Modelling an agent-based commercial vehicle transport system: a supply chain perspective

by

Quintin van Heerden
04361741

A dissertation submitted in partial fulfilment of the requirements for the degree of

MASTER OF ENGINEERING (INDUSTRIAL ENGINEERING)

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA
PRETORIA
SOUTH AFRICA

November 2014

Abstract

Title:	Modelling an agent-based commercial vehicle transport system: a supply chain perspective
Author:	Quintin van Heerden
Supervisor:	Prof Johan W. Joubert
Department:	Industrial and Systems Engineering
University:	University of Pretoria
Degree	Master of Engineering (Industrial Engineering)

Many state-of-practice commercial vehicle transport models are not representative of actual road transport movements, since they do not integrate supply chain elements. The objective of this research is to model stakeholders in a supply chain as agents in an agent-based commercial vehicle transport model. Furthermore, the objective is to model these agents' decisions and interactions and to ensure that the model is sensitive to changes in the supply chain. To achieve this, various steps are followed. The literature on commercial vehicle modelling is reviewed and a distinction is made between three perspectives of commercial vehicle transport models: aggregate models; disaggregate, agent- and tour-based models; and behaviour-based models. A base case agent-based commercial vehicle model, that consists of both intra- and inter-provincial commercial vehicles, is developed using a complex network and GPS records. Utilising complex network metrics, supply chain stakeholders are identified and the most important nodes in the network are extracted. One of these important nodes, an organisation in the Fast Moving Consumer Goods (FMCG) industry, provides a dataset consisting of the details of distribution data over a 10-month period.

This dataset is used in a case study to show how to model stakeholders in a supply chain. More specifically, the Carrier agent is introduced and the Carrier-Receiver interaction is modelled. Demand is generated from the dataset and the Carrier's reaction to the demand is shown through its tour planning. The effect of different levels of traffic congestion as well as the order policy of customers on the Carrier's tour planning is evaluated by showing the changes in distance travelled, tonne-kilometers moved, costs incurred, and travel time for different scenarios.

The research is of value to both organisations that need to do fleet management and government who is responsible for infrastructure maintenance and development. Organisations can utilise these models to do fleet composition analyses and evaluate the impact of changes to their logistics decision making or the effect of government interventions on their operations. Government can benefit from these models by analysing the effect of infrastructure decision-making on tonne-kilometers moved and the impact on expected travel times.

Declaration

I, Quintin van Heerden, declare that the dissertation, which I hereby submit for the degree of Master of Engineering (Industrial Engineering) at the University of Pretoria, is my own work and has not previously been submitted by me for any degree at this or other institutions.

Quintin van Heerden

Acknowledgements

I would like to express my appreciation to:

Prof Johan W. Joubert: my supervisor at the Department of Industrial and Systems Engineering at the University of Pretoria. I have gained an appreciation for research and transport modelling in the 3.5 years that I have been involved in research with him. I am grateful for his insight, guidance, and mentorship during the past 3.5 years.

Prof Kai Nagel: Professor from the Technische Universität Berlin, for his technical assistance to get some of the models up and running as well as his guidance.

Mr Stefan Schröder and Dr Gernot Liedtke: for the examples and assistance to understand how to implement the freight framework in MATSim.

Mr Pierre du Plessis: at the Council for Scientific and Industrial Research, for his assistance with the drawing of the 3D map.

My family: my dad, Prof Michael van Heerden, for his excellent assistance with language and research methodology and to help me gain an appreciation for the value of sound education. Furthermore, my mother, Anne-Marie van Heerden, for the continual support that she gave me throughout this period. Finally, my sisters and brothers-in-law, for their support and continual belief in me.

My Heavenly Father: for the ability, insight, and perserverance that I was blessed with.

Completing my Masters degree has been an incredible journey. This Masters project means more to me than merely a degree. I have learned invaluable analytical, research, and modelling skills during this period of time. I have learned how to perservere after spending numerous late nights debugging code and mastering the modelling skills required to successfully complete this dissertation. I believe that success is only significant if it creates added value for both yourself and others. I hope to further my studies and with the skills and expertise that I acquire, make a significant difference in South Africa and ultimately in the world.

Contents

List of Figures

List of Tables

Acronyms and Abbreviations

3PL	3rd Party Logistics
CRS	Coordinate Reference System
EOQ	Economic Order Quantity
FMCG	Fast Moving Consumer Goods
FSM	Four Step Model
GDP	Gross Domestic Product
JIT	Just-In-Time
KPI	Key Performance Indicator
MATSim	Multi-Agent Transport Simulation
MSA	Moving South Africa
NATMAP 2050	National Transport Master Plan
NMBM	Nelson Mandela Bay Metropole
MRP	Material Requirements Planning
NDoT	National Department of Transport
NFLS	National Freight Logistics Strategy
O-D	Origin-Destination
OSM	OpenStreetMap
RFA	Road Freight Association
SANRAL	South African National Roads Agency Limited
TAZ	Traffic Analysis Zone
TCC	The Cold Chain
TSP	Transport Service Provider
VKT	Vehicle Kilometers Travelled
WGS	World Geodetic System
XML	Extensible Markup Language

Chapter 1

Introduction

Transport is a key component in our daily lives as it provides a means to move goods, persons, and services from an origin to a destination. Whereas people rely on transport to travel to their place of employment, access education, and participate in activities such as leisure and shopping, businesses utilise transport to move goods to places where it can be used or sold. The daily functioning of the world as we know it, relies on transportation.

Consider the following example: Mr Rogers owns a small fish-and-chips shop on a street corner. He needs to ensure that he has an adequate supply of frozen fish every week to prepare and sell to his customers. Based on the number of regular customers that buy fish and chips every week, Mr Rogers has a good idea of how much stock to keep on hand for these customers. There are also other customers that buy fish and chips on an ad-hoc basis, which causes some variability and uncertainty as to how much stock Mr Rogers should keep. On the one hand he does not want to keep too much stock of frozen fish since this would result in elevated storage and electricity costs. On the other hand, if he has too little stock, first-time customers may not experience his perfected dish and he may then lose potential customer sales.

Mr Rogers needs to do weekly planning on how much stock to keep. He places an order with the local distribution centre. The distribution centre in turn receives numerous orders every week from multiple customers, such as Mr Rogers, ranging from retailers to restaurants. The distribution centre needs to consider how much stock to keep to be able to supply its customers.

The source of products may span multiple locations, both local and international. For instance, fish is caught along the coast of South Africa and transported to a facility for processing and packing, whereafter it is transported again and consolidated at the distribution centre. Frozen chicken might be imported from Eastern countries as well as sourced locally and also transported to the distribution centre. Several other frozen products follow similar procedures and are consolidated at the distribution centre from where it must be distributed to the various customers (retailers, restaurants, and Mr Rogers), situated throughout the country. These end customers rely on on-time delivery to be able to serve their customers.

When shopping, people expect to find fresh and sufficient produce on the shelves without necessarily being aware of the complexities that organisations have to endure to ensure that an adequate supply is availabe in order to meet this expectation of customers. Restaurants need stock to ensure that customers can order anything on the menu. Small shops' existence may be influenced by deliveries. The supply chain, ranging from the original sources (fishermen and imports) to the final destination (end customers) includes multiple transport legs. Each of these transportation legs is influenced by the interactions between the stakeholders involved.

Businesses rely on proper road infrastructure to perform business activities timeously, safely, and at the lowest possible cost since it affects their bottom line. Since businesses are profit driven, they endeavour to optimise their activities so as to increase profits by reducing expenditure. In particular, businesses that use transportation to deliver products to customers try to reduce transportation costs through planning and scheduling of their vehicle fleets to perform their activities as efficiently and cost effectively as possible.

Government is responsible for providing adequate transport infrastructure to ensure the efficient movement of freight in the country. In doing so, government aims to reduce the total tonne-kilometres moved, reduce travel times and ultimately reduce the cost of moving goods to improve the country's competitiveness and ensure economic growth.

1.1 State of logistics in South Africa

With recent trends in supply chains such as globalisation and Just-In-Time (JIT), organisations face increasing pressure to be more reliable, flexible and efficient to sustain a high quality of service. Furthermore, the ability to adapt to factors, such as trade policy measures from government or interventions such as tolling, provides an organisation with a competitive advantage.

Commercial vehicles are essential in supply chains to ensure the timeous delivery of products. To deliver products on time, at the right place, and in an acceptable condition, organisations incur logistics costs, which according to the 9^{th} State of Logistics Survey for South Africa, comprises inventory carrying costs, transport costs, management and administration costs, and warehousing costs (Havenga and Simpson, 2013).

South Africa is known to have high logistics costs compared to the rest of the world. Even though South Africa's logistics costs, expressed as a percentage of Gross Domestic Product (GDP), have decreased from 13.5% to 12.7% during the 2009–2010 period, total logistics costs have risen ever since. This can be attributed to inefficiencies in the country's ports, port tariffs in South Africa being 360% higher than the global average (The World Bank Group, 2014a), and the lack of water-based transportation.

Transportation accounts for 53% of the above-mentioned logistics costs in South Africa, much higher than the global sub-40% (Simpson and Havenga, 2012). This can be attributed to, amongst others, the major increase in the oil price and accordingly the local fuel price as well as long distances between suppliers and customers. Furthermore, Havenga and Simpson (2013) report in the 9^{th} State of Logistics survey for South Africa that the road-rail split is 89:11. A large amount of cargo that can be transported by rail is currently still being transported by road. Ultimately, these added costs are passed on to the consumers when they purchase the products.

The World Bank Group (2014a) quantifies countries' logistics performance with the Logistics Performance Index. It rates the performance of infrastructure, logistics competence, and timeliness, amongst other factors. South Africa ranked a very competitive 23^{rd} out of 155 countries in 2012, but has since fallen to 34^{th} position in 2014 (The World Bank Group, 2014b), below the likes of Qatar and Malaysia, which could be detrimental to economic sustainability.

A country's logistics performance is greatly influenced by its transport infrastructure development and maintenance decisions. The location, number of, and size of hubs, ports, and logistics terminals facilitate and influence competitiveness and economic growth. One of the key drivers of efficient and effective logistics systems is that of appropriate transport infrastructure.

1.2 Transport planning in South Africa

Government, who is responsible for the planning, development, and maintenance of the road infrastructure, envisioned the integration of transport operations and infrastructure in a White Paper on National Transport Policy in 1996, as well as part of the twenty-year transport strategy plan, *Moving South Africa (MSA)* in 1998. The aim was to improve spatial development, road planning, and public transport as well as reduce commuter distances and congestion. Unfortunately, as is noted in the *National Freight Logistics Strategy (NFLS)* (Department of Transport, 2005), freight volume and traffic growth have already exceeded that of the forecasts made in the *Moving South Africa* plan, long before the 2020 horison.

The National Department of Transport (NDoT) then focused on improving the quality of government regulation through the NFLS. In addition, the National Transport Master Plan (NATMAP 2050) was developed with the vision to continually upgrade transport infrastructure for sustainable urbanisation and transportation systems. An example of this is the allocation of increased funding for transport infrastructure development as part of the National Infrastructure Plan, adopted in 2012 (South African Government Information, 2013). Among these plans are the 18 Strategic Integrated Projects (SIPs), which include the development of a number of freight corridors and port upgrades in KwaZulu-Natal and the Western Cape.

While these strategies serve as wonderful potential blueprints on how to increase efficiency, achieve integration, and enhance the country's competitiveness, there is a lack of scientific rigorous backing of statements. For instance, in the NATMAP 2050, one of the actions is to identify, assess, and propose land-use and spatial development models for sustainable investment in state-of-the-art transport systems. Unfortunately, as is the case in the NATMAP 2050, strategies tend to mention state-of-the-art transport systems, but fail to address the use of appropriate transport models to evaluate effects of certain decisions.

Decision makers in government rely on transport planning models to assist with the evaluation of policies and interventions and decisions on the best spending for transport infrastructure development. Unfortunately, decisions to fund multi-billion rand infrastructure projects are frequently based on outdated, non-representative, and suboptimal models, or even made without models. The limitations of models that are used cause the decision makers to overlook key risks and the unanticipated knock-on effects of certain infrastructure and policy decisions. Such oversight often results in unexpected expenses for government and/or cargo owners. Inadequate planning results in risks being overlooked. The need for a proper representative commercial vehicle transportation model is therefore paramount.

1.3 State of commercial vehicle transport modelling

For government to plan, develop, and maintain appropriate transport infrastructure requires evidence-based decision-making. To properly evaluate interventions and scenarios, accurate, relevant and realistic transport models are required. In the case of passenger transport, adequate transport models exist to emulate private vehicles. However, commercial vehicle transport models tend to lag behind.

Many commercial vehicle transport models are adaptations of the Four Step Model (FSM), which was initially developed for private vehicle transport modelling. The four steps of the FSM are: trip generation, trip distribution, modal split, and route assignment. *Trip generation* includes modelling of the number of trips produced and attracted by a region. Next, these trips are connected in the *trip distribution* step with the use of a gravity model. Finally, modal choice is modelled based on characteristics of the trip makers. Equivalent commercial vehicle models are then developed and based on these steps by including freight specific factors, such as the size of warehouses.

Various aggregate and disaggregate models exist that attempt to overcome the use of the FSM. Two main schools of thought exist in these models: commodity flow perspective (aggregate) and activity-based perspective (disaggregate).

The commodity based perspective involves modelling freight demand as commodity flows. This is also the underlying methodology of the freight demand model for Transnet in South Africa, which is used for port and rail infrastructure planning.

In many studies, vehicular flows are often inferred from the commodity flows. Tan et al. (2004), for instance, transpose commodity flows into Origin-Destination (O-D) data and use it in a discrete event simulation. While commodity flow models are useful for aggregate economic analyses, they often overly simplify or neglect vehicle loads, load factors, and explicit activity chain structures. These models cannot accurately predict vehicle movement. Hensher and Figliozzi (2007) and Samimi et al. (2009) argue that to capitalise on logistics decision-making benefits, the movement of commercial vehicles on the road network should be understood.

The activity based perspective adressess this shortcoming through disaggregate analyses of activity chains. An activity chain is a representation of multiple stops by a commercial vehicle along its route; this typically comprises starting at a depot, conducting various pickup and delivery activities, and ending at a depot again. Good examples of activity chains that are analysed both spatially and temporally and then used to model a disaggregate synthetic commercial vehicle population are found in Joubert and Axhausen (2011) as well as Van Heerden and Joubert (2012). Joubert et al. (2010) model and simulate an agent-based commercial vehicle population and show the potential of such a model through accurate time-dependent results. This model, however, only considers intra-provincial traffic. They mention, though, that inter-provincial traffic should also be considered to improve the accuracy of results.

While activity based methods focus on the vehicle movement in much detail, which is useful for the analysis of traffic patterns, it often neglects the commodities that are carried in the vehicles. Furthermore, many of the disaggregate, agent-based, and tour-based models are based on historical data, which makes forecasting of travel behaviour difficult since it lacks behavioural aspects. For instance, logistical changes in the system such as relocation of warehouses and fleet composition cannot be captured.

Chow et al. (2010) argue that good freight demand models should contain a strong behavioural foundation. To capture behaviour, requires that each stakeholder be modelled independently to enable the emulation of its behaviour. This will ensure more realistic representation of stakeholders' actual behaviour, which will lead to more accurate modelling of the interactions between these stakeholders.

Roorda et al. (2010), Ruan et al. (2012), and Schröder et al. (2012) present frameworks on how to include multiple stakeholders in freight demand models. Joubert et al. (2013) utilise one of these frameworks, and describe how one can model different stakeholders in the supply chain and shed light on the different decisions that are made by the stakeholders. A recent implementation by Liedtke et al. (2014) was done for the Tokyo region in Japan. They tested the impact of various measures in the transport sector on logistics decision-making. They successfully indicate how the model is sensitive to toll being introduced and how shippers react by changing their vehicle size and lot sizes to account for the change in the system. However, such an implementation does not exist for South Africa, yet.

1.4 Research question

Multiple stakeholders exist in supply chains, each endeavouring to achieve their own objectives. While government, on the one hand, should foster efficient goods movement, it needs to evaluate the impact of certain infrastructure development choices or policy implications and can only adequately do so if the commercial vehicle transport models have such functionality. Companies, on the other hand, should be able to evaluate what effect traffic congestion and policies have on their logistical operations and bottom line. There is a lack of proper commercial vehicle transport models, which integrate and consider the supply chain and transportation planning domains. For these reasons, a representative commercial vehicle transport model is required that captures behavioural factors in the supply chain and that is also sensitive to changes in the logistics system.

The associated research question is:

> *What are the benefits of including supply chain stakeholders as agents in a commercial vehicle transport model?*

A secondary question is:

> *Can a commercial vehicle transport model, that includes supply chain stakeholders as agents, be used as a decision support tool in a South African context?*

1.5 Research design & methodology

To answer the research questions, the objective of this research is twofold. The first objective is to extend the work of Joubert et al. (2010) by developing an agent-based commercial vehicle transport model that consists of both intra- and inter-provincial commercial vehicles. This could be achieved by developing a probability-based sampling methodology to generate activity chains. Vehicle movements are inferred from GPS logs and used as follows.

Activity chains will be extracted from the GPS logs using the methodology of Joubert et al. (2010) as well as Van Heerden and Joubert (2012).

A complex network will be generated using the methodology of Joubert and Axhausen (2013), from the inter-activity movements of commercial vehicles. This network describes the connectedness between different facilities by considering the number of trips between each pair of facilities.

A 3-dimensional probability matrix will be generated consisting of the number of activities per chain, the start time of the chain, and the duration of the chain.

Synthetic commercial vehicle populations will be generated for both intra- and inter-provincial commercial vehicles by sampling from the complex network and 3-dimensional matrices.

Validation will be done for both populations in terms of Vehicle Kilometers Travelled (VKT).

This model should provide an accurate representation of commercial vehicle travel patterns since it considers both intra- and inter-provincial commercial vehicle activity chains. Such a model is useful for the purposes of adding a whole freight population to a combined private and commercial vehicle simulation model.

The second objective is to apply the freight framework of Schröder et al. (2012) in a commercial vehicle model that incorporates the interactions and decision-making by different stakeholders (the carrier and receiver) in a supply chain. This will be achieved as follows.

Cutomers will be modelled from actual customer orders in the distribution data of one company.

The Carrier will be modelled by supplying it with a vehicle fleet with which it creates scheduled and routed tours to attempt to deliver all the orders to its customers within a given time-window.

Sensitivity analysis will be done to evaluate the effect of traffic congestion and changes in the order policy of the customers on the fleet utilisation, tonne-km moved, distance travelled, travel time, and transport costs.

Interpretation of the results will be done to reflect the benefits of such a transport model instead of a model that is based only on observed movements of vehicles. Furthermore, to determine whether such a model can be used as a decision support tool.

1.5.1 The study area: Nelson Mandela Bay Metropole (NMBM)

The study area selected to use in this dissertation is the NMBM. The Metropolitan area has more than 1.5 million inhabitants and is developing rapidly. Decision support tools for transport infrastructure decision-making could be of major benefit to the NMBM.

The NMBM is home to the city of Port Elizabeth, and the towns of Uitenhage and Despatch, as well as the lower income communities of Motherwell and Kwa-Nobuhle. Figure 1.1 depicts the location of the NMBM in the Eastern Cape province of South Africa as well as a zoomed-in view of the road network in the NMBM area.

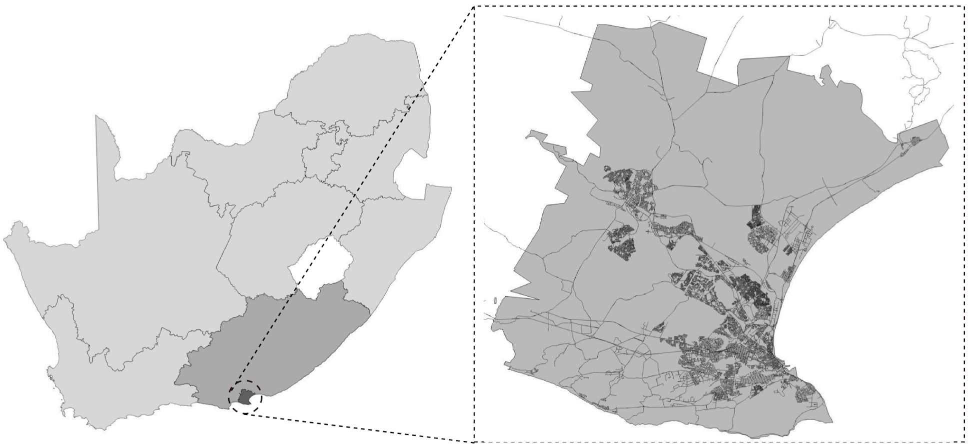


Figure 1.1: The NMBM inside the Eastern Cape Province of South Africa.

1.6 Document structure

The literature on different perspectives on freight modelling is reviewed in Chapter 2 to identify shortcomings in the different techniques. Particular reference is made to behaviour-based freight modelling. Chapter 3 demonstrates how freight movement can be modelled from historical data in the form of GPS logs. Also, a complex network is generated from the vehicle movements between facilities and this complex network is used to develop a commercial vehicle transport model that consists of both intra- and inter-provincial vehicles. Using complex network theory metrics, important nodes in the network are identified. One of the most important (most central) nodes (a company in the FMCG industry) is used as a case study in Chapter 4 to introduce the required factors to take into account in a behaviour-rich freight model. Chapter 5 covers the effect of various changes to the system on logistics decision-making. Finally, Chapter 6 contains conclusions from all the previous Chapters and sets a research agenda on how to enhance the model explained in this dissertation. It also identifies possible further avenues of research that may emanate from this research.

Chapter 2

Literature Review

The state-of-practice in commercial vehicle transport modelling appears to need to be enhanced as it is not representative of actual road transport movements. Transport modelling techniques include adaptations from the classical Four Step Model (FSM) for private vehicle movement and vehicle movements that are inferred from aggregate commodity flows.

Hunt and Stefan (2007) confirm that among the common methods utilised to account for commercial vehicles in transport models is the use of inflation factors: a mere increase of private vehicles in the model. The assumption is made that commercial vehicles' behaviour is similar to that of private vehicles. Private vehicles normally travel for business, education or leisure activities and have only a few activities per trip. Commercial vehicles travel for different reasons than private vehicles, namely to conduct business by means of pickups or deliveries. These vehicles normally have a larger number of activities to conduct, carry high value goods, and consequently affect the economy differently than private vehicles. This raises some concern about the validity of models utilising inflation factors to account for commercial vehicles.

Commercial vehicles are often regarded as a minority group in transport models, although they contribute disproportionately to the deterioration of road conditions, green-house gas emmissions, and traffic congestion. Commercial vehicles should therefore be an integral part of any transport model. Commercial vehicle movements should, though, be modelled to be representative of actual commercial vehicle behaviour and not be associated with private vehicle behaviour.

The rest of this chapter describes the types of commercial vehicle transport models as contained in the literature. Some shortfalls in the modelling techniques are also identified.

2.1 Aggregate freight transport modelling

Aggregate freight transport models are primarily based on aggregate commodity flows or the FSM. The four steps of the FSM, as described by Ortuzar and Willumsen (2011), follow with a graphical representation in Figure 2.1.

Firstly, *trip generation* consists of the calculation of the number of vehicle trips produced by each region (the origin of a trip) as well as the number of vehicle trips attracted by each region (the destination of a trip). A *region* is typically modelled as a Traffic Analysis Zone (TAZ); that is, a demarcated area of similar land-use, socio-economic status of the people, or population density.

Vehicle trip productions and attractions are normally predicted from land-use forecasts, socio-economic factors, and household demographics. The factors that are normally included in the trip generation models include: household income, car ownership, value of the land, accessibility, and residential density. A dense residential area may produce more work trips than an industrial area since people depart from their homes to go to their places of employment.

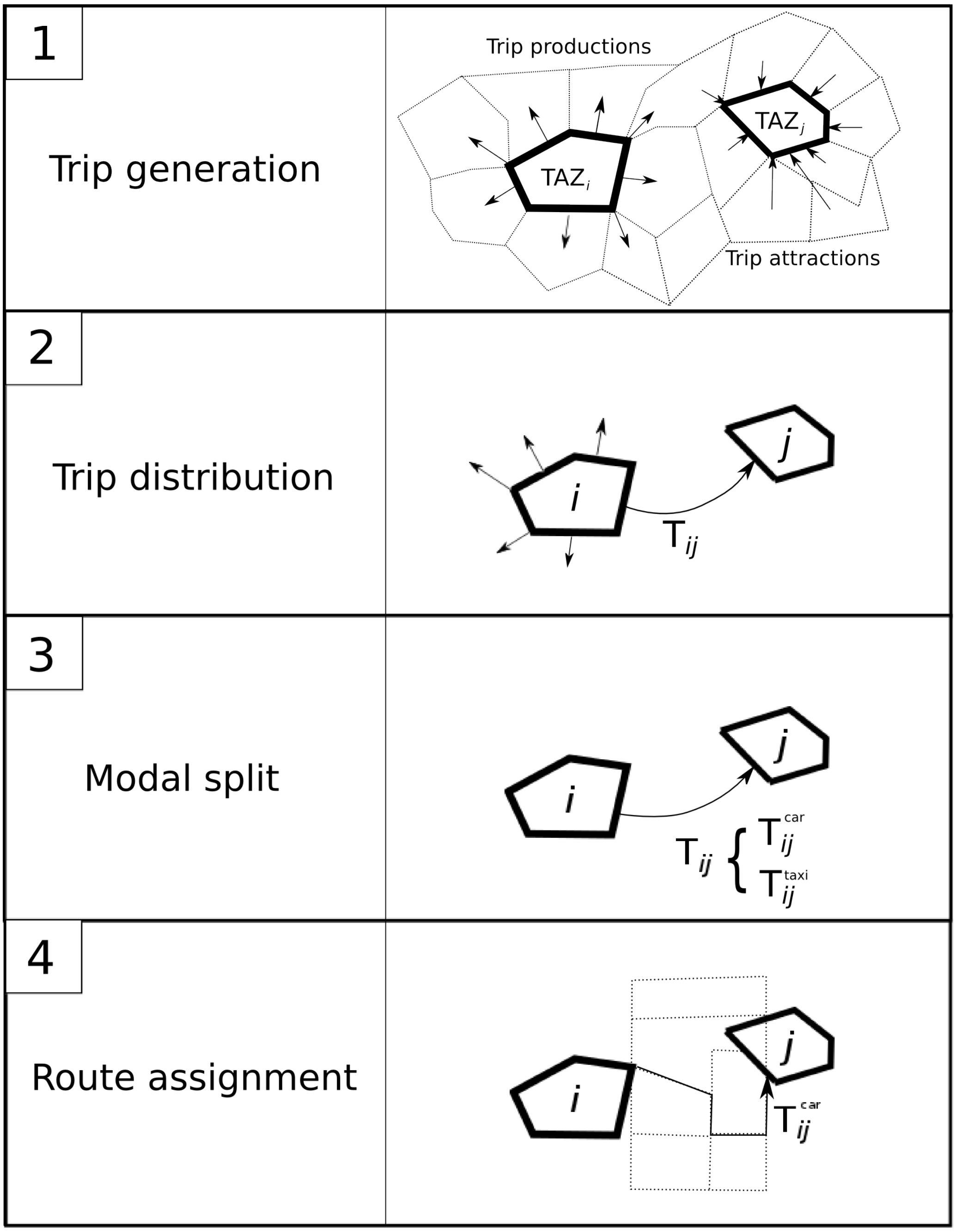


Figure 2.1: The Four Step Model. Adapted from Meyer and Miller (2001).

Conversely, an industrial area may attract more work trips than a residential area because people are more likely to go to their places of employment in an industrial area.

Secondly, the *trip distribution* step is used to estimate the fractions of vehicle trips connecting each pair of TAZs. Among the methods of estimation are: uniform growth factor methods, singly or doubly constrained methods, logit models, or the mostly used gravity model. The factors taken into account in the gravity model include: population density of both the production and attraction zones, the distance between the two zones, and costs associated with the trip. The result of this step is an Origin-Destination (O-D) matrix, describing the number of trips between each zone pair.

Thirdly, commuters' choice of transport mode is modelled to obtain a *modal split*. Factors considered in this step include: characteristics of those making the trip, the journey, and the transport facility. For instance, people who own cars are more likely to choose a car over public transport for convenience sake. Binary logit models are normally developed from data from a stated preference survey. These models then describe the probability of a specific mode being chosen for a specific trip.

Fourthly, in the *route assignment* step, road users are incrementally assigned to links on the road network between the origins and destinations, reducing the available capacity each time more vehicles are assigned to the links. An increase in vehicles on a link reduces the travel speed, which is calculated using speed and flow curves. Therefore, as links are loaded, other links may become more preferential in the choice of a route. This process continues until all vehicles are assigned to the network and an equillibrium is reached.

The FSM was originally developed for private vehicle transport modelling. However, private vehicle and commercial vehicle behaviour differ distinctly in terms of vehicle mix, the reason for travel, and the number and type of activities they perform (Hunt and Stefan, 2007). In this regard, aggregate freight modelling often adapts the FSM to cater for these differences. Ortuzar and Willumsen (2011) explain how to include freight specific factors in adaptations of the FSM for freight models by considering the number of employees of the organisation, the roofed area of organisation, sales, and the total land area of the organisation.

Marker Jr and Goulias (1998), for example, make use of a three step zonal approach consisting of trip generation, trip distribution and traffic assignment. Trips are estimated from the number of employees working at each organisation. These trips are then aggregated to TAZs. A doubly constrained gravity model is then used for the trip distribution step and traffic assigment is achieved through network loading until user equilibrium is reached. Vehicle movements are thus based on zonal and organisational characteristics and attributes.

Among the other common methods found in aggregate freight modelling, is the inference of vehicle flows from commodity flows. Tan et al. (2004), for instance, transpose commodity flows into O-D data and use it in a discrete event simulation. This method is adequate for aggregate economic analyses, but for the sake of traffic congestion and travel time prediction, the vehicle movements should be analysed and understood and their behaviour not merely inferred from commodity movements (Liedtke and Schepperle, 2004). Commodity movement cannot accurately predict vehicle movement since the characteristics of the activity chains are unknown.

Hensher and Figliozzi (2007) as well as Samimi et al. (2009) support this argument by comparing both aggregate and disaggregate freight models in their review of freight modelling techniques. They highlight the gap in the understanding of freight vehicles' behaviour and how this hinders the ability to capitalise on the decision-making benefits that disaggregate freight modelling holds.

2.2 Disaggregate, agent- and tour-based freight modelling

Hunt and Stefan (2007) model commercial vehicle movement in three parts: external-internal movements, fleet-allocation, and tour-based movements.

External-internal movements are modelled by means of a singly-constrained gravity model. Trips that originate outside the study area and end within the study area are assigned as cross-border trips. Next, fleet-allocation is done by generating aggregate vehicle trips between origins and destinations. These trips do not deal with specific shipments, but rather cover areas of interest or road links. An example of such an allocation is that of courier services. Tour based movements are included in the model using a tour-based micro simulation of which the output is a zone-to-zone trip table. A household travel model is used in conjunction with the zone-to-zone trip table to find network equilibrium. Although specific tours are modelled, the tour generation step is still based on an aggregate trip generation model. This method utilises both aggregate and disaggregate modelling techniques.

De Jong and Ben-Akiva (2007) consider the detailed movement of commercial vehicles. They study the frequency of activity chains, distribution centre use, and the mode of each of the activity legs. This approach simultaneously considers activity chain structure and management philosophies such as Economic Order Quantity (EOQ), load consolidation, and the use of distribution centres and trans-shipment hubs. The model converts commodity flows into disaggregate firm-to-firm flows. Micro-simulation is used to simulate logistic decisions such as shipment size or mode choice. Although logistic decisions are incorporated in this model, vehicle movements are still inferred from commodity flows.

Joubert and Axhausen (2011) improve the understanding of disaggregate commercial vehicle movement. They consider both the temporal- and spatial-characteristics of commercial vehicle activities. They extract activity chains from raw GPS data and analyse these chains by considering the number of activities per chain, activity and chain start-and-end times, and chain durations.

Joubert et al. (2010) expand on this by generating intra-provincial traffic from the extracted chains. They test their model using the Multi-Agent Transport Simulation (MATSim) toolkit by simulating an agent-based transport model which consists of both private and commercial vehicles, and obtain accurate time-dependent results. They emphasise, however, that modelling of commercial vehicles should also consider inter-provincial vehicles.

Van Heerden and Joubert (2012) follow the extraction method of Joubert et al. (2010) to extract inter- and intra-provincial commercial vehicle activity chains of 41 711 vehicles over a six-month period. In their analysis, they illustrate that for different days of the week, vehicles behave differently. They distinguish both temporally and spatially between inter- and intra-provincial vehicles as well as considering differences in activity chain characteristics on different days of the week. The study then further focuses on inter-provincial vehicles by analysing where these enter and exit the study area (the Gauteng Province of South Africa) on an hourly basis, and the number of activities that these perform inside the study area.

Although some of the models discussed are promising to represent commercial vehicle movements, there are still various assumptions to be challenged and limitations to be addressed. Many of the aggregate models contain assumptions that are too simplified in the inference of vehicle flow from commodity flow. Also, specific movements of vehicles are difficult to predict, since the reasons for vehicle movements are not always understood.

Many of the disaggregate, agent- and tour-based models neglect the commodities that are carried. Some models utilise simplified load factors or disregard these entirely. There are hybrid models that attempt to utilise both methods to cater for both commodities carried and model the vehicle movements, yet many of the underlying methodologies are still based on the FSM, and this leads to the accuracy and ability of such models being questioned.

Activity-based commercial vehicle transport models are often based on historical data, which makes forecasting of travel behaviour difficult since it lacks behavioural aspects. Furthermore, logistical changes in the system, such as relocation of warehouses and changes in fleet composition cannot be accurately captured and these models are not sensitive to such changes.

Holguin-Veras (2000) argues that two types of flow exist: commodity flow and vehicle flow. Commodity flow represents actual demand and vehicle flow represents traffic. He too argues that commodity flow models, which capture economic mechanisms, often neglect vehicle flow elements, such as empty trips (when a commercial vehicle returns to the depot without any commodities loaded). He also mentions that trip-based models take empty and loaded trips into account, but do not capture economic mechanisms associated with the commodities. To address this issue, he presents a framework for the integration of elements of both commodity and vehicular flow models. Producers are responsible to create or transform the commodities that are demanded by consumers. Carriers are responsible to move the commodities from the producer to the consumer. Lastly, government is responsible to regulate transport infrastructure.

Chow et al. (2010) reiterate that good freight demand models should contain a strong behavioural foundation. The commercial vehicle flow that is observed on the roads is actually the result of a much more complex set of interactions between different stakeholders in supply chains.

2.3 Behaviour-based freight transport modelling

In a review of city logistics modelling trends, Anand et al. (2012) reveal that recent commercial vehicle transport modelling efforts have increasingly started to incorporate stakeholder behavioural elements. Modelling techniques now tend to capture different stakeholders in a supply chain as well as the dynamics of city logistics processes.

City logistics, as described by the Institute for City Logistics (2013), constitutes the optimisation of logistics and transport activities in urban areas as exercised by private companies. These companies need to endeavour to reduce the number of their commercial vehicles on the city roads by taking into consideration, among others, fuel consumption and traffic conditions.

Holguin-Veras and Patil (2005) argue that commercial vehicles perform interrelated, multi-trip tours, which are the results of underlying logistical decisions. Ruan et al. (2012) similarly argue that commercial vehicle deliveries are the results of the decisions of the stakeholders in the system: the shipper, carrier, and receiver.

Roorda et al. (2010) explain that the understanding and representation of different stakeholders in a freight transport system is fundamental. They argue that the interactions between the stakeholders and the changes in their actions and interactions over time are important behavioural aspects that need to be included in transport models. They present a framework of an agent-based simulation model that captures the various functions and interactions in the market between business establishments, firms, and logistical facilities.

A business establishment is an organisation at a single location that has internal resources to process, produce, or store commodities. This establishment could also provide logistics services such as warehousing or transportation. Business establishments interact with other businesses through contracts. Roorda et al. (2010) further explain that firms constitute one or more business establishments, where business establishments within the same firm are more likely to do business with one another than with external customers. For these logistical networks to function properly, requires various decisions by the different actors. The authors divide the decision-making in firms and business establishments into four distinct groups as follows.

Fundamental business decisions are typically long term decisions that include decisions such as whether or not to start the firm or business establishment within the firm, which commodities to produce and which services to provide, and decisions on the location or relocation of facilities.

Supply chain management decisions are medium to long term decisions to enhance supply chain efficiency. These decisions are based on how to best utilise internal resources, whether or not to sell excess capability, and to form relationships with other business establishments.

Market interaction decisions span between short and medium in term. These decisions include any interaction with the market such as to supply certain commodities and services or to obtain commodities and services that are required for the proper functioning of the business establishments.

Operational decisions are short term decisions with which to optimise processes to perform services as efficiently and effectively as possible.

Considering the Fast Moving Consumer Goods (FMCG) scenario in Chapter 1, the organisation that owns the distribution centre would make *fundamental business decisions* to determine the best location from which to serve its customers. Furthermore, this organisation may try to establish long term contracts with its suppliers (the local fishermen and international suppliers) to capitalise on economies of scale. *Market interactions* would include taking orders from customers and supplying in their needs. Finally, operational decisions would include how to best service the customers by optimising the vehicle fleet composition and doing routing and scheduling of vehicles with the aim of minimising cost.

Similarly, Schröder et al. (2012) allow, within the MATSim setting, for more realistic modelling of freight agent behaviour, where an agent is modelled as a representation of a stakeholder in a supply chain. They distinguish between three agents: the Shipper, the Transport Service Provider, and the Carrier.

The *Shipper* is introduced as an entity wishing to move goods to a customer. This commitment to serve its customers is in the form of a contract. These contracts determine the type, value and quantity of commodities to be delivered to the customers. Shippers can own warehouses, which are subject to long term decision-making, similar to that of Roorda et al. (2010). Finally, the Shipper divides the orders of its customers into smaller shipments and schedules deliveries to each of its customers and then has the option to choose a *Transport Service Provider* to carry out the transportation.

Transport Service Provider agents compete for the shippers' contracts in terms of specific shipments by providing rates at which they can fulfil the business obligation. Once a Transport Service Provider secures a contract, it establishes the freight chains that will be required to transport all the commodities from the sender to the receiver. These chains could comprise multiple facilities, trans-shipment facilities, and multiple (multimodal) legs.

Carrier agents are contracted by the Transport Service Provider to execute the different freight chains. The Carrier has an obligation to execute the transportation legs by adhering to both the pickup and delivery time-windows imposed on it and utilising its vehicle fleet as optimally as possible. The carriers have the ability to perform the modal choice (also the associated fleet sizing) and the vehicle routing and scheduling.

To model freight agents at this level, requires extensive knowledge of each agent's behaviour as well as adequate data to capture such behaviour. It is also evident from the literature review that the various stakeholders make decisions and that these decisions ultimately lead to vehicle flow. It is important to capture all types of commercial vehicle traffic (both intra- and inter-provincial) in commercial vehicle transport models to ensure that commercial vehicle movements are represented more realistically than in traditional transport models. Furthermore, it is important to identify the key role players in a supply chain and incorporate their behavioural aspects to adequately evaluate the effect of changes in their behaviour on the transport system. In Chapter 3 a closer look is taken at the MATSim Toolkit, modelling intra- and inter-provincial commercial vehicle populations and how to identify key roleplayers in a freight network using complex network theory metrics. The Chapter is concluded with a description of how to incorporate logistics behaviour into a MATSim model.

Chapter 3

Freight modelling in MATSim[1]

Agent-based transport modelling is a modelling paradigm in which each person in real-life is represented by an agent in the transport model. Each person is therefore modelled individually with decision-making capabilities. Accordingly, each commercial vehicle can be modelled as an agent as well.

Agent-based modelling has shown much potential over the past few years to accurately model large-scale transport scenarios. The Multi-Agent Transport Simulation (MATSim) Toolkit is a simulation package that can be used to simulate such models. Using MATSim, two promising implementations have been made in South Africa. Firstly, Fourie (2010) compares MATSim with the state-of-practice simulation package, EMME/2, for the Gauteng Province in South Africa. EMME/2 is a transport forecasting model based on the traditional four step modelling paradigm. By using the same dataset in both packages, he shows that MATSim delivers much more realistic results by means of accurate travel time predictions.

Similarly, Joubert et al. (2010) show how commercial vehicles can be simulated in conjunction with private vehicles over a 24-hour time-period, using MATSim. They also show accurate time-dependent results. Although the freight population in this model had some assumptions and draw-backs, such as that the population only includes intra-provincial traffic, it served as a proof of concept that an integrated passenger and commercial vehicle transport model for South Africa could be simulated.

The aim of this Chapter is to address the shortcoming in Joubert et al. (2010) by including both intra- and inter-provincial commercial vehicles in an agent-based model. The MATSim toolkit is the simulation package selected to be used in this dissertation to extend the population of Joubert et al. (2010) and also because of its ability to model individual agents. When endeavouring to incorporate logistics behaviour and decision-making in the transport model, one needs to model each stakeholder separately as an agent. MATSim's structure is described in more detail in the following section. Thereafter, a commercial vehicle synthetic population consisting of both intra- and inter-provincial commercial vehicle activity chains is developed.

[1] Parts of the content in this Chapter appeared in:

- Joubert, J. W. and Van Heerden, Q. (2013). Large Scale Multimodal Transport Modelling, Part 1: Demand Generation. In 32^{nd} *Annual Southern African Transport Conference*, pages 644–652. ISBN: 978-1-920017-62-0. Available from: <http://hdl.handle.net/10204/6963>
- Van Heerden, Q. and Joubert, J. W. (2013). Large Scale Multimodal Transport Modelling, Part 2: Implementation and Validation. In 32^{nd} *Annual Southern African Transport Conference*, pages 653–661. ISBN: 978-1-920017-62-0. Available from: <http://hdl.handle.net/10204/6972>
- Van Heerden, Q. and Joubert, J. W. (2014). Generating intra- and inter-provincial commercial vehicle activity chains. In *Procedia - Social and Behavioral Sciences*, volume 125, pages 136–147

Finally, a description is given how to incorporate logistics behaviour using MATSim.

3.1 The MATSim Toolkit

The MATSim Toolkit is an open-source, agent-based simulation environment in which one can model large-scale transport models (MATSim Development Team, 2012). Figure 3.1 depicts the simulation process applied in MATSim. A description of each step follows.

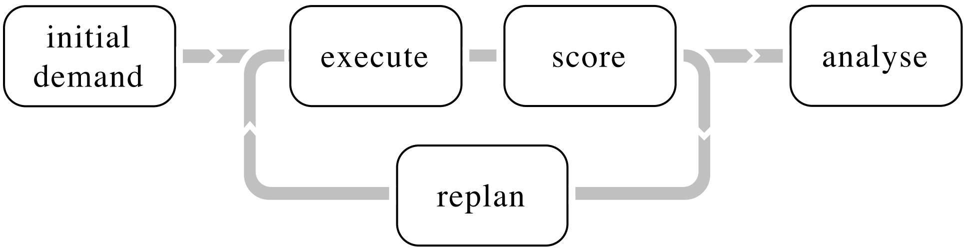


Figure 3.1: The simulation process in MATSim.

3.1.1 Initial demand

To be able to run a simulation model, one needs a population that makes use of transport. This modelled population, also known as a *synthetic population*, represents a collection of private individuals, also known as *agents*.

Agents have attributes that include, among others, income level, car ownership, and age. Each agent aspires to participate in a sequence of *activities* and has the freedom to do activity timing. To attend to the activities, the agent needs to travel between activities. This travelling is known as *transport legs*. The sum of all activities and transport legs during the day constitutes an agent's *plan*.

As in real life, agents in the simulation need a road network to travel on. The road network is generated from OpenStreetMap (OSM) (OpenStreetMap Contributors, 2013) and consists of a set of links, described by the attributes: length of the links, capacity (in vehicles/hour), free-speed, and number of lanes.

The simulation is set up using a configuration in the form of a *config file*, which contains details pertaining to the simulation. The different groups of configurations are seperated into modules. Among the modules are those of *strategy, road pricing, counts, households*, and *network*. The *strategy* module describes the different options given to agents for replanning purposes. The *counts* module allows the modeller to add a file containing actual traffic counts on certain road links, which can be used to validate the accuracy of the simulated traffic counts. The *households* module allows the modeller to add demographic data at a household level. The *network* module allows the modeller to add a road network. Various other modules exist to further customise a scenario, some of which are optional and others which are necessary for a simulation model to run.

3.1.2 Execute

MATSim has the capability to model and simulate full 24-hour days. A day of 24 hours is one iteration in the simulation.

The input files from the *initial demand* step are fed into the simulation and each agent has a choice set consisting of different plans. Each agent selects one plan from from the choice set. All agents simultaneously execute their chosen plans on the road network provided.

Since agents do not initially expect traffic congestion when selecting an initial plan, the expected plan and executed plan often differs. To try and quantify the difference between the expected and executed plans, the agents need to score their plans.

3.1.3 Score

Scoring of plans involves utility theory. Positive utility is earned for spending time at work and home, whereas negative utility is earned for being delayed in traffic or being late for an activity. In MATSim, a plan is scored by primarily using the *Charypar-Nagel* scoring function:

$$U_{total} = \Sigma_{i=1}^{n}\left[U_{performance,i} + U_{late,i} + U_{travel,i}\right] \tag{3.1}$$

The total utility for the executed plan is given by U_{total}. n refers to the number of activities in the plan. Whereas $U_{performance,i}$ is the positive utility earned for performing an activity, $U_{late,i}$ is the negative utility earned for arriving late for an activity. Finally, $U_{travel,i}$ is the negative utility earned for travelling to an activity. Total utility can be increased by reducing time spent on travelling to activities and increasing actual time spent performing the activities.

Each agent therefore scores his/her day's activities and transport legs, which gives the agent an indication of the "success" of the day's execution. Based on this score, the agent may replan for the following iteration of the simulation.

3.1.4 Replan

In the configuration file of the simulation, the modeller can provide the agents with different strategies for replanning purposes. One strategy allows agents to depart from home earlier or later with the expectation of increasing utility by avoiding peak hour traffic. Another strategy is to change the choice of route or transport mode with the expectation of increasing utility by driving on a route or mode less used. Another strategy is that an agent is allowed to substitute the executed plan during the current iteration with another plan from memory. The strategies that are available can be set in the *config* file with weightings to ensure that a certain strategy is chosen more often.

The process iteratively continues while agents endeavour to maximise individual utility. The system eventually reaches a steady state when the average score of executed plans converge.

3.1.5 Analyse

The analysis process involves validating and verifying whether the model reacts correctly and how accurate the simulation is in comparison to real life traffic counts. Here the input files from the initial demand file play a role again. For certain links in the road network, the traffic counts could be validated against simulated traffic counts.

The next section focuses on the *initial demand* step in the MATSim process. From Van Heerden and Joubert (2012) it is evident that there is a difference in the activity chain characteristics of intra- and inter-provincial commercial vehicle activity chains. In the next section, a commercial vehicle synthetic population is generated that consists of both intra- and inter-provincial activity chains.

3.2 Generating intra- and inter-provincial commercial vehicle activity chains

The dataset used in this section is a 40Gb flatfile containing the detailed GPS logs of 41 711 commercial vehicles for the 6-month period of 1 January 2009 – 30 June 2009. The file contains 6 fields of data per row. These fields, as described in Van Heerden (2011), consist of data in the format: (1) a unique vehicle identification number, which does not reveal any information about the customer; (2) a Unix time stamp (measured in seconds from the epoch, 1 January 1970); (3) a longitude value in World Geodetic System (WGS) 84 Coordinate Reference System (CRS) decimal degrees; (4) a latitude value in WGS 84 CRS decimal degrees; (5) a vehicle status identifier; (6) the vehicle's speed.

Such a file with raw data is cumbersome to work with and therefore the object oriented programming language, Java, was used to access and rework the data into a more usable format. This process follows.

3.2.1 Activity chain extraction

Joubert and Axhausen (2011) distinguish between different types of commercial vehicle activities: major activities are activities in excess of 300 minutes (or 5 hours) in duration and minor activities are activities shorter than 300 minutes in duration. Major activities are typically depot stops, whereas minor activities represent pickups or deliveries. Their activity chain extraction method is followed in this dissertation. The process is depicted in Figure 3.2 and a description follows.

Splitting of original file: Individual files were created for each vehicle by extracting only the GPS data pertaining to the specific vehicle. This resulted in 41 711 individual flatfiles containing GPS logs, with the *VehicleId* as filename.

Sorting the records: The GPS records in each file were not necessarily sorted chronologically. The records were subsequently sorted chronologically according to the time stamp field.

Extracting activity chains: For each vehicle file, all minor activities are ignored to the point where the first major activity is reached. From this point onward, minor activities are included until the next major activity is reached, and this forms an activity chain. The last major activity is then used as the first major activity of the next activity chain, since the location is the same. An example of this is a truck returning to the depot and eventually departing from this depot again in a subsequent activity chain.

This process is continued until all activity chains are extracted. Whereas Joubert and Axhausen (2011) only considered chains with at least one minor activity, in this dissertation chains consisting of only two major activities and no minor activities, are also considered. Such a chain is depicted in row C in Figure 3.2. The reason for this approach is that two consecutive major activities could be a relocation to another facility for maintenance, refuelling or other purposes than its usual business commitments. The movement of the vehicle will have an effect on traffic since it departs from the depot.

Clustering of activities: The clustering approach of Joubert and Axhausen (2013) is used to cluster activities in the activity chains. When at least 15 activities were observed within a radius of 30 meters, a cluster was created for those activities and an artificial *facility* was created. An artificial facility is any location where a vehicle stopped, thus it could be a depot, petrol station, or even alongside the road. The locations of the activities forming the cluster were changed to the centroid of the cluster/facility. For each pair of subsequent activities, where the distance between the two activities is less than 100 meters, the activities were merged into one activity, with the duration of the activity being the sum of the two activity durations.

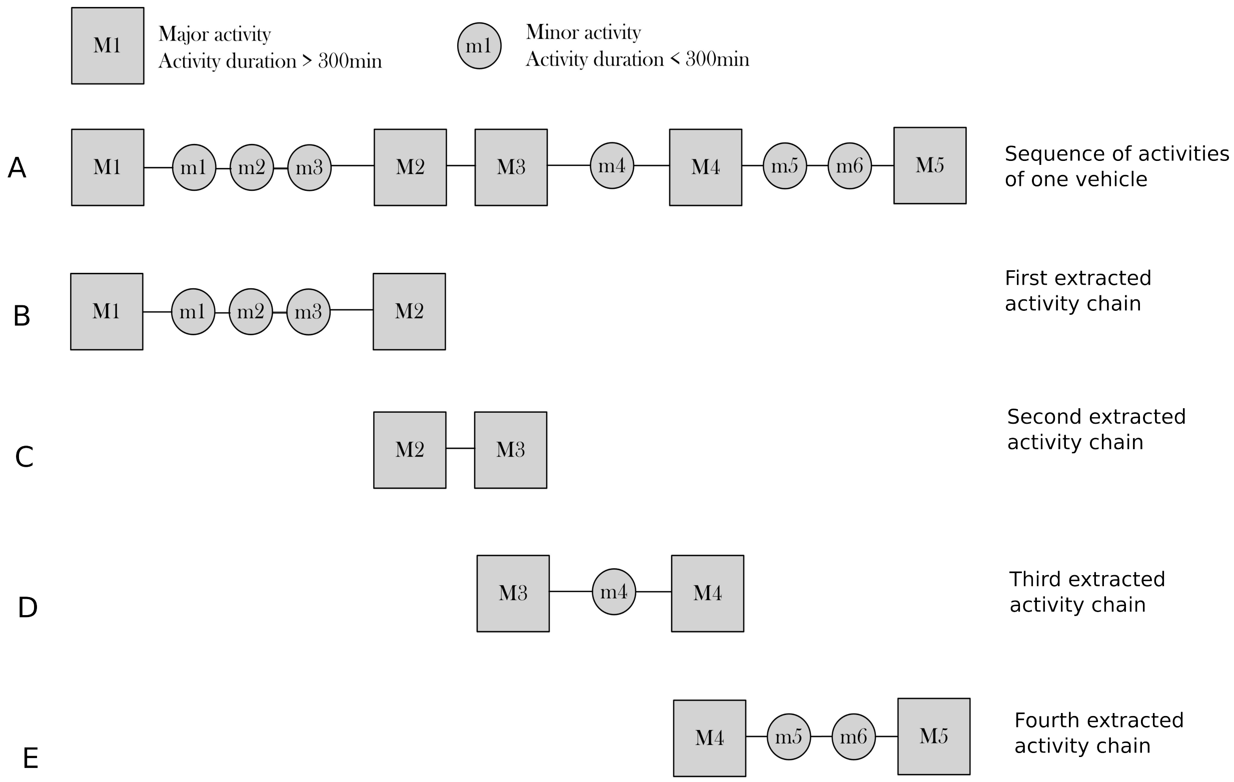


Figure 3.2: Extraction of activity chains (Adapted from: Van Heerden (2011))

3.2.2 Complex network generation

Joubert and Axhausen (2013) present the methodology to generate a complex network from commercial vehicle GPS logs. A complex network, as described in graph theory, is a set of nodes, also known as *vertices*, that are connected by a set of links, also known as *dyads*. This network describes the connectivity between the nodes. The methodology to generate such a network follows.

Each activity chain of each vehicle file was considered. For each activity pair in each activity chain, if both activities were associated with a facility, the origin and destination locations were noted. If the locations did not exist, they were created as new nodes in the complex network. The connection, or *dyad*, between the two nodes are specified by a link with an initial weight of 1. If the dyad between these two nodes existed, the weight was incremented by 1. This resulted in a network of vertices, connected by weighted dyads, that indicated the vertices between which the most activity took place.

Following this methodology, a complex network was generated for both intra- and inter-provincial commercial vehicles while using the Nelson Mandela Bay Metropole (NMBM) as a study area. Based on Joubert and Axhausen (2011), the distinction between intra- and inter-provincial vehicles are: intra-provincial commercial vehicles are those vehicles that perform more than 60% of their activities within the study area; and inter-provincial commercial vehicles are those vehicles that perform less than 60% of their activities within the study area. Vehicles that do not perform any activities within the study area were omitted.

3.2.3 Population synthesis

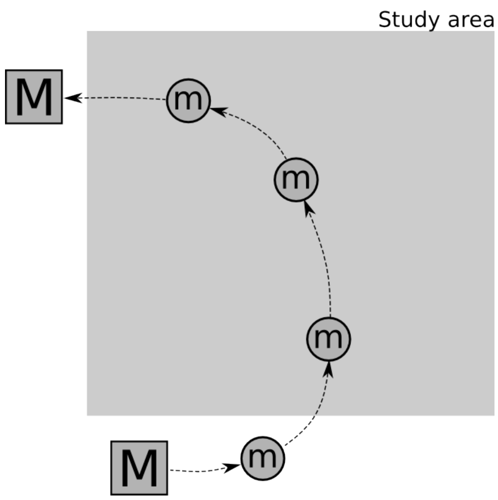
Using the activity chains and complex network, the synthetic commercial vehicle populations were generated as follows.

Intra-provincial traffic: Using the activity chains of each vehicle that is classified as an intra-provincial vehicle, each chain's start time, number of activities in the chain, and chain duration were recorded and a three-dimensional conditional probability matrix was populated with these results. Using this matrix, a start time, number of activities per chain, and chain duration were sampled.

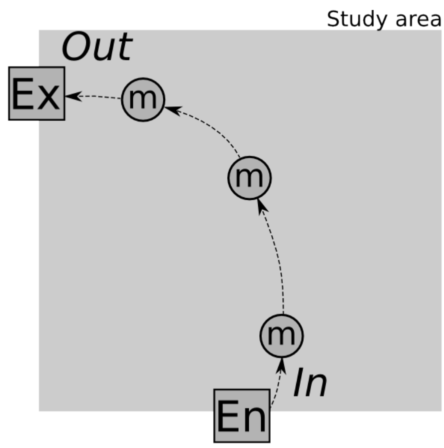
One complex network theory metric that describes the importance of a node, is degree centrality, which is a value describing the total weight of connections to that node, both incoming and outgoing. Only locations where major activities took place were considered. Using the degree centrality as a weight, a random location was sampled as a starting point for the chain. All outgoing links were considered and using their weights, the next node's location was sampled. This process was continued until the number of activities sampled was reached, adding minor activities to the activity chain. For the final activity, only major activities were considered and this ended the activity chain.

Inter-provincial traffic: Using the activity chains of each vehicle classified as an inter-provincial vehicle, the three-dimensional matrix was again constructed and used for sampling.

A distinction was made between two types of inter-provincial activity chains. Firstly, *in-out* chains are those chains that start outside the study area, enter the area, perform one or more activities inside the area before leaving the area again. Such a chain can be seen in Figure 3.3a. The portion(s) of a vehicle's chain that occurs outside the province was replaced with a single entry (or exit) activity and the activity start time was changed to the time of entry. An example of such a chain is shown in Figure 3.3b. From these activity chains, a conditional probability matrix was again populated from the start time, number of activities per chain, and the chain duration. Additionally, a matrix was constructed, noting the entry and exit gates of activity chains to and from the study area.



(a) An *in-out* chain

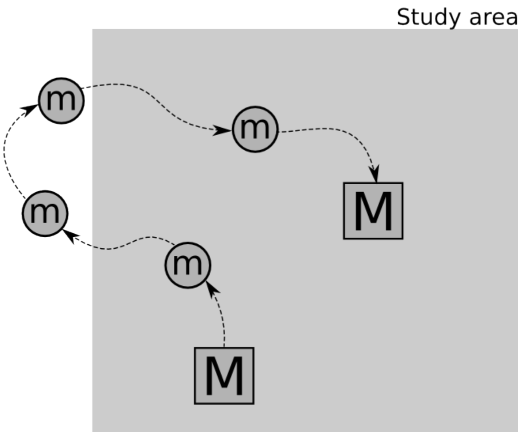


(b) *In-out* chain with exit and entry activities

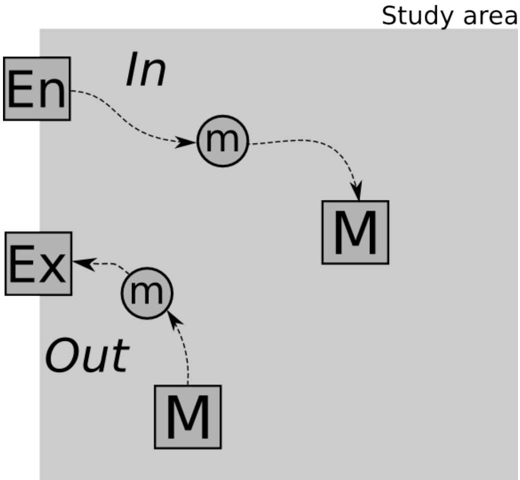
Figure 3.3: Removing activities of an in-out chain that are outside the area

Using the three-dimensional matrix, gate-pair matrix, and complex network, *in-out* chains were generated as follows. A gate-pair was sampled from the gate-pair matrix and the entry and exit locations were used as the start and end of the activity chain. A start time, number of activities, and chain duration were sampled from the three-dimensional matrix. For the number of activities sampled, activity locations were sampled from the complex network following the same procedure as in the previous section.

Next, *out-in* chains are those chains that start inside the study area, leave the area, perform a number of activities outside the area before returning and ending inside the area. An example of such a chain is shown in Figure 3.4a, where it can be seen that one activity chain essentially consists of three portions, of which only two portions are within the study area. The activities outside the area are excluded and an entry and exit activity is generated where the vehicle crossed the border of the study area, as is shown in Figure 3.4b.



(a) *Out-in* chain



(b) *Out-in* chain divided into two parts

Figure 3.4: Removing activities of an out-in chain that are outside the area

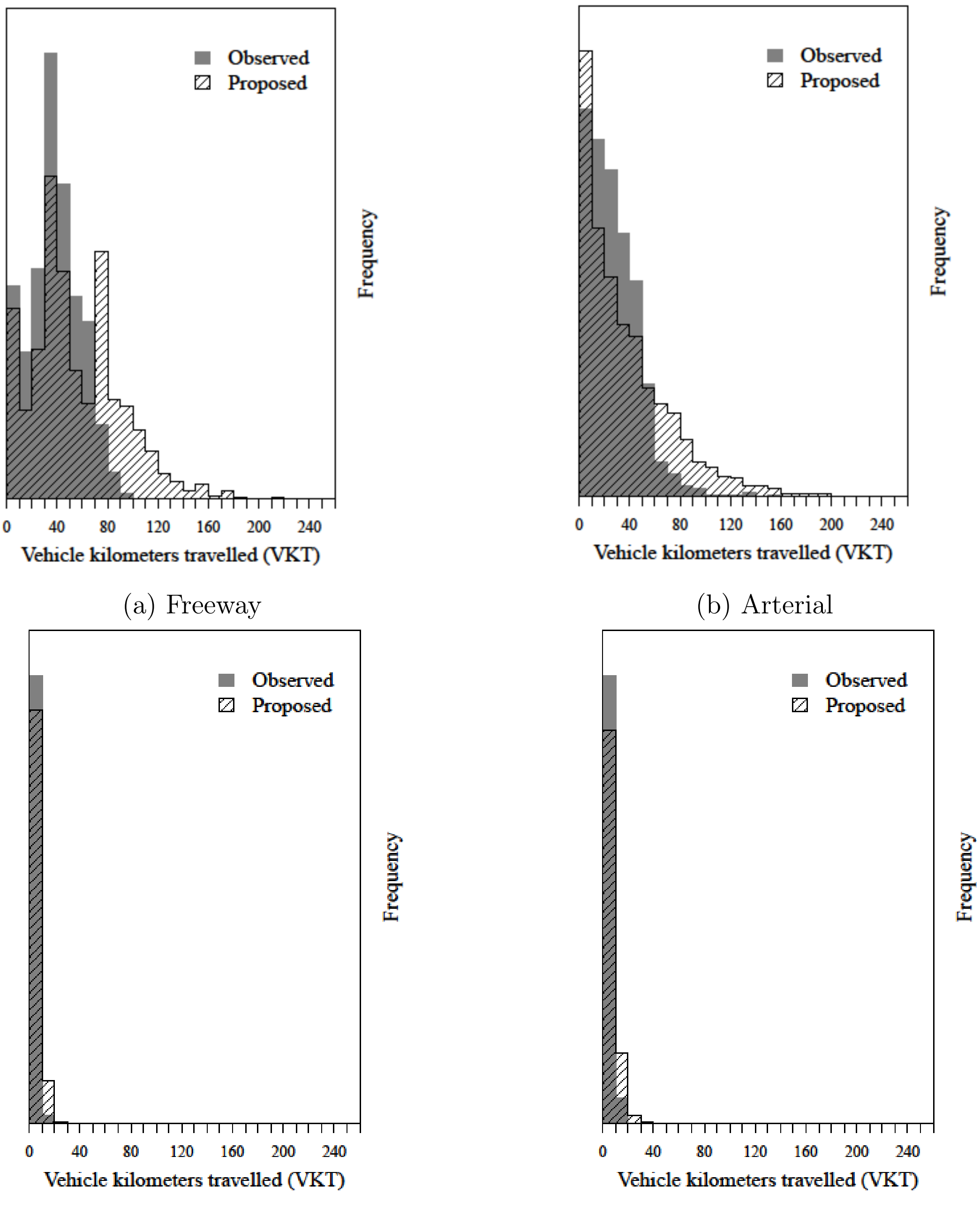
A distinction is made between the first (before leaving the area) and second (after returning to the area) portion of the chain separately. A conditional probability matrix was again generated from the start time, number of activities, and the chain duration. The first major activity was sampled from nodes that were flagged as possible major activity locations. The required number of minor activities was then sampled from the complex network. A gate pair was also sampled, and the first gate was used as the final major activity, i.e. exit activity that completed the out portion of the chain.

For the returning part of the *out-in* chain, the first major activity was the second gate from the gate-pair. Another conditional probability matrix was generated of the start time, number of activities, and duration. Minor activities following the entry activity were sampled from the complex network, and the final major activity was sampled from the nodes flagged as possible major activity locations.

3.2.4 Population validation

The accuracy of the population was evaluated in terms of the Vehicle Kilometers Travelled (VKT). VKT is a proxy to measure the impact of commercial vehicle activity on the roads and environment. Van Heerden and Joubert (2012) show that the activity chain characteristics of intra- and inter-provincial commercial vehicles differ and should be addressed separately. Accordingly, the VKT for both intra- and inter-provincial commercial vehicle activity chains were analysed.

Since the activity chains were generated from GPS logs that are infrequent and with consecutive signals not necessarily consistent, accurate map matching is not possible and the VKT was estimated. The estimated observed VKT was the result of an A*-landmark routing algorithm that determines the shortest path between successive activities (Hart et al., 1968). This VKT was compared to the VKT from the synthetic commercial vehicle population for different road types, based on the road types classified in the Highway Functional Classification System on *OpenStreetMap*. Figure 3.5 depicts the comparison.



(a) Freeway (b) Arterial

(c) Street (d) Other

Figure 3.5: Comparison of VKT on different road types

A favourable comparison between VKT from observed data and the synthetic population for the different road types was obtained. Slight under-estimation is present in the shorter activity chains, whereas over-estimation is present in the longer activity chains. This could be due to the fact that only locations associated with a facility were used when sampling the next location from the complex network. This means that from a specific location, there may not have been outgoing links to other nodes. In such a case, a random node was selected within a certain radius from the node and the sampling of the remaining activities' location continued from that node.

The commercial vehicle population generated in this section consists of both intra- and inter-provincial activity chains. This population is an accurate representation of actual commercial vehicle movements. Furthermore, this population is a better representation of commercial vehicle movements than models that account for commercial vehicles by inflating private vehicles or models that only include intra-provincial traffic. This population can be used in a large-scale combined private and commercial vehicle simulation model to assess policy and infrastructure decisions by evaluating travel times and traffic volumes on certain routes. To model at such a large-scale is often critisized because it requires vast amounts of data. In this section, it is illustrated how to model a commercial vehicle population for a whole region, using only GPS logs.

Commercial vehicle movements were modelled from observed movements. If, for instance, a warehouse is relocated or a new road is built, this commercial vehicle population, unfortunately, will not be able to predict future vehicle movements since the model does not incorporate logistical behaviour aspects. In the same vein, if, for instance, a carrier serves a number of customers and certain customers change the frequency with which they order products, a difference will not be visible in this model since this type of interaction is not captured in the model. It would require a lot of data to model a whole region's commercial vehicle movements with logistical behavioural aspects and therefore the focus in the rest of this dissertation is on introducing a behaviour-rich population for one company. This new population can then be used, in conjunction with the population that was generated in this section, in a large-scale transport simulation. To accurately capture logistical behavioural interactions in a commercial vehicle model, requires modelling of supply chain stakeholders as agents.

3.3 Incorporating stakeholder behaviour with the MATSim freight framework

Schröder et al. (2012) created a *freight framework* within the MATSim Toolkit, which allows for more realistic modelling of freight agent behaviour that captures logistics decisions and activities such as pick-ups and deliveries. Utilising this framework, two new freight agents are introduced: the Carrier agent and the Receiver agent.

3.3.1 The Receiver

The Receiver agent represents the Shipper's customers in a supply chain. The Receiver needs to ensure that it has adequate stock levels to best serve its customers. Consider the example from Chapter 1 again. The restaurants, the fish and chips shop, and the retailers are all considered to be Receivers that place orders to replenish their stock.

Each Receiver needs to plan the amount of stock to keep in order to not lose potential sales, but also to not incur unnecessary storage costs. Many organisations do Material Requirements Planning (MRP) in which they forecast their sales and based on the forecasts plan backwards to determine by what time they should order what amount of what product so as to have it available for sale in time. Based on this MRP, they place orders with their suppliers, which typically do last-mile deliveries from a distribution centre.

What adds complexity to this distribution channel is the fact that the Receivers may order different products from different suppliers and on different days of the week, which have different lead times based on the distance to be covered. To properly understand the influence of the Receiver on the Shipper agent, and utimately the Carrier, an intensive study needs to be conducted on consumer behaviour of different types of Receivers. Retailers' customers behave differently to those of restaurants in that they tend to buy larger volumes in shorter times.

For the purpose of this dissertation, the Receiver is described as follows. Each Receiver has a physical location, which is described by a coordinate. A Receiver can place an *Order* from a supplier on a given day. An *Order* contains the following information:

- The date.
- The type of commodity purchased.
- The mass of the commodity purchased.
- The sales value of the order.
- The description of the customer.
- A time-window during which the Receiver takes deliveries.

Furthermore, a customer may opt to order products once a certain minimum level of inventory is reached or decide to order every certain number of days and replenish stock to a certain level. This is known as an *order policy*. All of these factors play a role in how the Carrier reacts to the orders.

3.3.2 The Carrier

The Carrier is responsible for the actual movements of goods from the Shipper to the Receiver, and this is normally specified in a contract. The contract contains the set of orders from the Receivers. The Carrier responds to these orders by using its logistical capabilities, which can be divided into two categories: warehousing and transportation.

Firstly, the Carrier normally operates from one or more distribution centres, which gives it the ability to consolidate products before final distribution. This capability allows the Carrier to capatilise on economies of scale and reduce the price of products. The location of the distribution centre also plays a role in how responsive the Carrier can be to orders.

Secondly, the Carrier owns a fleet of vehicles with which it does the deliveries of orders. Multiple orders can be loaded onto a vehicle and considered to be delivered in the same *tour*. Considering the destinations of all the orders loaded on the vehicle, a *tour* is constructed using route optimisation and scheduling. Routes are determined using a routing algorithm, such as the Dijkstra or A*Landmark algorithms, which determine the shortest path in a road network from an origin to a destination. Given the demand of customers and the Carrier's vehicle fleet, tours can be constructed, in MATSim, to deliver all products to the customers, using the *Jsprit* toolkit (Schröder, 2014). *Jsprit* is an open-source Java toolkit that can solve multiple versions of the Travelling Salesman and Vehicle Routing Problems and is also incorporated into the MATSim freight framework.

The ultimate result is a Carrier *Plan*. This plan describes detials of the Carrier and its actions:

- The start time i.e. when the vehicle will depart from the depot.
- The pickup time-window.
- The estimated travel time based on the route.

- The delivery time-window.
- All orders, modelled in MATSim as *services*.
- The locations of all pickup, delivery, start and end activities.
- And finally, the vehicle fleet available.

For each vehicle in the fleet, the following information is required:

- The type of vehicle (e.g. refrigerated 3 tonne truck).
- The time-window during which the vehicles can be operated. This is based on the organisation that owns the vehicles.
- The capacity of the vehicle.
- The cost per distance unit.
- The maximum legal freespeed of the vehicle.

These Carrier plans can then be used in the initial demand step of the MATSim process, by adding these vehicles to the other commercial vehicle population that was developed in the previous section. These populations could then be used in a large-scale combined private- and commercial- vehicle transport simulation to evaluate the effect of traffic congestion or other policies on the vehicle movements. However, other populations are not necessarily required, since this behaviour-rich population can be used on its own, for instance, to evaluate what effect changes in the order policy of the Receiver has on the vehicle utilisation of the Carrier.

This subsection described the information that is required to model Carrier and Receiver agents and how they interact with each other. These agents need to be modelled based on a representation of a real Carrier and its Receivers. The following section shows how important supply chain stakeholders can be identified in a real freight network.

3.3.3 Identifying agents in a complex freight network

To model logistics behaviour at this level, one needs to understand stakeholders' behaviour to capture this behaviour in a commercial vehicle transport model. For the purposes of this dissertation, only the outbound logistics were considered and emphasis is placed on the Carrier and Receiver stakeholders in the supply chain. A description of how important stakeholders can be identified in a freight network, follows.

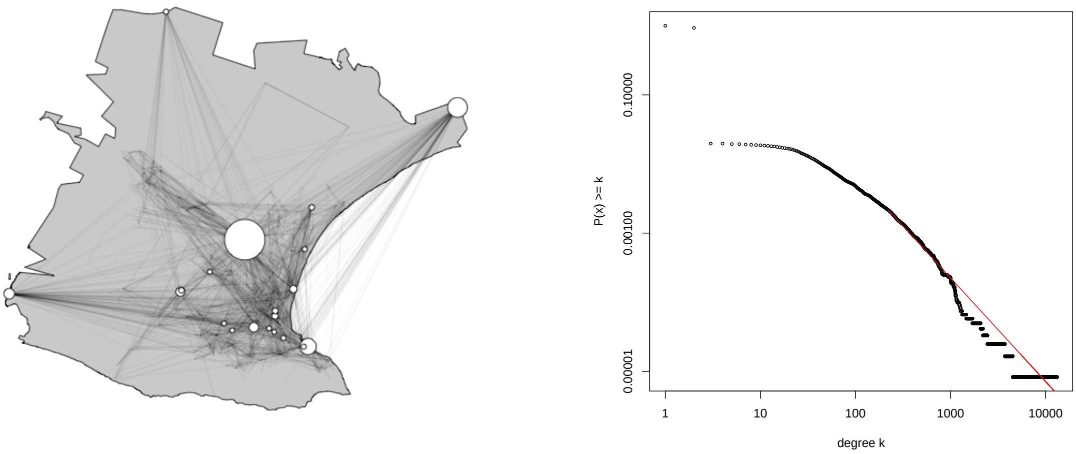
The complex network that was generated in section 3.2.2 for the NMBM is used here again. The network consists of a set of 120 913 vertices and 245 212 dyads. The set of vertices constitutes all locations at which an activity took place, i.e. a vehicle stopped. In this dissertation, the vertices are also called *facilities*, since vehicles tend to stop at facilities. It is important to note, though, that these are artificial facilities since not all facilities are necessarily actual facilities such as shopping centres or distribution centres. A facility could also represent a petrol station or merely an informal stop next to the road.

Using complex network theory metrics, one can describe various attributes of the network. For the purposes of this dissertation, only the most important vertices/facilities need to be identified from the network. The chosen metric was degree centrality: the number of links incident on a node. The degree centrality of each vertex is measured as the total number of links entering and exiting the vertex. The cumulative degree distribution of the network was fitted and tends to exhibit the behaviour of a power law function from a minimum degree of

219 towards the tail of the distribution, as can be seen on the log-log plot in Figure 3.6b. The Kolmogorov-Smirnov statistic was also calulcated as 0.04624 (which is below 0.05), which classifies this distribution as being power law according to the Kolmogorov-Smirnov test. The lower bounds, up to about a degree of 50, which is skewed off, may suggest that there are numerous vertices with very few vehicle stops and which are very insignificant in locating the most important vertices in the graph.

Therefore, only a few nodes in the graph have very high degree centrality and the rest of the nodes are not very well connected at all. This could possibly be attributed to the fact that a few major companies dominate the freight market in the NMBM area. Another possibility is that a large number of freight movements originate in the NMBM, but that customers are situated outside the NMBM area.

The 20 vertices with the highest degree centralities were identified. Figure 3.6a depicts the complex network with lines indicating a connection between two vertices and the circles depict the top 20 vertices, sized according to their relative degree centralities. Only the top 20 vertices are shown on the graph. For the sake of good visibility, link weights are not taken into consideration and all links are depicted with the same thickness.



(a) The top 20 vertices in the complex network.

(b) Cumulative degree distribution exhibiting the power law.

Figure 3.6: Major roleplayers in a complex freight network

From these 20 vertices, 3 were gateways into and out of the NMBM and only 3 organisations could be identified with map matching. The 3 organisations were the 1st, 8th, and 10th most important vertices in the network. The 3 organisations were contacted in an effort to establish a working relationship and to gather the necessary data required to model logistics behaviour in a transport model by means of a case study. The selected organisation for the case study is The Cold Chain (TCC), a company owned by the Imperial Group. In Chapter 4 a description is given of the supply chain structure of TCC, the data obtained from TCC and the methodology used to prepare the data for modelling of logistical behaviour. In Chapter 5 the data are used to model the Carrier and Receiver interactions as described in this Chapter. The model is tested for sensitivity to changes in the behaviour of the Receiver as well as changes in levels of traffic congestion.

Chapter 4

A case study in the FMCG industry

The organisation used as a case study in this dissertation, is a distributorship, located in an industrial area near the town of Motherwell in the Nelson Mandela Bay Metropole (NMBM), in the Eastern Cape Province of South Africa, namely the The Cold Chain (TCC). The data used in this case study are used for descriptive statistical analyses. Furthermore, the format and type of data were not sufficient to use as-is in modelling of stakeholder behaviour, and had to be augmented with additional fields of data. A background of TCC, the data augmentation methodology as well as the analyses follow.

4.1 Background of The Cold Chain

TCC started out as ICS Distributors, a subsidiary fully owned by ICS Foods. The company was renamed as TCC in 1984. Foodcorp, that also owned Enterprise Foods, bought a 50% shareholding of TCC in 1991. TCC was at that point in time responsible for sales and distribution activities of Enterprise Foods. TCC managed to retain major customers such as County Fair Foods and Tiger Brands Perishables from its customer base. In 2000, TCC was completely taken over by Imperial Logistics, an Imperial Holdings company (Imperial Logistics, 2013).

TCC has various capabilities that allow differentiated service offerings. It has a distribution centre where it can consolidate and redistribute chilled and frozen perishable products through secondary distribution channels to the Fast Moving Consumer Goods (FMCG), wholesale, and retail industries. It also has its own in-house vehicle fleet, which can be utilised to do the deliveries. Furthermore, it has the capacity and skills to deliver sales, merchandising and debt management from retailers on behalf of manufacturing brands.

TCC has a division that offers customers end-to-end supply chain solutions where food manufacturers can benefit from shared services and lowered costs because supply chain costs are channeled through a sole service provider. It also offers reverse logistics from customers back to TCC.

In relation to the agents that were described in Chapters 2 and 3, TCC can be described as being both a Transport Service Provider (TSP) and a Carrier agent. Its Mercor Solutions division fulfils the role of a TSP since it is responsible for the warehousing and distribution functions as well as Material Requirements Planning (MRP). It has to ensure that orders are placed on time to ensure delivery on the nominated day and that products are packed on the shelf with the maximum forward share possible. TCC then utilises an in-house vehicle fleet to do the actual deliveries, which relate to the Carrier agent.

TCC serves some top-end customers in the South African retailer environment including: Dairybelle, Enterprise Foods, Heinz Foods, Natures Choice, the Pick 'n Pay Group, Shoprite Checkers, Tiger Brands, and U-Save (Imperial Logistics, 2013).

4.2 Identifying key customers and products

The data file obtained from TCC was in the form of a flatfile, which includes the distribution of goods to all customers for a ten month period during 2012–2013.

Initial screening of the data revealed that 299 unique customers and 116 unique products existed in the dataset. The pareto-analysis technique was used to subset the data to a reduced, usable number of customers and products as customer locations were not provided and had to be sourced manually. The pareto principle is also known as the 80/20 rule. Using the customers as an example, the pareto principle indicates that 80% of sales value would be generated by only 20% of the customers. Figures 4.1 and 4.2 depict the pareto plots for the customers and products with an indication of the number of customers that account for 80% and 90% of sales value, respectively.

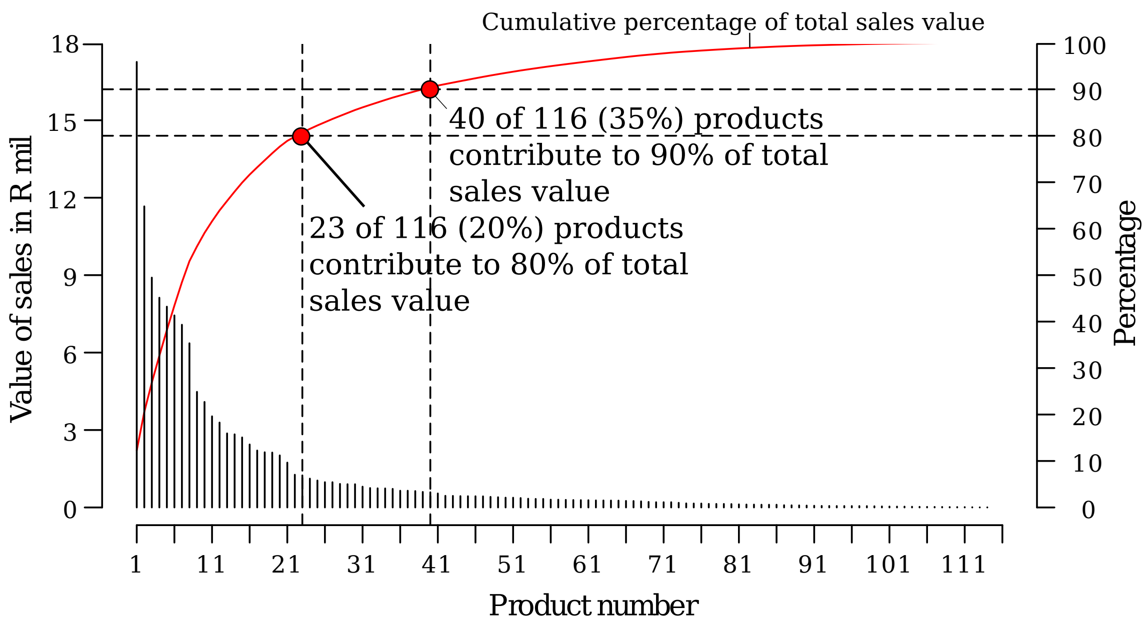


Figure 4.1: Pareto analysis of TCC's products

From these plots it was identified that 23 of the 116 products (20%) account for 80% of the sales value of products. For the purposes of this dissertation it was decided to only work with the products that contribute 80% of the sales value in order to reduce unnecessary complexity in the models. Similarly, it was identified that 36 of the 299 customers (12%) account for 80% of the sales value, and 53 of the 299 customers (18%) account for 90% of the sales value. For the purposes of this dissertation it was decided to work with the customers that account for 90% of sales value and not all 299, since the locations of customers were unknown.

To identify the customer locations, various methodologies were used. Some of the major retailers, such as the Pick 'n Pay, SPAR, and Checkers groups have the coordinates of their franchises available on their websites. For other customers, the customer name was Googled and if it was situated in a mall, the mall was identified and the location of the mall found on OpenStreetMap (OSM). All 53 customers' locations were identified and Figure 4.3 depicts the locations of both the distribution centre of TCC and its customers in NMBM and the greater Eastern Cape Province.

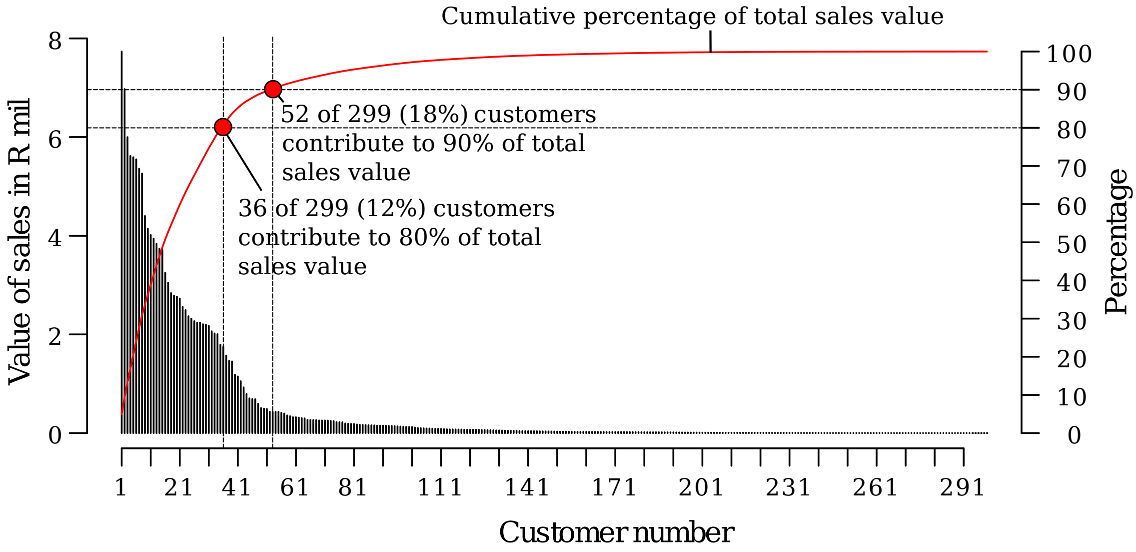


Figure 4.2: Pareto analysis of TCC's customers

4.3 Augmenting the data file

The original flatfile, containing the distribution data, was limited for analyses purposes, since very little insight could be drawn from it in its original format. The file contained 8 fields of data according to the following business rules:

Business Month: Indicating the month of the financial year.

Business Day: Indicating the business day of the month and not the calendar day. Operations only take place during weekdays, thus Monday in week 1 would be referred to as day 1, and Monday in week 2 as day 6. Also, calendar months often end in the middle of a week and accordingly that part of the week forms part of week one of the following business month.

Supplier: The supplier of the product of the shipment entry.

Customer: The customer to which the shipment was delivered.

Group: The retail group to which the customer belongs. For example, retailers U-save, Checkers, Shoprite, and OK Foods all form part of the Shoprite group.

Product: The product that was shipped. This is given as product groups, such as frozen chicken, frozen fish, and milk products.

Mass: The mass of the product (in kg) that was shipped.

Sale: The sales value (in Rand) of the product that was shipped.

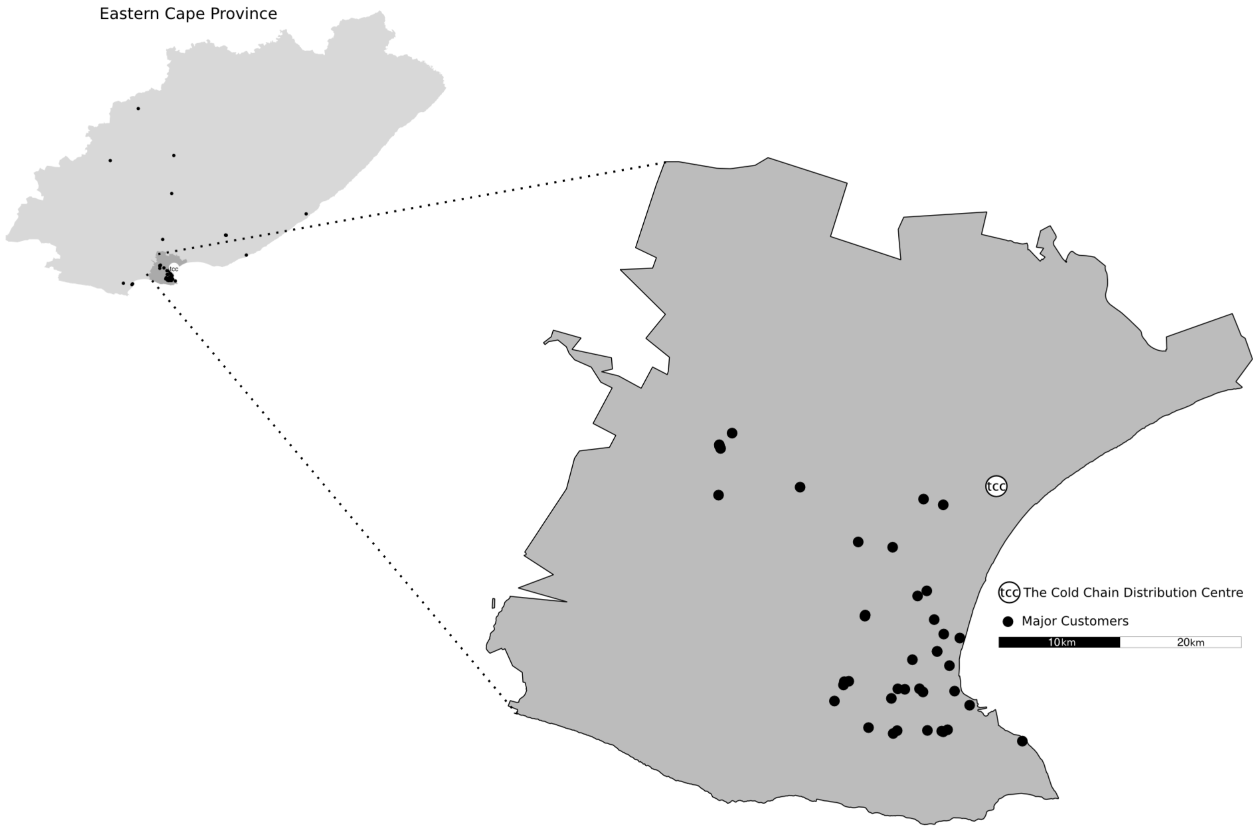


Figure 4.3: Location of the major customers of TCC

4.3.1 Calendar month and calendar day

Since the Business Month and Business Day fields could be confusing as to the actual month in which and day on which operations took place, the calendar month and day had to be calculated and the file be augmented with the information.

Business Month 1 and Business Day 1 in the file correspond with the first Monday of TCC's financial year, which was determined to be 2 July 2012. Business Month 1 and Business Day 2 would then be 3 July, etc. Accordingly, the dates for each business day and month were determined and two fields were added to the file:

Calendar Month: The Calendar Month to which the Business Month corresponds, where January through December would be mapped as Calendar Month 1 through Calendar Month 12.

Calendar Day: The Calendar Day of each of the Calendar Months, corresponding to the available days in each month.

4.3.2 Day-of-week value

To enable day-of-week analyses, the day of the week had to be determined. Utilising the Calendar Month and Calendar Day values as well as the year, a value for the day of the week was determined. An additional field was added to the file:

Day-of-week: An Integer value indicating the day of the week, where 1=Sunday, 2=Monday, . . . , 7=Saturday. In this dataset, however, weekends do not contain any data because deliveries only take place on weekdays, thus only values 2 through 6 are used.

4.3.3 Sequential day

To ease later calculations of typical logistical measures such as *the number of days since the last order*, a sequential day value was calculated. The first working day of the financial year was coded as sequential day 1, incrementing per Calendar Day until the final day of operations in the data file. The distribution data file was augmented with the field:

Sequential Day: An integer value indicating the day on which a shipment was made. Values are between 1 and 300 both inclusive, which correspond to 2 July 2012 and 23 April 2013, respectively.

4.3.4 Aggregating same-day shipments

In the data file, multiple entries of the same product, being delivered to the same customer, on the same day, exist. These would typically be added onto the same load in an actual delivery. Under this assumption such multiple entries were aggregated into one entry, with the mass and sales values being the sum of the entries' mass and sales values. The file was therefore only updated with the aggregated shipments replacing the original shipments in the file, but new fields were not introduced.

4.3.5 Days since last order

Next, for each product type being delivered to each customer, record was kept of when the last order for that product was delivered to that customer. For each entry, a new field was introduced and the file augmented with:

Days Since Last Order: An Integer value indicating how many sequential days have elapsed since the last order of the product for the specific customer. This could shed light on order policies of customers.

4.3.6 Mass and sales value of last order

From the previous step (where the number of days since last order was calculated), record was also kept of the mass and sales value of the last order. Two additional fields were added for each entry:

Last Mass: Mass (in kg) of the last order made by this customer for this product.

Last Sale: Sales (in Rand) of the last order made by this customer for this product.

4.3.7 Distance from distribution centre

Finally, the distance of each customer from the distribution centre was calculated and the file augmented with the field:

Distance from DC: A real number indicating the straight line distance (in meters) between the customer and the distribution centre.

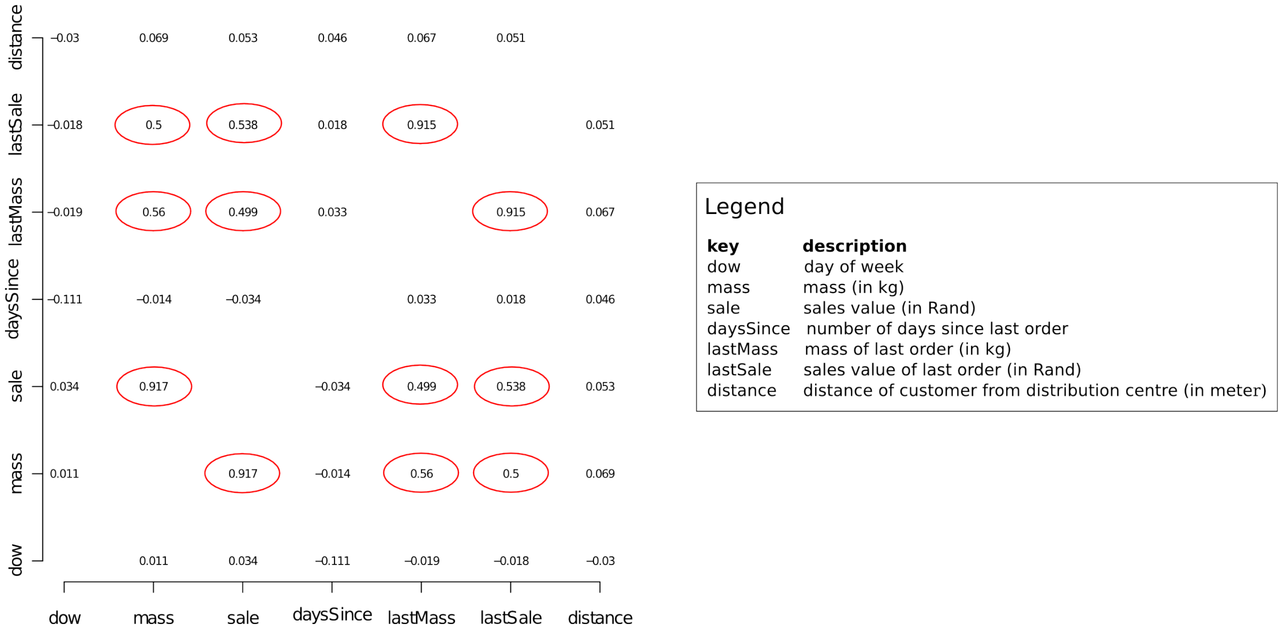


Figure 4.4: Correlation coefficients per pair of variables

4.4 Supply chain insights from the data

To identify trends or correlations between variables in the distribution data, would potentially be beneficial for a demand generation model. Correlation coefficients were determined for each pair of variables in the dataset and the coefficients are depicted in Figure 4.4.

The most prominent correlations between variables are those of the pairs of mass-sale with a correlation of 91.7% and lastMass-lastSale with a correlation of 91.5%. One would expect a high correlation between these values since it indicates that the prices of the products remained reasonably constant throughout the period. A correlation of less than 100% is an indication of some fluctuation, though. The only other prominent pairs are those of mass-lastMass with a correlation of 56% and sale-lastSale with a correlation of 53.8%.

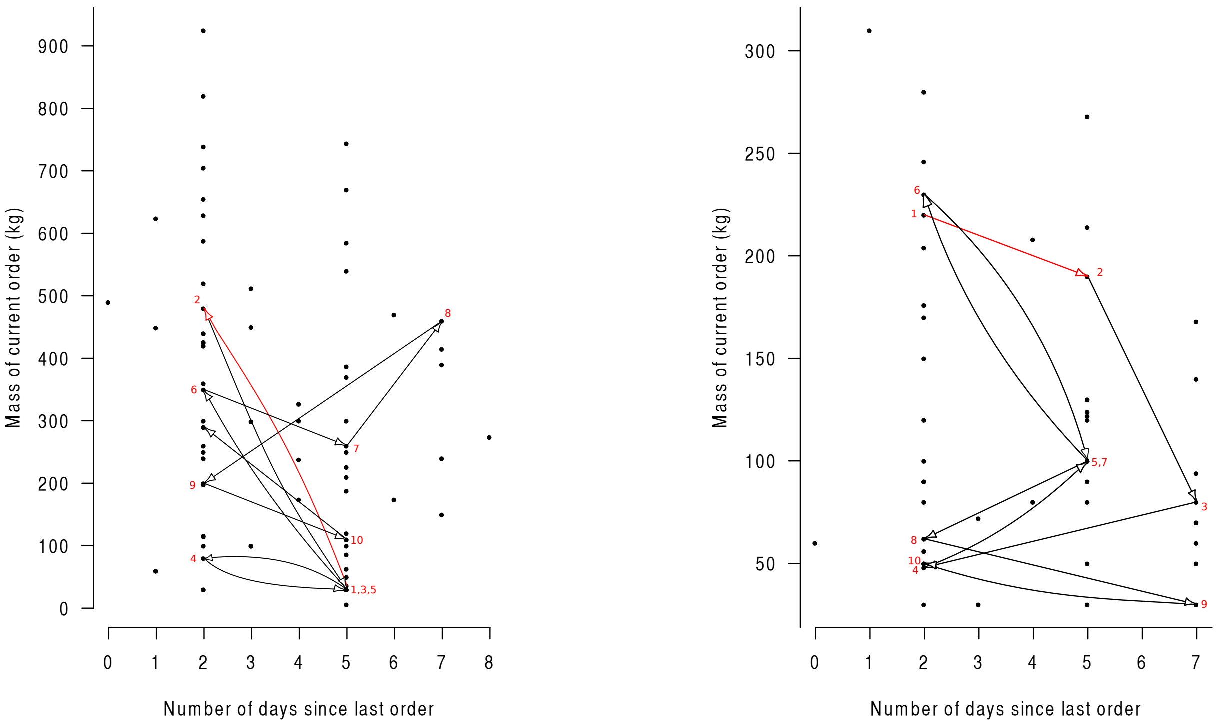
It was decided to evaluate 3 pairs of variables for further insight into the distribution of goods at TCC. For each analysis, only 2 customers and 2 products are selected to show in the dissertation, since there are 1 219 (53 customers x 23 products) possible combinations.

4.5 Order frequency

From a supplier perspective, it is useful to know how frequently orders are placed by each customer as well as the size of each order. If the frequency of orders from a customer is consistent, new orders can be forecasted with higher accuracy. This could assist to determine optimal stock and replenishment levels for these products.

The customers of TCC are grouped into *customer groups*. For instance, Shoprite and Checkers belong to the Shoprite group. Two customers (customers 2 and 4) were selected from the same customer group and two products were evaluated for each customer. The order frequency of products 9 and 23 of customer 2 were determined and are depicted in Figures 4.5a and 4.5b.

From these figures there are trends as to how often orders are placed. Arrows were added for a sequence of 10 orders in the plots to show how many days pass before the following order is placed. The sequence is indicated using red numbers and the red arrow depicts the starting point in the sequence. Product 9 appears to be ordered every 2nd and 5th day, whereas product



(a) Product 9 for Customer 2

(b) Product 23 for Customer 2

Figure 4.5: Order size and frequency of two products for Customer 2

23 appears to be ordered every 2nd, 5th, and 7th day. The other dots on the plots may indicate rush orders or orders on atypical days caused by a public holiday on the normal delivery date.

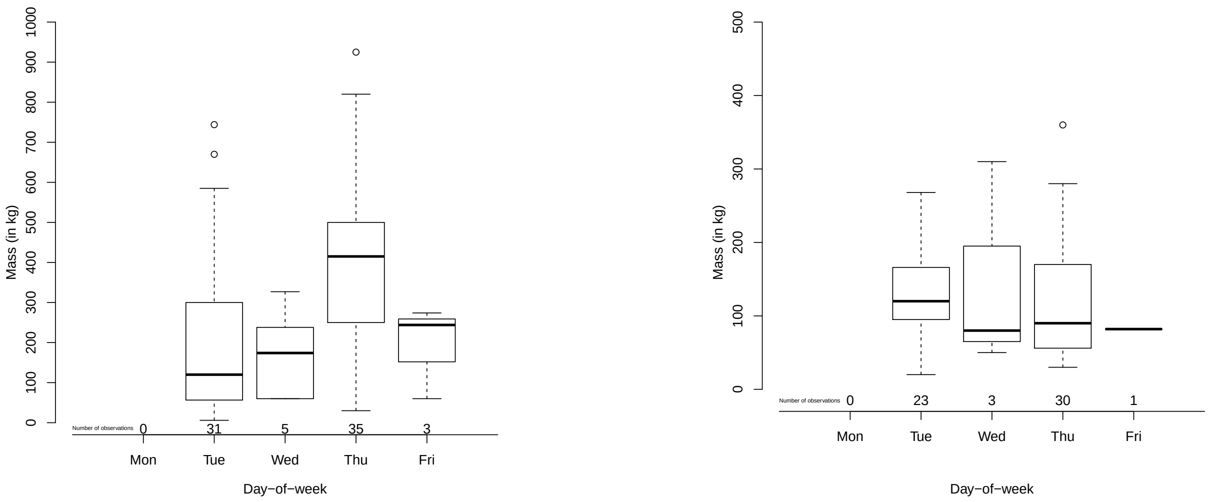
Subsequently, the day-of-week of the orders for products 9 and 23 of customer 2 were determined and are depicted in Figures 4.6a and 4.6b. The boxplot captures the variance in the size of the orders for each day of the week and underneath each boxplot is the number of times an order was placed on that given day. From the plots, it can be seen that both products are ordered primarily on a Tuesday and Thursday. Based on the frequency plots, Product 9 is therefore primarily ordered on a Tuesday, then 2 days later on a Thursday, and then 5 days later on the next Tuesday, etc. Product 23, follows a similar pattern, however, every now and again it is only ordered on Thursdays, i.e. every 7 days, possibly since stock on hand is adequate.

Given that this customer orders both products (and other products) on Tuesdays and Thursdays, it can be deduced that this customer typically follows the (R,Q)-Targets inventory policy, where the Re-order point, R, and the Reorder Quantity, Q, are time-based instead of quantity based.

The same analyses are shown for Products 2 and 3 of Customer 4. The order frequency of the Products were determined and are depicted in Figures 4.7a and 4.7b. These figures also show trends as to how often orders are placed. Both products appear to be ordered every 3rd and 4th day, with some extra orders every 2nd day at times and also some atypical orders.

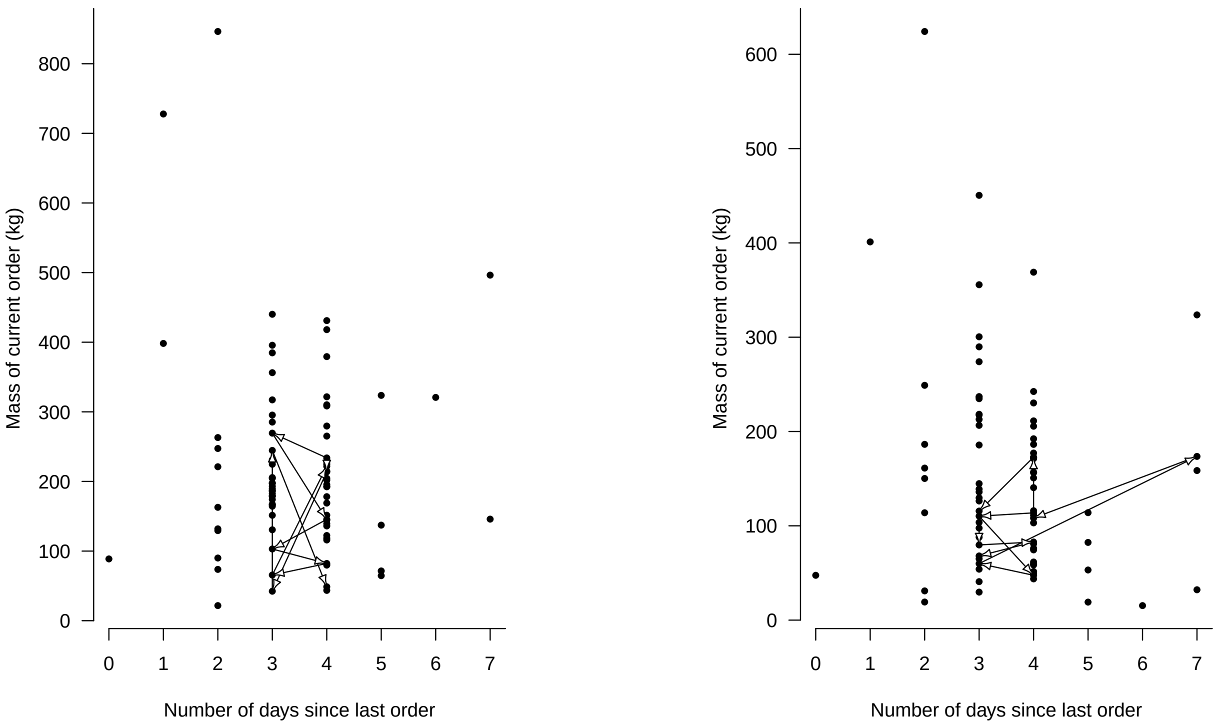
Subsequently, the day-of-week value of the orders for these Products was also determined and all these are depicted in Figures 4.8a and 4.8b. From the plots, it can be seen that both Products are ordered primarily on a Tuesday and Friday. Based on the frequency plots, both Products are ordered on a Tuesday, then 3 days later on a Friday, then 4 days later on the next Tuesday, etc. Sometimes Products are ordered on Thursdays, which correlates with the frequency of 2 days.

Given that this Customer orders both Products (and other products) on Tuesdays and



(a) Product 9 for Customer 2 (b) Product 23 for Customer 2

Figure 4.6: Order size per weekday of two products for Customer 2

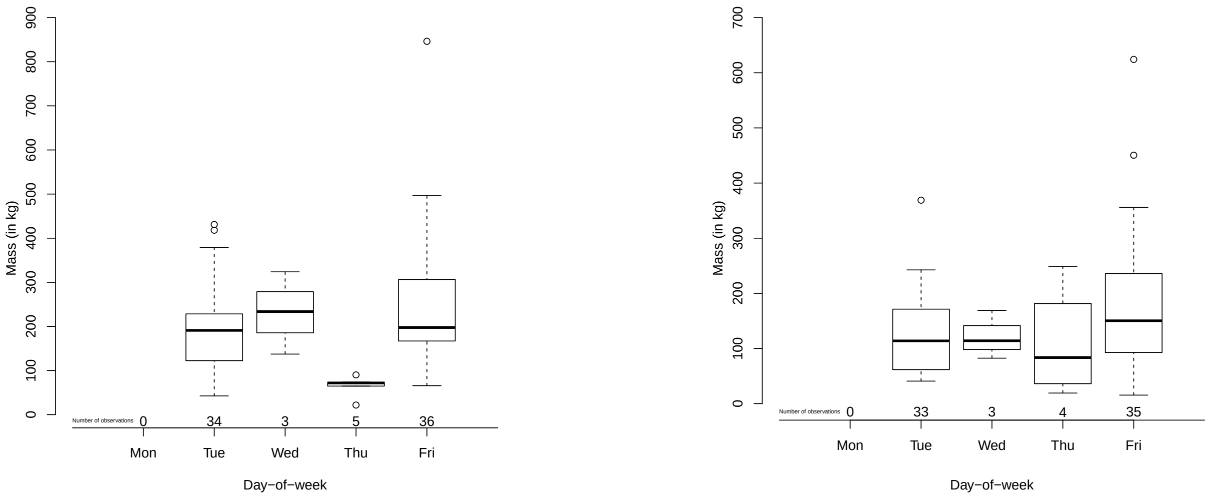


(a) Product 2 for Customer 4 (b) Product 3 for Customer 4

Figure 4.7: Order size and frequency of two products for Customer 4

Fridays, it can be deduced that this customer typically also follows the (R,Q)-Targets inventory policy.

What is interesting to note, though, is that both customers belong to the same customer group, yet their order policies differ. This reiterates that stakeholders' behaviour should be understood and modelled independently to capture a true representation of each stakeholder in the supply chain. A simple blanket procedure will not be accurate for all customers as order



(a) Product 2 for Customer 4

(b) Product 3 for Customer 4

Figure 4.8: Order size per weekday of two products for Customer 4

frequencies, days of the week on which orders are placed, as well as order sizes differ. Furthermore, some retailer groups may impose designated order days on their franchises, whereas others may allow them to choose order days themselves.

4.6 Order mass as a function of the previous order's mass

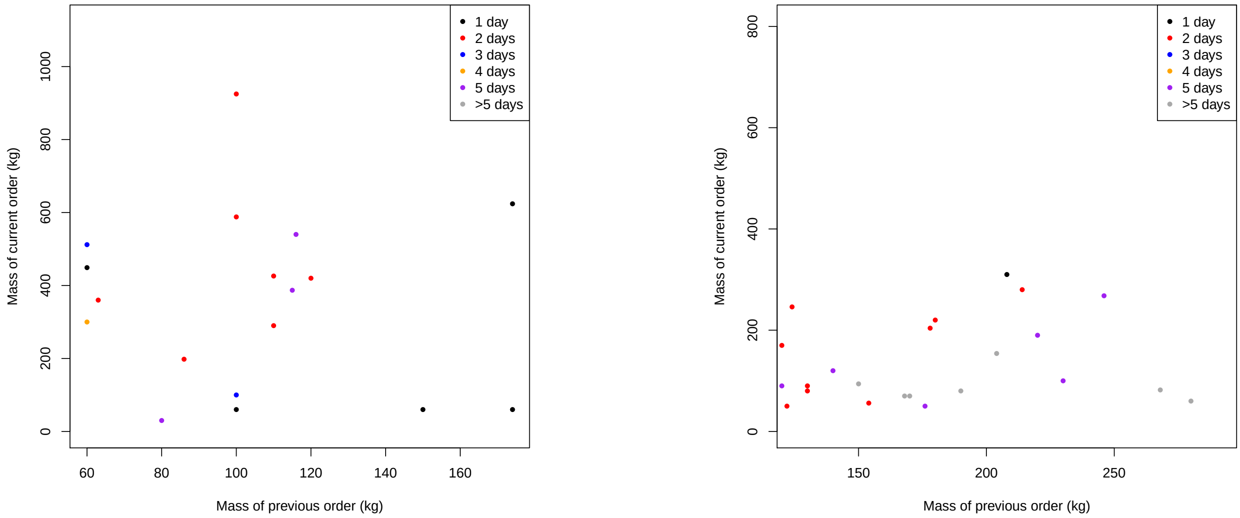
Even though the variable pair of mass-lastMass had a correlation of only 56%, it was decided to determine whether any trends could be identified. Figures 4.9a and 4.9b depict the relationship between these variables and data entries are colour coded according to the number of days that has passed since the previous order, giving an indication whether the order mass may be dependent on the combination of the mass of the previous order and the number of days that has passed since the previous order. Insights could not be drawn from the plots and trends could not be identified, which confirms the low correlation coefficient.

4.7 Influence of distance on order mass

The distance of all customers from the distribution centre was plotted in relation to the size of the orders that customers placed. This was to determine whether distant customers tend to order in larger quantities or less frequent due to larger transport costs, or not. Figure 4.10 reflects the order sizes against different distances from the distribution centre. The points are coloured according to the number of days since the previous order for the specific product and customer combination, but presented at aggregated level.

Emanating from the figure, it appears that all customers tend to vary the size of their orders, while disregarding distance for both order size and frequency. If one selects a distance on the distance axis, there is much variation visible in the mass values in the mass axis' direction for that given distance from the distribution centre.

For the purposes of this dissertation, the emphasis is on the interactions between the CARRIER and RECEIVER stakeholders in the supply chain. In this Chapter, the order policy of the customers were analysed and there were differences in the day-of-week on which customers order as well as the number of days that elapse since the last order for different customers



(a) Product 9 of Customer 2 (b) Product 23 of Customer 2

Figure 4.9: Correlation between mass and last mass variables of two products for Customer 4

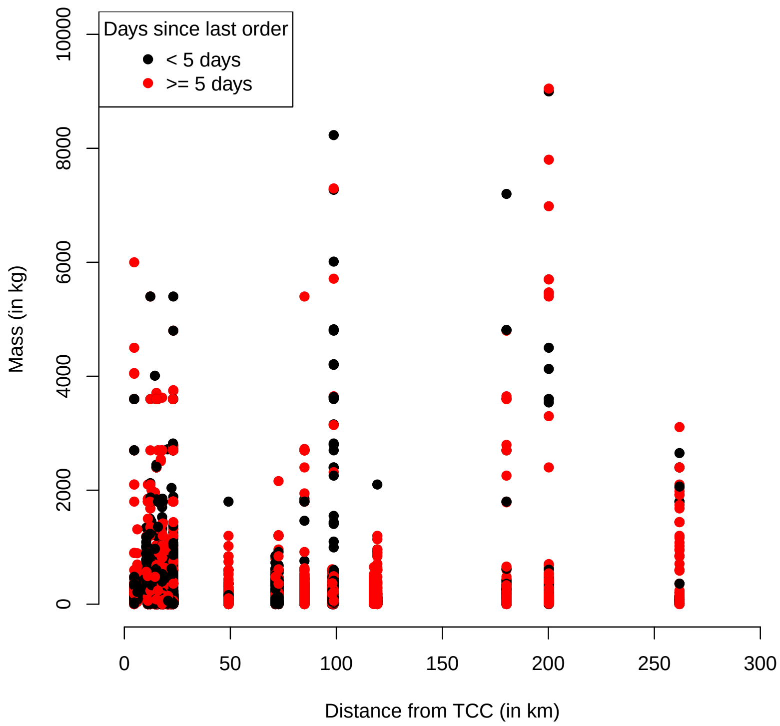


Figure 4.10: Distance of customers vs the mass of orders

and different products. Thus, different customers have different logistical behaviour. In the next Chapter, the CARRIER agent and the interaction between the CARRIER and Receivers are modelled. Furthermore, the influence of the RECEIVER's order policy and the influence of traffic congestion on the CARRIER are evaluated.

Chapter 5

Modelling and sensitivity analysis

In this Chapter, logistical behaviour is incorporated into a commercial vehicle transport model using the Multi-Agent Transport Simulation (MATSim) freight framework and the data of The Cold Chain (TCC). In this instance, the SHIPPER and CARRIER agents are the same organisation, i.e. TCC. RECEIVERS place orders with TCC and TCC responds to the demand by assigning vehicles to transport products to the various retailers.

In this Chapter, the preparation of the road network for the Nelson Mandela Bay Metropole (NMBM) area is discussed. Furthermore, the characteristics and costs associated with the delivery of products by the CARRIER to the RECEIVERS are illustrated. Three distinct scenarios are used to test and analyse the logistics reaction of the CARRIER to changes in the behaviour of RECEIVERS and transport conditions. Firstly, two implementations of the freight contribution in MATSim are developed to show a detailed description of the CARRIER's plans when orders are placed by the RECEIVERS. This also serves as the base model against which other models are evaluated. Secondly, the effect of traffic congestion on the CARRIER's choice of fleet composition and routing of vehicles is evaluated. Finally, the order policy of selected customers will be changed to evaluate the effect of a customer's order policy on the fleet size and composition of the CARRIER. Each scenario is described in detail, followed by modelling and analyses.

5.1 Implementation of logistics behaviour in MATSim

The first implementation of logistics behaviour in MATSim was done by considering two days in the order data of TCC: day 33 and day 38, which relates to Thursday, 2 August 2012, and Tuesday 7 August 2012, respectively. All orders for these two days were extracted from the data to be used further. The rest of this section describes how the road network for the NMBM was extracted, followed by how the demand for these two days were generated, followed by a description of the CARRIER's available fleet of vehicles and finally the CARRIER's response to the demand by generating CARRIER Plans.

5.1.1 Preparation of the road network

The road network, which represents an actual road network, was obtained from OpenStreetMap (OSM) (OpenStreetMap Contributors, 2013), an open source platform similar to Google Maps, but sources information from its users. Users of OSM can sign up to edit or update the road network using local knowledge. All such changes that are uploaded throughout the day are compiled into what is known as *snapshots* of an area.

Some websites or servers may host daily snapshots and others weekly or monthly. Similarly, there are also different snapshots that cover different sizes of areas, such as at metropolitan level and all the way up to provinces, countries or even a global snapshot. In this dissertation, the Geofabrik GmbH (2013) server is used, from where an OSM network for South Africa was obtained.

As was illustrated in Section 4.2, some of the customers are situated outside the study area. In many studies, vehicle routes would merely be modelled up to the border of the study area, thus focussing on the movement of vehicles inside the area. This, though, requires some workaround to cater for the extra travel time which would occur outside the area. In this dissertation the road network is extended to allow vehicles to travel to customers outside the area. In this way more realistic travel times can be obtained. Since only the NMBM area and greater Eastern Cape is considered in the modelling in this Chapter, the network was clipped to a smaller area in order to reduce the size and memory intensiveness.

Utilising *Osmosis* (Osmosis, 2014), a command-line Java application which allows one to process OSM networks, the final road network was generated in four steps. Firstly, the South Africa network was clipped to the region of the Eastern Cape Province, using the *bounding box* function in *Osmosis* and all road types were kept. Similarly, the road network was clipped to the region of the NMBM. Finally, only the major roads were extracted from the Eastern Cape network. The major roads in the Eastern Cape Province as well as all the roads in the NMBM are depicted in Figure 5.1. These two networks were merged in the fourth step to obtain the network to be used in the rest of the modelling.

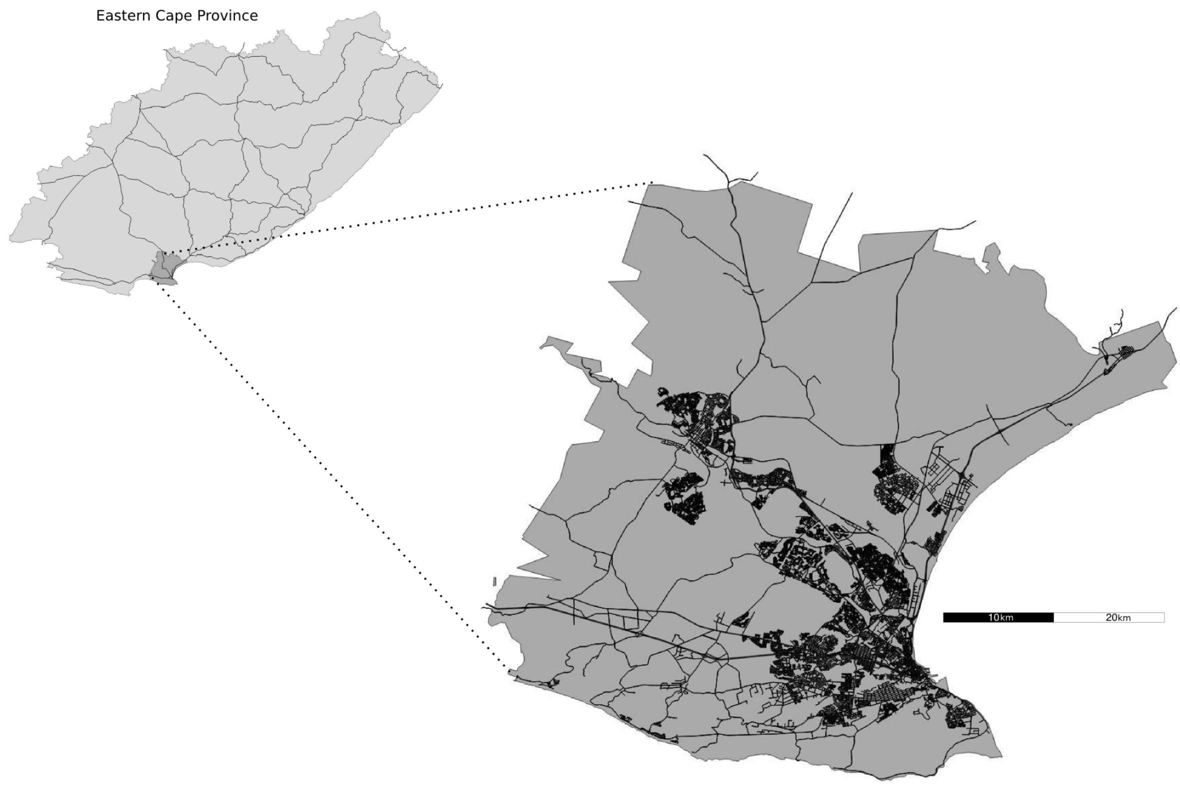


Figure 5.1: Combining the major roads in the Eastern Cape province with all roads in NMBM

Since a detailed road network can be very memory intensive when used in transport models, the road network was subsequently 'cleaned'. Wherever two road links on both sides of a node share the same attributes (such as number of lanes) it was simplified by merging the two links into one link with the same attributes. Finally, the OSM road network was converted into a MATSim road network. The network links' freespeed and capacity were also set to the corresponding freespeeds of different road types in South Africa. If the value of *number of lanes* was not available for a specific link, a default value per the OSM definitions were used, summarised in Table 5.1.

Table 5.1: Freespeeds and lane capacity of different road types in South Africa

Road Type	Number of lanes	Freespeed (km/h)	Capacity (vehicles/hour per lane)
Trunk	1	120	2000
Motorway	2	120	2000
Motorway link	1	80	1500
Primary	1	80	1500
Primary link	1	60	1500
Secondary	1	80	1000
Tertiary	1	60	1000
Unclassified	1	60	800
Residential	1	45	600

5.1.2 Demand generation

Customers place orders for products with TCC. Actual orders for day 33 and day 38 in the data file were used to illustrate how customers' behaviour differ in terms of order sizes and which products are ordered.

For both of these chosen days, for each customer and product combination, the mass of the orders were noted. For each of these combinations, an order was generated in MATSim by using the *Service* object. The customer's location, product type ordered, mass of the order, duration of the drop-off activity, and a time window during which products can be delivered, were specified. For all customers, a standard time-window to take deliveries of 08:00-17:00 was chosen and modelled accordingly. Also, a duration of 300 seconds (5 minutes) was used for each drop-off activity, since these durations were unknown.

The demand for all products on a day was subsequently aggregated per customer for visualisation purposes. This would give an indication as to where in the NMBM and greater Eastern Cape products needed to be delivered as well as to which customers the most products needed to be delivered. The demand for days 33 and 38 are shown in Figure 5.2 and Table 5.2.

From Table 5.2 it is evident that certain customers only order on certain days of the week. Those customers who ordered on both days, ordered different amounts of products. On day 33, the mass of all orders totalled 25.9 tonne, whereas on day 38, the mass of all orders totalled only 8.98 tonne. Furthermore, some customers ordered up to 7 tonne of products and others more than 3 tonne. This should have an effect on the size of vehicle allocated to deliver products to customers as well as the vehicle fleet composition for deliveries on the particular day.

5.2 Carrier capabilities

A Carrier delivers products to customers using an available fleet of vehicles. The fleet may be heterogeneous in size, capacity, and operating costs. In the data of TCC, ten vehicles were identified belonging to five different classes, also known as *vehicle types*. The Carrier agent was accordingly modelled as having a vehicle fleet of ten vehicles. The five *vehicle types* used are referred to as a 3 tonner, 6 tonner, 7 tonner, 12 tonner, and 15 tonner.

To determine different operating costs of the different vehicle types, the *vehicle cost schedule* of 2011 of the Road Freight Association (RFA) was consulted (The Road Freight Association, 2012). The vehicle cost schedule had limited information on refridgerated truck operating costs, which is required since products carried are frozen. From those available refridgerated truck costs in the vehicle cost schedule, it was calculated that the operating costs are roughly 1.15 times higher than the operating cost of equivalent sized vans in the cost schedule. Thus, a van

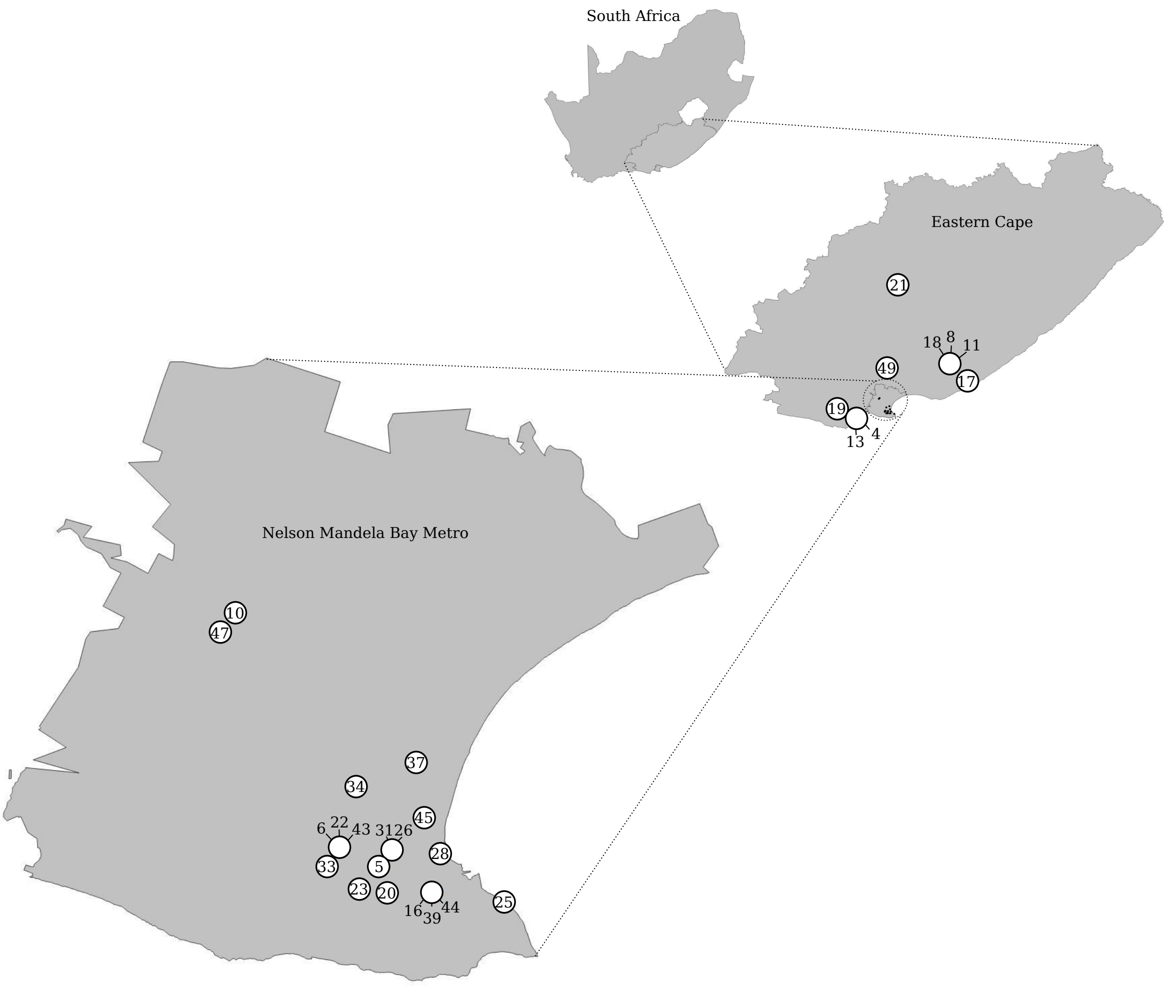


Figure 5.2: Customer locations in NMBM and the Eastern Cape

with a capacity close to the vehicle type was chosen and an operating cost value accordingly derived for an equivalent refridgerated truck. The classification of trucks in the vehicle cost schedule is known as concepts and trucks of different capacities have different concept numbers. Figure 5.3 depicts the five chosen concepts from the vehicle cost schedule [1].

Table 5.3 provides a summary of the costs for each of the ten vehicles as well as from which concept in the Vehicle Cost Schedule the values were derived. The first column contains an ID for each truck. The second column indicates the truck type, which also indicates the capacity of the truck since the name relates to the capacity. The third column indicates from which concept number the values were derived for the truck. The fourth column indicates the maximum freespeed of the vehicle. According to the National Road Traffic Act of 1989 for South Africa, vehicles of which the Gross Vehicle Mass exceeds 9 tonne, are limited to 80km/h (Arrive Alive, 2014). The columns thereafter in the table indicate the costs in South African currency.

Actual vehicles were modelled by specifying a unique ID for each vehicle, the vehicle type, the capacity, and the cost per distance unit, and a time window during which it can operate. Each truck was given a unique ID, for instance *truck_6_1*; this means it is a 6 Tonner truck and the first instance of a 6 Tonner truck to be generated in the model. A subsequent 6 Tonner truck was given an ID of *truck_6_2*. Next, each truck was given a capacity according to the

[1]These values are not industry standards. The vehicle cost schedule is a copyrighted document. Permission to use certain information from the schedule in this dissertation was obtained from the CEO of the Road Freight Association. The letters are in Appendix A

Table 5.2: Orders for customers on day 33 and day 38

	Total order size (kg)	
Customer	Day 33	Day 38
4	1 886.29	0
5	857.66	565.56
6	932.7	0
8	1 028.34	0
10	0	284.11
11	1 338.32	0
13	6 915.29	0
16	1 172.87	0
17	0	511.58
18	865.89	0
19	1 409.82	0
20	420.23	987.99
21	8.8	0
22	0	183.4
23	546.53	389.95
25	1 677.09	1 276.47
26	1 572.68	439.93
28	57.2	0
31	2 202.84	629.18
33	669.81	554.75
34	1 045.2	529.02
37	1.38	0
39	898.3	696.22
43	359.59	706.08
44	0	61.2
45	0	1 105.26
47	0	58.92
49	49	0
Total	**25 915.83**	**8 979.62**

vehicle type it relates to. For instance, a 6 Tonner truck was given a capacity of 6 000 units, where each unit is equivalent to one kilogram. The time window during which vehicles can operate was set to be between 06:00 and 21:00, both inclusive.

Next, the operating cost was set for each truck. Two costs were considered based on the vehicle cost schedule: a variable per kilometre cost, and a fixed per kilometre cost, from Table 5.3. Two different files were created: one containing only the variable cost per kilometre, and one containing the total cost per kilometre, where the fixed cost is included. The latter being a derived estimation, but useful to evaluate the effect of multiple cost sources on fleet composition and tour planing. The vehicles are used in the next step.

5.2.1 Tour planning and scheduling

Given the demand and the Carrier's vehicle fleet, tours were constructed to deliver all products to the customers, using the *Jsprit* toolkit (Schröder, 2014). *Jsprit* is an open-source

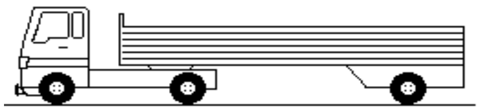
Concept number	Concept	Total Cost - cpk	
03	 4x2 Rigid (5001 to 7500 kg)	Fixed Cost - cpk	693.7
		Variable Cost - cpk	367.6
		Total Cost - cpk	1 061.3
04	 4x2 Rigid (7501 to 10000 kg)	Fixed Cost - cpk	816.8
		Variable Cost - cpk	452.8
		Total Cost - cpk	1 269.6
05	 4x2 Rigid	Fixed Cost - cpk	858.1
		Variable Cost - cpk	569.6
		Total Cost - cpk	1 427.7
07	 Three Axle Artic (4x2 TT+Single Axle ST)	Fixed Cost - cpk	1075.6
		Variable Cost - cpk	646.6
		Total Cost - cpk	1 722.2
08	 Four Axle Artic (4x2 TT+Tandem Axle ST)	Fixed Cost - cpk	1494.7
		Variable Cost - cpk	724.0
		Total Cost - cpk	2218.7

Figure 5.3: Vehicle concepts used from the RFA

Table 5.3: Summary of vehicle costs

Truck Id	Truck Type	Concept	Max Speed (km/h)	Van Cost Var (c/km)	Van Cost Fix (c/km)	Van Cost Total (c/km)	Fridge Truck Cost Var (c/km)	Fridge Truck Cost Fix (c/km)	Fridge Truck Cost Total (c/km)
truck_3_1	3-tonne	3	100	693.70	367.60	1061.30	797.71	422.79	1220.50
truck_6_1	6-tonne	4	100	816.80	452.80	1269.60	939.28	520.72	1460.00
truck_6_2	6-tonne	4	100	816.80	452.80	1269.60	939.28	520.72	1460.00
truck_6_3	6-tonne	4	100	816.80	452.80	1269.60	939.28	520.72	1460.00
truck_6_4	6-tonne	4	100	816.80	452.80	1269.60	939.28	520.72	1460.00
truck_7_1	7-tonne	5	100	858.10	569.60	1427.70	986.84	655.02	1641.86
truck_7_2	7-tonne	5	100	858.10	569.60	1427.70	986.84	655.02	1641.86
truck_12_1	12-tonne	7	80	1075.64	646.60	1722.28	1236.97	743.64	1980.61
truck_15_1	15-tonne	8	80	1494.67	724.00	2218.67	1718.86	832.59	2551.45
truck_15_2	15-tonne	8	80	1494.67	724.00	2218.67	1718.86	832.59	2551.45

Java toolkit that can solve multiple versions of the Travelling Salesman and Vehicle Routing Problems and is also incorporated in the MATSim freight framework.

The Carrier chose the number of vehicles to utilise and constructed tours for each vehicle. Subsequently, the vehicles were routed on the road network provided. Eventually, it sets the tours to the Carrier as the selected plans to execute, known as the Carrier *Plan*.

Using this functionality, a Carrier Plan was created for day 33 and day 38. The plans contain detailed descriptions of the tours that each vehicle would execute. Summaries of the plans are captured in Tables 5.4 and 5.5.

Table 5.4: Analysis of Initial Plans for day 33

VehicleId	**Distance** (km)	**Travel Time** (hours)	**Activity Time** (hours)	**Tonne-km moved** (tonne-km)	**Cost** (ZAR)
truck_3_1	500.01	6.25	0.42	18.61	6100.12
truck_6_3	81.81	1.14	5.58	243.77	1194.36
truck_6_4	68.21	0.88	5.58	153.77	995.9
truck_7_1	247.95	3.39	6.83	643.49	4071.41
truck_12_1	151.68	2.86	3.92	366.16	3004.72
Total	**2 665.78**	**38.12**	**21.51**	**4 513.94**	**36 906.09**

Table 5.5: Analysis of Initial Plans for day 38

VehicleId	**Distance** (km)	**Travel Time** (hours)	**Activity Time** (hours)	**Tonne-km moved** (tonne-km)	**Cost** (ZAR)
truck_3_1	55.15	0.76	4.92	59.94	672.85
truck_6_1	256.77	3.94	3.83	254.05	3748.89
truck_6_4	102.72	1.48	8.17	223.59	1499.67
Total	**1 280.10**	**20.11**	**16.08**	**792.66**	**18 124.43**

From these Tables, there is a difference between the number of commercial vehicles utilised on each day. On day 33, 5 of the 10 available commercial vehicles are utilised and on day 38 only 3 of the 10 available commercial vehicles in the fleet are utilised. The total distance of each commercial vehicle as well as the total expected travel time, activity time, tonne-km moved and cost of each commercial vehicle are also shown. Day 33 has many more orders to deliver, thus the higher number of vehicles used. From the demand figure in section 5.1.2, day 33 has a few customers that are located outside the NMBM, thus the longer distances travelled. This information is useful to evaluate the effect of other factors on the distance travelled, the expected travel time, and tonne-km moved.

No difference in vehicle utilisation could be found by using only variable costs as opposed to using variable and fixed costs per kilometre. This can be attributed to the fact that the difference of costs between different vehicle sizes don't change much when adding fixed cost per kilometre. Henceforth, both variable and fixed costs per kilometre are combined into a total cost per kilometre as per the costs in Table 5.3.

The Carrier plans that were created in this section were based on a free flowing network with no traffic congestion. When traffic congestion is introduced, the Carrier may need to change its routes and scheduling of vehicles to still be able to deliver all the products within the given time-windows.

5.3 The effect of traffic congestion

In this section, the effect of traffic congestion on the vehicle fleet composition and expected travel times are evaluated. In MATSim there are two methods with which to evaluate the effect of traffic congestion. Firstly, one can add the modelled commercial vehicles into a simulation model containing private and other commercial vehicles and observe the effect of all vehicles using the same road network. This method depends on actual vehicle movement in the model and the interaction between the various road users. Secondly, one can emulate traffic congestion by reducing the freespeed value on the road network between a given time window and for any of the links in the network.

The network was analysed firstly to obtain an overview of the fraction of links in the network with different levels of freespeed. The NMBM and Eastern Cape areas were considered separately. Recall that the Eastern Cape part of the network contains only the major roads, with no residential roads included.

Table 5.6: Analysis of the original road network for NMBM

Road type	**Max freespeed** (km/h)	**Fraction of links** (NMBM)	**Fraction of links** (Eastern Cape)
Trunk/Motorway	120	2%	44%
Primary/Secondary	80	17%	52%
Tertiary/Unclassified	60	13%	4%
Residential	45	68%	0%

Table 5.6 reflects that 68% of the network inside the NMBM area is residential with low freespeeds. Even low occurrences of traffic congestion should have an effect on the routes and vehicle composition of the CARRIER. The Eastern Cape area has 44% Trunk and Motorway roads and 52% belonging to Primary roads.

Emulating congestion involves considering each link in the network, noting its freespeed value (obtained from OSM or set as a default value, described in section 5.1.1), and reducing the value by a certain factor, using the *network change events* functionality in MATSim. For the NMBM area of the network only (the dense urban part), five levels of congestion scenarios were modelled that represent typical morning and afternoon peak periods. Between 07:00-10:00 and 16:00-19:00, the maximum freespeed of each link in the network was used as the base case (100%). Thereafter, the maximum freespeed of each link was reduced to 80% of the maximum freespeed, for the second scenario. This process was repeated for 60%, 40%, and 20%, respectively.

Using the same demand from the previous section for day 33 and day 38, new CARRIER Plans were generated for each of the scenarios with different levels of congestion and a discussion follows.

5.3.1 Day 33

For each of the scenarios, the time-of-day, during which each vehicle travelled and delivered orders, was analysed. Figure 5.4 shows a summary view of the scheduling and utilisation of vehicles in each scenario.

In all five scenarios, no 15-tonne trucks were utilised. As congestion increased, the 12-tonne truck is used more, almost to full capacity, except for the 20% scenario. In the 20% scenario, orders were shifted from the 12-tonne truck to the 3-tonne truck, which was utilised to full capacity, as can be seen in row J under the 20% column. This is expected since the 3-tonne truck can travel faster than the 12-tonne truck and is cheaper to operate.

For each of the scenarios, the distance, total travel time, tonne-km moved, and cost were calculated. Figure 5.5a shows the total distance of all vehicles in each scenario. A clear pattern is not visible, yet under extreme traffic congestion conditions (the 20% scenario), the total distance did increase by more than 200km from the base case scenario (100%). A cause could not be found for the reduction in the total distance in the 60% scenario, except that the algorithm employed to solve the vehicle routing problem may have been run for too few iterations in all scenarios. This could have caused sub-optimal solutions in some of the scenarios. Running the algorithm for more iterations, though, will increase the computational time and will not necessarily find a better solution.

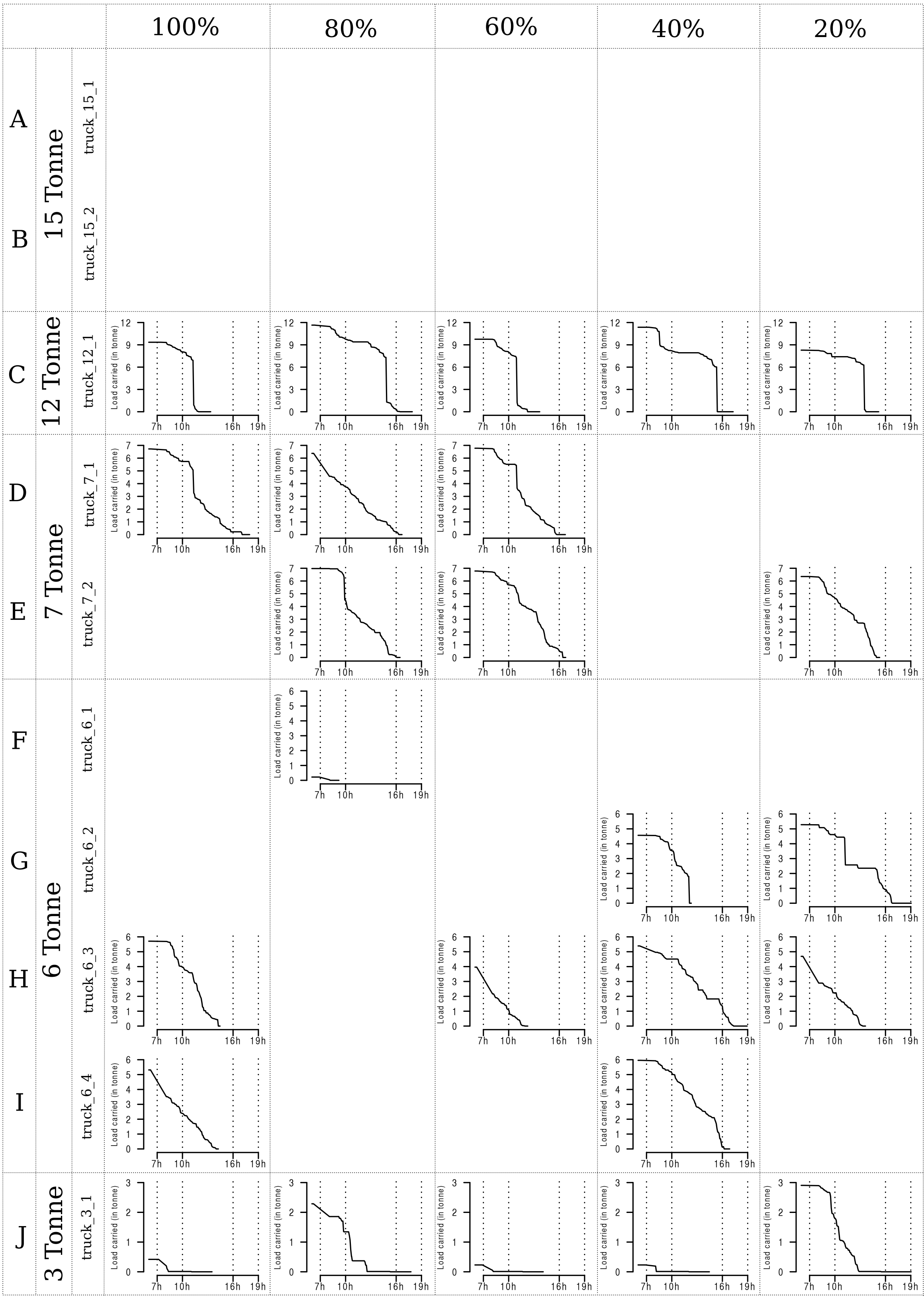
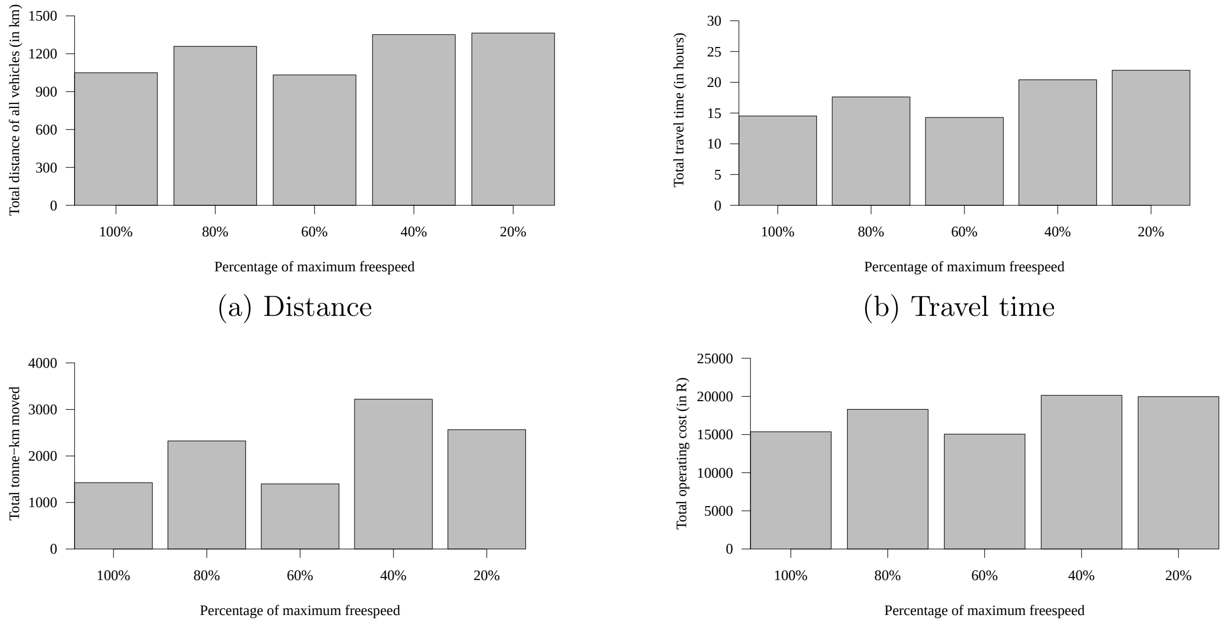


Figure 5.4: Results of different levels of traffic congestion on vehicle utilisation on day 33

From the figures, the base case scenario sees five vehicles being utilised: one 12 Tonner, one 7 Tonner, two 6 Tonners and a 3 Tonner. As congestion levels increase, the vehicle utilisation changes.

The cost in Figure 5.5d has a similar pattern to that of the distance. This was expected since the cost function in the algorithm only had operating cost of the vehicles to consider.



(a) Distance (b) Travel time

(c) Tonne-km (d) Cost

Figure 5.5: Effect of different levels of traffic congestion on day 33

Recall that a fixed and variable operating cost per kilometre was used for modelling purposes. This may cause that vehicle usage is only constrained by a per-distance cost and the total travel time could vary as long as the deliveries are made in the time-window. The effect of using a time-based cost is considered in following subsections of this dissertation.

From Figure 5.5c it can be noted that the tonne-km moved in the 40% and 20% scenarios surpassed that of the tonne-km moved on a free network. This can be useful for goverment to evaluate what effect traffic congestion has on the total tonne-km moved by each organisation. Also, the total distance travelled may not have increased that much, but the total travel time did increase significantly. Travel time predictions are also important to consider when government makes decisions for infrastructure investment.

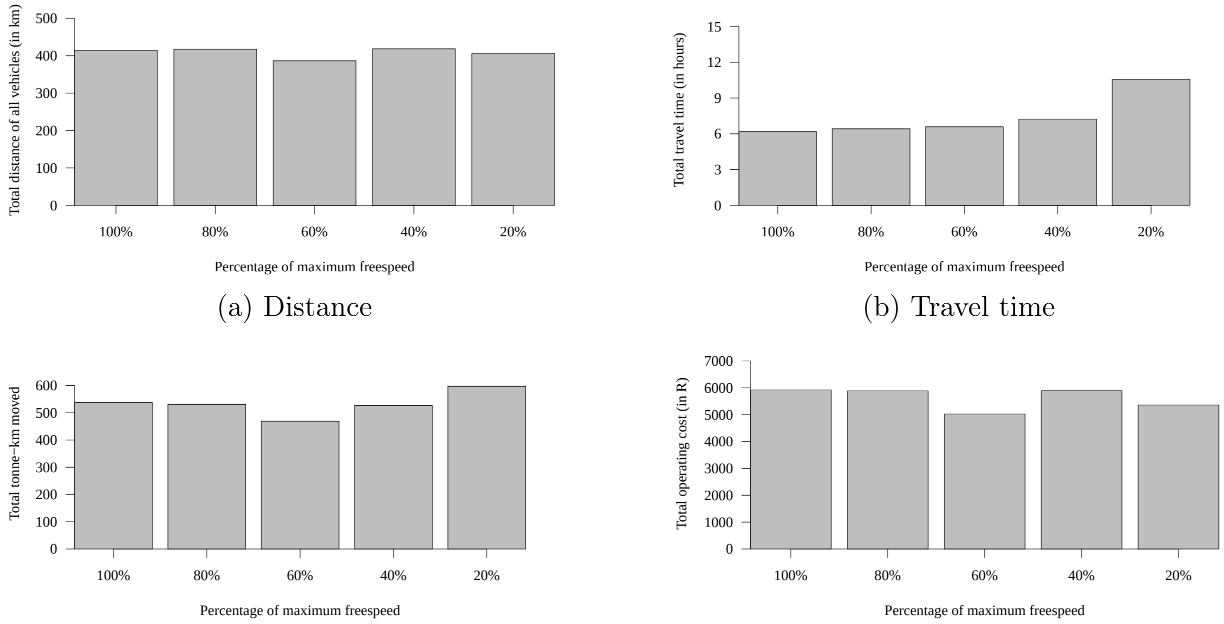
5.3.2 Day 38

All analyses were repeated for day 38. For all the scenarios, the CARRIER chose one 3-tonne and two 6-tonne trucks with which to do deliveries and only changed the sequences of customers visited and the routing. The total distance, as can be seen in Figure 5.6a did not change much in the scenarios. Since cost is distance-based and travel time is not penalised, cost did not increase, as is visible in Figure 5.6d. Travel time, as shown in Figure 5.6b, increased for each scenario, as was expected. In the most congested scenario the travel time was almost double that of the base-case scenario.

From the results obtained for day 33 and day 38, the model is sensitive to changes in the levels of traffic congestion since it changed the fleet utilisation for the different scenarios, which is useful to do fleet planning and scheduling and routing. The distance-based cost function may have caused some of the odd results, thus in the next subsection the effect of using a time-based cost function is evaluated.

5.3.3 Day 33 with cost of time

All analyses were rerun for day 33 with the difference being the cost function in the algorithm being given a time-based cost. The same per-kilometre cost, as was used previously, were



(a) Distance (b) Travel time

(c) Tonne-km (d) Cost

Figure 5.6: Effect of different levels of traffic congestion on day 38

used to allocate a per time-unit cost. Thus the 15-tonne trucks still cost more to operate per time-unit than the 12-tonne truck, etc. The five traffic congestion scenarios were rerun.

Figure 5.7 shows the summary of the vehicle utilisation and timing of all vehicles for each scenario. The main difference between the results of the distance-based cost and time-based cost is that the time-based cost scenario sees the larger trucks being utilised more often. One 15-tonne truck is used in the 60% and 40% scenarios as can be seen in row B. The 12-tonne truck, in row C, is used in each of the scenarios as was the case in the distance-based scenarios. From the 60% scenario onwards, the 3-tonne truck was introduced, which is due to the 3-tonne truck having a higher operating speed and can cover greater distances than the larger trucks in the same amount of time, which will be a cheaper option. The vehicles are also loaded almost to full capacity in all of the scenarios.

The total distance increased as traffic congestion increased, as is seen in Figure 5.8a. The travel time, in Figure 5.8b, again increased as traffic congestion increased, which was expected. The cost, in Figure 5.8c increased accordingly since it was linked to the travel time.

From these scenarios, the model is sensitive to changes in traffic congestion levels, but that a time-based cost delivered more realistic results. Ideally, the cost function should be supplied with a distance-based operating cost, a fixed cost per use of a vehicle, as well as a time-based cost. Furthermore, since drop-off durations were unknown and all drop-off activities were modelled with the same duration, results may differ when using more realistic drop-off durations.

The differences in the scenarios of day 33 and day 38 show that different combinations of customers that order different amounts of products on different days, coupled with increasing traffic congestion, changes the way in which the Carrier do its routing and scheduling. In the next section, the effect of the customer's order policy on the Carrier is investigated further.

5.4 The effect of the customer's order policy

To evaluate the effect and whether the model is sensitive to changes in the order policy, the order policy of a number of the customers of TCC was changed. This was done as follows.

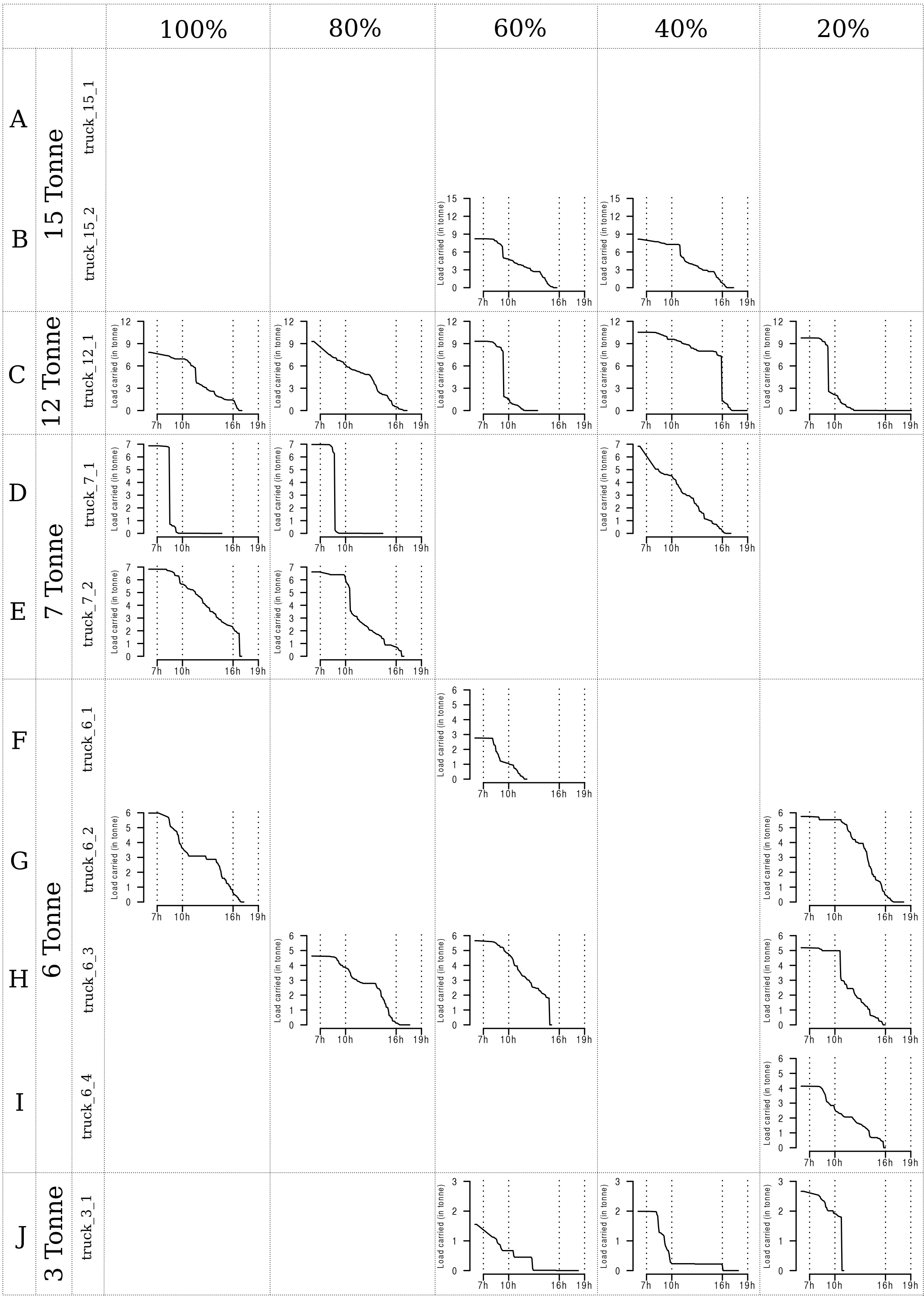
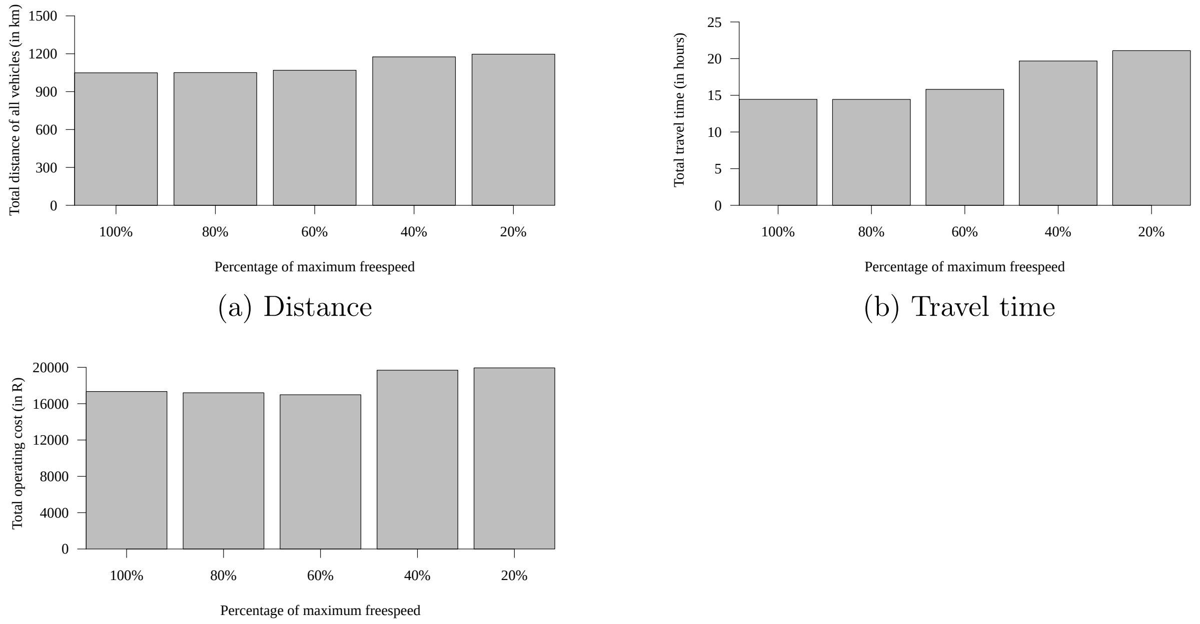


Figure 5.7: Results of different levels of traffic congestion on vehicle utilisation when using a time-based cost

Firstly, for all customers of TCC, the average frequency of orders was determined by considering the number of times a customer places an order for a specific product per week. This was averaged over the time period of the data used. A cumulative probability distribution of the average frequency of orders of customers is depicted in Figure 5.9.



(a) Distance

(b) Travel time

(c) Cost

Figure 5.8: Effect of different levels of traffic congestion using a time-based cost

5.4.1 Average frequency

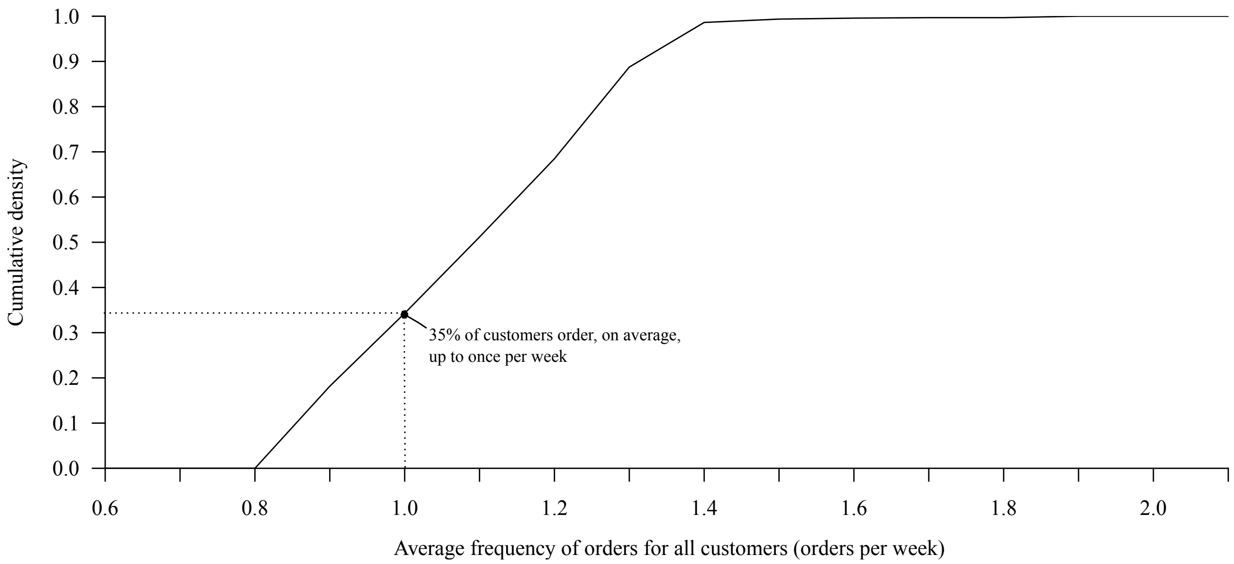


Figure 5.9: Average frequency

Similarly, the average weekly order size was determined for each customer and product combination. This average order size as well as the average frequency of orders were then used to generate demand.

The probability that a customer would place an order on a given day, was calculated as the average number of orders per week for that customer divided by the number of days available in the week (in this case only 5 because TCC only delivers on weekdays). This probability is described in (5.1).

$$\text{P(Order)} = \frac{\#\ \text{orders/week}}{5} \tag{5.1}$$

Next, a random number between 0 and 1 was generated and if this random number exceeded P(Order), which was calculated in (5.1), an order was generated for the specific product and for that customer, with an order size of the average weekly demand divided by the average order frequency per week, as is given in (5.2).

$$\text{Order Size} = \frac{\text{Average Weekly Demand}}{\text{Average \# of orders per week}} \tag{5.2}$$

This was repeated for all customer and product combinations to obtain new demand based on average order size and frequency. This demand was the base case to work from. To test different order policies, for all customers, the average frequency of orders per week were reduced by 0.4. The average demand was recalculated as before, using the new frequency of orders. Because these customers now order less frequently, the average demand increased. This was the second scenario. Three more scenarios were developed by increasing the average order frequency of the customers from average by 0.4, 0.8, and 1.2, respectively. Because all customers order on average between 0.8 and 2.2 times per week, these five scenarios would never cause an order frequency of less than 0 or more than 5. Table 5.7 contains a summary of the total demand for each scenario described. As was expected, the lower-than-average frequency resulted in higher demand and the higher the order frequency, the lower the demand.

Table 5.7: Total demand for each order policy scenario

	Scenarios				
	−0.4	**Base Case**	+0.4	+0.8	+1.2
Demand (in tonne)	60.53	39.88	29.81	23.82	19.84

The Carrier created tours again based on these five new levels of order frequency and order size and the results are shown in Figure 5.10.

The base case scenario had a demand of 39.88 tonne and six vehicles were utilised to deliver the orders to customers, which consisted of one 15-, one 12-, two 7-. and two 6-tonne trucks. For the scenario where the frequency was lowered by 0.4, the demand increased to 60.53 tonne and all vehicles were utilised to do the deliveries. For the scenarios where the order frequencies increased and the demand decreased, fewer and smaller vehicles were utilised. In most of the cases, the trucks finished their tours well within the time-window specified and was mostly only influenced by the morning peak hours.

Since the demand increased when order frequency decreased and the demand decreased as order frequency increased, it was expected that the tonne-km moved would decrease. Figure 5.11c shows this trend. The same holds for the operating cost since fewer vehicles were used as the frequency decreased, as can be seen in Figure 5.11d.

From a costing perspective, total costs deviated by about R10 000 in the different scenarios, with a saving of about R10 000 when the average order frequency increased by 1.2 orders per week. Delivering a larger order once a week instead of two smaller deliveries twice a week, would increase the transport costs on the day of delivery, but the second delivery's cost will disappear since there would not be a second delivery during the week. The implication, however, is that if the Carrier's fleet composition is optimised to serve the customers as efficiently as possible and a few customers suddenly change their order policy, the vehicles will not necessarily be able to deliver all products on a given day. Also, the vehicle utilisation may reduce drastically on the other days of the week, which causes vehicles not to be used. These vehicles could then be rented out to other companies, but this in turn adds security risk and possibly extra costs.

This analysis is useful to the TCC in that they can evaluate what effect different order policies will have on their vehicle utilisation as well as distance travelled and costs incurred.

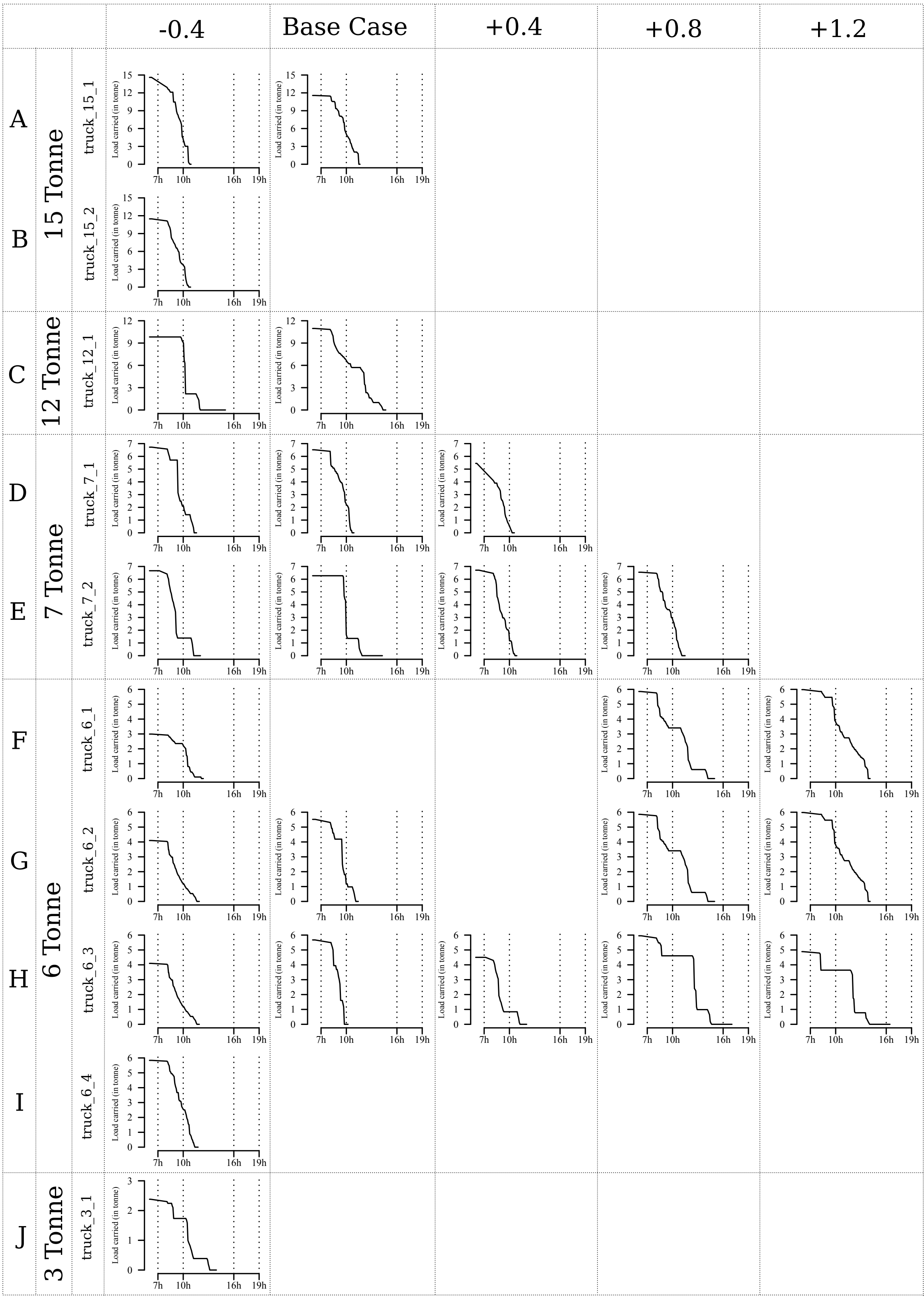
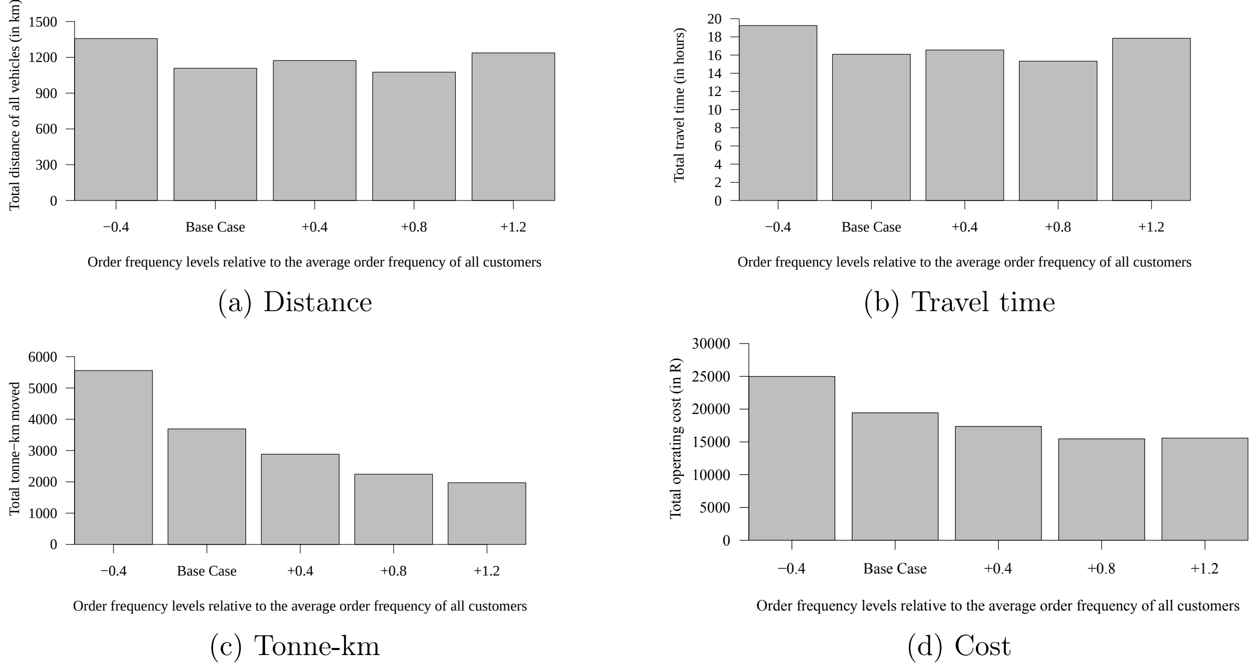


Figure 5.10: Impact of different order policies on vehicle utilisation

This could help them to develop an incentive for retailers if they change their order policy for TCC to operate more efficiently. TCC is vulnerable to the risk of higher costs as was noted in the scenario of day 38. These analyses are useful for TCC to evaluate the effect of such drastic changes to the frequency of orders on its operations. To be able to test different scenarios, they could plan how to best deal with such situations. For instance, it may be



(a) Distance (b) Travel time

(c) Tonne-km (d) Cost

Figure 5.11: Effect of different order policy frequencies

beneficial to utilise a 3rd Party Logistics (3PL) service provider to deliver some goods in some situations. Furthermore, TCC could introduce other vehicles in the model to determine the effect of changing the vehicle fleet composition.

In this Chapter, logistics behaviour was implemented in a MATSim freight model and from the experiments, one could note that the model is sensitive to changes in the traffic congestion levels as well as the order policy of a customer.

Chapter 6

Conclusion

Transport modelling is an important tool for both government and the private sector to evaluate the effects of a changing transport environment on the economy or on their businesses. In Chapter 1, the role of transport in businesses and private commuting was discussed. The state of logistics of South Africa and how transport influences the competitiveness of a country were included in this discussion.

Efficient transport relies on adequate transport infrastructure, which should be provided by government. Through transport planning, government makes decisions on infrastructure development and maintenance and require adequate transport models to base these decisions on. The same holds for commercial vehicle transport models. These should be representative of actual commercial vehicle movements and contain the necessary elements with which to model and analyse changes in logistical behaviour.

Although various commercial vehicle transport models exist in practice, it was identified that many of these models are based on private vehicle models or contain assumptions that render it unrepresentative of reality. Furthermore, it was identified that multiple stakeholders exist in a supply chain, all with their own objectives and decision-making capabilities. The more stakeholders involved, the more complex the interactions between the stakeholders in the supply chain. Commercial vehicle models should capture stakeholder behaviour to adequately evaluate the effect of changes in the logistics system on the behaviour of the stakeholders. This led to the research question on what the benefits are to capture and integrate supply chain decision-making into commercial vehicle transport models by modelling stakeholders as agents. A follow up question was posed as to the ability of the model to be used as a decision support tool in a South African context. This Chapter attempts to answer the research questions.

6.1 Intra- and inter-provincial commercial vehicle activity chains

The first objective of the research was to expand on the model of Joubert et al. (2010) by including both intra- and inter-provincial commercial vehicles in an agent-based commercial vehicle population. This model was developed by using a complex network to sample locations for activities and considering intra- and inter-provincial activity chains separately. The population was validated against observed data (GPS logs) in terms of Vehicle Kilometers Travelled (VKT). The results indicated that the VKT of the modelled population compared favourably to the observed population, but with some over- and under-estimation still evident. This could be due to the fact that only locations associated with a facility were used when sampling the next location from the complex network. This could possibly be improved by including all locations when building the complex network, but this would increase the computational complexity and memory required to manage such a large graph.

Such a model that incorporates both intra- and inter-provincial commercial vehicle activity chains is a better representation of actual commercial vehicle movements than what is found in many state-of-practice models. This commercial vehicles population could be added to a private vehicle model and simulated in a large-scale agent-based simulation. Such a simulation could then be used to evaluate the impact of certain policies and interventions, such as upgrading of a road or building of new transport infrastructure.

Commercial vehicle movements were modelled from observed movements. If, for instance, a warehouse is relocated or a new road is built, this commercial vehicle population, unfortunately, will not be able to predict future vehicle movements since the model does not incorporate logistical behaviour aspects. For this reason, commercial vehicle models should incorporate stakeholder behaviour to accurately capture such behavioural interactions. This requires modelling of supply chain stakeholders as agents.

6.2 Modelling logistics behaviour in MATSim

To be able to model at the level required in behaviour-based transport models, requires a thorough understanding of supply chain stakeholder behaviour. TCC was used as a case study to gain insights into the factors in the supply chains that influence stakeholders' decisions.

The order frequency of each customer and for each product was analysed. Order mass and the order mass as a function of the order mass of the previous order were also considered. The influence of distance on order mass was also among the analyses. From the analyses, it was determined that the order frequency for different customers and products differ. Retailers have set days on which they order and the order size differs for each order. For this reason, the impact of the order frequency on the Carrier was evaluated.

The average order frequency was changed for all customers. This had an effect on the average order size too and the Carrier reacted differently to each scenario. This analysis is useful to the TCC in that they can evaluate what effect different order policies will have on their vehicle utilisation as well as distance travelled and costs incurred.

Including logistical behaviour in commercial vehicle models is useful for government as well. Different levels of traffic congestion were emulated on the road network and the Carrier Plans were sensitive to these changes. The vehicle utilisation, tonne-km moved, distance covered, travel time, and costs changed as traffic congestion levels changed.

From government's perspective, the influence of road infrastructure decision-making can be evaluated by evaluating the increases or decreases in distance travelled and tonne-km moved by the commercial vehicles. The aim of government is to reduce the total tonne-km moved on roads as well as predict travel times and this model allows the modeller to evaluate the changes in these factors when there are changes in traffic congestion. Road maintenance can be modelled using the same congestion emulation technique as in Chapter 5. For instance, if road maintenance would take place on a specific corridor to upgrade it, the current situation can be modelled, the situation where certain links and capacity in the network is removed or reduced can be modelled, and the resulting upgraded road can be modelled. These three instances could then be evaluated to determine the effect of proposed maintenance on tonne-km moved and the effect on travel time.

By incorporating both intra- and inter-provincial commercial vehicles in a model, the model is representative of actual commercial vehicle movements. Furthermore, by incorporating logistical behaviour found in the supply chain into the commercial vehicle model by modelling the stakeholders as agents and capturing their behaviour, yields a model that is sensitive to changes in order policies of customers as well as different levels of traffic congestion. Such a model can be used as a decision support tool, since the effect of the changes on distance travelled, costs, tonne-km moved and travel time can be evaluated as described above.

There are, however, still some assumptions that need to be challenged as well as some shortcomings to be addressed in future research to enhance this modelling technique. The following section highlights some of the assumptions and shortcomings as well as how they can be addressed.

6.3 Future research

6.3.1 Consumer behaviour

In this dissertation, the demand of customers was determined from the distribution data supplied by TCC. Some supply chain insights could be gained from the data. Firstly, the order policies of customers differ for different products. Also, the distance of the customers from the distribution centre of TCC did not seem to have an effect on the order policies of the customers. Customers located farther away still had different order policies as opposed to ordering fewer times per week and in larger quantities to save on transport costs, as one would intuitively think.

It is imperative to understand retailer behaviour in detail so as to model the behaviour more accurately. Consumer behaviour drives the Material Requirements Planning (MRP) of retailers, since they need to keep an adequate supply of stock to provide in the consumers' needs. The retailers' MRP again influences the distributor's MRP. Thus, one avenue of future research could be to analyse consumer behaviour in a specific industry, determine how this influences the distributor's MRP, introduce a Receiver agent with its own decision-making capability, and then model the interaction with the Carrier agent.

Furthermore, this dissertation focussed on the Fast Moving Consumer Goods (FMCG) industry. Organisations in different industries would have different behaviour.

6.3.2 As-is vs To-be

In this dissertation it was shown that the transport model is sensitive to changes in the system, be it from a network or customer perspective. Various data sources were not available, such as the duration of delivery activities, the operating costs of vehicles, and the locations of all customers and had to be determined from other sources. Also, the plans generated do not necessarily relate accurately to that of the actual tours that TCC execute daily.

An enhanced model would include more accurate costs and duration values. Thus, an as-is case that emulates the exact tours of TCC could be modelled and compared to the suggested optimal routes and schedules proposed by the model.

6.3.3 Other stakeholders in the supply chain

In this dissertation, the interactions between various stakeholders in the supply chain are discussed. The modelling, however, focusses on Carrier-Retailer interaction where the Shipper, TSP, and Carrier agents belong to the same organisation. This is largely due to limitations in the functionality of the freight framework in MATSim, which is under continual development.

An avenue for further research could be to develop the framework with which to model a complete freight market with market dynamics and Shipper-Transport Service Provider (TSP) interaction.

6.3.4 Outsourcing of logistics functions

Outsourcing of logistics functions is a common phenomenon in organisations that wish to increase their customer base and establish long term larger contracts so as to justify owning an in-house fleet. It also reduces risk since risk is transfered to or shared by other organisations. In this dissertation, the focus was on an organisation with an in-house fleet.

Another option for modelling would be to model the outsourcing of the transport and warehousing functions to a 3PL service provider and evaluating the effect of such actions on the logistics costs of the organisation.

6.3.5 Total logistics cost

In this dissertation, the cost of commodities carried was not considered in analyses. Some of the other costs used in the models were estimations, since the data were not available.

In future research, trade-offs between inventory carrying costs and transportation costs can be evaluated by including various logistics costs in analyses, given that all cost factors and values are known.

6.3.6 Incorporating supply chain metrics

Supply chain metrics, such as SCOR metrics described by the Supply Chain Council, Inc. (2014) could be included to measure and evaluate the changes in Key Performance Indicator (KPI) performance of the organisation. For instance, the percentage of on-time deliveries, perfect order fulfilment, and order fulfilment cycle time. These could enhance the strategic decision-making capability of the models.

Bibliography

Anand, N., Quak, H., van Duin, R., and Tavaszy, L. (2012). City logistics modeling efforts: Trends and gaps - a review. In *Procedia - Social and Behavioral Sciences*, volume 39, pages 101–115.

Arrive Alive (2014). Speed Limits and the Law. Available online from: <`http://www.arrivealive.co.za/pages.aspx?i=25`> (Retrieved January 2014).

Chow, J. Y. J., Yang, C. H., and Regan, A. C. (2010). State-of-the art of freight forecast modeling: lessons learned and the road ahead. *Transportation*, 37(6):1011–1030.

De Jong, G. and Ben-Akiva, M. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B: Methodological*, 41(9):950–965.

Department of Transport (2005). National Freight Logistics Strategy. Available online from: <`http://www.portsregulator.org/images/documents/National_Freight_Logistics_Strategy.pdf`> (Retrieved August 2013).

Fourie, P. J. (2010). Agent-based transport simulation versus equilibrium assignment for private vehicle traffic in Gauteng. In *29th Annual Southern African Transport Conference*, pages 12–23.

Geofabrik GmbH (2013). Available online from: <`http://download.geofabrik.de/africa/south-africa-and-lesotho.html`> (Retrieved November 2013).

Hart, P., Nilsson, N., and Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics SSC4*, 4(2):100–107.

Havenga, J. and Simpson, Z. (2013). *The global state of logistics cost surveys. 9th State of Logistics Survey for South Africa 2012: Connecting neighbours, engaging the world.* Council for Scientific and Industrial Research. Viljoen, NM (Editor), Council for Scientific and Industrial Research, Pretoria, South Africa. Pages 2-19. ISBN: 978-0-7988-5609-6.

Hensher, D. and Figliozzi, M. A. (2007). Behavioural insights into the modelling of freight transportation and distribution systems. *Transportation Research Part B: Methodological*, 41(9):921–923.

Holguin-Veras, J. (2000). A framework for an integrative freight market simulation. In *3rdIEEE Intelligent Transportation Systems*, pages 476–481.

Holguin-Veras, J. and Patil, G. (2005). Observed trip chain behavior of commercial vehicles. *Transportation Research Record*, 1906:74–80.

Hunt, J. D. and Stefan, K. J. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological*, 41(9):981–1013.

Imperial Logistics (2013). Available online from: <http://www.imperiallogistics.co.za/company/305/cold-chain> (Retrieved October 2013).

Institute for City Logistics (2013). Institute for City Logistics. Available online from <http://www.citylogistics.org/> (Retrieved August 2013).

Joubert, J. W. and Axhausen, K. W. (2011). Inferring commercial vehicle activities in Gauteng, South Africa. *Journal of Transport Geography*, 19(1):115–124.

Joubert, J. W. and Axhausen, K. W. (2013). A complex network approach to understand commercial vehicle movement. *Transportation*, 40(3):1–22.

Joubert, J. W., Fourie, P. J., and Axhausen, K. W. (2010). Large-scale agent-based combined traffic simulation of private cars and commercial vehicles. *Transportation Research Record*, 2168:24–32.

Joubert, J. W. and Van Heerden, Q. (2013). Large Scale Multimodal Transport Modelling, Part 1: Demand Generation. In *32nd Annual Southern African Transport Conference*, pages 644–652. ISBN: 978-1-920017-62-0. Available from: <http://hdl.handle.net/10204/6963>.

Joubert, J. W., Van Heerden, Q., and Van Schoor, C. (2013). Complexities in moving from commodity to vehicular flows. In *32nd Annual Southern African Transport Conference*, pages 568–576. ISBN: 978-1-920017-62-0. Available from: <http://hdl.handle.net/10204/6962>.

Liedtke, G., Matteis, T., Wisetjindawat, W., and Schröder, S. (2014). Impacts of urban logistics measures on multiple decision layers - a case study. *Working paper*.

Liedtke, G. and Schepperle, H. (2004). Segmentation of the transportation market with regard to activity based freight transport modelling. *International Journal of Logistics: Research and applications*, 7(3):199–218.

Marker Jr, J. T. and Goulias, K. G. (1998). Response freight model under different degrees of geographic resolution — geographic information system application in pennsylvania. *Transportation Research Record*, 1625:118–123.

MATSim Development Team (2012). The Multi Agent Transportation Simulation Toolkit. <http://matsim.org> (Retrieved December 2013).

Meyer, M. and Miller, E. (2001). *Urban transportation planning - A decision-oriented approach*. McGraw-Hill.

OpenStreetMap Contributors (2013). Available online from: <www.openstreetmap.org> (Retrieved November 2013).

Ortuzar, J. D. and Willumsen, L. G. (2011). *Modelling Transport, Edition 4*. John Wiley and Sons.

Osmosis (2014). Available online from: <http://wiki.openstreetmap.org/wiki/Osmosis> (Retrieved August 2014).

Roorda, M. J., Cavalcante, R., McCabe, S., and Kwan, H. (2010). A conceptual framework for agent-based modelling of logistics services. *Transportation Research Part E: Logistics and Transportation Review*, 46(1):18 – 31.

Ruan, M., Lin, J., and Kawamura, K. (2012). Modeling urban commercial vehicle daily tour chaining. *Transportation Research Part E: Logistics and Transportation Review*, 48(6):1169–1184.

Samimi, A., Mohammadian, A., and Kawamura, K. (2009). Behavioral freight movement modeling. Resource Paper, Workshop 2: Behavioral paradigms for Modeling Freight Travel Decision-Making. In *12th International Conference on Travel Behaviour Research*, Jaipur, India.

Schröder, S. (2014). Jsprit-project. <https://github.com/jsprit/jsprit/>.

Schröder, S., Zilske, M., Liedtke, G., and Nagel, K. (2012). A computational framework for a multi-agent simulation of freight transport activities. In 91st *Annual Meeting of the Transportation ResearchBoard (TRB)*, Washington, D.C.

Simpson, Z. and Havenga, J. (2012). *Logistics costs - Lower relative costs, higher risk. 8th State of Logistics Survey for South Africa 2011: Gearing up for change.* Council for Scientific and Industrial Research. Viljoen, NM (Editor), Council for Scientific and Industrial Research, Pretoria, South Africa. Pages 27–41. ISBN: 978-0-7988-5601-0.

South African Government Information (2013). National Infrastructure Plan. Available online from <http://www.info.gov.za/issues/national-infrastructure-plan/#SIPs> (Retrieved 1 August 2013).

Supply Chain Council, Inc. (2014). The Supply Chain Operations Reference (SCOR) Model. <https://supply-chain.org/scor/11> (Retrieved January 2014).

Tan, A., Bowden, R., and Zhang, Y. (2004). Virtual simulation of statewide intermodal freight traffic. *Transportation Research Record*, 1873:53–63.

The Road Freight Association (2012). The Road Freight Association Vehicle Cost Schedule. <http://www.rfa.co.za/rfa/index.php?option=com_content&view=article&id=30&Itemid=193> (Retrieved December 2013).

The World Bank Group (2014a). Logistics Performance Index, 2012. <http://lpi.worldbank.org/> (Retrieved December 2013).

The World Bank Group (2014b). Connected to compete 2014 - Trade logistics in the modern economy. Available from: <http://www.worldbank.org/content/dam/Worldbank/document/Trade/LPI2014.pdf> (Retrieved April 2014).

Van Heerden, Q. (2011). *An analysis of inter- and intra-provincial commercial vehicle activity chains*. Mini-dissertation. Department of Industrial and Systems Engineering, University of Pretoria.

Van Heerden, Q. and Joubert, J. W. (2012). Commercial vehicle behaviour: analysing GPS records. In *42nd International Conference on Computers & Industrial Engineering*. Paper 99. ISSN: 2164-8689. Available from: <http://conferences.sun.ac.za/index.php/cie/cie-42/paper/viewFile/99/156>.

Van Heerden, Q. and Joubert, J. W. (2013). Large Scale Multimodal Transport Modelling, Part 2: Implementation and Validation. In *32nd Annual Southern African Transport Conference*, pages 653–661. ISBN: 978-1-920017-62-0. Available from: <http://hdl.handle.net/10204/6972>.

Van Heerden, Q. and Joubert, J. W. (2014). Generating intra- and inter-provincial commercial vehicle activity chains. In *Procedia - Social and Behavioral Sciences*, volume 125, pages 136–147.

Appendix A

Permission from the RFA

quintin.vh@gmail.com
9 April 2014

Ms Sharmini Naidoo
CEO: Road Freight Association of South Africa

Dear Ms Naidoo

I am a Masters student at the University of Pretoria and the subject of my dissertation is commercial vehicle transport modelling from a supply chain perspective. I model the interaction between the Carrier and Receiver. The Carrier needs to do tour planning by scheduling its vehicles to deliver all the products to its customers. The utilisation of vehicles that are available in the fleet is dependent on the operating cost of the vehicles. It is in this regard that I need to use the RFA Vehicle Cost Schedule values per kilometer for different vehicle sizes. The idea is not at all to present the values in the VCS as an industry standard, but rather use a credible source for an estimate of the costs used in my models. I then apply different scenarios such as evaluating the effect of traffic congestion on vehicle utilisation as well as the effect of the customer's order policy on vehicle utilisation, tonne-km, and total transport cost.

What I want to include from the 2011 VCS is the following of five of the vehicle concepts:

- the concept number
- the picture
- fixed cost per km
- variable cost per km

This is to visually show which truck concepts I chose when I refer to them in the dissertation as well as how I obtained the values. The sizes of the vehicles in my models aren't all available in the VCS. I thus calculate estimated values from the "Van" column for the closest equivalent refrigerated truck.

Since the Vehicle Cost Schedule is a copyrighted document, I hereby request the Association's permission to use parts of the Vehicle Cost Schedule in my Masters dissertation as explained above. I will appreciate the permission in writing on the Association's letterhead.

Regards
Quintin van Heerden

The Road Freight Association
Reg. No.1981/010367/08 Non Profit Company (NPC)
Unit E1: Isando Industrial Park
corner Gewel Street and Hulley Roads, Isando
P O Box 511, Isando, 1600
Telephone (011) 974-4399
Fax (011) 974-4903
e-mail: enquiries@rfa.co.za
website : www.rfa.co.za

Die Padvragvereniging
Regnr.1981/010367/08 Vereniging sonder winsbejag (NPC)
Eenheid E1: Isando Industriële Park
h/v Gewel en Hulley Str, Isando
Posbus 511, Isando, 1600
Telefoon (011) 974-4399
Faks (011) 974-4903
e-pos: enquiries@rfa.co.za
webwerf: www.rfa.co.za

12 May 2014

Quinton van Heerden

By email: quintin.vh@gmail.com

Dear Quinton

RE: RFA Vehicle Cost Schedule

The Road Freight Association gives permission to Quinton van Heerden to use our Vehicle Cost Schedule in his dissertation. The following information may be used from the 2011 VCS:

- The concept number
- The picture
- Fixed cost per km
- Variable cost per km

Kind Regards,

S. Naidoo (Miss)
Chief Executive Officer
The Road Freight Association

Without Trucks South Africa Stops

l