfC team attempted looking for documentation or user guides for the software but couldn't find any online, especially since the software codebase has been unmaintained for the past 8 years[6].

3.2.3 Counting App for Frequency Surveys

OMT does not currently conduct frequency surveys for buses. Frequency surveys are important to capture temporal data on the level of service for bus routes. By using an off-the-shelf "counting" app, OMT would be able to get headway figures for bus services by counting vehicles from their point of origin, the bus route terminal.

We provide in section 4 an approach to integrate this survey and subsequently the frequencies data output, into OMT's mapping process.

4 Recommendations & Technical Support

The assessment done by TfC for OMT's previous mapping projects gave insights on how to design the most suitable package of recommendations and support tools to achieve a better mapping process. The following design strategies were followed:

1. **Development approach**: After consultations with OMT, TfC decided to adopt a "development" approach to the technical support. Rather than proposing a mapping methodology that was applied previously by TfC, the team made sure to build on existing knowledge and previous efforts made by OMT.
2. **More data**: Capturing more data points by adding more information to capture in the field, while making use of every data point captured automatically by mapping apps (i.e., timestamps of GPS points can be used to create speed profiles)
3. **Standardized data pipeline**: To ensure data sustainability, data must be collected, stored and managed in an organized and rule-oriented database. To achieve that, a data pipeline was designed and documented, and supplementary data loading and transformation software tools were developed in the form of a plugin for QGIS, an open-source desktop GIS platform. The plugin was named "Paratransit Mapping Methodology Toolkit" or "PMM Toolkit".
4. **Local Capacity**: Stakeholders involved in mapping were given training workshops and capacity building sessions to embrace the updated mapping methodology and gain the ability to apply it. Extra focus was given to documentation of knowledge materials (step-by-step tutorials, videos, lecture-format slides, etc.)

The technical support provided by TfC to OMT comprised the following:

[5] https://transitec.net/fr/actualites/item/11029-launch-of-the-sustainable-urban-mobility-plan-in-greater-maputo.html
[6] https://github.com/conveyal/transit-wand

1. Comprehensive documentation of "Paratransit Mapping Methodology" (PMM) developed by TfC, covering all the conceptual and practical aspects of getting from no data to detailed GTFS and analysis outputs.
2. Software tools needed to implement PMM:
    a. TfC developed a QGIS plugin that implements the complex data loading and transformation procedures within PMM.
    b. Route Snapper, a desktop app previously used by TfC for snapping GPS traces to road network, was shared with OMT.
3. A 2-day training organized with OMT to raise the capacity of the local mapping team to implement PMM.

4.1 Paratransit Mapping Methodology

The first part of the TfC's support to OMT was creating comprehensive documentation of the mapping method that TfC has developed. The PMM process from start to finish is outlined by the diagram in Appendix A: Paratransit Mapping Methodology Flowchart. The process includes three main phases, three types of surveys and multiple data transformations, as well as manual and automated processing and cleaning steps.

The following sections summarize the PMM method and its implementation using open tools. Further implementation details and tutorials are found in the documentation shared with OMT.

The overview flowchart in Appendix A: Paratransit Mapping Methodology Flowchart shows three main phases in the mapping process: identification, data collection and data processing. The last block shows examples of the applications that the produced datasets in previous phases enable, such as creating GTFS feed.

4.1.1 Identification Phase

The methodology was designed with the assumption of little to no prior knowledge about the existing transit network. In such a case, the identification survey is used to discover the routes and terminals in the network.

Terminals, in PMM, are defined as transit hubs where a transit route starts or ends. Figure 6 shows an example of a terminal in Nkobe, Maputo. It demonstrates common characteristics of paratransit terminals. Different vehicles serving different routes line up and are dispatched with in a demand-based manner, i.e. vehicles are dispatched as soon as they are full with passengers.

Figure 6: Satellite Image (Google Maps) of Nkobe Terminal in Nkobe, Maputo.

A route is defined by multiple attributes: its origin and destination terminals, its operating agency, and the type of vehicle. While the notion of a route may differ slightly by local context, this strong definition within PMM is necessary for streamlining data transformations and analyses.

In the case that the terminals and routes to be mapped are known to the mapping team, the identification survey can be substituted with desk research or skipped.

The outputs of the identification phase are datasets of terminals, routes, agencies, and vehicle types that exist in the network to be mapped.

4.1.2 Data Collection

From the routes identified in the identification phase, a subset is selected for mapping based on constraints such as the project's budget, time frame and team size and efficiency.

TfC provided OMT with its "FR Parameters" tool, an excel-based tool used for planning a mapping project based on the constraints defined above, estimating the time and budget required.

Afterwards, mobilization of the data collection team commences, conducting the remaining two types of surveys: onboard surveys and frequency surveys. A field research manager (FRM) assigns surveys to field researchers (FRs), communicates directly with FRs and monitors the data collection progress. Data is received from the field as soon as a survey is finished and validated to mitigate errors promptly.

4.1.2.1 Onboard Surveys

Onboard surveys capture data points related to the trip's itinerary and stops. The field research boards the vehicle from its origin terminal and logs the GPS trace of the full trip to the destination terminal.

Along the way, they record locations of stops and how many passengers boarded and alighted at each stop.

4.1.2.2 Frequency Surveys

Frequency surveys attempt to capture the average headway for a route by observing the departure times of vehicles for said route for a long enough period (usually 45 min).

4.1.2.3 Understanding The Surveying Approach

Paratransit routes generally do not adhere to a consistent headway (demand-based dispatch) or stop locations (boarding and drop-off passengers anywhere). However, the raw data points captured in onboard and frequency surveys, can surface insightful patterns. For example, the raw stop locations from many routes on a main corridor show hotspots of passenger activity. The average headway observed from multiple frequency surveys of the same route enables comparing differences in demand between routes and terminals.

Hence, it is recommended that each route is surveyed more than once at different times of the day (peak-demand, off-peak, etc.) to increase the confidence in the insights created from the data.

In summary, the surveys are designed to maximize the data points observed from the field. Those data points are aggregated to extract network-level and route-level insights.

4.1.3 Data Processing

In this phase, the raw data from onboard and frequency surveys goes through both manual and automated processing steps that clean and aggregate it into the following outputs:

- GIS layer of trip itineraries
- GIS layer of main stops in the network
- Headway per route
- Speed profiles for road network segments

The former two outputs are created manually with the aid of the GIS raw data collected. The raw stops are consolidated into a smaller set of stops that represent a virtual network covering the mapping area. They are created with the aid of the raw data and experience with local transit. Raw trips, on the other hand, are easier to manually process; the raw GPS traces are "snapped" to the road network (extracted from OpenStreetMap using tools like HOT Export)

Figure 7: Raw stops (green) and digitized/processed stops (red) from a mapping project in Alexandria, Egypt.

The latter two outputs are produced using Extract-Transfrom-Load (ETL) scripts. Headways are calculated as averages from the raw data, while speed profiles rely on spatially relating the stops with the time-stamped GPS trackpoints to get average speeds along road-network segments.

These output datasets enable a variety of applications, such as creating a time-aware GTFS feed, or visualizing passenger activity levels across the network, or passenger flow through the network. Example visualizations are shown in Figure 8.

The data processing activities need to be performed technically oriented team members. They require experience in QGIS and understanding of data management concepts, such as relational joins and spatial joins.

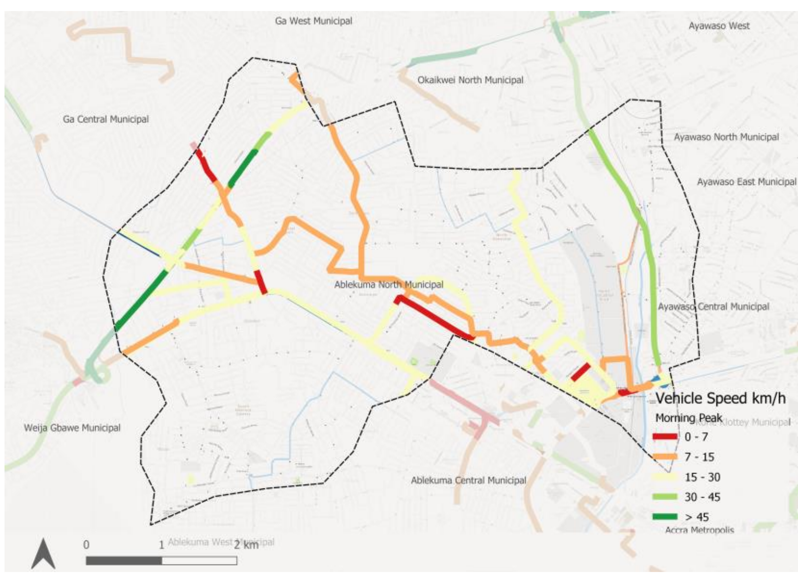


Figure 8: Visualization of road-network speed profiles in Accra, Ghana.

4.1.4 Scaling The Process

Risks in the field and management challenges increase with the scale of the mapping, in terms of team size or number of mapped routes and terminals. Such risks include FRs personal safety, inadequate back-office capacity (FR management & Data processors) against number of FRs, and/or inefficient assignment distribution (i.e., who is best fitted for a survey? Who's living closer to a given bus terminal? etc.).

TfC introduced to OMT the process of creating and utilizing a formal field research protocol (FR Protocol). The FR Protocol is a protocol that details to FRs how to deal with common or risky situations in the field. It also defines the channels of communication between FRs and the FR manager.

Procedures introduced during the training also focused on using the received data to measure FR progress KPIs and individual FRs efficiency. Using off-the-shelf tools such as QGIS or a dashboard-building tool such as Google's "Looker Studio", progress reports can be generated daily. Real-time monitoring of FRs will still be unavailable however, causing a delay of 24 hours at best between errors occurring in the field and the back-office capturing them.

4.1.5 Implementation Tools

TfC's documentation of PMM as well as the training sessions provided to OMT (described in 4.2) explain the implementation of PMM using open tools.

Table 2: PMM Implementation Tools

Tool	Description	Purpose
OSMTracker	Android app developed by OpenStreetMap community	Onboard surveys
Google Drive	Cloud Storage from Google	Cloud storage for sharing onboard surveys
KoboToolbox + KoboCollect App	Open-source surveying platform	Frequency surveys
QGIS	Open-source GIS program	Manipulating and visualizing GIS layers
PMM Toolkit	QGIS Plugin developed by TfC for this project	Used to manage and transform raw data in multiple stages of the mapping process.
Route Snapper	Desktop app developed by TfC in 2018	Route matching (snapping to road network) of the raw trips' GPS traces.
Open Source Routing Machine	Routing Engine	Routing Engine used as a backend for Route Snapper
GTFS Validator		Debugging GTFS feed

Tutorials explaining the implementation details and the tools are in the PMM documentation and the training sessions (links to the recordings can be found in the linked documentation). There are tutorials for the mappers explaining how to use the OSMTracker and KoboCollect mobile apps, and tutorials for the data processing personnel that go step-by-step from identification to the final outputs.

4.2 Capacity Building

Developed mapping methodology must be implemented locally by local mappers in Maputo to ensure the sustainability of the intervention. To that purpose TfC worked closely with OMT to organize a 2-day training for local mapping stakeholders. The training is supported by technical documentation and step-by-step tutorials.

4.2.1 Planning

A 2-day mapping training course was designed to introduce concepts and apply those concepts promptly afterwards. The agenda followed a sequence of steps to do mapping, where the first day focuses on the field mapping and how to plan a mapping project, and the second day focuses on data processing and using the field survey data.

The training was done in a hybrid mode where TfC team members are giving the training from Cairo, Egypt and OMT are attending and facilitating physically on the ground. Almost all attendees were together in one room while one or two additional attendees joined online.

After some drafts and revisions, the agenda for the 2 days was finalized and shared with attendees before the sessions.

Time	Session Title	Objectives
9:00 am - 9:15 am	Round of Introduction	

9:15 am – 10:00 am	#1.1: Survey Plan	• Introduction to the technical support initiative
		  ○ Objectives
		  ○ Methodology
		• Survey Types & Data Outputs
		  ○ Onboard Surveys
		  ○ Frequency Surveys
		• Field Research Parameters

10:00 am – 11:00 am	#1.2: Workshop: Setting up Mapping Tools	
11:00 am – 11:30 am	Tea Break	

11:30 am – 1:00 pm	#1.3: Software & Data	• Data Lifecycle: From field to GTFS
		• Software tools
		• Working with KoboCollect
		• Working with OSMTracker
		  ○ Boarding & Alighting
		  ○ Sharing Data
		  ○ Encoding attributes in file submitted
		• GPX2GIS: Trackpoints, raw stops, and raw trips

Take-home Assignments		• Onboard surveys • Frequency surveys
Day 2		

9:00 am – 10:00 am	#2.1: Working with Data	• Field Research progress monitoring
		• Data Acquisition
		• Data cleaning
		  ○ OSM routing for GPS trace cleaning
		  ○ Processed "virtual" stops
		• GIS to GTFS

10:00 am – 11:30 am	#2.2: Workshop: Data Transformation	• Download raw survey data • Transform data
11:30 am – 12:00 pm	Tea Break	
12:00 – 2:00 PM	#2.2: Workshop: Data Transformation	• Visualize on QGIS • Processed "virtual" stops • Q&A: Feedback on process

OMT were proactive in organizing and planning logistics for the training, giving TfC team more time to focus on developing the content and related material. They leveraged their on-ground presence and existing connections to diversify the attendees which would result in constructive discussions during the training.

About 20 Attendees came from several organizations:

- OMT
- AMT
- Mapeandomeubairro: A local NGO
- Eduardo Mondlane University
- São Tomas Univeristy
- VOID - Tecnologia e Comunicação: A local tech startup

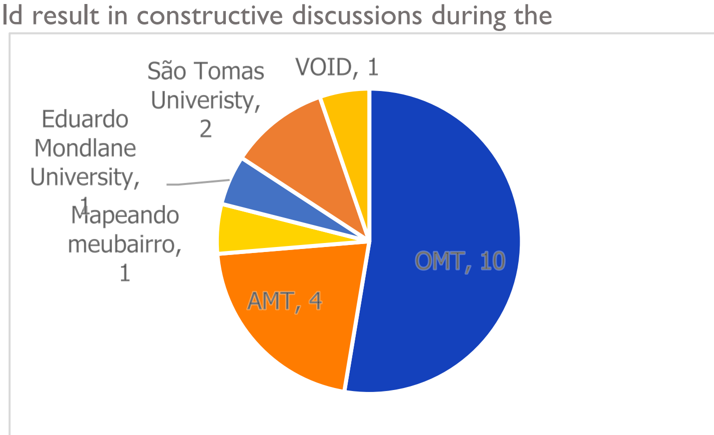


Figure 9: Number of Attendees by Organization

4.2.2 Day 1: Mapping

The day started with a quick round of introductions, then proceeded with the **first session** titled "**Survey Plan**". The session was given in a lecture format with pauses for interpretation in Portuguese by OMT staff for some of the more complicated parts of the session, and questions by attendees.

The first session gave an overview of mapping and field surveys, by first describing the different types of field surveys on public transport and their data outputs. This started questions on the meaning of some of the attributes of the data outputs and their potential usage, it also raised questions on how a survey such as the frequency survey can be "localized" to fit the unique way Chapas operate.

The session then goes into the human resource aspect, the team. It went into detail on team composition, required and preferred profiles to have, the best practices in training. Lastly it looked at mapping from a project management perspective. This meant defining activities, coming up with a timeline for those activities, estimating budget, and accounting for potential risks based on TfC's previous mapping experience.

Right after the first session is the **first workshop** "Setting up Mapping Tools" where the attendees start getting their phones ready to install and configure mobile apps for field surveys, as well as cloud storage to manage and export received surveys from the field. The attendees were following a step-by-step tutorial prepared by TfC, as a group.

The setup process consisted of the following:

a. Set up a Google Drive folder and share with the attendees Gmail accounts so they could later upload their onboard survey GPX files to the shared Drive.
b. Set up a KoboToolbox project and upload the ready-made frequency survey.
c. Install OMSTracker & KoboCollect on the attendees' android phones.
d. Set up KoboCollect apps to connect to the project created on the KoboToolbox web dashboard earlier.
e. Configure OSMTracker GPS interval to 5 seconds to avoid draining the battery and bloated GPX files.

The process went smoothly with minor debugging assistance from OMT's facilitators.

The workshop was followed by a short tea break to allow attendees some rest and time to digest the information shared during the past couple of hours.

Figure 10: Day 1 Training Session Proceedings

Second session was titled "Software & Data" and delved deeper into explaining the proposed data pipeline, how the surveys fit in the pipeline and each survey's data outputs in detail.

Afterwards, both mapping apps "OSMTracker" & "Kobo Collect" are explained in terms of functionality and how they're fitted to conduct the surveys required. Also, the desktop GIS software "QGIS" and python scripts developed by TfC to facilitate data transformation and cleaning are explained.

The attendees are then asked to test run the apps, while they're in the building, to get a feel for the UI and go through the step-by-step mapping tutorials prepared by TfC beforehand. Finally, OMT facilitation took over to organize the Take-home assignments for attendees, specifically to assign them nearby bus routes on which to conduct one onboard survey and one frequency survey. The received data would be displayed the following day to the attendees for further examination.

4.2.3 Day 2: Data Processing

Day 2 started with examining the received data from the take-home assignments. As expected, there was varying degrees of quality of received data. Some members mis-interpreted some details such as the codification of stop naming in the OSMTracker app, or how to share the files. While others submitted a perfect set of surveys. Received assignments were displayed geographically on QGIS for the attendees to show what went right and what could be improved.

TfC then proceeded onto the first session of the day titled "**Working with Data**". The use of data in monitoring field research progress was explained, and how there are differences between the traditional "manual" and the data-driven automated approaches. KoboToolBox is used to display some basic statistics for the assignments received from the day before such as number of surveys submitted by user.

Then the session went into the exact KPIs to be measured during field research, their role in decision making, and the available tools to monitor KPIs. Such tools include Excel, KoboToolBox, and QGIS. The latter was used in a live demonstration over the onboard surveys received from the take-home assignments.

The following topic was data acquisition, how they can import and export data into different formats. The database structure developed and used by TfC to store their field survey data was used as an example to illustrate the importance of aggregating data to gain insights. Namely the creation of heatmaps from raw stops data received from multiple onboard surveys was shown to the attendees as an example of raw data organization.

Next TfC explains data cleaning, why do we need to clean the received raw survey data and how. This explanation included the matching of raw onboard survey GPS traces to the OSM road network, but also the data validation that happens in parallel. Validation of frequency surveys, for example, would include comparing the number of recorded vehicles departures to the survey duration and doing a sanity check on the figures.

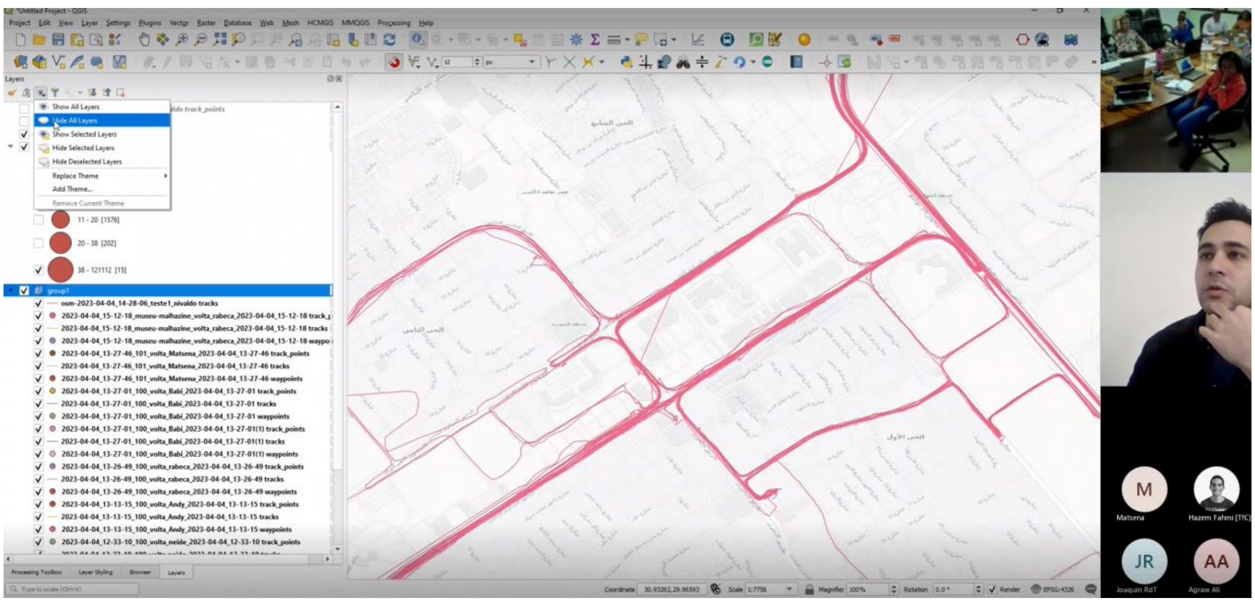


Figure 11: Day 2 Training Proceedings

Finally, the expected data outputs were explained so as to show the final results the attendees would arrive at once they are done with all the steps of the process. The digital datasets of a typical transport network are listed and detailed. (i.e. Routes, stops, road network, bus terminals, GTFS)

After a small break, the hands-on "Data Transformation" workshop commenced. The workshop went on for almost 3 hours with a break in-between and encompassed all the lessons learned during the previous sessions. The workshop targeted data specialists who have had previous experience working with digital transport data and GIS. The workshop flow was based on the paratransit mapping methodology diagram, whereas the full process was divided into 4 sequential phases: Identification → Data Collection → Data Processing → and Data Transformation phases.

The workshop began with the building of a database from scratch, starting by the "Identification" phase, TfC team displayed how to use QGIS to create and fill in the basic data tables that will act as the reference tables for the incoming data, namely the following tables:

- Agencies
- Vehicle Types
- Terminals
- Routes

The database schema and the fill-in form for each was developed by the TfC team beforehand on QGIS and a QGIS map document was shared with the attendees to follow with the steps.

Next, TfC team used previously collected data for the city of "Boane" by OMT to demonstrate the "Data Collection" phase. Attendees were shown how to inspect and validate the raw data, with two scenarios of errors in onboard and frequency surveys, and how to handle them.

Afterwards the "Data processing" was applied for both raw stops and trips data. The processing of raw stops was basic digitization of stops based on a heatmap created for the raw clusters of stops received from the Boane onboard survey on QGIS. Whereas for the processing of the raw trips, a legacy tool created by TfC in 2018 "Route Snapper" was used to align the raw traces to the road network.

Finally, an example of one of the most complex "Data Transformation" aspects was demonstrated, which was the conversion of raw GPS points, processed stops, and processed trips into "road segments" which would contain average speed profiles. These segments could be useful for geographic and temporal analysis on their own but could also be an integral input to the creation of an advanced GTFS feed that carries accurate travel time estimations.

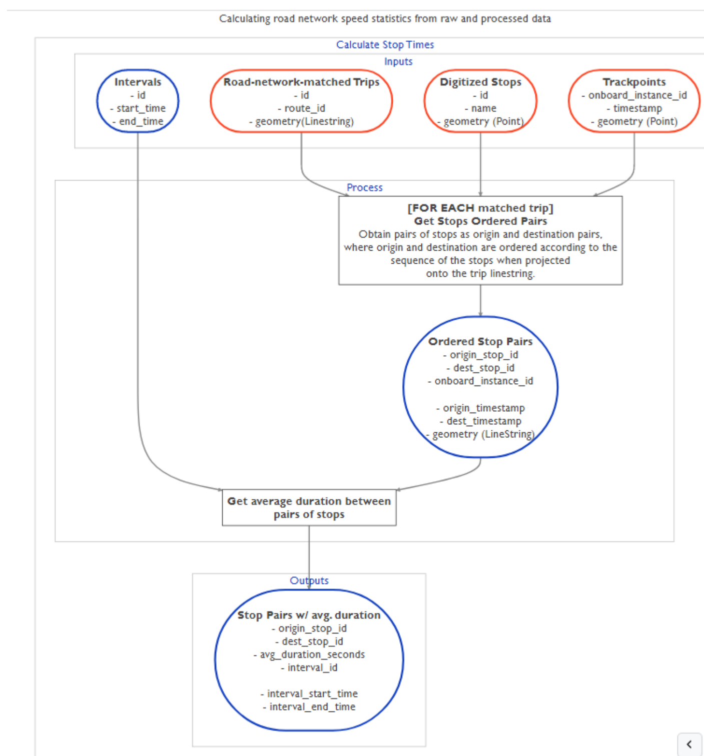


Figure 12: Speed Profile Calculation workflow designed by TfC

4.2.4 Reflections on the Training

Both TfC and OMT agreed on what went right and point to improve after the training. The main takeaway is that there was great teamwork dynamics in setting up and delivering the training. TfC brought in the experience and the content, while OMT intervened to translate to attendees in Portuguese some parts which they deemed necessary to pause at, their staff physical facilitation on the ground was also necessary for the success of the training.

OMT wish for a deeper dive-in on the conversion of survey data to GTFS, as they feel this is an important aspect to their future work. TfC committed to continue supporting OMT on this critical step beyond the scope of this assignment if needed.

Even with OMT's physical facilitation, some complex topics required the trainer's physical presence, namely in the data processing workshop. There was a lot of debugging needed for the different attendees' laptops while installing or configuring software, this could have gone much more smoothly if the trainers could treat those bugs with the attendees in-person.

There was an issue in presenting the full training course to all attendees due to their different backgrounds. Specifically, the "Data processors" could grasp all topics, while the "mappers" were lost on complex technical topics such as data management and processing. In the future it would be better to have attendees selected on a session-by-session basis to make sure team morale stays high.

5 Use Case: Chapas & The BRT

To illustrate the importance of digital data on transport, TfC team sourced and used existing data from OMT's 2017 mapping of Maputo city chapas routes, as well as the digitization of proposed BRT lines. The goal of the exercise was an experimental implementation of basic spatial analysis to demonstrate to Maputo's Local Stakeholders the importance of mapping.

The results of this exercise were displayed over a virtual call to staff members from OMT and AMT.

5.1 Available OSM Data

TfC team downloaded available digital public transport data in the MMA from OSM (266 routes in total), effectively for:

- Chapas (84 route legs, mapped by OMT)
- REM (164 route legs)
- Corujas (18 route legs)

As well as road & railway networks (20 routes in total) namely:

- SADC – Regional Transport Corridors
- National Roads (N1,N2,N4)
- CFM Rail Transport

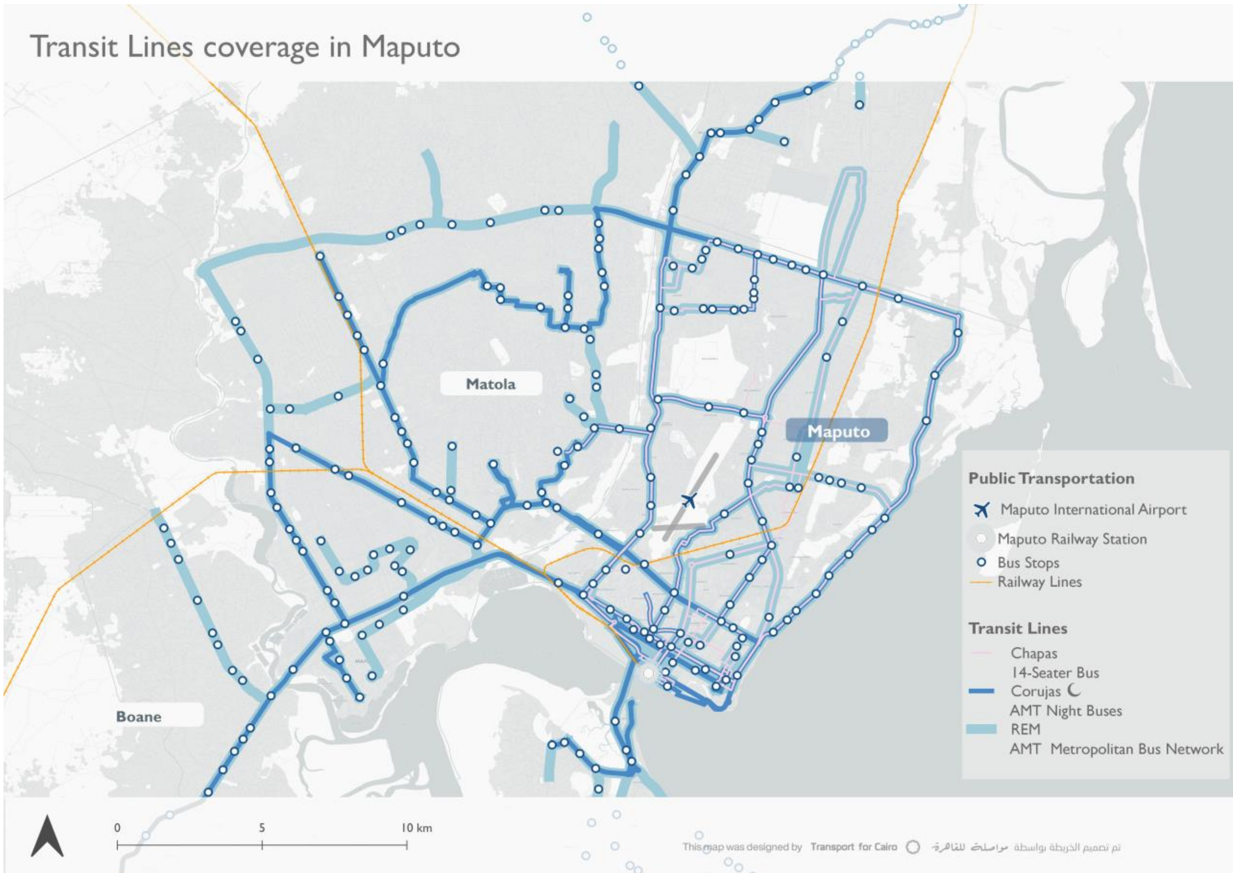


Figure 13: Spatial Distribution of Public Transport Services

Figure 13 visualises the combined dominance of the 4 main PT intercity routes. Although frequency & number of vehicles illustrates relative densities between those modes, yet clearly, we observe AMT dominated network coverage over Maputo, Matola & Boane through primary roads, secondary roads & rail. Chapas mapped indicate the intracity network service in Maputo alone, covering some tertiary roads unserved by AMT.

However, it is worthy to note that Chapas have been mapped within Maputo city only, and OSM data doesn't have information yet on Boane and Matola Chapa routes.

5.2 BRT Lines Digitization

Available information on BRT Lines in MMA provided by T-SUM (Transition to Sustainable Urban Mobility) Project in 2022 and Comprehensive Urban Transport Master Plan for the Greater Maputo Report in 2014, was used to support the initial mapping to the provisional BRT Line Network in the city.

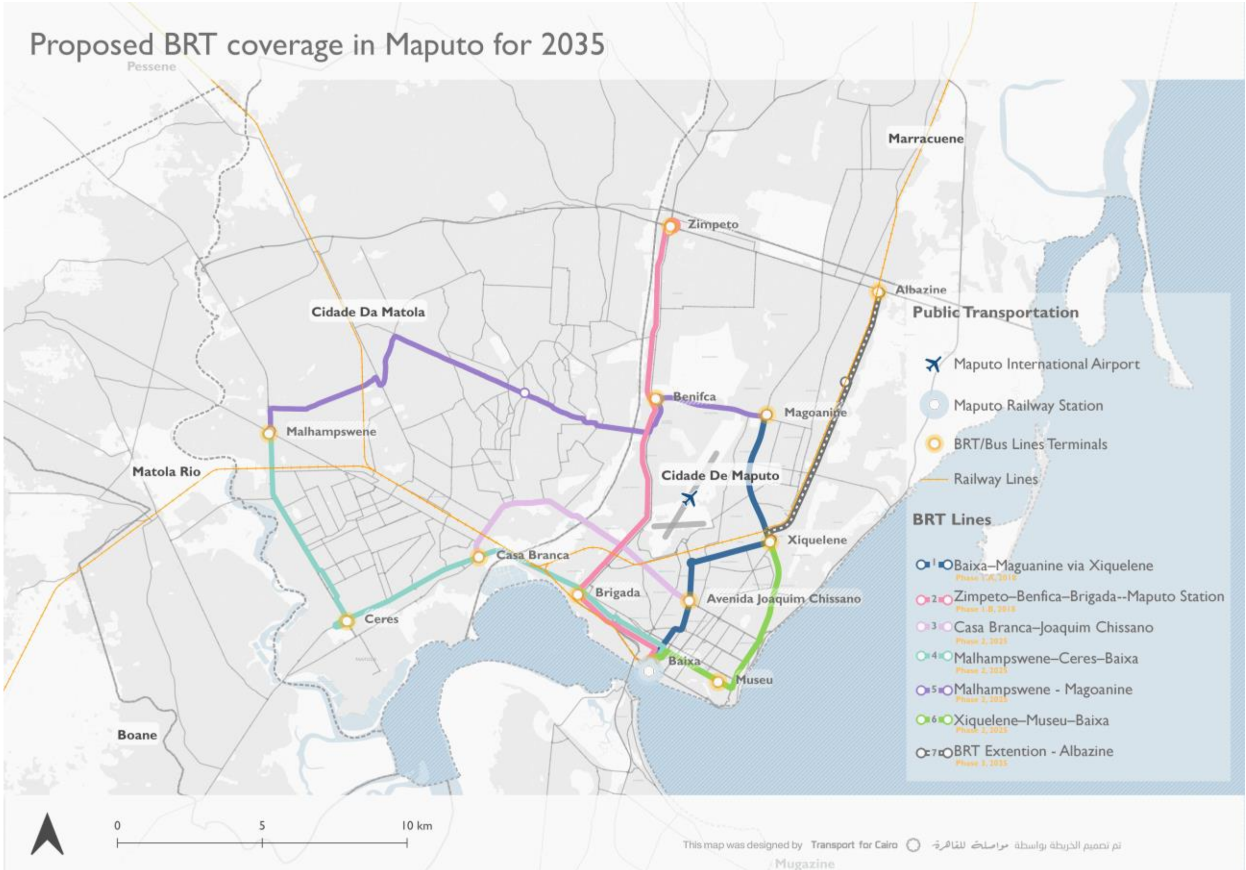


Figure 14: BRT Lines in Maputo

The network is characterized by a *radial pattern*, identified with **3 radials,** and traversed by **2 circular** routes. The plan for BRT development foresees 3 main phases, with major deployment through the 3 radial corridors:

- *Line 2:* Zimpeto - Baixa (along N1)
- *Line 4:* Malhampswene - Baixa (along N4)
- *Line 6 & 7:* Baixa - Albazine

In reference to Figure 14, BRT project phases are detailed below by T-SUM project & lines are coded.

Table 3: BRT Project Phases

Phases	Line No:	Route Origin-Destination	Total Distance
Phase 1: 2018	1	Baixa–Maguanine via Xiquelene	12.9 km
	2	Zimpeto–Benfica–Brigada-Maputo Station	19.1 km
Phase 2:	3	Casa Branca–Joaquim Chissano–J. Nyerere	13 km

2025 Matola-Maputo Connections	**4**	Malhampswene–Ceres–Baixa	21.2 km
	5	Malhampswene – Maguanine	23.8 km
	6	Xiquelene–Museu–Baixa	10 km
Phase 3: 2035 BRT Extensions	**7**	Albasine–via Cardeal A Santo (BRT1 extension)	11 km

Based on OSM & JICA reported data, the following 12 BRT terminals are defined & therefore mapped:

Baixa	Brigada	Malhampswene
Museu	Benifca	Ceres
Magoanine	Casa Branca	Xiquelene
Zimpeto	Avenida Joaquim Chissano	Albazine

Like Baixa, Magoanine, Ceres, Casa Branca, Brigada, Malhampswene, Benifca, Zimpeto & Museu, some terminals exist as Chapas or REM bus stations. While others like Albazine & Xiquelene intersect with Railway stations.\Definition of BRT-Competition / BRT-Complementary

For the BRT system to capture the bulk of existing public transport demand, all passengers using routes that are heavily concentrated on the corridors with high service frequency (BRT-Competitive) should be adequately served and targeted by the new system. The remaining routes (BRT-Complementary) should be left to operate outside of the BRT system.

To fully determine which routes among the existing public transport routes should be incorporated in the BRT system, two factors are to be taken into consideration:

1. The percentage of the existing PT routes traversing the BRT line corridors.
2. Frequency of services per route in a peak direction. (Not included in this case study due to lack of temporal data)

5.2.1 Methodology

The following methodology is applied to determine competitive versus complementary routes for the first factor only:

1. Overlap Analysis is done between BRT Routes & PT Transit trips. A 100-meter buffer area around individual trips are created & compared to calculate overlap area percentage between both BRT & PT Transit routes. (Map 6)
2. Trips which intersect the BRT corridors less than 50% of its length, are complementary trips.
3. Trips which intersect the BRT corridors more than 50% of its length are competitor trips.

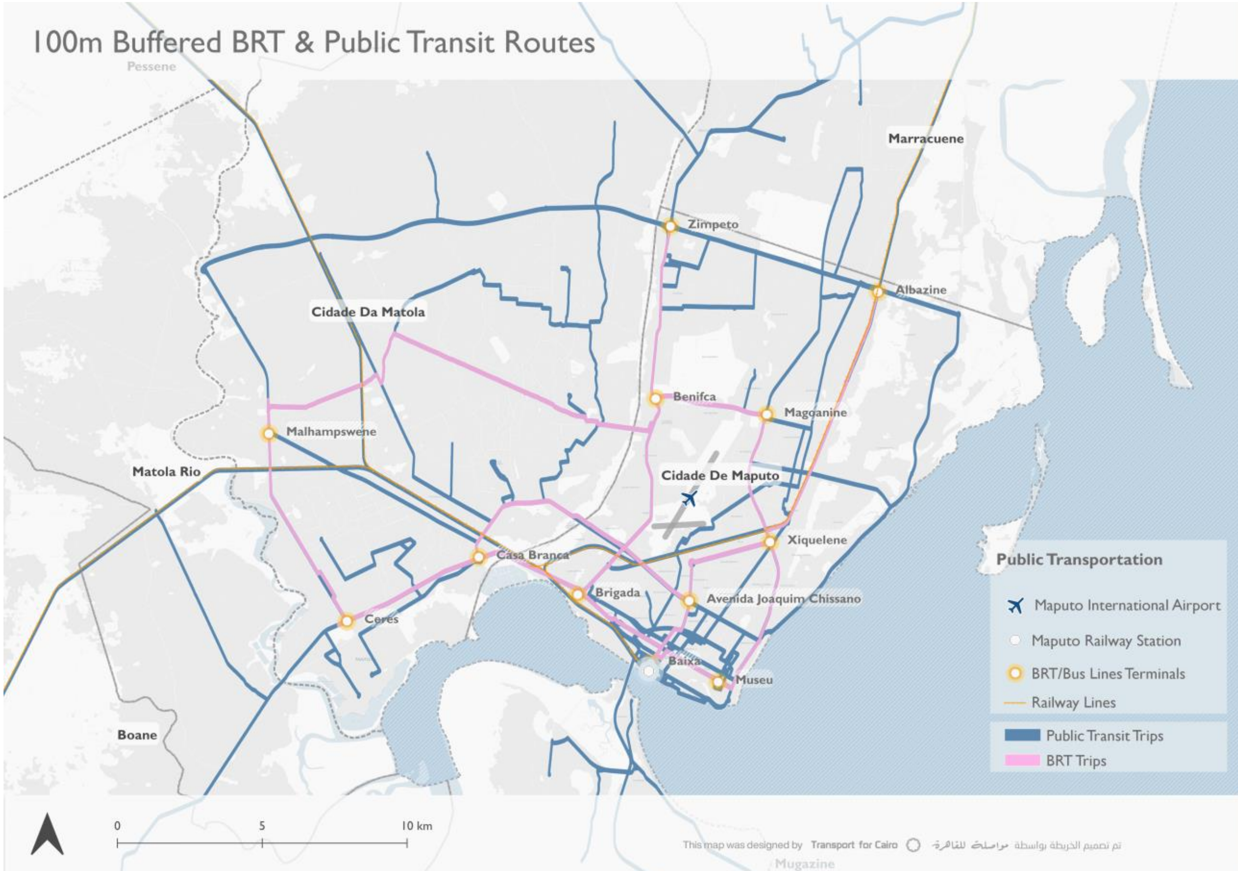


Figure 15: Area-overlap analysis between BRT & existing PT Lines

5.2.2 Results

A total of 266 trips are analyzed, 164 REM, 84 Chapas & 18 Corujas:

- **All trips**: 150 total trips (56.4%) are complementary while 116 total trips (43.6%) are competitor trips. (Figure 16). BRT-Competition is observed around the 3 main radial corridors proposed by BRT while its complementarity observed around northern edge (ring-road), southern (towards Catembe) & eastern (coastal end) where BRT least covers the city.
- **Chapas**: 45 trips (53.5%) are complementary while 39 trips (46.5%) are competitor trips (Figure 18). BRT-Competition is observed around Lines 1 & 2 (main corridors) while its complementarity observed around Av. da Marginal, a coastal road & Maputo ring-road.
- **Corujas**: 14 trips (77.8%) are complementary while 4 trips (22.2%) are competitor trips (Figure 17). BRT-Competition is observed mainly around routes west towards Matola while its complementarity observed around the rest of the city corridors.

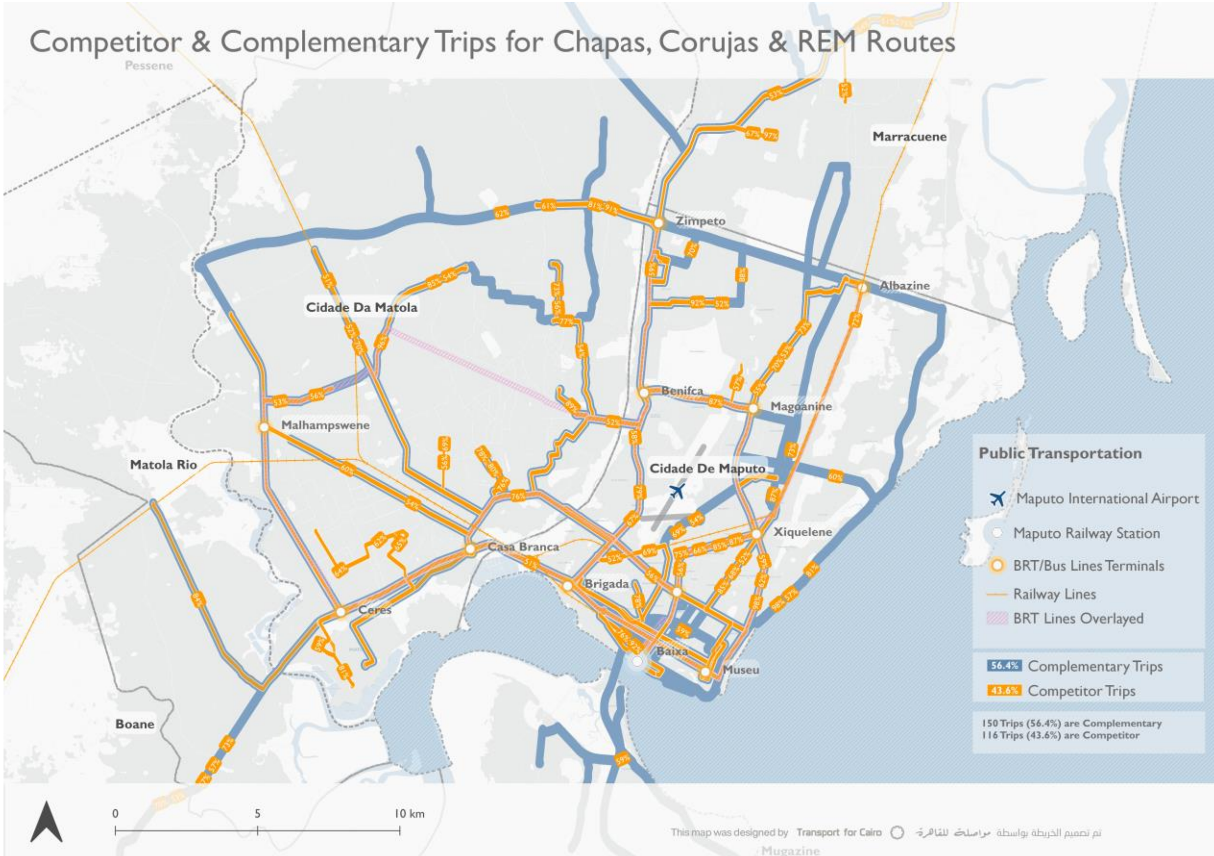
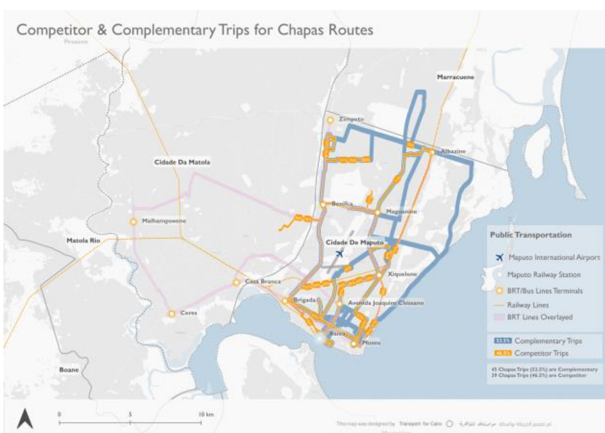


Figure 16: Competitor & Complementary Trips for All PT trips



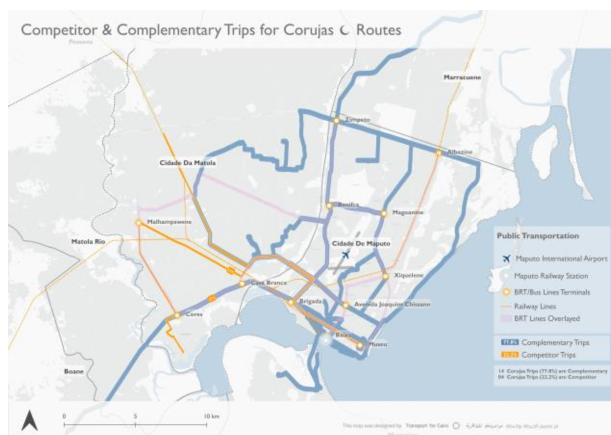


Figure 17: Competitor & Complementary Trips for Corujas	Figure 18: Competitor & Complementary Trips for Chapas

6 Conclusions & Future Steps

Maputo has a reliable local capacity in OMT and AMT to carry forward the mapping of Chapas and the rest of the paratransit network. The upcoming mapping of Matola and the rest of the greater MMA should bring in soon enough the full picture of the existing modes of transport.

Once the mapping capacity is guaranteed and the data exist, the next big challenge for Maputo would be for local stakeholders to go "beyond mapping" by using this data to inform decision making on planned and ongoing mass transit projects, namely the BRT corridors implementation.

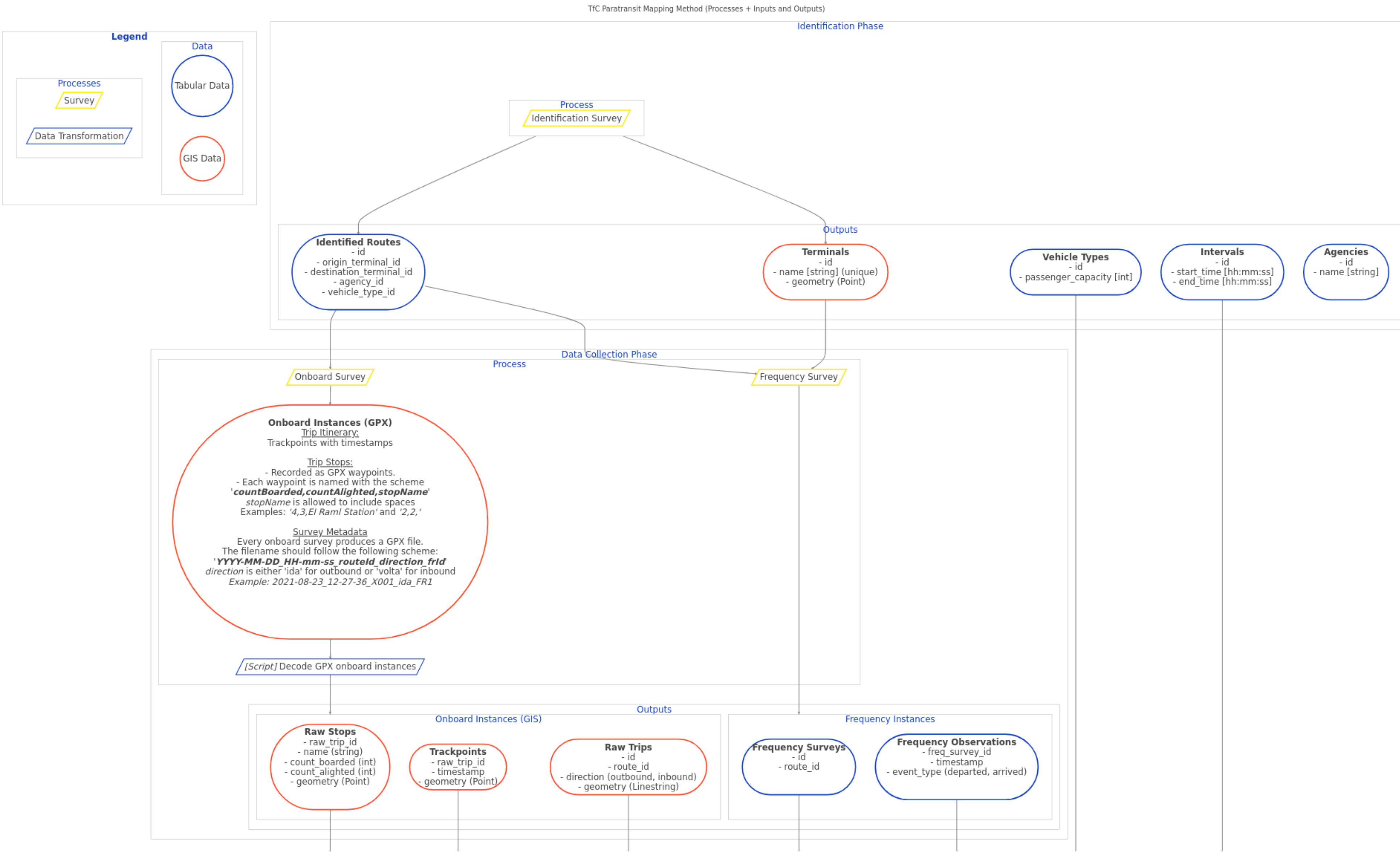
OMT are in a unique position to connect to young mappers from local OSM community, decision makers and engineers from OMT, multi-lateral development agencies such as the world bank, and international consultants. OMT would be able to advocate the use of data in discussions moving forward, starting with the basic questions such as: Where should BRT stations be located? And which Chapas routes are affected by a BRT corridor?

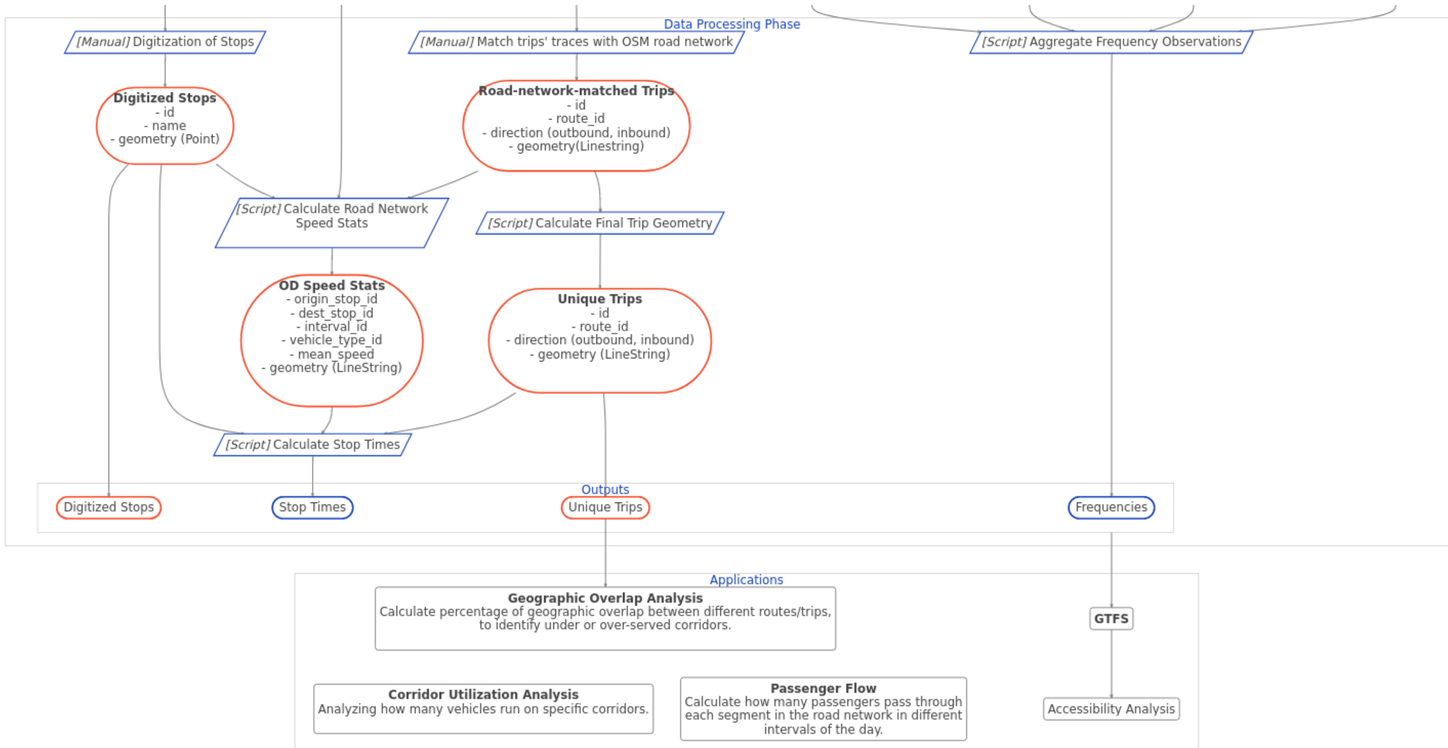
Mapping is one of the very first steps in Maputo's aim for the professionalization of transport in the city. But that is why it's crucial to invest in getting it right. Mapping data will have implications on the challenges to follow which will include infrastructure redesign, consultations with existing operators, procurement of assets, amongst others.

Therefore, ideally the next steps for Maputo can be summarized as follows:

1. Mapping of Matola: OMT's first application of lessons learned from training sessions
2. Mapping of the rest of the MMA
3. Gathering requirements for data analysis and insights: Converting non-technical questions from stakeholders such as AMT into technical data analysis tasks
4. Presenting a full report on the current situational analysis of the paratransit network to AMT and the World Bank
5. Engaging with future tenders and international consultants working within the scope of transport in the MMA

7 Appendix A: Paratransit Mapping Methodology Flowchart




Data Processing Phase
[Manual] Digitization of Stops
[Manual] Match trips' traces with OSM road network
[Script] Aggregate Frequency Observations
Digitized Stops
- id
- name
- geometry (Point)
Road-network-matched Trips
- id
- route_id
- direction (outbound, inbound)
- geometry(Linestring)
[Script] Calculate Road Network Speed Stats
[Script] Calculate Final Trip Geometry
OD Speed Stats
- origin_stop_id
- dest_stop_id
- interval_id
- vehicle_type_id
- mean_speed
- geometry (LineString)
Unique Trips
- id
- route_id
- direction (outbound, inbound)
- geometry (LineString)
[Script] Calculate Stop Times
Outputs
Digitized Stops
Stop Times
Unique Trips
Frequencies
Applications
Geographic Overlap Analysis
Calculate percentage of geographic overlap between different routes/trips, to identify under or over-served corridors.
GTFS
Corridor Utilization Analysis
Analyzing how many vehicles run on specific corridors.
Passenger Flow
Calculate how many passengers pass through each segment in the road network in different intervals of the day.
Accessibility Analysis