Public Transport Quality ?!

Assessing local public transport in Dar es Salaam, Tanzania

Kornelis (Niels) Fikse
August – November 2007
Dar es Salaam Rapid Transit – Dar es Salaam City Council

University of Twente
Enschede - The Netherlands

ITC
International Institute for Geo-Information
Science and Earth Observation

Public Transport Quality ?!

Assessing local public transport in Dar es Salaam, Tanzania

BACHELOR THESIS

Kornelis Fikse
University of Twente

Enschede, The Netherlands
February 2008

Supervisors:
Dr.Ir. M.H.P. Zuidgeest (ITC)
Eng. E. Kitandu (Dar es Salaam City Council)

Preface

In march 2006 I started my research into health care accessibility in Dar es Salaam, Tanzania as part of a GIS course. This small research provided me with basic information about Dar es Salaam, the possibilities of GIS and public transport modelling. Two years later I was ready to start my BSc thesis and I always wanted to do this thesis abroad. ITC provided me with the opportunity to finish my BSc in Dar es Salaam and to do a project on the development of a public transport assessment model.

First, I would like to thank Mark for providing me the opportunity to visit Dar es Salaam, and his advice and ideas about creating a good and useful research. I also would like to thank Alphonse, Daan and Ladis for their comments on the subject and their ideas about how I could model some requirements.
Also I would like to thank Mr Kleist Sykes, Ms Asteria Mlambo, Eng. Kitandu, Eng. Kuganda and Ms Aicha Msonde for their willingness to welcome me at their office, providing me with all data we requested and giving useful comments on my assumptions and methods. Also your interest in my research and your willingness to learn what I was doing was very helpful and stimulating.
Finally, I would like to thank PE Junji Shibata and his colleagues from JICA for discussing the research matters and providing new insights into (public) transport problems in Dar es Salaam and other developing cities.

I really enjoyed my stay at DART – DCC, and this period convinced me that I really would like to work in countries that are completely different and therefore more challenging than the Netherlands. I hope this research can be of use to ITC, DART and DCC and that it will challenge others to do further research in the subjects discussed.

February 2008
Enschede, The Netherlands

Niels Fikse

Abstract

Dar es Salaam, the largest city in Tanzania, has all the traffic and transport problems that are common in developing cities. In order to reduce congestion and air pollution and improve the quality of public transport, Dar es Salaam City Council decided to implement a new high quality Bus Rapid Transit system.
However, at this time there is very little knowledge about the current public transport system and the effects that the proposed Dar es Salaam Rapid Transit (DART) will have on the attributes that describe the trips that people make. It is suggested that the only way a public transport system can work efficient, is by having full knowledge about the system. This research therefore aims at providing a network based database that can store vital information about Dar es Salaam's public transport system.

First an inventory was made of possible characteristics (also called requirements or attributes) that could be used to assess local public transport. Because this list is quite extensive (it contains nearly two dozen different requirements) only three requirements are implemented in this research. These requirements are: accessibility, travel time and travel cost. However as all these requirements are ambiguous they are therefore split into different perceptions.

To make assessment of public transport possible, a network model was created in order to model travel times, travel cost and other attributes that are road or link dependent and can be accumulated during a trip. This model consists of multiple layers which contain information about origins, destinations and road or route links. Each mode has its own layer and the locations of public transport stops are used to transfer 'people' from one to another layer.
In order to model trips in the correct way, people have to wait for public transport and pay a flat fare, dummy links are introduced to assign these attributes to.

In order to model the three requirements that are mentioned before, data have been collected. From the Japanese International Cooperation Agency (JICA) information about population, road network and daladala routes was retrieved. Based on the preliminary reports from Interconsult Ltd. and Logit the routes and itineraries of the proposed DART were retrieved. Most of the data had to be altered in order to store the information or to use the data for analysis.

After this the analysis was performed. It appeared that the model is indeed capable of assessing public transport and can be used in future situations. However, because the input data is still a bit ambiguous, the results may not be absolutely true, but they give a good indication of the possibilities for using the model. The analysis does indicate that the implementation of DART improves travel time and reduces the generalized travel cost that people experience during their trips.

Finally it can be concluded that the proposed model and database can be used for further assessment and storage of information about public transport in Dar es Salaam. It also shows that GIS software can be used to assess transport systems in large cities.

Table of Contents

List of Figures

List of Tables

1. Introduction

First, a general description of the public transport situation in Dar es Salaam will be given and general problems described. Next the problems that this research aims to address are discussed and the objectives are defined. These objectives are translated into research questions. The chapter ends with an outline of the remainder of this report.

1.1. Background

This background aims at providing some basic knowledge about the area in which this research was conducted, the basics of urban transport problems and a short introduction to mass rapid transit in general and bus rapid transit (BRT) in particular.

1.1.1. Dar es Salaam

Dar es Salaam is the largest city of Tanzania, with a population of well over 2.5 million people (Japan International Cooperation Agency, Pacific Consultants International & Construction Project Consultants, 2007; The National Bureau of Statistics of the Republic of Tanzania, 2003) and used to be the capital of Tanzania. In 1973 Dodoma was made the capital, in order to reduce the enormous growth in traffic around Dar es Salaam (Hayuma, 1980), but that had very little effect, since most government organizations and international organizations still chose for Dar es Salaam as their base location.

Due to the fact that Dar es Salaam is such an important location it used to have a population growth of well over 4% per annum (The National Bureau of Statistics of the Republic of Tanzania, 2003) and recent research showed that this percentage is still correct (Japan International Cooperation Agency *et al.*, 2007). This population growth also results in spatial growth of the city, which causes congestion inbound in the morning and outbound in the evening peak hour. This is partially caused by the monocentric urban structure, in which the central business district (CBD) area acts as the single (main) attractor.

This problem is also worsened by the local public transport system that is in place, which consists of about 3000 to 6000 minivans and minibuses which in itself causes problems with greenhouse gasses and noise.

1.1.2. Urban transport problems

Dar es Salaam has more or less the same transport problems as every other large city in developing countries. Because of the increasing income level, which is a result from expanding activities, there is an increasing number of vehicles in urban areas (Dimitriou & Banjo, 1990). This problem is even worse than it would be in most developed countries, since the car is even more a symbol of status and is therefore the ultimate wealth for a lot of people (Wright, 2002). The problems that are caused by an overload of vehicles on the road are numerous and familiar to most people, e.g. congestion, air and noise pollution.

Especially the results of these problems are even worse. For instance people get late at work because of congestion, people get sick because of the air pollution and as a result areas in the city will deteriorate and even worsen the problem, because the people that have money for maintenance will move to other locations in the city, while the poor and underprivileged are doomed to stay in these unhealthy conditions.

This is also where public transport comes into the equation. From different research it becomes clear that over 50% of the trips is made by public bus and over 30% by foot (Japan International Cooperation Agency *et al.*, 2007; Rizzo, 2002). It can therefore safely be said that public transport is the transport mode for the masses. However due to the use of small minibuses huge numbers of busses are used in Dar es Salaam which also cause the effects that are discussed before.

Because public transport is the mode of transport for the people with less money and improving their situation would give them the possibility to improve their environment, thereby improving the economy and wealth of the whole city. Many researchers agree that good transport is vital to a country and that bad public transport slows down economic growth (Iles, 2005). It is therefore clear that in order to improve the quality of Dar es Salaam it is vital to improve the quality of public transport. When good (high speed, high quality, reliable) public transport is available, this might also reduce the number of people that pursue the possession of a car, because public transport is faster and cheaper, thus reducing congestion and improving life quality in Dar es Salaam.

It can be stated that public transport is very important in large cities, because given the amount of road space, private cars cannot fulfil all transport in the city centre. However, in order to have a good public transport system, it is vital to have good management (Dimitriou & Banjo, 1990). Good management however is highly depending on the amount and quality of information that is available to the managers (Daft, 2003)

1.1.3. Dar es Salaam Rapid Transit (DART)

In search of a solution to the urban transport problems a Bus Rapid Transit system, called Dar es Salaam Rapid Transit (DART) is proposed (Inter-consult Ltd. & Logit, 2006a). A BRT system is defined as *a bus-based mass transit system that delivers fast, comfortable and cost-effective urban mobility* (Wright, 2002).

Mass Rapid Transit

Mass transit systems are often in use in large urban areas, and are mainly used on high passenger demand corridors. In general mass rapid transit systems exist of heavy rail, light rail, commuter rail, metro, bus rapid transit, bus lane, bus way systems or a combination of any of these systems. In fact, it is very important to complement a railway system with a integrated bus system, in order to create fast, comfortable and cost-effective feeder systems. Which type of MRT is most suitable depends merely on the available space, passenger demand and available resources. However with a demand under 25,000 PAX/h it is very unlikely that a rail-based system would be the preferred one (Wright & Fjellstrom, 2003).

One of the most important factors of a good MRT system is, next to a fast and reliable system, a system that also considers comfort and commuter experience, thus making sure that the system provides the information and quality that passengers demand (Wright, 2002).

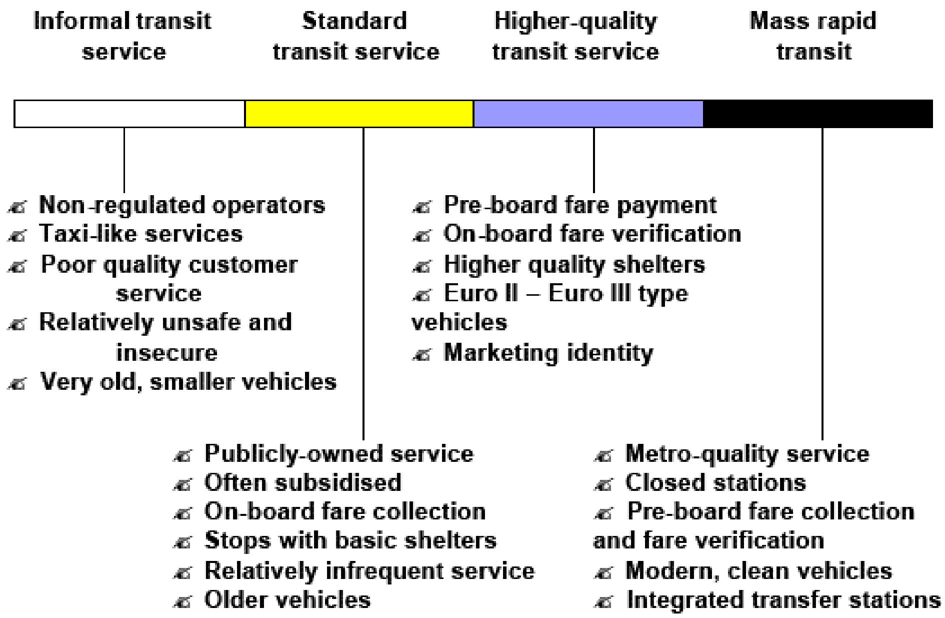

figure 1.1: MRT spectrum

Bus Rapid Transit

A Bus Rapid Transit (BRT) system is a relatively new mode of urban transport. Given its performance (and appearance) it is often referred to as a 'surface metro system'. There are four critical characteristics that make it different from an ordinary bus service:

- Segregated bus ways
- Rapid boarding and aligning (often with pre-boarding payment systems)
- Modal integration at stations and terminals
- High quality stations and customer service

The popularity of BRT systems for large cities started in 1974 in Curitiba (Brazil) (Wright, 2002) at the time a city with a population of 600,000 people, in 2002 there were already more then 2.2 million. However the biggest success story of BRT is without doubt the TransMilenio in Bogotá (Colombia). This large (7 million people) and densely populated city (over 240 people / hectare) showed that a BRT system is capable of delivering a fast, reliable and cheap urban public transport system.

1.2. Research Problem

Very little is known about the changes that will occur in local public transport in the near future. It is even more uncertain what will happen in the long run, when the full DART becomes available to Dar es Salaam. Although the JICA study team is working on a full master plan for the greater Dar es Salaam region (Japan International Cooperation Agency *et al.*, 2007), there isn't a system that focuses on the assessment of local public transport. Because the implementation of DART has not yet started, there still exist opportunities to perform a before-after study. A performance assessment of both the current public transport system and the expected performance of the future BRT system could therefore provide valuable information for the (operational) design of the expected BRT extensions. As stated by Daft (2003) information about systems is important to create good management. This research is a first attempt to create a GIS system that can store information about a local public transport system.

In order to perform a quality assessment for a public transport system, a set of indicators that describe the system is needed. Since most indicators are related to only a small part of the

public transport system, e.g. a road section or a bus route, a model is needed to link the different parts of the system together.
A GIS model of the public transport system would be the preferred way to store this network related information. An accompanying database can be used to store additional information. The advantage of such a GIS model is that two assessments can be made – one of the current and one of the future system – which can then be compared both spatially and temporally using illustrations instead of numbers. Predictions of future BRT extensions can then be compared in a similar way.

1.3. Research Objectives

The main objective of this research is therefore to design and create a GIS network model and database to assess public transport in Dar es Salaam. This database will contain all relevant characteristics of the local public transport system.

This objective is divided into three sub objectives.

- Obtain the indicators needed to perform a network based public transport system assessment, either by fieldwork or literature study.
- Create an environment – GIS model – to store all information, model certain indicators, analyse the indicators and assess the public transport system.
- Evaluate the database and public transport model by performing a performance comparison between the current public transport system and the future DART phase 1 system.

1.4. Research Questions

The research questions are divided between the three sub objectives.

Obtain the indicators needed to perform a network based public transport system assessment, either by fieldwork or literature study

- Which system indicators can be used to asses the public transport system?
- How could these indicators be converted to measurable network characteristics?

Create an environment – GIS model – to store all information, model certain indicators, analyse the indicators and assess the public transport system

- Which software package could be used to store all information and model the public transport network?
- How can the public transport network be modelled using GIS?
- How would the database structure for such a GIS look like?

Evaluate the database and public transport model by performing a performance comparison between the current public transport system and the future BRT phase 1 system

- Which illustrations give insight in the differences in performance indicators?
- What conclusions can be drawn from the performance comparison of the current local public transport system and the future DART phase 1 system?
- Is the proposed GIS model capable of supporting quality assessments for a public transport system?
- How could DART benefit from this assessment model in the future?

1.5. Outline

Chapter two describes how public transport can be assessed, based on a literature study. After showing which requirements should be considered, three different requirements are discussed in depth. These three requirements are used in the remainder of this report. It also discusses the principles of storing information at different levels, based on the type of information that is stored.

Chapter three discusses the construction of the model and the accompanying database. After the selection of the software package, the modelling approach is defined. This approach consists of a network model and a bus stop model. Last the database that is used to store all information is described.

In chapter four the data that is used and collected is described. Also recommendations for improving the data (which will improve the results) are given.

Chapter five describes the assessment of local public transport in Dar es Salaam and the methodology that is described in this report.

Last, the appendixes, describe information that is to extensive to include in the main report. In these appendixes information about terminology and transport system requirements can be found and it also contains the variables and parameters that are considered to be important to create a full public transport model. Last, the system requirements and a basic ‘user guide’ is added, which can be used to reproduce the results of this research.

2. Urban Transport Assessment

Because Dar es Salaam City is currently preparing a massive change in urban public transport a performance assessment of both current and future public transport could provide valuable information for the operational design of the expected BRT extensions as well as an assessment of urban public transport in general. This assessment could also be used to see whether the expectations are met, since often this is not the case (Mackett & Edwards, 1998).

2.1. Assessment Framework

This research adopts a comprehensive framework for assessment, which was created by Vuchic (2005). Although the framework was designed to compare different urban transport modes, it is also capable of assessing the improvement by implementing a new transport system. The framework consists of three main lines, which are Passenger Requirements, Operator Requirements and Community Requirements. Each of these 'parties' has its own interests which are shown below in figure 2.1.

Transit System Requirements		
Passengers	**Operator**	**Community**
Availability Frequency Punctuality Speed / Travel time Comfort Convenience Security and safety User cost	Area coverage Reliability Cycle speed Capacity Flexibility Safety and security Costs Passenger attraction Side effects	Service quality Passenger attraction System cost Reliability in emergencies Social objectives Environmental impact Energy consumption Long-range impacts

figure 2.1: Transit system requirements (Vuchic, 2005)

Each of the requirements is relatively vague and therefore a more precise and concrete description is created for each of the requirements. These descriptions can be found in Appendix A: Description of TS Requirements.

Because of the limited time available, only three requirements will be discussed in depth, but the GIS model that is created should be able to visualize all the requirements.

2.1.1. Availability

The description that is provided by Vuchic (2005) (see also: Appendix A: Description of TS Requirements) is unique in that sense that most authors use the term access to describe the same phenomenon (see also: Lyons, Marsden, Beecroft & Chatterjee, 2001; Murray, 2003; Murray & Wu, 2003).

This description is, however, quite limited since it only considers the distance towards the nearest public transport and leaves out other important factors like travel time and travel cost. This research therefore broadens the scope of availability to a level that is generally described as accessibility (see also: Baradaran & Ramjerdi, 2001; Lyons *et al.*, 2001; Murray, 2003; O'Sullivan, Morrison & Shearer, 2000) and consists of both access and geographic coverage (Murray & Wu, 2003).

Access

One of the most accepted approaches to determine the accessibility towards a public transport network is the Public Transport Accessibility Levels (PTAL) approach (Wu & Hine, 2003). The PTAL approach uses travel time towards the bus stop and waiting time to determine the access time of a certain route, following equation (2.1).

$$t_{acc;i;r} = \frac{d_{i;r}}{v_w} + t_{h;r} \tag{2.1}$$

$t_{acc;i;r}$ = access time from zone i to route r

$d_{i;r}$ = distance from zone i to the closest bus stop of route r

v_w = walking speed

$t_{h;r}$ = waiting time for a bus to arrive at route r

Then an Equivalent Doorstep Frequency (EDF) is calculated using equation (2.2). From which the Accessibility Index (AI) for each location can be determined, according to equation (2.3).

$$EDF_{i;r} = \frac{30}{t_{acc;i;r}} \tag{2.2}$$

$$AI_i = \sum_r EDF_{i;r} \tag{2.3}$$

However, the method as described by Wu & Hine (2003) states that a maximum distance threshold should be applied, so that only stops that people are really willing to walk to are taken into account. In the case of London City they therefore used a threshold of 400m, which is more or less equivalent to a five minute walk. In the case of Dar es Salaam however, a five-minute-walk threshold isn’t realistic, since people are used to walking long distances (Maunder & Fouracre, 1989). Even now people walk for quite a long time to reach a bus stop and even in western countries a 20 minute walk is considered acceptable to reach a facility (Halden, Jones & Wixey, 2005). Therefore a 30 minute walk time towards the bus stop is chosen as the constraint for the PTAL analysis, if there is no bus stop within the 30 minute range, the accessibility index is assumed to be zero. Based on the AI of each location, a PTAL can be assigned, as shown in table 2.1.

Accessibility Level	Range of (Corrected) Accessibility Index
Very Poor	0.00 – 5.00
Poor	5.01 – 10.00
Moderate	10.01 – 15.00
Good	15.01 – 20.00
Very Good	20.01 – 25.00
Excellent	> 25.00

table 2.1: (Corrected) Accessibility levels (Wu & Hine, 2003)

Geographic coverage

The geographic coverage is the amount of space that can be reached from a certain location given a travel time budget (Murray, 2001; Murray & Wu, 2003). However space is not the most important thing people want to be able to reach. Accessibility is about being able to reach facilities, work and other people. It seems therefore more logical to determine the percentage of people that one can reach from a certain location in one hour time. Then a Corrected Accessibility Index could be calculated using equation (2.4), so that the accessibility becomes dependent on the percentage of people one is able to reach in an hour travel time.

$$CAI_i = AI_i \frac{pop_{60}}{\sum_i pop_i} \qquad (2.4)$$

It then seems quite reasonable to use the same accessibility level scale as used in the PTAL approach, which is depicted in table 2.1.

2.1.2. Travel time

Travel time is according to some (Li, 2003) the most important factor for commuters, so it is worth investigating in depth. Travel time is usually defined as the time it takes for an individual for a door-to-door travel (Vuchic, 2005). Therefore the travel time itself can easily be calculated using equation (2.5).

$$t_T = t_a + t_w + t_{iv} + t_t + t_e \qquad (2.5)$$

$t_T =$ total travel time (from door to door)

$t_a =$ access time (from door to public transport stop)

$t_w =$ waiting time (at public transport stop until the vehicle arrives)

$t_{iv} =$ in-vehicle time (thus the time you spend in public transport)

$t_t =$ transfer time (between two modes or two different routes of public transport)

$t_e =$ egress time (from the final public transport stop to the destination)

However since travel time is so important, people tend to overestimate certain parts of their trip, which was to reason to develop a time perception approach. Therefore some extensive models have been created (Li, 2003), but due to large uncertainties in travel time perception in Dar es Salaam (and East-Africa in general) here a more simple model is used, which can be described using equation (2.6).

$$t_p = \alpha_a t_a + \alpha_w t_w + \alpha_{iv} t_{iv} + \alpha_t t_t + \alpha_e t_e \qquad (2.6)$$

When standardising with $\alpha_{iv} = 1.0$, it is estimated that α_a and α_e are 2.0 and α_w and α_t are 2.5 (Wardman, 2004). Thereby equation (2.6) could be rewritten to equation (2.7).

$$t_p = t_{iv} + 2.0 \cdot (t_a + t_e) + 2.5 \cdot (t_w + t_t) \tag{2.7}$$

Using this perceived travel time equation, more insight is gained in how people value the quality and travel time of their trip. This could then be used to check whether people are right when they complain that travel time is still (too) high.

2.1.3. User cost

User cost is usually not considered to be the most important factor for commuters (Li, 2003), but in developing countries this might very well be different. Therefore travel fare in itself is used to assess the urban public transport of Dar es Salaam where the total cost can be calculated using equation (2.8).

$$C_{f;T} = \sum_m C_{f;m} \tag{2.8}$$

$C_{f;m} =$ user cost of trip section m

$C_{f;T} =$ total user cost of the trip

However not only travel fare, but also the travel time is a cost function. The combination of both is generally indicated as generalized cost (GTC), which is defined by Raux (2003) as *a combination of the cost of using the mode and the cost of the travel time as assessed on the basis of the individual's value-of-time*. However different models are used to determine the generalized cost. The general idea is that each link has certain costs, which are valued dependent on the traveller itself (value of time, VOT), its socio-demographic characteristics and the purpose of the trip (Fiorenzo-Catalano, van Nes & Bovy, 2004). Fiorenzo-Catalano *et al.* (2004) propose a model as described in equation (2.9), where they recommend not to make a distinction between route attributes (e.g. walking, waiting, in-vehicle, etc.).

$$\begin{aligned} c_p^s &= \sum_{\forall a \in p} c_a^s \\ c_a^s &= \alpha_m^s \cdot C_a + VOT^s \cdot \beta_m^s \cdot X_a + CK_m^s \cdot D_a \end{aligned} \tag{2.9}$$

$c_p^s =$ the total perceived costs for people of user-class s using path p

$c_a^s =$ is the cost c for user-class s on link a

$\alpha_m^s =$ a factor for the costs perceived on that link for mode m and user-class s

$C_a =$ travel fare on link a

$VOT^s =$ value of time for user-class s

$\beta_m^s =$ a factor for the time perceived on that link for mode m and user-class s

$X_a =$ time attribute for link a

$CK_m^s =$ travel cost per kilometer for mode m and user-class s

$D_a =$ the length of link a

Here the GTC is calculated for each route section, and is summarised afterwards to calculate the GTC for the complete trip with purpose, user class combination s.

Other models emphasize the comfort level that is experienced by the commuters (Phani Kumar, Basu & Maitra, 2004), thereby using calculations like equation (2.10), where dl_T is used to describe discomfort and δ is used to describe other issues like security and safety. Note that the value of time is incorporated in the α's.

$$C_T = \alpha_1 t_{iv} + \alpha_2 \left(dl_T - 1 \right) \cdot t_{iv} + C_f + \delta \tag{2.10}$$

Finally there is a method as described by Currie (2005), which uses the perceived travel time as main component. Currie however decided not to model transfer time, but transfer penalties, which happens often, thereby calculating travel cost using equation (2.11).

$$GTC = \left[\left(\alpha_{a/e} t_{a/e} \right) + \left(\alpha_w t_w \right) + \left(\alpha_{iv} t_{iv} \right) + \left(\#_{transfer} \cdot penalty_{transfer} \right) + MSC_m \right] \cdot VOT + C_f \tag{2.11}$$

So generally speaking there are more or less two approaches:

- using realised time, VOT and travel fare
- using perceived time, VOT and travel fare

The preferred one would be the one using perceived time, since this emphasizes the way people perceive the costs; however this requires a good (realistic) perceived time function, which unfortunately does not yet exist. So this research aims at performing both GTC approaches, to make sure that when a good perceived time model is available, this can easily be used instead of the realised time approach. The GTC using realised time can then be described using equation (2.12), while the GTC using perceived time can be described as equation (2.13).

$$GTC_i = \frac{t_T}{VOT_i} + C_{f;T} \tag{2.12}$$

$$GTC_{p;i} = \frac{t_p}{VOT_i} + C_{f;T} \tag{2.13}$$

2.2. Assessment Attributes

The requirements that are described in Appendix A: Description of TS Requirements have to be converted towards measurable attributes, which can be stored in a GIS database. This has been done for all the passenger and operator attributes and the result is shown in Appendix B: Assessment Attributes. Not only the attributes are defined, but also the level at which the information will be stored. This model distinguishes six levels of storage (and two separate tables), which are discussed below.

Database level

The database storage level is used to store information that isn't location dependent. Examples of this type of information are fuel cost, personnel cost and walking speed. This information is stored in separate tables, to make it easier to find and modify this information if needed.

Hexagon level

The hexagon level consists of information about land-use, population and socio-demographic information that relates to the people in that area. This level is used to generate trips and to assess the influence of (e.g. noise and accessibility).

Road level

At road level mainly the distances of each road section and the corresponding travel time are stored. However a lot of link and route information can be directed towards the road sections, in order to display aggregated information.

Link level

The links represent the public transport routes between the stops and the waiting and transfer links for each bus stop. Therefore information about the number of passengers on certain links and flexibility is stored at this level.

Route level

The route level contains all information that is dependent on the route, like travel fare, average speed, frequency, etc. In fact information about the routes can be transferred through the links toward the roads, so that the information that is presented is spatially correct.

Vehicle level

Here all information that is dependent on the vehicle type is stored. So information about fuel consumption, maximum number of passengers, etc. is stored at this level.

Delay log table

The delay log table can (eventually) contain all information about the delay of BRT busses. Using this information bottlenecks and capacity problems can be identified and tackled, to keep the BRT running smoothly.

Accident log table

In the accident log table, information about incidents is stored, so that problem areas (e.g. difficult crossings, unsafe stations, etc.) can be identified and the government can take measures to prevent further problems.

By knowing the different levels at which information is needed, it is now possible to start looking at the possible methods of creating a database, that is able to perform all calculations and store the required information.

3. GIS Network Model

This chapter deals with the creation of the network model that is capable of assessing the local public transport system in Dar es Salaam. First a suitable software package will be selected, based on several criteria. Second the modelling approach will be defined, and finally the design of the database structure of the proposed GIS network model is discussed.

3.1. Software

Four different software packages (ArcGIS, Flowmap, OmniTrans and TransCAD) were considered to model the urban public transport in Dar es Salaam. In order to make a selection, a set of criteria was developed and used to assess each of the software packages. Finally a choice was made based on the assessment.

3.1.1. Criteria

Based on the discussion in the previous chapters, it has become clear which criteria have to be met, in order to make a good assessment possible. The criteria that are used are listed below.

- Capable of modelling public transport
- Capable of representing socio-demographic information
- Capable of representing spatially distributed information
- Capable of modelling waiting time, transfer time and flat fares
- Expansion of the model should be possible
- Available to researcher
- DCC/DART should be able to continue using the assessment model

All of these criteria are hard criteria, i.e. the software package has to meet all criteria in order to be selected. This is because each of the criteria describes a feature that is needed in order to create a good model for assessment.

3.1.2. Score

Based on the criteria above, a score is created for each of the four software packages. The scores are based on either manuals, brochures or personal experience (see: Caliper Corporation, 2001; ESRI, 2005; OmniTRANS International, 2005; Van der Zwan, Van der Wel, De Jong & Floor, 2005).

Criterion	ArcGIS	Flowmap	OmniTrans	TransCAD
Modelling Public Transport	+	+	+	+
Representing Soc-Dem. Inf.	+	+	-	+
Representing Sp. Dist. Inf.	+	+	O	+
Modelling t_w, t_t and fare	+	-	+	+
Expansion	+	+	+	+
Available to researcher	+	+	O	-
DCC/DART can use	+	+	O	+
TOTAL SCORE	**+**	**-**	**-**	**-**

table 3.1: Software package scores

Therefore, based on these scores, ArcGIS is selected to create a GIS based network model for assessing public transport in Dar es Salaam.

3.2. Modelling Approach

This section describes what an ordinary (commuter) trip looks like and how such a trip can be modelled using ArcGIS. Special attention is given to the way bus stops are modelled in order to provide a more accurate travel time and travel cost estimation.

3.2.1. Public transport trips

An ordinary public transport trip exists of an origin and destination (often home and work) two or more public transport stops and one or more in vehicle ride. A public transport trip could therefore be described by the model as depicted in figure 3.1.

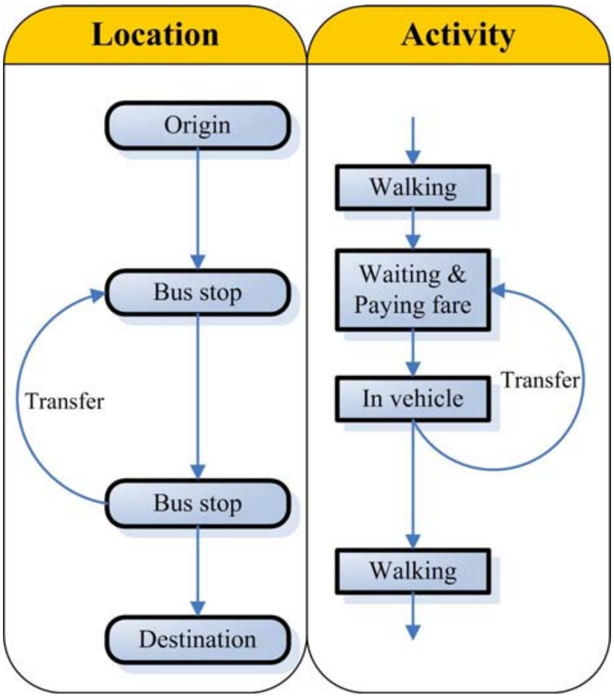

figure 3.1: Public transport trip

So when leaving home (origin) one first has to walk to the bus stop (t_a), waits for the bus to come (t_w) and pays the fare (C_f). Then he travels with the bus to another bus stop (t_{iv}) and aligns. Then there are two options, either one walks to his destination (t_e) or one changes to another bus and therefore has to wait and pay another fare (t_t, C_f).
In case of the daladala public transport system, no integrated fares are available, so there is little difference between waiting for your first daladala and waiting for a second or a third, each time a full fare has to be paid. However in the proposed BRT public transport system, an integrated fare is available, therefore in this case there should be a distinction between getting on the first bus, and getting on a second or a third (at least when the transfer takes place between BRT and feeder buses).

3.2.2. Network model

Two approaches were considered when constructing the network model for public transport assessment. The first approach is the modelling by geometry and the second is the modelling by attributes (Belal, 2002).

Modelling by geometry

This approach uses an extended geometry to model all different routes in a public transport system. Therefore dummy links are introduced that can be used to access a certain route from a bus stop (see figure 3.2). These dummy links are also needed to model transfers between routes and can be used to assign penalty attributes like (flat) fares and waiting time.

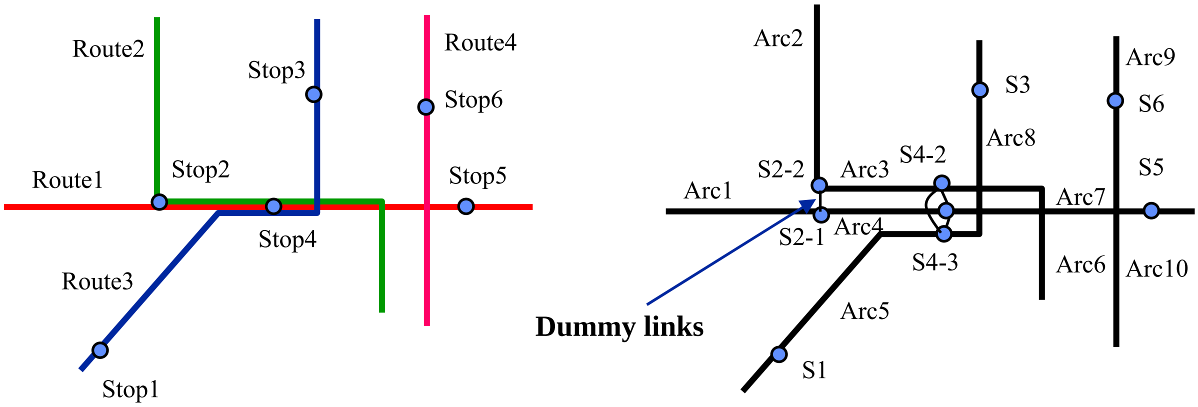

figure 3.2: Transport system and transport model based on geometry approach (Belal, 2002)

This approach results in a very straightforward database, that is easy to maintain, and nearly every GIS software package is capable of analysing such a model.
However the biggest disadvantage is that all route segments have to be created separately which is very labour intensive and takes up much storage space. Not only creating, but also maintaining and updating takes a long time, because each arc has to be edited manually.

Modelling by attributes

This approach can be divided into four different models (Belal, 2002) which are:

- Roads arc-node model
- Dynamic segmentation and linear referencing model
- Virtual network model
- Lane-based model

However, only the dynamic segmentation model and virtual network model are (in theory) capable of handling multi-modal / multi-route public transport networks. However modelling transfers is nearly impossible with the dynamic segmentation model (Luan, 2002) and therefore only the virtual network model is considered.

In the virtual network model (figure 3.3) a route based table is used to locate the stops along the route. A second table keeps track of all the possible options (i.e. continuing on the same mode, thus travelling from route 1 stop 1 to route 1 stop 2, or changing modes thereby travelling from route 1 stop 1 to route 2 stop 1) that a traveller has including all the penalty attributes.

Route ID	Bus type	operator
1		
2	xxxx	xxxxxx
3		
4	xxxxx	xxxxxx

Route Table

Route ID	Position	Stop id
1	xx	stop2
1	xx	stop4
1	xx	stop5
3	xx	stop1
3	xx	stop3
4	xx	stop6

Point event table

From route	From event	To route	To event	non attribute (ex. cost)
1	Stop2	1	Stop4	
1	Stop2	2	Stop2	
1	Stop2	3	Stop2	
2	Stop2	1	Stop2	
2	Stop2	2	Stop4	
2	Stop2	3	Stop2	

2 Vnet table

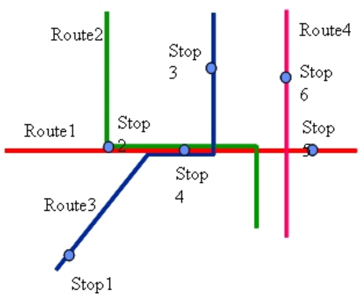

figure 3.3: Transport system and transport model based on attribute approach (Belal, 2002)

The main advantage of this approach is that very little geometry is needed to create a large urban public transport network model, thereby reducing the amount of labour that is needed to maintain and update the geometry.
However the disadvantage is that a very complicated database model is needed, which takes long to build and maintaining requires specialist insight in the modelling approach. Also it is quite difficult to use this approach to analyse the public transport system, since most GIS software packages cannot handle this way of data storage.

Conclusion

Based on the discussion above, the geometry approach is chosen as the desired one, because it is very straightforward and easy to understand. The drawback of being quite labour intensive to create is accepted because it does not outweigh the advantages.

3.2.3. Multiple layer approach

One of the main advantages of ArcGIS is the capability of handling geometry based networks and the ability to use multiple layers (figure 3.4). Using this approach, the trip is modelled more or less exactly as it is performed in reality. ArcGIS uses connectivity rules to determine which layers are connected through which points, thereby ensuring that only daladala routes can be reached through daladala stops, thereby blocking feeder or BRT routes.

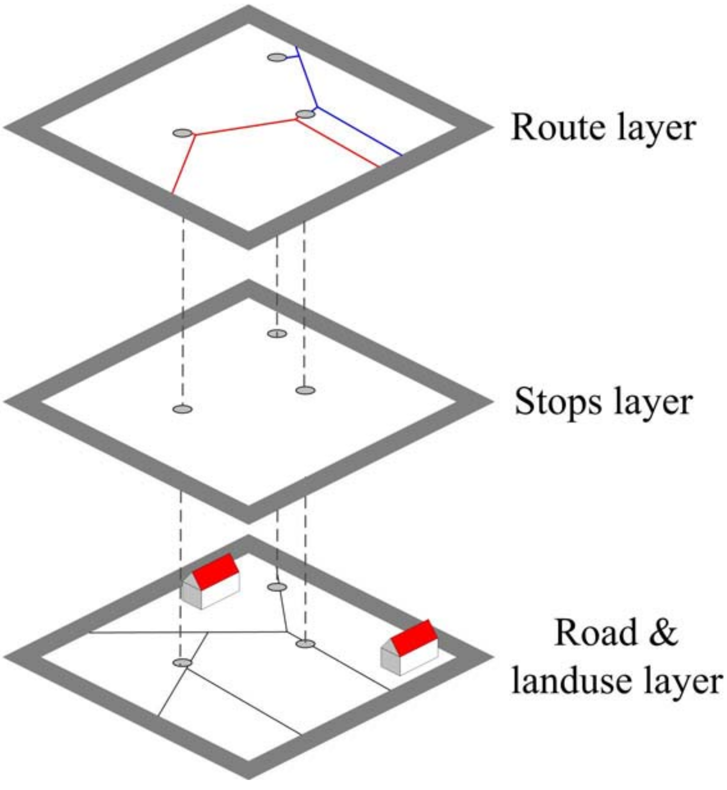

figure 3.4: Multi-level network approach

The lowest layer consists of the road network (walking / cycling network) and information about population and trip generation. Using the road network provided people will walk towards a bus stop and transfer through the bus stop to the public transport layer. There they will board onto a bus (or any other transit mode) and travel to a bus stop near their desired destination. At this station they go back to the road network layer and walk towards their final destination.

3.2.4. Bus stop model

In order to model both waiting time and flat travel fare, dummy links are introduced. These links are used to assign attributes that are not distance dependent (in contrast to for instance in-vehicle time and access time, which are distance dependent). A representation of a bus stop would then look like figure 3.5, which represents *New Posta*.

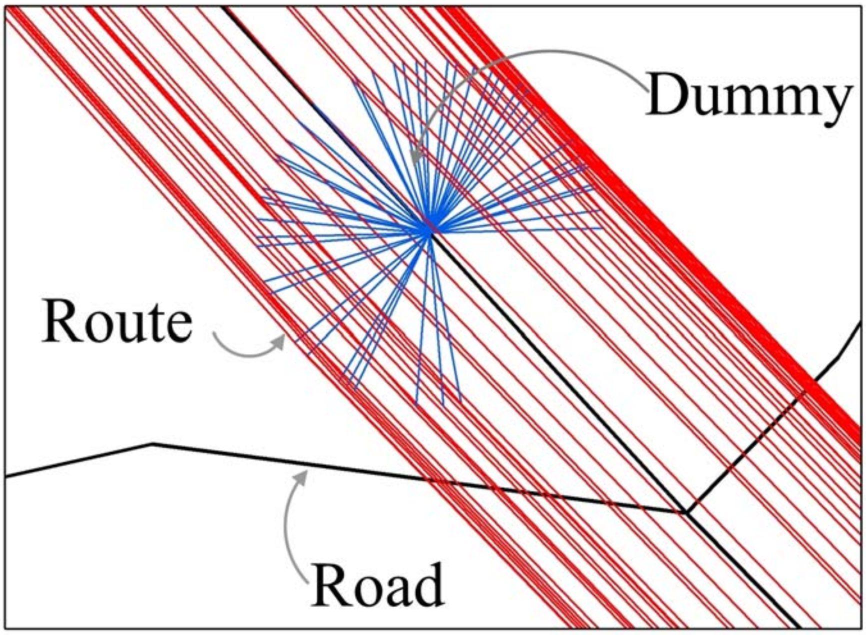

figure 3.5: Network representation of 'New Posta' in daladala scenario

The black lines represent the road network, which is more or less a true representation. The blue lines are used as dummy links. For each route there is a separate dummy link. In the direction from the road towards the route a fare and waiting time are assigned, the opposite direction has no penalties. The red lines are the routes that have a travel time assigned for both directions (which can be different).

Using this model, one can align and board a second bus, but not without waiting for this bus and paying a second fare. This gives therefore a true representation of the penalties one faces when changing buses.

3.2.5. Integrated transfer model

As described in the previous section, changing from one route to another will cost you another full fare. However in the proposed BRT system an integrated (flat) fare is used for the feeder and trunk. Therefore the BRT and feeder lines will be modelled in one layer, as is shown in figure 3.6, which represents *Magomeni Mapipa*.

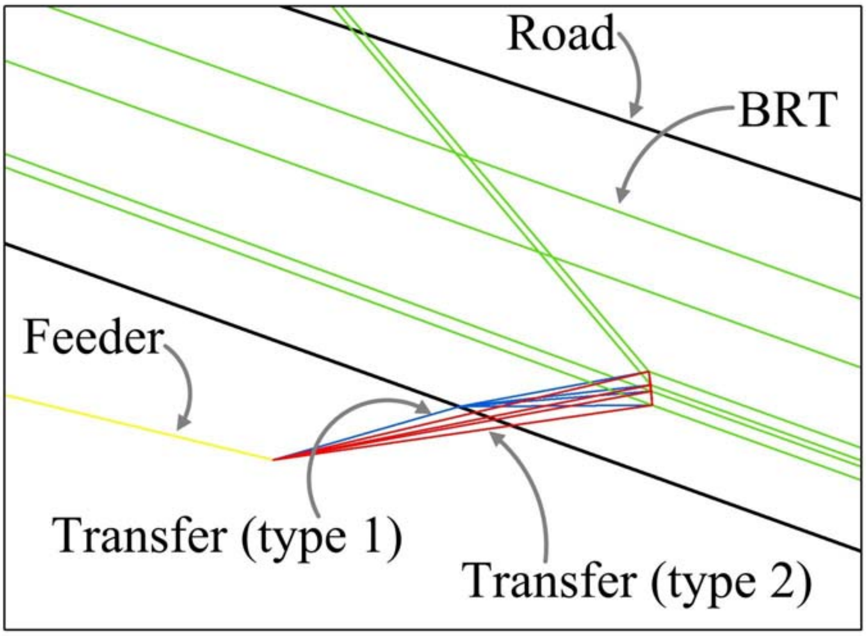

figure 3.6: Network representation of 'Magomeni Mapipa' in DART scenario

Here the black lines (again) represent the road network. The blue lines represent the transfer links (type 1) from the road to either the BRT system or the feeder system. These lines therefore contain the travel fare and waiting time penalty that is related to using that public transport mode. The red lines represent a transfer (type 2) between two integrated routes and therefore have an attribute which represents the extra fare that is charged when changing modes and an attribute that represents the waiting time. The green (trunk) en yellow (feeder) lines represent the different routes.

3.3. Database Structure

Based on the requirements, the derived attributes and the network model, a database structure can be designed. This structure is needed to store all available (and needed) information and to make calculations possible. The database structure that is proposed is based on the conceptual model that is depicted in figure 3.7. The concept is that public transport exists of routes, which are defined by a couple of network segments and a number of vehicles, thus having a certain length, speed, frequency and capacity. Each network segment consists of a couple of road segments, which define the exact route length, instead of the Euclidian distance.

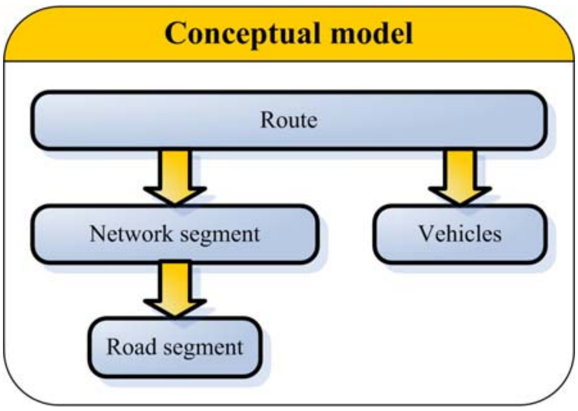

figure 3.7: Conceptual model of database structure

Based on this conceptual model and the multi layer approach as discussed in section 3.2.3. a database structure was designed. A representation of the (main) part of the database is shown in figure 3.8. Next to the elements that are present in the conceptual model (see figure 3.7) additional information about accidents, delay and population is stored.
Besides the tables that are mentioned in figure 3.8, the database also contains a couple of tables and fields that store information that is needed to summarize or calculate the certain requirements. Including these tables would, however, decrease the clearness of the figure.

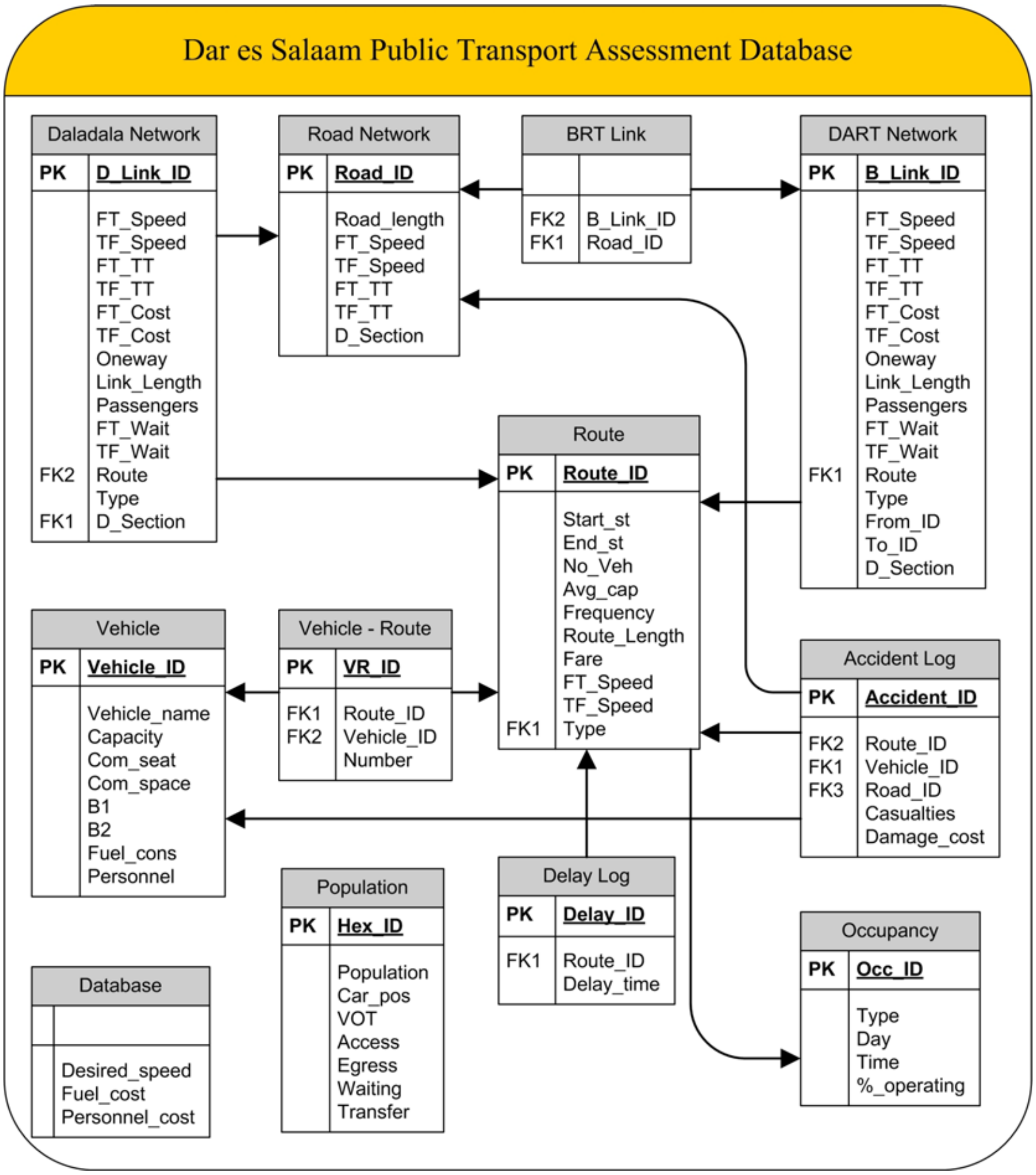

figure 3.8: Database structure Dar es Salaam Public Transport Assessment

From this figure it becomes clear that obtaining all data required to create a full assessment database based on figure 2.1 is not feasible within the time given. Therefore the whole database will be created, but only three requirements (being: accessibility, travel time and travel cost) will be analysed in this research.

4. Data Collection & Modification

This chapter aims at enlisting the data that was gathered for assessing local public transport in Dar es Salaam. Also the necessary modifications and assumptions are discussed, in order to create a clear picture of the validity of the entry data. Finally some recommendations will be made, in order to improve the dataset and thus improving the results of the analysis.
Most of the data was obtained from the research from the JICA study team (Japan International Cooperation Agency *et al.*, 2007). When the information was obtained through other sources, the reports will be mentioned explicitly.

4.1. Socio-economic Information

The socio-economic data is being used to determine the number of people that are living in a certain area, thereby defining the potential traffic flows. In order to give a good representation of the dispersion of people throughout the Dar es Salaam area, land use information was combined with census information.

4.1.1. Land use

The land use data was obtained using the information gathered and classified by the JICA study team. The land use was classified into six categories which were aggregated towards three categories, following table 4.1.

Classification by JICA	Aggregated Class
Residential area	Urban
Other urban area	
Industry	
Vacant / Agriculture	Rural
Water body	Uninhabited
Military	

table 4.1: Land use classifications

The new classifications were used to model the dispersion of people within a certain ward. Therefore it was assumed that within a land use class within a ward, the density of people would be equal.

4.1.2. Population

The population data was derived from the 2002 population census as stored by the JICA study team. The information was gathered at ward level and however for some wards a distinction was made between urban and rural area, this information was not used to determine the amount of people per area. To determine the distribution between urban and rural areas, for all wards that were classified as having both urban and rural areas (based on the information in the ward layer, not based on the land use layer) the population density for both urban and rural area was determined. It could be concluded that the population density in the urban areas was about 15 times as high as the population density in the rural areas.

The information of both the land use and population layer was combined in order to create one layer where land use, population and area size were present. Based on area size and land use, the population was recalculated using equation (4.1).

$$pop_i = \frac{f_{i;l} \cdot area_i}{\sum_i f_{i;l} \cdot area_i} \cdot pop_I \qquad \vee\, i \in I \tag{4.1}$$

$pop_i =$ the weighted population or area i

$f_{i;l} =$ factor that represents the land-use of area i

$area_i =$ size of area i

$pop_I =$ population of ward I

4.1.3. Hexagons

In order to model public transport movements, the areas that contain population information had to be sufficiently small, otherwise it is impossible to tell which route is shorter, but could not be too small, because then it would take too long to calculate all the different routes. Therefore a 250m hexagon was used for the city centre; while a1000m hexagon was used for the areas that where farther from the city centre. This was possible, because the population density in those areas is relatively low. Furthermore there was very little information about the road network and bus stops, so enlarging the hexagons wouldn't diminish the quality of the analysis. Using this method 3666 hexagons (or portions) were identified, which can be used for analysis (see figure 4.1).

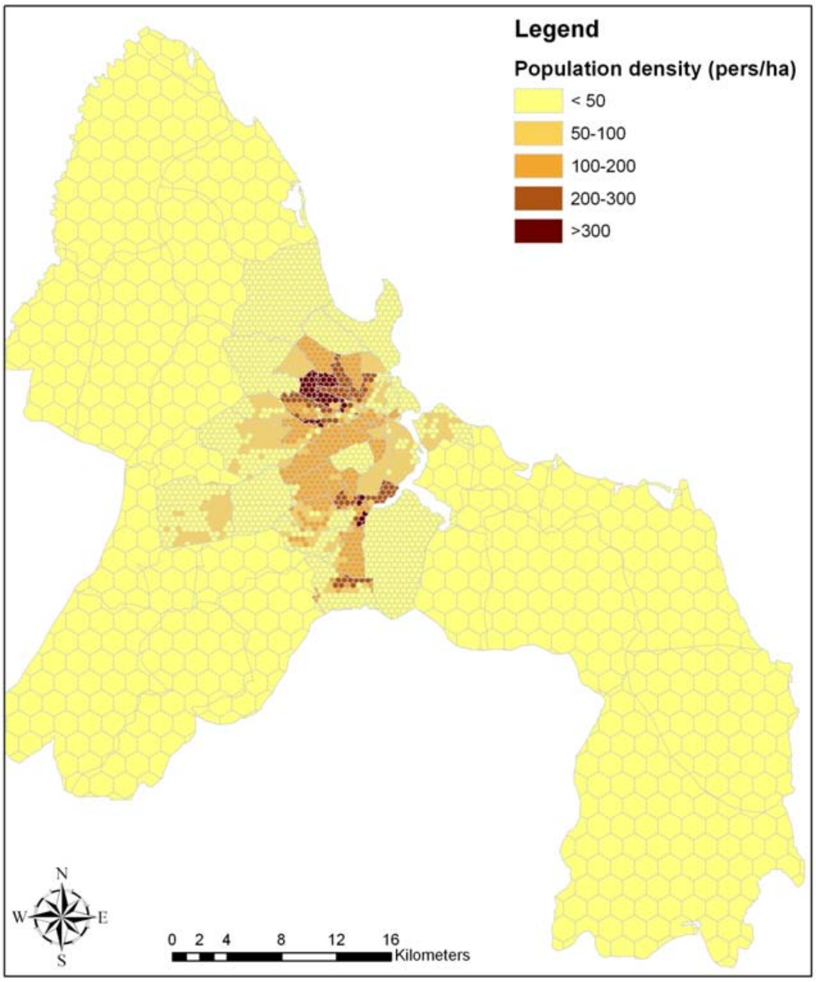

figure 4.1: Population density in hexagon representation

4.1.4. Value of time

A good value of time (VOT) for the Dar es Salaam region was not available. Preliminary results of the JICA study team showed a value of 1.88 Tsh/min which seemed quite low, cause this would mean that people are willing to walk over an hour to save the travel fare of one daladala (Tsh 250).

Preliminary results of Mestrum's research (2008) showed that the walking value of time was expected to be about 12 Tsh/min, while the in vehicle value of time was only 3 Tsh/min, thus indicating that people would value walking time four times worse then in-vehicle travel time. This difference in travel time value seems quite high, because usually walking time is valued only two to two-and-a-half times the in vehicle time (Wardman, 2004).

Therefore the raw data that was collected by the JICA study team (a stated preference survey to assess the willingness to pay for the proposed BRT system) at bus stops was used to estimate a value of time (see also: Japan International Cooperation Agency *et al.*, 2007). It was assumed that the average value of time could be estimated when the stated modal split between daladala and BRT was 1:1. This resulted in an average value of time of 7.4 Tsh/min.

However these values of time seem rather high and apparently didn't represent the value of time that citizens would have. Therefore a value of time of 5 TSh/min is used, which was defined in a previous research, according to Eng. Kitandu.

4.2. Road Network

For the road network there were several sources available, though each one used its own spatial reference. Therefore the road network as provided by the JICA study team was used, since its reference matched with the land use and population layers. The road network was checked for flaws using the daladala routes and other road maps that were available (e.g. road network provided by dr. Sherif Amer and road network provided by ing. Christian Lindner). However it appeared that the JICA study team road network was the most comprehensive one. Based on own observations within the central business district (CBD), the proposed BRT routes (Inter-consult Ltd. & Logit, 2006e) and existing daladala routes (Japan International Cooperation Agency *et al.*, 2007), some minor adjustments were made to the road network, in order to make all routes able to follow the road network. Only in the remote areas (e.g. Kigamboni) no additional roads were created, because of the uncertainty of the location of these roads.

4.2.1. Daladala Routes

The daladala routes were digitized in TransCAD (Inter-consult Ltd. & Logit, 2006e) and obtained through the JICA study team research. These files were converted from TransCAD via MapInfo to ArcGIS shape-files. However it appeared that their spatial reference was wrong, therefore the bus stops as well as the routes had to be reproduced manually.

4.2.2. Bus stops

The bus stops were manually relocated towards the nearest road in order to create connectivity between the bus stops and the road network. However, after a while it became clear that the locations might not be correct all the time, because they were not on the exact route. Because the error was limited to 5 to 10 meters, they were not corrected, since that would take a lot of time while the results wouldn't improve much. In some cases it was not clear on which road the bus stop was (e.g. in the remote rural areas, where there were no roads within a 1 kilometre radius). In those cases the bus stop was left at its original location.

4.2.3. Routes

In order to create a public transport model, each route consists of multiple route segments. Therefore each route was created manually. This was done by creating dummy links from each bus stop and then connecting al the dummy links to each other. For routes with a circular shape two routes were digitized, one for the from-to direction and one for the to-from direction. Each route segment has attribute values for in-vehicle time, waiting time and travel fare in both directions. The result of the network model is depicted in figure 3.5.
Unfortunately not all routes were available and the ones that were available weren't always correct. It was however not feasible to check this information within the limited time available. Also drivers can change routes whenever they want, simply because there is no fixed schedule or enforcing government.
For the route segments that could not be linked to a road, a length of 1.3 times the Euclidian distance was assumed, since this usually approximates the detour factor.

4.2.4. Speeds

It was quite hard to get good information about the daladala operational speeds. Therefore a small survey was carried out on twelve different daladala routes (see table 4.2). Each route was travelled between six and eleven times, in both directions. The time at certain locations was noted down and based on these locations the average speed was calculated. The results show that especially outside the centre area, high speeds (>25 km/h) are realized. Inside the densely populated area's the speed drops to somewhere between 10 and 15 km/h.

Route	Average speed (km/h)
Kibamba – Kariakoo	26.6
Mbezi – Muhimbili	24.3
Tabata Mawenzi – Muhimbili	11.8
Kigamboni – Vijibweni	14.3
Vingunguti – Kivukoni	15.2
Mabibo – Kariakoo	18.3
Kigamboni – Kibada	22.3
Gongo la Mboto – Kivukoni	25.5
Buguruni – Posta	13.4
Mtoni – Kariakoo	26.7
Mwenge – Pugu	25.3
Mwenge – Gongo la Mboto	22.4

table 4.2: Average daladala speeds

Therefore a speed of 25 km/h was used for all roads, exept for the ones in the densely populated city centre are. In this are a 10 km/h speed was assumed for the daladala's.

4.2.5. Frequencies

The frequencies for each route were first estimated by the number of vehicles that was licensed on each route, based on the SUMATRA information (SUMATRA, 2007). However it appeared that a lot of information was missing or could not be linked to a certain route. Therefore the information gathered by Logit (Inter-consult Ltd. & Logit, 2006c) was used in order to complete the frequency information. Based on the number of vehicles the frequency could be calculated using equation (4.2), while the expected number of vehicles could be calculated form the frequency based on equation (4.3).

$$f_r = \frac{30 \cdot v_r \cdot n_r}{l_r} \tag{4.2}$$

$$n_r = \frac{f_r \cdot l_r}{30 \cdot v_r} \tag{4.3}$$

After combining both datasets only eight routes were lacking information about frequencies and number of licenses. It is assumed that there are no vehicles currently operating on these routes.

4.3. DART Routes

The information about the proposed DART routes (both feeder and trunk services) was obtained through the consultant reports (Inter-consult Ltd. & Logit, 2006e, 2007). Because the digital files were only available in TransCAD and had another projection than the road network and ward shapefiles the routes were digitized manually based on the pictures in the report. For the main trunk route (Ubungo – Kivukoni) the exact location of stops was derived from the detailed draft design (Inter-consult Ltd. & Logit, 2006b). The complete DART routes are depicted in figure 4.2.

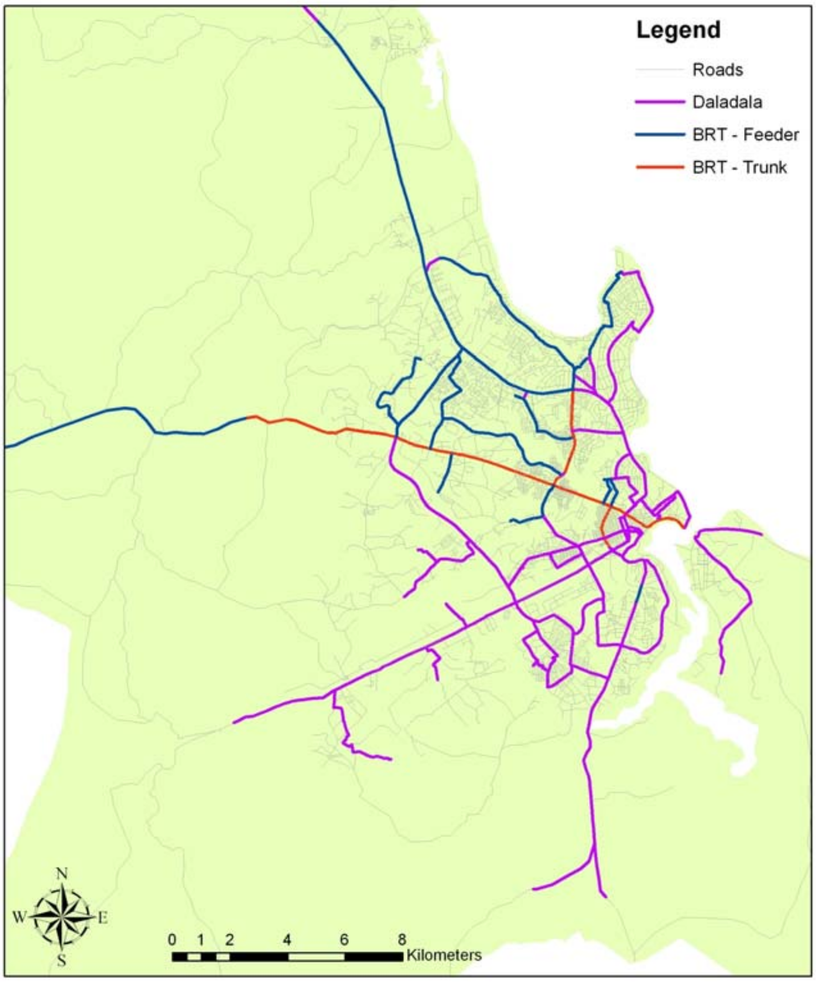

figure 4.2: Public transport routes (DART phase 1)

4.3.1. Feeder stops

The design of the feeder routes wasn't as extensive as the trunk route design. Therefore some assumptions had to be made. Since most feeders would follow the former daladala routes, it was assumed that they would serve the same stops as the daladala's did. However it was assumed that the feeders don't stop on trunk roads (roads where trunk services operate).
For the routes that didn't follow the former daladala routes the stops were located manually, by attempting to approximate the normal daladala stop behaviour.

4.4. Recommendations

From the previous sections it becomes clear that some assumptions had to be made, either because the data wasn't available or the data available was ambiguous or falsely. This section therefore discusses some work that can be done to improve the available data and thereby improving the results of this (and future) analysis.

4.4.1. Road network

Within the city centre the available road network data is quite extensive and appears to be correct (only a few flaws were discovered). However outside the centre very little is known

about the road network (only the main roads are located). Not only does this make the results for the greater Dar es Salaam region less reliable, it also makes planning for the greater Dar es Salaam more difficult since very little is known about these premises.
It is therefore suggested that the Urban Planning Department of Dar es Salaam City Council tries to obtain this information, e.g. through satellite images or measuring them using GPS.

4.4.2. Bus stops

A problem that is related to the uncertainty about the road network is the problem of missing bus stop information. Outside the city centre it is very unclear (e.g. for the Kigamboni, Kimara and Charambe areas) where the bus stops are located. Probably they will just stop whenever there are customers, but for feeder routes this should not be the case. It is therefore recommended to define the locations of bus stops (even if this would imply determining where the settlements are) more precise. This would increase the modelling and analysis results and improve the opportunities for urban and public transport planning.

4.4.3. DART design

From the research it becomes clear that DART does not eliminate the daladala routes in the coverage area (e.g. Kijitonyama, Msasani, Mikocheni and Kawe area). Given the freedom that daladala drivers now have to alter their routes, it might be that the daladala will compete with the feeders on certain routes (see figure 4.3). It is therefore recommended to reassess the daladala routes that could compete with DART or enforce the daladala drivers to stick to their schedule.

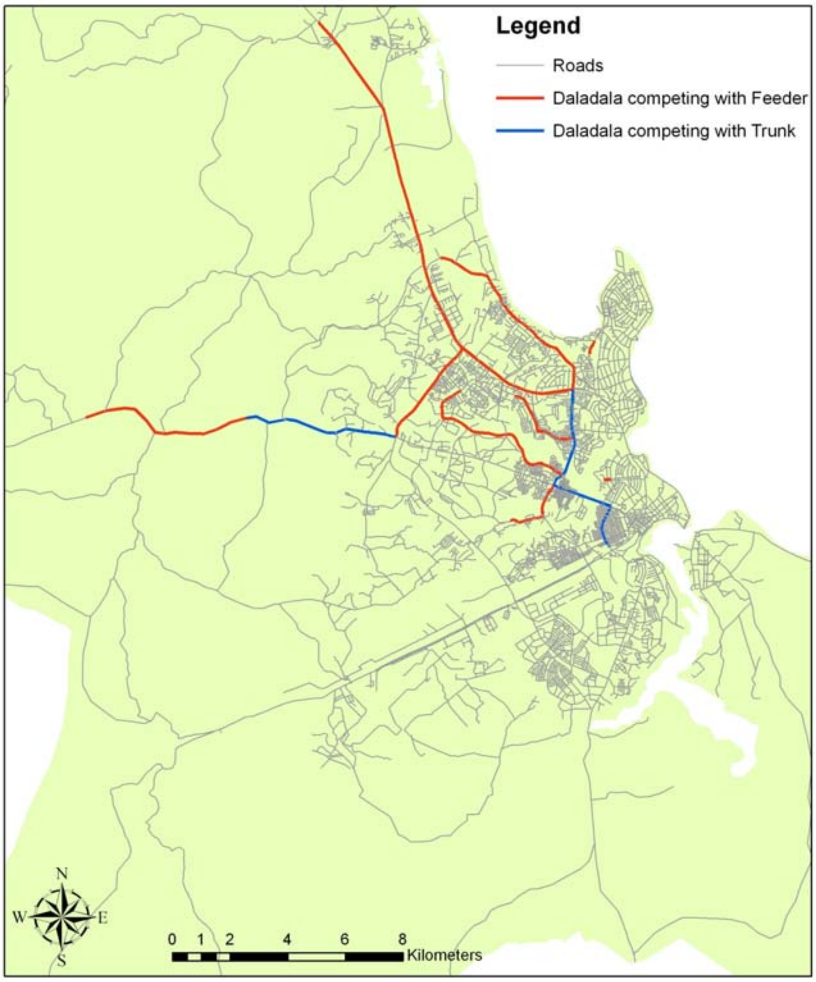

figure 4.3: Competing daladala routes

4.4.4. Value of time

The value of time that is used in this research is a very rough estimate since both the labelled stated preference research (JICA study team) and the research by Mestrum (2008) are not able to produce good (reliable) results on its own. It is therefore recommended to perform a true and extensive value of time research in order to be able to estimate the values of walking time and in vehicle time (for different modes) for different demographic groups.

5. Dar es Salaam Public Transport Assessment

In order to show the possibilities of the model, three different requirements (from Vuchic's list as shown in figure 2.1) will be assessed using the proposed urban transport model. For these indicators, a comparison will be made between current and future public transport in Dar es Salaam.

5.1. Accessibility Assessment

The accessibility indicator, as discussed in section 2.1.1. indicates how close people live to a public transport route and how often this area is being served. Therefore it uses the Accessibility Index which can be converted to a Public Transport Accessibility Level. First the official PTAL shall be discussed, in which only geographical accessibility is taken into account. Next also the corrected PTAL will be estimated, which also considers the number of people that can be reached within an hour travel time.
Appendix D: Creating Accessibility Analysis Layers contains a step-by-step manual that describes how to perform this analysis.

5.1.1. Geographical PTAL

In order to determine the geographical PTAL, first the access time from each origin to all the bus stops has to be determined using equation (5.1), which is identical to the equation (2.1) that is discussed in section 2.1.1.

$$t_{acc;i;r} = \frac{d_{i;r}}{v_w} + t_{h;r} \tag{5.1}$$

So in order to determine this value for each origin *i* and route *r* the distance from this origin to the bus stop has to be calculated. This was done using the road network, and for the distance from the origin to the road network Euclidian distances were used. Next the minimum value of $d_{i;r}$ for each route *r* had to be determined, in order to find the bus stop from route *r* that is closest to origin *i*. Using this information in combination with the average frequency and the walking speed, the value of $t_{acc;I;r}$ can be determined for each *i*, *r* couple. Each value of $t_{acc;I;r}$ that was over 30 (minutes) was left out, because it didn't meet the maximum threshold constraint.
When all these values are determined, the Accessibilty Index for each origin *i* can be determined using equation (2.2) and (2.3). These values were translated into accessibility levels using table 2.1.

For the current daladala public transport system, this analysis shows that the Accessibilty Indices range from 0 to 39, which indicates that there are areas where the public transport could be described as *Excellent*. The results of this analysis are shown in figure 5.1. It is clear that the city centre is served quite well, and around the main corridors the public transport scores at least *moderate*.

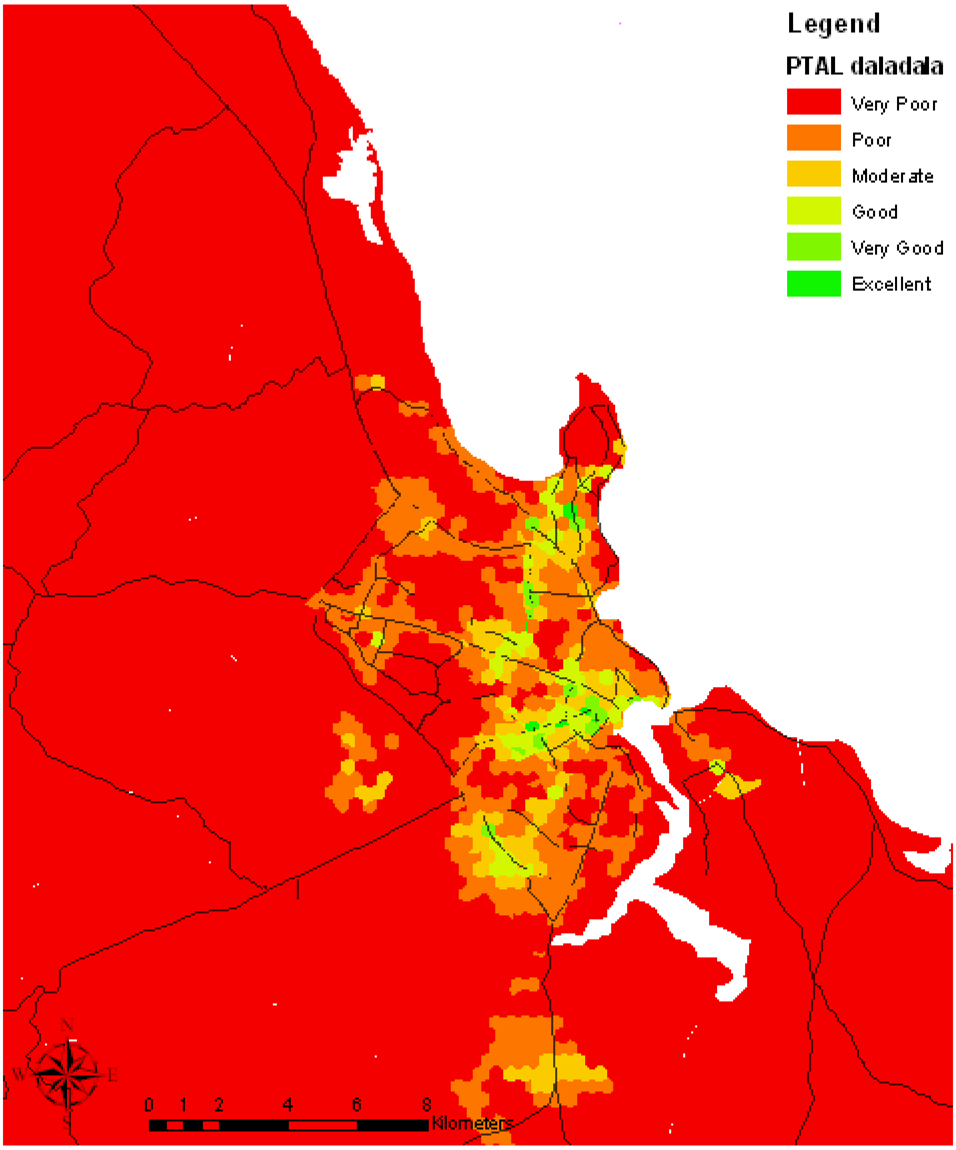

figure 5.1: PTAL for daladala system

The same analysis was performed for the phase I DART system, and it showed Accessibility Indices ranging from 0 to 52. The results are depicted in figure 5.2. It shows that a lot of areas are served much better then they were in the daladala system.

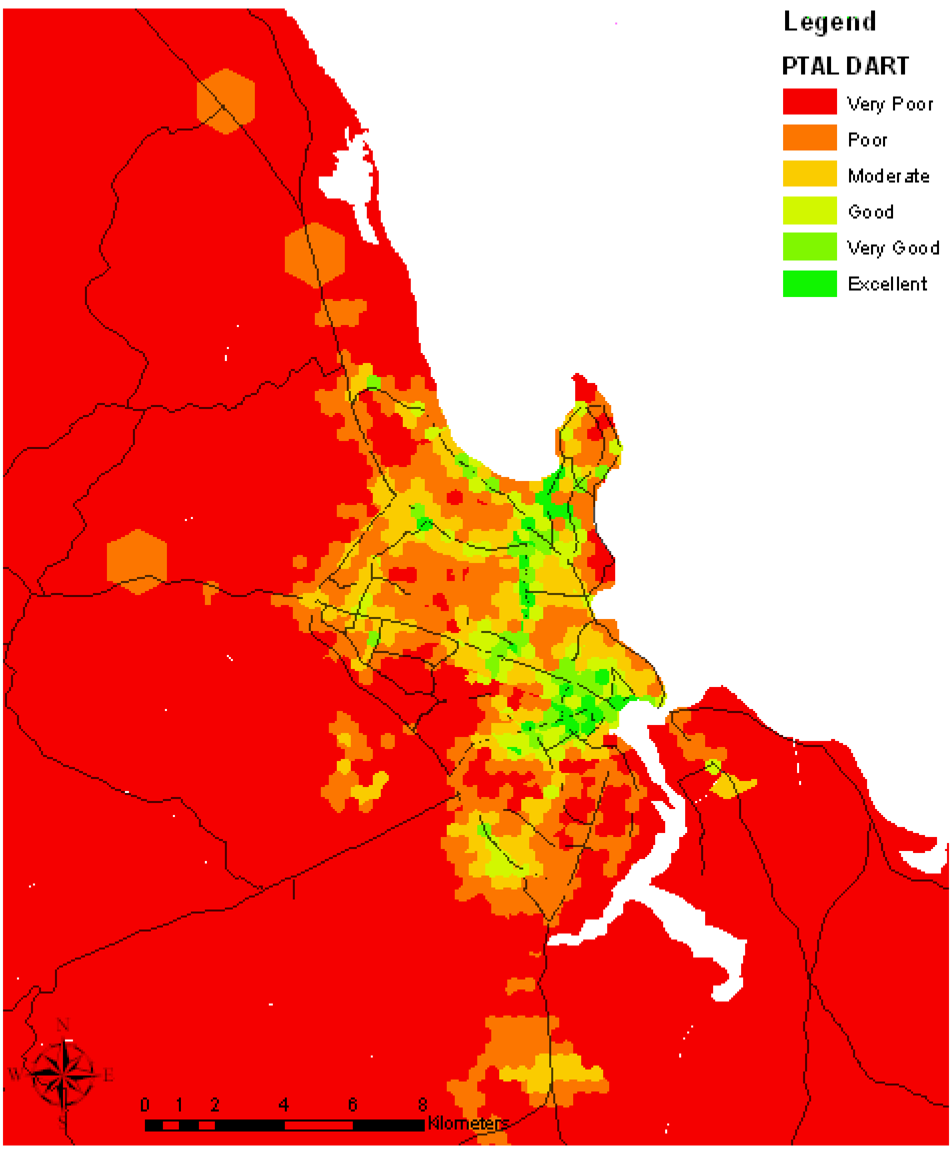

figure 5.2: PTAL for DART phase I

It is quite clear that some significant improvements in geographical accessibility can be achieved by the new public transport system. These improvements are depicted in figure 5.3 and it becomes clear that the geographical accessibility around the DART phase I and accompanying feeders improves significantly. Although a lot of routes disappear from this area, the frequency of the new routes is higher, so that the accessibility improves. Especially the Msasani and Masaki area (Msasani peninsula) accessibility improves, since one feeder is serving an area, where before no daladala route was located. The reason that the PTAL improves in the city centre is that nearly all routes still exist or are replaced by trunk routes and new trunk routes are implemented.

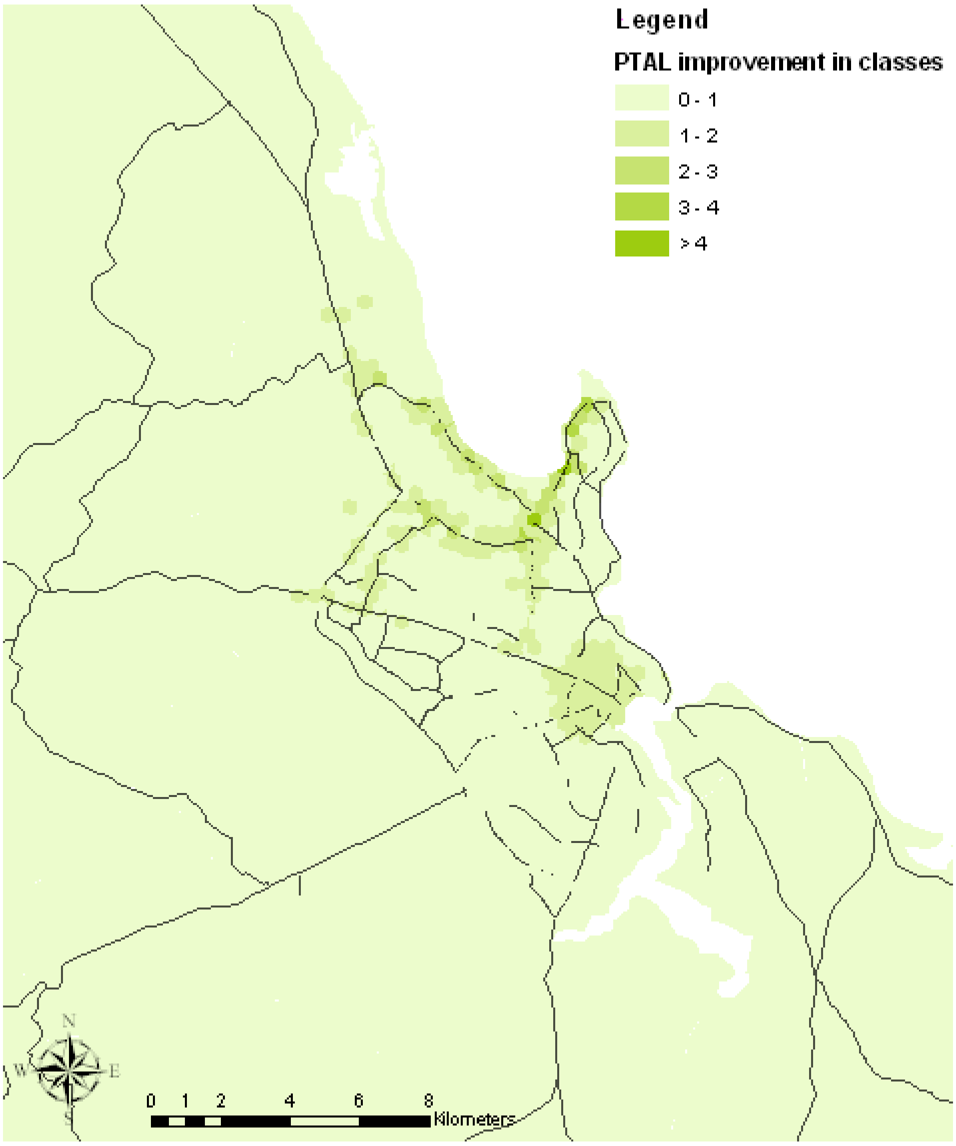

figure 5.3: Improvement in PTAL after DART phase I

5.1.2. Corrected PTAL

From figure 5.3 it becomes clear that the PTAL itself might be a distorted image, because adding new routes improves the PTAL level, but does not improve the quality and accessibility of public transport. Therefore the CPTAL was introduced which corrects the calculated PTAL with the population that can be reached within one hour travel time (see also: section 2.1.1. Availability and especially equation (2.4)). This means that if a new route is introduced the CPTAL increases more when this route services a area that wasn't served before than when the route services an area that was already served.

In figure 5.4 the CPTAL for the DART phase I is depicted.

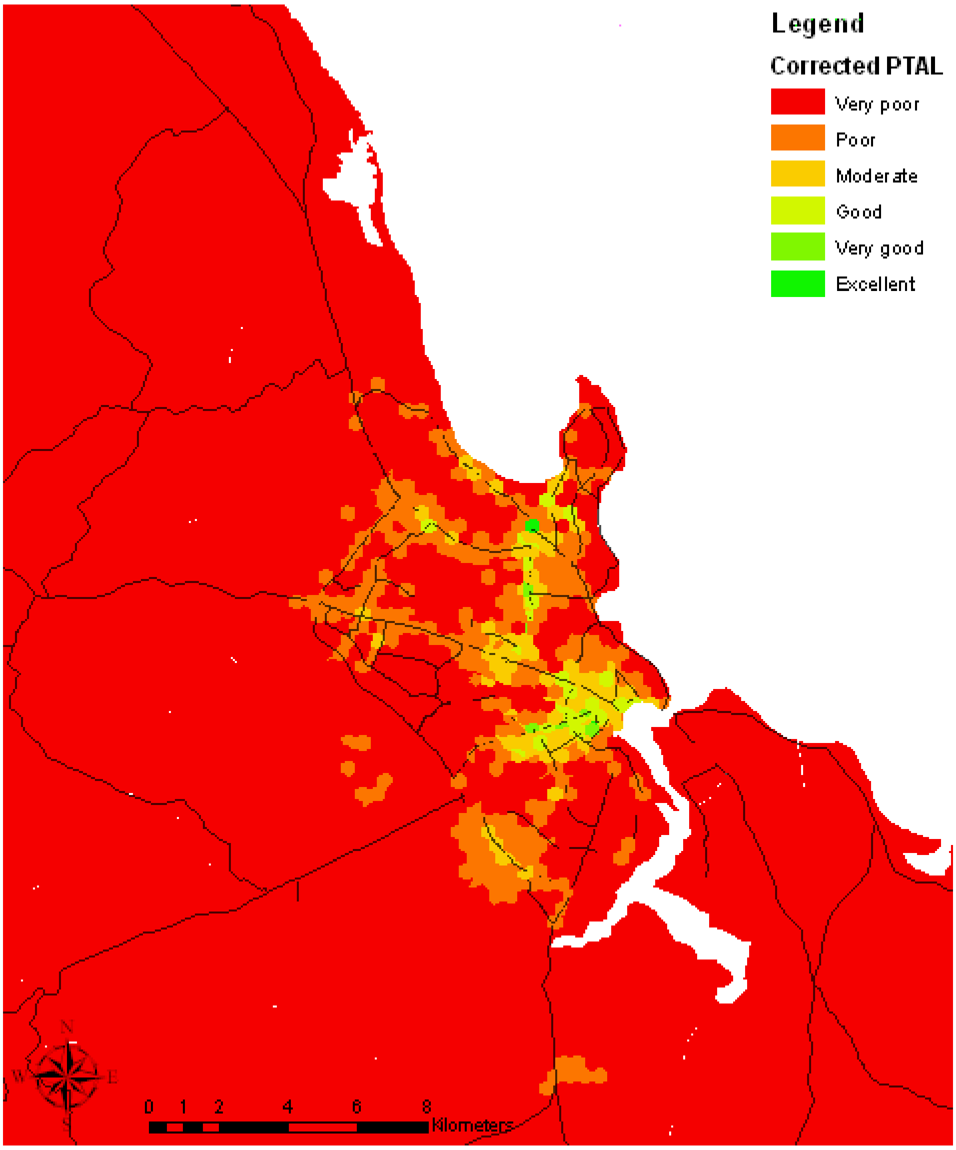

figure 5.4: Corrected PTAL after DART phase I

It is clear that the quality of accessibility is suddenly a lot lower. However it also decreases the differences between adjoining zones and gives a more realistic image.

5.1.3. Conclusions

From figure 5.2 and figure 5.3 it becomes clear that in the Msasani and Masaki area the accessibility improves significantly after the implementation of DART phase I. This is merely because of the introduction of new feeder lines, with very high frequencies. However this research also indicates that the CPTAL approach that is discussed, might give far more valuable information than the original PTAL approach, especially in areas where there are not that many public transport routes.

5.2. Travel Time Assessment

The travel time assessments are performed using an optimization for (perceived) travel time. This means that people will shift routes, if that improves their travel time, no matter how much the travel fare would be. Since a travel time assessment from all origins to all destinations would take too much time (this would mean calculating over 13 million routes), the destinations are limited to seven important locations in Dar es Salaam. The selected locations are:

- ❖ Kariakoo Market
- ❖ Kivukoni Ferry
- ❖ Muhimbili Hospital
- ❖ Mwenge
- ❖ Old Posta
- ❖ Ubungo Terminal
- ❖ Tandika

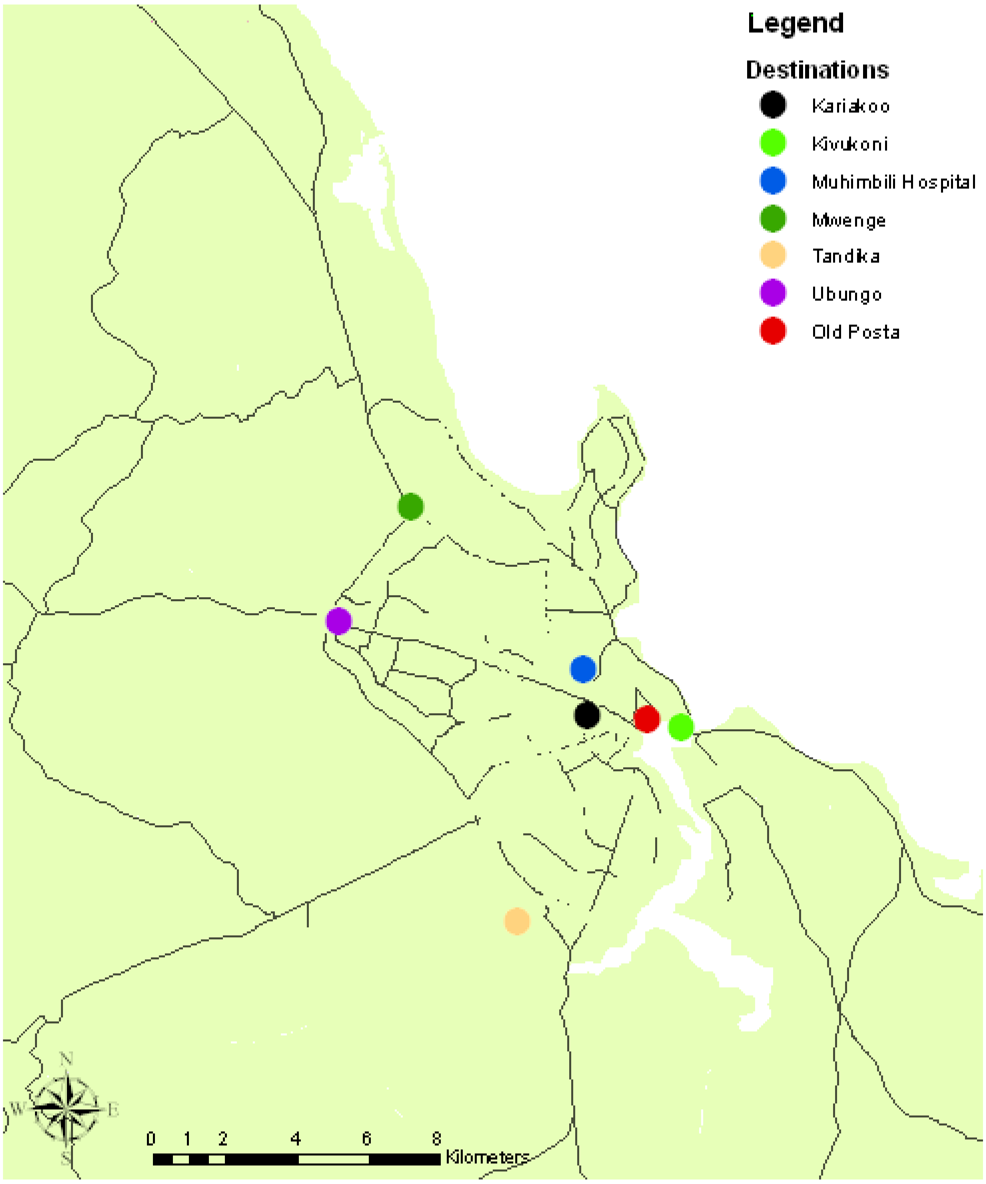

figure 5.5: Destinations for analysis

Appendix E: Appendix E: Creating Travel Time and Travel Cost Assessment Layers described the method that was used to perform the analysis and create the figures below.

5.2.1. True travel time

When analyzing the travel times, the travel times toward the seven different locations as mentioned before are averaged, in order to get one figure (see figure 5.6 for the greater Dar es Salaam analysis) instead of seven separate ones. This makes comparing the different situations more convenient. For some reason the analysis didn't work out for some area's. Though effort has been put into solving this problem, the solution hasn't been found. In the figures below these area's are clearly marked by the *No Data Available* label.

This analysis optimizes travel time regardless of other potential costs (for instance additional travel fares that are caused by switching between routes). This analysis therefore shows the potential travel time gain that could be realised if people would have an unlimited budget. The analysis can therefore be described by equation (5.2), which is similar to the one in section 2.1.2.

$$t_T = t_a + t_w + t_{iv} + t_t + t_e \tag{5.2}$$

This equation emphasises the fact that user costs are not part of this analysis.

From figure 5.6 it becomes clear that the analysis outside the city centre is quite arbitrary. Because the locations of roads are unclear (and therefore bus stops are located where no roads are drawn) some people will always walk to a road (which could take much longer to reach a bus stop) instead of walking to a bus stop that might be 100 meters further, but is not connected to a road. Therefore the further analysis will aim at the 'city centre', meaning the area where more information is known about roads and public transport routes.

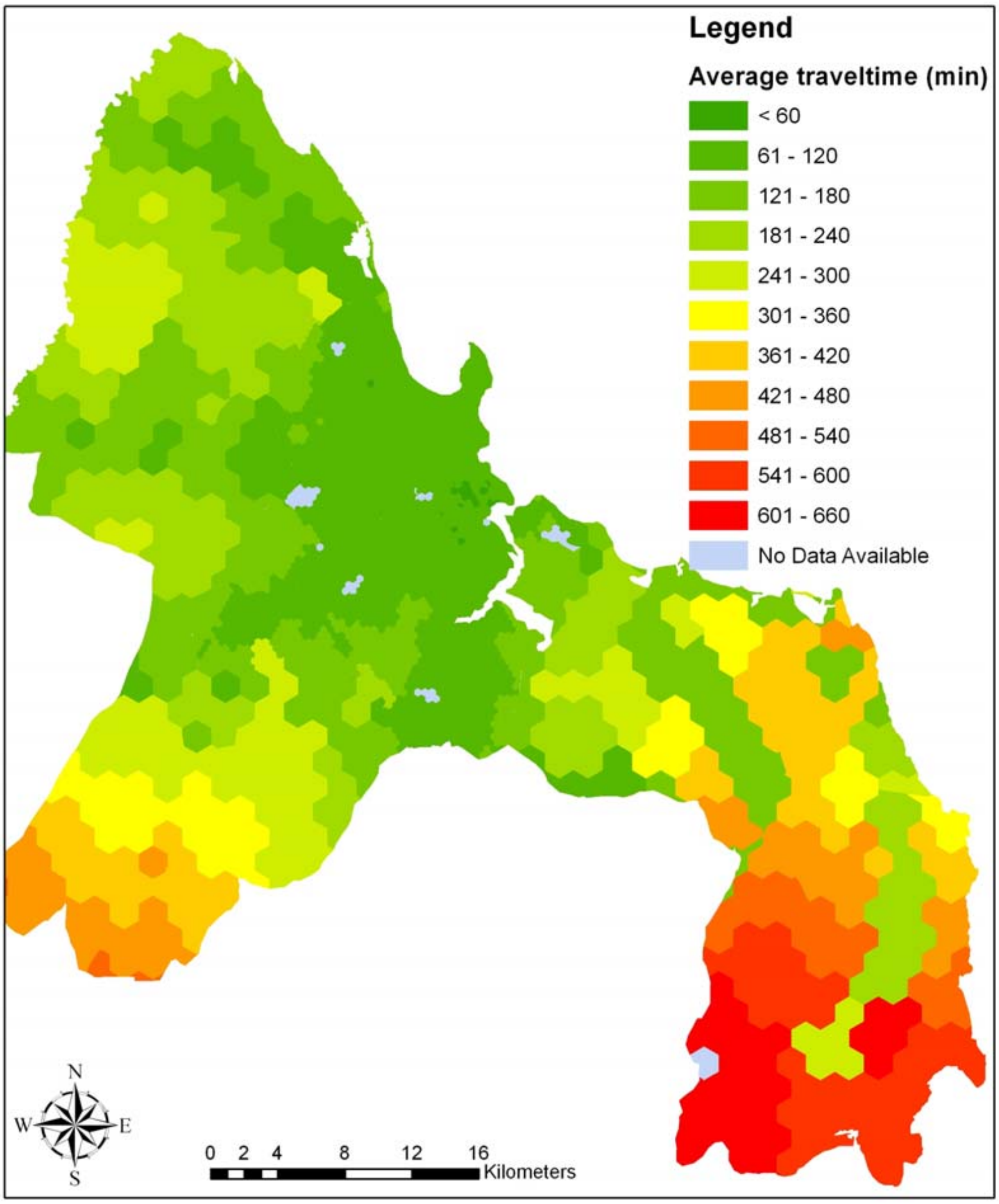

figure 5.6: Average travel time towards seven locations for Dar es Salaam for daladala system

In figure 5.7 the average travel time for the central part of Dar es Salaam is shown (note the different scale) and this figure shows that the estimations are better in this region, since the colours do not differ very much in most area's (except the Tegeta region). However from this figure it appears that it easily can take up to an hour-and-a-half to reach the important destinations.

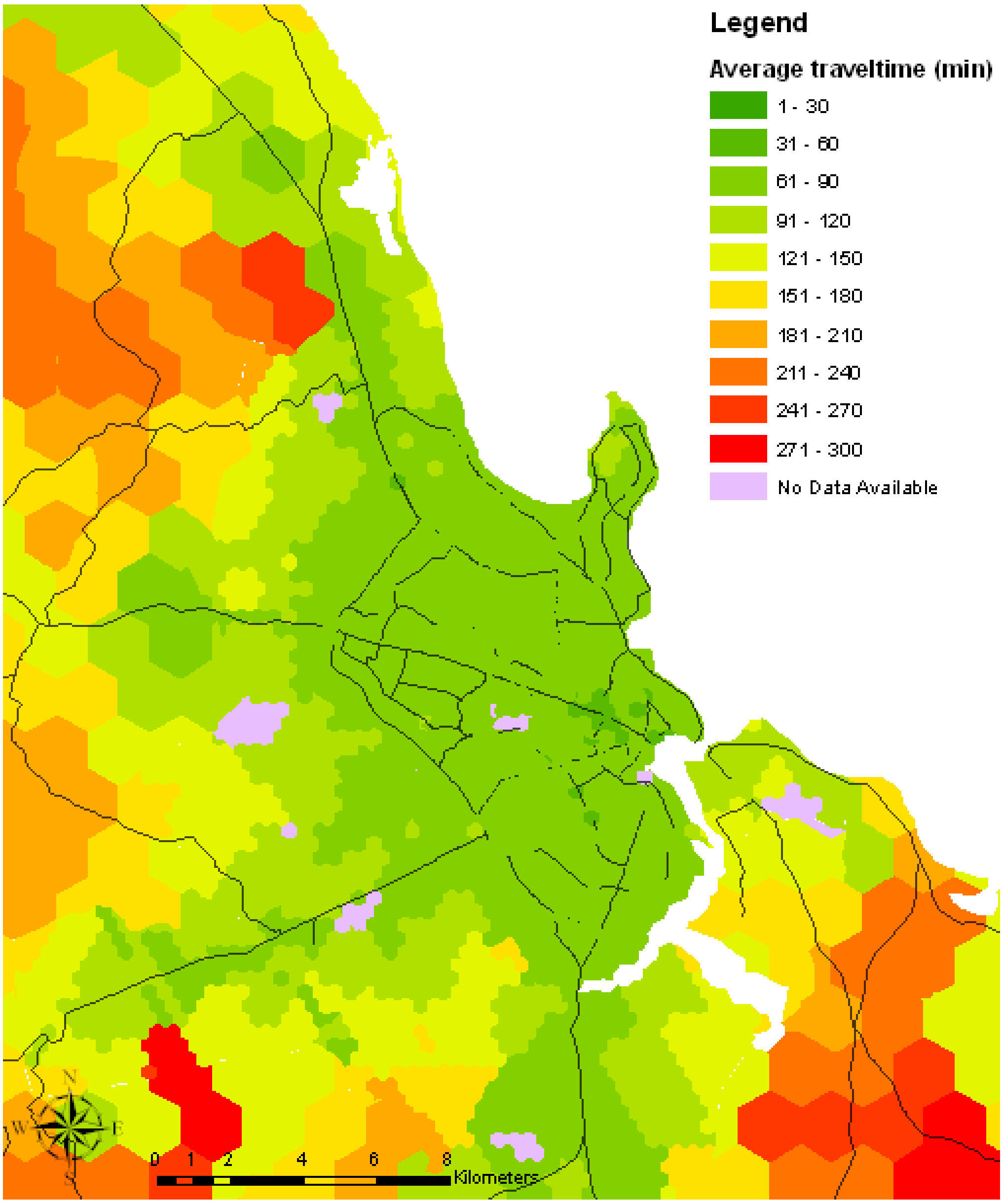

figure 5.7: Average travel time towards seven locations for daladala system

The same figure was created for the DART phase I system, which is depicted in figure 5.8. This figure looks more or less the same, but the lowest average travel time is now below 30 minutes, where in figure 5.7 the lowest average was over 30 minutes.

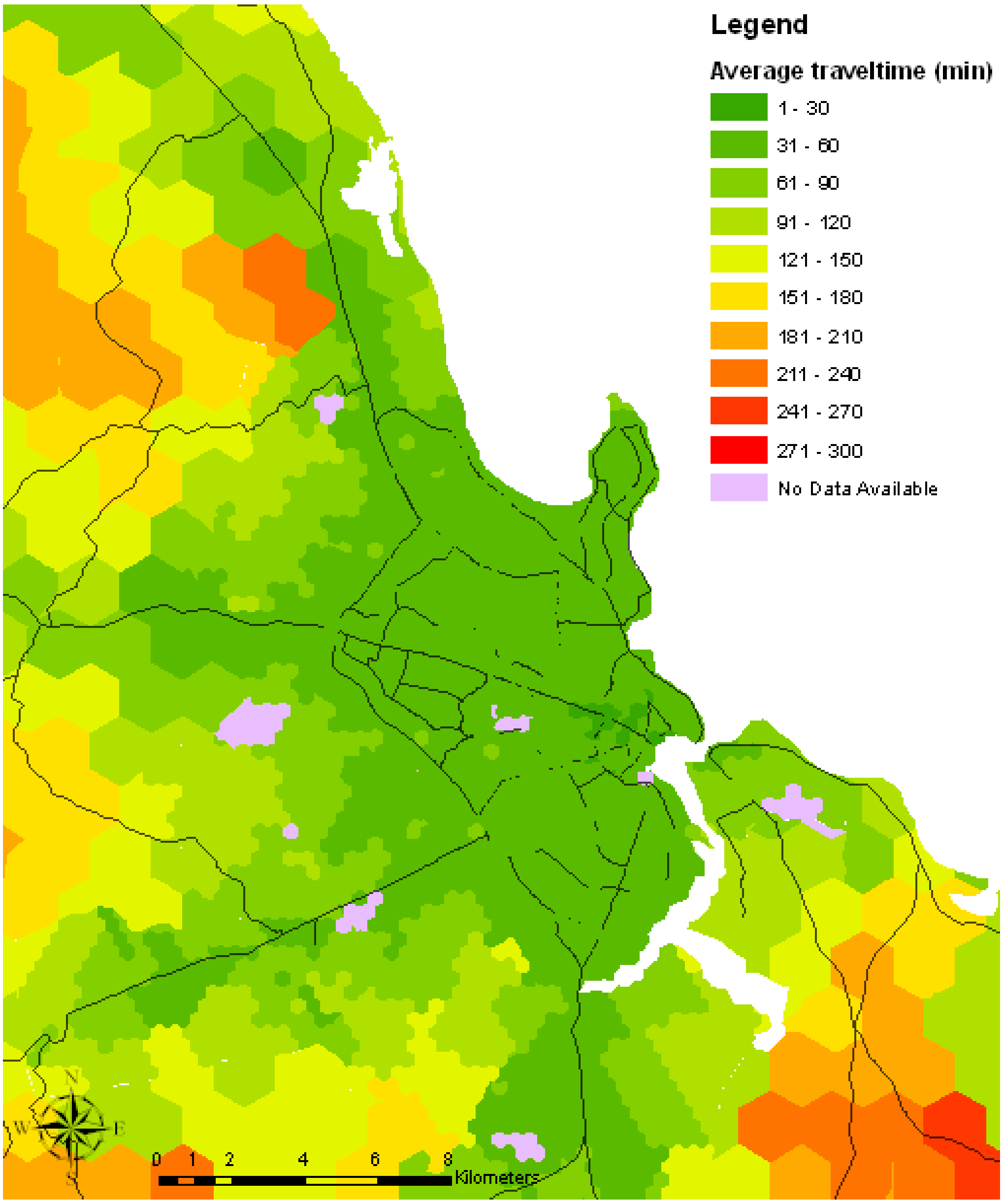

figure 5.8: Average travel time towards seven locations for DART phase I

From looking at figure 5.7 and figure 5.8 it appears that implementation of DART phase I could improve travel time for nearly everyone with about 30 minutes. This potential travel time improvement is depicted in figure 5.9 which indeed shows that for most of the area the travel time could indeed improve with 20 to 40 minutes.

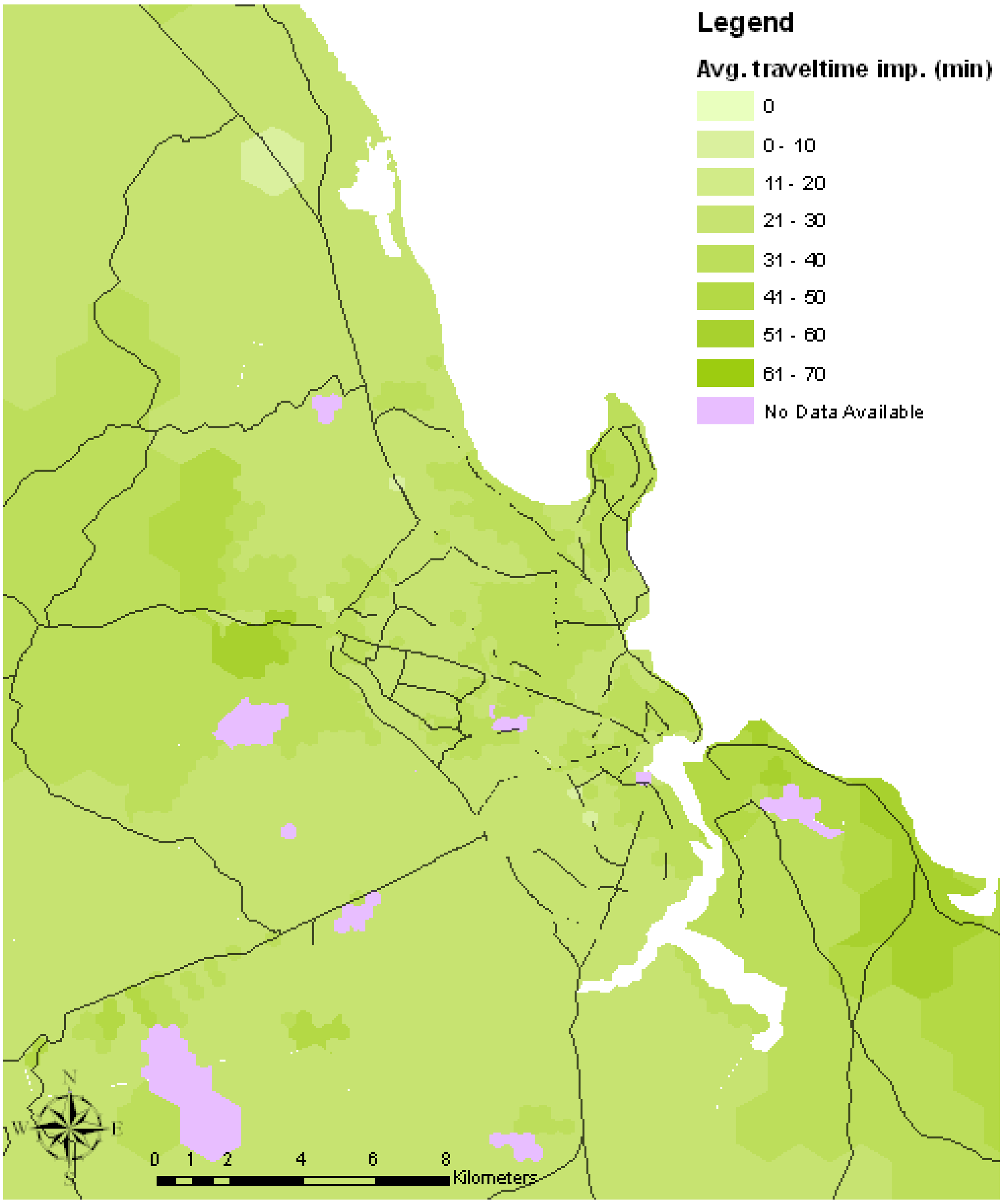

figure 5.9: Improvement in average travel time after DART phase I

5.2.2. Perceived travel time

Perhaps even more important than the true travel time, is the travel time as perceived by the public. Therefore the access-, egress-, waiting and transfer time are weighted in the travel time calculation, according to equation (5.3) which is the same as equation (2.7). It is more likely that people would use perceived travel time to decide which route they choose, so this assessment might be even more important.

$$t_p = t_{iv} + 2.0 \cdot (t_a + t_e) + 2.5 \cdot (t_w + t_t) \quad (5.3)$$

The perceived travel time is significantly higher than the true travel time, especially in the more remote areas. This becomes very clear when one compares figure 5.7 and figure 5.10. This makes sense, because the further one lives from a public transport stop the longer the access time. Also is it very likely that frequencies are lower in remote area's thus increasing the waiting time.

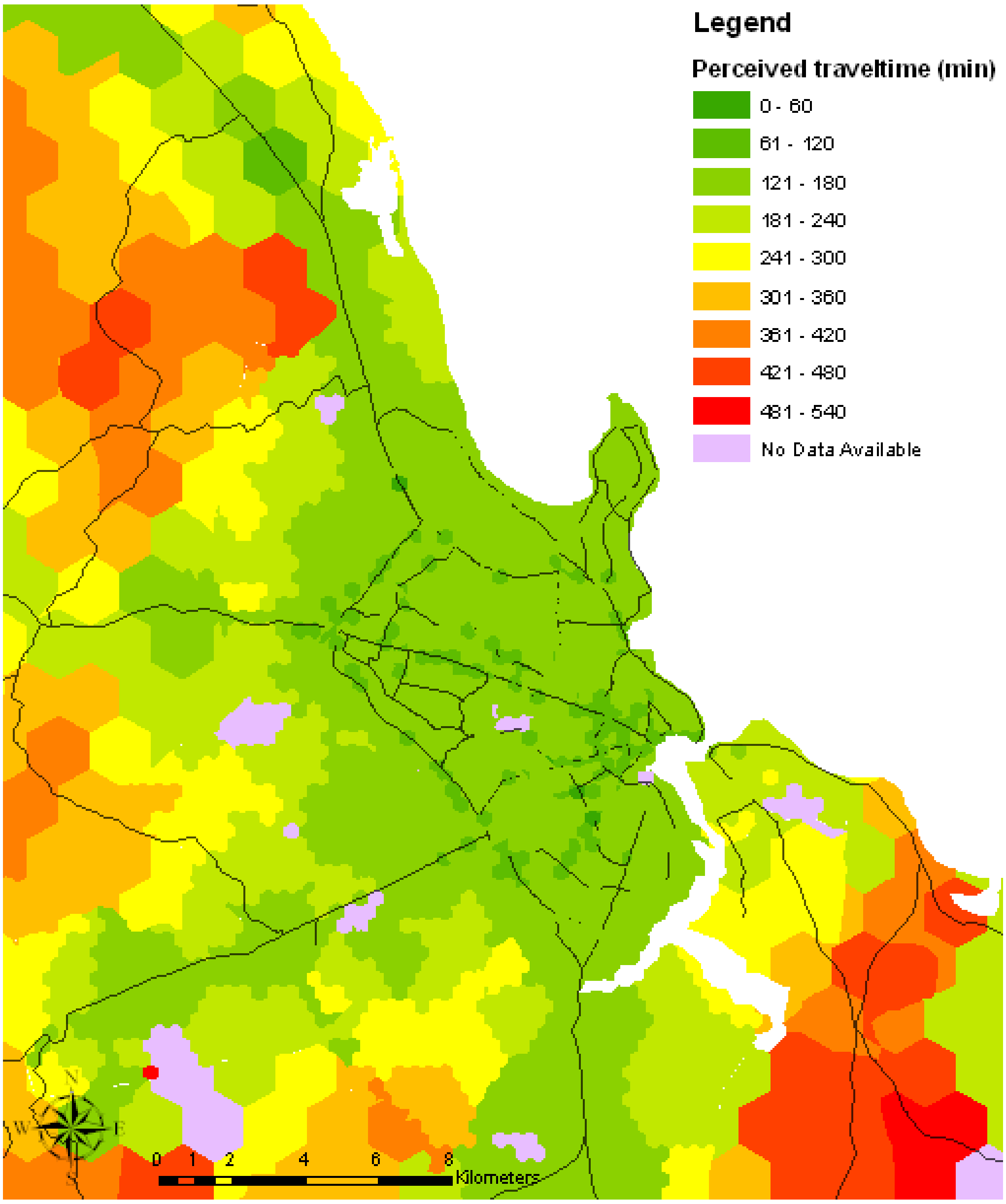

figure 5.10: Average perceived travel time towards seven destinations for daladala system

The DART phase I has often higher frequencies than the current public transport system, thus significantly decreasing the perceived travel time (since waiting is valued higher than the in vehicle travel time). This becomes clear when one looks at figure 5.11 which indicates that a lot more areas have a perceived average travel time under 60 minutes.

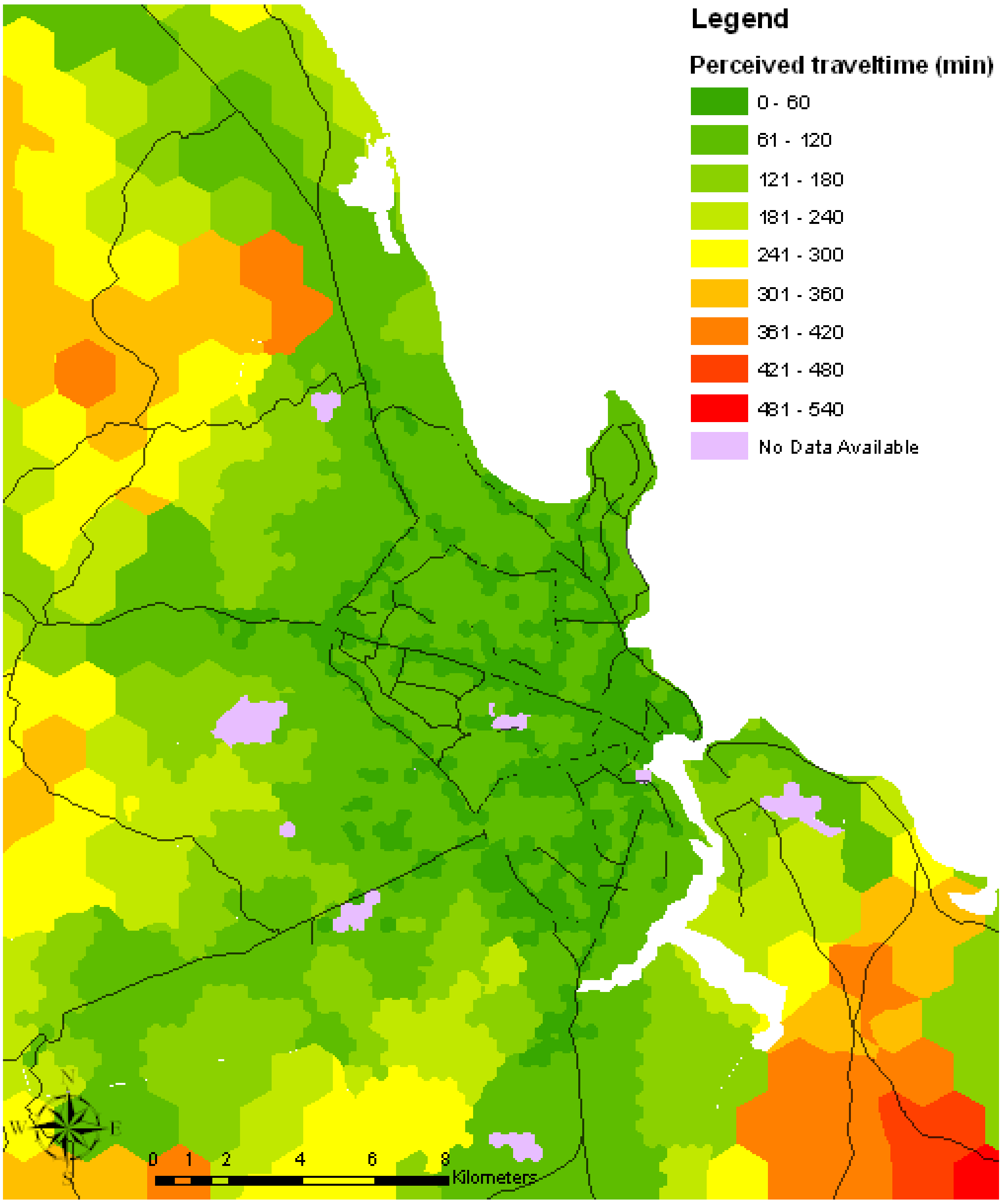

figure 5.11: Average perceived travel time to seven destinations for DART phase I

Ffigure 5.12 supports the assumption that DART phase I significantly improves the perceived travel time. Nearly all areas have an improvement of 45 up to 90 minutes. Especially the areas where the new feeder routes are planned improve a bit more and the same goes for the area around Ubungo terminal, which gets a very fast connection to the city centre.

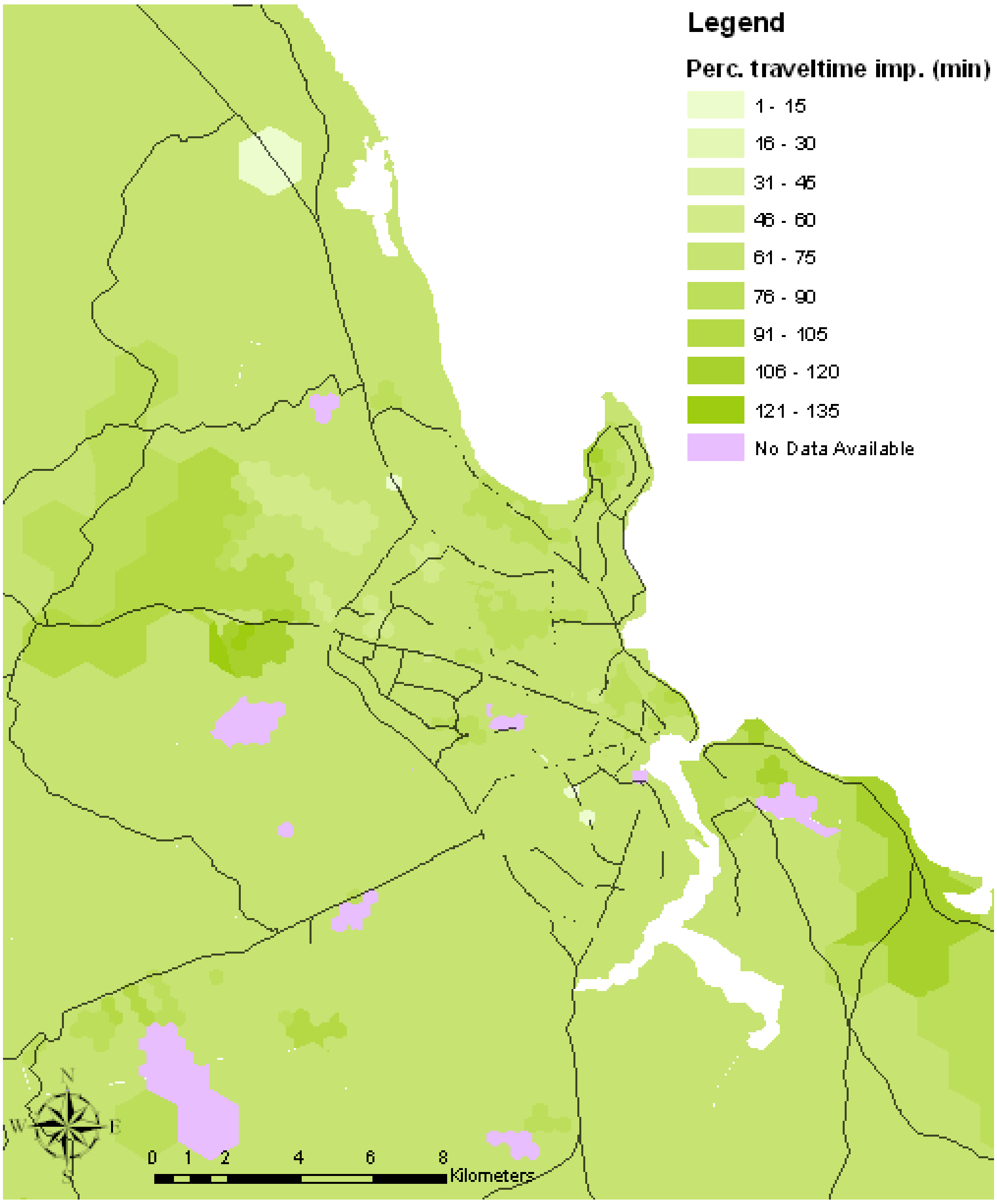

figure 5.12: Improvement in average perceived travel time after DART phase I

5.2.3. Conclusions

When looking at both travel time analysis it becomes clear that nearly every citizen of Dar es Salaam could improve his or her travel time, as long as they are willing to spend some money to change routes every now and then. Especially the area around Ubungo and Kigamboni (and also Msasani and Masaki) improve quite significantly which could indicate that these areas become more attractive in the coming years.

5.3. User Cost Assessment

The user cost assessment is done using three different criteria. The first is the true cost that is incurred due to paying travel fares, the second uses the generalized cost method and the third uses perceived generalized cost. Again only seven destinations (as described in paragraph 5.2.) are used, because a full Origin-Destination analysis would take too much time.

5.3.1. Travel cost

When one would try to reduce travel cost to a minimum, the solution would be simple: everyone walks everywhere, since no monetary costs are involved. Therefore the travel cost that are related to the routes that were calculated by the perceived travel time analysis (section 5.2.2.) are used to analyse the potential additional monetary cost that are related to the potential travel time improvement that could be realised. (see also section 5.2.1. up to section 5.2.3.)

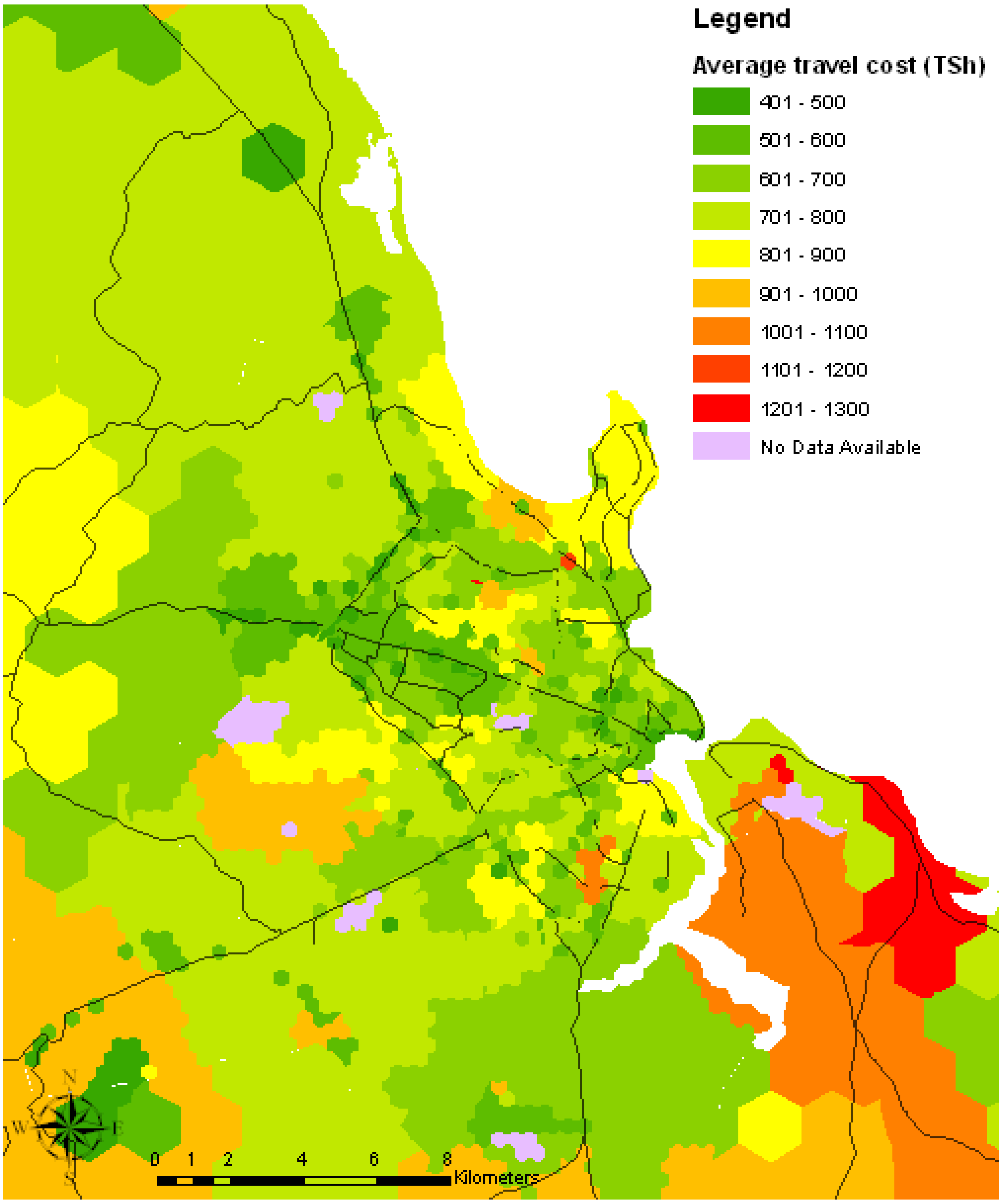

figure 5.13: Travel cost after perceived travel time optimization for daladala system

From figure 5.13 it becomes clear that there is some error, which exhibits itself by the large green areas in the south-west corner. This is probably a flaw which is caused by the fact that there are routes digitized that have no information about the routes attached, which causes problems when modelling.

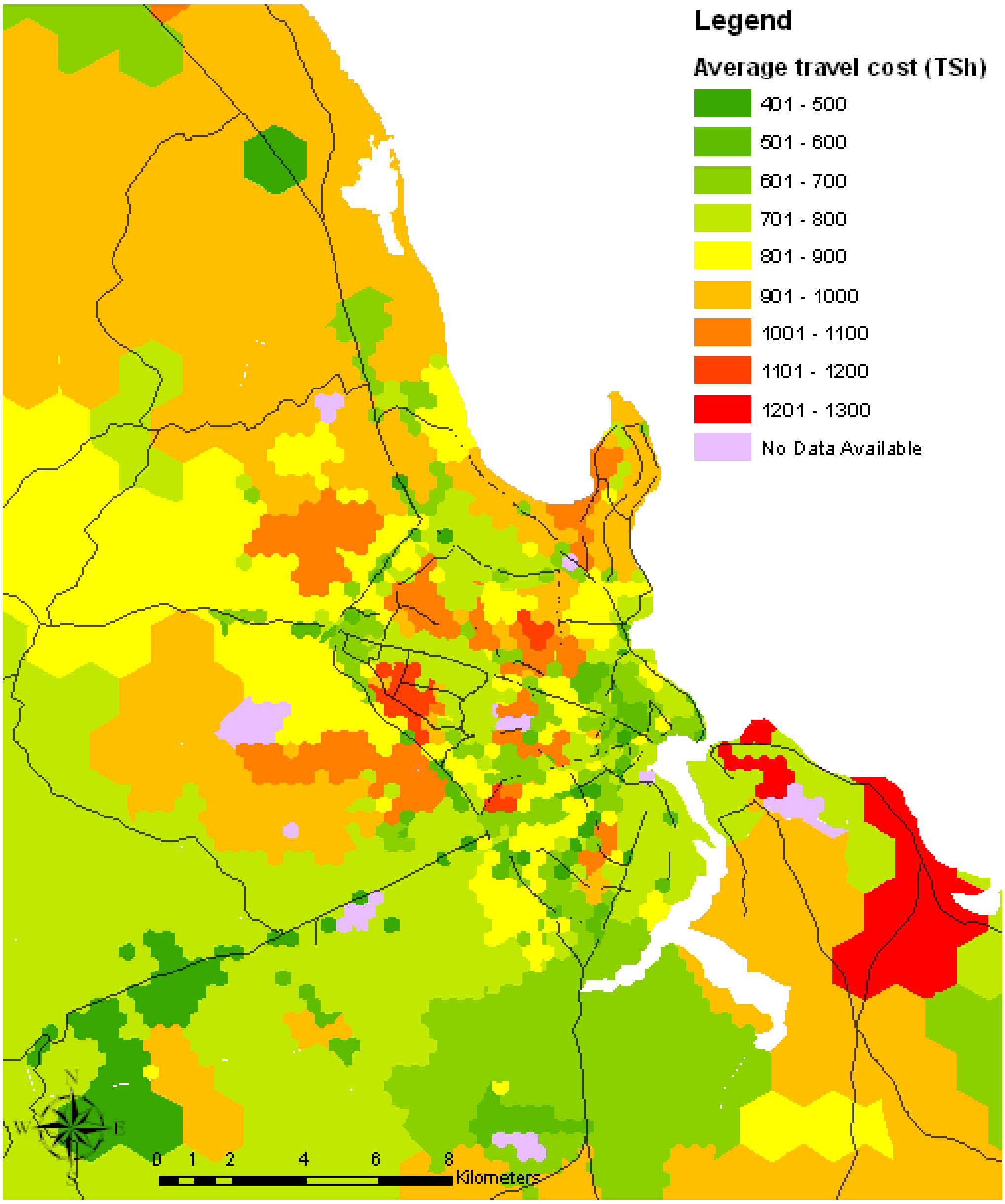

figure 5.14: Travel cost after optimization for perceived travel time after DART phase I

Around the corridors there are areas that suddenly pay a lot more fares if they try to optimize for perceived travel time. This makes sense, since they will use the new mode of public transport, even if it costs more than the travel time benefits they incur.

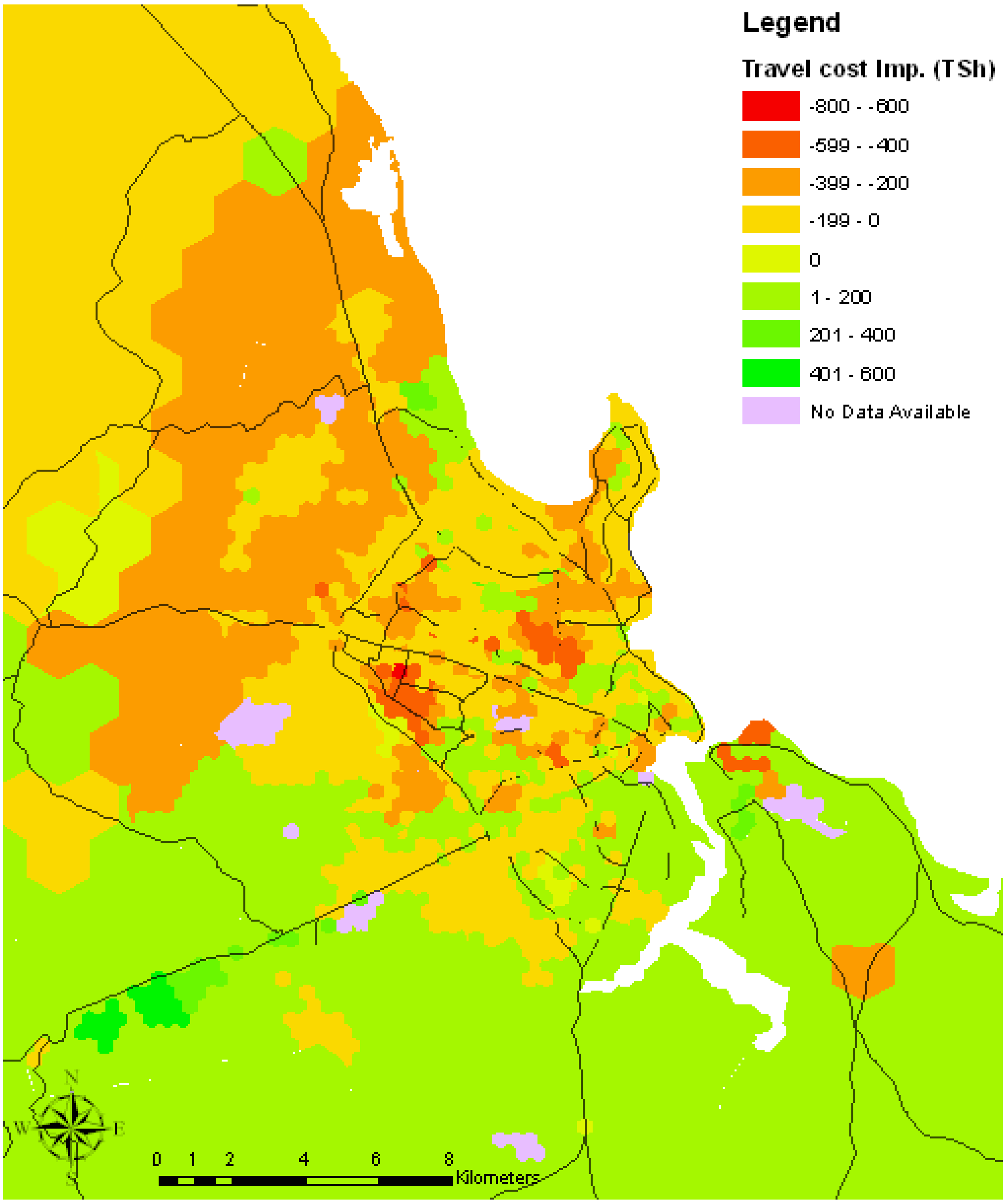

figure 5.15: Improvement in travel cost after DART phase I

It is clear that people that are willing to reduce their travel time would have to pay some money for it, which is about 200 TSh. It also indicates that the areas that are not yet served by DART for the time being benefit from the changes, but this could also be caused by people not using public transport for the last part of their trip, because the routes changed.

5.3.2. True generalized travel cost

The true generalized cost exists of the travel fares and the value that people give to the time they are in transit. To calculate the GTC equation (5.4) is used (which is identical to equation (2.12) which is explained in section 2.1.3.).

$$GTC_i = \frac{t_T}{VOT_i} + C_{f;T} \tag{5.4}$$

So in order to determine the GTC for a certain trip, for each link that is used the assigned travel time is weighted. The sum of all these weighted travel times and travel fares that incur, is the GTC.

From figure 5.16 it becomes clear that, on average, a trip to any of the seven selected locations costs about 1000 TSh. This value represents for instance a 200 minute trip with no fares, or a 100 minute trip with a travel fare of 500 TSh.

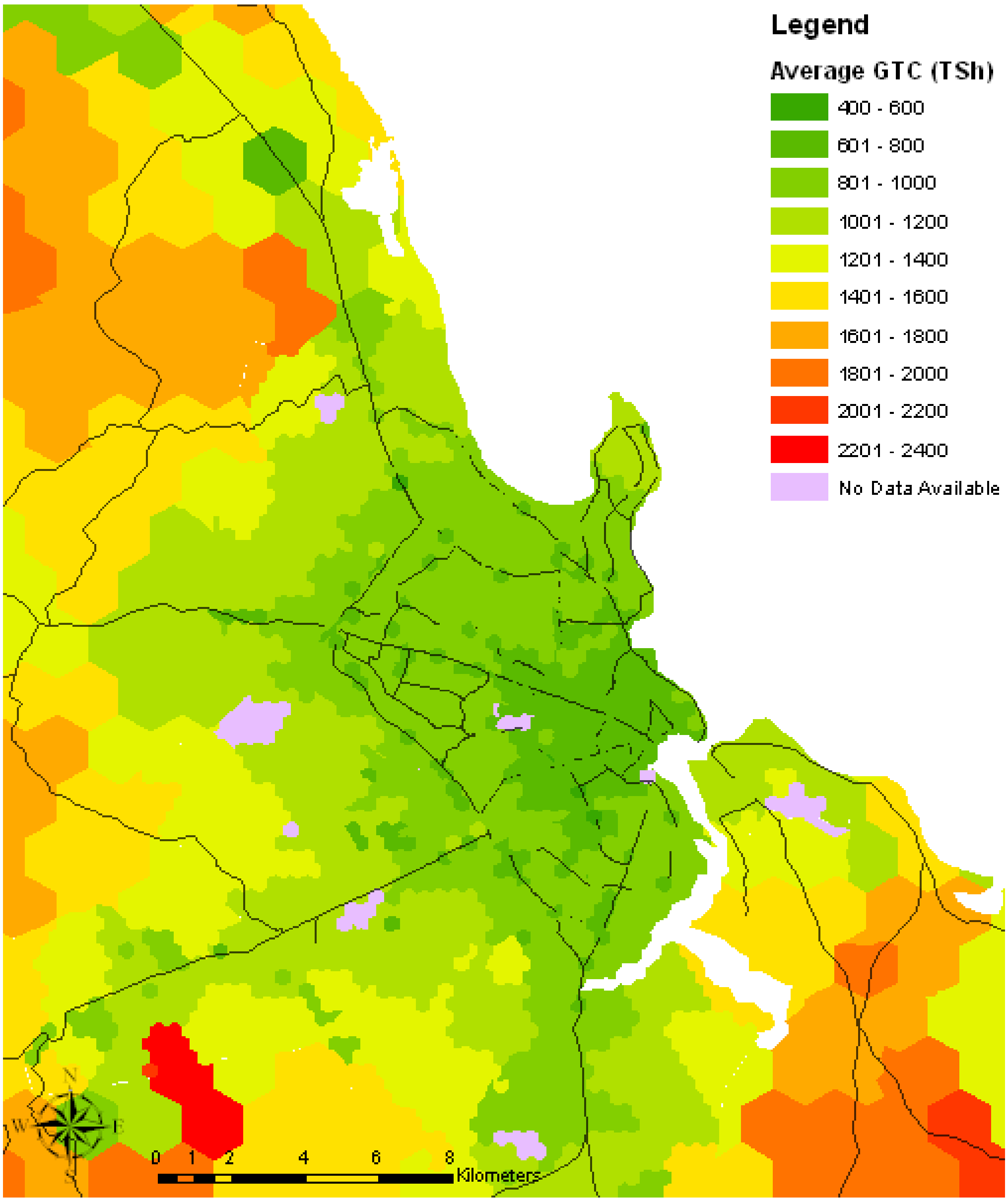

figure 5.16: Average GTC to seven destinations for daladala system

When figure 5.16 and figure 5.17 are compared, it becomes clear that for the area's in the city centre the improvement in GTC is very small, since the colours haven't changed.

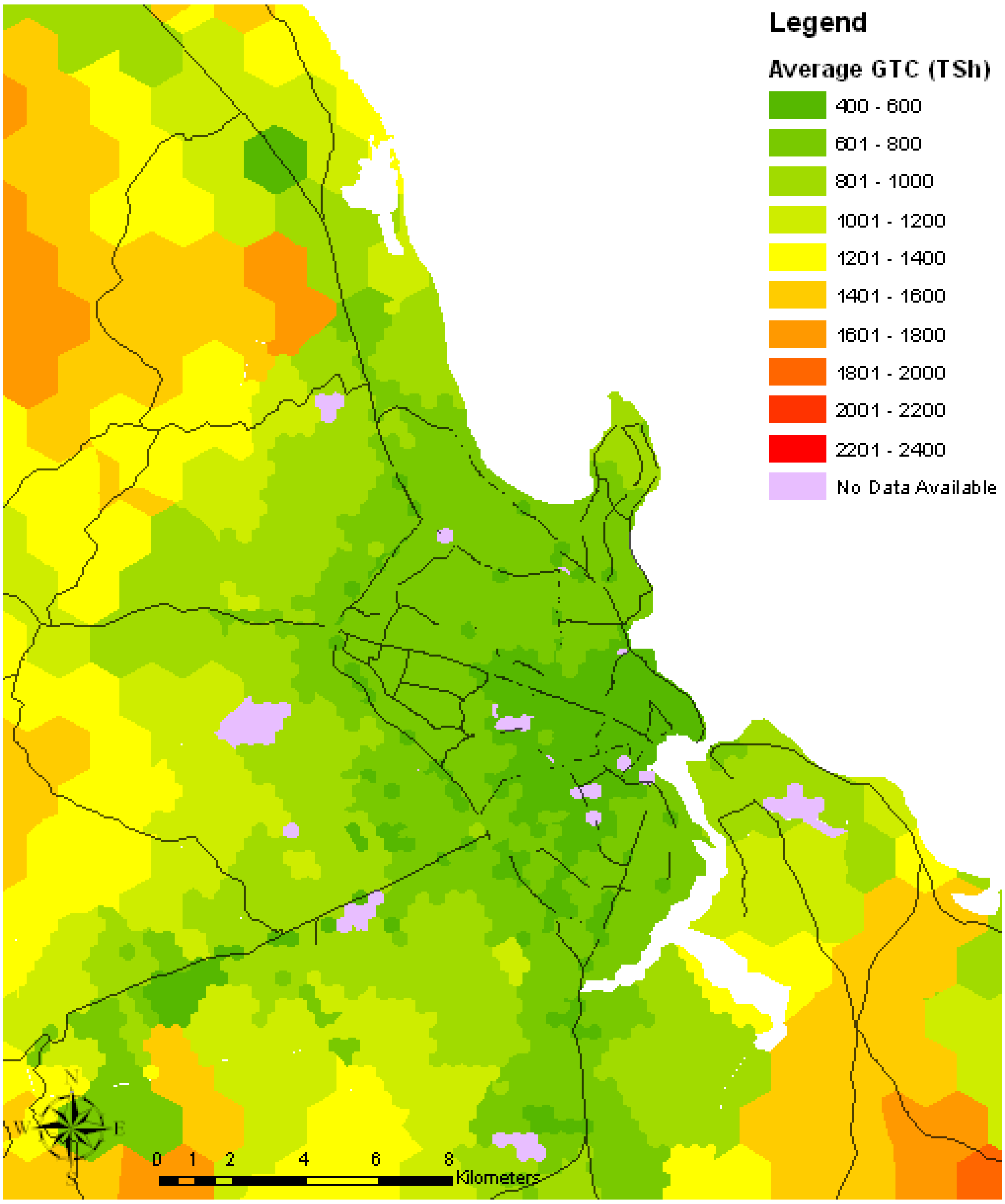

figure 5.17: Average GTC to seven destinations for DART phase I

The conclusions from above are confirmed by figure 5.17, as it shows that the improvements for most areas are about 50 TSh up to 200 TSh, which is about 10% of the median daily income in Dar es Salaam.

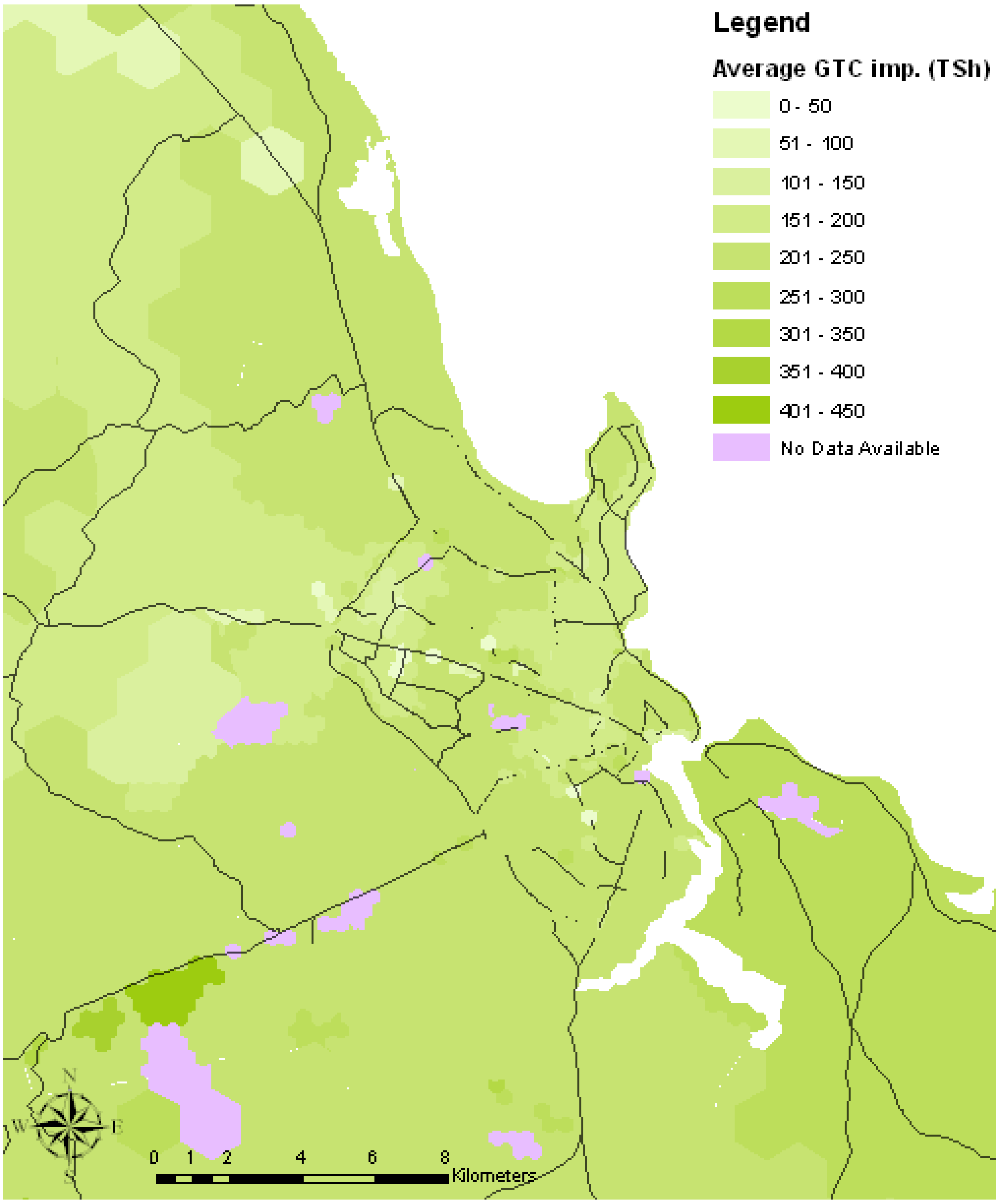

figure 5.18: Improvements in average GTC after DART phase I

5.3.3. Perceived generalized travel cost

The perceived travel cost is based on the same principle as the generalized travel cost, but are based on the perceived travel time. So in this case the travel time is calculated using equation (5.5) (which is identical to equation (2.7)) and the PGTC can then be calculated using equation (5.6) (which is in turn equal to equation (2.13)).

$$t_p = t_{iv} + 2.0 \cdot (t_a + t_e) + 2.5 \cdot (t_w + t_t) \tag{5.5}$$

$$GTC_{p;i} = \frac{t_p}{VOT_i} + C_{f;T} \tag{5.6}$$

When the PGTC (figure 5.19) and the GTC (figure 5.16) are compared, it appears that in the city centre the differences are small, while in the more remote areas the differences are large. This makes sense, because in the city centre the travel time is small and does not consist of long waiting times and transfers, whilst in the remote areas transfers and long waits are quite common.

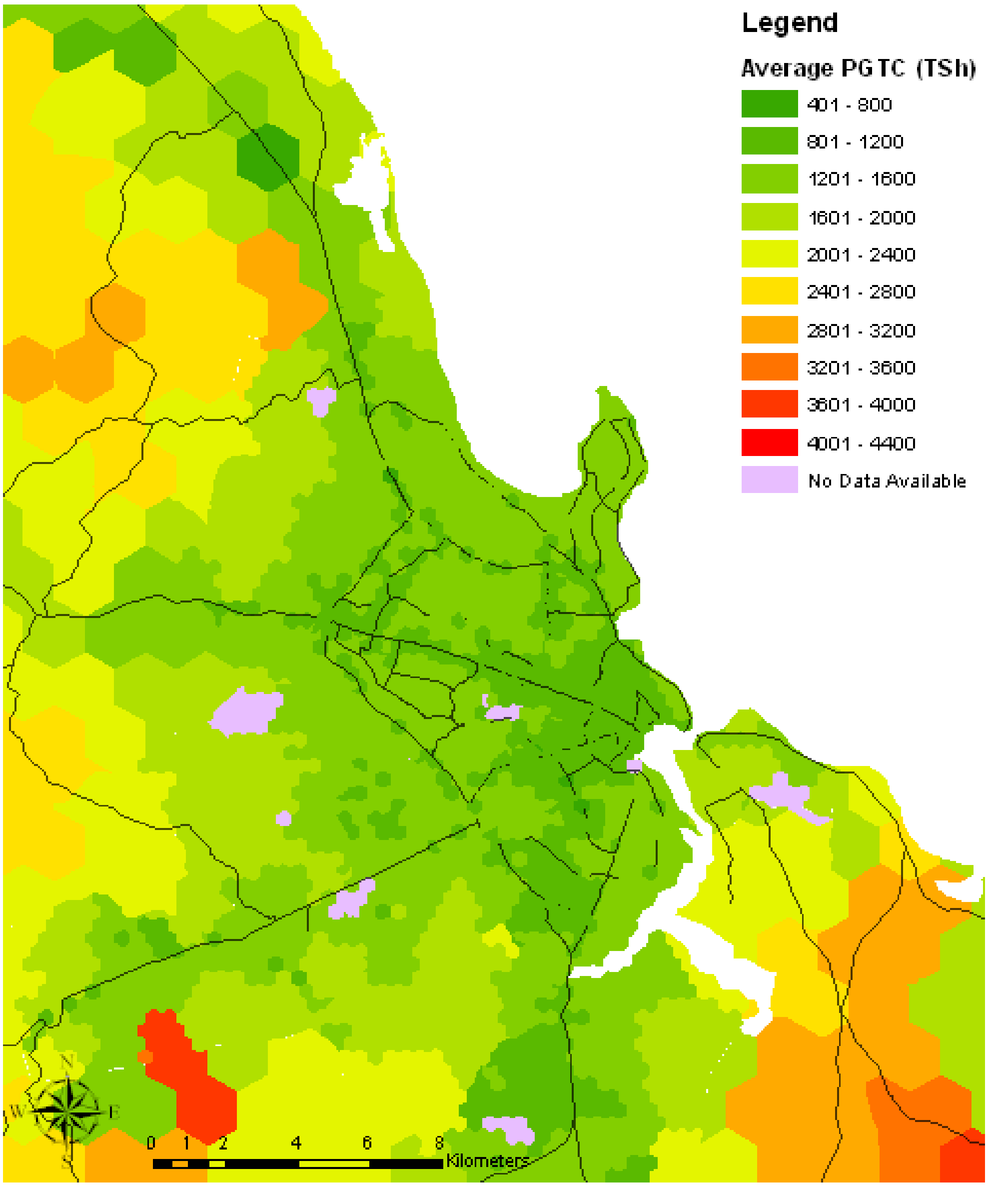

figure 5.19: Average PGTC to seven destinations for daladala system

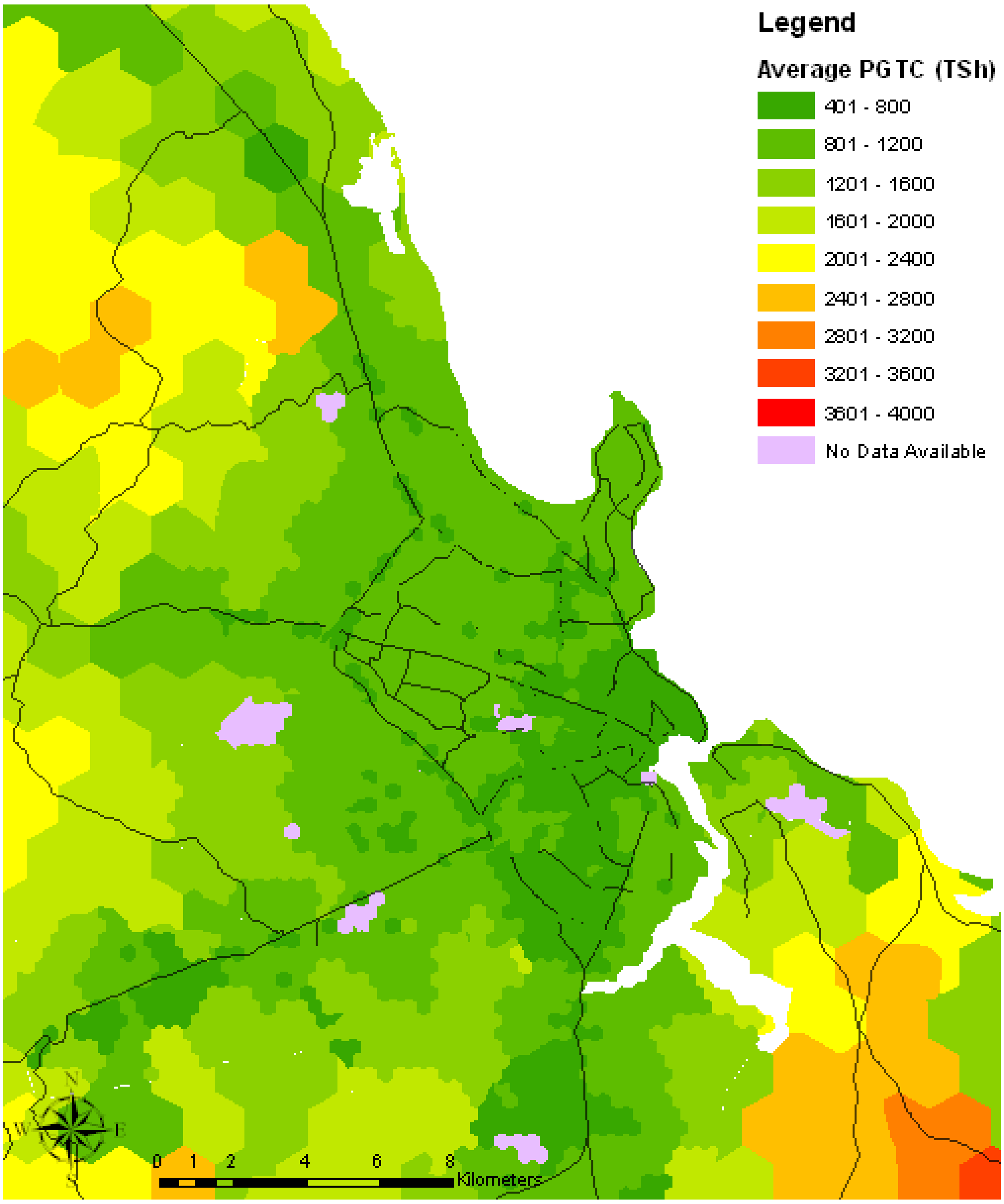

figure 5.20: Average PGTC to seven destinations for DART phase I

In figure 5.21 the improvement in perceived generalized cost is depicted. The areas close to the main daladala routes don't improve much, but around the feeder routes and the more remote areas the improvement in PGTC is about 300 TSh.

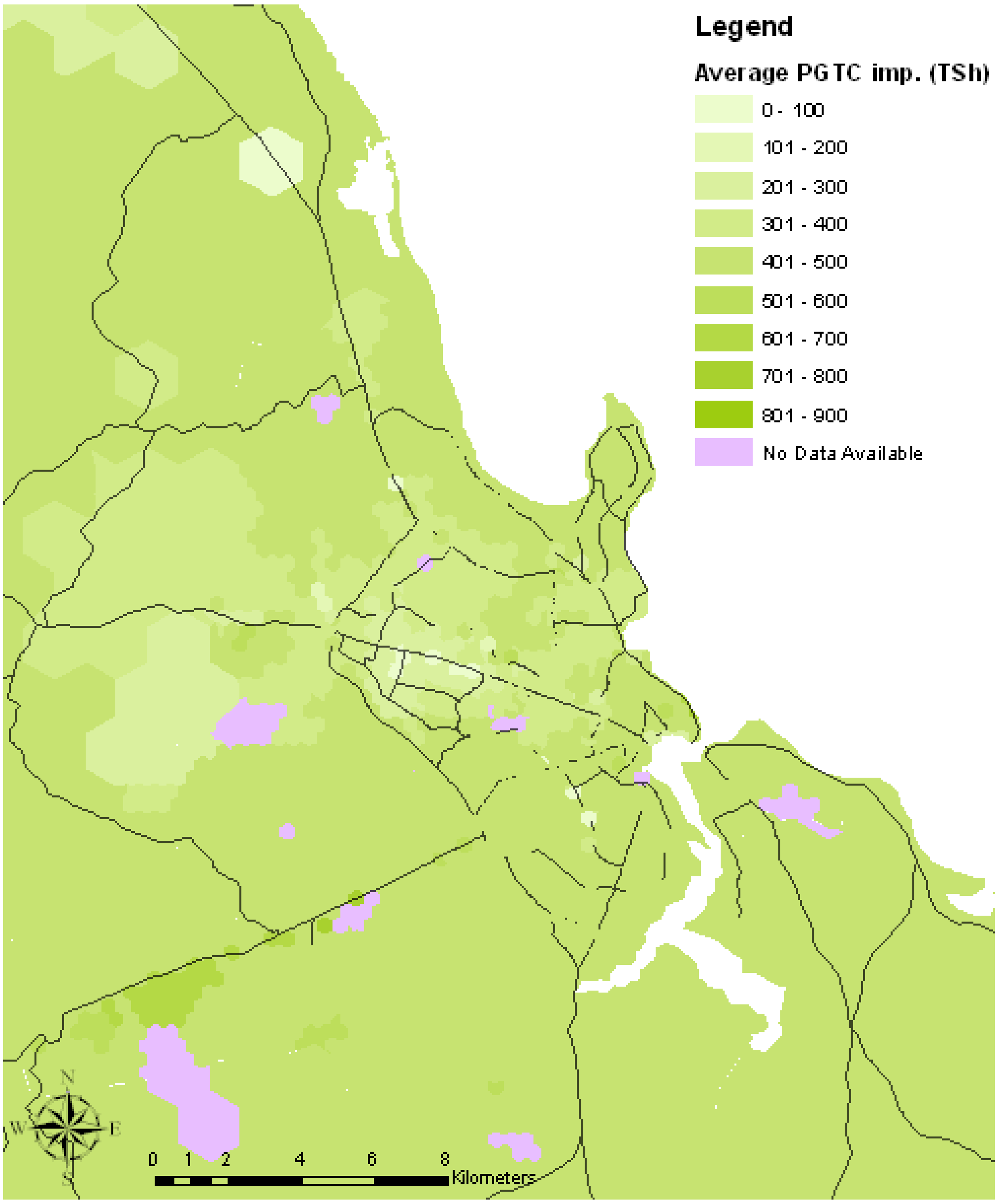

figure 5.21: Improvements in average PGTC after DART phase I

5.3.4. Conclusions

Both the generalized travel cost and perceived generalized travel cost improve for most areas, but with very small amounts. It is however quite remarkable that around the main corridors the GTC doesn't improve as much as in the more remote areas. This could however be explained because the access time around these areas is much smaller than the access times for the more remote areas.

5.4. Conclusions

For nearly every area the travel time and generalized travel cost improve, although the GTC doesn't improve with the amount one would expect based on the travel time improvement. It can therefore be concluded that (based on the value of time of 5 TSh/min) the travel time improves significantly for nearly everyone, but that this travel time improvement is largely nullified by the additional (or higher) fares.

It seems that this analysis can give insights in the effects that new urban public transport has on the lives of people, however to make a good evaluation a good value of time estimation and origin – destination matrix is needed. This analysis therefore proves that the assessment in itself is possible, but is highly dependent on the input variables, this became clear when a value of time of 7.5 TSh/min was used, the results were very ambiguous.

It seems especially interesting to do a more detailed research into the creation of a CPTAL approach that truly represents the accessibility for certain areas. This research proposed one (very simple) solution, but it is very interesting to see which other factors could be included to define accessibility.

6. Conclusions & Recommendations

This research shows that it is possible to use a GIS software package (with extensions) to model a local public transport system. Using network analysis information can be stored on different levels to be aggregated to the required level. When using common sense and local information, it appears that the results that are produced by the model when using the GTC and/or PGTC analysis give a realistic and true image of the situation.

The analysis that was performed on the proposed public transport network in Dar es Salaam, Tanzania, suggests that the proposed system improves both travel time and generalized travel cost. This goes for both the true as the perceived analysis. It is clear that the areas that are serviced well right now doe not benefit as much from the new public transport system as the more remote areas or the areas where high frequency lines are introduced (e.g. Msasani en Masaki).
However it also appears that the main roads are serviced very well, but e.g. areas like Sinza or Kijitonyama do not benefit as much as could be done with a high quality feeder system. It might be a good idea to look into this matter and perform a sensitivity analysis.

This research suggests that GIS analysis can prove valuable in assessing and planning local public transport. However to make the assessments more reliable, better information is needed about value of time, car possession, waiting times per route, etc. Some information is gathered at this time by JICA, some information can only be obtained by an in depth investigation by both SUMATRA and DCC.

In order to be able to process the data that is gathered, it is strongly suggested that more people get familiar with the possibilities of GIS and geographical information. Also investments in equipment (e.g. GPS devices) and data (e.g. aerial photographs) are needed to get a solid base for further analysis and research.

Last, it would be a good idea, to do an extensive research into what people define as accessibility and to find a way to model this using either GIS or transport software. This because it seems that accessibility is a very important term to assess how people perceive public transport.

7. References

Baradaran, S. & Ramjerdi, F. (2001) Performance of accessibility measures in Europe. *Journal of Transportation and Statistics, 4*(2/3), 31-48.

Belal, W. (2002). Transport data models [Presentation]. Enschede, The Netherlands: International Institute for Geo-information Science and Earth Observation.

Caliper Corporation. (2001) TransCAD. Retrieved September 19, 2007, from www.caliper.com

Currie, G. (2005) The demand performance of Bus Rapid Transit. *Journal of Public Transportation, 8*(1), 41-55.

Daft, R. L. (2003) *Management*. Mason, OH, USA: Thomson Learning.

Dimitriou, H. T. & Banjo, G. A. (1990) *Transport planning for third world cities*. Londen, UK: Routledge.

ESRI. (2005) ArcGIS network analyst. Retrieved September 19, 2007, from www.esri.com/networkanalyst

Fiorenzo-Catalano, S., van Nes, R. & Bovy, P. H. L. (2004) Choice set generation for multi-modal travel analysis. *European Journal of Transport and Infrastructure Research, 4*(2), 195-209.

Halden, D., Jones, P. & Wixey, S. (2005) *Accessibility analysis literature review*. Westminster, UK: University of Westminster.

Hayuma, A. M. (1980) Dodoma: The planning and building of the new capital city of Tanzania. *Habitat International, 5*(5-6), 653-680.

Iles, R. (2005) *Public transport in developing countries*. Oxford, UK: Elsevier.

Inter-consult Ltd. & Logit. (2006a). Dar es Salaam Planning and Design, Final Report [Presentation]. Dar es Salaam, Tanzania: Dar es Salaam City Council.

Inter-consult Ltd. & Logit. (2006b) *Dar es Salaam Rapid Transit: Draft Detailed Roadwork Designs*. Dar es Salaam, Tanzania.

Inter-consult Ltd. & Logit. (2006c) *Dar es Salaam Rapid Transit: Draft final report - Annex volume 3: Field survey and data collection*. Dar es Salaam, Tanzania.

Inter-consult Ltd. & Logit. (2006d) *Dar es Salaam Rapid Transit: Draft final report - Annex volume 4: Demand modelling and forecasting*. Dar es Salaam, Tanzania.

Inter-consult Ltd. & Logit. (2006e) *Dar es Salaam Rapid Transit: Draft final report - Annex volume 5: Operational design*. Dar es Salaam, Tanzania.

Inter-consult Ltd. & Logit. (2007) *Dar es Salaam Rapid Transit: Draft final report - Annex volume 5: Operational design*. Dar es Salaam, Tanzania.

Japan International Cooperation Agency, Pacific Consultants International & Construction Project Consultants. (2007) *Dar es Salaam Transport Policy and System Development Master Plan: Progress Report*. Dar es Salaam, Tanzania.

Li, Y. W. (2003) Evaluating the urban commute experience: A time perception approach. *Journal of Public Transportation, 6*(4), 41-67.

Luan, H. C. (2002) *Public transport data modelling in Cairo*, International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands.

Lyons, G., Marsden, G., Beecroft, M. & Chatterjee, K. (2001) *Transportation requirements*. Southampton, UK: University of Southampton.

Mackett, R. L. & Edwards, M. (1998) The impact of new urban public transport systems: will the expectations be met? *Transportation Research Part A: Policy and Practice, 32*(4), 231-245.

Maunder, D. A. C. & Fouracre, P. R. (1989) *Non-motorized travel in Third world cities.* Paper presented at the Institute of British Geographers' Annual Conference on Managing our Environment, Coventry, UK.

Mestrum, D. (2008) *Analyzing people's preferences to support Better Rapid Transit in Dar es Salaam (Preliminary thesis title)*, University of Twente, Enschede, The Netherlands.

Murray, A. T. (2001) Strategic analysis of public transport coverage. *Socio-Economic Planning Sciences, 35*(2001), 175-188.

Murray, A. T. (2003) A coverage model for improving public transit system accessibility and expanding access. *Annals of Operations Research, 123*, 143-156.

Murray, A. T. & Wu, X. (2003) Accessibility tradeoffs in public transit planning. *Journal of Geographical Systems, 5*(1), 93.

O'Sullivan, D., Morrison, A. & Shearer, J. (2000) Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science, 14*(1), 85-104.

OmniTRANS International. (2005) OmniTRANS 4.2 manual. Retrieved September 19, 2007, from www.omnitrans.nl

Phani Kumar, C. V., Basu, D. & Maitra, B. (2004) Modelling generalized cost of travel for rural bus users: a case study. *Journal of Public Transportation, 7*(2), 59-72.

Raux, C. (2003) *A system dynamics model for the urban travel system.* Paper presented at the European Transport Conference, Strasbourg, France.

Richards, F. J. (1959) A flexible growth function for empirical use. *Journal of Experimental Biology, 10*, 290-300.

Rizzo, M. (2002) Being taken for a ride: privatisation of the Dar es Salaam transport system 1983 - 1998. *Journal of Modern African Studies, 40*(1), 133-157.

SUMATRA. (2007) *Daladala licenses 2005-2007*. Dar es Salaam.

The National Bureau of Statistics of the Republic of Tanzania. (2003) *2002 population and housing census [2002PHC]*. Dar es Salaam, Tanzania.

Van der Zwan, J., Van der Wel, R., De Jong, T. & Floor, H. (2005) *Manual Flowmap 7.2*. Utrecht, The Netherlands: Utrecht University.

Vuchic, V. R. (2005) *Urban transit: operations, planning and economics*. Hoboken, NJ, USA: John Wiley & Sons.

Wardman, M. (2004) Public transport values of time. *Transport Policy, 11*(4), 363-377.

Wright, L. (2002) *Bus rapid transit*. Eschborn, Germany: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ).

Wright, L. & Fjellstrom, K. (2003) *Mass transit options*. Eschborn, Germany: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ).

Wu, B. M. & Hine, J. P. (2003) A PTAL approach to measuring changes in bus service accessibility. *Transport Policy, 10*, 307-320.

Appendix A: Description of TS Requirements

This appendix gives a more concrete description of the requirements as proposed by Vuchic (2005). However some indicators are discussed in depth in section 2.1. Assessment Framework.

Passenger Requirements

Passengers are considered the most important party, because if the system does not meet their standards, very few people will use the system and all (passengers, operator and community) will suffer losses (Vuchic, 2005). However it has to be said that there are also quite a number of people who have little choice and who will use the public transport as long as the price is low. This goes especially in developing countries and in places where people have very little money and cannot afford other transport options.

Availability

Availability describes whether people can indeed access the system both spatially (is the transit stop close enough) and temporal (is the transit facility available at the moment when it is needed) (Vuchic, 2005). This however seems a quite limited approach of the subject since there are numerous factors that influence the people's possibilities to use the system. Hence, not only access time is important, but also whether the destination can be reached, how long the trip takes, etc. Therefore the term accessibility (see also: Lyons *et al.*, 2001; Murray, 2003; Murray & Wu, 2003; O'Sullivan *et al.*, 2000) is used to describe the proposed comprehensive term, whereas availability will be used to describe the Vuchic requirement.

Frequency / Headway

The frequency is the number of vehicles operating at one route. The headway is the time between two consecutive vehicles and can therefore be used to describe the expected waiting time, if a constant approach of passengers is assumed. The headway can be calculated from the number of vehicles per route, route length and realised speed using equation (A.1).

$$t_{h;r} = \frac{2 \cdot l_r}{v_r \cdot n_r} \tag{A.1}$$

$t_{h;r}$ = headway time for route *r*

l_r = route length of route *r*

v_r = vehicle speed on route *r*

n_r = number of vehicles on route *r*

Punctuality

This is usually described as the on-time performance of a transit route. Therefore the average delay factor can be used to describe punctuality. This average delay factor can be described using equation (A.2), which is the average delay divided by the headway. But due to the nature of the current system (daladala) it is not possible to measure this variable, it can however be used when DART is in operation, to assess the performance.

$$f_{d,r} = \frac{\sum_{n} t_{d,r}}{n \cdot t_{h,r}} \tag{A.2}$$

$f_{d;r} =$ delay factor for route r

$t_{d;r} =$ delay for a vehicle on route r

$n =$ number of measurements

$t_{h;r} =$ headway time for route r

Speed / Travel time

The speed, or in fact travel time, is probably the most important factor concerning passenger satisfaction, at least in Europe and North America. In this research travel time is therefore defined as the time it takes from the moment one leaves its origin (e.g. its house) up to the moment one arrives at its destination. Not only the true travel time is important, but especially the perceived travel time is of great influence when assessing public transport. This perceived travel time can be modelled like equation (A.3), where α_n is a factor that describes the perception of time (this equation is explained in detail in section 2.1.2. , see also equation (2.6)).

$$t_p = \alpha_a t_a + \alpha_w t_w + \alpha_{iv} t_{iv} + \alpha_t t_t + \alpha_e t_e \tag{A.3}$$

Comfort

A comfortable transit can be described using many variables, among which walking environment, attractive stations, temperature and ventilation. However especially in a city like Dar es Salaam, where public transport is lacking capacity and primarily consists of small minivans, comfort could most likely be described in a more straightforward manner.

In this research comfort will therefore be described using the time in a certain transit mode, the availability of a seat and the space available per passenger. Then for each transit mode the total comfort level could be calculated using equation (A.4).

$$Com_t = \left(\alpha_{1;t} Com_{seat;t} + \alpha_{2;t} Com_{space;t}\right) \cdot \left[1 - \frac{1}{1 + e^{-\beta_{1;t}\left(t_t - \beta_{2;t}\right)}}\right] \tag{A.4}$$

$Com_t =$ comfort level for a trip in mode t

$Com_{seat;t} =$ comfort level for having a seat in mode t

$Com_{space;t} =$ comfort level for having a certain amount of space per passenger on mode t

$t_t =$ time spend on mode t

This equation is a modification of Richard's curve (Richards, 1959) so that it models a decrease instead of a increase in value. The α's represent a factor to obtain the maximum comfort level, the β_1 describes the maximum decrease and the β_2 pinpoints at which time this decrease is realised. Then the total comfort of the whole trip can be estimated using equation (A.5).

$$Com_T = \frac{\sum_i t_{t;i} \cdot Com_{t;i}}{t_T} \tag{A.5}$$

$Com_T =$ comfort level for the total trip T

$t_{t;i} =$ time spend on trip section i in mode t

$t_T =$ total time for trip T

Convenience

Convenience is even harder to define than comfort, but Vuchic (2005) emphasises the choice of people for a certain transport mode and the information they have about the transit mode. In the city of Dar es Salaam however, very few people have alternatives (besides walking) so convenient travel would never be reached. However he also mentions the directness of travel which is a good indicator, when modelling in a network based GIS environment. The only question that remains is which travel time can be used to scale the performance of urban public transport. In western countries, the travel time that is realised by private car could be used, however in cities like Dar es Salaam, this is not feasible, since most people do not possess a car. Therefore three possibilities can be devised:

- walking time
- public transport time without waiting / transfer time
- calculated travel time based on Euclidean distance and a desirable speed

Walking time seems a good benchmark for people living near their destination, since they will decide to go and walk when public transport travel time is too high. However for people that travel a large distance (e.g. Kimara – CBD) walking is no realistic option. Using the public transport without delay times, is a good indicator for delay factors in public transport, but doesn't say anything about the time you are on the road, which could be inconvenient as well. Therefore the last option will be used in this research, where the convenience can be calculated using equation (A.6). This also provides local government with the possibility to create a benchmark for an average travel speed throughout the city.

$$Con_T = \frac{t_p \cdot v_{des}}{1.34 \cdot ed} \tag{A.6}$$

$Con_T =$ convenience of trip T

$t_p =$ perceived travel time of trip T

$v_{des} =$ desired speed for a certain trip T

$ed =$ Euclidian Distance from the origin to the destination of trip T

However some thresholds may be needed to prevent people living very close to their destination from not complying with the convenience requirement. Also mention the fact that perceived travel time could give more insight into the convenience, because it incorporates uneasy travel.

Security and safety

Safety from a passenger point of view could perhaps best be described as the absence of accidents. However also the perception of safety (how likely is an accident going to happen) might influence people's opinion on safety. However, to make representations in GIS possible, the first definition shall be used.

User cost

The cost of travel usually consists of only the travel fare. However in this research also the generalized travel cost shall be considered, since this might differ from area to area. This is also important for transit modelling, since very few people base their transit mode on monetary of time performance only. The generalized cost can be calculated using equation (A.7) which is explained in depth in section 2.1.3.

$$GTC_i = C_f + \frac{t_T}{VOT_i} \tag{A.7}$$

Transit Operator Requirements

The requirements of the operator do sometimes overlap the passenger requirements, but look at the matter from a different (more economical / business) point of view.

Area coverage

The coverage of a public transport system is usually defined by a five minute (primary) and ten minute (secondary) accessibility circle. A more precise method is to calculate the percentage of people that have access to public transport within five or ten minutes. These figures however, are based upon western standards, which most likely have to be adapted for African (Dar es Salaam) standards. Therefore no coverage circles will be created, but the distance towards the closest bus stop will be determined, which can then be used to create 'coverage circles'.

Reliability

This is strongly correlated with the change of failure and the punctuality as described in de passenger requirements section. These requirements will therefore be assumed to be more or less the same as the punctuality, since both describe the delay of a certain line and therefore equation (A.2) holds. However, also the accident log (see Security and safety) is a good indicator for the reliability of the system. Reliability could therefore be represented in a layer where both delay and accidents are shown.

Cycle speed / Line capacity

This is the time it takes to make a complete run on a certain route, which, in case of a straight line, means from A to B and back to A. When the average speed and route length are known this requirement can be calculated using equation (A.8). The cycle speed is of great influence on the operators' cost, because on faster routes, fewer vehicles are needed.

$$t_{c;r} = \frac{2 \cdot l_r}{v_r} \tag{A.8}$$

$l_r =$ length of route r

$v_r =$ vehicle speed on route r

The capacity of a route is important for the same reason, and is depending on the headway and the vehicle capacity. Therefore the line capacity can be calculated using equation (A.9).

$$C_r = f_r \, \overline{Cap}_v \tag{A.9}$$

$C_r =$ passenger capacity on route r
$f_r =$ frequency of route r
$\overline{Cap}_v =$ average capacity of the vehicles on route r

Flexibility

A system is regarded flexible, when it can change either in space or in schedule without any (or with minor) expenses. This is however very difficult to describe, especially in a GIS environment. Therefore flexibility is defined as 'whether the public transport vehicle can move spatially, whether the vehicle can overtake others and whether the capacity can easily be expanded'. This is therefore defined for each route section which is than translated to each road section, where the worst case holds.

Safety and security

This is regarded to me more or less the same as the safety and security from a passengers point of view.

Costs

The costs of the urban transport system can be divided into operational cost and capital cost (involved in the construction and maintenance of infrastructure). In this research this requirements aims at estimating the operating costs of each route by estimating fuel and personnel costs and tries to assess the costs per passenger kilometre. The total cost per vehicle of type v per hour can be calculated using equation (A.10). Therefore the total cost per route per day can be calculated by equation (A.11) and the cost per passenger can then be derived from equation (A.12).

$$C_v = \left(n_{p;v} \cdot C_p\right) + \left(v_v \cdot u_d \cdot C_d\right) \tag{A.10}$$

$$C_r = \sum_v \left(n_{r;v} \cdot C_v\right) \cdot \sum_t O_{r;t} \tag{A.11}$$

$$C_{pkm;r} = \frac{C_r}{\sum_l n_{p;l} \cdot l_l} \tag{A.12}$$

$C_v =$ cost of running vehicle type v for one hour
$n_{p;v} =$ number of staff on vehicle type v
$C_p =$ cost of one staff member for one hour
$v_v =$ average speed of vehicle type v
$u_d =$ fuel consumption per kilometer of vehicle type v
$C_d =$ fuel cost
$C_r =$ cost of route r
$n_{r;v} =$ number of vehicles type v on route r
$O_{r;t} =$ percentage of vehicles working on route r during time period t
$C_{pkm;r} =$ cost per passenger kilometer for route r
$n_{p;l} =$ number of passengers on route r per link l
$l_l =$ length of link l

Passenger attraction

The attraction is the total number of people that use the public transport system. This is however quite difficult to determine, but could be estimated using a route choice model. One of the most important factors is therefore, especially in Dar es Salaam, whether people have a car at their disposal. It is therefore presumed that people who do have a vehicle, are not likely to use public transport.

Side effects

The side effects of a urban public transport system are numerous. Vuchic (2005) mentions e.g. physical effects, noise, vibrations and air pollution. Because this research aims merely at public transport, it is quite hard to determine the influence of these factors, since all traffic contributes to them. Therefore only noise and air pollution are considered side effects in this research and the necessary attributes will be stored in the database.

Community Requirements

The most important requirements of the community are somehow related to passengers or operators. Here all parties are united to optimize the public transport system.

Service quality / Passenger attraction

Here the previously determined passenger attraction and area coverage meet a benchmark, to see whether political objectives are met. The service quality is therefore defined as the area coverage, whereas the passenger attraction remains the same as in the transit operator requirements.

System cost

The system cost are considered to be the total lifetime cost, therefore existing of the construction cost, maintenance cost en operating cost during the complete lifecycle. It is estimated that a BRT system has a lifetime of about 30 years, before major reconstruction has to take place. Therefore the life period is set at 30 years.

Reliability in emergencies

This is very difficult to define, since it is completely dependent on the situation whether a system is capable of coping with it. So the definition limits itself to the duration of evacuation of a certain area.

Social objectives

Social objectives can be set at different requirements, for instance area coverage and availability. Therefore this requirement can be met for every requirement that is defined before. Also it is possible to model the accessibility of certain large events and other social happenings.

Environmental and energy aspects

The environmental and energy aspects are already mentioned in the noise and air pollution (side effects) requirement. Based on the fuel consumption that is estimated in the operator cost requirement also energy aspects can be modelled. In the future, a model could be developed to estimate future (public) transport demand based on for instance fuel prices.

Long-range impacts

The long range impacts become primarily visible when assessing the change in land use overtime. However this model could be able to predict where land use changes are likely to happen, based on the assumption that areas that suddenly improve significantly on the accessibility criteria are likely to develop.

Appendix B: Assessment Attributes

This appendix starts with a matrix which 'describes' the relation between requirements and attributes. Next a brief description of each attribute is given, so that it is clear which information is needed for the assessment.

Requirement – Attribute Matrix

		Passenger									Operator								
	abbreviation	Availability	Frequency	Punctuality	Travel time	Comfort	Convenience	Security & Safety	User Cost		Area coverage	Reliability	Cycle speed	Line capacity	Flexibility	Safety & Security	Cost	Passenger attraction	Side effects
Population	Pop_i	X									X							X	
Vehicles per route	$n_{v;r}$	O	X	O	O	O			O			O		O			X		
Road length	l_{rs}	X			X	O			O		X							X	
Walking speed	v_w	X			X	O			O		X							X	
Factor walking	α_a/α_e				X	O			O										
Factor waiting / transfer	α_w/α_t				X	O			O										
Daladala link length	l_l	O	X	O	X	O			O			O	X	O					
Daladala speed	v_d	O	X	O	X	O			O			O	X	O			X		
BRT link length	l_l	O	X	O	X	O			O			O	X	O					
BRT speed	v_b/v_f	O	X	O	X	O			O			O	X	O			X		
BRT delay time	$t_{d;r}$			X								O							
Comfort of seat	$com_{t;seat}$					X													
Comfort of space	$Com_{t;space}$					X													
Comfort factors	$\alpha_{t;1}/\alpha_{t;2}$					X													
Comfort time factors	$B_1/\beta_{t;2}$					X													
Desired minimum speed	v_{des}						X												
Fares	$C_{f;t}$								X										
Value of time	VOT_i								X										
Vehicle capacity	Cap_v													X					
Fuel cost	C_d																X		
Fuel consumption	u_d																X		
Personnel cost	C_p																X		
Personnel per vehicle	n_p																X		
PT occupancy	O_t																X		
Passengers	$n_{p;l}$																X		
Land use	LU_i																		
Accident log	AL							X				O				X			
Flexibility constraints	$f_1/f_2/f_3$														X				
Car possession	p_c																	X	

Legend	
X	directly used in computations
O	uses a requirement that consists of this attribute

Attribute description

Attribute data	Abbreviation	Storage level	Unit	Description
Population	Pop_i	H	#	people that live in area i
Vehicles per route	$n_{v;r}$	R	#	vehicles of type v that are licensed to route r
Walking speed	v_w	Rd	km/h	Speed of walking people
Road length	l_{rs}	Rd	m	Length of each road section
Factor walking	α_a	H	-	Relative value of access time
	α_e	H	-	Relative value of egress time
Factor waiting / transfer	α_w	H	-	Relative value of waiting time
	α_t	H	-	Relative value of transfer time
Daladala link length	l_l	L	m	Distance between two daladala stops
Daladala speed	v_d	R	km/h	Average operating speed of daladala
BRT link length	l_l	L	m	Distance between two bus stops
BRT speed	v_b	R	km/h	Average operating speed of BRT
	v_f	R	km/h	Average operating speed of feeder
BRT delay time	$t_{d;r}$	D	min	Delay time of every BRT bus
Comfort of seat	$com_{t;seat}$	V	-	Comfort value of having a seat
Comfort of space	$Com_{t;space}$	V	-	Comfort value of riding an empty vehicle
Comfort factors	$\alpha_{t;1}$	L	-	Does one have a seat
	$\alpha_{t;2}$	L	-	Opposite proportional to # of passengers
Comfort time factors	$\beta_{t;1}$	V	-	Maximum decrease in comfort
	$\beta_{t;2}$	V	min	Time of maximum decrease
Desired minimum speed	v_{des}	Db	km/h	Desired minimum speed based on policy
Fares	$C_{f;t}$	R	Tsh	Fares of routes and transfers
Value of time	VOT_i	H	Tsh/h	Value that people give to time (savings)
Vehicle capacity	Cap_v	V	#	People that are allowed in a vehicle
Fuel cost	C_d	Db	Tsh/L	Price of fuel per litre
Fuel consumption	u_d	V	km/L	Average fuel consumption per vehicle type
Personnel cost	C_p	Db	Tsh	Salary cost per person
Personnel per vehicle	n_p	V	#	People that work are needed per vehicle type
PT occupancy	O_t	Db	%	Amount of vehicles that is in service
Passengers	$n_{p;l}$	L	#	People that travel on certain link
Land use	LU_i	H	-	The land use in area i
Accident log	AL	A	-	Accident description
Flexibility constraints	f_1	L	-	Is spatially movement possible
	f_2	L	-	Is overtaking possible
	f_3	L	-	Is capacity expansion possible
Car possession	p_c	H	%	People that own a car in area i

Legend		
A	Accident log	Table with accident information
D	Delay log	Table with delay information
Db	Database	One single value for the whole system
H	Hexagon	Location information
L	Link	Public transport network layer
R	Route	General information about routes
Rd	Road	Information at street level
V	Vehicle	General information about vehicles

Appendix C: Adding and Editing Geometry

To be able to add new data to the public transport model or to edit and improve existing data, a good understanding of how the data is stored is vital. This appendix therefore describes where the data is stored, and how to add new data without losing information from previous phases. The term phase is used to point at a certain stage in the evolvement of public transport over time. Thus the first stage is the current daladala system (in the model denoted by DN), and the next phase would be the implementation of dart phase 1 (in the mode denoted by D1).

Storage Levels

The database uses different ways to store information. The most simple one is the storage of general data, that is neither spatial nor temporal (in this case meaning: dependent on a certain phase). This information is stored in simple tables, because other methods would simply be too complex. The second type is the information that is spatial, but isn't temporal. This type of information is stored in feature classes, to make spatial storage possible. The third and final information type is information that is both spatial and temporal. This type of information in stored in feature classes, but they are grouped together per phase in feature datasets.
Because each feature class name can only appear once in the database, each feature class that is temporal has an indicator for the phase it belongs to appear in the name. Thus a route layer that is part of the DART phase I could be called Route_D1.
For the most important tables and feature classes, the information that is stored will be explained below, the phase that will be discussed is the DART phase I, because it is the most complex one (in terms of different feature classes).

General Tables

This paragraph discusses the tables that are used in this research. There are also other tables present, but they can be used to cover some of the requirements that were not discussed in depth and can therefore be altered.

BRT_Link

This table contains two columns, being: ***B_Link_ID*** and ***Road_ID***. This table is used to connect each link from the DART routes to a road segment, in order to make true representations possible (e.g. show vehicle intensities per road section). This table can also be used to calculate the true length of a route section because using the length of a route section would result in the Euclidian distance.

Route

The route table stores all information that is valid for the whole route. The first column ***Label*** indicates both the origin and destination of the route and can therefore be used to label the route. The ***Route_ID*** is a number to identify each route. The routes have unique numbers, meaning that the routes numbered from 1-500 are the daladala routes in the situation DN and the routes numbered 501-999 are the same routes, but now in the D1 phase. It is therefore possible to change the characteristics of a certain route per phase, whilst keeping all information together in one table. The routes numbered 1001-1999 are the DART routes and the routes numbered 2001-2999 are the feeder routes, that are used in the D1 phase.
The ***Start_st*** and ***End_st*** is derived from the label and can be used for easy selection of routes with certain start or endpoints. The ***No_Veh*** column states the number of vehicles that should

be operation on that route, according to SUMATRA (2007) or the consultants design reports (Inter-consult Ltd. & Logit, 2007).
The ***Route_Length*** is the length from start station to the end station in meters. When there are oneway restrictions present at a certain route, the length shown in the average length. The ***Fare*** is the flat fare that has to be paid when entering a vehicle, this is also why there is no fare set for the DART feeder and trunk routes, because they have integrated complex fares, that are dependent on the mode that was used before entering that route.
The ***FT_Speed*** and ***TF_Speed*** is used to calculate frequencies and are based on a very small speed survey. For the DART routes, the information from the Logit designs is used. The ***Freq_Logit*** is the frequency that each (daladala) route has based on the consultants research (Inter-consult Ltd. & Logit, 2006c, 2006d).
The ***Type*** column indicates the type of route it represents. The values and their meanings are listed below in table C.1.

Type	Description
1	Daladala
2	DART Feeder
3	DART Trunk (stopping service)
4	DART Trunk (fast service)
5	Ferry

table C.1: Route types

The ***Avg_Cap*** is a column that could be used in the future to represent the average capacity per vehicle on each route. Last, the ***Frequency*** column is based on the speed, length and number of vehicles using equation (A.1) or the frequency as determined by the consultants (Inter-consult Ltd. & Logit, 2006c, 2006d).

General Feature Classes

The general feature classes contain information about the city itself and although there are more layers stored only one is really important (the others are products of calculations or are used for that purpose).

HEX_Population

The HEX_Population feature class stores all spatial information that is related to the people living there. The first columns contain information that is needed by ArcGIS to create the polygons.
The ***Municipality*** indicates to which municipality this hexagon belongs, and in a similar way indicates ***Ward*** to which ward this hexagon belongs. This makes it possible to aggregate information to a higher level (note: the other way around is not possible, information that is stored at municipality level, cannot be detailed into ward information).
The ***Area*** column indicates the size of the hexagon in square meters and the ***Population*** column stores the estimate of number of people that live in that hexagon. The next columns could be used to store information about car possession, and their value of access, egress, waiting and transfer times. The ***Hex_ID*** is used to link information that belongs to a certain hexagon to that hexagon (e.g. when new accessibility analysis have been made).
The other columns in this feature class contain information about different analysis that are performed. They have a systematic naming, in order to make retrieving information easy. Nearly every column is made in the format *PHASE_OPTIMIZATION_ANALYSISMETHOD_ANALYSISTARGET*.

Here the *PHASE* is either DN, D1 or DN_D1 which indicates the improvement from phase DN to D1. *OPTIMIZATION* indicates which attribute was used to determine the result, e.g. one could minimize the true travel time to get an idea about the travel costs incurred, but one could also minimize the generalize travel cost to perform this analysis. The *ANALYSISMETHOD* indicates what type of statistical analysis is represented, often this is the average value, but it could also be the maximum, standard deviation, etc. Finally the *ANALYSISTARGET* indicates which attribute is displayed, this could be the travel cost (as mentioned before) but it could also be the generalized travel cost, perceived travel time, etc.
Thus a column DN_D1_GTC_AVG_TT contains information about the average travel time improvement from phase DN to D1 when people optimize for generalized travel cost.

Phase Dependent Feature Classes

The DART phase I feature dataset contains five different feature classes and a network dataset, that has to be created in order to make network analysis possible. This paragraph will therefore discuss the five feature classes, but also the basic settings for this network dataset.

Road_Network_D1

The road network layer table starts with a couple of columns that were present in the raw data file and were not removed, so that the basic data could still be viewed, even when a lot of new data was added. The first new column is the ***BRT_Dummy*** which was used to mark a couple of roads that were added, because DART is planning on moving some roads, or adding roads to access large bus terminals.
Road_ID is an identifier for each road link and the ***D_Section*** is created to link each road section to a daladala network section. Although the method used was quite complicated, this means that all road that have the same D_Section are part of the daladala route links that are marked with the same D_Section number. Again this is can be used to transfer information from roads to routes (e.g. distances) or information from routes to roads (e.g. intensities).
FT_Speed and ***TF_Speed*** can be used to calculate the travel time on each link, based on the road length, however since all speeds are equal the ***W_Speed*** column was used for this analysis (indicating that one can always walk in two directions with a certain speed, that is considered to be equal for each environment). However because this might be untrue, two columns were created in advance to store different walking speeds per road section.
The same goes for the ***FT_TT*** and ***TF_TT*** which could represent the walking time in both directions, but for the same reason, these are represented by the values in ***W_TT***.
The values for ***DD_DN***, ***DD_D1***, ***FF_D1*** and ***BB_D1*** are used to indicate whether there is a daladala, feeder or BRT present on that road section in phase DN or D1. For the D1 phase a summary is created in the ***Class_D1*** which forms a classification for the road (basically 1 indicates that a daladala is present, 10 indicates that a feeder is present and 100 indicates that a DART bus is driving on that link. This is summarized, thus a road link with Class_D1 = 111 is a road where daladala, feeder and DART are driving).

Daladala_D1

This feature class (layer) contains all bus stops for the daladala routes in the system. Further information about this stops could be stored (e.g. in what state the bus stop is, whether improvements are needed, the number of people that use that bus stop each day), but in this research that was not necessary.

Daladala_Network_D1

The daladala network layer contains all daladala routes and the dummies that are needed to model the bus stops. The ***D_Link_ID*** column contains unique values for all links, again to make data transfer possible.

The ***FT_Speed*** and ***TF_Speed*** indicate the speed in both directions, where the route is always modelled from the start station to the end station (as noted in the Routes table). This results in a travel time on each link, which is stored in ***FT_TT*** or ***TF_TT***. Because waiting and paying fare is modelled at dummy links, these values are 0 for normal route links.

The fares and waiting times are represented in ***FT_Cost***, ***TF_Cost***, ***FT_Wait*** and ***TF_Wait***. Because waiting and paying fares only happens when one gets onto a daladala, the values for the TF are always 0.

The ***Oneway*** column indicates whether that route can only be used in one direction. When that is the case, this direction is indicated with FT (when only from-to is possible) or TF (when only to-from is possible). The length of each link is represented in ***Link_length*** and the number of passengers on each link could be stored in the ***Passengers*** column.

The ***Route*** column indicates to which route this section belongs, and the ***D_Section*** is used to link the daladala routes to the roads, as discussed before. To indicate what the use is of each link, the ***Type*** column is used to determine whether the link is part of a route (R), is a dummy to model waiting and paying (D), is used to connect two one way routes (T) or whether it is the ferry (F). This column can be very convenient to make selections to update route attributes.

DART_stops_D1

All the bus stops of the DART trunk and feeder routes are stored in here. Because some of the bus stops have names and are classified (see: Inter-consult Ltd. & Logit, 2007) a column for names and a column to indicate the type is added (where 1 is a normal station, 2 is an feeder-trunk transfer station and 3 is a terminal).

DART_phase1_Network_D1

The DART network contains all route segments for the DART feeders and trunks, including the dummies and all transfer possibilities. The ***Route*** column indicates from which route this link is, when it is a transfer link it is assigned to route 0 (which is non-existent). The ***Type*** is used to determine the function of the link, therefore the codes as described in table C.2 are used.

Type	Description
BB	DART trunk link
FF	DART feeder link
DD	Transfer link between two trunks
FB	Transfer link between trunk and feeder
SB	Dummy link from street to DART trunk
SF	Dummy link from street to DART feeder
SS	Dummy link from street to street

table C.2: DART types

The last type (SS) is used to prevent daladala bus stops and DART bus stops to be at exactly the same location, because that causes problems in modelling. The attributes assigned to these links are therefore all 0, because both ends represent the same location (in reality).

The ***FT_Speed***, ***TF_Speed***, ***FT_TT*** and ***TF_TT*** have been discussed above, but the ***FT_Cost***, ***TF_Cost***, ***TF_Wait*** and ***TF_Wait*** are slightly different, because of the different transfer possibilities. For the dummy links from the street to the routes, not much has changed. However for the transfers between routes, the waiting time is identical to the waiting time for the route one is transferring to (there is no difference in arriving at a station by bus or by foot). The value of the additional cost is dependent on the fare structure that is proposed, here the structure as created by the consultants (Inter-consult Ltd. & Logit, 2007) is used. Note, that these transfer links **can** be travelled in two directions and that the values for both directions most likely differ.
The ***From_ID*** and ***To_ID*** is used to create one direction for the transfer links, which makes it more convenient to select and alter the information (the waiting time to a route can conveniently be changed by selecting all route links that have that route as From_ID or To_ID).
The DART trunk routes are connected to the road network with the previously discussed table, the feeders use the same structure as the daladala's and therefore the ***D_Section*** column is present in this feature class.

DART_phase1_Network_ND

The network dataset creates the actual network model, based on all the phase dependent feature classes discussed above. The most important of a good network model is the connectivity policy. This policy described between which layers people can travel, and at what location (e.g. intersections or only endpoints). The connectivity policy for this model is shown in figure C.1.

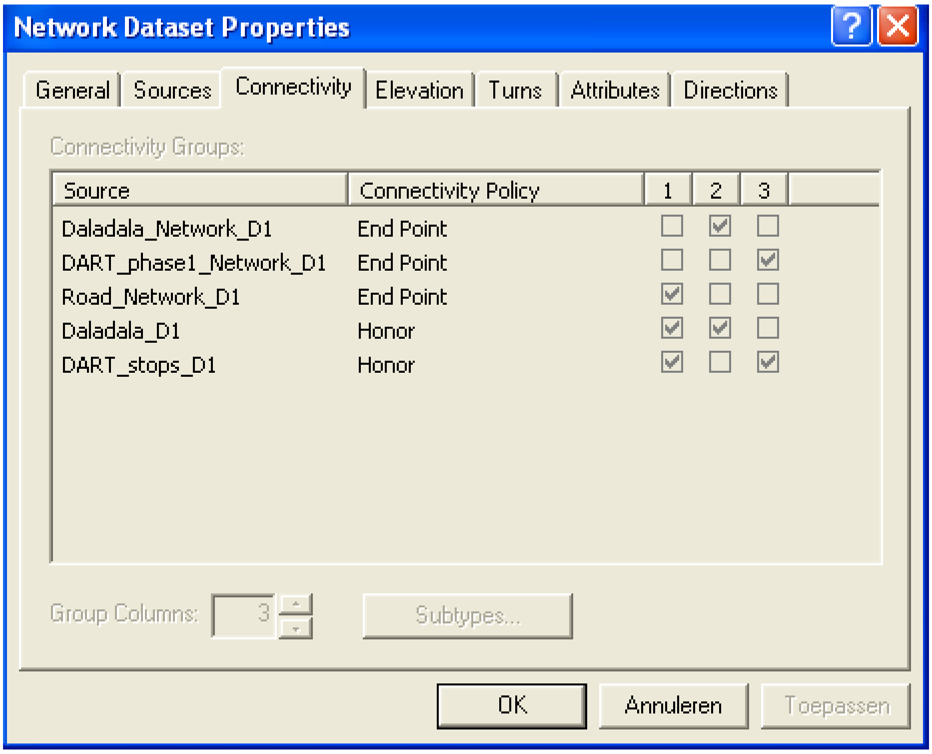

figure C.1: Connectivity policy for DART phase I

This policy describes that one can move from the road network to the daladala and DART stops, but not directly to their routes. From the daladala stops one can go to the daladala routes and from the DART stops one can enter the DART routes. This policy enforces people to move to the bus stop using the road network and from there transfer to a bus route (thus

resulting in waiting time and paying fares). When people transfer between different modes (thus between daladala's and DART busses) people will have to go back to the road network and walk towards the other bus stop in order to get onto that route.
In this network model also the attributes that are used to do the analysis are defined using so called evaluators. How these should be created is discussed in Appendix D and E.

How to Create a New Phase

When a new phase is created the most convenient solution is to copy the previous phase and alter al the route numbers (this can be done very fast by simply selecting all the routes and adding *x* to them. Remind that this should also be done in the route table.
When this is done the information can be altered or new routes can be drawn. It is important that new routes connect to bus stops exactly (use the snap function) because otherwise modelling is going wrong (the model thinks that the two objects are not connected). For each route link the information discussed above should be provided.
When all layers are finished a new network dataset should be created, and a new connectivity policy should be entered. Also all the attributes and their evaluators should be entered, which can sometimes be quite hard, because selecting the wrong column (e.g. the TF instead of the FT) can result in nice, but completely wrong results.

Appendix D: Creating Accessibility Analysis Layers

As goes for all the analyses, the most difficult part of performing analysis is to be able to link the information to the desired location. This appendix therefore describes the method that is used to perform the PTAL accessibility analysis, so that this analysis can also be executed at later times.

Data Requirements

In order to be able to make an accessibility analysis the following data is required:

- Point layer representing the area where people live [1]
- Point layer containing information about the location of public transport stops [2]
- Line layer containing the road network [3]
- Line layer containing the urban transport routes, in the format that is used in this research [4]
- Table containing information about the different transport routes [5]

Data Preparation

First the distance from the houses towards the closest public transport stops have to be determined. Also it is necessary to know which routes stop at which stops. When this data is readily available, e.g. because of the way the route data was collected, then this information can be used. Otherwise the steps below can be used to derive this information from ArcGIS.

Find closest public transport stops for each housing area

In ArcCatalog create a network dataset [6] based on only the road network [3]. Define as only *impedance attribute* TT_A which is based on the walking time for each road section.
Open ArcMap and perform a OD Cost Matrix analysis and load the point layer with housing information [1] as origins [7] and load the point layer with public transport stops [2] as destinations [8]. Use TT_A as *impedance* and set the *cutoff value* to 30 minutes. Make sure that each location is snapped to the nearest road, by setting the *search tolerance* to 5 or 6 kilometers. In order to make calculations go faster, set the *output shapefile* to none. Save the resulting lines [9].
Join the housing information [1] with the origins [7] using the *spatial join* and assign *all attributes* to the nearest point. Save the new layer [10]. Do the same with the public transport stops [2] and the destinations [8], also save this new layer [11].
Export the attribute table information of the two joined layers [10,11] and the resulting lines layer [9] into the database [12,13,14] naming them Origins, Destinations and Lines respectively.
Using MS Access open the PT assessment database and run the 'Create Table HEX_STOP' macro. This will create a new table called HEX_STOP [15]. This only works when the origins table [12] contains a column called Hex_ID and the Destinations table [13] contains a column called Id.
Using ArcCatalog remove the shape files that were created through spatial joining [10,11] and remove the tables that you created manually [12,13,14] as well as the network dataset [6].

Find routes that visit a certain public transport stop

Since you are already in ArcCatalog, create a new network dataset [16], consisting only of the public transport stops. Define again as *impedance attribute* TT_A, but use as evaluator a constant factor of 0.
Open ArcMap and load the public transport network layer [4]. Select only the dummy links and export these to a new file [17].
Select form the ArcToolbox *Data Management* → *Features* → *Feature to point* and transform the dummy links [17] to points [18].
Use the new network dataset [16] and load the points [18] as origins [19] for a new OD Cost Matrix analysis. Set the *search tolerance* to the length of the dummy links (usually 10 meters) in order to improve the calculation speed.
Again use the *spatial join* to assign *all attributes* from the origins [19] to the dummy points [18], save this new layer, containing the stop and route information [20]. Again export the attribute table and save this information as STOP_ROUTE [21].
Close ArcMap and open ArcCatalog to clean the database. Remove the newly created files except the attribute table, thus removing [16, 17, 18, 20].
Open MS Access and clean the STOP_ROUTE table [21] by removing everything except Route, FT_Wait and ObjectID_1, which should be renamed to Stop_ID.

Create network model for Corrected Accessibility Index / other analysis

First a net network dataset [22] has to be created in ArcCatalog. Here the complete road network [3], public transport network [4] and public transport stops [2] have to be used and the correct *Connectivity* policy has to be set. The *impedance attributes* should at least contain an attribute which is calculated using all travel time factors. This can be created by using field as *type* and *value* should be an expression calculating the summation of all time attributes. When the network dataset is also used for other purposes, other *impedance attributes* should be included.
Close ArcCatalog en open ArcMap to perform a OD-Cost analysis. Load the network dataset [22] and load the point layer with housing information [1] as both origins and destinations. When needed (for instance with slow computers) one can split up the origin and destination layer into smaller parts. Set the *impedance unit* to travel time, *range* to 60 minutes and set the *export* to no lines. Activate *oneway* and make sure that each location is snapped to either the road network or public transport stop, and ***not*** to the public transport network.
Run the OD-Cost analysis (multiple times if the origins and destinations are split up) and save the resulting lines [23] and join the origin and destination layer spatially to the original layer [1] and store all values for the point that is closest. [24,25] Now join the file with 'lines'[23] with the origin and destination layer that you just created [24,25] where the *Origin__1* should be joined with either the *OriginID* or *DestinationID*. Create two new columns with *OID* and *DID* and calculate the correct *Hex_ID* values. (This appears a bit difficult, but you have to make sure that the original *HexID* values are assigned to the correct origin destination pairs. When you have large files, it might be easier to do this task in MS Access using a query. Also you would then need to combine all the different line-layers that are created, in order to get one large file with all OD-pairs.
Join the housing information layer [1] with the Lines layer [23] based on the created *DID* field and the *HEX_ID*. Then create a new layer *Pop_Dest* and calculate its value based on the *Population* field from the housing layer [1]. Then summarize the information in layer [23] over the *OID* field thereby creating a sum value for *Pop_Dest*. This means that a new layer is created [26] where the first column exists of OID values (just as much as in the original housing layer) with in a second column the population that they can reach within one hour travel time.

Calculate Accessibility Index

The regular geographic Accessibility Index (PTAL) values can easily be obtained using the ‘Create Table HEX_AI’ macro, which creates a new table that contains the AI values. This information can then be joined with a polygon housing information layer, based on Hex_ID, using the regular *join table* command.

Calculate Corrected Accessibility Index

Join the housing layer [1] and the summation layer [26] based on *Hex_ID* and *OID*. Show the statistics of the *Population* field and remember the sum. Create a new field called *CPTAL* and calculate its value by the following equation: *CPTAL = AI * Pop_Dest / Total_population.*

Appendix E: Creating Travel Time and Travel Cost Assessment Layers

This appendix contains the information that is needed to perform all travel time and travel cost analysis, since the data preparation can be done at once.

Data Requirements

In order to be able to make an accessibility analysis the following data is required:

- Point layer representing the area where people live [1]
- Point layer containing information about the destinations [2]
- Point layer containing information about the location of public transport stops [3, n times]
- Line layer containing the road network [4]
- Line layer containing the urban transport routes, in the format that is used in this research [5, n times]
- Table containing information about the different transport routes [6]

Layer 2 could very well be the same layer as layer 1, however this might often be not necessary and probably takes a very long time to calculate (in the case presented here over 13 million routes would have to be calculated, which might take up to seven days).

Data Preparation

The most important thing is creating a network dataset that is able to perform all analysis. To make clear how multiple public transport modes should be handled, this appendix will discuss the method that is used to create the daladala and DART phase I combination. After that in ArcMap nearly every analysis can be made and showed on a map.

Create network model for travel time and travel cost assessment

Creating the network dataset [7] might be the most important factor in these analyses. The dataset should consist of all the point layers concerning the public transport [3, n times], the corresponding routes [5, n times] and the road network [4].
It is very important to set a good connectivity policy, which would usually mean that the road network is connected to all the public transport stops (connectivity set 1: point layers [3, n times] and road network [4]). All the other connectivity sets should consist of one point layer (e.g. [3, daladala]) and the corresponding route layer (in this case [5, daladala]). When there are two different transport modalities there are three different connectivity groups, one for the road network and one for every transport modality.
Next the *impedance attributes* should be set. These attributes are used to optimize the route choice, but can also be used to show side effects of this optimization. The different attributes will be discussed underneath. Every attribute is dependent on a set of attribute fields that are assigned to each road or network section. These fields are denoted by {*Fieldname*}. For each attribute only the from-to possibility will be discussed, since the to-from is essentially the same (in names, not in values !).

Cost

{FT_Cost} // Vehicle layers
{*none*} // Road layer

Generalized Travel Cost (GTC)

[VOT] * ({FT_TT} + {FT_Wait}) + {FT_Cost} // Vehicle layers
[VOT} * {FT_Walk} // Road layer

In vehicle time

{FT_TT} // Vehicle layers
{*none*} // Road layer

Perceived Generalized Travel Cost (PGTC)

[VOT] * ({FT_TT} + [Perceived_factor_wait] * {FT_Wait}) + {FT_Cost} // Vehicle layers
[VOT] * [Perceived factor_walk] * {FT_Walk} // Road layer

Perceived travel time

{FT_TT} + [Perceived_factor_wait] * {FT_Wait} // Vehicle layers
[Perceived_factor_walk] * {FT_Walk} // Road layer

Walking time

{*none*} // Vehicle layers
{FT_Walk} // Road layer

Waiting time

{FT_Wait} // Vehicle layers
{*none*} // Road layer

Total travel time

{FT_TT} + {FT_Wait} // Vehicle layers
{FT_Walk} // Road layer

Oneway

This attribute is not an impedance attribute, but a restrictive attribute. Usually ArcGIS sets it itself and also makes sure that FT directs the traffic in the from-to direction, while TF directs the traffic in the to-from direction.

Travel time and Travel cost analysis

The most important thing is to load the housing layer [1], destination layer [2] (which could be the same as [1]) and the network dataset [7]. Create a OD-Cost matrix and set the *impedance attribute* to the analysis you want to perform. Set the *export* to no lines and activate *oneway* and make sure that each location is snapped to either the road network or public transport stop, and ***not*** to the public transport network. In the tab in between you can select which attributes you want to store.
After running the analysis, save the resulting lines [8] and spatially join the housing layer [1] with the origins and the destinations layer [2] with the destinations assingning all attributes to the closest point. Store both layers [9,10]. Now join the file with ‘lines’ [8] with the origin and destination layer that you just created [9,10] where the *Origin__1* should be joined with either the *OriginID* or *DestinationID*. Create two new columns with *OID* and *DID* and calculate the correct *Hex_ID* values. (This appears a bit difficult, but you have to make sure that the original *HexID* values are assigned to the correct origin destination pairs.

Now you have for each OD pair a set of attributes that describe the route, given the impedance unit you selected. These figures can be summarised over origins or destinations (to calculate averages for instance) or can e.g. be used as input for a trip rate estimation.