

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

# CONTEXT SENSITIVE ROAD PLANNING FOR DEVELOPING COUNTRIES

By

EDWARD ANDREW BEUKES

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY  
IN CIVIL ENGINEERING

University of Cape Town

August 2011



UNIVERSITY OF CAPE TOWN  
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD



# CONTEXT SENSITIVE ROAD PLANNING FOR DEVELOPING COUNTRIES

By

EDWARD ANDREW BEUKES

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY  
IN CIVIL ENGINEERING

In the Department of Civil Engineering

Faculty of Engineering and the Built Environment

University of Cape Town

August 2011



UNIVERSITY OF CAPE TOWN  
IYUNIVESITHI YASEKAPA - UNIVERSITEIT VAN KAAPSTAD

# Copyright

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

© University of Cape Town 2011

University of Cape Town

# Declaration

I declare that the contents of this thesis are entirely my own work, except for the specific and acknowledged references to the published work of others made in the text.

To the best of my knowledge and belief, it contains neither material previously published by another person nor material to which a substantial extent has been accepted for the award of any other degree of the university.

Candidate:

Edward Andrew Beukes: .....

Signed at:..... this ..... day of ..... 2011.

Supervisor:

Associate Professor Marianne Vanderschuren: .....

Signed at:..... this ..... day of ..... 2011.

# Abstract

Improving mobility is key to facilitating the economic upliftment of the urban poor. In South Africa, the majority of the urban poor live on the periphery of cities. They travel long distances, at great cost, to go to work and school, and are dependent on public transport and walking or cycling (Non-Motorised Transport or NMT) for their travel needs. Legislation and policies introduced after the demise of Apartheid, emphasises the importance of public transport and NMT, but road planning practice in South Africa continues to be, largely, automobile-centric. The needs of other road users are often overlooked, even in areas where they are in the majority.

The motivation for this research was to understand why, despite an improved legal and policy environment, road infrastructure planning and design practice in South Africa still tends to be, mainly, focused on the needs of automobiles. Initially, investigations focused on local policies, and the design manuals that guide road planning and design practices, but soon moved to a deeper investigation of road planning theory, specifically to the influence and origins of road categorisation systems, and their effect on road design. Various alternative approaches to road categorisation from around the world were investigated, and this led to some research around sustainable transport practices and how they compare with current road planning practices. It was from here that the central premise of this thesis originated.

Whilst acknowledging that current road planning and design practices are rooted in established, scientifically sound methods and research, a need was nonetheless identified to incorporate more contextual information into the transport planning process, to produce a more comprehensive, holistic and multimodal approach to practice. A description of the context, defined in terms of land use, socioeconomic, environmental and transport information, was developed and forms the basis of a method for making recommendations for road infrastructure provision.

The technique used to incorporate these additional sources of information is called Spatial Multiple Criteria Assessment (SMCA), and is the spatial application of Multiple Criteria Assessment (MCA), a Decision Support System. The analysis is conducted in a Geographic Information System (GIS), since the context, as it is defined in this research, varies spatially. The results of this analysis are analysed and extracted before being clustered using a combination of partitional and hierarchical clustering algorithms. Clustering the results allows for the identification of a range of contextually unique areas, the characteristics of which can then be used as a basis for infrastructure recommendations.

The research identified three case study routes in Cape Town, South Africa, which were used to develop, test and calibrate the methods. The data used in the study was sourced from a number of (primarily) government agencies.

The research shows that the local context is an important factor to consider when planning roads. The method developed shows how this local context can be quantified and how the analytical treatment of contextual information can be used to make infrastructure recommendations. The method developed, thus, brings to transport planning a novel approach to dealing with the 'access' function of roads in a systematic way.

University of Cape Town

*“If anything is certain, it is that change is certain. The world we are planning today will not exist in this form tomorrow.”*

Phil Crosby (1926 - 2001)

This PhD is dedicated to my family  
for their love, support and unending faith in me.

University of Cape Town

# Preface

Upon completion of my Masters degree at the University of Cape Town (UCT), I had a definite desire to study further, although I was unsure as to exactly what it was I wanted to do. I had kept in contact with my Masters supervisor, A/Prof. Marianne Vanderschuren, and let her know that I was contemplating further research, possibly leading to a PhD. It was A/Prof. Vanderschuren, then, that contacted me during 2007, to let me know of an opportunity to do PhD research at UCT through the Cycling Academia Network (CAN), based in The Netherlands.

CAN was a research group that was one of the projects of the Interface for Cycling Expertise (ICE), also based in The Netherlands. The research group was newly formed at the time and, as the name implies, was explicitly focussed on cycling research. The research I had done towards my Masters degree, however, was on Intelligent Transport Systems and, although I had had some exposure to non-motorised transport theory through coursework, the match was, initially, not obvious. Nonetheless, A/Prof. Vanderschuren recommended me as a candidate for inclusion in the research group and, after an interview process and the submission of a preliminary research proposal, I was accepted into the program.

I am, therefore, very much so, indebted to my supervisor, A/Prof. Vanderschuren for her faith in me, and her continued guidance. The many opportunities that I have received as a direct result of her intervention, have enriched my life, both personally and professionally, in ways that I still continue to discover each day. Together, we have walked quite a path already, and I can only hope that we will accomplish much more in the years to come.

I am also very thankful for the continual support and thoughtful advice of Dr. Mark Zuidgeest, from the University of Twente in The Netherlands, who acted as a co-supervisor of my research. Your efforts have certainly played a defining role in this research, and the support I received from you over the last four years has been invaluable. I look forward to continuing our collaboration in future.

The research group that I joined, CAN, proved to be an invaluable support network over the years, and the contacts that I have made from across the world have all, in one way or another, influenced the final shape of my research. I am grateful for the camaraderie and advice that I have received from everyone at CAN. The relationships I have developed with my colleagues at CAN, both professional and personal, will, surely, only strengthen in future.

This PhD was funded, in part, by ICE. I would like to express my gratitude to the people at ICE for their intellectual support and financial contribution to my research. I am also extremely grateful for the financial support provided by the National Department of Transport via the Southern Transport Centre for Development. Without these sources of funding, this PhD would not have been possible.

This PhD research was carried out at UCT, in the Department of Civil Engineering. I am very grateful for all of the assistance and support provided by the staff members in the department and the faculty. In

particular, I want to thank Cheryl Wright for her assistance with editing my thesis. Your contributions are highly valued.

During the course of my PhD, I spent a significant amount of time doing research at the Institute for Geo-Information and Earth Sciences (ITC) at the University of Twente in The Netherlands. The assistance and guidance I received from various members of staff at the ITC has played an important role in the development of the research work, and I am very thankful for the inputs I received there. I hope to visit in future again.

I was very fortunate to have received a large amount of the data required to conduct this research from various departments at the City of Cape Town Municipality. I was impressed by the professionalism of the staff and by the quality of the work being carried out there. I am extremely grateful for the assistance I received in this regard, since it has saved me an enormous amount of time and effort, making it possible to complete this research in a reasonable amount of time.

I would not have been able to complete this research without the patient support of my colleagues at ERO Engineers (Pty) Ltd. The time and resources given to me to be able to complete this research was invaluable and is highly appreciated.

Lastly, I would like to communicate my deepest gratitude and respect to my family, who have supported me throughout my life. To my dearest wife, Kaashifah, your patience, love, support and understanding has helped me overcome the difficulties I have encountered, and has inspired me to be the best I can be. To my mother and father, your keen interest and enduring support and encouragement has allowed me to pursue my ambitions and to take on challenges that would, otherwise, not have been achievable. I am forever grateful to you all.

Edward Andrew Beukes  
August 2011

# Executive Summary

An overview of the major findings of each chapter of the report is provided below.

The motivation for this research was to understand why, despite an improved legal and policy environment, road infrastructure provision in South Africa still tended to be mainly focused on the needs of automobiles. The belief was that planning and design practice had not fully embraced the ideals of policies, and that this played a significant role in the poor quality and piecemeal nature of the infrastructure provided for the other road-based modes.

Investigations initially focused on local policies, and the design manuals that guide road planning and design practices, but soon moved to a deeper investigation into road planning theory, specifically on the influence and origins of road categorisation systems and its effects on road design. Various alternative approaches to road categorisation from around the world were investigated, and this led to some research around sustainable transport practices and how they compare with current road planning practices. It was from here that the central premise of this thesis originated.

Acknowledging that current road planning and design practices are rooted in well established, scientifically sound methods and research, a need was nonetheless identified to incorporate the principles of sustainable transportation to produce a comprehensive, holistic approach to practice. A problem statement was developed as:

*Each mode in use in a road has its own specific characteristics and needs, and these determine the design parameters for that mode. Also, each location in a network or along a road is defined by a set of contextual parameters that determine how and by whom it is most often used. It is in the intersection between the modal characteristics and the location specific factors, or the needs of a mode and the use of a location, that an ideal planning solution can be found. Accordingly, certain modes are better suited to a particular set of contextual circumstances than others. Therefore, under a given mix of contextual circumstances, certain modes should be given a higher priority than the rest. This modal priority ranking can then be used to inform road planning recommendations, using established practice and guidelines to inform decisions.*

The research outlined in this thesis, therefore, investigates whether it is possible to rethink the road planning process to incorporate a broader range of factors than is currently used so that a more sustainable, contextually sensitive road facility can be developed. Several questions need to be answered before this main question is addressed.

1. How are roads currently planned and designed, what are the reasons for this approach, and what are the benefits and disadvantages of the current approach to road planning?
2. What are the legal, policy and regulatory frameworks that affect road planning practice, and to

- what extent do these influence the planning practice?
3. What is the role of context in the transport environment?
  4. What are the factors that best describe the context of a route and to what extent do these factors currently play a role in road planning practices?
  5. What would be an appropriate platform to incorporate multiple additional factors in the planning process?
  6. What are the data requirements for a context analysis?
  7. Which corridors would be best suited for developing and testing the method?
  8. What are the crucial elements of the method that need to be developed and validated?
  9. How could the results of a contextual analysis be used to make road planning recommendations?

Internationally, there is now an established appreciation for the problems caused by rampant motorisation and automobile dependence. Issues as varied as climate change, public health, and the economy have driven perceptions towards a more balanced approach to the provision of transport infrastructure. This, coupled with South Africa's recent political transition to democracy, has given impetus to a review of transport policy and legislation. This process has only recently reached some level of finality with the passing of the National Land Transport Act in 2009. Aside from aligning local practices with current international norms, the new policies were developed with the important final objective of achieving the social and economic transformation required to address the disparities caused by the Apartheid system.

The objectives of policies and legislation can, however, only be realised if they are effectively translated into regulations and infrastructure provision, and are universally adopted by practice. The progress achieved on this front has not been satisfactory, and the result is that infrastructure provision has tended to be in conflict with policy objectives, although, it would appear from spending priorities that this may be slowly changing.

In terms of road safety, one of the most important factors is speeding, and since speed is one of the primary considerations in the design of infrastructure, there are a large range of tools available to designers to plan roads with speeds that are appropriate to the circumstances. The design speeds selected during the planning stage must reflect these circumstances as best as possible, and the final road design must ensure that the operational speeds on the road accord with the design speed. The posted speed limit must reflect the drivers perception of what is a safe and reasonable speed to travel at, and this information must be provided to the driver from the road design itself.

The design speed, operational speeds and the speed limit, then, should to a large extent be informed by the circumstances along the route. Currently, design speeds are selected based upon the road classification which, as was shown in Chapter 2, relates more to the functional hierarchy of the road network and the traffic volumes the route carries than to the specific circumstances along the route itself. It must also be considered that these circumstances may vary from place to place along the route, and so, if it is to be influenced by local circumstance, the design speed should vary in step. However, this is not practical with current planning norms.

Context Sensitive Design (Context Sensitive Design (CSD)) has in recent years has been promoted by some of the major transportation associations in the USA as being able to provide outcomes that harmonize the transportation requirements expected of a project with the needs and values of the

community it serves or passes through. The need for such harmonisation arose as a result of the growing awareness of the oft unaccounted for costs, the externalities, that the construction and the operation of road infrastructure can exact. Communities began to put a greater value on the environmental, cultural and social assets that may have been threatened by a project than on the transport benefits offered by it.

CSD developed in response to these challenges. It can be described as a project management approach driven by a set of principles that addresses the conceptualisation and development, the role and participation of stakeholders, the design and planning considerations, the construction and financial management and the operations and maintenance planning of a project. However, comprehensive though it may be, CSD has struggled to gain traction amongst planners and designers, in part because it suffers from considerable ambiguities, and because of concerns around professional liability issues regarding exploiting the flexibility in design standards to meet its principles. Although there are clear benefits to the approach, it has not managed to find support beyond the USA yet.

The complete streets movement has many similarities to CSD, but focusses more on the need to create adequate facilities for all modes, with a focus on promoting walking and cycling amongst urban communities. This aspect has garnered it support amongst community organisations, advocacy groups and in certain academics circles concerned with the consequences of low physical activity on health, especially for children. Further support for the movement is due to its promotion of social justice or equity issues relating to improving accessibility for the poor. The movement has been limited to a number of pilot projects in communities in the USA, and, like CSD, suffers from an ambiguous, ill-defined scope and application protocol, and despite growing interest, has thus far not been widely integrated into practice.

The ARTISTS project was initiated to address the perceived shortcomings of planning for arterial streets, where the primary functional character of the route is approximately evenly split between access to properties and mobility for through vehicles. The project developed a new approach to road categorisation that differentiated between roads based upon the relative importance of locale-specific aspects along it, and the links importance to mobility in the network. The scheme is notable in that it highlights the importance of balancing aspects related to the 'place', and those related to mobility, and in that it recognises that these qualities may vary independently both in time and in space along the route. The method is, however, limited insofar as it is configured only for arterial routes and fails to develop a convincing, quantitative method for comparing either the 'place' or 'mobility' aspects of the route, as well as the segmentation of the route.

CSD, complete streets, the ARTISTS project and the planning, design and transport road safety problems highlights the need to consider the context of the location along the route as well as the mobility needs when planning the road. The link between this context and transport infrastructure lies in the modes of transport that the infrastructure is designed to accommodate. Therefore, identifying which mode is most suitable given the context, and prioritising the needs of that mode would ensure that the infrastructure provided is context-sensitive, thereby ensuring more equitable, potentially safer infrastructure.

This context can be defined as being related to who is using the road and why they are there. Demographic information on the road users, and information on the activities being performed at the location is required to answer these questions. This information is already recognised as being important to the prediction of travel behaviour, and in particular modal split. Its application in the description of the context is thus particularly apt. The information that best describes the context can be divided into four categories. These include land use information, socioeconomic information, environmental information

and transportation information.

Variables that are commonly cited in transportation literature were considered for these categories, with the emphasis placed on variables that impact travel behaviour, and in particular, modal split, since modal priority is being used to express context in terms of infrastructural choices. Variables identified as being important include the land use type, population or household density, employment density, income, age, culturally and ecologically sensitive areas, the relative demand for each mode in the area and the location of Public Transportation (PT) stops.

Since the definition of the context involves numerous attributes, a method capable of taking into account multiple inputs, all of which may vary spatially, was required to assess the context. A special application of Multiple Criteria Decision Analysis (MCDA), spatial multiple criteria analysis (SMCA) was employed to order and assess the criteria identified in Chapter 4 in an formalised evaluative framework capable of producing measurable results. SMCA is typically used to analyse and develop solutions for decision problems with spatially varying attributes.

The criteria maps used in the assessment have to be standardised to be able to perform the arithmetic operations used in the aggregation step on them, but also to express the criteria's relationship to the alternative (whether the criteria is a cost or a benefit in the context of the assessment). Typical with SMCA studies, there is usually only one standardisation protocol developed, but the particular application of the method for this assessment required that separate standardisation protocols be developed for each alternative. The method develops five level of suitability maps, one for each mode of transport. The framework developed for the standardisation protocols relies on the identification of distinct value statements that are then translated into impact vectors by interpreting the value statements in terms of indicators related to the criterion, such as speeds or volumes.

The standardised criterion maps must then be converted into raster format. The rasterised maps can then be weighted. The weighted map scores are then spatially aggregated using a map summation algorithm. The output is a spatial level of suitability map. This map is a representation of the spatial preference scores for that mode. One level of suitability map is, therefore, generated for each mode. The level of suitability maps must then be reprocessed using a spatial averaging technique to draw in information from areas beyond the road centreline into the analysis. The averaged level of suitability map scores along the road are then extracted and can be plotted in a spreadsheet to compare the mode suitability scores and identify the modal ranking at every point along the road.

Three case study routes were selected, Voortrekker Road between Salt River and Bellville, Lansdowne Road between Claremont and Macassar and Koeberg Road between Maitland and Table View. These routes were selected based upon a combination of the the road safety problems and the variety of land uses and community types along their courses. Primary data sources included various departments of the City of Cape Town municipality and the 2001 National Census.

Suitability rasters were generated for the whole city, and an analysis of these results was conducted. With the criteria used in this SMCA, bicycle, pedestrian and PT modes were found to have somewhat similar cumulative suitability scores, which was generally higher than that of freight and car. There were, however, still noticeable differences between the scores for each mode.

The suitability of a particular mode to a specific location can be influenced by weighting the criteria to reflect the value judgements of the decision makers, and weighting can be used to explore various preference scenarios. Although weights have been used in this way in other SMCA studies before, using weights in this application presented some theoretical difficulties since the spatial performance

of each mode in relation to each criterion was assessed by defining the relationship between the mode and the criterion as either a spatial cost or a benefit. Therefore, an identical increase in the value of all the alternatives for any criterion at any one location would have different effects on the relative performance of the alternatives for that location. The use of weights is also somewhat problematic since the standardisation process already involves the application of value judgements to assess the performance of an alternative relative to a criterion. The influence of additional weighting on the final score would then be a sort of 'double weighting', since value judgements are being introduced at two points in the procedure.

Since weighting strategies act so as to impose additional value judgements on this interpretation of the context, the values statements expressed in the standardisation framework, therefore, become the distinguishing feature against which such weighting strategies are evaluated. Therefore, the stated values would be the drivers of the weighting scheme selected, and not the individual criteria per se. Following this approach would imply that the weights across all the criteria used in the evaluation will not sum to 1, and so an additional normalisation step must be used to ensure that the weights used sum correctly. The land use variable was standardised separately, since the different modes each have different relationships to the various land uses. This approach to weighting overcomes the theoretical difficulties described above.

The effects of three weighting schemes were tested on the suitability scores for the Voortrekker Road case study. The effects of different weighting schemes on the results were found to be significant. The bulk of the variation in the results lies in the suitability of the car and freight modes, which were found to vary dramatically along the route. The different weight schemes were found to produce different rank profiles along the routes in different places. Some areas of the route have a different rank profile under each scenario, whereas others always have the same rank profile.

Areas of rank stability were found to be important because they have contextual features that make them highly suited to certain modes irrespective of weighting. Areas of rank instability are also important, since depending on the weight scheme used (from the decision makers perspective or priorities), the rank profile in these areas may change. It is, therefore, important to explore how these changes occur, and which criteria are responsible for the reversal in rankings. Weighting is, therefore, an important consideration in the analysis, and results should be explored under a range of weighting scenarios to explore the context as thoroughly as possible before making decisions about infrastructure provision.

Although the analysis of the SMCA data can provide the planner with useful information, the variability or 'noise' within the output makes it necessary to simplify the results that are produced by clustering similar sections into groups that can then be analysed further and applied as proxies for the actual scores. These proxy results can be thought of as rank types, for which road design templates can be developed that would best suit the context. Clustering the data assists with ascertaining the underlying structure of the data sets in order to gain clearer insight into the data, evaluate the hypothesis, detect anomalies, and identify salient features. It also helps to establish a natural classification by establishing the degree of similarity between sections of the case study routes. It is also used to simplify the data by compressing the original data into clusters that can be used to organize the data into route typologies.

The two most popular approaches to clustering are partitional and hierarchical. Partitional algorithms construct partitions in the data such that each cluster optimizes a pre-stated clustering criterion. Hierarchical algorithms create a hierarchy amongst the sets of data points. Hierarchical algorithms can be either agglomerative or divisive. Agglomerative algorithms start with each object being a separate cluster itself, and successively merges groups according to a distance measure.

Although a large range of clustering techniques were identified, K-means clustering was used since it is a well established method, is flexible, and is able to produce reliable results very quickly. The clusters produced were verified against the results of an agglomerative hierarchical clustering, using a range of techniques to identify the most appropriate number of clusters for each dataset.

An initial investigation of the data revealed that the modes could be split into three groups according to their suitability scores along the route; car and freight, and pedestrian and bike obtained similar scores, and could be combined into mode groups. Public transport formed the third group. These similarities amongst the modes are problematic when clustering the data, since a strong clustering relies on there being distinct differences in the elements being clustered. The less differences there are, the less distinct are the clusters that are formed. The data was, therefore, first merged into three groups to identify the appropriate number of clusters for each case study. These cluster numbers were then used to cluster the original data set. The clustering techniques used produced 6 clusters along Voortrekker Road, 7 clusters along Lansdowne Road, and 8 clusters along Koeberg Road.

Clustering the data reveals which sections of the route are contextually similar to each other. Clusters can be seen to be repeated within all of the data sets, which shows that there are contextually similar areas along the route that are spatially separate from each other. These areas should, according to the hypothesis, have similar infrastructure, since the modal priorities, in terms of the context, are very similar. Clusters range from approximately 500m to 3000m in length, and in general, cluster means tend to change in a stepwise fashion, meaning that cluster mean scores do not generally change erratically, instead they increase and decrease in increments. This indicates that contexts also tends to change incrementally, and that modal priority tends to increase or decrease incrementally as well.

The method used to develop infrastructure recommendations relies on descriptive statements of the mode suitability. The mechanism used to translate mode suitability into such a descriptive statement uses the rank of the mode, and the score of the mode in relation to the others in the cluster. The cluster means are mapped to a linear scale that sets the highest score of any mode in any cluster along the route (the route maximum) to 1 and the lowest score (the route minimum) to 0. The transformed cluster means can then be classified into 4 score bands, highest suitability, medium suitability, low suitability and unsuitable.

For each suitability band, the appropriate operational limitations of the mode are described in terms of three parameters: the access afforded to the mode, the right of way or priority of movement afforded to the mode, and the level of independence of operations afforded to the mode. These operational characteristics are then used to determine the appropriate combination of infrastructural and regulatory interventions that would be most likely to produce the desired service levels.

The work conducted towards this thesis made a number of novel findings that contribute to knowledge, and have the potential to improve the status quo. The basis of the work was that there are a range of factors that can be used to describe the nature of the people using a road, and the activities they are involved in along it. This information should play a more direct role in the planning of these roads. These factors, collectively termed the context, have always in some sense been considered in transport planning, but have tended to have been overshadowed by concerns around efficiency and cost, and there has not been a comprehensive framework within which the context could be evaluated, and its implications investigated. The way in which infrastructure interfaced with or suited the context has, thus, always been left to the discretion or judgement of the engineer or planner of the facility. The development of a method to systematically evaluate the context and interpret the results in terms of infrastructure is, therefore, the primary contribution of this research.

The research identified which factors could be used to describe the context of a location, and demonstrated that it is possible to quantify the context in terms of its effects on the suitability of the various modes of transport. Quantification has a number of advantages for planning and designing infrastructure. Being able to quantify the suitability of a mode of transport to a particular location given the context allows one to prioritise between modes in terms of infrastructure provision, which can then be used as the basis for planning and design.

Quantification also has the benefit of facilitating accurate comparisons between different locations along the route. This is useful for planning and design in that the subtleties in variation of the context along the route are retained during the analysis, allowing for a fine grained tailoring of the required infrastructure. It also demonstrates that context varies spatially, that it is not static, and that, therefore, for infrastructure to be contextually sensitive, it must vary to suit.

The method developed to assess the context employs a novel application of the principles of multiple criteria assessment to conduct the evaluation. Whereas SMCA had previously been used to compare the suitability of a number of sites for a development, or to identify routing alternatives in an analysis space, the application developed here assesses suitability of a number of alternatives in a constrained space, and uniquely, no alternative is necessarily taken as being the one correct solution, instead, it may only be the correct solution at that location. Also, none of the other alternatives are abandoned if they do not rank highest, instead, their relative suitability is used to determine their priority in terms of the operational aspects of the road being planned.

The standardisation framework used in the SMCA is also novel in that for any one criterion the alternatives being assessed cannot logically be standardised using the same standardisation method. A separate system of assessment values, therefore, needed to be employed to guide the standardisation used for each criterion for the various alternative modes. This introduces an element of subjectivity into the analysis at an early stage, and complicates the application of weighting using traditional weighting methodologies. To overcome this, the same system of assessment values was used to derive appropriate weights, thereby ensuring that a consistent set of values is used throughout the assessment.

The context, being an amalgam of a range of disparate factors, does not have any intrinsic meaning by itself. Instead, it is the implications of the context that has meaning, and, therefore, context can only be understood in terms of its implications for other aspects of the facility. In this research, context is defined in terms of its implications for the suitability of the various modes of transport. The translation of contextual suitability into infrastructure recommendations, therefore, relied on defining the operations of the modes of transport on the road. This approach is somewhat similar to the definition of Level of Service (LOS), in that it also describes service levels in terms of operational characteristics and performances, however, in this instance these descriptions are prospective, in that they define what should be, instead of retrospective, or defining what is.

The methodology developed in this research, therefore, successfully develops a definition of the context, demonstrates its importance, explores its characteristics and implications and translates these into descriptives that can be used to inform infrastructure provision.

The findings of this thesis indicate that it is important that the contextual setting be considered in road planning. Currently, a systematic analysis of the contextual variation along the route does not form part of road planning practice. In addition, current planning practice results in infrastructure provision that is biased towards the needs of Private Motorised Transportation (PMT). Policy priorities are, therefore, not being met as effectively as they should be, which impacts negatively on the success of efforts to improve the lives of the underprivileged, and contributes to the road safety problems on the countries

roads. Contextually sensitive planning could help to correct the disjuncture between the objectives of transport policy, and transport planning. It is, therefore, recommended that an analysis of the context be included in the current transport planning processes.

This has implications for the project cycle. Road projects typically include a feasibility or preliminary investigation stage, the findings of which is then presented in a report. This is followed, after some deliberation between stakeholders, by a detailed design phase during which construction drawings and contract documentation is produced. The major decisions regarding the extent and focus of the project are, therefore, finalised before the detailed design phase, and it is, therefore, during this initial stage of the project cycle that an analysis such as the one developed in this research would be most beneficial. It is recommended that planning authorities require that a contextual assessment be included as part of the initial project planning investigations.

Context studies could also be conducted prior to any project being initiated, as part of the overall strategic planning for an area. This information would be useful to investigate different planning scenarios, and could assist in identifying current deficiencies in the existing network. Context investigations could help provide an explanation for localized road safety problems, and could be used to identify interventions that may improve conditions in these areas. It is recommended that authorities conduct an analysis of the type developed in this research on their road networks.

Studies such as these are data intensive, and although geocoded data is expensive to generate, this data can be used to analyse a range of problems, and is very useful for planning in general. One of the primary shortcomings of this research is that the data used was all at different levels of aggregation, and not all from the same reference year. This has negative implications for the precision with which such studies can be conducted. It is recommended that more effort is expended in geocoding data sets and bringing existing data sets up to date.

The work conducted in this research also raises a number of opportunities for further research. Firstly, it is conceivable that, depending upon data availability, certain criteria may not be able to be included in the analysis. It would be beneficial to establish what the minimum or 'core' data requirements for a reasonable description of the context should be. The impacts of excluding certain information should be examined in more detail. Also useful would be a better understanding of the implications of including additional, non-core data in the analysis, and the development of guidelines for criteria selection that would outline these findings.

There is also opportunity for developing a more streamlined process for the cluster analysis, and in particular the selection of the correct number of clusters to use for the final results. This part of the methodology requires a detailed understanding of clustering theory that may complicate the widespread adoption of the methodology. Data clustering is a constantly evolving, rapidly developing field, and there is definitely scope for a more comprehensive investigation of the different clustering techniques that have been developed.

All of the case studies investigated during the development of the methodology were urban arterial routes in Cape Town. The rationale for this decision has been outlined previously, but there is definitely opportunity for investigating the application of the method on non-arterial or even rural routes. It is suspected, based upon a cursory investigation of the SMCA results, but has not been verified, that the recommendations for these types of roads will be in accordance with the findings for the case studies used. It could also be useful for the same reasons to investigate the results for other areas in South Africa and elsewhere.

There is a definite opportunity, beyond the scope of the work conducted for this thesis, to adapt the methodology for an area-wide analysis. Given that the SMCA was conducted across the city as a whole, such an adaptation will most likely involve clustering the SMCA results across the whole city. It is hoped that this work will be able to provide further insight into the contextual structure of the city as a whole, and urban areas in general, and open up opportunities for developing a generalised suite of context types.

University of Cape Town



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Road Planning, Sustainability and Context . . . . .	1
1.2	Problem Statement and Research Questions . . . . .	4
1.3	Scope and Limitations . . . . .	5
1.4	Research Approach and Methodology . . . . .	5
1.5	Outline of the Thesis . . . . .	7
<b>2</b>	<b>Planning, Policy and Practice</b>	<b>9</b>
2.1	Transport Planning - Past and Present . . . . .	9
2.2	Transport Modelling . . . . .	17
2.3	Road Planning and Policy in South Africa . . . . .	25
2.4	Guidelines and Transport Policy . . . . .	33
2.5	Résumé . . . . .	46
<b>3</b>	<b>Transportation and Road Safety</b>	<b>49</b>
3.1	Road Accidents and Driver Behaviour . . . . .	49
3.2	Policy Responses to Road Accidents: Theory and Myth . . . . .	50
3.3	Speeding and Road Safety . . . . .	55
3.4	Traffic Calming and Road Safety . . . . .	63
3.5	Résumé . . . . .	70
<b>4</b>	<b>Contextualising Road Design</b>	<b>71</b>
4.1	Context Sensitive Solutions . . . . .	71
4.2	Complete Streets . . . . .	74
4.3	ARTISTS . . . . .	75
4.4	Defining Context . . . . .	78
4.5	Résumé . . . . .	97
<b>5</b>	<b>Multiple Criteria Decision Making</b>	<b>99</b>
5.1	Background . . . . .	99
5.2	Spatial Multiple Criteria Assessment (SMCA) . . . . .	104
5.3	SMCA and Context . . . . .	107
5.4	Standardisation Framework . . . . .	108
5.5	Aggregation . . . . .	113
5.6	Résumé . . . . .	121

<b>6</b>	<b>Data Sources and Case Study Selection</b>	<b>123</b>
6.1	Data Sources, Verification and Limitations . . . . .	123
6.2	Case Study Selection . . . . .	131
6.3	Case Study Route Descriptions . . . . .	134
6.4	Résumé . . . . .	136
<b>7</b>	<b>Case Study Results and Assessment</b>	<b>139</b>
7.1	Level of Suitability Rasters . . . . .	139
7.2	Extracted Raster Scores . . . . .	142
7.3	Weighting, Values and Scenario Planning . . . . .	151
7.4	Sensitivity to Weighting . . . . .	152
7.5	Résumé . . . . .	158
<b>8</b>	<b>Cluster Analysis</b>	<b>163</b>
8.1	Introduction to Cluster Analysis . . . . .	163
8.2	Clustering the SMCA Data . . . . .	167
8.3	Résumé . . . . .	182
<b>9</b>	<b>Infrastructure and Context</b>	<b>183</b>
9.1	Analysis of Route Clusters . . . . .	183
9.2	Infrastructure Recommendations from Cluster Analysis . . . . .	185
9.3	Operational characteristics and Infrastructure . . . . .	187
9.4	Interpreting Modal Suitability . . . . .	189
9.5	Résumé . . . . .	195
<b>10</b>	<b>Conclusions and Recommendations</b>	<b>197</b>
10.1	Answers to the Research Questions . . . . .	198
10.2	Reflection . . . . .	203
10.3	Recommendations . . . . .	205
	<b>Bibliography</b>	<b>206</b>
	<b>Appendices</b>	<b>231</b>
	<b>Appendix A Planning, Policy and Practice</b>	<b>231</b>
	<b>Appendix B Transportation and Road Safety</b>	<b>233</b>
	<b>Appendix C Multiple Criteria Decision Making</b>	<b>235</b>
	<b>Appendix D Data Sources and Case Study Selection</b>	<b>237</b>
	<b>Appendix E Case Study Results and Assessment</b>	<b>247</b>
	<b>Appendix F Cluster Analysis</b>	<b>251</b>
	<b>Appendix G Infrastructure and Context</b>	<b>261</b>

# List of Figures

1.1	Overview of research method . . . . .	6
2.1	Central channel street in Victorian London . . . . .	10
2.2	Functional categorisation of roads according to Buchanan . . . . .	13
2.3	The relationship between ‘mobility’ and ‘access’ . . . . .	13
2.4	Traditional functional classification system compared to the tri-partition system . . . . .	14
2.5	Traditional ‘predict and provide’ transport modelling process . . . . .	17
2.6	The traditional transport model . . . . .	19
2.7	Diversion curves for decision choice behaviour . . . . .	22
2.8	Provincial roads and transport expenses 2007-2008 . . . . .	33
2.9	Cape Town roads and transport expenses 2008-2012 . . . . .	34
2.10	South African economic growth rates 2001-2008 . . . . .	36
2.11	Main mode of travel to work in South Africa . . . . .	37
2.12	Cape Town accident statistics by accident type . . . . .	38
2.13	Hierarchy of Internal Township Roads . . . . .	40
2.14	Red Book proposed configurations of roadway systems . . . . .	43
2.15	Red Book illustration: Typical elements of the cross-section . . . . .	44
2.16	Draft UTG 1 illustration: Typical elements of the cross-section . . . . .	45
3.1	RSA monthly accident rate and average monthly accident rate . . . . .	51
3.2	Time spent driving over lines . . . . .	59
3.3	Average speed change for different calming devices . . . . .	66
3.4	Speed profiles recorded in four streets in a 200 m range . . . . .	67
3.5	Average traffic volume change for different calming devices . . . . .	68
3.6	Average reduction in accident frequency for different calming devices . . . . .	69
4.1	ARTISTS Functional Classification . . . . .	77
4.2	Household access to cars by income group . . . . .	86
4.3	Weekday trip-making, by age group of household members . . . . .	88
4.4	Percentage of urban NMT trips in the USA, Germany and The Netherlands . . . . .	90
4.5	Use of transport modes by age group . . . . .	91
4.6	Sources of brown haze in Cape Town, 1997 . . . . .	94
5.1	General Schema of MADM Methods . . . . .	101
5.2	Decision Paths in Spatial Decision Problems . . . . .	105
5.3	Extract of standardised land use scores for cars for Cape Town . . . . .	111
5.4	Extract of standardised land use scores for PT for Cape Town . . . . .	111

5.5	Extract of benefit standardised household density scores for Cape Town . . . . .	112
5.6	Extract of cost standardised household density scores for Cape Town . . . . .	113
5.7	Extract of cost standardised scores for proximity to PT stops in Cape Town . . . . .	114
5.8	Extract of equally weighted level of suitability map for PT in Cape Town . . . . .	116
5.9	Example of point value extract locations along a road . . . . .	117
5.10	Extract of averaged level of suitability map for PT in Cape Town . . . . .	118
5.11	Extract of averaged distance weighted level of suitability map for PT in Cape Town . . . . .	119
5.12	The effect of influence area size on the average suitability score . . . . .	120
6.1	Number Of Polygons and the Percentage Area Per Zoning Category . . . . .	126
6.2	Income versus Education Level . . . . .	128
6.3	Ten worst-known roads for accidents in Cape Town and education level . . . . .	133
6.4	Ten worst-known roads for accidents in Cape Town and land use zones . . . . .	134
6.5	Typical views at points along Voortrekker Road . . . . .	135
6.6	Typical views at points along Lansdowne Road . . . . .	135
6.7	Typical views at points along Koeberg Road . . . . .	137
7.1	Public Transport Suitability Scores from SMCA for Case Study Routes . . . . .	140
7.2	Frequency Distribution of Suitability Scores by Mode . . . . .	141
7.3	Cumulative Frequencies of Suitability Scores by Mode . . . . .	141
7.4	Spatial Distribution of Reclassified SMCA Scores for Pedestrians . . . . .	142
7.5	Frequency Distribution of Suitability Scores in Range 3 for Pedestrian Mode . . . . .	143
7.6	Suitability Scores along Voortrekker Road . . . . .	143
7.7	Freight Suitability Raster on Voortrekker Road, Km. 0 - 4 . . . . .	144
7.8	Freight Suitability Raster on Voortrekker Road, Km. 5.5 - 9.5 . . . . .	145
7.9	Suitability Scores along Lansdowne Road . . . . .	147
7.10	Suitability Scores along Koeberg Road . . . . .	149
7.11	Aerial Overview of Freight Suitability along Koeberg Road . . . . .	149
7.12	Comparison of Public Transport Suitability Using Different Weight Schemes . . . . .	155
7.13	Comparison of Pedestrian Suitability Using Different Weight Schemes . . . . .	156
7.14	Comparison of Car Suitability Using Different Weight Schemes . . . . .	157
7.15	Comparison of Histograms for Different Weight Schemes Profiles for Car . . . . .	158
8.1	Steps performed by the K-means algorithm . . . . .	165
8.2	Mode score relationships for Voortrekker Road . . . . .	168
8.3	Similarities between modes along Voortrekker Road . . . . .	168
8.4	Merged suitability scores along Voortrekker Road . . . . .	169
8.5	K-means clustering results for three cluster centres along Voortrekker Road . . . . .	171
8.6	K-means clustering results with 4 cluster centres along Voortrekker Road . . . . .	171
8.7	Total Sum of Distances Versus Curvature . . . . .	173
8.8	Silhouette Plot for Three-Cluster Scenario for Voortrekker Road . . . . .	174
8.9	Mean Silhouette Score versus Number of Clusters . . . . .	175
8.10	Dendrogram of Clusters along Voortrekker Road . . . . .	176
8.11	Mean Silhouette Score versus Number of Clusters - Lansdowne Road . . . . .	177
8.12	Dendrogram of Clusters along Lansdowne Road . . . . .	178
8.13	Mean Silhouette Score versus Number of Clusters - Koeberg Road . . . . .	179
8.14	Dendrogram of Clusters along Koeberg Road . . . . .	180

9.1	Comparison of Cluster Means and SMCA Scores for Voortrekker Road . . . . .	184
9.2	Comparison of Cluster Means Voortrekker Road . . . . .	185
9.3	Comparison of Cluster Means and Cluster Location on Voortrekker Road . . . . .	186
9.4	Adjusted Cluster Means for Voortrekker Road . . . . .	190
9.5	Voortrekker Road in a cluster 1 area . . . . .	191
9.6	Damrakstraat, Amsterdam . . . . .	191
9.7	Existing Layout and Proposal for Context Sensitive Upgrade . . . . .	193
9.8	Voortrekker Road in a cluster 4 area . . . . .	194
9.9	Industrial area along Voortrekker Road designated as cluster 5 . . . . .	194
9.10	Undeveloped area along Voortrekker Road designated as cluster 5 . . . . .	195
10.1	Overview of methodology . . . . .	204
A.1	The ‘Radburn Plan’ for Radburn, New Jersey. . . . .	231
B.1	A typical layout of a woonerf. . . . .	234
C.1	Conical distance weighting . . . . .	235
D.1	Location of case study roads in Cape Town . . . . .	237
D.2	Grouped Zoning Map for Cape Town . . . . .	238
D.3	Household Densities in Cape Town . . . . .	240
D.4	Percentage of Vulnerable Road Users . . . . .	241
D.5	Employment Densities in Cape Town . . . . .	242
D.6	Location of Heritage Features in Cape Town . . . . .	243
D.7	Location of Wetlands and Environmentally Sensitive Areas in Cape Town . . . . .	244
D.8	Public Transport versus Car Trips per TAZ in Cape Town . . . . .	245
E.1	Bicycle Suitability Scores from SMCA for Case Study Routes . . . . .	247
E.2	Car Suitability Scores from SMCA for Case Study Routes . . . . .	247
E.3	Freight Suitability Scores from SMCA for Case Study Routes . . . . .	248
E.4	Pedestrian Suitability Scores from SMCA for Case Study Routes . . . . .	248
E.5	Frequency Distribution of Suitability Scores for Pedestrian Mode . . . . .	249
E.6	Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 1 . . . . .	249
E.7	Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 2 . . . . .	250
E.8	Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 3 . . . . .	250
F.1	A CF-Tree used by the BIRCH algorithm . . . . .	252
F.2	Mode score relationships for Lansdowne Road . . . . .	254
F.3	Mode score relationships for Koeberg Road . . . . .	254
F.4	Silhouette Plot for Four-Cluster Scenario for Voortrekker Road . . . . .	255
F.5	Total Sum of Distances Versus Number of Clusters for K-means on Voortrekker Road . . . . .	255
F.6	Mean Silhouette Score versus Number of Clusters . . . . .	256
F.7	Perspective views of k-means clustering results for 6 cluster centres along Voortrekker Road . . . . .	256
F.8	Perspective views of k-means clustering results for 7 cluster centres along Lansdowne Road . . . . .	257
F.9	K-means clustering results for 8 cluster centres along Koeberg Road . . . . .	257

F.10	K-means clustering results for 5 and 10 cluster centres along Koeberg Road . . . . .	258
F.11	Dendrogram of 8 and 9 Cluster solutions along Koeberg Road . . . . .	259
G.1	Cluster locations along Lansdowne Road . . . . .	261
G.2	Comparison of Cluster Means and SMCA Scores for Lansdowne Road . . . . .	261
G.3	Cluster locations along Koeberg Road . . . . .	262
G.4	Comparison of Cluster Means and SMCA Scores for Koeberg Road . . . . .	263
G.5	Comparison of Cluster Means for Lansdowne Road . . . . .	263
G.6	Comparison of Cluster Means for Koeberg Road . . . . .	264

University of Cape Town

# List of Tables

2.1	Customer-based Key Performance Indicators (KPI's) for the NLTSF 2006-2011 . . . . .	32
2.2	Residential road categorisation and design parameters . . . . .	41
2.3	Comparison between classification systems . . . . .	42
3.1	South African traffic offences rates for 2005 . . . . .	50
3.2	Percentage Contribution to Accidents per Human Factor . . . . .	51
3.3	Number of Fatalities per Road User Group . . . . .	52
3.4	Ten worst road sections for pedestrian accidents in Cape Town in 2005 . . . . .	53
3.5	Summary of the effects of lowering speed limits . . . . .	61
3.6	Summary of the effects of increasing speed limits . . . . .	62
4.1	Main mode to work - % of commuters . . . . .	85
4.2	Trip generation in the RSA by household income . . . . .	85
4.3	Trip purpose, by age group of household member . . . . .	89
4.4	Impact of Age on Modal Choice in the USA . . . . .	90
4.5	Annual costs of environmental non-monetary externalities of motor vehicle use . . . . .	93
5.1	The performance table . . . . .	101
5.2	The Fundamental Scale for Pairwise Comparisons . . . . .	102
5.3	Safety maximisation qualitative ranking . . . . .	109
5.4	Mobility maximisation qualitative ranking . . . . .	110
5.5	Criteria, value statements, indicators and standardisations . . . . .	110
5.6	Criteria, value statements and indicators . . . . .	114
5.7	SMCA criteria tree . . . . .	115
5.8	Comparison of weighting scheme results . . . . .	119
6.1	Criteria, Aggregation Level, Number of Features and Data Sources . . . . .	124
6.2	Reliability of zoning labels . . . . .	127
6.3	Pearson Correlation Coefficient for Income and Education Level . . . . .	129
6.4	Education level index weighting scheme . . . . .	130
6.5	Ten worst-known road sections for accidents in Cape Town . . . . .	132
7.1	Descriptive Statistics for Voortrekker Road Suitability Scores . . . . .	145
7.2	Mode Suitability Correlation Matrix for Voortrekker Road . . . . .	146
7.3	Descriptive Statistics for Lansdowne Road Suitability Scores . . . . .	148
7.4	Mode Suitability Correlation Matrix for Lansdowne Road . . . . .	148
7.5	Descriptive Statistics for Koeberg Road Suitability Scores . . . . .	150
7.6	Mode Suitability Correlation Matrix for Koeberg Road . . . . .	150

7.7	A Hypothetical Value Driven Weighting Scheme . . . . .	153
7.8	Hypothetical Value Driven Weighting Scheme for Land Uses . . . . .	153
7.9	Value Rankings for Scenarios for Voortrekker Road . . . . .	154
7.10	Weighting Schemes for Voortrekker Road . . . . .	154
7.11	Descriptive Statistics for Weighted Analysis Results for Public Transport . . . . .	155
7.12	Correlation Between Various Weight Analysis Results for Public Transport . . . . .	156
7.13	Descriptive Statistics for Weighted Analysis Results for Car . . . . .	157
7.14	Descriptive Statistics for Case Study Routes . . . . .	159
8.1	Mean distance to closest centroid for increasing values of $k$ - Voortrekker Road . . . . .	175
8.2	Mean Distance to Closest Centroid - Lansdowne Road . . . . .	178
8.3	Mean Distance to Closest Centroid - Koeberg Road . . . . .	179
8.4	Comparison of silhouette score means for merged data sets and original data sets . . . . .	181
8.5	Cluster centroid details for case study roads . . . . .	181
9.1	Mode suitability in terms of operational permissions . . . . .	187
10.1	Criteria and Data Sources . . . . .	201
B.1	Operational Effects of Freeway Lane Widths. . . . .	233
B.2	Operational Effects of Lane and Shoulder Width on Two-Lane Highways. . . . .	233
C.1	Themes, criteria and the explanation for use in the Via Baltica corridor study . . . . .	236
D.1	Zoning Categories and Base Zones for Cape Town. <i>Source: Adapted from City of Cape Town (2007)</i> . . . . .	239
F.1	Properties of Various Clustering Algorithms . . . . .	253

# List of Acronyms

<b>AASHTO</b>	American Association of State Highway and Transportation Officials
<b>AHTCG</b>	Ad Hoc Technical Committee on Geometries
<b>AIC</b>	Akaike Information Criterion
<b>ALbD</b>	Active Living by Design
<b>APA</b>	American Planning Association
<b>ARTISTS</b>	Arterial Streets for People
<b>BIC</b>	Bayesian Information Criterion
<b>BMJ</b>	British Medical Journal
<b>BPR</b>	Bureau of Public Roads
<b>BRT</b>	Bus Rapid Transit
<b>Caltrans</b>	California Department of Transportation
<b>CAN</b>	Cycling Academia Network
<b>CBD</b>	Central Business District
<b>CEA</b>	Cumulative Effects Assessment
<b>CI</b>	Confidence Interval
<b>CoCT</b>	City of Cape Town
<b>COLTO</b>	Committee of Land Transport Officials
<b>CSD</b>	Context Sensitive Design
<b>CSIR</b>	Council for Scientific and Industrial Research
<b>CSS</b>	Context Sensitive Solutions
<b>CUTA</b>	Committee of Urban Transport Authorities
<b>DoEAT</b>	Department of Environmental Affairs and Tourism
<b>EAN</b>	Equivalent Accident Number
<b>EU</b>	European Union
<b>FHWA</b>	Federal Highway Administration

<b>GIS</b>	Geographic Information System
<b>HCM</b>	Highway Capacity Manual
<b>HOV</b>	High Occupancy Vehicle
<b>ICE</b>	Interface for Cycling Expertise
<b>ITC</b>	Institute for Geo-Information and Earth Sciences
<b>ITE</b>	Institute of Traffic Engineers
<b>KPI</b>	Key Performance Indicator
<b>KYTC</b>	Kentucky Transportation Cabinet
<b>LOS</b>	Level of Service
<b>MADM</b>	Multiple Attribute Decision Making
<b>MAVT</b>	Multiple Attribute Value Theory
<b>MCA</b>	Multiple Criteria Assessment
<b>MCDM</b>	Multiple Criteria Decision Making
<b>MEC</b>	Member of the Executive Council
<b>MNL</b>	Multinomial Logit
<b>MODM</b>	Multiple Objective Decision Making
<b>MSA</b>	Moving South Africa
<b>NCSC</b>	National Complete Streets Coalition
<b>NDOT</b>	National Department of Transport
<b>NEMA</b>	National Environmental Management Act
<b>NEPA</b>	National Environmental Policy Act
<b>NHTSA</b>	National Highway Traffic Safety Administration
<b>NLTA</b>	National Land Transport Act
<b>NLTSF</b>	National Land Transport Strategic Framework
<b>NLTTA</b>	National Land Transport Transition Act
<b>NMT</b>	Non-Motorised Transportation
<b>OD</b>	Origin-Destination
<b>PAYE</b>	Pay As You Earn
<b>PM</b>	Particulate Matter
<b>PMT</b>	Private Motorised Transportation
<b>PT</b>	Public Transportation

<b>RSC</b>	Regional Services Council
<b>RTMC</b>	Road Traffic Management Corporation
<b>RWJF</b>	Robert Wood Johnson Foundation
<b>SANRAL</b>	South African National Roads Agency Limited
<b>SARTSM</b>	South African Road Traffic Signs Manual
<b>SDSS</b>	Spatial Decision Support System
<b>SMCA</b>	Spatial Multiple Criteria Assessment
<b>SOV</b>	Single Occupancy Vehicle
<b>SWOV</b>	Stichting Wetenschappelijk Onderzoek Verkeersveiligheid
<b>TAZ</b>	Travel Analysis Zone
<b>TDM</b>	Travel Demand Management
<b>TRB</b>	Transportation Research Board
<b>UCT</b>	University of Cape Town
<b>UE</b>	User Equilibrium
<b>UK</b>	United Kingdom
<b>USA</b>	United States of America
<b>UTG</b>	Urban Transport Guidelines
<b>VKT</b>	Vehicle Kilometres Travelled
<b>VMT</b>	Vehicle Miles Travelled



# Chapter 1

## Introduction

### 1.1 Road Planning, Sustainability and Context

Transportation is a critical component of everyday life and, as such, access to a good transportation system is a key factor to improving quality of life. The reality is, though, that most people are quite far removed from the planning of these systems. They rely on specialists to ensure that the transportation systems they use are efficient, safe and affordable. It has become increasingly apparent, though, that in many instances, the results have not always met these expectations.

The planning of transportation systems is often a politically sensitive topic. It plays such a pivotal role in peoples lives, and it affects the economies, societies and environments in which people live in such dramatic fashion, that politics is often the deciding factor when it comes to developing transportation systems (see Owens, 1995; Richardson, 2001). This is, albeit undoubtedly proper, also problematic, in that transportation planning is a highly technical discipline, the complexities of which have really only recently begun to be appreciated. These complexities mean that a thorough understanding of the field is required to appreciate the reasons motivating a particular decision, which may not necessarily align with what, at first glance, seems to be common sense.

In particular, the realization that, not only does transportation affect the economy, the society and the environment but that these, in turn, also affect transportation, has introduced large areas of unexplained complexity into the study of transport planning. It is no longer appropriate to impose a transport plan that other urban systems must conform to, because these types of plans inevitably disrupt these systems in unacceptable ways, leading to accidents, congestion, social and urban decay and environmental degradation. Transport planning must be sustainable, and to be sustainable, it must be sensitive to the complexities inherent in the urban systems it affects.

Pearson (1985) indicates that, “... *the core of the idea of sustainability is the concept that current decisions should not damage prospects for maintaining or improving living standards in the future.*” Another key aspect of sustainable development is that of social justice. The Charter of European Cities and Towns Towards Sustainability (European Conference on Sustainable Cities and Towns, 1994) (the ‘Aalborg Charter’), states that the objective of sustainable development is “... *to achieve social justice, sustainable economies, and environmental sustainability.*” Applying these definitions to transportation, it can be concluded that a sustainable transport system is one that is considerate of the needs of all (the social justice aspect) and that, in meeting these needs, does not place at risk the livelihoods or living standards of current and future generations.

This definition implies a vast range of possible factors that could impact on the sustainability of a transport system, many of which may not be easily quantifiable. This presents decision makers with a fundamental problem when assessing sustainability. To this end, much of the research effort around sustainable transport has focused on identifying suitable indicators for measuring sustainability (see Black et al., 2002; Litman, 2007; Richardson, 2005). Other methods, such as the Sustainable Livelihoods approach (Scoones, 1998), have attempted to interpret sustainability in terms of a life function - in this case, maintaining a livelihood.

In general, whatever the framework used to define sustainability, and whichever the indicators used to interpret it, these approaches all address transport planning at the network-wide level, that is, they are primarily concerned with the sustainable movement of people, goods and services. Transportation infrastructure, and roads in particular, must, however, do more than accommodate movement. Marshall (2004b) notes that streets are the locations for a variety of economic, civic, social and ceremonial activities that have very little to do with movement. These activities aside though, one of the primary functions of roads is to provide access to properties. This is, of course, the classic paradox of road planning, that many roads must fulfil two roles that are intrinsically in opposition to each other, as described in the influential *'Traffic in Towns'* by Buchanan (1963). That is, that on the one hand, roads are conduits of movement, and should ideally be planned to facilitate the fast and efficient movement of travellers along it, and on the other hand, that roads provide access to properties, a function that hinders the efficient flow of traffic.

This clash of purpose is often the cause of many of the problems encountered in the operation of road infrastructure. Accidents, congestion, increased noxious emissions and the decay of the urban and social environment<sup>1</sup> can all, in part, be attributed to the friction between through road users and access road users. Complicating matters is the fact that various users, for both mobility and access functions, all differ in terms of the activities they are conducting and the movements they want to make, and their ability to do so speedily and skillfully. Furthermore, nearly all roads must accommodate the needs of multiple modes of travel, each with its distinct operational characteristics and limitations.

Inevitably, conflicts arise, and it can be argued that, in order for any road infrastructure to be considered sustainable, it should be able to manage these conflicts sustainably. Road infrastructure provision that is biased towards the needs of motorised transport, has had a negative effect on the safety of Non-Motorised Transportation (NMT) modes. This is reflected in the very high proportion of pedestrian fatalities on South African roads (see Mabunda et al., 2008). Furthermore, this has had an influence on people's perceptions of the safety of these modes, resulting in an underutilisation of cycling. Walking is only chosen as mode by those who have no other alternative (the poor). The National Household Travel Survey, (NDoT, 2005a), lists concerns about road safety, due to conflicts between modes<sup>2</sup>, as being amongst the most significant deterrent factors for the use of NMT.

Managing these conflicts, then, in terms of the definition for sustainable transport derived previously, entails ensuring that the requirements, with respect to social justice are met and that no current or future livelihoods are negatively affected. Of course, conflict management is always an exercise in compromise. It may, therefore, be more appropriate to identify a balance of needs, something of a sustainable pareto optimum, where no road user can be made better off without making another worse off.

In summary, when assessing the sustainability of road infrastructure it is important to look at both the

<sup>1</sup> With reference to Appleyard (1980), who showed how, as traffic levels increased along residential streets, levels of social interaction amongst neighbours decreased, along with community cohesion.

<sup>2</sup> (see also City of Cape Town, 2005)

macro (network level) and micro (locale level) effects of the facility. Traditional planning techniques, however, are unable to describe these micro effects holistically. Road planning tends to focus on variables such as volumes, speeds and the operational geometric requirements of the various modes. Typically, land uses and socioeconomic variables are only important, in that by investigating these, planners are able to predict the probable number of trips attracted to, and generated from, an area through a combination of analytic and heuristic techniques, but these outcomes still feed only into the estimation of traffic volumes.

Although optimising operations to decrease congestion does contribute to improving the overall sustainability of a facility, the range of impacts that can be considered using traditional planning methods is very limited, especially below the neighbourhood level. However, significant variation in the indicators used to assess sustainability occurs at these micro levels, since a route's character can change within the distance of one or two neighbourhood blocks. It is, therefore, important that alternative planning methods be developed that can incorporate a suitable range of indicators to holistically assess the sustainability of road infrastructure at a much less aggregate level than is currently possible.

Typically, indicators for estimating sustainability fall into three broad categories: economic indicators, social indicators and environmental indicators (Gudmundsson, 2001; Litman and Burwell, 2006). As mentioned, transport planning traditionally considers aspects of land use and transportation or traffic to assess plans. A holistic approach would combine the traditional road planning indicators with those identified as suitable for measuring sustainability.

In recent years, CSD has been promoted by both the American Association of State Highway and Transportation Officials (AASHTO) and the Federal Highway Administration (FHWA) as a best practice (AASHTO, 2004; FHWA, 1997) to promote more sustainable road infrastructure. CSD provides a systematic and comprehensive approach to project development from inception and planning through operations and maintenance. Its goal is to achieve a project development process that provides an outcome harmonizing transportation requirements with community needs and values (Stamatiadis et al., 2009).

Complete streets initiatives are another example of CSD. The National Complete Streets Coalition (NCSC) highlight that Complete Streets are “... *designed and operated to enable safe access for all users*” (NCSC, 2010). The approach to implementing Complete Streets policies vary, but generally encompass the adoption of a Complete Streets policy, the generation of a set of guidelines or manuals, such as that published by the American Planning Association (APA) (McCann and Rynne, 2010) and, possibly, legislation to support the policy.

Current road and network design methods that rely almost exclusively on traffic information to recommend service levels and design parameters have been found to significantly impact mode choice (Cervero and Radisch, 1996) and vehicle miles travelled (Holtzclaw, 1990; Kitamura et al., 1997). The need for CSD arose, in part, from the realisation that the majority of urban streets serve multiple roles, having to accommodate the needs of multiple modes of transport, and needs related to mobility (through users) and access (local users). As a result, a certain amount of flexibility is required in design to meet all of these needs (see FHWA, 1997; Hebbert, 2005). As mentioned, roads often serve multiple functions in urban areas. This multiplicity of roles implies that the activities served by the road, and the needs of the people who will use it, must be thoroughly evaluated and understood before an appropriate planning recommendation can be made. Contextual factors, typically, relate to who is using the road, why they are using the road and how they can be expected to use the road.

The set of information required to answer these questions matches well with the set of information typically used to assess sustainability. Factors such as aspects related to the adjacent land uses, the

socioeconomic profile along the route, the environmental (ecological and cultural) landscape along the route, are equally as important as traffic and transportation factors when describing the context of a particular location along a road. It is evident that these contextual factors will vary spatially and temporally. Similarly, certain locations along a road may comply better with the definition sustainability, as previously derived, than others. This is especially true for long arterial routes.

## 1.2 Problem Statement and Research Questions

The original motivation for this research was to understand why, despite an improved legal and policy environment, road infrastructure provision in South Africa still tended to be, mainly, focused on the needs of automobiles. The belief was that planning and design practice had not fully embraced the ideals of policies, and that this played a significant role in the poor quality and piecemeal nature of the infrastructure provided for the other road-based modes.

Initially, investigations focused on local policies, and the design manuals that guide road planning and design practices, but soon moved to a deeper investigation of road planning theory, specifically to the influence and origins of road categorisation systems and their effect on road design. Various alternative approaches to road categorisation from around the world were investigated, and this led to some research around sustainable transport practices and how they compare with current road planning practices. It was from here that the central premise of this thesis originated.

Whilst acknowledging that current road planning and design practices are rooted in established, scientifically sound methods and research, a need was nonetheless identified to incorporate more contextual information into the transport planning process, to produce a more comprehensive, holistic and multimodal approach to practice.

A problem statement, or hypothesis, was developed as:

*Each mode used in a road has its own specific characteristics and needs, which determine the design parameters for that mode. Also, each location in a network, or along a road, is defined by a set of contextual parameters that determine how, and by whom, the road, at that location, is most often used. It is when these modal characteristics and location specific factors, or the needs of a mode and the use of a location align, that an ideal planning solution is found. Accordingly, certain modes are better suited to a particular set of contextual circumstances than others. Therefore, under a given mix of contextual circumstances, certain modes should be given a higher priority than the rest. This multimodal priority ranking can be used to inform road planning recommendations, using established practice and guidelines to inform decisions.*

The research described in this thesis, therefore, investigates whether it is possible to rethink the road planning process to incorporate a broader range of factors than is currently used so that a more sustainable, contextually sensitive road facility can be developed. Several subsidiary research questions, shown below, need to be answered to investigate the hypothesis:

1. How are roads currently planned and designed, what are the reasons for this approach, and what are the benefits and disadvantages of the current approach to road planning?
2. What are the legal, policy and regulatory frameworks that affect road planning practice, and to what extent do these influence the planning practice?
3. What is the role of context in the transport environment?

4. What are the factors that best describe the context of a route and to what extent do these factors currently play a role in road planning practices?
5. What would be an appropriate platform to incorporate multiple additional factors in the planning process?
6. What are the data requirements for a context analysis?
7. Which corridors would be best suited for developing and testing the method?
8. What are the crucial elements of the method that need to be developed and validated?
9. How could the results of a contextual analysis be used to make road planning recommendations?

### **1.3 Scope and Limitations**

The research was borne out of observations of South African road planning practices, and so is limited in scope to South Africa. However, the analysis method developed in this research is generic, and so, although this has not been explicitly tested, the techniques developed should be applicable anywhere in the world. Various case studies were conducted, the results of which are presented in Chapter 7 of this thesis. All of the case study roads are in Cape Town in the Western Cape, South Africa.

The analysis is data intensive, with a large range of data sources being drawn upon to complete the various stages of the method. Data availability was, therefore, a significant limitation in the project, and sufficient funds and time was only available to test the method on one data set. This was also the primary reason for limiting the case study roads to roads in Cape Town. However, these case studies were selected based upon the variation in context along the route, and so there is enough variation within the case studies to demonstrate the validity of the analysis.

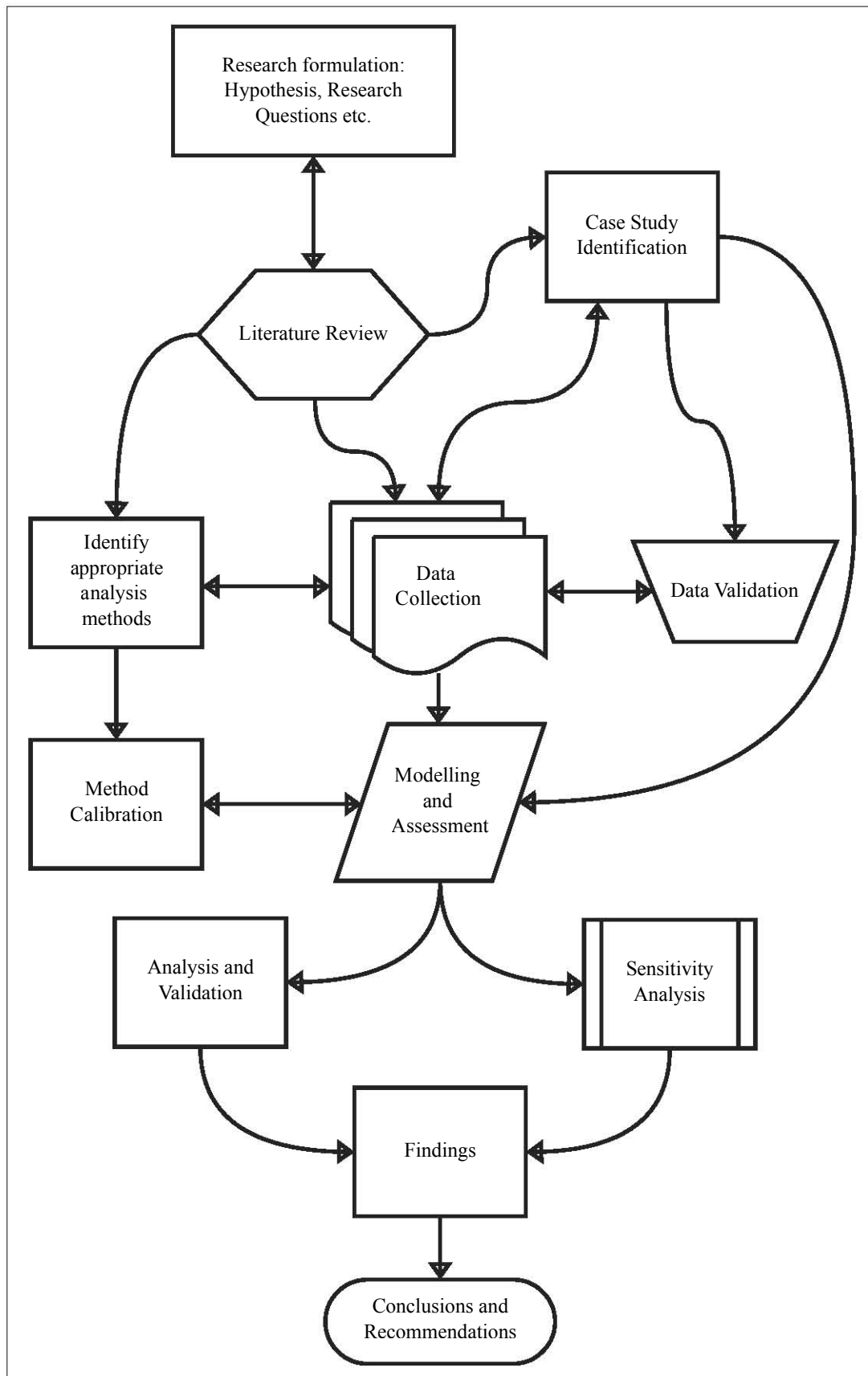
### **1.4 Research Approach and Methodology**

The research undertaken for this thesis consisted of a number of processes, often conducted in parallel. Figure 1.1 shows an overview of the research methodology used.

The first goal of the research was to formulate the research problem by developing a hypothesis and deriving research questions. This process was supported by a preliminary literature review on road planning and design, and discussions with practitioners and academics. An expanded literature review was then done, and was used to identify the required data sources, the selection criteria for suitable case study roads, and the various tools and techniques that could be used for the analysis and development of the method arising from the research.

The data collected for use in the analysis needed to be validated and optimised for use in the analysis. A range of methods was used, including in-situ verification in the field, supported by satellite image analysis and various statistical methods. The data that was collected also, to a certain extent, informed the choice of analysis method, since the techniques selected for use were dependent on the data availability. Data was collected for three case study roads in Cape Town. One case study was used to develop the methodology and to calibrate the various input parameters, and the remaining two case studies were used to validate the method.

The results were analysed in relation to the data collected from field investigations and satellite imagery, and a sensitivity analysis was conducted. This information was also used to calibrate and validate the approach. The outcomes of the analysis were used to develop the method proposed in the research, and to draw conclusions and make recommendations.



**Figure 1.1:** Overview of research method

## 1.5 Outline of the Thesis

This thesis consists of ten chapters detailing work that has been carried out over the course of a three year research period started in 2008.

Chapter 1 contains the introduction to the thesis including some background to the research, a description of the research problem and a listing of the research questions. A section on the limitations and scope of the research is then followed by a brief discussion on the research methodology and an outline of the document.

In Chapter 2, road planning practice in South Africa, and elsewhere, is investigated. The chapter opens with a history of road planning, and then examines the current South African transportation legislation and policy frameworks and their implications for practice. The chapter concludes with a discussion of local design guides and manuals and their influence on practice.

Chapter 3 details an overview of the state of road safety in South Africa, the plight of vulnerable road users, and a discussion on the effectiveness of existing methods to improve road safety.

Chapter 4 focusses the importance of context. The chapter opens with a review of the work conducted in the fields of CSD and the Complete Streets Movement. This is followed by an investigation into the findings of the ARTISTS project. A description of the term ‘context’ is then formulated for use in the research conducted for this thesis. This includes an investigation into the criteria that can be used to assess the context.

Chapter 5 includes a review on decision support systems and their role in planning. The chapter opens with an introduction to multiple criteria evaluation techniques and methods with a specific focus on spatial multiple criteria evaluation. This is followed by a discussion on the use of SMCA in this research which includes the details of the adaptations made to address the specific requirements of using SMCA in this application.

Chapter 6 describes the data used in the study, the case study selection rationale and an overview of the case studies themselves. The chapter opens with a description of the data used in the analysis and the data sources. This is followed by a discussion on the data verification and optimisation processes done on each data set. The chapter concludes with discussions of the three arterial routes used as case studies.

In Chapter 7 the results of the SMCA analysis on the case study roads are analysed and discussed. The chapter also includes an investigation into the sensitivity of the results to alternative weighting schemes.

In Chapter 8 the evaluation results are subjected to a clustering analysis. The chapter provides a general overview of clustering theory, specifically focusing on the clustering methods used in this research.

Chapter 9 includes a discussion on translating the results of the analysis into road. The chapter opens with an analysis of the clustering results before describing the methodology proposed for developing infrastructure recommendations.

Chapter 10 contains the concluding remarks. The chapter summarises the research findings and the results of the analyses. The contributions to knowledge made by the research are highlighted and recommendations for suggested future work and a discussion on the remaining open problems concludes the thesis.



## Chapter 2

# Planning, Policy and Practice

### 2.1 Transport Planning - Past and Present

Historically, streets have developed to accommodate all types of human activities. Aside from their role as being the places where people, animals and carts moved between buildings, streets were also places where people socialised, met with friends and relatives and where merchants traded goods or rendered services. They were the places that children played and travellers rested, the main venues where communities celebrated and mourned and where civic debates and protests were held. Today, in many places around the world, streets still perform many of these functions. However, in modern times, many streets have been given over to serving only one function: that of accommodating automobiles. In this section of the text the origins of modern day planning and the role of the automobile in shaping towns and cities today, is explored.

#### 2.1.1 The Origins of Transport Planning

Despite the automobile being such an integral aspect of city life, it is a relatively new addition to the urban milieu. According to Clark (1998), people have been living in urban settlements for over eight millennia. The automobile, however, has only been in general use for the last one hundred years. Since its invention, it has become a dominant influence on the planning of urban areas. It has allowed people to live further from their places of work, where land prices are lower and they could afford the larger and better appointed homes idealised by the proponents of the Garden City Movement (Moughtin, 1991).

Although the automobile facilitated the emergence of suburbia in the twentieth century, the process of suburbanization was driven by the desire of many urban dwellers to retreat to the tranquility of the countryside, or put differently, to escape the unpleasant living conditions in cities created by industrialisation. The high densities of nineteenth century cities were seen as being a primary contributory cause for many of the social ills that beset urban communities of the time. Boyer (1983) notes that, towards the end of the nineteenth century, a movement for urban improvement was developing in American cities rooted in annoyance at the 'ugly and chaotic city conditions' and the belief that urban areas were unnatural and unhealthy living spaces for human beings.

Frost (1991) explains that, with the rapid advances in medicine and agriculture in Europe during the eighteenth and nineteenth centuries, and growing revenues from intercontinental trade, came a rapid increase in population. These people, mostly living in rural areas at the time, faced with a life of rural poverty, often elected to seek opportunities in the rapidly growing towns and cities. As these cities

grew with the influx of rural migrants, satellite suburbs developed on their peripheries. The streets that were created in these new peripheral areas were often simply extensions to the existing street network, developing organically as settlements grew. Very little formal planning went into their development (Clark, 2003).

Taylor (2002) notes that population growth, coupled with the increasing industrialisation and transportation improvements in cities towards the end of the eighteenth and the beginning of the nineteenth centuries, meant that streets were increasingly unable to cope with the demands placed upon them. Of all the functions of streets at the time, two in particular were seen as being cause for concern.

The first function of the street that was seen as being problematic was the streets role in promoting public health and, in particular, the disposal of sewage. Rethinking the medieval street, which was generally laid to fall to a central channel (as shown in Figure 2.1), was seen as vital in the battle against the scourge of the day, cholera. As a result, the older central channel designs were replaced by a doubly cambered surface, which drained to channels on either side of the road. The footways became separated from the carriageway by the channels and raised using kerb stones. This provided pedestrians with a suitable level of separation from the drainage water (Peterson, 1979).



**Figure 2.1:** Central channel street in Victorian London.

*Source:* BBC News, UK, England: The changing face of London. Retrieved September 20, 2010, from <http://news.bbc.co.uk>

Additionally, urban mortality in the mid nineteenth century was seen as being unacceptably high, and one of the solutions put forward was to build wide and straight streets that would bring fresh air and light into slum areas. As a result, standards for housing were developed to give minimum road widths of forty to fifty feet for new developments (Hebbert, 1999). The effect of these measures, along with the systematic closure of old slums, contributed to greatly improved life expectancies for town dwellers (see Wohl, 1983, page 329).

It is interesting to note that many of the geometric elements of streets, originally developed to address the practical requirement of controlling issues of health, are still influencing the way streets are designed and built in the twenty-first century, albeit for very different reasons. For instance, the kerbstone, originally intended as a containment device for water, has become a medium for controlling movement.

### 2.1.2 Transport Planning and the Rise of the Motor Car

The second purpose of streets, which was seen as being problematic, was the streets role as a conduit of movement. Originally this was primarily by foot, by horse or by carriage, but from the early twentieth century became increasingly motorised. During the nineteenth century concerns around traffic centered mostly on the problems presented by the huge number of horses required to ferry people and goods around, and the consequent hygiene and disease problems presented by having so many horses plying the streets (see Morris, 2007).

The spread of the automobile, however, added a new dimension to the traffic problems experienced in cities, that of speed and, consequently, of safety. Discussions around speed limits and appropriate driving speeds on different roads were taking place as early as 1906 (see BMJ Publishing Group, 1906). Although the dawn of the motor car era was heralded by many for ridding cities of the unpleasantness inherent to horse transport, as early as 1913 commentators were voicing their concerns regarding the road safety problems presented by cars (see Crum, 1913). The post war years saw the first organised efforts to address these problems. The solutions that were developed have come to shape urban areas in such fundamental ways that they define, to a large extent, how people conduct their lives in cities.

Goodwin (1999) points out that, in the UK, since the end of the second world war, there have been two major lines of argument on how to manage the car and the infrastructure required to accommodate it. The first view has been to accept the growth of automobile use as inevitable and provide the road capacity necessary to accommodate it (leading to the so-called predict and provide approach to transport planning). The other has been to control or moderate its use so that it is kept within the scope of broader traffic or societal objectives. Both Tripp (1942) and Buchanan (1961) argued for the latter approach, but the dominant argument has been that it is necessary to provide road capacity to match the traffic levels.

One of the principle ideas that developed at the time was that of segregation of function, in particular the segregation of vehicles and pedestrians (Eden, 1947; Rockey, 1983). With the emphasis on the fast movement of cars, safety was a prominent consideration. The priority, however, was not given to people on the street, but to the smooth and rapid flow of traffic (Hass-Klau, 1993). Planning horizons were decades long, a result of the massive levels of investment required to complete projects. Planners were, therefore, forced to envision future operational conditions for a rapidly changing world. They relied on past trends showing a rapid uptake of automobiles, and then built in robust margins of error to account for any unforeseen developments.

Solesbury and Townsend (1970) provide an overview of the major transportation studies conducted in metropolitan centers in the UK in the sixties. All of the nine studies they investigated employed variations on the familiar four-step modeling procedure to identify existing road capacities, to predict future travel demand, and then to make recommendations on infrastructure improvements to accommodate the predicted traffic growth.

The methods used in these studies were not without critics. Hillman et al. (1973) and Plowden (1972), and in the US, Schaeffer and Sclar (1980), described them as 'self-fulfilling'. They argued that, even when household car ownership was high, it was not universal, and that the people who lost most from the recommendations made from these kinds of studies, were precisely those who already had the greatest travel problems, namely children, the elderly, the poor, and women. Schaeffer and Sclar noted that child development psychologists, before the 1950s, were able to use 'independent travel' (by bus) as a key measure of an eight to fourteen year olds ability to orient and master space. However, as bus services disappeared from suburbs, even the affluent young could only leave their immediate

neighbourhood if someone drove them, and independent travel was delayed until they were old enough to get a driving license and access to a car.

Transport planning today still follows many of the same procedures to make recommendations. Vuchic (1999) argues that transportation planning should rather rely upon the use of a set of incentives and disincentives to either increase or decrease the cost of travel for each mode, in order to achieve a predetermined objective that is aimed at minimising the total cost that travel represents to society.

At the heart of transport planning, as it is practiced today, lies the idea that motorised traffic is best kept separate, and that its interface with people can be controlled by a hierarchy of roads defined by a road classification system (Forbes, 1999; O'Flaherty, 1997). The traditional functional classification system, developed in the UK and the USA, can be traced back to the '*Radburn Layout*', developed in the USA in the late 1920s for the unincorporated new town of Radburn in Fair Lawn, New Jersey (Figure A.1<sup>3</sup>) by architects Clarence Stein and Henry Wright (Lillebye, 1996). One of the major features of the plan was a hierarchical road network, developed to separate pedestrians and motor vehicles. All fast-moving traffic was restricted to feeder roads, and the houses were clustered on cul-de-sacs with one side accessible from the street, while the other side opened onto communal gardens that had pathways leading to a central park.

The concept was formalised in 1963, when the report, '*Traffic in Towns*', was published (Buchanan, 1963). This report contained a comprehensive vision for the design of towns and villages in a highly motorised society. A distinction was made between roads having a traffic flow function, and roads that give access to destinations. The elaboration of these ideas resulted in a proposal for a route hierarchy, built up from primary, district and local distributors and access roads to destinations (Figure 2.2). Buchanan argued that, within access roads, traffic should be of minor importance relative to the environment (which relates to his concept of the 'environmental area') and, in every area, at least the maximum acceptable traffic capacity had to be determined.

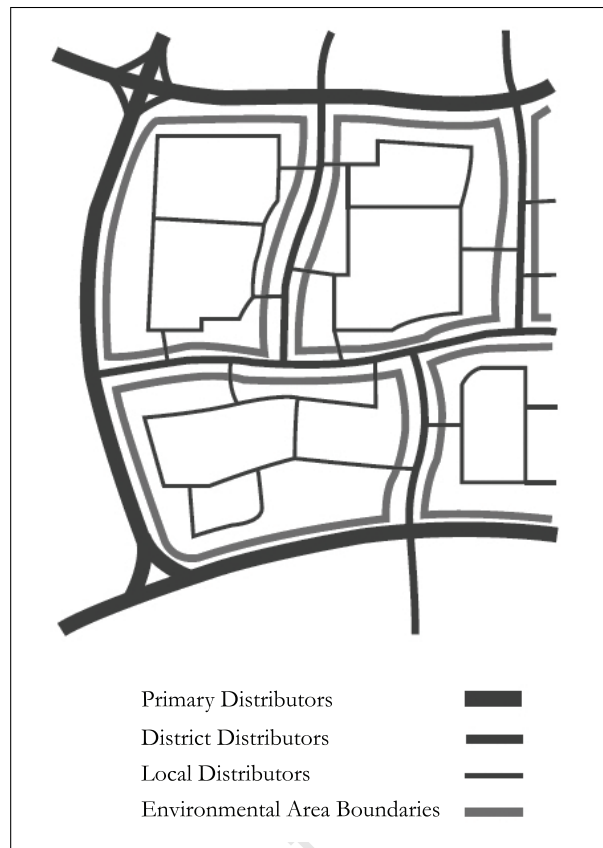
These ideas were widely adopted, and today play a critical role in transport planning all over the world. Marshall (2004a) notes that road hierarchy has been an influential and often dominating factor in determining the character of modern urban layouts. In *Streets and Patterns* (Marshall, 2005), he argues that road hierarchy has often been criticised for resulting in dull or dysfunctional road-dominated layouts lacking in 'urbanity or sense of place', but that this has more to do with the inappropriate application of the concepts behind road hierarchies than with Buchanan's vision of the environmental area.

Marshall's point is illustrated by the approach to road hierarchy described in the American Association of State Highway and Transportation Officials (AASHTO) '*A Policy on the Geometric Design of Highways and Streets*' (AASHTO, 2004). This document describes the relationship between 'mobility' and 'access' as an inversely proportional sinusoidal curve (see Figure 2.3). In this conceptualization, all streets fall somewhere along a continuum between mobility only streets and access only streets. This conceptual view of streets defines the role that a street plays in the network, and forms the theoretical basis for the hierarchy developed within this paradigm, and that is now used extensively all over the world.

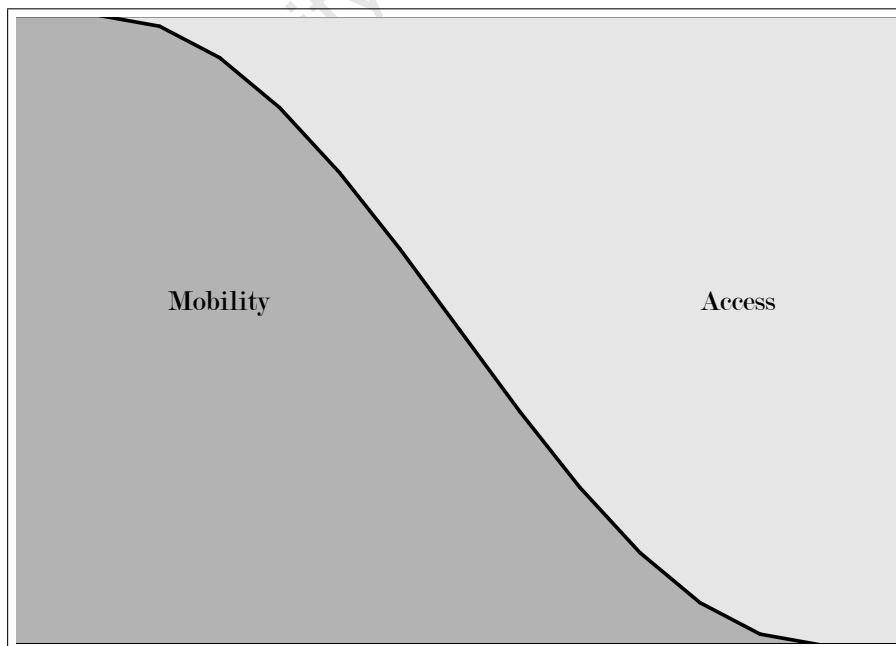
### 2.1.3 Transport Planning and the Provision of Infrastructure

The extrapolation of Buchanan's view of road hierarchy into a continuum of function is problematic in so far as streets, that do not fall neatly along this curve, cannot be accommodated by the classification

<sup>3</sup> All references denoted with an alpha-numeric designation can be found in the Appendices to this thesis.



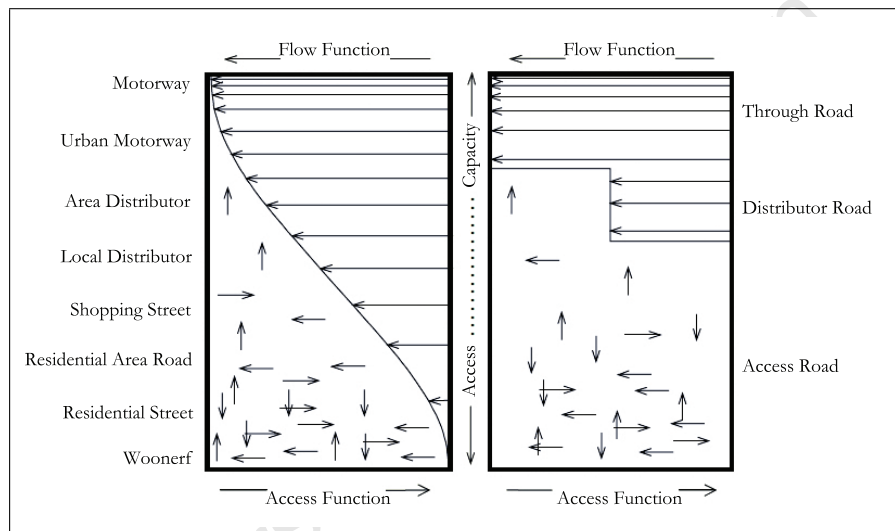
**Figure 2.2:** Functional categorisation of roads according to Buchanan.  
 Source: Buchanan (1963)



**Figure 2.3:** The relationship between 'mobility' and 'access'.  
 Source: AASHTO, 2004

system. Furthermore, although the function of the street and, consequently, the infrastructure that is required to serve that function, is clear towards the extremities of the curve, many streets fall somewhere towards the center of the curve, where the precise implications of the functional relationship is harder to pin down.

More recently, in an attempt to address these concerns, various alternate road categorisation schemes have been developed. Notable amongst these is the concept of ‘Sustainable Safety’ (SWOV, 2005), developed in The Netherlands, which asserts that a street can only have one function, either mobility, or access. This implies that, where a street is designated as a mobility route, all other street functions are actively suppressed, and where it is an access route, mobility is given low priority. To ease the transition between the two extremes, a third ‘distributor’ category has been introduced (Figure 2.4). The method also advocates the importance of the complete separation of modes along mobility routes, emphasizing that each mode receive the highest LOS practical in any location.



**Figure 2.4:** Traditional functional classification system (left) compared to the tri-partition system used in Sustainable Safety (right).

Source: SWOV, 2005

The sustainable safety method is interesting in that it emphasises the role of engineering design in crash avoidance. It takes the position that a large proportion of accidents that are otherwise attributed to some form of driver error, could actually have been made less likely to occur, or avoided completely, with more careful design.

The extent to which road categorisation influences the provision of transport infrastructure cannot be understated. Road classification, as it is conventionally applied, takes much more from its theoretical underpinnings than is required to produce a generic description of the facility. Instead, it uses this theory as the basis for a more prescriptive interpretation of the function of the road. Conventionally, it concentrates on the traffic functions of streets, and the relationship between the categories developed from the scheme in terms of mobility on the network. The categorisation scheme, therefore, generates a hierarchy of levels of mobility that can be reinterpreted as parameters that define the infrastructural norms for that facility type. The nature of the categorisation system, and its interpretation in terms of mobility, means that the primary parameters that are developed are operating speeds and target volumes. The hierarchical nature of the system lends itself to restrictive norms in terms of which level of road

can link to which and, also, by implication, what quality of service must be targeted for which modes along which links.

The approach dovetails very well with other practices in transportation planning, and this is arguably not coincidental, since many of these ideas were being formalised around the same time. Notably, the concepts behind, and the implementation of, road categorisation work very well with the concept of Level of Service (LOS), first formally published in the USA in the 1965 edition of the Highway Capacity Manual (HCM) (Highway Research Board, 1965), soon after the publication in the UK of the Buchanan report, '*Traffic in Towns*'.

LOS provides planners with a way to describe the quality or efficiency of operations on a roadway or a road facility (such as an intersection) as one of six increasing ranks, 'A' through 'F', with 'A' being best. It can also be used to describe the desired operating conditions of a given facility, and is used in this way in most parts of the world. Transport authorities, using LOS, can set benchmarks for the efficiency of operations on their networks, and can then use these benchmarks to justify the need for modifications or upgrades to the facilities in their jurisdiction.

Over the years, the analytical methods behind LOS have been refined and expanded to include much more than the originally intended use, which was to succinctly describe the efficiency of operations on a section of freeway. Lower order roads, signalised and unsignalised intersections, sidewalks and bicycle lanes have all, with time, come under the ambit of the methodology. However, this has not happened without a fair amount of criticism.

Some of the common criticisms levelled at the use LOS in transportation planning are listed below:

1. LOS Reflects the convenience of motor vehicle travel measured as either free flow speed or time delay. The higher the free flow speed or the lower the time delay, the higher the LOS provided. Improving service levels is thus limited to implementing measures that impact operating speeds or traffic flow. However, this is often only achieved at the expense of other road users, or to the detriment of the neighbourhood within which the facility is located.
2. Research has shown that increasing roadway capacity is not always an effective long-term strategy to manage roadway congestion, since improvements in LOS are often offset by increased traffic flow induced as a result of the lowered overall cost of travel along the upgraded route (Cervero, 2001; Fröhlich, 2003; Nielsen and Hovgesen, 2004; Noland, 2001).
3. LOS analysis does not account for the modal shift induced when reduced motor vehicle capacity encourages auto trips to shift to other travel times, routes or modes of transport. Reductions in LOS can lead to changes in the route of a journey, changes in the time at which trips are made, changes in the mode of travel, changes in the frequency of travel, changes in the destination of travel, trip elimination, and consolidation of trips to serve several destinations with one journey (Cairns et al., 2002).
4. It is not clear what the effect of improved LOS, in the form of increased speeds and volumes, implies for vehicle emissions and fuel consumption. Although there is evidence to suggest that improved service levels result in lowered emissions and fuel consumption (Cobian et al., 2009; Diana et al., 2007; Unal et al., 2003), these results do not account for the effects of induced demand. The research also suggests that, although marked improvements are noticed when transitioning between lower LOS's, the effects are less significant between higher LOS's.
5. Previous versions of the method did not take into account relationships and conflicts among modes. For example, higher traffic speeds, higher flows, broader roadways, and reduced lateral separation reduce the LOS for pedestrians. In fact, the methodologies used to assess pedestrian,

bicycle and, to some extent, transit LOS were flawed in that the parameters used to assess these modes did not adequately reflect the unique needs or operational requirements for these modes (Florida Department of Transportation, 2009). The methods used to assess service levels for these modes have, recently, been dramatically improved (see Dowling et al., 2008), but the new methodologies have not yet been widely adopted and will, in any event, need to be calibrated to local conditions before they can be used internationally.

6. Some authors have suggested that the number of service levels available (six according to the HCM methodology) does not adequately reflect the range of conditions encountered in reality and needs to be expanded (Brilon, 2000; Cameron, 1996; Maitra et al., 1999). Other authors have countered that to the average road user and, for that matter, transport planner, there is very little perceptible difference between adjacent service levels, and so the grading should, instead, reflect the users perceptions of LOS (Choocharukul, 2004; Nakamura et al., 2000; Pfefer, 1999).

#### **2.1.4 The Challenges of Progress**

The concerns around transport planning and the effects it was having on society, the economy and the environment, led to a series of policy and legislative changes, particularly in the USA during the 1990s, that had important implications for practice (Kane and Del Mistro, 2003). The so called 'predict and provide' methodologies and policies that informed earlier schools of planning thought were increasingly being challenged by alternate policy approaches, collectively referred to as 'new realism' in the UK, endorsing travel pricing and integrated land use and transport planning as being more appropriate (Owens, 1995).

Van Exel and Rietveld (2009) call attention to the fact that the new focus on managing demand and reducing the need for travel has resulted in the need for planners to understand the motivations behind peoples travel choices becoming more pressing. A substantial literature has developed over the last twenty years dealing with this precise issue. It seems, however, that there are significant caveats to the successful implementation of the new policies.

Owens (1995) remarks that, often, the ideals of policy and the realities around the steps to be taken to implement them, are at odds with each other. Simply put, price increases in the form of additional taxation or land use measures, such as densification, are often unpalatable to public and political sentiments, despite the general consensus that current trends in transportation are unsustainable. Walton and Shaw (2003) assessed a new project appraisal approach that attempted to implement the principles of the 'new realism'. They mention that after initial enthusiasm for the principles enshrined in the policy, the UK government was increasingly forced into a position they term, pragmatic multimodalism, once public sentiment in favor of the approach started to wane.

Implementing these new policy approaches necessitates a re-envisioning of societal and environmental priorities, that will exact a toll on the average middle class car user and that will, therefore, require substantial political will to see through. In societies such as those in the UK and the USA, where the majority of the voter base are heavily reliant on car use, the political will to implement such policies may not be forthcoming. In countries like South Africa, where the majority of voters (especially of those who are likely to vote for the current ruling party) are captive to PT, you would expect to see these policies being implemented post-haste, but, at least in South Africa, this is not really the case.

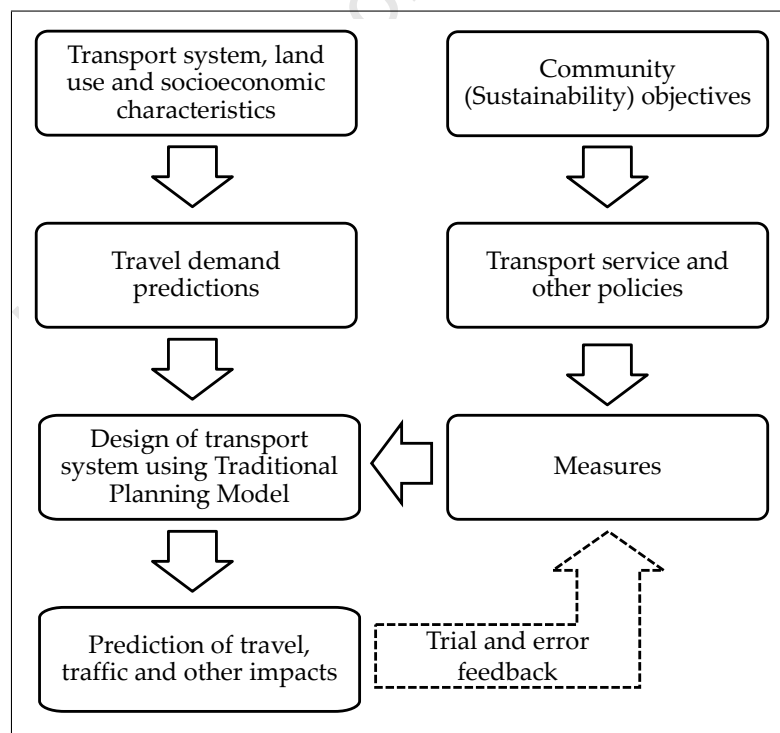
## 2.2 Transport Modelling

Although it was developed over 50 years ago, the traditional transport model, also called the four-step model, still dominates much of the work done in traffic and transportation planning. The various steps in the model may have seen a number of incremental improvements over the years, but the process, as a whole, is still the de facto standard for modelling transportation flows.

The current approach to analytical transport planning and modelling was developed during the 1950s and 1960s, when the first Chicago and Detroit transport studies were performed in the USA (Black, 1990). The methodology developed at the time can be characterised as being reactive, in that increases in road and PT capacities were proposed to accommodate expected future increases in volumes calculated, based upon estimates of existing vehicle and passenger volumes. For this reason, the methodology is often referred to as the ‘predict and provide’ approach.

Whereas the early models were, typically, unimodal, modelling only PMT trips, Bates (2000) notes that transport models have since evolved to accommodate multiple modes. Transport modelling has derived its theoretical basis from, primarily, economic theory. Increased computing power, and improved data collection and estimation techniques, has allowed the models to become more detailed and to be applied at larger scales. The traditional transport modelling process is graphically depicted in Figure 2.5. Projections of travel demand, that are based on predicted transport system and traffic characteristics, socioeconomic development and land use planning which are, in turn, derived from a transport policy objective, form the basis for the design of the transport system using the traditional transport planning model (Zuidgeest, 2005).

Typically, transport systems are depicted as networks, constructed from nodes (representing cities,



**Figure 2.5:** Traditional ‘predict and provide’ transport modelling process.

*Source:* Adapted from Zuidgeest (2005).

zones, intersections etc.), which are joined together by capacity-constrained links (representing roads, paths, railways, etc.). Each zone is represented by a zonal centroid, which consists of origins and destinations. Trip origins and destinations are located at the centroids, which generate the traffic flows along the links between the zones in the network. The physical, demographic and socioeconomic variables needed to define the system of activities are assigned as additional attributes of these centroids. Furthermore, a mode-set for vehicles as well as origin-destination specific route-sets for routes are defined (Zuidgeest, 2005).

Conventionally, the model is divided into four sequentially linked sub-models, which Zuidgeest describes as follows:

**Trip generation**, which is an estimate of the number of trips associated with a zone, and consists of the trips produced and the trips attracted to that zone;

**Trip distribution**, which is an estimate of the allocation of trips between each pair of zones in the study area, resulting in the production of an origin - destination (OD) trip table;

**Modal split**, which is a determination of the number of trips by each mode of transport between each pair of zones;

**Trip assignment**, which allocates all trips by origin and destination zone to the links that comprise the road network. Separate allocations normally take place for each mode. Trips are usually converted to vehicle trips using an average vehicle occupancy rate. Hence, this sub-model is better named 'traffic assignment'.

The first three sub-models deal with the calculation of travel demand (in person trips), whereas the fourth sub-model converts this demand to vehicle trips on the road network. Zuidgeest (2005), summarising Bates (2000), gives a thorough description of the first three sub-models, which forms the basis for the model summary presented here. The modelling of travel demand implies a procedure for predicting the travel decisions people would tend to make, given the generalised cost of all the alternative travel options available to them. The decisions include route (trip assignment), mode choice (modal split), destination (trip distribution), and frequency (trip generation). Often, the choice of time of travel is also added. These choices can be linked together using choice hierarchies, by using discrete-choice theory (see also Ben-Akiva and Lerman (1985)). Lower level choices are made conditional to higher level choices in a theoretically consistent way. A schematic of this choice-hierarchy is shown in Figure 2.6 (Zuidgeest, 2005).

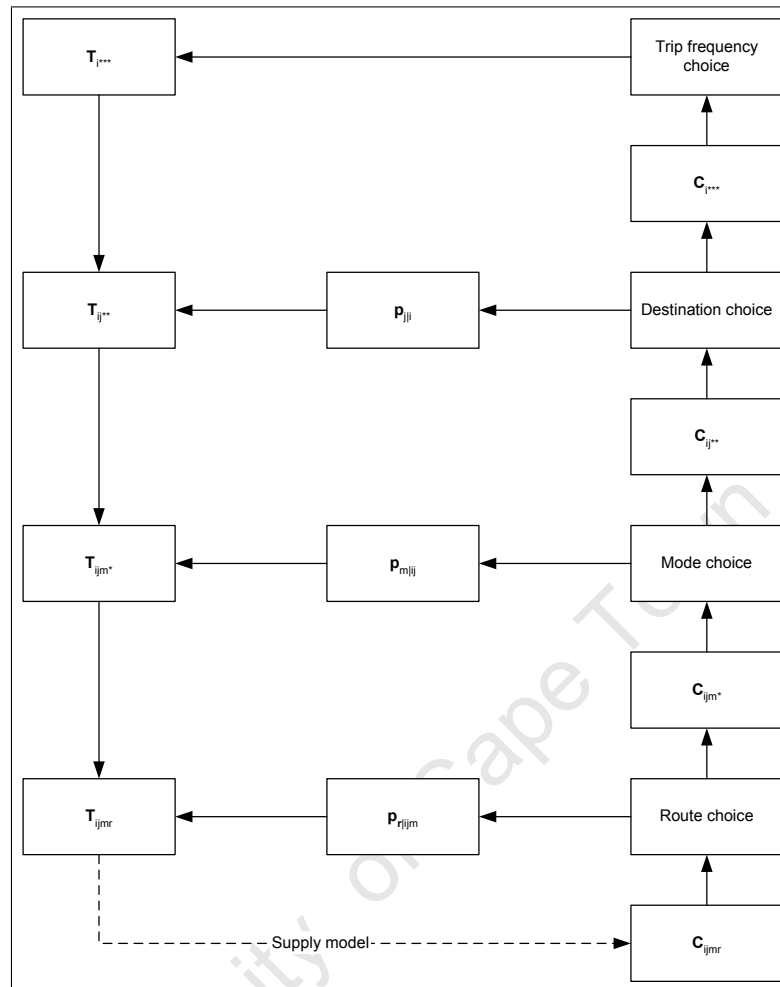
### 2.2.1 Trip Generation

The first sub-model, the trip generation model, can be split into a trip production and a trip attraction model. In general, the trip production model can be estimated quite accurately, using:

$$T_{i|k} = f(x_{1i|k}, x_{2i|k}, \dots, x_{ni|k}; [c_{i|k}^*]) = f(\vec{x}_{i|k}; [c_{i|k}^*]), \quad (2.1)$$

where  $k$  is the population segment (usually related to trip purpose and socioeconomic background),  $i$  is the origin zone in the study-area, while the vector:  $\vec{x}_{i|k} = x_{1i|k}, x_{2i|k}, \dots, x_{ni|k}$ , represents  $n$  socioeconomic characteristics for population segment  $k$  in zone  $i$ , and  $c_{i|k}^*$  is the generalised, or composite cost, of travelling from the origin zone  $i$ .

Zuidgeest notes that the composite cost of travelling (here presented between brackets) is rarely used in practice - trip production is usually treated as being only dependent on variables that are exogenous



**Figure 2.6:** The traditional transport model.

Source: Adapted from Bates (2000).

to the model. However, it is very conceivable that the number of trips generated at a zone is influenced by the transport system to and from that zone (trip generation is related to accessibility).

The trip production equation is used to obtain the total number of trip-ends in the zones of study. Most trip production models are household or person-based, implying that a zonal aggregation has to be performed, generally, by using information on the total number of households per segmentation  $k$  in zone  $i$ ,  $H_{i|k}$ , that is:

$$T_i = a_i^0 + \sum_k H_{i|k} f(\vec{x}_{i|k}; [c_{i|k}^*]), \quad (2.2)$$

where  $a_i^0$  is an intercept, usually resulting from the model estimation, e.g. from applying multiple linear regression techniques.

Trip attraction models for  $T_j$  can have a similar structure as in Equation (2.1), where the explanatory variables for the different attractions in zone  $j$  are related to the type of land uses attracting the produced

trips. However, it is actually rather complicated to find reliable estimates for these variables. So, since trip productions can be more accurately determined, trip attractions are usually balanced against the total number of trip productions. To balance the total number of trip productions and trip attractions, a balancing factor,  $f$ , is applied to all trip attractions  $T'_j$ , to obtain the balanced  $T_j$ :

$$f = \frac{T}{\sum_{j \in \mathcal{J}} T'_j}, \quad (2.3)$$

with the total number of trips being:  $T = \sum_{i \in \mathcal{I}} T_i$ .

### 2.2.2 Trip Distribution

In the trip distribution model, an OD - table, matrix  $T_{ij}$ , relates the number of trips in the matrix cell  $(i, j)$  to the characteristics of the origin/production zone  $i$ , the characteristics of the destination/attraction zone  $j$  and the characteristics of the generalised cost of travel, between zones  $i$  and  $j$ . This relationship is named the gravity model due to its similarity to the Newtonian law of gravitation, and has the general form:

$$T_{ij} = \mu Q_i X_j f(c_{ij}), \quad (2.4)$$

where  $Q_i$  is the production potential of zone  $i$ ,  $X_j$  the attraction potential for zone  $j$ ,  $\mu$  the 'gravity' constant<sup>4</sup> and  $f(c_{ij})$  is a distribution function<sup>5</sup>. Therefore, if the only deterrent to travel is assumed to be the travel distance between zones  $i$  and  $j$ , indicated as  $d_{ij}$ ,  $f(c_{ij})$  takes the form:

$$f(c_{ij}) = \frac{1}{d_{ij}^2}, \quad (2.5)$$

The distribution function can also be used in an exponential form, with parameter  $\lambda$  relating it to the generalised costs of travel  $c_{ij}$  between zones  $i$  and  $j$ , such that:

$$f(c_{ij}) = \exp(-\lambda c_{ij}). \quad (2.6)$$

The exact calculation of  $T_{ij}$  usually assumes that trip productions and/or trip attractions from the trip generation sub-model are known. In the production constrained trip distribution model, the number of trip productions  $T_i$  are imposed as a set of constraints:  $\sum_j T_{ij} = T_i$ , on the general trip distribution model, after some calculations giving:

$$T_{ij} = T_i \frac{X_j f(c_{ij})}{\sum_{j' \in \mathcal{J}} X_{j'} f(c_{ij'})}. \quad (2.7)$$

<sup>4</sup> Parameter  $\mu$  is interpreted here as a measure of average trip intensity in an area, being the number of travellers  $P$  divided by the number of and variability in trip alternatives  $k$ , i.e.:  $\mu = \frac{P}{k}$  (Bovy and Van der Zijpp, 1999).

<sup>5</sup> Also called deterrence, or impedance function.

The production constrained trip distribution model is, therefore, a proportional model, that splits the known trip production numbers in proportion to the attraction potential  $X_j$ . An attraction constrained trip distribution model can also be constructed using a similar methodology.

If both the number of trip productions and trip attractions are known, the calculation of  $T_{ij}$  is usually performed as an iterative process, known as the *Furness* method. In the method, two sets of constraints are given, namely, the numbers of arrivals:  $\sum_i T_{ij} = T_j$ , and departures:  $\sum_j T_{ij} = T_i$ , as well as two balancing parameters  $a_i$  and  $b_j$ , changing Equation (2.4) to read:

$$T_{ij} = a_i T_i b_j T_j f(c_{ij}). \quad (2.8)$$

In Figure 2.6 this destination-choice model is depicted as a conditional probability:

$$p_{j|i:k} = f(c_{ij|k}^*, c_{i\{j\}|k} : \vec{x}_{i|k}, \vec{z}_j), \quad (2.9)$$

where  $k$  is a segment of the population,  $i$  and  $j$  are the origin and destination zones,  $p_{j|i:k}$  is the proportion of all travellers of type  $k$  from zone  $i$ , who travel to zone  $j$ ,  $c_{ij|k}^*$  is the composite cost of travel between zones  $i$  and  $j$ , and  $c_{i\{j\}|k}$  is the associated cost of travel to all zones, with  $\{j\}$  the set of possible destination zones  $\mathcal{J}$ .  $\vec{x}_{i|k} = x_{1i|k}, x_{2i|k}, \dots, x_{ni|k}$ , is the vector of  $n$  socioeconomic characteristics for segmentation  $k$ , while:  $\vec{z}_j = z_{1j}, z_{2j}, \dots, z_{n'j}$ , is a vector of  $n'$  zonal characteristics.

### 2.2.3 Modal Split

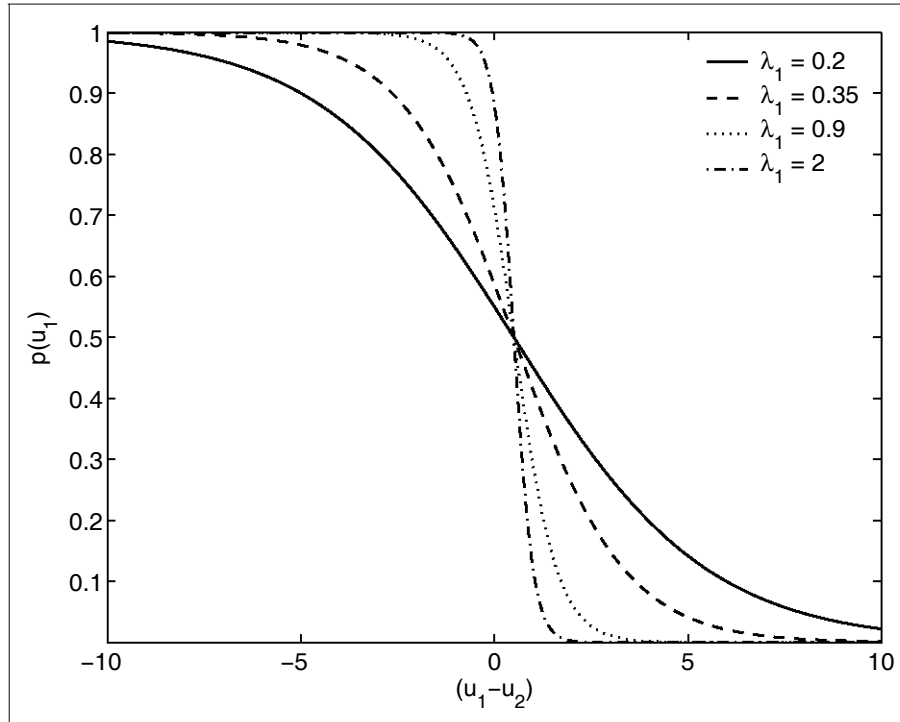
As with Equation (2.9), the mode choice model can also be formulated as a conditional probability:

$$p_{m|ij:k} = f(c_{ijm|k}^*, c_{ij\{m\}|k}), \quad (2.10)$$

where  $k$  is a segment of the population,  $i$  and  $j$  are the origin and destination zone,  $m$  the mode,  $p_{m|ij:k}$  is the proportion of all travellers from population segment  $k$  moving between zone  $i$  and zone  $j$  who use mode  $m$ ,  $c_{ijm|k}^*$  is the composite cost of travel between zones  $i$  and  $j$  by mode  $m$ , and  $c_{ij\{m\}|k}$  is the associated cost of travel by all modes, with  $\{m\}$  the set of modes  $\mathcal{M}$  being considered.

According to Bates (2000), the main sources of variation in the mode choice models that are used, in practice, are the number and type of modes actually distinguished, and the detail of the generalised cost functions  $c_{ijm|k}^*$ . Mode choice models originally only distinguished between PMT and PT modes. They used sigmoidal curves, whereby the probability of choosing the mode vanishes when the costs of using it are greatly in excess of the costs of using the other mode, but which allows for reasonable sensitivity when the costs of using either mode are comparable. For multimodal models, commonly used nowadays, the discrete-choice, or multinomial logit (MNL) formulation of Equation (2.10) is used:

$$p_{m|ij:k} = \frac{\exp(-\lambda_{1|k} f(c_{ijm|k}^*))}{\sum_{m' \in \mathcal{M}} \exp(-\lambda_{1|k} f(c_{ijm'|k}^*))}, \quad (2.11)$$



**Figure 2.7:** Diversion curves for decision choice behaviour.

Source: Adapted from Zuidgeest (2005).

where  $\lambda_{1|k}$  is the ‘spread’, or scale parameter, reflecting the degree of substitutability between the modes  $m$ .  $\lambda_{1|k}$ , therefore, represents the sensitivity of choice of mode to changes in generalised cost, or utility  $c_{ijm|k}^*$ . Often,  $\lambda_{1|k}$  is chosen to be unity, as in the Maximum Likelihood Estimation of the model, and gets integrated in the parameters for the utility function  $f(c_{ijm|k}^*)$ , denoted as utility or disutility,  $u_{ijm|k}$ . The effect of variations in the value of  $\lambda_{1|k}$ , for the binomial case, is depicted in Figure 2.7. At lower values of  $\lambda_{1|k}$ , changes in decision choice, in this case choosing for alternative 1 with disutility, or travel impedance  $u_1$ , in relation to alternative 2 with disutility  $u_2$ , are smoother.

Utility is considered to be the value that individuals derive from choosing a certain alternative and is, therefore, related to the (relative) attractiveness of the modes of transport. The net-utility for mode alternative  $m$ , held by an individual or a population segment  $k$  consists of a measurable, or systematic part  $v_{m|k}$ , and a random part  $\epsilon_{m|k}$ , representing particular taste-values, but also observational errors made in the modelling:

$$u_{m|k} = v_{m|k} + \epsilon_{m|k}. \quad (2.12)$$

The measurable part  $v_{m|k}$  may be a function of several attributes, such as travel characteristics, travel time and the monetary travel costs per mode. As in neoclassical economic theory, the alternative with the highest utility is assumed to be chosen. Therefore, the probability that alternative  $m$  is chosen by decision-maker type  $k$  within choice-set  $\mathcal{M}$  is:

$$p_{m|k} = p \left[ u_{m|k} = \arg \max_{m' \in \mathcal{M}} u_{m'|k} \right]. \quad (2.13)$$

The mode-specific travel demand  $T_{ijm}$  can now be calculated using the discrete-choice model (2.11):

$$T_{ijm} = \theta_m T_{ij} p_{m|ij:k}, \quad (2.14)$$

where  $\theta_m$  is the vehicle occupancy factor.

#### 2.2.4 Trip Assignment

In the fourth step of the model, the travel demand,  $T_{ijm}$ , between points  $i$  and  $j$  for mode  $m$ , is estimated using a traffic assignment sub-model, which draws upon information from a supply model representing the transport system itself. The trips  $T_{ijm}$  are converted into vehicle trips and are (iteratively) loaded onto the network of shortest paths, with a route  $r$ 's specific generalised costs  $c_{ijmr}$  between the origins  $i$  and destinations  $j$  for the modes  $m$ , either allowing or not allowing for congestion to develop. The simplest all-or-nothing traffic assignment model reads:

$$T_{ijmr} = \begin{cases} T_{ijm} & \text{for the minimum cost route } c_{ijmr}, \\ 0 & \text{for all other routes.} \end{cases} \quad (2.15)$$

In Equation (2.15), all trips between points  $i$  and  $j$  are assigned to the route between points  $i$  and  $j$  with the minimum cost, on the basis that this is the route that travellers would want to use. However, this situation is only realistic when there is no congestion, and when there is only one route with a distinct minimum cost advantage. If there is congestion, the assignment model should, at least, be capacity-restrained, leading to the so-called *Wardrop's First Principle*, or *User-Equilibrium (UE)*, in which all travellers between an origin-destination pair perceive equal costs. If congestion is considered, a nonlinear congestion curve representing the speed-flow relationship is, usually, applied. This is often used in conjunction with the Bureau of Public Roads' (BPR) travel time Equation (Bureau of Public Roads, 1964) :

$$\tau_l = \tau_l^0 \left[ 1.0 + \alpha_1 \left( \frac{V_l}{C_l} \right)^{\beta_1} \right], \quad (2.16)$$

where  $\tau_l$  is the travel time on link  $l$  in the transport network,  $\tau_l^0$  the free-flow travel time on link  $l$ ,  $V_l$  and  $C_l$ , respectively, the link volume and link capacity. In addition,  $\alpha_1$  and  $\beta_1$  are positive-valued parameters. Naturally, a route, generally, comprises of several links. Therefore, route travel time  $\tau_{ijmr}$  is calculated as the summation of individual link travel times, with each link comprising a certain route  $r$ , between points  $i$  and  $j$  (by mode  $m$ ). When not considering intersection delay, as is sometimes done in these types of models, the route travel time is given by:

$$\tau_{ijmr} = \sum_{l \in r \{ \subset \mathcal{R}_{ij} \}} \tau_l^0 \left[ 1.0 + \alpha_1 \left( \frac{V_l}{C_l} \right)^{\beta_1} \right]. \quad (2.17)$$

The mode choice and traffic assignment model can also be combined in a simultaneous mode, destination and route choice model. Combining equations (2.7) and (2.11) yields:

$$T_{ijmr|k} = \theta_m T_i \frac{X_j \exp(-\lambda_{1|k} f(c_{ijmr|k}))}{\sum_{j' \in \mathcal{J}} \sum_{m' \in \mathcal{M}} \sum_{r' \in \mathcal{R}_{i,j'}} X_{j'} \exp(-\lambda_{1|k} f(c_{ij'm'r'|k}))}, \quad (2.18)$$

where the total utility  $f(c_{ijmr|k})$  of a destination-mode-route choice combination, for a given origin, is expressed as a function of individual measurable choice-utility components  $v$  and accompanying random elements  $\epsilon$  (compare equation (2.12)), noting that directly, utility consists of time and cost elements  $c$ :

$$f(c_{ijmr|k}) \equiv u_{ijmr|k} = v_{ij|k} + v_{jm|k} + v_{ijmr|k} + \epsilon_{ij|k} + \epsilon_{jm|k} + \epsilon_{ijmr|k}. \quad (2.19)$$

This formulation only holds when the individual choice-utility components are assumed to be *non-correlated*, and parameter  $\lambda_{1|k}$  is unique for the combined choice<sup>6</sup>. A multinomial logit model can then be formed, in which the argument is the conditional utility (based on travel time  $\tau_{ijmr|k}$  and travel costs  $\kappa_{ijmr|k}$ ) for a given simultaneous origin  $i$ , destination  $j$ , mode  $m$  and route  $r$  alternative, i.e.:

$$u_{ijmr|k} = \alpha \tau_{ijmr|k} + \beta \kappa_{ijmr|k}. \quad (2.20)$$

From this simultaneous mode, destination, route choice model or the standard traffic assignment sub-model, the vehicle link flows  $q_{lm}^v$  and link travel times  $\tau_l$  and, thus, link speeds  $s_l$ , can be obtained through link-route analysis.

Several forms of traffic impact analysis can now be performed using the information generated by the different sub-models. Analysts use the information for purposes ranging from selected link analysis, to investigate, for example, congestion levels, to different types of environmental impact studies by applying environmental models, thus generating estimates for energy consumption, traffic emissions and noise pollution.

### 2.2.5 Discussion

The four-step method is the predominant model used (even if not always used as a whole) in transport planning today. The detailed discussion of the mathematical foundations of the method is important so that the assumptions made in the method, and the implications this has for infrastructure provision, can be seen in light of the discussion around transport policy, presented in the next section (Section 2.3).

Firstly, in the trip generation step, productions are often simpler to estimate than attractions, because the data required can be reliably sourced from (typically) census information, whereas attractions require a range of metrics for each land use type to arrive at a similarly robust estimate. Trip attraction is, therefore, often balanced by using a proportionality factor applied to the trip production estimates. Although factoring the attractions as a proportion of the productions simplifies the data collection and estimations required, it introduces errors between the modelled results and the actual measured traffic flows. This step of the method is, therefore, often replaced or supplemented by information collated empirically and distributed in tabular form in trip generation manuals, such as that discussed in Section 2.4.2. However, this introduces further problems, as will be shown. Assumptions around

<sup>6</sup> see Oppenheim (1995) for a discussion on such formulation

attraction are further complicated in situations of high unemployment, uneven distributions of income levels and, of course, inequitable levels of accessibility.

The trip distribution model is, essentially, introduced to distribute trip ends between zones in a way that is consistent with production and attraction values at these zones. Trip distribution, of course, must be related to the quality and ease of access between two zones, a factor which is almost never considered when developing OD-matrices.

Similarly, the modal split model relates the mode choice to various cost elements. The underlying assumption is that all modes are available to all travellers, albeit at different cost levels. This assumption is flawed since, of course, in the developing world, all modes are not available to all travellers, and the generalised costs of travel by a particular mode cannot be used to account for mode choice. Furthermore, in most instances in practice, this step is completely excluded from the analysis which, mostly, exclusively models PMT. Where PT is included as an alternate mode, it is devolved into equivalent private cars. NMT modes are almost never considered as a viable option and so the mode choices of a large proportion of travellers are not realistically represented.

The assumptions made in the trip assignment sub-model, generally, do not include information about traveller familiarity with the network, and cannot account for unpredictable behaviour (like traveller propensity to use a congested link despite the option of better alternatives). The model explicitly assumes that people will always seek the lowest cost route, which is not necessarily the case. The fundamental flaw in the assumption is that people have access to all the information they need to judge which route is the lowest cost (which is almost always false) and that, given this information, they are able to make the assessment efficiently and accurately (also false), and that their decisions, based upon their analysis of this information, are rational (in general, also false). This is the basis for interventions such as Variable Message Signs and other information technology services, which provides travellers with some of the additional information they need to make more informed travel decisions. These, however, are not able to provide 100% of the information of the status of the network, and cannot ensure that travellers will make rational route choice decisions.

### **2.3 Road Planning and Policy in South Africa**

The political transition that South Africa underwent in the early 1990s brought with it a series of policy and legislative changes in the transport sphere, aimed at bringing to effect the vision of a new, more equal, South African society. The National Department of Transport (NDoT) established the National Transport Forum in 1992, bringing together a range of stakeholders, which identified shortcomings in the existing policies. The forum compiled a number of policy proposals, which were included as inputs in the National Transport Policy Review, a consultative process initiated by the NDoT aimed at formulating a transport policy for the new South Africa. This process culminated in the publication of a Green Paper which, after an extensive period of public review, led to the publication of a White Paper on transport policy for South Africa.

The NDoT began the Moving South Africa (MSA) project in June, 1997, to take the vision developed in the 1996 White Paper and develop a twenty year strategy to realise it. The project was given a mandate to “... *develop a strategy to ensure that the transportation system of South Africa meets the needs of South Africa in the 21st century and therefore contributes to the country's growth and economic development.*” (NDoT, 1999).

Since the White Paper had already set out the vision for transportation in South Africa, MSA's mission was to determine how - in an environment of limited resources, capacity, and time - to implement that

vision in a way that was consistent with the key thrusts identified in the White Paper. It was, therefore, necessary for the strategy to verify the White Paper objectives on the basis of hard data, and to prioritise amongst some of the sometimes competing objectives of the White Paper.

The National Land Transport Transition Act (NLTTA) was promulgated in 2000 to provide the measures necessary to transform and to restructure the Republic's land transport system, to give effect to the national policy concerning the first phases of the process and to achieve a smooth transition to the new system applicable nationally (Republic of South Africa, 2000). The NLTTA was superseded by The National Land Transport Act (NLTA), promulgated in 2009, to further the process of the transformation and restructuring of the national land transport system initiated by the NLTTA.

The NLTA was formulated to give effect to national policy, to prescribe national principles, requirements, guidelines, frameworks, national norms and standards that are to be applied uniformly in the provinces, and other matters contemplated in Section 146(2) of the Constitution. The act also aims to consolidate land transport functions and locate them in the appropriate sphere of government (Republic of South Africa, 2009).

The National Land Transport Strategic Framework (NLTSF) is a legal requirement in terms of Section 21 of the NLTTA. Section 21 of the NLTTA required that *"The Minister must annually, by a date to be determined by the Minister after consultation with the MECs and published by notice in the Government Gazette, prepare a national land transport strategic framework for the country for a five-year period corresponding with the Department's financial years. . ."* (Republic of South Africa, 2000, page 22-23). The framework gives guidance on transport planning and land transport delivery by national government, provinces and municipalities for a five-year period.

These various policy documents are discussed in more detail in the following sections.

### 2.3.1 White Paper on National Transport Policy (1996)

The White Paper (NDoT, 1996) represented a significant departure from previous transport policy in South Africa. It placed a strong emphasis on redressing the wrongs of Apartheid and specifically highlights the need to promote PT over PMT. The mission statement proposed by the White Paper with respect to land passenger transport is for the Department of Transport to provide leadership in:

*"The promotion of a safe, reliable, effective, efficient, coordinated, integrated, and environmentally friendly land passenger transport system in South African urban and rural areas, and the Southern African region, managed in an accountable manner to ensure that people experience improving levels of mobility and accessibility."*

The document lists a number of 'strategic objectives' that should be reached to achieve the mission. Amongst these, the document lists ensuring sustainable and dedicated funding for passenger transport infrastructure, operations, and law enforcement. It also lists encouraging more efficient urban land use structures, correcting spatial imbalances and reducing travel distances and times for commuting, specifically mentioning a limit of around 40 km or one hour in each direction as an objective. It explicitly requires the promotion of the use of PT over PMT, specifically setting a goal of achieving a ratio of 80:20 between PT and PMT usage.

The list of objectives also incorporates statements regarding the requirement for equity in addressing the needs of all road users. It highlights the importance of PT services being affordable, with a stated objective of commuters spending less than 10% of their disposable incomes on transportation. The centrality of NMT is confirmed by the objective of encouraging, promoting and planning for the use of NMT wherever appropriate.

The objectives listed imply an approach to transportation that is very much driven by issues of social justice and equity which is, of course, key to righting the wrongs of Apartheid planning. The strong focus on PT prioritisation, by proposing a modal split and placing targets on household expenditure on transport, coupled with the focus on limiting travel times and improving accessibility and mobility, signals a planning paradigm very much aligned with the needs of the poor. Mechanisms are already suggested for achieving these objectives, being primarily related to improved regulation of PT services and service providers, and a more holistic, multimodal approach to planning. The document also emphasises the importance of better integrating land use planning and transportation planning.

### 2.3.2 Moving South Africa: The Action Agenda (1999)

The Moving South Africa (MSA) project (NDoT, 1999) was commissioned by the NDoT to, “... produce a data-driven program for strategic action that extends the short to medium-term policy formulation documented in the White Paper into a long-term strategic formulation embodying the sets of trade-offs and choices necessary to realise the vision as set out in the White Paper.” The terms of reference for MSA were, therefore, very much linked to the recommendations of the White Paper that preceded it. As such, the strategies detailed in the MSA document align closely with the objectives of the White Paper.

The document identifies a number of distinct customer segments, being urban passenger transport, rural passenger transport, tourist and long-distance customers, special needs customers, and freight transport. For each market segment a vision is developed, strategic challenges are identified and actions are proposed, specifying key targets to achieve that vision.

With respect to urban passenger transport, the vision developed places a very strong focus on the promotion of PT over other modes.

*“By 2020, urban customers will be able to participate fully in the various activities of city life by using a PT network that provides as much city-wide coverage as possible and which is affordable, safe, secure, fast and frequent.*

*The core of the PT system will be a network of high volume, high frequency corridors in which PT will be the priority. Customers’ need for improved access and short trip times will be met by having regular feeder services to the high volume corridors, user-friendly transfer facilities, short wait times due to high corridor frequencies and the possibility of differentiated services for customers with specific needs.*

*... to achieve this vision, PT provision must be planned and regulated at the local level, with local control over stable funding sources for both operations and infrastructure, detailed research into local customer needs and close co-operation with local land use planning and other relevant local functions.”*

A number of core challenges are identified. These include a lack of affordable basic access, an ineffective PT system, increasing car dependence and inefficient spatial planning. The drivers of these problems are identified as poor subsidy targeting, the effects of past land use patterns, the lack of financial sustainability of the PT system and the lack of PT infrastructure investment, poor PT planning, operation and regulation and high road investment.

The document then suggests a number of strategies to achieve the vision. Amongst these, the document identifies the densification of corridors as an appropriate strategy for improving PT in the context of the dispersed, low density city structures, resulting from the Apartheid era planning ideals of single use developments and resettlement of communities to peripheral areas to further the ends of racial

segregation. Densification along corridors could be achieved by halting any further dispersion of development away from corridors, by creating intermediate employment nodes on the corridor and through building residential areas closer to the CBD. It is argued that, in fact, corridors are already naturally occurring in many metro areas and the strategy would, therefore, simply encourage these to grow.

The optimal deployment of modes to meet customer service requirements is listed as a strategy and defined through a combination of ‘customer-based transport planning’ and tough road-space management. Customer-based transport planning involves the identification of the appropriate (PT) mode along a corridor, given information on ridership levels and service preferences. Road-space management is used to support PT and NMT modes through “*a combination of controls and pricing, backed up by improvements in the PT system.*”

The document also highlights specific actions with respect to road planning. Building on the argument that roads are used by all the sectors distinguished in the document, it is suggested that road investment and maintenance strategies be oriented towards the strategies identified for each group. This necessitates the designation of specific corridors of movement that should be prioritised for each sector. Furthermore, within this sector specific network, road investments should be prioritised according to customer objectives and national objectives. A further strong theme is to correct the pricing of facility use. The policy calls for a full-user cost recovery, including all externalities associated with the use of the road. This, it is argued, will assist with reducing the severity of congestion and pollution.

Moving South Africa made a definitive contribution to the policy and strategic transport planning landscape in South Africa, building upon the recommendations of the 1996 White Paper, and setting the stage for the promulgation of legislation. Regarding urban passenger transport, the report set out in clear terms where the priorities should lie, and began to describe strategies to achieve the objectives it defined.

### **2.3.3 The National Land Transport Transition Act (2000) and the National Land Transport Act (2009)**

The first major revision of the legislation that governed transportation came in the form of the National Land Transport Transition Act of 2000 (Republic of South Africa, 2000), which has subsequently been superseded by the National Land Transport Act of 2009 (Republic of South Africa, 2009). The purpose of the transitional act is described in the introduction to the act as being to provide the measures necessary to:

1. Transform and to restructure the Republic’s land transport system.
2. Give effect to the national policy concerning the first phases of the process.
3. Achieve a smooth transition to the new system applicable nationally.

The act prescribes those policies, principles, requirements, guidelines, frameworks, standards and other matters of the Constitution required to deal effectively with the transformation and restructuring of the land transport system of the country, in the process introducing and establishing the new land transport system for the country as a whole.

The 2000 act furthers the visions and objectives of the 1996 White Paper and the 1999 MSA policy document, by placing a very strong emphasis on the importance of PT, and issues of social justice and the upliftment of the poor (ostensibly through improved mobility and access).

The act sets out the principles that should apply to the determination, formulation, development and application of land transport policy. It states that PT services are aimed at providing affordable transport

to the public, and should be designed so as to achieve integration of modes, cost-efficiency and service quality. Furthermore, the services should have value to the customer and have the least harmful impact on the environment.

The act requires that PT services be designed such that appropriate modes are selected and planned for on the basis of where they will have the highest impact on reducing the total systems cost of travel, and this decision should be informed by an appropriate assessment of the impact on the customer and the anticipated customer reaction to such change. They should be planned so that customer needs are met by facilitating customer reactions to system changes in the planning process and by maximising the integration of such services. They should also be planned, where possible, so that subsidies are aimed at assisting currently marginalised users and those who have poor access to social and economic activity.

The act requires that all role-players must strive to achieve an effective land transport system through integrated planning, provision and regulation of infrastructure and services and diligent and effective law enforcement.

Continuing the emphasis on modal integration, the act states that PT services, facilities and infrastructure must be provided and developed so as to integrate the different modes of land transport. Safety and effective law enforcement must be promoted as vital factors in land transport management and regulatory systems and the efforts, in this regard, of all competent authorities and functionaries must be coordinated to prevent duplication.

As in the MSA document, the act specifically states that for the purposes of land transport planning and the provision of land transport infrastructure and facilities, PT must be given higher priority than PMT<sup>7</sup>.

The act also furthers the objectives of the national strategy by requiring that land transport functions must be integrated with related functions, such as land use and economic planning and development, specifically mentioning the development of corridors. Densification and infilling, and transport planning must guide land use and development planning.

The act also emphasises the equity issues raised in the policy documents. It requires that the needs of special categories of passengers must be considered in planning and providing PT infrastructure, facilities and services, and that these needs should be met, as far as possible, by the system provided for mainstream PT.

The act emphasises the principle that user charging, or direct cost recovery from users, must be applied wherever appropriate and possible, in that such users should pay for all or most of the costs related to the service or activity in question. This furthers the policy objective of eliminating the indirect subsidisation of PMT modes.

In Part 3 of the act, it further requires that the Minister must facilitate the increased utilisation of PT, strive to ensure that in the promotion of integrated PT modes, due consideration is given to the needs of all transport users and promote effective integrated transport planning.

With regard to transport planning, in Part 7 the act requires that land transport planning must be integrated with the land development process. Subject to this section, land transport planning must be carried out so as to cover both PT and PMT and all the modes of land transport relevant in the area concerned, and must focus on the most effective and economic way of moving from one point to another in the system. Transport plans must be developed so as to enhance the effective functioning of cities,

<sup>7</sup> Some practitioners and academics have noted that the act fails to specify what this means in practical terms.

towns and rural areas through integrated planning of transport infrastructure and facilities, transport operations including freight movement, bulk services and PT services. Transport plans must direct employment opportunities and activities, mixed land uses and high density residential development into high utilisation PT corridors, interconnected through development nodes within the corridors, and discourage urban sprawl where PT services are inadequate. Plans should give priority to infilling and densification along PT corridors and, as mentioned elsewhere, give higher priority to PT than PMT by ensuring the provision of adequate PT services and applying travel demand management measures to discourage PMT. Transport plans should enhance accessibility to PT services and facilities, and transport functionality, in the case of persons with disabilities, and plans should be developed so as to minimise adverse impacts on the environment.

A number of themes are repeated throughout the act. Firstly, as mentioned, the important role envisioned for PT in the transport system is emphatically stated on a number of occasions. Not only does the act compel the Minister to ensure the increased use of PT, which is considered the only sustainable way forward for South Africa, but it also enforces that PT is considered in all areas of transport planning.

On that point, there is also a very strong emphasis placed on integrated land use and transport planning, not only multimodal planning, also a thematic carry-over from the MSA. With regards to the recommendations of the MSA, the act mentions the importance of corridors to planning, and highlights densification and infill development (now making this a legal imperative) as strategies for encouraging the uptake of PT use.

The Act can be read as being very prescriptive in its attempts to bring to fruition the vision developed for South Africa in the White Paper and the MSA project. The 2009 National Land Transport Act is less elaborate overall in its description of the principles for national transport policy, stating only that the Minister must prescribe the principles applying to "... *the determination, formulation, development and application of land transport policy in the Republic.*" This allows the Minister more flexibility in determining policy than the transitional act did - in essence, policy documents, such as the MSA document and the NLTSF, now constitute the NDoT's policy on transportation in South Africa.

However, the new act still requires the Minister to facilitate the increased use of PT and to promote effective integrated transport planning. The Act also still requires that "... *land transport planning must be integrated with the land development and land use planning processes.*...", and so many of the principles developed in the preceding legislation are carried forward in this new act.

#### **2.3.4 The National Land Transport Strategic Framework 2006-2011 (2006)**

The National Land Transport Strategic Framework (NLTSF) (NDoT, 2006a) for the period 2006 to 2011 is a legal requirement in terms of Section 21 of the NLTA and gives guidance on transport planning and land transport delivery by national government, provinces and municipalities.

The document distinguishes between fifteen 'functional areas' within transport which it addresses individually, describing policies, strategies and actions for each. The functional areas that are covered include, amongst others, PT, land use restructuring, roads, cross-border road transport, freight transport, rural transport, traffic safety and enforcement, transport for persons with disabilities, NMT, the environment, tourism and intermodalism and integration of transport planning.

The framework also addresses implementation mechanisms and measures for monitoring the implementation of the NLTSF by means of key performance indicators. The policy is to be updated in each of the first four years following its introduction, with a completely revised version being published in the fifth year, 2011.

As regards land transport policy, the first item the framework addresses is that of PT and NMT stating that:

*“For the purposes of land transport planning and the provision of land transport infrastructure and facilities, PT must be given higher priority than PMT. This will entail the implementation of effective Travel Demand Management (TDM) measures to promote more efficient private car usage and to free up resources for PT upgrading and promotion.”*

Also, with regards to NMT:

*“Land transport planning and provision must pay greater attention to promoting the safe and efficient use of NMT modes such as walking and cycling.”*

As with the documents discussed previously, the NLTSF also specifically states that PT must be prioritised over PMT, that transport planning must be multimodal and integrated with land use planning, that planning must be directed to encourage densification along corridors and to discourage urban sprawl. The document further states that walking and cycling will be promoted as the preferred modes in South Africa for their appropriate distances.

Key performance indicators (KPI's) are identified, divided into customer-based KPI's focusing on PT usage, access to PT and road safety, and NLTSF-based KPI's covering restructuring, regulatory, enforcement, freight transport and funding issues. The document lists data sources for the indices, but does not specifically identify measurement intervals, stating only that MEC's will be required to submit performance measurements at a date set by the Minister, presumably being at least once during the five year update cycle. The customer-based KPI's mentioned in the document are given in Table 2.1.

The data for the KPI's comes primarily from the National Household Travel Survey. The last such survey was conducted in 2003, and although the intention was to conduct it regularly (every five years at least - as with the national census), a follow up survey has yet to be conducted. This calls the practicality of these KPI's into question, since a household travel survey is a large, complicated and, therefore, expensive undertaking. No solution to this problem has been proposed by the NDoT yet.

### **2.3.5 Policies in Summary**

KPI's notwithstanding, all of the themes developed in the policy and legislative documents discussed previously are confirmed in the NLTSF. The South African governments objectives in transport are clearly stated and, at least with respect to roads and transport planning, unequivocal. Clearly, PT is a primary concern, so much so that the provisions dealing with it were originally legislated in great detail in the NLTTA. The vision is presented explicitly under the banner of *'Public Transport First'*, and all policy and legislative references state clearly that PT is to be prioritised over PMT in all aspects of transport planning and infrastructure provision. Policy documents also specifically identify the use of Travel Demand Management (TDM) measures to discourage the excessive use of PMT, and emphasise the principle of *'user pays'* to reconcile the economic realities of the cost of travel and account for the externalities that are currently subsidised for private motorised travel.

Another important area of focus is the integrated planning of transport and land use. There is a stated acknowledgment that land use and transport are interrelated and, as such, affect each other explicitly. Planning in either realm should, therefore, be done with reference to the effects on the other.

Multimodalism is another important policy theme. All the documents specifically require that planning be done from a multimodal perspective, with all modes being considered and provided for wherever appropriate. Furthermore, the most preferred mode must be identified and the use of that mode be encouraged, given the situation.

**Table 2.1:** Customer-based Key Performance Indicators for the NLTSF 2006-2011

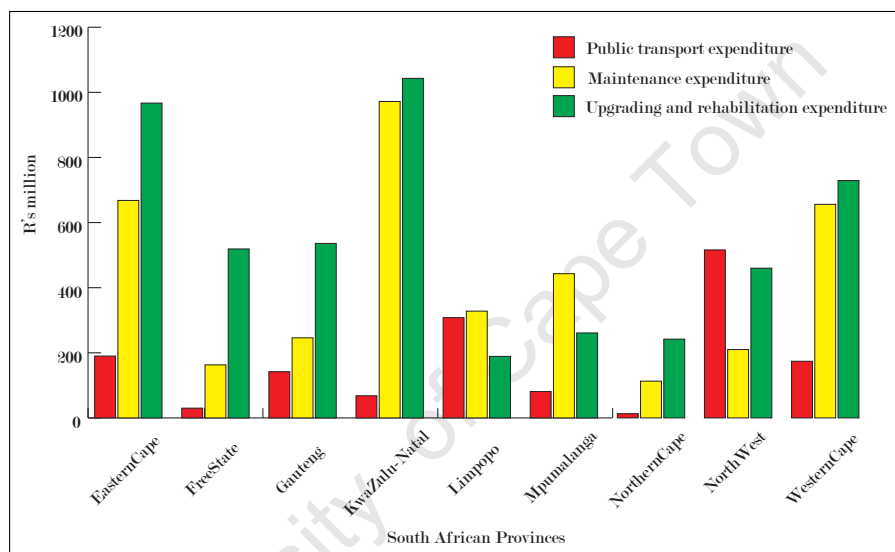
Key policy area	Customer-based KPI	Definitions
Promotion of PT usage	1. Average travel time to work, for all PT commuters.	1.1 Travel time to work: Total one-way door-to-door time from the time of leaving the home until arrival at the regular work location. This includes time taken for intermediate stops along the way to work, and excludes those working at home. In surveys, travel time should be recorded for the 'regular' daily trip to work. 1.2 Public transport: Bus, minibus-taxi, rail and other transport for reward as defined in NLTTA.
	2. Percentage of motorised transport users using PT to work.	2.1 Motorised transport users: People using motorised PT or PMT modes at any stage during the trip to work, during the morning peak period. Morning peak period should be defined for each area, or a uniform 5:30 to 8:30 period can be applied. 2.2 Public transport: see 1.2.
	3. Average age of subsidised bus, mini/midibus-taxi, and commuter rail coach fleet.	3.1 Age: Time since date of manufacture (rebuilt and rehabilitated vehicles deemed 3 and 8 years old respectively). 3.2 Subsidised bus: Bus services receiving operating subsidies from national or provincial government. 3.3 Rail coach fleet: Coaches regularly used for commuter rail operations.
Promotion of access to public transport	4. Percentage of rural people living within 2 km of access to regular PT services.	4.1 Rural: Residents of magisterial districts outside metropolitan areas, metro fringes, and major towns as defined by the NDoT's Rural Typology Study. 4.2 Access: Public transport station or stop is within a 30-minute walk or 2 km of the residence (self-reported). 4.3 Regular: Operating with a frequency of at least once per week. 4.4 Public transport: see 1.2.
	5. Percentage of households spending more than 10% of disposable income on PT	5.1 Household: A person or group of people living together for at least 4 nights per week, who eat together and share resources. 5.2 Disposable income: Monthly take-home income per household after deductions. 5.3 Monthly spending on PT: Households total monthly expenditure on public transport, excluding money spent on holiday travel.
Traffic safety	6. Number of road traffic fatalities per vehicle type.	6.1 Road traffic fatalities: Road users (including drivers and passengers of motorised modes and pedal cycles) dying within 6 days of being involved in a road traffic accident. 6.2 Vehicle type: Disaggregated for users of car, bus, minibus, light delivery, heavy vehicle, pedal cycles, and other vehicles.
	7. Number of road traffic pedestrian fatalities.	7.1 Pedestrian fatalities: Pedestrians dying within 6 days of being involved in a road traffic accident.
	8. Number of road traffic fatalities per 100 million vehicle km per vehicle type.	8.1 Fatalities: see 6.1. 8.2 Vehicle kilometres: Annual vehicle kilometres travelled. 8.3 Vehicle type: Disaggregated for users of car, bus, minibus, light delivery, and heavy vehicles.

Source: Adapted from NDoT (2006a)

NMT is specifically mentioned as being the preferred mode for short trips, and the use of these modes is to be encouraged wherever possible. This implies that infrastructure for NMT modes should be provided to accommodate short trips wherever required.

Another important aspect of policy is that of corridor planning. The MSA document notes that many South African areas lend themselves to this type of city structure, and that corridor planning offers the only viable solution to improving conditions to support PT, given the low densities typical of cities in South Africa. Policy documents specifically mention actions to promote densification, especially along high priority corridors, and to discourage urban sprawl.

Since these policy measures have been in place for at least 15 years already, to what extent are they being translated into tangible action by provincial and local authorities? Given that infrastructure provision is, to a large extent, the end result of transport planning, it should reflect these priorities.



**Figure 2.8:** Provincial roads and transport expenses 2007-2008. A comparison of the expenditure on PT specific projects, road infrastructure maintenance, road upgrades and rehabilitation projects for the financial year 2007 to 2008 for each of the provinces in South Africa.

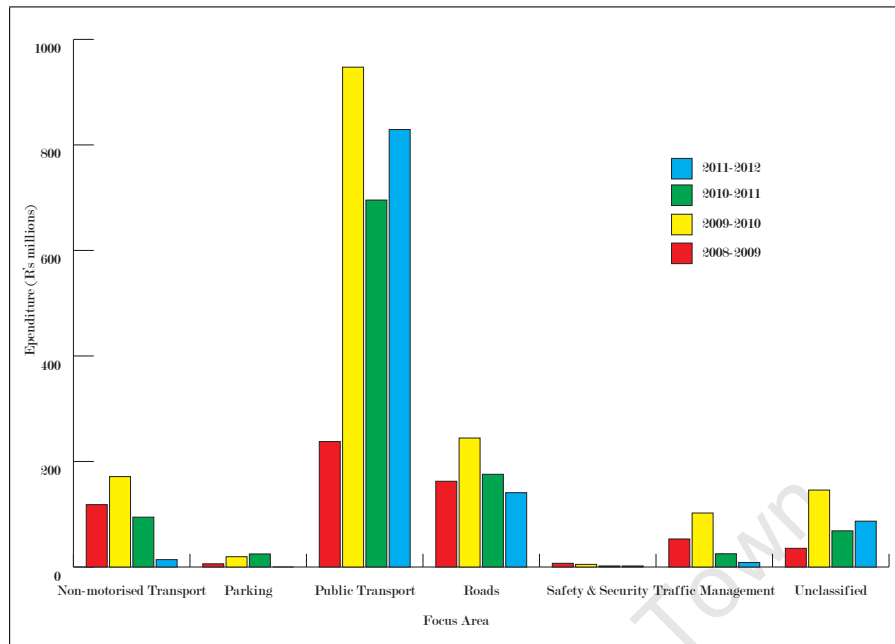
*Source:* South African National Treasury (2009)

Expenditure on PT is not a priority for the majority of provinces, although there are notable exceptions, specifically in Limpopo and in the North West Province (Figure 2.8). The remaining provinces tended to spend very little of their budgets for this particular year on PT. Road upgrades and rehabilitation projects formed the majority of infrastructure expenditure in six of the nine provinces, and was the second largest expense in a further two. It could be argued, however, that PT is more of a municipal function than provincial one. PT related projects feature very prominently in the City of Cape Town's budgetary planning (Figure 2.9), which is very encouraging in terms of the directives developed in the guiding policy and legislative documents.

## 2.4 Guidelines and Transport Policy

### 2.4.1 The Role of Guidelines in Transport Planning

Policy and legislation drive the provision of transport infrastructure, in that they set the priorities, strategies and actions to be taken when planning transportation. In order to successfully enact policy,



**Figure 2.9:** Expenditure on roads and transport projects, both completed (2008), under way and projected (2009 - 2012) for the City of Cape Town administrative region for 2008 through 2012.  
*Source:* City of Cape Town (2009b)

however, guidance is often required to assist practitioners when developing plans. This guidance is most often provided in guidelines developed under the auspices of government and guidelines and, therefore, represents the translation of policy and legislative objectives into practical actions. They advise practitioners as to the best practice with regards to the implementation of these policy objectives and strategies. Guidelines, therefore, play an integral role in the practical interpretation of government policies, being the development of plans and infrastructure and so, ideally, should be revised in light of any policy changes or revisions. If not, practitioners are left to their own devices when interpreting policy, possibly leading to plans and infrastructure that is not in line with policy objectives. The worst case scenario is that policy changes are simply ignored, because there is no practical guidance as to how to implement it.

The development of guidelines presents an ideal opportunity for a thorough investigation into the practical issues around implementing new policies, and an assessment of where current standard practice may still be appropriate, or not. It provides practitioners with the opportunity to interrogate current thinking and systems, and identify where these must be adapted, or abandoned and replaced, if needs be, to achieve policy objectives. Often, a good starting point is to look abroad to see what else has been done under similar circumstances which is, to a large extent, how the current range of guidelines originated. This international experience may then be adapted as befits local circumstances, where appropriate, or a requirement for new research into a specific issue may be identified if no suitable solutions are found elsewhere.

In any event, the method used to implement policy, which may be data driven itself, and which forms the content of a guideline, should be grounded in sound, scientific reasoning which, in the field of transportation, is often verified by solid empirical evidence. Most often, research findings from the

academic realm, after a period of dissemination, peer review and debate, work their way into policies and eventually form the basis for the practical guidance given in guidelines.

It is because of this that deviations from the recommendations of guidelines, despite the fact that these only represent some form of general advice to practitioners, are avoided. The scientific standing, and the general acceptance of the recommendations in guidelines as being the best practice, means that they offer a measure of indemnity from professional or legal challenges to the product produced or the effects it has.

Guidelines, however, are developed with a significant level of generalism. They cannot address, with any degree of specificity, all the possible situations that may be encountered in practice. Their use is subject to the acknowledgement that the practitioner must apply the principles outlined in the guideline, as best as possible, to his unique situation. In fact, there is very seldom a situation where an example given in a guideline fits the reality of a particular circumstance exactly. The overly rigorous adherence to the letter of any guideline is, therefore, foolhardy, even more so when the guideline may, in some respects, be outdated.

However, as mentioned, the disincentives to deviating from guidelines are great. Complicating the issue, many roads authorities have developed their own sets of design standards that have been based upon existing guidelines (see SANRAL, 2009; Visser et al., 2003), and these tend to be enforced with even greater compliance requirements. This would not be so much of a problem if the guidelines that are currently in use were regularly revised to reflect the current policy objectives, but unfortunately, many of them are not. Some of the guidelines commonly used or referred to in practice are now examined to see to what extent they reflect those policy ideals described in the previous section.

#### **2.4.2 South African Trip Generation Rates Manual, 2nd Edition**

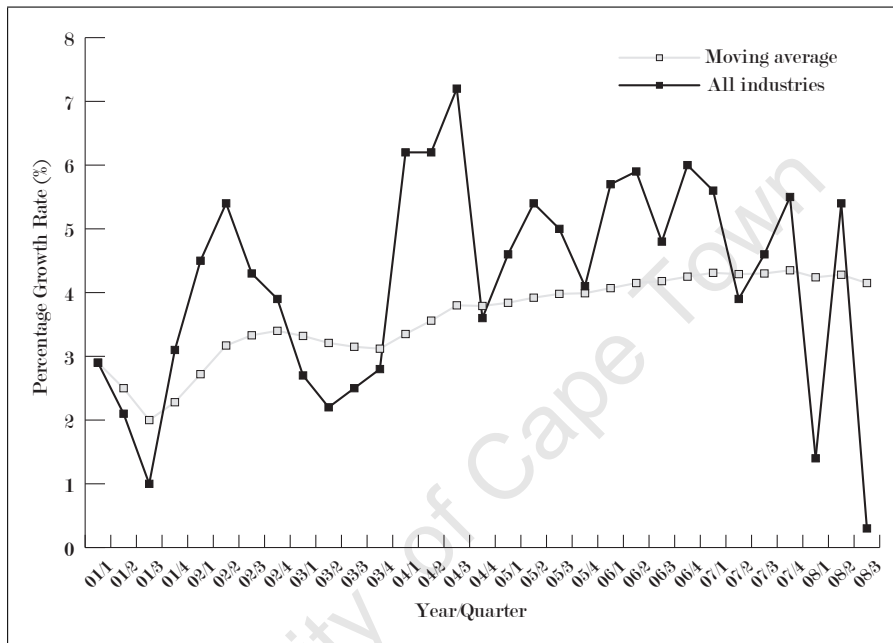
The first step in the four-step model (see Section 2.2), trip generation provides an estimate of the frequency of trips in relation to the propensity to travel, given a set of land use and population characteristics (McNally, 2000). Put more simply, it is an estimate of the demand for travel. An accurate estimate of the number of trips generated is vital, as this forms the basis for all further steps in the transport modelling process.

It is necessary to adopt an empirically verifiable approach to estimate of the number of trips generated. To this end, surveys of existing facilities and land uses are conducted and statistically reliable trip generation indices are developed for various land use categories and socio-economic circumstances. These can then be applied by practitioners with a fair degree of confidence on the projects they may be working on.

As the demand for travel grew in South Africa, the need arose for South Africa to develop its own set of indices. This need was fulfilled with the release of “*The Effect of a Change in Land Use on Traffic Volumes*” issued by the Department of Transport in 1980. This document was updated in 1989 with the release of “*South African Trip Generation Rates, 1989*” to address shortcomings in the original publication. The 1989 volume has, subsequently, been replaced by the current “*South African Trip Generation Rates, 2nd Edition*” (Stander et al., 1995) released in 1995.

The current document covers an expanded range of land uses, and benefits from a larger data set, from which the indices are derived. A total of 193 surveys are included in the report, conducted in all the major cities, by officials and consultants (Stander et al., 1995). The document is still in general use in the practice and is often cited in reports. It is the de facto standard for estimating trip volumes in South Africa.

Over the last decade, South Africa has enjoyed an almost constantly improving economic growth rate (Figure 2.10). Economic growth encourages increased travel demand and, over time, changes in the modal split. This is especially true in South Africa, where mode choice and trip making frequency are very much a function of income (NDoT, 2005a). The recent growth in the South African economy has been driven by a growing middle class and the emergence of a newly empowered economic population group, the so-called emerging black middle class (Investec (Pty) Ltd., 2006). Since the economic landscape in South Africa has changed quite dramatically since 1995, there is reason to question the accuracy of the 1995 indices in today's transport environment.



**Figure 2.10:** South African economic growth rates 2001-2008. The decade 2000 to 2010 has seen the South African economy growing steadily.

Source: Statistics South Africa (2008)

### Trip Generation Manual and Modal Split

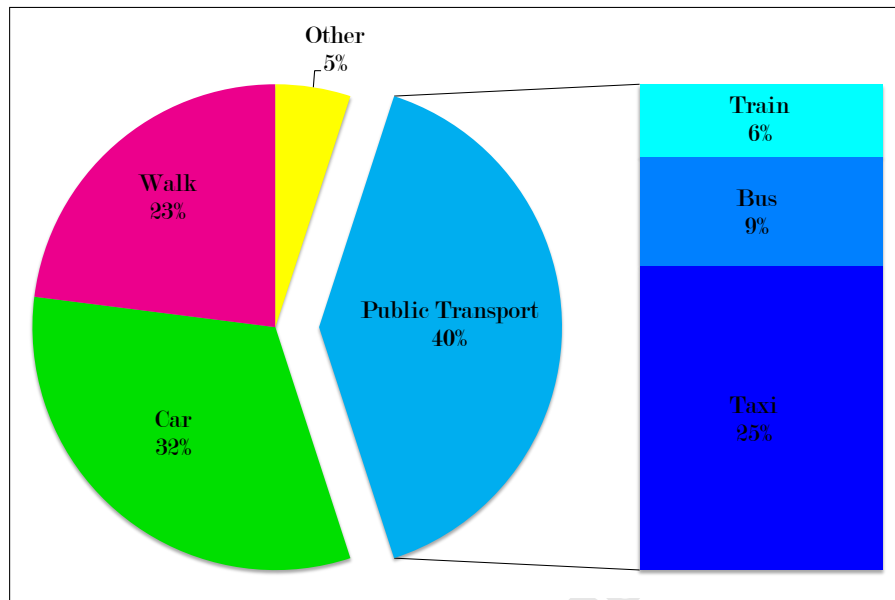
The third step in the four step model, namely modal split, encompasses an attempt to distribute the estimated demand amongst the various modes in a manner representative of the expected reality. Once again, empirical data is often collected to inform the choices made at this step.

The 1995 Trip Generation Rates Manual (Stander et al., 1995) does not provide estimates for trip generation by mode, instead, generally, providing vehicle trips and person trips<sup>8</sup> only. There is no indication of the number of PT trips, or how the number of person trips is split by mode. Consequently, it is not surprising that the resultant infrastructure is often heavily skewed towards motorised transport.

South African travel behaviour is characterised by high levels of walking and PT use (Figure 2.11), primarily amongst the poorer sectors of the population. The implication is, therefore, that in many instances, nearly two-thirds (63%) of all trips are not explicitly considered during traffic assessments.

<sup>8</sup> A person trip is generally defined as being a trip made by one person in any mode of transportation. For example, four persons travelling together in one car make four person trips.

Furthermore, it must also be considered that a PT trip, by necessity, implies at least two further walking trips (access and egress trips).



**Figure 2.11:** Main mode of travel to work, nationally, in South Africa. Most South Africans use PT or NMT to get to work.

Source: NDoT (2005a)

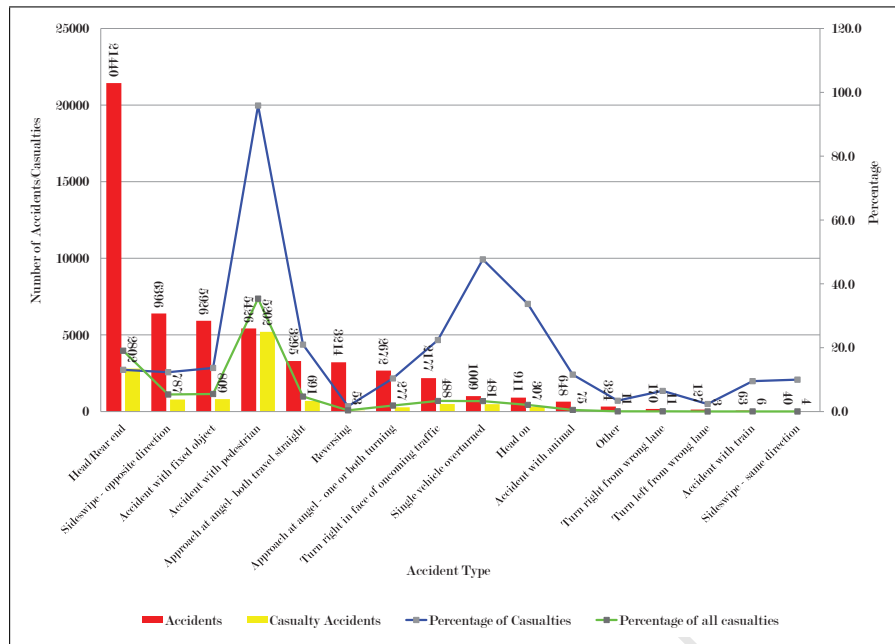
The need for a more complete view of multimodal trip generation becomes all the more urgent in light of the road safety statistics for South Africa (Figure 2.12<sup>9</sup>). Of particular interest is that pedestrian casualties accounted for 35% of all traffic related casualties in Cape Town during 2005, and that of all accidents involving pedestrians that occurred that year, nearly 96% resulted in casualties (although this could also be attributed to under-reporting of non-fatal accidents). This highlights that pedestrians are a particularly vulnerable grouping. Since accidents involving pedestrians almost invariably result in casualties, it becomes imperative that proper care be taken when designing facilities that are likely to be used by multiple modes.

The current trip generation manual can be said to be in danger of basing its indices on outdated data, and will need to be revised. It only provides indices for motorised trips, thereby omitting a significant fraction of trips made, and leading to an underestimation of the needs of NMT and PT road users. This also places it in direct contravention of all of the guiding transport policy documents and legislation.

### 2.4.3 Urban Transport Guidelines Series

The Urban Transport Guidelines (UTG) is a series of documents written for practitioners, which describes recommended practice, based on South African experience and research, in selected aspects of urban transportation. The series was developed and produced by the Committee of Urban Transport Authorities (CUTA), which was formed in 1982 to provide a forum for discussion to promote coordination and, where appropriate, uniformity on technical standards for, and approaches to, the road and transport systems of urban areas in South Africa.

<sup>9</sup> Accidents where the type was given as 'unknown' have been excluded from the chart. This accounts for approximately 34% of all accidents.



**Figure 2.12:** Cape Town accident statistics by accident type. Pedestrians are clearly a particularly vulnerable grouping, accounting for 35% of all traffic related casualties in the area during 2005. *Source:* City of Cape Town (2005)

At the time, the various agencies responsible for the design of urban roads were concerned about the wide range of geometric design standards and policies relating to the design of urban roads, not only between, but even within the various metropolitan areas. At the first meeting of CUTA in August 1982, it was decided to establish an Ad Hoc Technical Committee on Geometries (AHTCG). This committee decided to produce guidelines for the geometric design of urban roads with the following objectives:

1. To promote a uniform approach to the adoption of geometric design standards for urban roads.
2. To recommend dimensions for geometric design elements to provide adequate standards of safety and convenience on urban roads under South African conditions.
3. Recognizing that, in the upgrading of urban roads and in the construction of new roads within built-up urban areas, restrictions in space so often prevent the provision of geometric design elements to ideal dimensions.
4. To provide guidelines for the adoption of reduced dimensions which would, under the prevailing circumstances, still provide reasonable levels of safety and convenience within economic, environmental, social and political constraints under South African conditions.

The committee produced a total of 11 separate guideline documents, one for each class of urban road, the last being updated in 1991 (Rust et al., 2008). The practices promoted in the UTG's were drawn quite explicitly from British and American planning practices in the 1920s and 1930s. This is evident in the adoption of the 'neighbourhood unit' concept (Perry et al., 1929), and ideas about functional road hierarchies and through-traffic elimination developed in the 1940s and 1960s described in the 'environmental area' concept developed by Buchanan (1963).

Both Perry and Buchanan's movement network concepts were defined by a cellular city of local areas bounded by arterials. Both concepts focused on eliminating through traffic from local neighbourhoods

by the use of a closed internal road network with limited access points onto bounding, high volume arterials. The arterials were widened to allow for uncongested traffic flows, whereas internal roads were narrower, using curvilinear alignments and cul-de-sacs to discourage through traffic. Public amenities were located in central locations within neighbourhood cells, which were sized according to the population required to support a primary school.

These ideas were carried forward in the UTG series of manuals. This is aptly demonstrated in the Draft UTG7 manual (Committee of Urban Transport Authorities, 1989) which notes that with regard to residential roads:

*“Residential roads and footpaths are an integral part of housing layout where surroundings free from traffic nuisance are of prime importance and where the patterns of movement around buildings should, with due regard for safety and convenience, give equal priority to the needs of pedestrians and vehicles. To achieve this balance it is important that extraneous traffic be minimised by actively discouraging route continuity.”* (see Committee of Urban Transport Authorities, 1989, page 5)

All of the UTG manuals place a strong emphasis on the function of the street in terms of its relationship with traffic flow, and its role in the network. The eventual design specifications for the road are largely determined according to this functional categorisation of the road, which is defined in terms of its level of provision for mobility and access to property. This focus on road hierarchy, and the rigidity of function in terms of road hierarchy, is carried through even within the individual categories of roads (Figure 2.13).

Figure 2.13 shows a generic residential area layout, used in the UTG7 manual to demonstrate what, in the committee's opinion, were the ideal network features for residential areas. The layout, as with the conceptualisation of residential areas in Perry and Buchanan's plans, is characterised by curvilinear roads, cul-de-sacs and loop routes. Access to the internal residential network is very limited, and the only routes through the area follow convoluted, winding alignments.

The manual then attempts to define each type of residential route in terms of a set of quantitative parameters that define its place in the network, and that begins to define its ultimate design. Table 2.2 illustrates the translation of the hierarchical ideals described in the manual into design parameters.

It is interesting to note that the various categories are already defined in terms of criteria that relate mostly to the needs of automobiles. Design speeds, of course, relates to the width of lanes (since for low-speed residential roads, vehicle turning radii is more important in determining geometric curves than travel speed), and the number of dwelling units related to the expected traffic volumes the link will serve. The needs of other modes of transport is only dealt with superficially. Very little specific information is given on planning for these modes.

The situation is very similar across all the UTG manuals that deal with the geometric design of streets (for urban areas this is manuals 1, 7 and 10), and there is a lot of repetition across the series (although this is to be expected since there is a lot of overlap in the design principles that are applied to each type of road). UTG1, which gives guidance for the geometric design of urban arterials, is the only manual in the series that provides any information on the design requirements for PT vehicles. It only deals with buses, however, and concentrates on the operations of bus services in the traffic stream as it relates to merging movements and the provision of embayments, making reference to *'Bus Terminals and Bus Stations: Planning and Design Guidelines'*<sup>10</sup>.

<sup>10</sup> Department of Transport. *Bus Terminal and Bus Stations Planning and Design*. Draft Report for Use and Comment. Pretoria, July 1983.



**Table 2.2:** Residential road categorisation and design parameters

Road Class	Units served	Max length	Residential uses permitted	Max speed km/h
a) Major residential access link	Up to 200	Less than 500 m	Res. I, II, III & IV	40
b) Access loop	Up to 120	300-500 m	Res. I & II	40
c) Access cul-de-sac	6-60	150 m	Res. I & II	30
d) Access way	Up to 60	50 m between speed restricting elements	Res. I & II	20
e) Access court	Up to 30	50 m	Res. I & II	20
f) Access strip	Up to 4	Depth of erf	Res. I & II	N/A

Dwelling Unit Types:

Residential I: detached houses

Residential II: multiple dwelling units, group housing

Residential III: blocks of flats, low rise

Residential IV: high rise blocks of flats

*Source:* Adapted from Committee of Urban Transport Authorities (1989)

In terms of current policy, with its clear emphasis on PT, the needs of NMT road users, its focus on integrated land use and transport planning and multimodalism, the UTG manuals are outdated, and can no longer be said to reflect the priorities of government. Their continued use in industry is a reflection of the impact they had on the provision of infrastructure, and the level of detail and utility they provide with regards to planning for motorised modes. This one-sided approach to road design is, however, no longer appropriate.

#### 2.4.4 Guidelines for Human Settlement Planning and Design

The *Guidelines for Human Settlement Planning and Design* (CSIR, 2000), also known as the 'Red Book', was compiled by the Council for Scientific and Industrial Research (CSIR) under the auspices of the Department of Housing and published in 2000. Various governmental, academic, professional institutions and private sector bodies contributed to its contents, which covers the full range of planning and design issues related with settlement planning and infrastructure provision.

The guideline positions itself as being, on the one hand, a natural progression from the guidelines that preceded it, and that have been found to have shortcomings under the new policy dispensation and, on the other hand, as embodying a substantial revision of the principles that informed those prior documents. As regards transportation, it attempts to incorporate many of the concepts highlighted in the White Paper on national transport policy and the NLTTA, making significant progress on the role of integrated transport planning, planning for multiple modes and the importance of NMT, especially walking.

Two sections in the document deal, specifically, with roads and transport. The first covers movement

network planning<sup>11</sup>, and deals broadly with the quantitative and qualitative aspects of settlement systems. The second focusses on the engineering issues around road design and construction, including the technical aspects of geometric design. The document still refers the reader to the earlier UTG series of manuals in places, as some of the more detailed technical aspects are not covered. This is not ideal, since it could leave the reader with the impression that the UTG's are still a generally applicable resource, whereas many of the principles they put forward have, to a large extent, been revised in this more current guideline. It would perhaps have been prudent to simply include that material in the UTG's that is still useful in this new guideline.

The manual starts off by highlighting the importance of pedestrian movement in settlement planning, relating it to the scaling of settlements and their constituent elements. Public and PMT is said to become important once the pedestrian scale is exceeded. This chapter presents a very theoretical overview of settlements and their planning, as the authors of the document envision them.

The manual then deals with movement networks and their layouts in more detail, covering aspects of theory and relating them to best practice in planning in a mostly qualitative manner, providing a range of sketches and diagrams to illustrate the concepts it proposes (Figure 2.14).

It is interesting to compare the approach to network layouts promoted in the Red Book to that in the UTG manuals, since both plans are essentially an attempt to achieve similar goals, that is, to create better living environments for people by discouraging high volumes of through traffic. The Red Book plan, however, focusses attention on the primacy of the pedestrian by creating two layouts, one for pedestrians and another for vehicles, from one 'movement network'. The curvilinear roads of the UTG illustration have been replaced by a grid network, affording much more direct routes for pedestrians and shortening walking distances. The focus on road hierarchy has been replaced by a dual network concept, with roads simply being either mixed mode or unimodal.

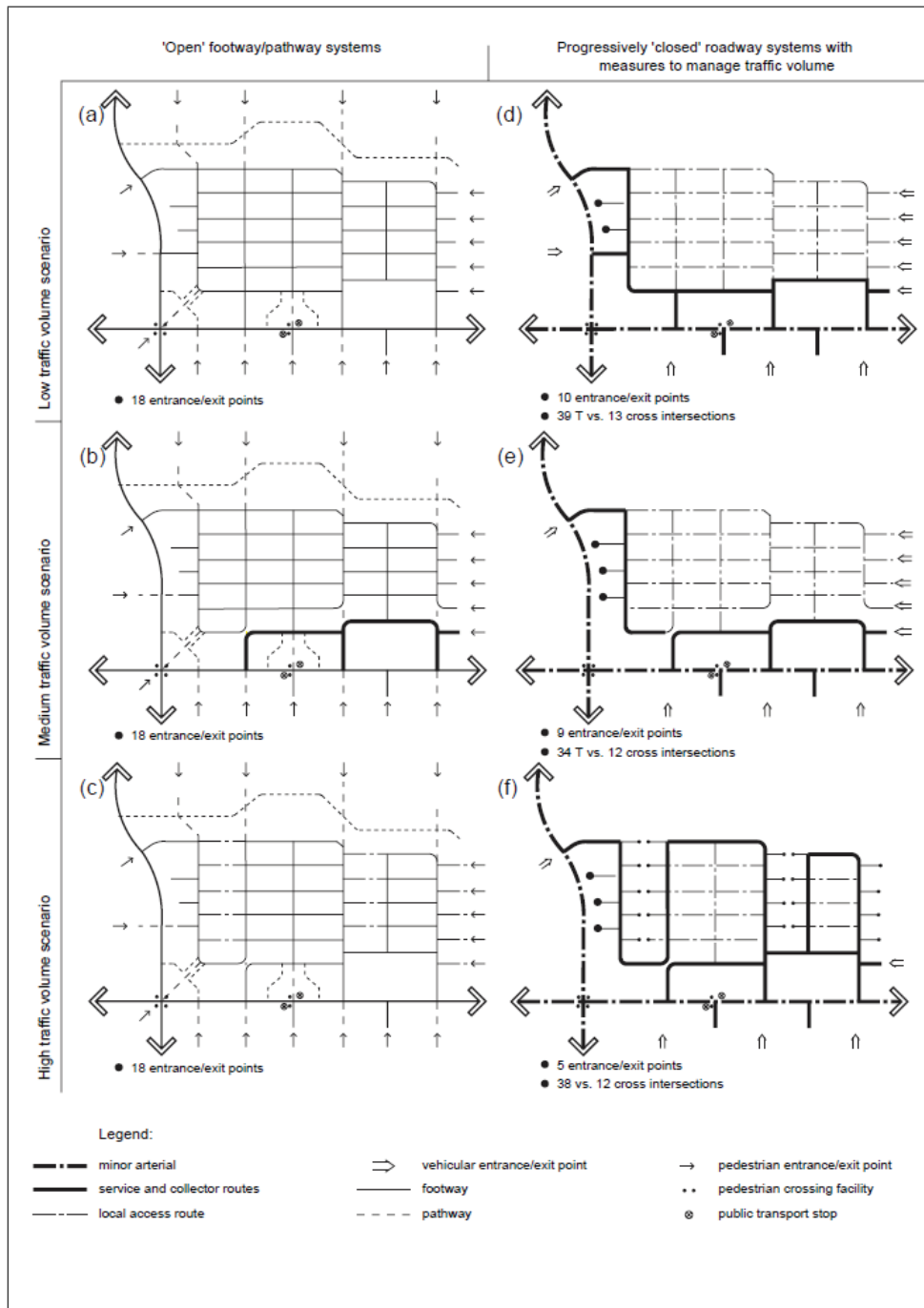
In the Red Book's approach to road hierarchy (Table 2.3), the traditional notions behind road hierarchy have been abandoned. The guideline notes that the traditional hierarchy has been abandoned "... because it placed an over-emphasis on the vehicular movement function of the street system. The concept of a hierarchy also implicitly carried with it the notion of one part of the network being more important than another. The network comprises a system of interlinking streets serving different functions, and often serving these different functions differently."

**Table 2.3:** Comparison between classification systems

<b>Movement Network</b>	<b>Five-Tier System</b>	<b>Urban Transport Guidelines (UTG Series)</b>
Vehicle-only route	1 Regional distributor	(Freeways not included)
	2 Primary distributor	Major arterial (UTG1)
Mixed pedestrian and Vehicle route	3 District distributor	Minor arterial (UTG1)
	4 Local distributor	Collector (UTG15)
	5 Access street	Local street (UTG7 &10)
Pedestrian-only route	5 Access street	(not applicable)

*Source:* Adapted from CSIR, 2000

<sup>11</sup> The term 'movement network' is used to denote that these networks should accommodate multiple modes, and is a nod to the multimodal principles found in transport policy and legislation.



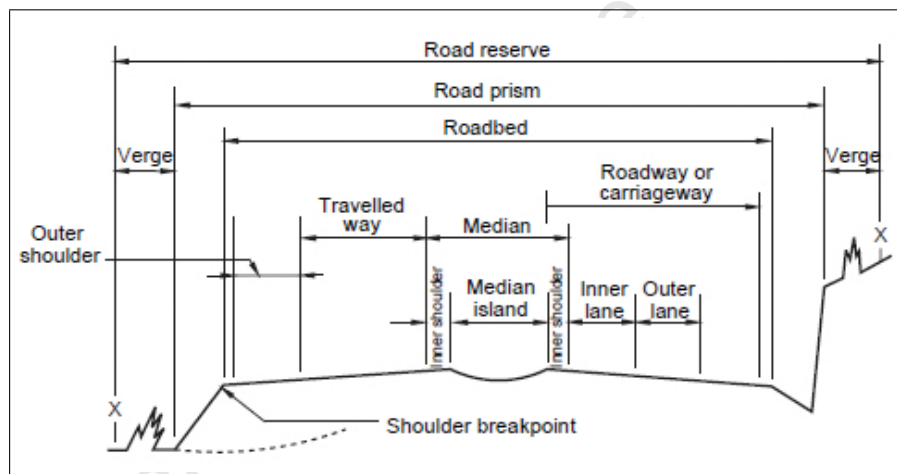
**Figure 2.14:** Red Book proposed configurations of pathway and roadway systems within a public right-of-way network under different traffic volume situations. On the left, the 'open' pedestrian network remains relatively constant whereas, on the right, the vehicular network becomes progressively more restrictive to discourage through traffic.

Source: CSIR, 2000

The wisdom of including this table should be questioned since, at first glance, it appears to be drawing parallels between the UTG-era functional hierarchy approach and the movement network approach. The authors indicate that the table was included to assist designers in developing an understanding of the new classification system, with the caveat that there is no one-to-one relationship between the previous and the new approach. However, this table lends itself to an interpretation of this kind, and will inevitably lead to the new system being understood in terms of the old one.

The manual goes on to cover many of the more technical aspects of geometric design, such as sight distances and turning radii, information which is very similar to that found in the UTG series of manuals, albeit updated somewhat according to more recent research results. Of particular interest to this discussion, however, is the section dealing with cross-sectional elements, in particular, the allocation of road space.

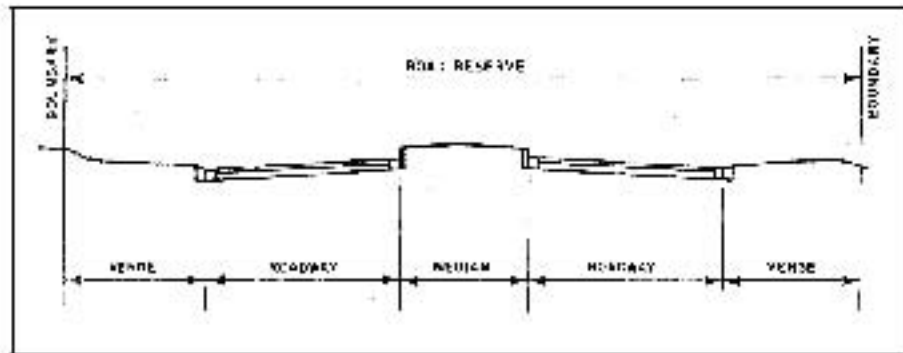
Here, there could be said to be some discrepancies between the theoretical underpinnings of the guideline, the policy framework it subscribes to, and the recommendations made to practitioners. Considering the care with which concepts, such as multimodalism and the importance of NMT and PT, were introduced to the practice, the ultimate recommendations made do not convey a similar sense of urgency as regards provision of infrastructure for these modes. Very little innovation is found in the suggestions the manual makes over what was presented in the UTG manuals.



**Figure 2.15:** Red Book illustration: Typical elements of the cross-section.  
 Source: CSIR, 2000

Figures 2.15 and 2.16 show generic illustrations from the new Red Book guideline and UTG series of manuals of the typical elements of a cross-section. Both illustrations fail to indicate any sidewalks or cycle lanes, and there is no mention made of PT. Although these are simply generic illustrations, they do not carry forward the stated intentions of the new manual, or the ideologies of the policies that drive it.

The guideline innovates in that it addresses the possibilities for the inclusion of High Occupancy Vehicle (HOV) lanes, but it does not suggest any warrants for the provision of these facilities, suggesting instead that from an operational perspective, the provision of a HOV lane should ensure that a street's capacity to move people is increased for it to be warranted. The manual notes that "... where an existing lane is converted to an HOV lane, i.e. when it is no longer available for use by other vehicles, the provision of the HOV lane could lead to a decline in the total throughput of passengers." This statement does not seem to sit well with the policy objectives of decreasing reliance on PMT, since the



**Figure 2.16:** Draft UTG 1 illustration: Typical elements of the cross-section.  
 Source: Committee of Land Transport Officials (COLTO) (1986)

provision of HOV lanes is also related to travel demand management measures, and promotes the use of higher occupancy modes such as PT. The manual refers the practitioner to the Transportation Research Board (TRB) *Special Report 209* (Transportation Research Board, 1994) for further information with regard to the analysis of mass transit facilities.

Regarding the provision of cycle lanes, the guideline makes the following recommendations:

*“Ideally, cycle lanes should be located in the verge area, as the speed differential between bicycles and pedestrians is likely to be less than that between bicycles and motorised vehicles. Where this is not possible and either there is significant cycle traffic or it is desired to encourage bicycles as a mode of travel, a cycle lane can be added outside those intended for motorised vehicles.”*

Recommendations are then made as to appropriate lane widths, being of between 1.5 m to 2.0 m. Regarding the provision of sidewalks and footways, the guideline recommends that *“Wherever there is significant usage by pedestrians, the shoulder is replaced by a sidewalk.”* The reader is then referred to the *Pedestrian facility guidelines: Manual to plan, design and maintain safe pedestrian facilities. Department of Transport Report 92/126, Pretoria* for further information, and a short discussion on the LOS criteria<sup>12</sup> for sidewalk width follows.

Both recommendations rely on the interpretation of the word ‘significant’ in the determination of the necessity for a service and the type of facility provided. The introduction of subjective judgement of the necessity for this type of infrastructure has often led to sidewalks, footways and cycle lanes not being provided in residential suburbs because they add to the cost of the project and they are not, in the strictest sense, required infrastructure. This was true even when the UTG 7 held a rather unambiguous statement on when and where such infrastructure is required.

*“In the case of certain cul-de-sacs or short loops it may be acceptable to design the roadway as a shared surface. On these shared surfaces the width of the roadway should be used to control vehicle speeds and, as such, the needs of pedestrians have a negative influence on roadway width. On other classes of road a separate footway should be provided in the road reserve.”*

Although this statement makes use of the rather unfortunate wording, ‘negative influence on roadway width’, which was clearly meant to indicate that the needs of pedestrians on shared surfaces dictates the

<sup>12</sup> The LOS referred to in the Red Book uses the, now, outdated pedestrian density as a measure of service levels (Dowling et al. (see 2008) for more information).

use of a narrower cross-section, there is little doubt that the authors intended that footways should be provided in the road reserve everywhere, aside from in a few isolated instances, such as in cul-de-sacs. However, since no quantitative parameters are provided as to what constitutes a 'short' loop, a measure of subjective flexibility is introduced that could be exploited to avoid having to pay for the construction of pedestrian services. Arguably, this is what has happened in many instances, although a measure of ignorance around the needs of pedestrians and cyclists is surely also to blame.

The *Guidelines for Human Settlement Planning and Design* makes significant strides in bringing the principles and objectives of policies and legislation into the realm of planning recommendations for practitioners, but in relying too much on prior (often outdated) references, even if only in part, has failed to differentiate itself sufficiently from these references and, as such, may have inadvertently retarded its universal adoption as the standard reference for practitioners.

The document states upfront that it is explicitly not intended to be an administrative 'check list' to aid official scrutiny. The contents are intended solely as a guide, not as specifications and, therefore, are only meant to act as an aid in the preparation of project plans and specifications. The positioning of the document, therefore, does not confer on it any legal status, and its provisions cannot be legally enforced.

The professional responsibility for the success of a project, therefore, remains with the practitioners who, using the document as a planning aid, should use their professional expertise and experience to make decisions regarding design choices. Of course, there is nothing unique in this. All previous guidelines made similar assertions as to their role in practice. The rote application of the recommendations in a guideline cannot be accepted as good practice and, as stated in the guideline, will create more problems than it will solve.

However, it must also be appreciated that newer, more appropriate practices, should be encouraged as far as possible, and that subsequent iterations of guidelines should supersede previous versions. This is especially true for the Red Book, that differs so dramatically in its approach to the planning of urban areas and their transportation infrastructures when compared to what was considered best practice previously. The continued use of outdated references in practice cannot be condoned in light of the problems these lead to, which is precisely the reason for the development of a new set of guidelines in the first place.

## 2.5 Résumé

This chapter aimed to provide an overview of transport planning practice, its origins, the policies and legislation that guide it, and the references used to implement these policies. The issues around current practice were explored in some detail, from the perspective of current international research, the political changes that have led to shifts in policies and legislation, and the often slower, less agile responses to these changes made by practitioners.

Internationally, there is now an established appreciation for the problems caused by rampant motorisation and automobile dependence (Newman et al., 1995). Issues as varied as climate change (Blanco et al., 2009), health (WHO, 2004), and economic costs (Norley and Peters, 2010) have driven perceptions towards a more balanced approach to the provision of transport infrastructure.

This, coupled with South Africa's recent political transition to democracy, has given impetus to a review of transport policy and legislation, a process that has only recently reached some level of finality, with the passing of the National Land Transport Act in 2009. Aside from aligning local practices with current international norms, the new policies were developed with the important final objective of

achieving the social and economic transformation required to address the disparities entrenched by the Apartheid system.

However, the objectives of policies and legislation can only be enacted if they are effectively translated into recommendations for infrastructure provision, and universally adopted by practice. The progress achieved on this front has not always been satisfactory, and the result is that infrastructure provision has tended to be in conflict with policy objectives, although, it would appear from spending priorities that this may be slowly changing.



## Chapter 3

# Transportation and Road Safety

Countries in the developing world face particularly arduous transportation challenges, and South Africa is no exception. South Africa, however, is different from its peers in certain respects. Transport planning in South Africa has followed an interesting trajectory as a result of the apartheid ideology of the previous regime, and the country's recent transformation to democracy. This poses planners here with specific challenges, but also presents unique opportunities to exploit. Some of these challenges and opportunities for transport planning in South Africa are explored in the following sections.

### 3.1 Road Accidents and Driver Behaviour

In the developing world, where the first and the third world interface, the problems brought about by inappropriate transport planning are amplified. According to Iaych et al. (2009), South Africa has the third highest number of road deaths per 100 000 population in the world. In South Africa, pedestrians make up the bulk of road accident fatalities (Mabunda et al., 2008).

According to the Road Safety Strategy developed by the South African Department of Transport (NDoT), on the order of 95% of road traffic accidents happen as a result of one or more traffic offences committed (NDoT, 2006b). By implication, therefore, in South Africa, 95% of all traffic accidents happen as a result of some measure of driver negligence or ignorance. This statistic, however, also speaks to the efficacy of traffic policing in South Africa (Mohammed and Labuschagne, 2008), something that is regularly highlighted as being one of the major reasons for the high accident rate (see Castle and Kamy-Lukoda, 2006; Matzopoulos et al., 2008; Mohan, 2008; Nyarirangwe, 2009). A comparison between offence rates for 2005 and the target year 2010, compiled by the NDoT, is presented in Table 3.1. The table shows the scale of the road safety challenges faced by traffic authorities. Comparing the 2005 statistics against the target rates for 2010 also shows the extent of the ambition of the NDoT in addressing the problems caused by lawlessness.

According to the Road Traffic Management Corporation (RTMC (2009)), based upon the 2006 Millennium Development Goals, one of the goals of the 2015 Road Traffic Safety Management Plan is to reduce, by half, the rate of accident fatalities arising from road and other transport by 2015. Figure 3.1 shows that, despite the ambitious targets set for reducing road traffic offences, success on the scale required to achieve this goal has been illusive. Although there is a slight overall decrease in fatalities over the three year period between 2006 and 2009, notwithstanding a dramatic change in the status quo, these decreases are unlikely to be sufficient to achieve the goal of a 50% drop in the accident rate by 2015.

**Table 3.1:** South African traffic offences rates for 2005

<b>Offence</b>	<b>2005 Rate</b>	<b>Target Rate by 2010</b>
<i>Speed:</i>		
Percent of drivers exceeding the speed limit	17	5
<i>Alcohol:</i>		
Percent of drivers exceeding the legal limit	4.3	0.4
<i>Barrier Line:</i>		
Percent of offences per hour per barrier line	0.9	0.1
<i>Traffic Signals:</i>		
Percent of red phase offences	28	5
<i>Seat Belts:</i>		
Percent of vehicle occupants not wearing seat belts		
Drivers:	17	5
Front Passengers:	36	5
Rear Passengers:	97	10
<i>Driving Licence:</i>		
Percent of drivers not holding legal licence	2.3	1
<i>Professional Driving Permit:</i>		
Percent of drivers not holding valid permit	15.6	1
<i>Vehicle Tyres:</i>		
Percent of vehicles with defective tyres	21.3	2.5
<i>Vehicle Lights:</i>		
Percent of vehicles with defective lights	3.5	1

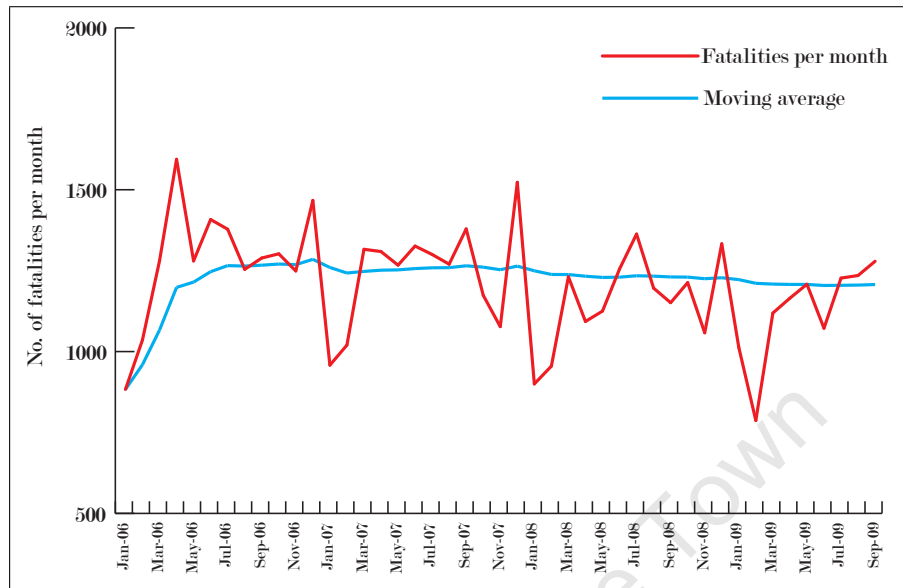
Source: (NDoT, 2006b)

This data shows that, although there is a general awareness that road safety is a major problem in South Africa, the efforts being made to address the problem are not achieving the goals set for them. The problem is twofold; either the goals being set are unrealistic, or the measures put in place to reach the goals are not extensive enough.

The contributory factors to the high accident rate are often split into three categories, namely, human factors, vehicular factors and road factors. As mentioned previously, human factors are often cited as the primary reasons for accidents occurring. Table 3.2 explores these contributory human factors further. According to Table 3.2, jay walking and speeding constitute the largest types of human factors contributing to road accidents in South Africa, being factors in between 35 and 45% of all fatal crashes. This matches well with the fact that pedestrian fatalities account for around 35% of the total fatalities each year, as shown in Table 3.3 (Botha, 2009), since most accidents involving pedestrians would have occurred where pedestrians were jay walking and the driver was speeding.

### **3.2 Policy Responses to Road Accidents: Theory and Myth**

Having identified that road safety is a major problem in South Africa, and having looked at some of the factors that contribute to the problem, what efforts are being made to address the problem, and how



**Figure 3.1:** RSA monthly accident rate and average monthly accident rate Jan2006 - Sept2009. Although there is a slight overall decrease in fatalities, notwithstanding a dramatic change in the status quo, these decreases are unlikely to be sufficient to achieve the goal of a 50% drop in the accident rate by 2015.

*Source:* Adapted from RTMC (2009)

**Table 3.2:** Percentage Contribution to Accidents per Human Factor: 2005 - 2009

Human Factors	2005	2006	2007	2008	2009
Pedestrian: jay walking	44.26	37.69	37.81	32.21	36.06
Speed too high for circumstances	29.51	35.15	36.9	44.25	34.64
Hit-and-run	11.74	8.28	7.2	7.69	9.71
Overtake when unlawful / unsafe	4.09	4.83	5.03	5.55	8.18
Turn in front of oncoming traffic	3.07	4.6	4.82	3.63	3.99
Disregard: red traffic light / stop sign / yield sign	2.13	3.33	3.12	2.74	3.48
Intoxicated Driver	2.03	2.62	2.05	1.9	1.89
Intoxicated Pedestrian	1.28	1.55	1.26	0.58	0.69
Fatigue / Driver fell asleep	1.88	1.96	1.82	1.45	1.36

*Source:* Adapted from Gainewe and Masangu (2010)

**Table 3.3:** Number of Fatalities per Road User Group

User Group	2007-2008	%	2008-2009	%
Drivers	4321	29.4%	3923	28.6%
Passengers	5067	34.4%	4950	36.1%
Pedestrians	5325	36.2%	4833	35.3%
<b>Total</b>	<b>14713</b>	<b>100.0%</b>	<b>13707</b>	<b>100.0%</b>

*Source:* Adapted from Botha (2009)

effective can these be? In this section of the thesis, the policy responses to the road safety problem are assessed in relation to the literature on the subject.

### 3.2.1 Education, Enforcement and Engineering

The National Road Safety Strategy (NDoT, 2006b) puts forward a number of strategies to address these problems, the majority of which fall into the following categories:

1. A general improvement of law enforcement measures.
2. Enhanced road user education campaigns.
3. The expanded implementation of traffic calming schemes.

In terms of the engineering and road design opportunities for improving road safety, the document has the following to say (see NDoT, 2006b, page 35):

*There is not a single site in South Africa where more than 1% of crashes occur. Therefore, even if that site is remediated by engineering methods, only a maximum of 1% of crash reduction will occur. Putting effort into behaviour and attitude change is, therefore, more beneficial.*

*Identification of hazardous locations (stretches of road) are, however, still a priority, so that enforcement activities can be concentrated on those areas, during the most dangerous times of the day, and engineering solutions can be explored.*

Considering that more than 10 000 people are killed on South African roads each year, if there were a single location that accounted for 1% of all of these fatalities, let alone all fatal and non-fatal crashes, this would surely be a major cause for concern, and possibly legal inquiry relating to professional negligence. Surely the benchmark for considering the possibility of a design flaw cannot be that a location must account for more than 1% of all crashes, or even 0.1% of all crashes. These considerations notwithstanding, the statement also dismisses the potential of road design to effect a change in road user behaviour, and to pro-actively limit the number of fatalities that occur. The position implied is that there is very little wrong with the way roads are planned or designed. The problem, instead, lies with the way roads are used (or misused).

This view, however, is in conflict with the reality of both trip making and driver behaviour in South Africa. Walking, as both a main and a secondary mode, is very important in South Africa, and a large proportion of road users walk long distances to reach workplaces, public service centers and PT stops (see NDoT, 2005a). Behrens (2005) notes that both mean and 95th percentile walking trip lengths in South African cities are considerably longer than the maximum walking trip length conventionally

assumed in practice (+800 m). He holds that this contradicts assumptions regarding the localised nature of walking trips within local neighbourhoods (as initially proposed by Buchanan and assumed in the traditional road hierarchy philosophies) because a significant proportion of observed South African walking trip lengths exceed conventional parallel arterial or district distributor frequencies (of 1500 and 2000 m).

**Table 3.4:** Ten worst road sections for pedestrian accidents in Cape Town in 2005

Route Name	Type	Fatalities	Injuries	EAN
Lansdowne Rd - Lansdowne	Arterial	10	208	744
Voortrekker Rd - Cape Town	Arterial	3	101	339
N1 - Cape Town	Freeway	19	17	279
N2 - Cape Town	Freeway	10	51	273
R300 - Delft	Freeway	12	31	237
N7 - Cape Town	Freeway	7	45	219
Mew Way - Khayelitsha	Arterial	5	42	186
Klipfontein RD - Cape Town	Arterial	3	42	162
Halt Rd - Elsies River	Arterial	2	43	153
Main Rd - Delft	Arterial	1	45	147

*Source:* City of Cape Town (2005)

Table 3.4 lists the ten worst road sections for pedestrian accidents by Equivalent Accident Number (EAN<sup>13</sup>), calculated using the World Bank's recommended weighting of 12 for fatality accidents and 3 for injury accidents<sup>14</sup>. All of the ten routes are either arterials or freeways. All the routes carry high volumes of fast moving traffic and all, other than Voortrekker Road, also pass through or beside low income neighbourhoods for substantial sections of their lengths. Considering Behrens' findings, these conditions lend themselves to a high incidence of both jay walking and speeding, helping to explain the large numbers of pedestrians killed and injured along these roads.

Road safety policies cite traffic calming as one of the possible measures to address the problem (NDoT, 2006b). However, municipal traffic calming policies state that traffic calming may only be used on lower order roads, not on arterials or freeways (see City of Cape Town, 2008). Therefore, barring major engineering intervention, already ruled out by the Road Safety Policy, the primary methods available to improve road safety along these routes are escalated education campaigns and improved law enforcement. The question then is, to what extent can these measures be expected to improve road safety?

Zaidel (2002) compiled a review of studies investigating the potential and empirical impacts of enforcement on accident rate. He found that although indications are that extensive enforcement that ensures a near 100% road law compliance rate could lower accident frequencies by up to 50%, the empirical evidence for improvements in the overall accident rate are less dramatic. His review yielded estimates of nearer to a 10% reduction in accident frequency. He also notes that there is no statistical relationship between the number of traffic law enforcement officers and the accident rate, meaning

<sup>13</sup> The Equivalent Accident Number (EAN) is used to compare locations in terms of the number and severity of accidents that took place there. The score is a summation of the number of accidents weighted according to accident severity. Weights are calculated in relation to the generalised cost of accidents of different severity levels.

<sup>14</sup> [http://www.worldbank.org/transport/roads/saf\\_docs/haz\\_locs.htm](http://www.worldbank.org/transport/roads/saf_docs/haz_locs.htm)

that although the presence of law enforcement is important, the absolute number of officers is less important than their operational effectiveness.

Mohammed and Labuschagne (2008) note that, despite the high accident rate, the number of traffic law enforcement officers in South Africa is actually comparable to that of many European nations, and so an increase in the number of officers patrolling the streets is not necessarily going to lead to a proportional decrease in the accident rate. They suggest that improvements in the judicial system that supports the operations of traffic law enforcement would do more to improve the public perception of the risk of prosecution, and consequently help to lower the incidence of lawlessness, thereby leading to a lower accident rate.

Although education programs are commonly listed as one of the ways of promoting road safety, the effectiveness of these programmes in reducing accident rates is the subject of much debate. In fact, several exhaustive studies on the relationship between driver and pedestrian education and road safety have been conducted in recent years, all finding that there is no relationship between education programs and road safety.

Researchers at the John Hopkins School of Public Health (Vernick et al., 1999) recently reviewed nine studies that met their quality criteria and concluded that there is no convincing evidence that high school driver education reduces motor vehicle crash involvement rates for young drivers, either at the individual or community level. In fact, the study finds that by providing an opportunity for early licensure, there is evidence that these courses are associated with higher crash involvement for young drivers.

Christie (2001), in an Australian study, examined the international literature on the effectiveness of driver training as a road safety measure. His focus was broader than the Vernick et al. study on high school based driver education or training programs in that he investigated the safety value of such programs for three distinct groups: learner drivers; young or recently licensed drivers; and experienced drivers. His comprehensive review suggested that, for learner drivers, pre-license driver training contributes little to post-license reductions in casualty crashes or traffic violation rates. In addition, mandatory pre-license training, or even formal pre-license education, such as high school driver education programs in the USA, may contribute to increased exposure-to-risk for young drivers, particularly females, by encouraging early solo licensing. He also found that there is considerable evidence that driver training that attempts to impart advanced skills, such as skid control to learner drivers, may contribute to increased crash risk, particularly among young males.

His review of the evaluation studies of post-licensing training programs for novice drivers also resulted in a similar conclusion, that there is no clear evidence that post-license training for novice drivers leads to reductions in crash or violation involvement. Moreover, he found no sound evidence that either advanced or defensive driving courses reduce the accident involvement of experienced drivers who attend them.

Duperrex et al. (2002) conducted a systematic review of randomised controlled trials of safety education programmes for pedestrians of all ages. They identified 15 randomised controlled trials of safety education programmes for pedestrians. Fourteen of these trials targeted children, and one targeted adults. None of the trials assessed the effect of safety education on the occurrence of pedestrian injury, but six trials assessed its effect on behaviour.

The review found that the effect of pedestrian education on behaviour varied considerably across the studies. They concluded that pedestrian safety education can change observed road crossing behaviour but were unable to say, with any confidence, whether this reduced the risk of pedestrian injury in road

traffic crashes. They found that there is a lack of good evidence of the effectiveness of safety education for adult pedestrians, especially for elderly people. Importantly, none of the trials was conducted in low or middle income countries, where behaviour could reasonably be expected to differ from high income countries.

Behrens (2010) conducted a series of studies between 2004 and 2008 that observed pedestrian crossing behaviour on selected arterials and freeways in Cape Town. The studies investigated whether the frequency of illegal crossings increased with distance from a formal crossing facility. He found that, instead, there was no relationship between the frequency of illegal crossings and the distance of these crossings from the crossing facility. Significant numbers of pedestrians were observed to cross arterials and freeways unassisted at small distances from crossing facilities. While a greater use of freeway crossing facilities was observed, relative to arterial crossing facilities, he posited that crossing behaviour would be better explained by the location of the crossing facilities in relation to dominant pedestrian desire lines.

Behrens's studies also found a relationship between the time a person has been living in the city, the gender of the crosser and the frequency of illegal crossing behaviour. He conducted a roadside intercept survey with crossers, which included pedestrians who crossed illegally at-grade and those who crossed legally on a grade-separated facility. He found that females used grade-separated crossings more frequently than men, and that pedestrians who had lived in a city for longer were less likely to cross at grade than those who had moved to the city more recently.

Behrens makes a number of important conclusions from the findings of the study. He suggests that, since illegal crossing frequency is not necessarily related to crossing facility frequency, crossing facilities should be provided more often and should be better aligned to pedestrian demand lines. Furthermore, he suggests that, in addition to intensified policing, ongoing awareness and education programmes could be employed to accelerate the learning experience and appreciation of traffic risk for newly arrived migrants.

The synthesis of the findings of these studies on the effects of enforcement and education or awareness campaigns on accident rates is that, although each approach has merit and can contribute to lowering accident rates, they are unlikely to have a marked effect on accident frequency without engineering and planning interventions to support them. Policies that rule out the use of engineering or planning interventions on certain stretches of road force an over-reliance on education and awareness programmes and increased enforcement as the only tools with which to confront problems such as speeding and jay walking.

### **3.3 Speeding and Road Safety**

Regulations ensuring low vehicular speeds have a long history. The '*Red Flag Act*' was passed in England in 1865, restricting the speed of horseless vehicles to 4 mph in open country and 2 mph in towns. The act also required that there were three drivers for each vehicle - two to travel in the vehicle and one to walk ahead carrying a red flag. The act was not repealed until 1896, and as could be expected, compliance with its requirements was extremely low (Garder et al., 2002). Regulation and legislation put into effect through enforcement alone is only partially effective. Instead, more persistent measures must be put in place to ensure that rules are not broken. Engineering measures refer to the physical structures put in place to directly, or indirectly, ensure that road rules are adhered to.

As seen in Table 3.2, speeding is a major road safety problem in South Africa. The reasons for speeding have been well researched. Greaves and Ellison (2010), in Australia, compared self-reported driver

speeding characteristics to measured speeding characteristics in the context of personality profile (e.g. aggression, thrill-seeking, altruism) and aversion to risk. They note that while crash information, enforcement records and self-reported driving behaviour provide an indication as to the prevalence of speeding across drivers, they are limited by:

- The number of measuring locations.
- The inability to monitor the same drivers in multiple situations across time.
- The lack of detailed data on the frequency of speeding by driver.
- The perceived under-reporting of speeding when using self-reporting surveys.
- The limited possibility of collecting demographic and psychological data about drivers recorded using speed cameras.

The development of GPS technology has allowed continuous monitoring of driving behaviour, and so has enabled studies of the sort undertaken by Greaves and Ellison. They found that personality correlations with self-reported speeding behaviour suggests that aggression, excitement, and efficacy are all associated with more speeding, per se, while aversion to risk is associated with less speeding. However, they also found that different personality traits drive speeding in different speed limit zones. Whereas aggression and efficacy drove speeding in the lower speed limit zones, excitement was the primary driver in higher speed limit zones. This indicates that where speed limits are lower, such as in residential areas, drivers speed because they perceive the limit to be too low, and believe that they can control the vehicle safely at a speed higher than the posted limit. They found that overall, 19% of the total distance driven was spent above the speed limit. However, they note that this result disguised the substantial heterogeneity in speeding, there being a small but notable number of drivers who regularly exceeded the speed limit by large magnitudes.

This startling result confirms suspicions that speeding is, indeed, a very common behaviour. Greaves and Ellison specifically elected to measure the distance covered above the speed limit, and not the time spent speeding since at slower speeds the distance covered is less than at higher speeds, for the same amount of time.

In a similar study, Ogle (2005) collected individual driver speeding behaviour data from 172 instrumented vehicles from the Commute Atlanta program. The results from her investigations were even more startling. She found that, on average, nearly 40% of all driving activity by the sample population was above the posted speed limit. According to her results, the amount and extent of speeding was highest for young drivers and that speeding behaviour decreased in amount and extent as age increased.

Speeding, therefore, is very often a causal factor in accidents, and is very prevalent amongst drivers. Bester and Makunje (1998) note that the injury accident rate in South Africa is 36,5% higher than that in the USA, whereas the fatality rate is about 1000% higher indicating that, although the number of accidents in South Africa is problematic, the seriousness of the accidents is even more so. They contend that one of the reasons for the high fatality rate is the high number of pedestrian accidents on South African roads.

What then are the engineering measures that can be used to address problems such as speeding and jay walking? As could be expected, these vary according to the type of road in question and its operating conditions. Freeways require different measures than arterials and other lower order roads.

Martens et al. (1997) compiled a literature review of the efficacy various speed reducing measures. They note that currently the largest reductions in driving speed are achieved with measures that physically restrict driving speeds, such as chicanes and road humps. The use of these measures are clearly not

appropriate under all circumstances and so alternative measures must be employed to make drivers lower their speeds. They argue that by designing a road that provides a speed image that is in accordance with the actual speed limit, drivers will choose the appropriate driving speed more or less automatically.

The measures they list range from the relatively straightforward to the unconventional, even controversial. They note that driving speed can be reduced by providing a warning signal to drivers, that alerts them for an upcoming dangerous location. This can be done by using transverse road markings or transversely placed rumble strips. If the markings or rumble strips increase in number while approaching the dangerous location, they find that this usually leads to an extra reduction in speed, since it creates the illusion of acceleration. They caution that the effect strongly decreases over time and may even disappear.

More controversially, they contend that another way to reduce driving speed is to decrease visibility distances. They argue that, in this way, the drivers uncertainty is increased and, in order to achieve better anticipation, he has to slow down. This decreased visibility distance can be achieved by increasing the amount of curvature, rising and falling gradients, buildings and overgrowth. They do mention some of the disadvantages of these measures that should be taken into account before using them. They note that driving safety could decrease when people do not reduce their speed as much as required. Therefore, it would be wise to combine measures like these with, for instance, road markings or transverse rumble strips to warn drivers to slow down.

Furthermore, they list decreasing driving comfort for higher speeds as a speed reduction measure. Road pavement treatments use this principle as well, but the problem with such measures is that they also increase driving discomfort for relatively low speeds. They note that the best way to apply the principle of decreasing comfort would be to decrease driving comfort only for drivers that are actually exceeding the speed limit. Rumble strips, placed longitudinally on the road, are rather effective in reducing driving comfort for high speeds. Drivers will try to avoid these strips, something that requires more accurate lane keeping, which is usually only possible with decreased driving speed. The rumble strips should be chosen in such a way that avoiding the strips is possible when driving at the posted speed limit. Other measures that influence driving speed via driving comfort are roughness of the road surface and the amount of peripheral information. If the road surface contains a certain level of roughness, higher speeds will lead to more vibration and noise. Increasing the amount of information in the visual periphery, so that angular velocity in the peripheral field of view exceeds 2 rad/s, will lead to speed reductions, since drivers will try to avoid such values due to the experienced visual discomfort.

Most road design adaptations lead to the best speed reducing results if they are combined with other adaptations in road design. In this way, speed reductions can be larger since the measures work in the same direction. By providing drivers with the idea of an increased risk for high speeds, driving speeds can also be reduced. Preferably, perceived and actual risk should be related to one another. Reducing road width requires accurate steering behaviour and increases the perceived risk of running off the road or hitting other vehicles. Placing obstacles alongside the road works much in the same way, increasing the risks of running off the road and requiring improved lane keeping. The problem with these measures is that they also increase the risk and severity of accidents involved with high speeds, which might reduce the positive effect of speed reductions. Therefore, extra precautions should be taken to minimise the consequences in case a driver actually leaves the road or hits an object, or to influence the perceived risk only, for instance by reducing lane width without reducing pavement width.

Macaulay et al. (2004), in Australia, investigated the effects of two kinds of these 'perceptual measures' on speed reduction around curves and at intersection approaches. Around curves, laterally diverging

timber posts were installed on the outside of the curve that increased in height from the beginning of the curve to the centre of the curve, and decreased again to the end of the curve. At the ends of the curve the posts were placed around 1,2 m from the road edge, moving to 3,0 m from the road edge at the centre of the curve. This treatment makes the curve appear sharper than it actually is, in theory causing drivers to use greater caution when navigating through it. At intersection approaches peripheral transverse lines were painted with high contrast yellow paint in the shoulder, starting approximately 400 m from the intersection. Three treatment sites for each measure were selected in Melbourne and Sydney, and these were compared to three control sites for each measure in each city, bringing the total number of sites evaluated to 24.

They found that at both the curve and intersection treatment sites the measures used were not conclusively effective in reducing average speeds. The results they gathered indicated that, while at some sites speed reductions of between 5 and 8 km/h were obtained, at others no effect could be measured, and in some instances increases in speeds were observed. Their study also evaluated the efficacy of the measures over the long term. They collected results at all the sites 12 months after installation. They noted that long term effects could be observed at certain sites, but that these were generally of a lower order than the initial results.

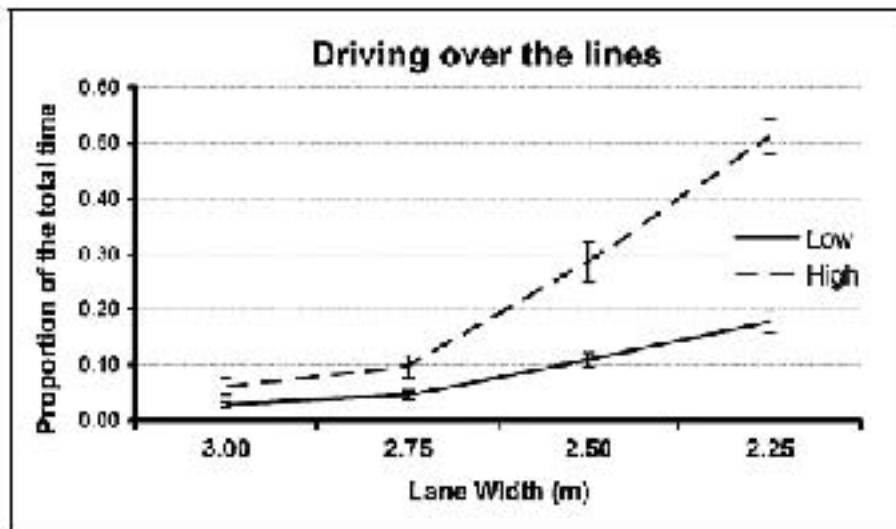
They also note the relationship between lane width and operating speeds, citing Yagar (1984) and Jacobs and Sayer (1983). Decreasing lane widths tend to lead to decreasing vehicle speeds, although the effect is not always straightforward. They cite Van Der Horst (1983) who observed an increase in operating speeds, at the study location, after the lane width was decreased from 4.6 m to 3.6 m. In this instance, the lane width was effected by the introduction of a central kerbed median, thus providing improved visual guidance, as well as physical clearance from oncoming traffic, assisting the driver to better maintain higher driving speeds. They also mention that decreasing lane widths should only be considered when it will not increase the accident rate.

The FHWA, on their website, lists the operational effects of decreasing lane widths (see Table B.1). They note that a 0.6 m reduction in lane width can lead to a 10.6 km/h reduction in free-flow speeds. According to their findings, the interaction of lane width with other geometric elements, primarily shoulder width, also has an effect on free-flow speeds (see Table B.2). Decreasing the shoulder width also leads to a reduction in free-flow speeds.

Dijksterhuis et al. (2010) used a driving simulator to determine changes in mental effort amongst 30 study participants in response to manipulations of steering demand. He measured changes in mental effort by using subjective effort ratings, physiology, and the standard deviation of the lateral position. Steering demand was increased by exposure to narrow lane widths and high density oncoming traffic, while speed was fixed in all conditions to prevent a compensatory reaction.

They found that for increasing levels of lateral demand, extra effort was required for steering, indicated by a decrease in the standard deviation in the lateral position of the vehicle in the lane. In other words, since participants were forced to focus more on the steering task, there was less deviation from the ideal driving line amongst the participants. No clear trend was observed for lateral displacement of the vehicle as a result of lane width variations, although an increase in oncoming traffic was associated with a position to the right of the lane centre<sup>15</sup>. Although lateral control, as measured by the standard deviation in the lateral position of the vehicle in the lane, improved for every level of steering demand, subjective ratings were only sensitive to different levels of lane width under conditions of high demand. Of particular interest is his assessment of time spent driving over the lines (see Figure 3.2).

<sup>15</sup> Since this is a Dutch study, participants drove on the right hand side of the road.



**Figure 3.2:** Time spent driving over lines. As lane widths decrease to approximately the width of the vehicle, an ever increasing amount of time is spent crossing over the lines, and the effect is even more pronounced at higher speeds.

*Source:* Adapted from Dijksterhuis et al. (2010)

As lane widths decrease to approximately the width of the vehicle, an ever increasing amount of time is spent crossing over the lines. Since the subjects were not allowed to decelerate, their only option to maintain a comfortable distance between themselves and oncoming traffic was to spend an increased amount of time driving onto the shoulder. Given that this ‘evasive’ manoeuvre is still not ideal, it seems reasonable, therefore, that given the option, participants would have lowered their driving speeds to compensate instead.

Reduced lane widths, however, are also commonly associated with increased accident rates, with many studies investigating the effect of lane width, shoulder width and combinations of the two on accident rates under different circumstances. Bester and Makunje (1998) summarised the findings from three studies conducted in Southern Africa (two in South Africa and one in Malawi) analysing the effect of road geometry on accident rates on rural roads. They found that an increase in lane width, shoulder width and traffic volumes would lead to a decrease in the accident rate. Also, as could be expected, more adverse terrain will result in higher accident rates.

In the USA, Goldstine (1991) conducted a study to assess the effect of shoulder and road widening on accident rates. Twenty-five projects covering 152 miles of road were selected for analysis. The sampled roads had been widened to one of four widths: either 32, 36, 40, or 44 feet. Accident rates were compared before and after the construction period. Reductions of between 38 and 53% were observed in accident rates, although the amount of reduction varied with traffic volume and the roadway width after construction. There is a great deal of evidence supporting these results. Similar findings were made by Belmont (1954); Cope (1955); Council and Stewart (2000); Dart and Mann (1970); Heimbach C. L. and Chang (1983) and Hadi et al. (1995). There are, however, a few important caveats that should be noted with regard to these results.

Despite the large volume of work done on the subject of lane and shoulder widths and their effect on road safety, the bulk of the studies pertain to two-lane rural roads. Little is known about the effect of lane width on multilane roads or urban roads. Those studies that have concentrated their efforts on

urban arterials have almost invariably not managed to find any significant relationship (see Harwood, 1990; McLean, 1997; Potts et al., 2007). This is because in urban areas, and especially on urban arterials, there tends to be many more confounding factors that cannot be accounted for in the analysis. Furthermore, when road sections differ in lane width they tend to differ also in other important respects, especially in urban areas. This makes the isolation of the safety effect of lane width difficult.

### 3.3.1 Speed Limits and Road Safety

Central to the complexities surrounding speeding on public roads is the question of speed limits. Speed limits on public roads directly influence the mean speed on the link and dictate the operating speeds of the link.

Speed limits are primarily determined by the geometric alignment of the road, but a number of other factors play an important role when the posted limits are set. These include, but are not limited to:

- The number of lanes.
- Whether it is a divided or undivided road.
- The presence or absence of shoulders.
- The driver visibility (e.g. line of sight for cars entering the road at much lower speeds).
- The superelevation in curves.
- The number of and radius of curves.
- The number of access points.
- The expected volume of traffic.
- The location of the road (e.g. urban in neighbourhood with children, limited access without pedestrians or slow vehicles).

The setting of speed limits in urban environments can become a contentious issue. In certain instances, the decision may be influenced by political concerns. Inappropriate speed limits run the risk of low driver compliance and of losing overall driver support. Additionally, they can create dangerous situations arising from driver frustration.

However, speed limits undoubtedly have an effect on traveling speeds, and since traveling speeds are related to accident frequency, speed limits affect the number of accidents that occur on a road. Research conducted by the CSIR in South Africa from the mid 1970s to the mid 1980s found that lowering the speed limits on the rural road network (resulting in lower operating speeds) had a significant effect on the occurrence of road accidents. A reduction in the speed limit from 120 km/h to 80 km/h resulted in a decrease in the casualty crash rate (the number of casualty crashes per million vehicle kilometres travelled) from about 0,59 to about 0,44<sup>16</sup>.

Stuster et al. (1998) conducted a review of the available literature on the effect of lowering and raising the speed limits on speeds and accident frequency for the FHWA in the USA. Tables 3.5 and 3.6 provide the list of results they compiled.

The effects listed in Table 3.5 indicate quite strongly that there is a link between lowering speed limits and decreasing the accident rate. In Table 3.6 the results from studies assessing the effects of increasing the speed limits, as compiled by Stuster et al. (1998), are listed.

<sup>16</sup> Arrive Alive South Africa - Speed Factor. Retrieved January 17, 2011, from <http://www.arrivealive.co.za/pages.aspx?i=1670>

**Table 3.5:** Summary of the effects of lowering speed limits

Reference	Country	Change	Results
Nilsson (1990)	Sweden	110 km/h to 90 km/h	Speeds declined by 14 km/h; Fatal crashes declined by 21%
Engel (1990)	Denmark	60 km/h to 50 km/h	Fatal crashes declined by 24%; Injury crashes declined by 9%
Peltola (2000)	UK	100 km/h to 80 km/h	Speeds declined by 4 km/h; Crashes declined by 14%
Sliogeris (1992)	Australia	110 km/h to 100 km/h	Injury crashes declined by 19%
Finch et al. (1994)	Switzerland	130 km/h to 120 km/h	Speeds declined by 5 km/h; Fatal crashes declined by 12%
Scharping (1994)	Germany	60 km/h to 50 km/h	Crashes declined by 20%
Newstead and Mullan (1996)	Australia	5-20 km/h decreases	No significant change (4% increase relative to sites not changed)
Parker (1997)	USA (22 states)	8-32 km/h decreases	No significant changes

Source: Adapted from Stuster et al. (1998)

Once again, the effects appear to be quite conclusive: increasing speed limits seems likely to lead to an increase in the accident rate. Moore et al. (1995), in Australia, examined the relationship between the speed of passenger cars and risk of involvement in a severe crash, in an urban setting, using a case-control study. They considered the cases of 45 accidents involving vehicles in severe crashes in the Adelaide metropolitan area and determined their pre-crash speeds using accident reconstruction techniques. For each case they measured the speeds of 10 controls using an amphotometer<sup>17</sup>. The controls used were cars not involved in crashes that passed through the crash location at the same time of day, day of the week, and season. They found that the risk of involvement in a severe crash increased as vehicle speeds increased. Within 60 km per hour zones, compared with vehicles travelling at about the posted limit, vehicles travelling at 75-84 km/h had an odds ratio of 7.8 (95% confidence interval (CI) = 1.4-38.8) for a severe crash, whereas vehicles with speeds in excess of 84 km/h had an odds ratio of 39.0 (95% CI = 9.3-170.5). Elvik (2010) notes that:

*“Speed is regulated by means of speed limits in all highly motorised countries. Speed limits would not be needed if drivers were able to choose speeds that are optimal from a societal point of view without the guidance given by speed limits. Thus, the rationality of driver speed choice is an important criterion for assessing whether speed limits are justified or not in terms of normative welfare economics. Speed limits would not be justified*

<sup>17</sup> A device that is used to measure the passage of a vehicle.

**Table 3.6:** Summary of the effects of increasing speed limits

Reference	Country	Change	Results
NHTSA (1988)	USA	89 km/h to 105 km/h	Fatal crashes increased by 21%
McKnight and Klein (1990)	USA	89 km/h to 105 km/h	Fatal crashes increased by 22%; Speeding increased by 48%
Garber and Graham (1990)	USA (40 States)	89 km/h to 105 km/h	Fatalities increased by 15%; Decrease or no effect in 12 States
Streff et al. (1990)	USA (Michigan)	89 km/h to 105 km/h	Fatal and injury crashes increased significantly on rural freeways
Pant et al. (1992)	USA (Ohio)	89 km/h to 105 km/h	Injury and property damage crashes increased; but not fatal crashes
Sliogeris (1992)	Australia	100 km/h to 110 km/h	Injury crashes increased by 25%
Lave and Elias (1994)	USA (40 states)	89 km/h to 105 km/h	Statewide fatality rates decreased 3-5%; (Significant in 14 of 40 States)
Parker (1985)	USA (Michigan)	Various	No significant changes
Newstead and Mullan (1996)	Australia (Victoria)	5-20 km/h increases	Crashes increased overall by 8%; 35% decline in zones raised from 60-80
Parker (1997)	USA (22 states)	8-24 km/h	No significant changes

Source: Adapted from Stuster et al. (1998)

*if driver speed choice was perfectly rational and resulted in outcomes that are optimal from a societal point of view."*

As regards the possibility to influence drivers to adhere to speed restrictions, Goldenbeld and van Schagen (2007) conducted a study that tried to identify factors to operationalise the concept of credible speed limits and to make some progress to describe these in such a way that road authorities can put the concept into practice. Their study focused on the credibility of an 80 km/h limit for different rural roads, and assessed the effects of various characteristics of the road and its environment, as well as the effects of personality characteristics in how respondents perceived these. One of the outcomes from this study was that the authors found that credibility of a speed limit is influenced by identifiable features of the road and its surroundings. They, therefore, conclude that it is indeed possible to make a limit more credible by either fitting the limit better to the features, or fitting the features better to the limit, depending on road function and the desired safety level required.

Retrofitting infrastructure to improve safety levels is common practice around the world. When the aim of the retrofit is to lower vehicular speeds or volumes, this type of retrofit is commonly referred to as

'traffic calming'. In the next section, some of the details of traffic calming, and the potential it holds for improving road safety, are discussed.

### 3.4 Traffic Calming and Road Safety

As mentioned, one of the most successful ways to combat excessive speeding, and a method that is common across the world, is to implement some form of traffic calming. In South Africa, the National Guidelines for Traffic Calming (Schermers and Theyse, 1996) defines traffic calming as:

*Traffic calming has the objective of moderating traffic behaviour through physical and legislative measures aimed at the reduction of either vehicle speeds and/or traffic volumes in order to improve traffic safety and quality of life in the built environment, but with due regard to mobility and accessibility, so as to ensure a balance between the environment and traffic.*

The Institute of Traffic Engineers (ITE) (Lockwood, 1997) defines traffic calming as:

*Traffic calming is the combination of mainly physical measures that reduce the negative effects of motor vehicle use, alter driver behaviour and improve conditions for non-motorised street users.*

Russell (1990), in an editorial in the *Town Planning Review*, listed the following as the objectives of traffic calming:

1. To improve road safety.
2. To reclaim space for pedestrian and 'non-traffic' activities.
3. To improve pedestrian mobility and reduce traffic barriers.
4. To promote feelings of greater security, in particular amongst residents, pedestrians and cyclists.
5. To create an improved environment.

The list, although probably not exhaustive, in making reference to the indirect aims of traffic calming certainly captures some of the spirit of traffic calming, despite not making specific reference to the direct objectives, such as lowering speeds and decreasing volumes. What is immediately apparent are the parallels that can be drawn between the aims of traffic calming and some of the South African transport policy objectives discussed in Chapter 2. This would imply that methods and principles of traffic calming are of particular relevance to the South African situation.

The history of traffic calming is commonly traced back to the development of so-called 'woonerven' in the Netherlands. In the 1960s, residents in the Dutch city of Delft began to fight cut-through traffic in their areas by converting them into woonerven or 'living-yards'. These were highly altered street environments that turned streets from channels for the movement of cars into shared spaces that emphasised the equal rights of pedestrians (Ewing, 1999).

Woonerven employed many specific design features to achieve the desired effect. The features that typify woonerven are:

1. The integration of roadways and footways.
2. The narrowing of roadway widths to slightly more than a single lane, incorporating widenings every 30-40 m to enable vehicles to pass.
3. The use of physical speed reduction measures (e.g. tables, surface changes, chicanes) and shifted horizontal roadway alignments every 30-40 m.

Extensive use of vegetation, creating parking bays that jut out into the street and limiting the line of sight of the driver by creating sharp chicanes while the street furniture forced vehicles to slow down to walking pace in order to avoid all the obstacles and enhanced the overall effect (Figure B.1) (Hass-Klau et al., 1992).

However, woonerven were not always the ideal solution. They were only suitable for streets with low traffic volumes and, although popular, were very expensive, costing up to 50% more than normal reconstructed streets (Ewing, 1999). Soon, the Dutch government ran out of funds and started seeking a cheaper alternative that would have similar effects.

After some experimentation, 30 km/h speed limit zones were introduced in the 1980s as a cost-effective alternative to woonerven (SWOV, 2004). Many other European countries also adopted the idea and it soon became almost standard for residential streets. Debate arose around whether the concept is as effective when only speed signs, as opposed to signs along with physical calming measures, are implemented. Opinions are still divided on this issue, but some sort of a compromise has been found. In many towns there is a blanket 30 km/h speed limit in residential areas, with specific problem streets receiving physical calming measures. This intermediate approach between the woonerf and the 30 km/h zone allows much larger areas to be treated than before.

Germany started experimenting with traffic calming in the 1970s. They soon learnt that calming individual streets only resulted in increased volumes in neighbouring streets. The only approach was to consider the entire area, including the collector streets. This area-wide approach was tested with positive results being obtained. This led the way for many other European cities to adopt the area-wide approach. Lately, however, an even broader scale view of traffic calming is being adopted in German towns and cities, with city-wide vehicular restraint programs being adopted.

Buchanan's *Traffic in Towns* is often credited with starting the modern traffic calming movement. However, the British were slow to adopt the innovations made by their European neighbours, citing their good safety record as making traffic calming unnecessary. Buchanan's ideas were implemented all across Britain up until the 1980s, with street closures and one-way streets enforcing volume controls in problem areas. The results were, however, not as impressive as what was being achieved in European calming projects, and it took changes in laws and regulations, along with a new street design manual, to bring the UK in line with European practice. Current regulations are much more liberal, allowing for almost any sort of measure to be implemented and changing the hierarchy of roads as well.

Other countries like Australia, the USA and Japan, have also been very actively incorporating traffic calming into their new developments, as well as redeveloping older areas with much success. Lessons to be learnt from these countries and their European counterparts include:

1. Traffic calming must happen from the top down and supportive legislation must be in place to ensure the successful implementation of traffic calming projects.
2. Rigorous before and after studies must be done to ensure continuing understanding of the effects of different measures.
3. Traffic calming must be approached on an area-wide basis, incorporating higher order roads as well, to avoid excessive congestion developing in nearby areas.
4. Area-wide projects must be tested before they are permanently implemented to ensure that the best solution is put in place.
5. Public support for the project must be established before it can be implemented.

6. Emergency services and PT services must be consulted to address their concerns regarding accessibility.

There are many tools available to the designer when considering which traffic calming measures to implement. The list is so extensive that it would be impossible to provide a complete index of the measures used around the world. Each country develops its own traffic calming policy and, as such, would have its own design codes and regulations. Although these are, in many instances, very similar, the detailed specifications vary from country to country and, in fact, from project to project.

Each measure is more suited to a specific design objective, and in most cases there are wide ranging cost implications involved as well. However, the different calming measures can be broken up into a number of categories. Examples of the most common measures are given below, separated into five categories for simplicity:

1. **Vertical Deflections:** Speed humps, Speed tables, Raised intersections

These types of measures consist of raised sections of the carriageway and are designed to physically enforce reduced speeds, but are used to discourage cut-through traffic as well. In some countries they are named by their specific design speeds e.g. 55 km/h speed humps, or they may be referred to by their height or width e.g. 14' humps. As is implied, many different designs have been developed for different situations. In South Africa, most speed humps are round-topped asphalt designs, with different paintwork patterns to warn motorists of the hazard. For this reason they are also often accompanied by hazard signs to alert motorists of their presence. Other types of humps include flat topped humps, which are often surfaced with paving blocks or bricks, and an assortment of patented designs manufactured privately off-site and purchased by the developer.

2. **Horizontal Deflections:** Chicanes, Circles

As the name implies, these types of measures force vehicles to shift their travel paths horizontally and are used either as speed controls or volume controls. Traffic circles also play an important role in regulating and directing traffic flows. Traffic circles are gaining popularity in South Africa as alternatives to stop streets and traffic lights, although chicanes are very rarely found. Abroad, chicanes come in a number of forms, primarily being simple kerb extensions, either angular or rounded, and may include side-channels to assist drainage. Circles are found in a number of forms as well, but can be roughly divided between mountable or un-mountable, the former often being referred to as mini-circles. Their size depends primarily on the volumes that they are required to handle and the geometric characteristics of the road.

3. **Narrowings:** Neckdowns, Chokers, Centre island narrowings, Bulbouts

Narrowings are commonly found in conjunction with other measures, especially vertical deflections. They are sometimes used at pedestrian crossings or around schools and at intersections or gateway areas. They serve a number of purposes, amongst others, to reduce speeds and to increase driver awareness and to discourage cut-through traffic by reducing the lane widths. They are commonly constructed as simple kerb extensions and may include side-channels to facilitate drainage.

4. **Closures:** Full closures, Partial closures, Diagonal diverters

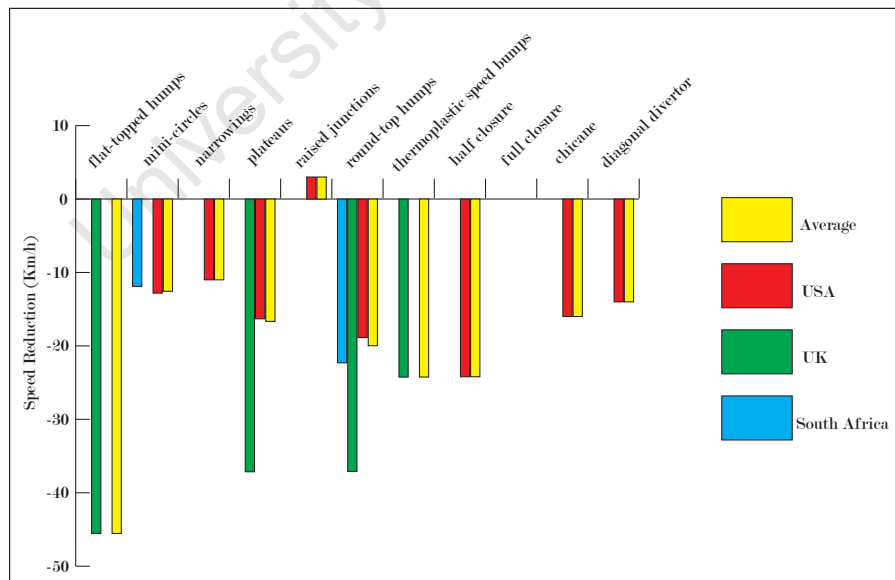
Closures are amongst the earliest calming measures used. Their primary function is to reduce volumes of traffic and, consequently, they are found almost exclusively in residential areas. Full closures block flows in both directions, whereas partial closures block flows only in one direction. Diagonal diverters are used to channel traffic into an adjoining street, and serve similar purposes as closures. Although closures have been used quite frequently in South Africa, the

use of diagonal diverters is still limited. Closures are also not very popular with residents, or the emergency services, because they can drastically reduce accessibility. Extensive public participation is often required to obtain permission to proceed with these sorts of projects.

##### 5. Surface Treatments: Paint work, Coloured pavers, Rumble strips

Painted islands, chokers, narrowings and crossings are all extensively used in South Africa. They are primarily used as regulatory devices instead of calming devices. They are cheap to install, although they have to be maintained and their effectiveness has been questioned. Residents also may see them as unsightly. Coloured pavers are often used in conjunction with other measures and are another method used to increase driver awareness. They highlight the presence of other physical measures and suggest that caution must be taken because pedestrians use the area. They also form a major part of urban renewal schemes that often go hand in hand with traffic calming schemes. Although they are effective, if used correctly, they do increase project costs significantly. Rumble strips are used primarily to alert drivers of approaching hazards, and to encourage drivers to slow down. They have been used extensively in South Africa, especially on high-speed roads before unexpected traffic lights or intersections.

The various types of traffic calming measures have wide ranging effects on speeds and volumes, and a sizeable literature has developed investigating these around the world. Beukes (2002) conducted a review traffic calming practice in South Africa, focussing on identifying the extent to which different traffic calming measures are used in South Africa, their approximate costs and their effectiveness at lowering speeds and volumes, in comparison to the rest of the world. He found that, in general, traffic calming devices' performance was comparable to that in the USA, and that certain devices, most notably vertical deflections, tended to have more dramatic effects in the UK than in South Africa and the USA (Figure 3.3). This almost certainly is a result of different design specifications around hump profile and spacing.

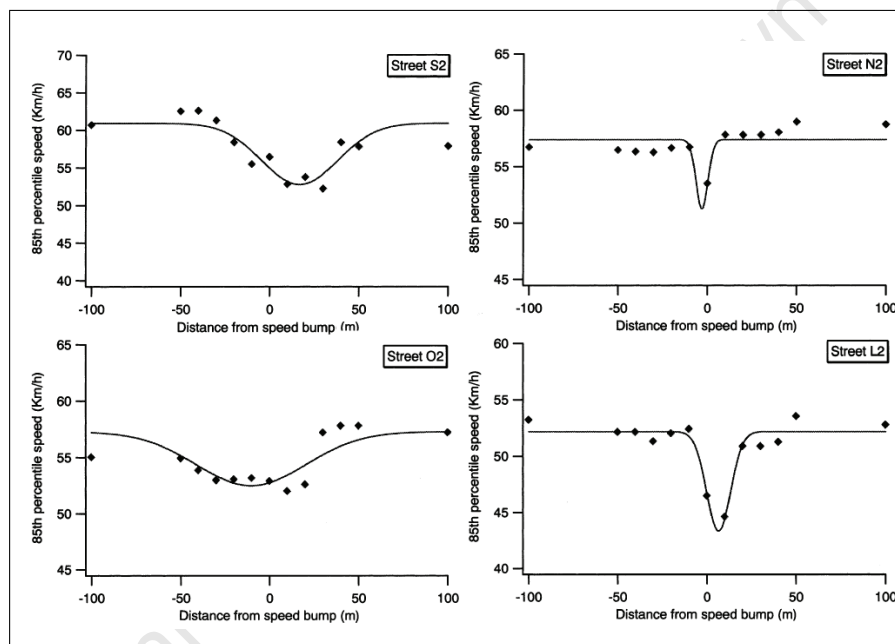


**Figure 3.3:** Average percentage speed change in the USA, the UK and South Africa for different calming devices. The most dramatic speed reductions are generally achieved for vertical deflection devices such as speed humps and plateaus.

*Source:* Adapted from Beukes (2002)

In general, the most dramatic speed reductions are achieved through the use of vertical deflection devices such as speed humps and plateaus. This result is confirmed by numerous other sources (see Ewing, 1999; Hummel et al., 2002; Jobanputra, 2010; Tester et al., 2004).

Pau and Angius (2001), in Italy, investigated the effectiveness of speed humps in lowering vehicle speeds. Their results highlight an important cautionary note with regards to the use of speed humps: the effects of these types of measures tends to be very localised. The study measured speeds in the vicinity of locations where rubber speed humps were installed to protect crossings. The humps were produced by 3M<sup>®</sup>, and were single modules of black vulcanised rubber covered with stripes of a yellow high-reflective tape which must be clamped together to compose the entire undulation. Fixture to the road surface is made using steel nails. The modules are 30 mm high and 600 mm wide with a circular profile and are suitable for streets where the posted speed limit is 50 km/h. They present speed profiles across humps at four study locations (Figure 3.4).

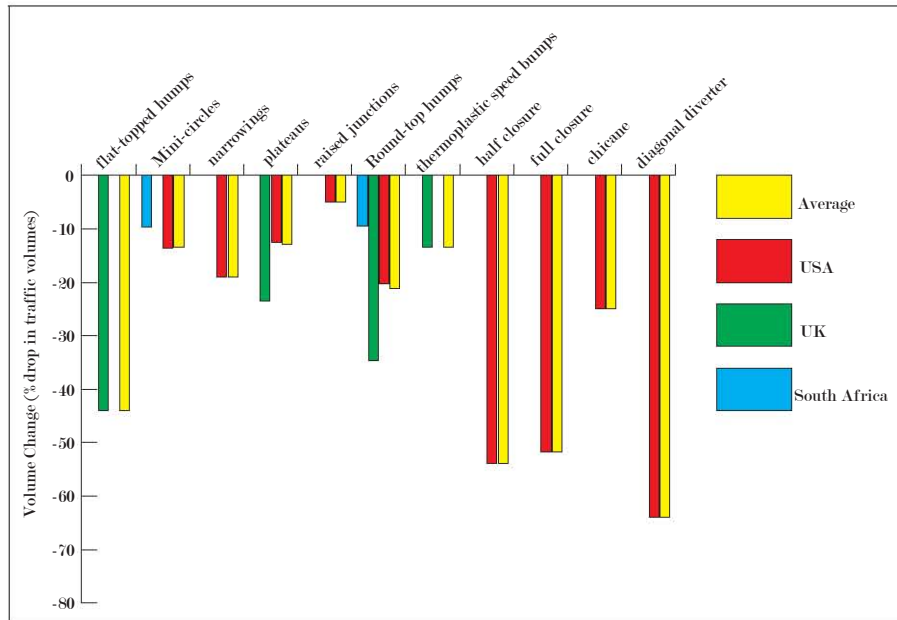


**Figure 3.4:** Speed profiles recorded in four streets in Italy over a 200 m range (100 m before and 100 m after the speed bump position). The effect of the humps is extremely localised. Speeds return to the before state within approximately 50 m of the humps and, in some instances, even less.

*Source:* Taken from Pau and Angius (2001)

The localised nature of the effect is clear. Speeds return to the before state within approximately 50 m of the humps, and in some instances even less. The authors also noted that speeds across the humps themselves are abnormally high compared to similar studies conducted previously, averaging around 50 km/h. The efficacy of these particular hump designs must, therefore, be questioned, and could also help explain why the effects of the humps were so localised. This raises another interesting point, especially in South Africa, where humps tend to be constructed using asphalt. If the hump profile, either by design or as a result of construction methods, is inadequate, the effect of the hump could be greatly diminished or, alternatively, overly severe. Special care must be taken not only in the selection of a profile, but in the overall planning and construction of the scheme to ensure the device acts as intended.

With regards to the effects on traffic volumes, closures naturally have the most dramatic effects, however, evidence from the UK indicates that flat-topped humps also yield significant reductions in traffic volumes<sup>18</sup>. Once again, the US and South African results are more closely matched to each other than to the UK results (Figure 3.5).



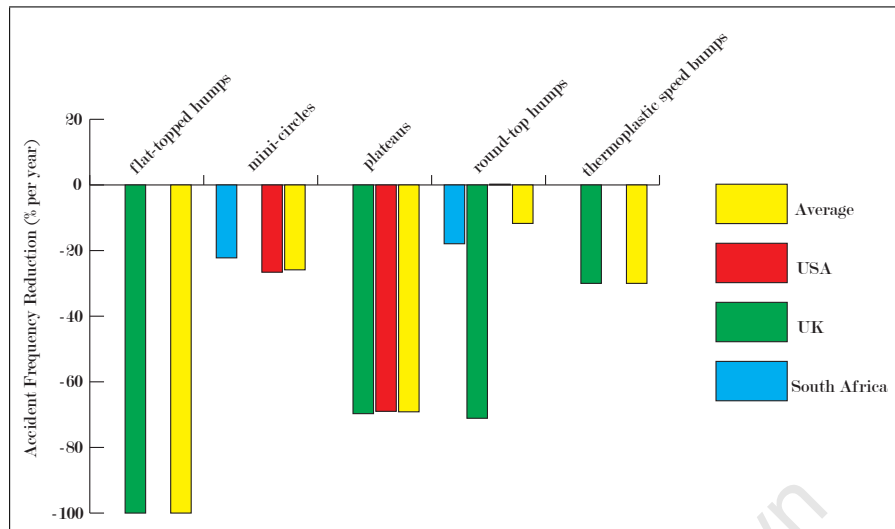
**Figure 3.5:** Average percentage traffic volume change in the USA, the UK and South Africa for different calming devices. US and South African results are more closely matched to each other than to the UK results.

*Source:* Adapted from Beukes (2002)

As highlighted by Russell (1990), one of the major aims of traffic calming schemes is to lower accident rates, mostly those involving pedestrians and cyclists. In this sense, lowering speeds and volumes is simply a means to an end, although there are other benefits (such as environmental and liveability benefits). Much of the research effort on traffic calming has focussed on the benefits of traffic calming on lowering the frequency and severity of accidents (see Bunn, 2003; Bunn et al., 2003; Elvik, 2001; Pucher and Dijkstra, 2003; Retting et al., 2003; Zein et al., 1997). In general, although the magnitude of the results vary from study to study, the effects have been found to be positive, i.e. traffic calming does lead to a reduction in accidents (see also Figure 3.6).

Many authors highlight the importance of applying traffic calming on an area-wide basis. Much of the research effort has focussed on the effect of area-wide traffic calming schemes on accident rates. Bunn (2003); Bunn et al. (2003); Elvik (2001); Liabo (2003); Morrison et al. (2003); Retting et al. (2003) and West (1998) all found that area-wide traffic calming schemes are effective in lowering the frequency of accidents, especially accidents involving pedestrians. Elvik (2001) found a 15% reduction in the accident rate amongst the studies he reviewed. A similar result was found by Liabo (2003) for streets overall, but in differentiating between residential and main roads the effect was found to be 25% and 10% respectively. Bunn (2003) found that area-wide traffic calming schemes produced an 11% reduction in the rate of road traffic injuries (fatal and non-fatal). Brilon and Blanke (1993) analysed accident data from before and after studies of 'extensive traffic calming' in six towns in Germany in

<sup>18</sup> Although in this data set the UK generally outperformed South Africa and the USA for most devices



**Figure 3.6:** Average percentage reduction in accident frequency in the USA, the UK and South Africa for different calming devices. Flat-topped humps and plateaus were found to be the most effective measure to reduce accident frequency in all the countries.

*Source:* Adapted from Beukes (2002)

the 1980s. They found that for motorbikes, involvement in accidents could be reduced by up to 78%. Additionally, mean reductions of 17% for cyclists and 25% for pedestrians were observed.

Clearly, traffic calming, and especially area-wide traffic calming, hold much promise for lowering accident rates in South Africa, and the use of these devices should be encouraged wherever possible. However, considering the information in Table 3.4, it would appear that most of the effort should be concentrated on higher order roads, since they account for such a disproportionately large number of pedestrian accidents. However, municipal traffic calming policies rule out the use of traffic calming measures on these roads. The question then is, to what extent has traffic calming been applied on higher-order roads, and what were the outcomes of these projects?

Garder et al. (2002) found that there are only a few truly traffic calmed arterials worldwide. They note that in North America, most traffic calming of arterials has occurred as a byproduct of mobility improvements, such as reducing roads from four lanes to three lanes with center turn lanes and constructing roundabouts. They found that the overall effectiveness of arterial traffic calming schemes has been moderate even when there were clear reductions in pedestrian injuries. They come to the conclusion that speed control through other means may, in the long run, be preferable to construction of humps and chokers on arterials, citing public opposition as one of the main reasons for avoiding the approach.

Galante et al. (2010) investigated drivers' speed behaviour in a section of a rural highway crossing a small urban community in the existing scenario without any traffic calming device and in two different design scenarios with traffic calming in the urban community. Two 'gateways' and four 'integrative traffic calming devices' along the route within the urban area were tested using a driving simulator for both directions of travel. The gateways were aimed at slowing down the vehicles entering in the built-up area, while the traffic calming devices were aimed at complementing the gateway effect inside the built-up area. Analysis of simulation results showed a different behaviour of drivers approaching the urban community in the existing scenario and in the design scenarios. In the south direction, mean

speed reduction ranging between 16 and 17 km/h, with 5% level of significance, was observed. In the north direction, mean speed reduction equal to 11 km/h, with 10% level of significance, was observed.

Leden et al. (2006) investigated the effects of traffic calming along an arterial through the community center of Storuman, Sweden, which was reconstructed in 1999 and 2000. Pedestrian walkways, traffic islands, chicanes of a type referred to as 'Danish buns', a roundabout and a two-directional cycle track along the E12 were installed. The purpose of the reconstruction was to improve safety for pedestrians and bicyclists, primarily for children, the elderly and the disabled, and to reduce the barrier effect of the E12 thoroughfare.

In May 2000, the code governing the conduct of drivers at marked crosswalks in Sweden became stricter to improve safety and mobility for pedestrians. Ledén et al. (2006) analysed the combined effect of the reconstruction of the arterial and the change of the drivers code. They found that yield behaviour towards pedestrians changed significantly. The difference was even greater with respect to yielding to child bicyclists - from 6% before to 84% after - even though the code change only related to pedestrians. Measures of speed, behavioural studies, questionnaires, face-to-face interviews and crash data analysis suggest that safety has increased, not only along the E12, but also along adjacent roads.

### 3.5 Résumé

In this chapter it was shown that excessive speed plays a critical role in road safety, both abroad and in South Africa. It was shown that speeding is a very common transgression, and have considered the policies currently in place in South Africa to improve road safety, with specific reference to speeding. Some of the tools available to the designer to address the problem of road safety were then explored, focussing on driver perception, traffic calming and, lastly, the effects of speed limits on road safety.

The conclusions that can be made are that there are a large range of tools available to designers to plan roads with speeds that are appropriate to the circumstances. Furthermore, the design speeds selected must reflect these circumstances as best as possible and, probably more importantly, the final road design must ensure that the operational speeds on the road are in accordance with the design speed. Also, the posted speed limit must reflect the drivers perception of what is a safe and reasonable speed to travel at, and this information must be provided to the driver from the road design itself.

The design speed, operational speeds and the speed limit then are very much reliant on the circumstances along the route. Currently, design speeds are derived from the road classification which, as was shown in Chapter 2, relates more to the functional hierarchy of the road network than to the specific circumstances along the route itself. It must also be considered that these circumstances may vary from place to place along the route and so, by this logic, the design speed should vary in step. However, this does not seem practical under current planning norms. Furthermore, how would one describe these changing circumstances in a manner that can be translated into useful information for the designer, and how would this relate to concepts such as design speed selection? These questions will be further explored in Chapter 4.

## Chapter 4

# Contextualising Road Design

In this chapter the role of context in road design practice is explored. The chapter opens with an overview of current practices incorporating context, such as (Context Sensitive Design (CSD)) and the Complete Streets in the USA, and looks at other attempts at incorporating contextual elements into road planning, such as the ARTISTS project in the EU. A definition for context is then derived from the perspective of this research, and the elements that potentially suit this definition of context are explored, and either selected or rejected, depending upon whether they are sensible under the definition. In this way, a better understanding of how context can be defined, in terms of this research, is developed and this definition can be carried forward as forming the theoretical basis upon which the later chapters are built.

### 4.1 Context Sensitive Solutions

According to the FHWA, CSD (interchangeably referred to as Context Sensitive Solutions (CSS)) is “... a collaborative, interdisciplinary approach that involves all stakeholders to develop a transportation facility that fits its physical setting and preserves scenic, aesthetic, historic and environmental resources, while maintaining safety and mobility” (FHWA, 2007). The New York State Department of Transportation defines CSD as “... a philosophy wherein safe transportation solutions are designed in harmony with the community” (De Cerreño and Pierson, 2004).

The principles of CSD and methods for quantifying and measuring the performance of projects in terms of CSD principles are outlined in seven principles in the NCHRP Report No. 642: *Quantifying the Benefits of Context Sensitive Solutions* compiled by Stamatiadis et al. (2009) for the TRB. Amongst others, it includes the use of interdisciplinary teams, the need to address all alternative modes, the need to maintain environmental harmony and the need to utilise the full range of design choices. CSD is, therefore, mainly a project development and project management approach to use existing design standards to develop socially and environmentally sensitive infrastructure.

During the 1960s the general public began to express concerns about the adverse environmental impacts of mans’ developments on the environment, including, but not limited to road building. In the USA, these concerns culminated in the passing of National Environmental Policy Act (NEPA) in 1969 (United States Code, 1969). From this point forward, roadway design and construction became more than a matter of building the most economic, shortest, widest, or fastest facility. Rather, engineers and planners are now required to consider features and effects such as wetlands, threatened and endangered species, adverse noise, and other environmental considerations (Neuman et al., 2002).

There is also a renewed interest and concern with the cultural, historic, and other values that define a community. Communities have become increasingly more vocal and protective over these assets. Projects that may impact community assets, whether to develop open space, tear down a long-standing building with unique architecture, or build a new road are now increasingly viewed as potential threats.

Public authorities and professional engineers, trained to provide a certain quality of design using traditional approaches, began to run into resistance from the public when projects were perceived as having adverse impacts on the communities through which they passed. Benefits such as faster travel times, greater safety and less delay were no longer widely accepted or perceived as worth the costs in relation to environmental and community disruption. Whereas the public in the past had unquestioningly accepted the proposals of engineering professionals, these were now increasingly being called into question, regardless of how well thought out they may have been. Roads, in particular, despite being broadly recognised as necessary to the public health and economic well-being of a community, were increasingly viewed as permanent intrusions on the landscape.

Defining context remains a challenge, since it refers to all the external influences and factors that play a role in the success of a project - that success itself being dependent upon how it is measured. The various sources reviewed, all from the US, include an array of factors, and there does not appear to be consensus on where the line should be drawn.

In Utah, in the USA, the state Department of Transport includes physical, social, economic, political and cultural impacts in its definition<sup>19</sup>, although it does not give specifics as to what these include, or how to assess these, supposedly relying on literature such as Stamatiadis et al. (2009) to support the application of the policy.

The California Department of Transportation (Caltrans) notes that in the development of transportation projects, the social, economic, and environmental effects of projects must be considered fully along with technical issues, so that final decisions are made in the best overall public interest (Caltrans, 2005). They mention that attention should be given to considerations such as the need for safe and efficient transportation, the needs of low mobility and disadvantaged groups and the costs of eliminating or minimising adverse effects on natural resources, environmental values, public services, aesthetic values, and community and individual integrity.

According to Olszak et al. (2008), the Maryland Department of Transportation, FHWA, and AASHTO co-sponsored a workshop in 2008 called "*Thinking Beyond the Pavement: A National Workshop on Integrating Highway Development with Communities and the Environment*", which refined the definition of the CSD process. The workshop envisioned CSD as follows: "*Context sensitive design asks questions first about the need and purpose of the transportation project, and then equally addresses safety, mobility, and the preservation of scenic, aesthetic, historic, environmental, and other community values. Context sensitive design involves a collaborative, interdisciplinary approach in which citizens are part of the design team.*"

The workshop led to the development of seven principles for CSD. These are:

- The project satisfies the purpose and needs agreed to by the full range of stakeholders. This agreement is forged in the earliest phase of the project and amended as warranted.
- The project is a safe facility for both the user and the community.
- The project is in harmony with the community and preserves environmental, scenic, aesthetic, historic, and natural resource values of the area.

<sup>19</sup> See <http://www.udot.utah.gov/main/F?p=100:pg:0:::::V,T.,144> for further information.

- The project exceeds expectations of designers and stakeholders and achieves a level of excellence in people's minds.
- The project involves efficient and effective use of resources (time, budget, community) of all parties involved.
- The project is designed and built with minimal disruption to the community.
- The project is seen as having added lasting value to the community.

The Minnesota DOT<sup>20</sup> and the Kentucky Transportation Cabinet<sup>21</sup> (KYTC) have developed similar CSD principles. Amongst the key concerns mentioned is the need to balance safety, mobility, community, and environmental goals in all projects, the need to involve the public and affected agencies early and continuously, the need to address all modes of travel and to apply flexibility in interpreting and implementing design standards. They also mention the importance of incorporating aesthetics as an integral part of good design.

The emergence of CSD highlights the growing trend amongst transportation planners and designers to see that the impacts of the infrastructure they build are much broader than simply the transportation function they are planned for. What CSD attempts to do is to ensure that road infrastructure is viewed more holistically, as being part of a more complex system, not apart from it. The road, therefore, has to sit comfortably in its environs, managing to fulfil its transport role and still be sensitive to the particularities of the environment it is in, whether that is an urban centre, a suburban township or a rural hinterland. These sentiments are captured in the various philosophies underpinning the practice, such as emphasising the importance of safety for the user of the road and the community through which it passes and stressing the need for harmony with the environmental, aesthetic and historic resource values of the area.

The need for a more holistic view of the contextual setting of a roads project arose, in part, from the realisation that the majority of urban streets serve multiple roles, having to accommodate the needs of multiple modes of transport and needs related to mobility (through users) and access (local users). As a result, a certain amount of flexibility in design is required to meet all these needs (see FHWA, 1997; Hebbert, 2005).

The mechanics of CSD, however, have proven more difficult to apply. Road planning and design is often typified by a mechanistic process, involving the application of established mathematical relationships and engineering norms (in the form of parameters and coefficients) to site specific data (such as land use and population statistics, traffic counts, geological and hydrological information and geographic data), yielding repeatable results that can be unambiguously interpreted into designs. The philosophies and terminologies of harmony, aesthetics and value are quite alien to the usual practices employed when planning and designing roads.

In response, CSD has been formulated, not as an attempt to create new design standards, but rather to build on the available flexibility in current design standards and guidelines. Training in CSD often employs case studies to showcase project strategies and highlight potential dangers or pitfalls in the implementation of the principles, but also, in part, to give a practical interpretation of what is ostensibly quite an abstract, ambiguous philosophy (Parkany, 2006; Townsend et al., 2005). In the FHWA publication, *Flexibility in Highway Design* (FHWA, 1997), Jane F. Garvey, the Acting Federal Highway Administrator at the time, states in the introductory message of the publication that:

<sup>20</sup> <http://www.cts.umn.edu/contextsensitive/index.html>

<sup>21</sup> <http://www.ktc.uky.edu/csd.html>

*“This guide does not attempt to create new standards. Rather, the guide builds on the flexibility in current laws and regulations to explore opportunities to use flexible design as a tool to help sustain important community interests without compromising safety. To do so, this guide stresses the need to identify and discuss those flexibilities and to continue breaking down barriers that sometimes make it difficult for highway designers to be aware of local concerns of interested organizations and citizens.”*

This emphasis placed on identifying the inherent flexibility in the laws, regulations and best practices for road planning and design, speaks to the policy of many designers to apply design guidelines as rigidly as possible, using them as minimum standards so as to avoid exposure to tort liability (Jones, 2004). It also, however, highlights the need for an alternative mechanism to harness the flexibility inherent in the guidelines more effectively, so as to achieve some of the objectives of CSD using the quantitative assessment and evaluation skills already commonplace in the practice.

## 4.2 Complete Streets

In contrast to CSD, which has migrated from academia to find acceptance in more formal planning bodies, and subsequently policy and legislative circles, the Complete Streets movement originated amongst social advocacy groups, and has largely been a community driven and community focussed development, lately also attracting the support of academic interest groups in the fields of preventative healthcare, social justice and integrated multimodal transport.

The Complete Streets movement has become closely linked with promoting a more active lifestyle amongst urban populations (see Bors et al., 2009; Deehr and Shumann, 2009; Geraghty et al., 2009; Glasgow and King, 2009), and in combating obesity, especially amongst children (Sallis and Glanz, 2011). The interest generated around this aspect of the movement has been very beneficial for it. The Active Living by Design (ALbD) program, a community grant program of the Robert Wood Johnson Foundation (RWJF), which was established to help 25 communities create environments that support active living (Bors et al., 2009), was a key initiative in promoting Complete Streets. Each funded community established a multidisciplinary community partnership and implemented so called ‘5P’ strategies dealing with preparation, promotions, programs, policy and physical projects.

Policy strategies around infrastructure were aimed at lowering or removing the physical barriers to adopting a more active lifestyle. Typical interventions included adopting a trail master plan, enhancing a land use plan, allocating government funds to projects and improving street design guidelines to better accommodate cycling and walking. Projects undertaken as part of the initiatives included building new parks and walking trails, painting crosswalks and bike lanes, and improving stairway visibility and access.

Complete Streets has also been championed by advocates for social justice. Deehr and Shumann (2009) note that equity and social justice were key values when carrying out the community group Active Seattle’s work in the ALbD program. McCann and Rynne (2010) note that lower-income residents who rely more heavily on transit, bicycling and walking for transportation, yet often do not have the time or resources to fight for better facilities, on a project-by-project basis, are a key constituent of Complete Streets programs. They quote Mike Piscitelli, transportation director for New Haven, Connecticut, as saying that the city’s Complete Streets policy has *“... been a way to create an identity around something that’s been around the city for a while as an important priority.”*

Geraghty et al. (2009) documents the application of Complete Streets principles in promoting walking and cycling in Sacramento, California, in the USA. Once again driven by a community partnership, The

Partnership for Active Communities, brought together a diverse range of organisations to create a five year project supporting NMT. Using a community action model, the partnership focused on programs and promotions to expand walk- and bike-to-school programs. The focus was placed on policy and physical projects through conducting systematic reviews of development projects to influence land use. A comprehensive communications plan united diverse partnership interests to advocate for Complete Streets policy changes and the improvement of transportation infrastructure.

As a result of these efforts, walk and bike to school programs grew in size and number, and community design workshops helped leverage more than US\$12 million in additional support, including Safe Routes to School grants. The partnership delivered more than 150 project reviews to city planners, architects and developers with recommendations for improved pedestrian and bicycle infrastructure. Complete Streets is now included as a policy in the regions transportation plan, in the mobility element of the city's updated general plan and the county's draft circulation plan and in the regional transit master plan.

Laplante and McCann (2008) notes that transportation projects, typically, begin with an automobile-oriented problem, either an increase in average daily traffic or a deteriorating LOS (LOS). Solutions, therefore, tend to be framed in an automobile centric manner, with the performance of the right of way for bicyclists, pedestrians and transit riders or transit vehicles often not being a primary consideration. Laplante and McCann identify the intent of Complete Streets as being "... to change the everyday practice of transportation agencies so that every mode should be part of every stage of the design process in just about every road project - whether a minor traffic signal rehabilitation or a major road widening. The ultimate aim is to create a complete and safe transportation network for all modes."

Pucher et al. (2010) conducted an international review of existing research on the effects of various interventions on levels of bicycling. Their study lists Complete Streets initiatives amongst those identified, stating that, "*Complete Streets policies had been adopted by 25 local and regional governments in the US and by 10 states as of 2007. The US Congress is considering a federal Complete Streets policy. The number of projects built according to Complete Streets principles is growing.*"

Both CSD and Complete Streets initiatives, such as those mentioned above, have found traction in the political sphere through the work of community organisations and advocacy groups. This has led to funding being allocated to the various projects, and has forced planners from both the consultancy world and the authorities responsible for oversight to change their approach to road planning. Without this continued community pressure, it is doubtful that much progress could have been made. In many parts of the world, communities are not as empowered, either through illiteracy, ignorance, oppression or simple complacency, and these issues may not be seen as being of primary concern. In these places, in the absence of such communal unity, the principles of CSD or Complete Streets must be driven from within the engineering and planning fraternities. It is, therefore, important that engineers and planners understand the issues and embrace the solutions on offer. Given the difficulties of garnering support for them in the USA, which has vastly more financial, logistic and organisational capabilities than most developing world regions, not to mention community support, it is uncertain as to whether this approach will spread with sufficient rapidity to elsewhere in the world.

### 4.3 ARTISTS

Complete Streets and CSD represent two convergent developments in transportation planning that highlight the need to consider more than the traditional mechanistic, mobility concerns when planning transport infrastructure. Urban streets may perform a variety of civic, ceremonial, political, cultural and social roles, as well as commercial and economic roles, in addition to their movement roles (Svensson,

2004). This multiplicity of roles implies that the functions performed by the road, and the needs of those who are expected to use it, must be thoroughly evaluated and understood, before an appropriate planning recommendation can be made.

The Arterial Streets for People (ARTISTS) project (Svensson, 2004), a research project in the European Commission Fifth Framework Programme, was set up to address the perceived gap in expertise in designing and planning urban arterial streets. The authors contend that conventional guidance on the design and management of urban roads (roads being primarily focussed on mobility) and streets (being primarily focussed on access to properties) has tended to focus on either arterial roads or local access streets. They hold that there is a lack of a clear and consistent approach to the design of arterial streets, which combine both significant through traffic and urban place functions. The project sought to develop a new approach to road planning that addressed these problems.

The approach developed was novel, and was explicitly developed to be 'people-centric', considering the function of the road, both in terms of the specific location along the road being considered, and the route as a whole's place in the network. The authors note that there are often several uses or activities competing for the available urban street-space. These activities may sometimes be in conflict with each other, and so where they coincide in space they may need to be controlled so that different activities use the same space but at different times. They note that it is the task of street design and management to mediate between competing activities and afford them an appropriate share of space and time.

The novelty of the approach adopted is in the method used to find an appropriate balance between the so-called place specific functions, and the movement function at any point along the road. The system is based on two unique principles:

- Any street section has a combination of link status and place status. These are independent of each other, rather than one being the inverse of the other, as with traditional functional classification systems.
- Link status and place status will depend, not only on the immediate attributes of the street section (including physical form and use), but also on their role with respect to the wider street and urban system considered as a whole.

The method defines the link status as the relative significance of a street section as a link in the network. It is based on the link's scale of significance (local access street, district distributor, city arterial) within the network. Care is taken to note that link status may vary by mode, so that a link may be particularly important in the pedestrian network, but not so in the vehicular network, although these may overlap. However, the definition to be applied, in this instance, is actually a restatement of more conventional practice, in that the link status being referred to is actually the 'network function' or 'strategic function' of the link.

In determining the link status, the method acknowledges that the current link status is either predefined by design, or assigned by the relevant authorities, and that these authorities probably remain the most suitable parties to decide this input. However, the report suggests that community inputs should also be sought to understand how road users perceive the strategic importance of the link to their journeys and this information be used to supplement expert judgement. This information can probably also be gleaned from revealed behaviours or trends in trip making data.

Place status is defined as denoting the relative significance of a street section as an urban place in the whole urban area. A street or square may perform a city-wide role or a more local role. Therefore, the place status is, like link status, related to geographical scale with regard to frequency and type of use, and in principle can take on national or international scale significance.

These ideas are then interpreted as a two dimensional classification scheme, with one axis representing the link status, divided into five categories from local through national, and the second axis representing the place status of the section being considered, also divided into five categories from local through national (Figure 4.1).

Link status	National	Ie	Id	Ic	Ib	Ia	<div style="display: flex; flex-direction: column; gap: 5px;"> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #444;"></span> Arterial roads/ways</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #222;"></span> Arterial streets</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #ccc;"></span> Non-arterial roads/ways</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #eee;"></span> Non-arterial streets</div> </div>
	City	Ile	Ild	Ilc	Ilb	Ila	
	District	IIle	IIId	IIlc	IIlb	IIla	
	Neighbourhood	IVe	IVd	IVc	IVb	IVa	
	Local	Ve	Vd	Vc	Vb	Va	
		Local	Neighbourhood	District	City	National	Place status

**Figure 4.1:** ARTISTS Functional Classification.

*Source:* Svensson (2004)

Envisioning road categorisation in this manner has several benefits. The character of different links can be described in terms of its importance to the functioning of the broader road network, without losing sight of the various locations along it and the relative importance each has in terms of their urban or place status. This allows for a more holistic view of the functioning of the road, by not separating it from the communities and areas it passes through. This aspect of the framework also implicitly recognises that a single link may have varying characters along its length, and that each of these may warrant a different treatment to ensure the optimal overall performance, such performance being measured in terms of more than simply mobility concerns.

The resultant classification table contains 25 different possible combinations of place and network status - each supposedly being a unique description of the overall character of the route. This contrasts starkly with the much more conservative five or six categories of road defined by traditional classification schemes. This many distinct categories is both a benefit, as well as a flaw. The subtle variations amongst roads can be better captured and expressed, but the multiplicity of categories can be confusing, and more importantly, as the number of categories increases, the distinctions between categories becomes less meaningful.

However, the important aspect of the ARTISTS approach to this research is that it introduces the concept of the importance of 'place', however loosely defined, into the classification process. Place is clearly an important aspect of any road, since, as discussed by the ARTISTS group, place is a very significant descriptor of the activities conducted at any location, and of the people who are conducting those activities. Also important is the fact that the ARTISTS approach recognises that 'place' importance varies between different locations, something that is not explicitly included in Traditional Functional Classification. The weakness of the method, however, stems from the subjectivity in describing 'place'

importance. A further shortcoming is that route segmentation is equally subjective, and relies upon local knowledge and the intersecting road network to delineate segments.

A better approach, then, would define place in more specific, preferably measurable terms, eliminating any subjectivity from the assessment. Segmentation, too, should be more granular, relying instead on the measures of place to differentiate between segments. A more rigorous definition of place would draw upon the various aspects defining the type of activities conducted at a location, and the metrics defining the people who can be expected to be using a location, and apply this information systematically to define place.

## 4.4 Defining Context

In a certain sense, the combination of the concepts of ‘place status’ and ‘link status’ is similar to the context, since the context of any location along a road is composed of both aspects of the place (activities), and aspects of the link (mobility). The challenge, however, is to define what this means in terms of provision for road functions. Infrastructure should be configured so as to provide the best balance between the needs related to the activities performed at any location, and the mobility needs along the route. Traditional classification methods attempt to find the best compromise between these needs by prioritising in relation to mobility (higher order roads allow for higher levels of mobility, lower order roads for less mobility). However, travel is a derived need, driven primarily by the need to conduct activities that are spatially separated and, in many instances, the travel associated with any activity is multimodal. One might drive from home or use PT to a commercial district, but the actual shopping is done on foot.

There is, therefore, a link between the context of a location, the activities performed at the location and the mode of travel used for the trip. Each mode used in a road has its own specific characteristics and needs, which determine the design parameters for that mode. Also, each location in a network, or along a road, is defined by a set of contextual parameters that determine how, and by whom, the road, at that location, is most often used. It is when these modal characteristics and location specific factors, or the needs of a mode and the use of a location align, that an ideal planning solution is found. Accordingly, certain modes are better suited to a particular set of contextual circumstances than others. Therefore, under a given mix of contextual circumstances, certain modes should be given a higher priority than the rest. Priority then, and by extension context, can be used to determine infrastructural needs.

There is still the question of defining the context, and how this relates to the various modes of transport. Since context is a description of who is using the road, and why they are there, the natural descriptors of context should be found amongst descriptors of land use (such as type and intensity), demographic or socioeconomic factors (such as densities, employment levels, income levels and age groups), environmental factors (such as environmentally or culturally sensitive areas) and transportation factors (such as the presence of PT facilities and the number and type of trips through the area). The specifics of these and other factors and how they might relate to the context, as they have been defined here, are explored in the next section.

### 4.4.1 Land Use and Context

There has been a large amount of research done into the relationships between various aspects of land use and transportation. The results from these studies tends to show that there is definitely some interaction between aspects of land use and travel behaviour, but there is disagreement as to precisely what this interaction is, how strong the effects are, and which variables are at play. Often,

methodological differences, or incomparable data sets, lead to differing results being reported between studies (Badoe, 2000).

The concept that mode choice could be influenced by land use factors, or put another way, that land use factors influence travel behaviour, first emerged in the 1960s, when some of the earliest work in the field was conducted by Levinson and Wynn (1963) for the Highway Research Board in the USA, into the effects of land use on transportation demands. The results from that investigation showed that if neighbourhood density and the distance to the central business district (CBD) were both to increase by one standard deviation, average household vehicle miles travelled (VMT) would drop by roughly one third.

Further progress was made during the 1970s with work by Pushkarev and Zupan (1977), who investigated the relationship between land use and transit patronage. This publication introduced the concept of 'land use thresholds' at which different types of transit were feasible investments. The methodology used single-equation ordinary least square regression analysis, their choice of method being dictated by the lack of data available at the time and their desire to present the results as nomograms. The nomograms were designed to facilitate a planner's choice of a feasible transit alternative, given values of current or expected density levels and other relevant variables. The authors used the size of the CBD, measured in non-residential floor space, the distance of the site from the CBD, and residential density, as the determinants of transit demand. The study also accounted for socioeconomic characteristics affecting transit patronage, such as vehicle ownership levels, household size and income.

The 1990s saw an increasing focus on the issue of the land use and transportation linkage, driven primarily by the growth of the New Urbanist movement in the USA (see Leccese and McCormick, 2000), an increasing awareness of the urban problems associated with automobile dependence (Newman and Kenworthy, 2000) and the growth of the environmental movement. Researchers began looking in more detail at specifics around the relationship between land use and transportation. The earlier studies mentioned previously were, generally, only concerned with work or commuting trips, and tended to concentrate on the split between automobile usage and transit usage, often with a focus on transit feasibility or cost effectiveness. Later research, such as that conducted by (Handy, 1996b) and (Kitamura et al., 1997), investigated the effect of urban form and neighbourhood design on mode choice, in particular walking trips (to investigate the New Urbanist agenda), and looked at a broader range of trip purposes, such as shopping and recreational trips.

More recently, interest has developed in the influence of land use and urban form on scholar travel behaviour. Ewing et al. (2004) conducted a study to examine the relationship between mode of travel to school and factors that might affect mode choice. Using data from Florida, they estimated a multinomial logit model to explain school mode choice for a sample of students from different groups. Their model indicated that students with shorter walk or bike times to school proved significantly more likely to walk or bike, and students travelling through areas with sidewalks on main roads were also more likely to walk. Just as interestingly, they found that land use variables, such as density and land use mix, were not significant indicators of mode choice. The travel behaviour literature, generally, highlights the importance of these variables in travel decision making but, according to this analysis, school trips are different (Black et al., 2001; Braza et al., 2004; Kouri, 1999). Ewing et al. theorise that, since school trips tend to be unrelated to other activities, the need for proximity to other land uses is reduced. In addition, these trips tend to be mandatory and so the walking environment is less important than it is with other trip types. School trips also involve children, who may be less sensitive to walking conditions than adults.

### **Residential Density and Context**

The literature, generally, attempts to relate either residential density, employment density, accessibility or neighbourhood design, or some combination of these, to mode choice. As regards residential density, the majority of authors seem to concur that an increase in density favours transit use, and that lower densities tends to coincide with increased automobile use (Cervero, 1996; Parsons Brinckerhoff Quade and Douglas Inc., 1996; Smith, 1983). However, despite this result seeming self-evident, many authors caution that density alone does not explain mode choice completely.

Levinson and Kumar (1997) argue that the relationship between density and transportation cannot be isolated from self-selection bias. They argue that individuals choose a density (or distance from the city center) based, in part, on how much they want to commute and what lifestyle they wish to lead. Creating additional high density areas may not increase the number of people with certain commuting and lifestyle preferences. Dunphy and Fisher (1996) note that, although their study confirms the previously established pattern of higher levels of transit use and lower automobile travel in higher-density communities, the effects are not clear-cut because of the intervening relationship between density and the demographic characteristics of households. Schimek (1996) found that members of households in higher-density areas make fewer automobile trips and travel fewer automobile kilometers than those in low-density areas. His analysis indicates that for 'trips', most of the effect of density is direct, but for 'distance driven' two-thirds of the effect of density on automobile use comes through the mechanism of lower rates of car ownership in high-density areas, despite the fact that his modelling controlled for the influence of household income on residential density. He speculates that lower rates of vehicle ownership in higher-density areas is more likely the result of the greater attractiveness of alternatives (walking and public transit) and the greater difficulties and higher costs of motor vehicle storage in high-density areas.

There does not seem to be a direct causal relationship between residential density and mode choice, despite residential density being an important contributing factor to the selection of a mode for an individuals trip. As a policy instrument then, increasing residential density through infill development, without complementary or supporting measures, may not achieve significant mode shifts.

In terms of its relationship to context then, residential density appears to be an important factor, since residential density does, in combination with socio-demographic, economic and other land use variables, play an important role in mode choice. In addition, residential density is an important factor in transit patronage and so, in terms of context, should be used to identify areas that are more or less suited to this mode. Of course, since transit trips also imply at least one NMT trip end, areas that are suited to transit should also be friendly to NMT modes.

### **Employment Density and Context**

Frank and Pivo (1995) included employment density in their assessment of the factors that influence mode choice in their study. They found that employment was significantly related to mode choice for both work and shopping trips, especially when both trip ends are considered. They found that an increase in employment density was negatively correlated to single occupancy vehicle use, and positively correlated to walking and transit use.

Miller and Ibrahim (1998) analysed data from Toronto, finding that for their data, vehicle kilometres travelled (VKT) per worker increased as one moved away from the city center, other major employment zones within the urban area, or both. Parsons Brinckerhoff Quade and Douglas Inc. (1996) found that

the employment density in the vicinity of stations was correlated with increased boardings at that station.

Cervero (1989) found that suburban employment centre density influenced the work trip mode split, and that the larger the size of the suburban employment centre, the greater the potential for ride-sharing. He quotes a figure of 3.5% for increased ride sharing for every 5 000 additional jobs in the suburban employment centre.

Although the relationship between employment density and mode choice has been less well investigated than residential density and mode choice, the results from the studies tends to be conclusive and consistent. This is, possibly, due to the more direct relationship which exists between employment density and PT service supply, since employment centres are generally the focus points for PT networks and the locations of terminuses and interchanges. Other modes also tend to enjoy higher levels of service in centres of employment, thereby facilitating their increased use. In addition, transit, ride-sharing and NMT use in higher density areas, particularly in CBDs, is influenced by higher parking costs and parking competition.

A higher employment density is, in general, correlated with increased used of transit modes and walking. In South Africa, where walking, transit and other forms of PT play such an important role in commuting trips (NDoT, 2005a), an argument could be made for prioritising these modes in relation to employment density. Employment density can, therefore, be said to be a descriptive factor for the local context, since areas with a higher employment density are better suited to certain modes than others.

### Accessibility and Context

Badoe (2000) defines the term accessibility relating to a variety of measures of ‘how well connected’ a given location is with activities, such as work opportunities, shopping destinations, and the like. The usual measure is taken in terms of how much of a given activity is located how close to the location in question, with ‘how close’ often being expressed in terms of travel time by the various modes. Accessibility is, typically, formulated using an analogy of the Newtonian gravity model, in which a function of travel time by mode is the denominator, and the relative attractiveness of an area (measured in terms of a function of the quantity of the attracting land use) is the numerator. Kockelman (1996) gives the following formulation:

$$Accessibility_i = \sum_j \frac{A_j}{f(t_{ij})} \quad (4.1)$$

where  $A$  is the attractiveness of zone  $j$  and  $t$  is the travel time from zone  $i$  to  $j$ .

Levinson and Kumar (1994) developed a similar approach.

$$A_{iEm} = \sum_{j=1}^J (E_j * f(c_{ijm})) \quad (4.2)$$

$$A_{iRm} = \sum_{j=1}^J (R_j * f(c_{ijm})) \quad (4.3)$$

where:

$A_{iEm}$  = accessibility to jobs (employment) from zone  $i$  by mode  $m$

$A_{iRm}$  = accessibility to houses (residences) from zone  $i$  by mode  $m$

$E_j$  = number of jobs (employment) in zone  $j$

$R_j$  = number of houses (residences) in zone  $j$   $f(c_{ijm})$  = function of cost/time between zones  $i$  and  $j$

$f(c_{ijm})$  was formulated separately for auto and transit modes, based upon work done in a previous study for metropolitan Washington D.C. For auto trips:

$$f(c_{ija}) = \exp(-0.97 - 0.08c_{ija}) \quad (4.4)$$

and for transit trips:

$$f(c_{ijt}) = \exp(-1.91 - 0.08c_{ijt} + 0.265c_{ijt}^{0.5}) \quad (4.5)$$

The indices used here were estimated using multiple regression. The dependent variable in the estimation of these equations was the number of trips divided by the number of opportunities (possible trip ends), to which a natural log transformation was applied. Travel time, and its transformations, served as independent variables.

In both examples, the cost of travel between locations is simplified to travel time. This is a useful approach since it takes into consideration more than simply the distance between locations, but also the travel conditions along the route (congestion, directness, levels of service). Other measures of accessibility include the ‘connectivity quotient’ (Bendavid-Val, 1991). This index is, essentially, a normalised average travel distance to other sites or cities, and its summation can be weighted by the destination size to better account for differences in destination ‘attractiveness’.

Although a large body of research exists exploring the links between accessibility and mode choice, the findings are not as conclusive as for density. In general, although researchers agree that there are definitely impacts, the extent of these impacts is in dispute. Kockelman (1996) found that accessibility is an important determinant of VMT and mode choice. As accessibility increased, VMT decreased and walking and transit use increased. Handy (1996a) found that residences and commercial areas must be close and barriers to walking must be minimised for shopping walk trips to occur. Handy (1996b) found that the distance from home to the store is important in the choice of destination for walk shopping trips. Messenger and Ewing (1996) found that the jobs-housing balance, in combination with auto ownership levels and transit service levels, affects transit mode split. Miller and Ibrahim (1998) found that VKT per worker increases with distance from the CBD and/or other major employment centres, but jobs-housing balance has little impact on VKT/worker. Ewing (1995) found that good regional accessibility to activities is more important than localised density or land use mix in reducing vehicular travel. He also found that the jobs-housing balance has little impact on vehicular travel, but that the accessibility of residences to a range of land uses does reduce vehicular travel.

The relationship between accessibility and context is unclear. There is consensus amongst researchers in the field around the value of a measure, such as accessibility, to predict modal split, but the relationship between accessibility and context may not always be as simplistic as that of density and context. According to Ewing; Kockelman; Messenger and Ewing, areas that are highly inaccessible, either as a result of their physical distance from job sites, or the poor service levels of the links between

them and job sites, would result in high levels of automobile use. Using the theoretical formulation, used previously in these contextual settings, an increase in the facilities for automobiles would be appropriate.

The problem with this approach is that it ignores the historic realities of South African cities. Many poor, historically disadvantaged communities were forcibly resettled to peripheral suburbs as a result of the Apartheid governments Group Areas Act (Act No. 41 of 1950). The result is that these communities, although highly inaccessible, are heavily reliant on PT to access employment centres elsewhere in the city. The approach used to characterise context in terms of accessibility would be contrary to the realities of the situation here. Using the inverse to define context in terms of accessibility presents similar theoretical problems.

In this sense then, despite being a useful indicator of modal split (at least for communities with high levels of car ownership) accessibility is not an appropriate descriptive of the context, and cannot be used to identify which mode should receive priority in a given location.

### **Neighbourhood Form and Context**

In recent years many researchers have focussed on questions concerning the role which local neighbourhood design plays in determining travel behaviour. In particular, advocates of 'neo-traditional neighbourhoods' and town planning have argued that such neighbourhoods encourage more walking and transit trips, as well as shorter trips, and so can be used to minimise congestion and environmental impacts of motorised travel.

Crane and Crepeau (1998) investigated the effects of urban form on travel behaviour, looking at the influence of street pattern connectivity, street type mix, street network density and land use mix in terms of the residential, commercial and vacant land share. Their empirical model was based on the idea that trip behaviour can be explained as a function of trip costs and benefits which, in turn, are the product of trip lengths, modes, purposes and individual characteristics. They found that their data did not support the argument that the neighbourhood street pattern has any significant effect on car or pedestrian travel when controlling for land uses and densities around the trip origin, trip costs, and traveller characteristics.

Cervero and Gorham (1995) compared commuting characteristics of transit-oriented and auto-oriented suburban neighbourhoods in the San Francisco Bay Area and in Southern California. In contrast to Crane and Crepeau, they found that the distinction between traditional neighbourhoods, originally laid out around transit stations and, more recent, automobile-centric neighbourhood patterns, does influence commuting behaviour. Specifically, they note that neighbourhood design seems to affect the degree to which people drive alone to work, and the degree to which they walk or bicycle. Transit neighbourhoods, in general, showed lower drive-alone modal shares and trip generation rates than did their automobile counterparts. Similarly, Transit neighbourhoods averaged higher walking and bicycling modal shares and generation rates than did their automobile counterparts.

The contrasting findings from these two studies, broadly, illustrates the inconclusiveness of the literature on this issue. Badoe (2000), noting these mixed findings, speculates that the key issue is that, in general, 'activity time-space prisms' extend well beyond the local neighbourhood, considering the levels of accessibility available to many people living in urban areas these days and their expectations or needs concerning the range of activities they can access. He mentions the fact that it is nearly impossible to achieve unity in the jobs-housing balance of an area, given the complexity of labour markets and the nature and range of work activities people take part in. As a result, although neighbourhood design is probably important in encouraging walking and cycling for local trips, such as shopping or school trips

(see Ewing et al., 2004; Handy, 1996a), many other trips (beyond the confines of the neighbourhood) are more strongly influenced by other factors.

The relationship between neighbourhood form and context is, therefore, also rather tenuous. The influence of neighbourhood form on local trips is quite clear for residential and even commercial areas, to a certain extent, but unclear for other land use types. The literature, generally, refers to residential land uses only, whereas the assessment of the local context should be applicable to all parts of the city.

#### **4.4.2 Socioeconomics and Context**

The socioeconomic profile along the route encompasses issues related to, amongst others, neighbourhood demographics, such as age and gender, income levels and employment levels. This information is critical to the route context, since it details the types of users, their levels of ability and the probable modal split at a location (NDoT, 2003).

Stead (2001) notes that one of the criticisms occasionally levelled at empirical studies of land use and travel patterns, is that the influence of socioeconomic factors is not considered. Many of the studies mentioned in Section 4.4.1 contend that higher densities are associated with less travel, and less Single Occupancy Vehicle (SOV) trips. This could also be attributed to variations in income, or automobile ownership (which is often highly correlated with income).

Handy et al. (2005) investigated the question of correlation versus causality in the relationship between the built environment and travel behaviour using data from Northern California. They looked at differences between neighbourhoods, but also investigated the effects of changes within neighbourhoods. They found that a simple comparison of different neighbourhood types showed significant variations in the levels of driving. However, a multivariate analysis of cross-sectional data showed that these differences were largely explained by attitudinal information of travel preferences. The effect of the built environment mostly disappears when attitudes and socioeconomic factors were accounted for. Their longitudinal analysis of changes in driving and changes in the built environment, however, showed significant associations, even when attitudes have been accounted for, providing support for a causal relationship.

Of the studies that have recognised the effects of socioeconomic factors and adopted research methods that attempt to hold socioeconomic variables constant (in order to observe the effects of land use characteristics), two main research methods have been used. The first, and more popular method, uses multiple regression analysis, in which socioeconomic variables and land use characteristics are treated as explanatory variables (see Cervero, 1989; Ewing, 1995; Frank and Pivo, 1995; Kitamura et al., 1997). The method allows for the identification of the main socioeconomic and land use characteristics that are associated with certain travel patterns. The method does not, however, allow the identification of causal relationships. A second method employed to hold socioeconomic variables constant involves the selection of case study areas that have similar socioeconomic profiles but different land use characteristics. In this way, socioeconomic differences are minimised and the variation in travel patterns is assumed to be the result of land use characteristics (examples include research reported by Curtis (1995) and Handy (1992)).

This section of the thesis explores some of the issues regarding which socioeconomic factors should play a role in a description of the context.

#### **Income and Context**

That socioeconomic factors are important indicators of travel behaviour is particularly true in South Africa, where socioeconomic characteristics, such as income, play a pivotal role in determining people's

travel behaviour. According to Peska and Venter (2009), income forms the basis for the differentiation of residential trip generation rates in South Africa. Table 4.1 lists percentage of South African commuters by main mode to work and income group.

**Table 4.1:** Main mode to work - % of commuters

Income Group	Train	Bus	Taxi	Car	Walk/cycle	Other
Up to R500	3	7	20.5	4.4	57.9	7.2
R501 - R1000	6.6	10.5	29	6.6	39.5	7.8
R1001 - R2000	10.4	12.4	37.9	13.8	19.4	6.2
R2001 - R3000	8.9	11.1	31.3	28.5	13.7	6.4
>R3000	2.4	5.5	15.7	65.4	8.4	2.6
RSA	6.2	9.2	26.6	27.7	24.6	5.7

Source: NDoT (2005a)

It is clear from the table how strongly income is related to mode choice. Walking and cycling (although walking is by far more common than cycling) is the main mode for approximately one quarter of all work trips, but amongst the poorest groupings accounts for between 40 and 60% of all work trips. PT use is also much more prevalent amongst the lower income groups, and amongst higher income groups private car use dominates all the other modes.

This information is corroborated by considering the number and income levels of households with access to one or more cars. Amongst households in the highest income groups, 82% have access to a car. This contrasts starkly with the figures for lower income groups, and is reflected in the mode choices made by the different groups (Figure 4.2).

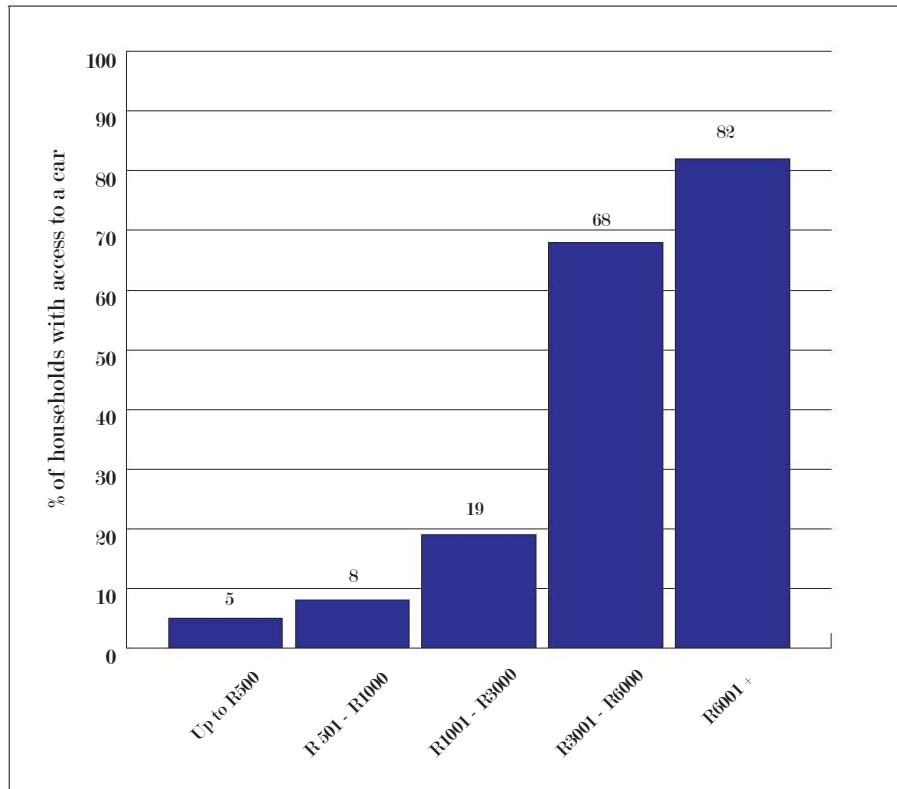
**Table 4.2:** Trip generation in the RSA by household income

Monthly household income	% of all persons				Average no of trips per person	Average no of trips per household
	0	1	2	3+		
Up to R500	28.8	47.4	12.4	11.4	2.34	7.79
R501 - R1000	30	48.8	11.7	9.6	2.18	8.41
R1001 - R3000	24.5	51.1	13	11.4	2.44	9.89
R3001 - R6000	17.8	54.2	14.3	13.7	2.74	10.74
>R6000	12.5	56	15.5	16	3.06	10.88

Source: NDoT (2005a)

Trip generation is also strongly related to income. Table 4.2 shows the number of trips made per person, and the average number of trips per household, for the different income categories. The data shows that, although approximately 50% of people make at least one trip per day (it was assumed that each trip from home generates a return trip), irrespective of income level, people in households with a higher income make approximately 30% more trips than those from the lowest income households.

The data shows that any examination of travel behaviour in South Africa must consider income as being an important variable. It is, therefore, reasonable to expect that income should play a role in



**Figure 4.2:** Household access to cars by income group.

*Source:* Adapted from NDoT (2005a)

describing the context of a route. Along sections of the route passing through higher income areas, the people from that area using the route may reasonably be expected to make different travel choices than those using the route in lower income areas.

### Gender and Context

The influence of gender on travel behaviour is a very well researched topic, with it being widely understood that men and woman tend to have different travel patterns as a result of the different roles they generally play in the household. Gordon et al. (1989) investigated the gender differences in metropolitan travel behaviour from data for the US. Their research found that women consistently have shorter worktrips than men, regardless of income, occupation, marital and family status, mode of travel or location. They also found that women undertake more non-worktrips than men.

These findings are confirmed elsewhere in the literature. Root et al. (2002) explains that women use transportation differently than men do. For instance, like Gordon et al. (1989), the study finds that women tend to make more non-work trips, and that these trips are, generally, shorter than mens (see also Beirao and Sarsfieldcabral, 2007). Similarly, Hjorthol (2000) notes that in the case of work trips, women are likely to have shorter trips than men. Woman also seem to have different attitudes to automobile use than men do. Polk (2003) found that women tend to make fewer car trips than men.

Many authors also note the importance of trip chaining in female travel patterns (see Bhat, 1997; Dobbs, 2005; Golob and Hensher, 2007). A good illustration of some of the issues at play here can be found in Zhang et al. (2008), who studied the travel behaviour of woman in Beijing. Their study results show that

the trip chains of high-income females are more complicated than that of low-income females. Children have a significant impact on a woman's travel behaviour. They found that the proportion of women who had complicated trip chains was 17.9 times greater for women who had children between the ages of 0 and 6 years, than for women without minor children. Interestingly, although not unexpected, the presence of supporting family members tended to simplify the home work trip chains of female commuters.

Gender differences are also often encountered in the perception of risk. Kim et al. (2008) modelled the crossing violations at 15 signalised and 15 unsignalised crossing locations that were identified as being particularly hazardous in Hawaii. They found that there were statistically significant differences between violators and non-violators in terms of gender. Among male pedestrians, a statistically significant proportion (52.5%) broke the law, compared to male non-violators (46.3%). While among female pedestrians, the opposite is true, a larger proportion were found to be compliant (53.8%), with 47.5% breaking the law. In South Africa, Behrens (2010) recorded similar findings for illegal crossings at arterials and freeways.

With these gender differences in mind, what are the implications for the contextual setting of a road? It is generally accepted that transport infrastructure should be gender neutral. However, to achieve this neutrality, infrastructure should be sensitive to the needs and concerns of all users. Often, in this regard, the specific needs of females have been overlooked or poorly understood. Dobbs (2005) argues that, given current social realities, access to PMT is essential for female participation in economic activities. She notes that it allows women to travel to employment outside of their immediate neighbourhoods and existing travel corridors. It also provides them with the opportunity to access employment within the time constraints and complexities of their domestic responsibilities, and it offers them an element of safety and security that is not always available on PT.

PMT does, in fact, assist women in meeting many of their daily transport needs, given the activity time space constraints they often face. It is this activity time space framework, in which many women operate, that could go some way to explaining the differences in travel behaviour between men and women. Many women restrict the length of their workdays so that they can be home in time to meet children, or to collect them from school. This limits the distance they can travel, but increases the number of individual trips they have to make to conduct all of the activities they must complete. Both Bhat (1997) and Zhang et al. (2008) found that the travel behaviour differences between single men and single woman were much smaller than for coupled men and woman. This finding seems to hold for a range of age groups (Golob and Hensher, 2007).

However, are these gender differences significant enough to influence modal priority at any location? Put differently, how would a significantly higher proportion of one sex, in a particular location, make a mode more or less suitable? Dobbs (2005) indicates that, if the gender balance in any particular location were skewed towards woman, the automobile should be favoured. It could be argued, though, that the conditions of other modes, such as PT, could be improved or adapted to better meet the underlying needs driving travel behaviour in women. For example, noting that females are substantially less likely than males to cycle in countries with low bicycle transport mode share, Garrard et al. (2008) investigated whether female commuter cyclists were more likely to use bicycle routes that provide separation from motor vehicle traffic. They found that, consistent with other studies that highlight gender differences in aversion to risk, female commuter cyclists preferred to use routes with maximum separation from motorised traffic. Improved cycling infrastructure, in the form of bicycle paths and lanes, that provide a high degree of separation from motor traffic is likely to be important for increasing transportation cycling amongst under-represented population groups, such as women.

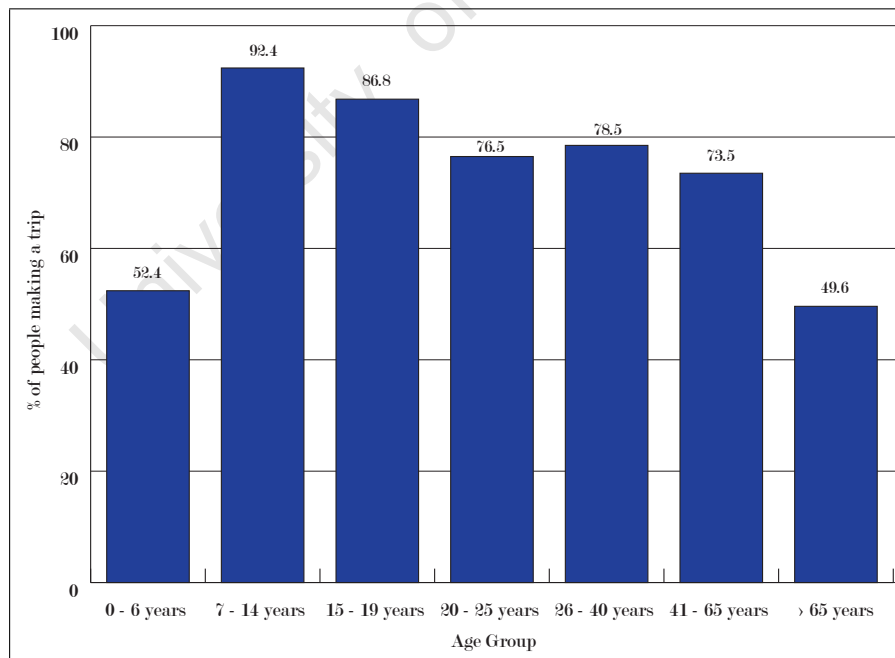
However, although these types of initiatives improve the friendliness of a mode for females, thereby making it more gender neutral, they do not imply that this would make the mode more suitable, given the gender ratio at the location. It seems that, despite gender being an important factor in travel behaviour, context is a gender neutral concept.

### Age and Context

The travel behaviour of different age groups varies, because the travel demands of the activities that different age groups typically participate in varies. This is partly due to the physical limitations that youth and old age places on people. Young people do not drive, and the very young cannot travel without being accompanied by someone older. High rates of travel are, physically and mentally, stressful for the elderly. These facts influence not only the number of trips people undertake, but also the modes of transport they use to make these trips.

Research on the effects of age on travel behaviour is often focussed on topics such as road safety and accidents (especially for children - often with the focus on the trip to school (see Brysiewicz, 2001; Clifton and Kremer-Fults, 2007; Kleinemas and Rudinger, 2010)), encouraging active lifestyles (see Kemperman, 2010; Mackett, 2010), and the social exclusion effects of the lack of adequate age appropriate or sensitive transport facilities (see Ahern et al., 2010; Delbosc and Currie, 2010; Dimitrov, 2010).

In South Africa, a very high proportion of household members in the 7-19 year age group make weekday trips, while fewer (but still about half) of the youngest and oldest groups make trips. The greatest number of trip-makers is, however, among the school-going age group (7-19 years) and the economically active age group (26-65 years) (Figure 4.3).



**Figure 4.3:** Weekday trip-making, by age group of household members.

*Source:* Adapted from NDoT (2005a)

The main demographic indicator of trip purpose in South Africa is age group (NDoT, 2005a). Table 4.3 shows that travelling to educational institutions is the main reason for weekday tripmaking by all age

groups up to 19 years of age. From 26 - 65 years of age, work trips predominate. Among the elderly and retired, visiting and shopping are the most common trip purposes.

**Table 4.3:** Trip purpose, by age group of household member

Age	% of household members naming purpose			
	Education	Shopping	Visiting	Work
0 - 6 years	58.5	19.5	32.1	0.5
7 - 14 years	95.9	17.5	16.4	0.4
15 - 19 years	82.8	25.3	23.2	3.3
20 - 25 years	25.7	37.7	36	28.1
26 - 40 years	2.6	37.9	32.4	57.5
41 - 65 years	1.3	35.9	32.7	53.9
>65 years	0.9	45.3	45.5	6

*Source:* Adapted from NDoT (2005a)

Table 4.4 shows some of the results from an analysis of the 2001 US National Household Travel Survey compiled by Pucher and Renne (2003). As in South Africa, mobility rates in the US declines with age. However, in contrast to the situation in South Africa, the choice of travel mode of elderly people is quite similar to the rest of the adult population there. As with other Americans, the elderly are overwhelmingly dependent on the car for their mobility. They rely on the car for a higher percentage of their trips (89.1%) than for any other age group.

In Pucher and Dijkstra (2003), the authors investigate the differences in travel behaviour with age between Germany, the Netherlands and the USA (Figure 4.4). They show that elderly Germans and Dutch make over half their trips by walking or cycling, whereas in the USA, those non-motorised modes account for only 9% of the trips that elderly Americans make. Dutch elderly who are 75 or older make 24% of all their trips by bike and Germans in this age group make 7% of their trips by bike. Americans who are 65 or older make only 0.4% of their trips by bike. The lack of feasible alternatives to the private car in cities in the United States forces the elderly to drive. The authors note that the reliance on the private car exposes elderly Americans to significant traffic dangers and it deprives them of the valuable physical exercise they would get from walking and cycling.

In South Africa, the working age population makes the greatest use of all travel modes (Figure 4.5). The use of motorised and other modes is lowest amongst those persons of 14 years or less. For the age group 66 and over, the use of PT modes was found to be relatively low compared to the other age groups. Clearly, age plays an important role with respect to trip generation, as well as modal choice. There also are distinct differences for the various age groups in different parts of the world. The predominant trip purposes amongst age groups also varies significantly. The context of a particular location is, therefore, influenced by the age groups that most often frequent it, since one of the aspects of context is who is using the location. Age should, therefore, play a role in the description of context.

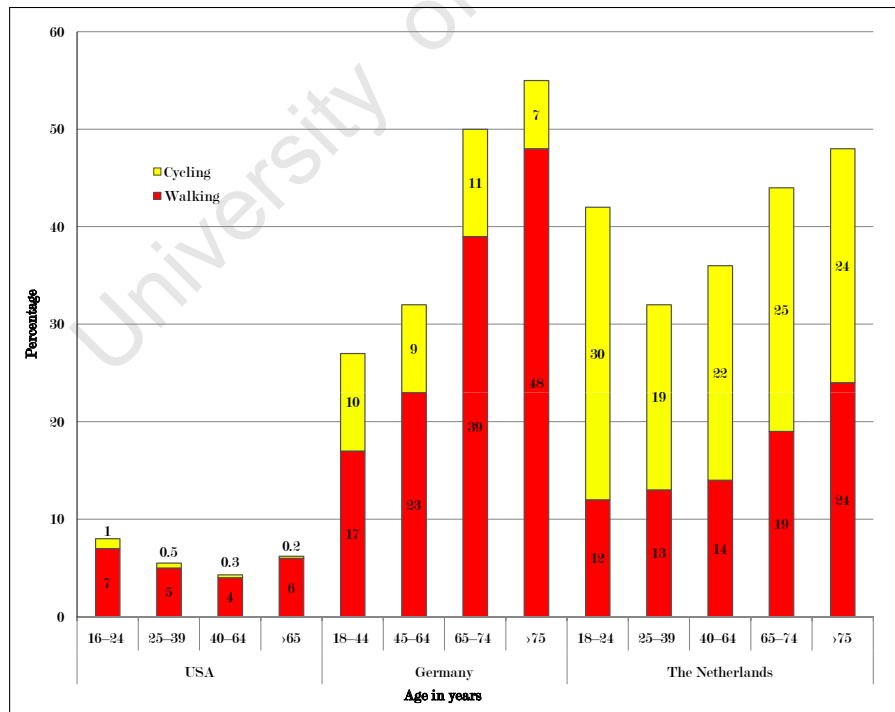
#### 4.4.3 Environment and Context

Probably the most comprehensively analysed aspect of transportation in the last 30 years is that of the effects of transport on the environment. Over the years, a number of high profile, academically

**Table 4.4:** Impact of Age on Modal Choice in the USA (% of trips by means of transportation)

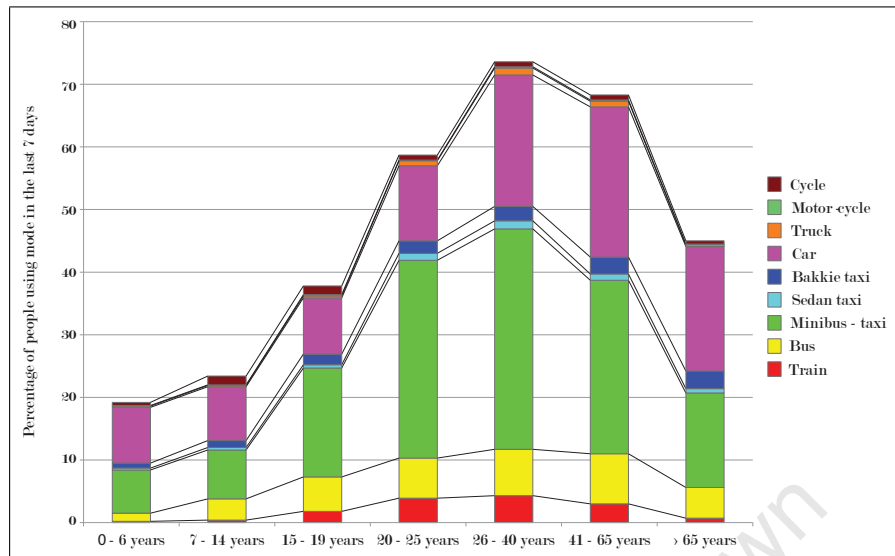
Mode of Transportation	Age					All
	5 to 15	16 to 24	25 to 39	40 to 64	65 & over	
Total Auto	70.7	85.3	87.4	89.8	89.1	85.8
SOV1	0.5	39.2	43.6	51.9	45.7	37.6
HOV2	70.2	46.1	43.8	38	43.4	48.2
Total Transit	1.1	2.9	2.1	1.5	1.3	1.7
Bus and Light Rail	0.9	2.1	1.2	1	1.2	1.2
Metro/Subway/Heavy Rail	0.1	0.6	0.7	0.3	0.1	0.4
Commuter Rail	0	0.2	0.2	0.2	0	0.1
Total Nonmotorised	18.4	10	9.8	8.2	9.3	10.5
Walk	15.2	9.3	9.2	7.8	8.9	9.6
Bicycle	3.2	0.6	0.6	0.4	0.4	0.9
School Bus	8.9	1.2	0	0	0.1	1.4
Taxicab	0.1	0.1	0.2	0.1	0.1	0.1
Other	0.8	0.4	0.3	0.3	0.2	0.4
All	100	100	100	100	100	100

Source: Adapted from Pucher and Renne (2003)



**Figure 4.4:** Percentage of trips in urban areas made by walking and bicycling in the United States, Germany and The Netherlands, by age group, for 1995.

Source: Adapted from Pucher and Dijkstra (2003)



**Figure 4.5:** Use of transport modes by age group in South Africa 7 days prior to the National Household Travel Survey day. The data does not include walking as a mode, since everyone would have walked at some stage during the reporting period.

Source: Adapted from NDoT (2005b)

rigorous and exhaustive studies have been published assessing every aspect of these effects from nearly every perspective.

In the UK, the Royal Commission on Environmental Pollution's 18th report on Transport and the Environment (Royal Commission on Environmental Pollution, 1994), and the follow up, similarly entitled 20th report (Royal Commission on Environmental Pollution, 1997), provides what, according to Banister (1995), "... is a landmark in the debate over transport and the environment." The report addresses the effects of motorisation thoroughly, noting that, "... the unrelenting growth of transport has become possibly the greatest environmental threat facing the UK, and one of the greatest obstacles to achieving sustainable development."

Discussion of transport's effects on the environment is often framed in terms of the concept of sustainability. It is unsurprising, therefore, that the report's list of objectives for a sustainable transport policy is squarely aimed at reducing the dominance of automobile.

1. To ensure that an effective transport policy, at all levels of government, is integrated with land use policy and gives priority to minimising the need for transport and increasing the proportions of trips made by environmentally less damaging modes.
2. To achieve standards of air quality that will prevent damage to human health and the environment.
3. To improve the quality of life, particularly in towns and cities, by reducing the dominance of cars and lorries and providing alternative means of access.
4. To increase the proportions of personal travel and freight transport by environmentally less damaging modes and to make the best use of existing infrastructure.
5. To halt any loss of land to transport infrastructure in areas of conservation, cultural, scenic or amenity value, unless the use of the land for that purpose has been shown to be the best practicable environmental option.

6. To reduce carbon dioxide emissions from transport.
7. To reduce, substantially, the demands which transport infrastructure and the vehicle industry place on non-renewable materials.
8. To reduce noise nuisance from transport.

The list echoes many of the sentiments of South African transport policy, placing an emphasis on promoting integration between land use planning and transport planning and promoting the use of PT and NMT. The list, however, also highlights many of the broader environmental concerns around the impact of transportation, and the infrastructures required to support transportation.

To get a sense of the scale of the environmental problems created by transportation, it is worth considering the findings of Delucchi (1998) on the annualised social cost of motor-vehicle use in the U.S., estimated for the period 1990-1991<sup>22</sup>. Table 4.5 provides an extract of the estimates for the annual costs of the non-monetary externalities of motor vehicle use. Only those costs directly related to environmental impacts are included.

Although these figures only represent a narrow interpretation of the environmental costs of transportation (amongst the costs not listed are the monetary and opportunity value of land taken up for transport uses, loss of scenic quality due to the intrusion of transport infrastructure, opportunity costs of non-renewable resources used for the operation of transport and the provision of transport infrastructure, loss of quality of life), it does place the challenges posed into perspective.

The total costs estimated by Delucchi, assuming no growth in the absolute quantities they are based on (which is, admittedly, an unlikely scenario) and adjusted for inflation, translates to between US\$57.3 billion and US\$859.3 billion (assuming an average annual rate of inflation of 2.45%).

Of the impacts listed, air pollution costs account for approximately 95% of the total. Particularly striking amongst the figures, is the large contribution to the total of the cost of human mortality and morbidity, due to particulate emissions from vehicles (accounting for approximately 50% of the total).

Wicking-Baird et al. (1997) conducted an investigation into the so-called brown haze (smog) in Cape Town. They found that the haze occurs mostly from April to September, due to strong temperature inversions and windless conditions that occur during these months, which leads to the build-up of pollutants emitted into the atmosphere. The main objective of the study was to determine the contribution of all major sources to the brown haze, and to obtain a better understanding of the mechanism of haze formation.

According to Wicking-Baird et al. (1997), in urban areas, particles less than 2.5 microns in size (PM2.5) are the single largest cause of smog. They are also the most harmful size range of particles to human health. As a result, the study focussed on identifying the main sources of PM2.5 in the city.

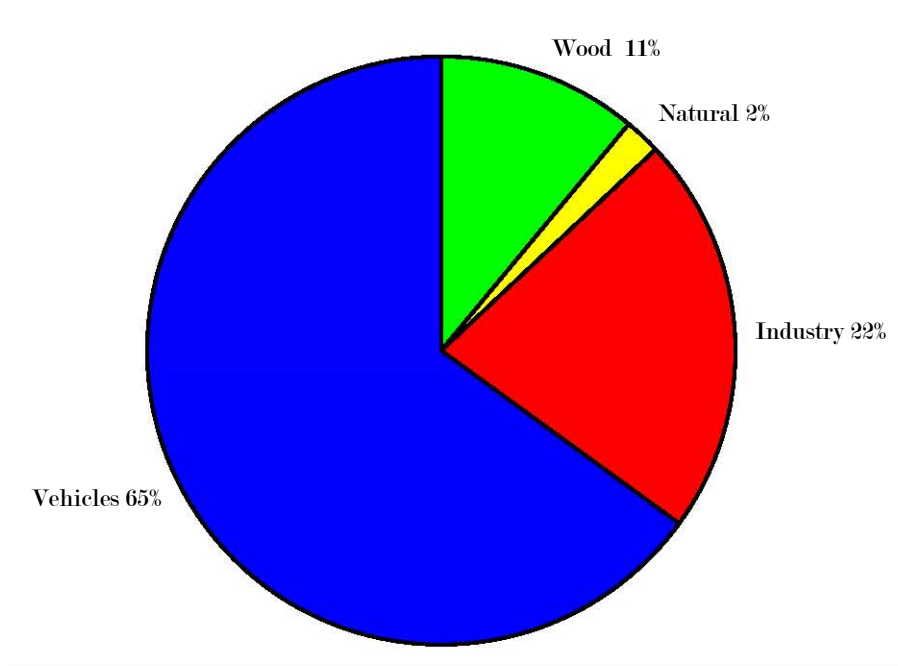
The results indicate that, by far, the major source of particulates in the atmosphere in Cape Town is from vehicles, contributing nearly two thirds to the composition of the brown haze (Figure 4.6). The next largest source is industrial (22%), with wood burning accounting for 11% and natural sources (including crustal sources and sea salt) making up the remainder. Although the data is somewhat dated, the link between motorised transportation and environmental pollution is undeniable. Moreover, it is likely that the situation would have worsened since 1997, as the number of vehicles operating in Cape Town has grown dramatically in the interim. The 2003 decommissioning, and subsequent demolition

<sup>22</sup> These figures, being nearly 20 years old now, may seem somewhat outdated, but it could be argued that considering their age, their magnitude only serves to emphasise the scale of the problems faced.

**Table 4.5:** Estimated annual costs of environmental non-monetary externalities of motor vehicle use, 1990-91 (Billion US\$1991).

<b>Cost Item</b>	<b>Low (US\$b)</b>	<b>High (US\$b)</b>
Air pollution: human mortality and morbidity due to particulate emissions from vehicles	16.7	266.4
Air pollution: human mortality and morbidity due to all other pollutants from vehicles	2.3	17.1
Air pollution: human mortality and morbidity, due to all pollutants from upstream processes	2.3	13
Air pollution: human mortality and morbidity, due to road dust	3.0	153.5
Air pollution: loss of visibility, due to all pollutants attributable to motor vehicles	5.1	36.9
Air pollution: damage to agricultural crops, due to ozone attributable to motor vehicles	3.3	5.7
Air pollution: damages to materials, due to all pollutants attributable to motor vehicles	0.4	8.0
Air pollution: damage to forests, due to all pollutants attributable to motor vehicles	0.2	2.0
Global warming due to fuel-cycle emissions of greenhouse gases (U. S. damages only)	0.5	9.2
Noise from motor vehicles	0.5	15.0
Water pollution: health and environmental effects of leaking motor fuel storage tanks	0.1	0.5
Water pollution: environmental and economic impacts of large oil spills	0.2	0.5
Water pollution: urban runoff polluted by oil from motor vehicles, and pollution from highway deicing	0.7	1.7
<b>Total</b>	<b>35.3</b>	<b>529.5</b>

*Source:* Adapted from Delucchi (1998)



**Figure 4.6:** Sources of brown haze in Cape Town, 1997. By far the major source of particulates in the atmosphere in Cape Town is from vehicles.

*Source:* Adapted from Wicking-Baird et al. (1997)

of the Athlone coal-fired power station in 2010, also means that the particulate share attributable to industry would have decreased somewhat.

The environmental profile along the route speaks to the environmental and cultural or heritage sensitivity along the route. In terms of the National Environmental Management Act (Act 107 of 1998) (Department of Environmental Affairs, 1998), infrastructure provision must give consideration to the physical, biological, social, economic and cultural aspects of the environment that may be affected by the proposed activity. Road infrastructure must, therefore, be planned so as to minimise the expected impacts it may have.

The act, specifically, requires all development to be socially, environmentally and economically sustainable, noting that such sustainable development will take into consideration that pollution and degradation of the environment should be avoided, or, where it cannot be altogether avoided, is minimised and remedied; and that the disturbance of landscapes and sites that constitute the nation's cultural heritage is avoided, or where it cannot be altogether avoided, is minimised and remedied.

Environmentally and culturally sensitive areas are clearly affected by the modes of transport that pass through or nearby them. The presence of environmentally sensitive areas, therefore, should be included as a descriptor for the context. Considering the polluting effects of vehicles, it seems clear that a reduction in the use of these vehicles, especially in the vicinity of environmentally sensitive areas, is prudent.

#### 4.4.4 Transportation and Context

Demand in terms of traffic volume, and supply in terms of capacity, as well as travel speed, are the primary parameters that inform road designs. Additional factors that are usually considered include

modal split, accident statistics, the ratio of through versus local traffic and existing and prior planning for the route and surrounding network. Inputs for these factors are, often, either collected empirically, or are derived from surrounding land uses and demographics. As context relates to who is using the road, and why they are there, some of these factors should be considered when describing it, since any location along a route is defined not only by the activities at the location, but also the mobility requirements at the location.

### **Traffic Volumes and Context**

Volumes and capacities are the major outputs of most traffic studies, since they determine, to a large extent, the scale of the facility required to satisfy the traffic needs at a particular location, or along a section of the route. Volumes are primarily estimated using two methods; backcasting or trip generation. Most often a combination of the two methods is used to predict the future traffic conditions at a location or along a section of route.

Existing traffic volumes are collected through manual or electronic counts, and compared to historic records for the area, or, if unavailable, for similar routes in similar conditions. A projected traffic growth rate is estimated by considering the historic traffic growth along the route, the economic growth of the region, and any significant land use developments that may have impacted the historic figures. This predicted growth is taken as a base, or background, growth scenario. That is, traffic growth along the route that would 'naturally' occur, should no significant developments alter the status quo (Mathew and Rao, 2007).

Of course, the majority of traffic studies are conducted, precisely, to evaluate the impact of changes to the status quo. So the next step in the evaluation would, usually, be to estimate the added impact of these changes. Such changes to the traffic flows are often caused by new developments that will feed onto the route, and their impacts need to be assessed. Traditionally, two methods are available to estimate the trips generated by these developments: travel demand forecasting models for implementing the four-step planning process (see Section 2.2), and travel rate or trip generation indices (see Section 2.4.2) for estimating congestion and delay information (Anderson, 2002).

Most often, a trip generation index is used, since this allows for a straight forward and reasonably reliable estimate of the impacts of any single new development. The use of these indices also dramatically reduces the amount of data required for demand estimation. This simplifies the computational effort required, thereby decreasing the possibility for error. Additionally, since indices are estimated from an analysis of case studies of similar local developments, their use ensures that estimates are locally relevant, and at least historically accurate (Withers and Bester, 2009).

A trip generation index relates the typical number of trips generated by a development to the development type, and some measure of its extent, such as the leasable area of a shopping complex, or the number of employees in an office block. In South Africa, the recognised source for these indices is the South African Trip Generation Rates manual (Stander et al., 1995), currently in its second edition.

Of course, the indiscriminate use of trip generation indices can be problematic. Indices are based on historic information, and represent the 'typical' or 'average' for a particular development type. The particular development being assessed may well differ quite significantly from what can be thought of as 'average' or 'typical', which would increase the error generated from using an index. In addition, if an index is not revised at regular intervals, the data used to estimate it may become outdated, further increasing the disparity between the number of predicted and the actual trips generated. Since every index has an associated error that is published with it, a standard deviation, using indices to estimate

traffic for mixed land use developments is problematic since these error margins become compounded if multiple indices for different land uses are applied to the same development.

Alternative methods for estimating trip generation relies on modelling the factors that influences trip generation and travel behaviour. Models are, usually, formulated in two ways: either a growth factor is estimated based upon a combination of socioeconomic factors and applied to existing traffic, or a regression analysis is performed using the factors as explanatory variables.

The main factors used to assess personal trip production include income, vehicle ownership, household structure and family size. Additional factors, such as the value of land, residential density and accessibility, are also often considered. Trip attraction, on the other hand, is influenced by factors such as space available for industrial, commercial and other services. Employment and accessibility are also often used. In trip generation modelling freight trips, in addition to personal trips, are also important. Although freight comprises about 20% of trips, their contribution to congestion and pollution is significant. Freight trips are influenced by number of employees, number of sales and area of commercial firms (Mathew and Rao, 2007). Trip generation model accuracy is often increased if trips are disaggregated by trip purpose, time of day and person type.

The number of trips generated from, or attracted to, a particular area could be considered an important indicator of the context, in that it directly impacts the type and extent of the infrastructure required to support the activities there. In as far as the context is related to the priority afforded to each mode, the number of trips generated should ideally be differentiated by mode. The inclusion of trip generation is complicated, however, by the fact that some of the factors that affect trip generation are already considered elsewhere in the description of the context. This has the potential to compound the effect of these factors on the description of the context.

This problem can be addressed by distinguishing the trips through an area, from the trips originating in, or destined for the area, and then only considering the through trips. Alternatively, the ratio of trips by mode generated in the area (the modal split) could be used, since this information captures the relative importance of each mode in terms of traffic requirements. However, this also has the consequence of losing the magnitude of the volumes being considered. In terms of describing the priority one mode should receive relative to another, however, it could be argued that the volume of traffic, per se, is of a lesser importance than the modal split.

### **The Location of Public Transport Stops**

People using a road may do so in order to access PT services, since people are attracted to PT routes, specifically, because of this need. In a sense, accessing PT can be thought of as an activity in itself, since people travel to PT stops (the destination) to fulfil a need (accessing PT) that cannot be satisfied elsewhere.

This indicates that the location of PT stops is an important descriptor of the context, since locations near PT stops would attract additional traffic, primarily in the form of pedestrians. The additional pedestrian traffic needs to be accommodated with a higher LOS in the vicinity of the stop, since people standing at the stop may interfere with people walking by, and more people can be expected to cross the road in the vicinity of the stop. PT vehicles can be expected to perform parking and merging manoeuvres in these areas, and at busier stops it may be likely that some informal trading occurs. The mix of people who use the area may differ somewhat from the surrounds, and so the infrastructural requirements to support these road users, and the activities associated with them, will differ from other areas along the route.

PT stops can, therefore, be considered to be important descriptors of the local context, since variations, albeit subtle, in the modal priorities would be necessary in these areas to accommodate the altered requirements of road users at the stops.

### **Accident Statistics and Context**

The location and severity of accidents along a route is often an indication of specific problems in certain areas along the route. The accident characteristics of these so-called ‘hazardous locations’ or ‘blackspots’ can be analysed to identify problematic traffic conflicts, or roadway features, that make them prone to a higher frequency or a greater average severity of accidents (see Park et al., 2010; Xianghai et al., 2010). Volume 4 of the *South African Road Safety Manual* (Committee of Land Transport Officials (COLTO), 1999) includes guidelines on the use of road safety audits to identify and remedy existing hazardous locations and to prevent the introduction of hazardous practices in new projects (Grosskopf et al., 2010).

The site specific nature of these blackspots lends itself to the concept of a location specific context, in that certain locations that are more prone to accidents, could be redesigned so as to decrease the frequency of accidents occurring. Reprioritising amongst the modes in the vicinity of the blackspot may be considered a suitable methodology to address such problems. However, this assumption is flawed, since prioritising amongst modes might only have an appreciable effect for accidents between different modes, and may not be as effective for accidents between the same mode.

The range of systemic errors that ultimately contribute to a crash occurring (Reason, 1990), implies that even if one element of the system is improved, there may be other elements of the system that still leads to crashes occurring. That being said, there is ample evidence to suggest that certain roadway features are associated with a higher frequency of accidents (see Garber and Kassebaum, 2008; Mackenzie et al., 2008; Mok et al., 2006). The question is, given these considerations, is accident frequency a good descriptor of the context?

The context has been defined as relating to who is using the road, and why they are using the road. Accidents are an unfortunate consequence of discord between who is using the road, why they are there, and the infrastructure that has been provided to support these users and their activities. Minimising the accident frequency or severity on a section of the route is, therefore, an outcome, or an objective, of reprioritising between modes. As a result, accident data cannot be included in a description of the context, but could be used as an indicator to identify roads that need further assessment.

## **4.5 Résumé**

In this chapter an examination of the contextual influences on road usage, and the tools that have been developed to account for them, was conducted. The chapter opens with a look at Context Sensitive Design (CSD), which, in recent years, has been promoted by some of the major transport associations in the USA as being able to provide outcomes that harmonise the transportation requirements expected of a project with the needs and values of the community it serves or passes through.

CSD has developed in response to these challenges. It is described as a project management approach driven by a set of principles that addresses the conceptualisation and development, the role and participation of stakeholders, the design and planning considerations, the construction and financial management, and the operations and maintenance planning of a project. However, comprehensive though it may be, CSD has struggled to gain traction amongst planners and designers, in part because there are no concisely outlined, practically implementable methodologies, and because of concerns

around professional liability issues regarding exploiting the flexibility in design standards to meet its principles. Although there are clear benefits to the approach, it has not managed to find support beyond the USA yet.

The Complete Streets movement has many similarities to CSD, but focusses more on the need to create adequate facilities for all modes, with a focus on promoting walking and cycling amongst urban communities. This aspect has found its support amongst community organisations, advocacy groups and in certain academic circles concerned with the consequences of low physical activity on health, especially for children. Further support for the movement is due to its promotion of social justice or equity issues, relating to improving accessibility for the poor. The movement has been limited to a number of pilot projects in communities in the USA and, like CSD, suffers from an ambiguous, ill-defined scope and application protocol, and despite growing interest, has thus far not found widespread acceptance amongst practitioners.

The ARTISTS project was initiated to address the perceived shortcomings of planning for arterial streets - streets where the primary functional character of the route is approximately evenly split between access to properties and mobility for through vehicles. The project developed a new approach to road categorisation that differentiated between roads based upon the relative importance of locale specific aspects along it, and the links importance to mobility in the network. The scheme is important, in that it highlights the importance of balancing aspects related to the 'place', and those related to mobility, and in that it recognises that these qualities may vary independently, both in time and in space, along the route. The method is, however, limited insofar as it is configured only for arterial routes and fails to develop a convincing, quantitative method for comparing either the 'place' or 'mobility' aspects of the route, as well as the segmentation of the route.

CSD, Complete Streets, the ARTISTS project and the planning, design and transport road safety problems discussed in the previous chapters highlights the need to consider the context of the particular location along the route, as well as the mobility needs, when planning the road. The link between context and transport infrastructure lies in the modes that the infrastructure is primarily designed to accommodate. Therefore, identifying which mode is most suitable, given the context, and prioritising the needs of that mode, would ensure that the infrastructure provided is context sensitive, thereby ensuring more equitable, potentially safer infrastructure.

In this chapter, the context, therefore, was defined as being related to who is using the road, and why. This definition implies that the context can be described by combining demographic information on the road users, and information on the activities being performed. This information is recognised as being important to the prediction of travel behaviour and, in particular, modal split. Its application in the description of the context is thus particularly apt.

The set of information used to describe the context divided into four categories. These include land use information, socioeconomic information, environmental information and transportation information. Variables were considered for these categories that are commonly cited in transportation literature, with a particular emphasis placed on information that impacts travel behaviour and, in particular, modal split, since modal priority is defined as the vehicle to express context via infrastructural choices. Variables identified as being important include the land use type, population or household density, employment density, income, age, culturally and ecologically sensitive areas, the relative proportion of each mode using the route and the location of PT stops.

## Chapter 5

# Multiple Criteria Decision Making

In Chapter 4, context was described as relating to the characteristics of the people using the location and the activities they are conducting. The linkage to infrastructure was made through the suitability ranking of the various modes of transport and the criteria, that could be used to express the context in quantitative terms, was identified.

The topic of this chapter is the ordering of these criteria into an evaluative framework capable of producing measurable results. The method used to do this is called Multiple Criteria Decision Making (MCDM), and the chapter opens with some background information on the field of decision making and the method, followed by a more in-depth exploration of the techniques used in the method.

The particular implementation of MCDM, namely Spatial MCDM, that is used in this research is introduced, and the details of the method explored. This is followed by an in-depth discussion around the technicalities and peculiarities of the use of this method for this research. This chapter, therefore, builds the methodological framework for testing the hypotheses and, in particular, for assessing the context as defined in Chapter 4.

### 5.1 Background

Prioritising amongst the modes, given a range of criteria, can be thought of as a decision problem. This problem can be stated as,

*Given a set of information regarding a particular location and information regarding the various modes of transport, which mode is best suited to that location?*

Implicit in this statement is that modal suitability is assessed as the performance of each mode in relation to the characteristics of the location. Such a problem, where there are multiple alternatives, each to be assessed in relation to their performance in terms of multiple criteria, is termed a multiple criteria problem.

Modern strategies to analyse and solve multiple criteria problems developed from work in nonlinear programming for management sciences. Some of the earliest work was done by Kuhn and Tucker (1951), who formulated optimality conditions for nonlinear programming (so-called KKT conditions). In their work they also considered problems with multiple objectives. Later, Charnes et al. (1955) published research that contained the essence of goal programming, even though the name goal programming was first used in 1961 by Charnes and Cooper in *Management models and industrial applications of linear programming*. Romero (1991) provides a comprehensive treatment of the subject.

Building on the earlier work in goal programming, Keeney and Raiffa (1976) published an important work that was instrumental in establishing Multi-Attribute Value Theory (MAVT) as a discipline. MAVT can be used to assess problems that involve a finite and discrete set of alternative policies, that have to be evaluated on the basis of conflicting objectives. The aim of MAVT is to generate a score for each alternative such that a preference order of the alternatives, that reflects the decision makers value judgements, can be developed. MAVT assumes that for every decision problem, a function,  $U$ , exists that represents the preferences of the decision maker. This function  $U$ , is used to transform the attributes of each alternative into a real score. The alternative with the largest score is then identified as the best.

In Europe, Roy (1968) and his colleagues developed ELECTRE, a family of Multiple Criteria Decision Analysis (MCDA) methods still popular today. The original work evolved into ELECTRE I and has subsequently continued to be adapted with ELECTRE II, III, IV, IS and TRI (Figueira et al., 2005). The ELECTRE methods rely on the construction of outranking relationships, in which actions are compared in a pairwise fashion. The method involves the construction of a network of preferences. Using the network, the method develops a set of outranking decisions, or decisions that should be considered as being the best.

Roy (1996) defines multicriteria assessment as “... a decision-aid and a mathematical tool allowing the comparison of different alternatives or scenarios according to many criteria, often conflicting, in order to guide the decision maker towards a judicious choice.” Voogd (1982) describes the aim of multicriteria assessment techniques as being “... to investigate a number of choice possibilities in the light of multiple criteria and conflicting objectives.” According to Janssen and Rietveld (1990), in doing so, it is possible to generate rankings of alternatives according to their relative attractiveness.

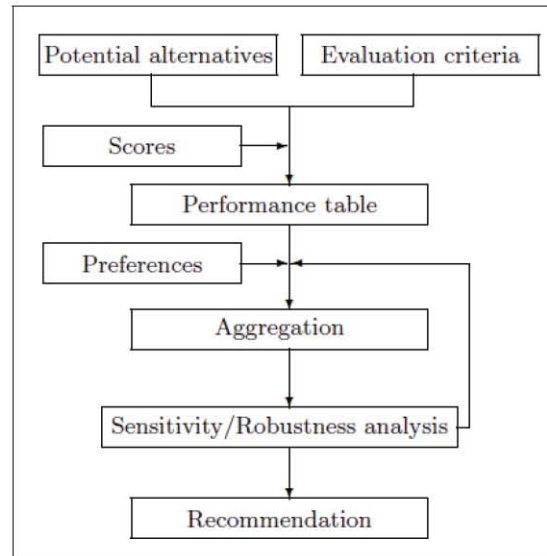
### 5.1.1 Multiple Attribute Decision Making

Multicriteria methods are classified depending on whether the set of possible alternatives is discrete or continuous. The former is referred to as multiple attribute decision making (MADM), and the latter as multiple objective decision making (MODM) (Chakhar and Mousseau, 2007). In the decision problem being considered in this research, the methods used to solve MADM problems clearly apply, since there are a limited number of modes to be considered.

The first step of discrete techniques is the construction of a performance table which contains the scores of a set of alternatives in relation to a set of criteria. The different criteria scores are then aggregated using a specific decision rule or aggregation procedure (Figure 5.1). This procedure takes into account the decision makers preferences, which are generally represented in terms of weights that are assigned to the various criteria. The aggregated criteria scores can then be used by the decision maker to compare the different alternatives. MADM methods can be further subdivided into two approaches, methods that compare the performance of alternatives against each other, so-called outranking methodologies, and methods that assess the performance of alternatives against a goal or a norm.

The uncertainty associated with any decision situation requires that a sensitivity analysis is conducted to enable the decision maker to test the consistency or robustness of a given result. This can be done by analysing the variation of the results in response to some modification in the input data or in the decision makers preferences (weights).

As mentioned, MADM methods rely on the construction of a performance table containing the scores of the alternatives for each criteria. The performance table typically takes the form of a matrix, as shown below:



**Figure 5.1:** General Schema of MADM Methods.

Source: Adapted from Chakhar and Mousseau (2007)

		$x_1$	$\cdots$	$x_n$
		$\mathbf{A}_1$	$\cdots$	$\mathbf{A}_n$
$w_1$	$\mathbf{C}_1$	$a_{11}$	$\cdots$	$a_{1n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$w_m$	$\mathbf{C}_m$	$a_{m1}$	$\cdots$	$a_{mn}$

**Table 5.1:** The performance table

In the performance table, each row belongs to a particular criterion and each column describes the performance of one of the alternatives. The score  $a_{ij}$  describes the performance of alternative  $A_j$  against criterion  $C_i$ . A higher score value, usually, means a better performance, since any goal of minimisation can be easily transformed into a goal of maximisation.

The score  $a_{ij}$  is given by the product of a standardisation function applied to the input value,  $x_j$ , that is associated with the alternative,  $A_j$ , and the weight for criterion  $i$ ,  $w_i$ .

$$a_{ij} = w_i \times f(x_j) \quad (5.1)$$

The evaluation of alternatives may be quantitative or qualitative. Several methods require quantitative evaluations, but in the literature, there are some exclusively qualitative methods mentioned as well, such as the median ranking method. Other methods, such as the ELECTRE family of methods (Roy, 1968), can accommodate both types of evaluations. Despite this, if possible, it is useful to use only either quantitative or qualitative methods in the assessment. When most of criteria are qualitative, quantitative criteria may be converted into qualitative ones and a qualitative method used. Otherwise, a quantification method (i.e., assigning numeric values to qualitative data) is applied; linear scaling is generally the most popular approach.

The application of a quantification method requires the definition of a measurement scale. A popular

measurement scale is the Likert-type. This scale is composed of approximatively the same number of favorable and unfavourable levels. An example with five levels would include very unfavourable, unfavourable, neutral, favorable and very favorable. Other more detailed measurement scales may also be used. The quantification procedure consists of constructing a measurement scale similar to that mentioned above. Numerical values are then associated with each level of the scale. A version of the Likert Scale is presented below:

**Table 5.2:** The Fundamental Scale for Pairwise Comparisons

Intensity of Importance	Definition	Explanation
1	Equal Importance	Two elements contribute equally to the objective
3	Moderate Importance	Experience and judgment slightly favour one element over another
5	Strong Importance	Experience and judgment strongly favour one element over another
7	Very Strong Importance	One element is favoured very strongly over another, its dominance is demonstrated in practice
9	Extreme Importance	The evidence favouring one element over another is of the highest possible order of affirmation

Intensities of 2, 4, 6, and 8 can be used to express intermediate values.

Intensities 1.1, 1.2, 1.3, etc. can be used for elements that are very close in importance.

Source: (Saaty, 2005)

In addition to continuous value criteria, some criteria are best modelled as constraints. A constraint (or admissibility criterion) represents natural or artificial restrictions on the potential alternatives. Constraints are often used in the pre-analysis steps to divide alternatives into categories such as 'acceptable' or 'unacceptable'. An alternative is acceptable if its performance on one or several criteria exceeds a minimum, or does not exceed a maximum. In practice, constraints are often modelled using conjunctive or disjunctive aggregation. With the conjunctive method, a minimum satisfaction level  $g_{min,j}$  is associated with each criterion  $g_j$ . If the performance of an alternative, with respect to different criteria, is equal or better to these minimum satisfaction levels (i.e.  $g_j(a_i) \geq g_{min,j}, \forall j \in F$ ), the alternative is considered as acceptable. Otherwise, the alternative is considered as unacceptable. With the disjunctive method, the alternative is considered acceptable as soon as at least one satisfaction level is exceeded.

The evaluation of alternatives may be expressed according to different scales (ordinal, interval, ratio). However, a large number of multicriteria methods require that all of their criteria are expressed in a similar scale. Standardizing the criteria permits the rescaling of all the evaluation dimensions to between 0 and 1<sup>23</sup>. This allows for comparisons between and within criteria.

There are a large number of standardization procedures. In all of them, standardization starts from an initial vector  $(g_j(a_1); g_j(a_2); \dots; g_j(a_m))$  to obtain a standardised vector  $(r_{1j}; r_{2j}; \dots; r_{mj})$  with  $0 \leq r_{ij} \leq 1; \forall j \in F$  and  $i = 1, \dots, n$  ( $n$  is the number of alternatives). The most common

<sup>23</sup> There is no reason, other than simplicity, for the choice of 0 to 1. Any range of values could be used. 0 to 100 is also common.

standardization procedure in multicriteria decision making is the linear transformation procedure, or maximisation. It associates with each alternative  $a_i$  and for each criterion  $g_j$  the percentage of the maximum over all alternatives. It is applied when the performance score of a particular alternative is assessed in relation to the maximum score for all the alternatives. Different approaches are used depending on whether the criterion expresses a cost or a benefit.

For a benefit criterion:

$$r_{ij} = \frac{g_j(a_i)}{\max_i g_j(a_i)}, \quad i = 1, \dots, n; \quad j \in F \quad (5.2)$$

For a cost criterion:

$$r_{ij} = 1 - \frac{g_j(a_i)}{\max_i g_j(a_i)} + \frac{\min_i g_j(a_i)}{\max_i g_j(a_i)}, \quad i = 1, \dots, n; \quad j \in F \quad (5.3)$$

As shown in the performance table (Table 5.1), weights  $w_1, \dots, w_m$  are assigned to the criteria. Weight  $w_i$  reflects the relative importance of criteria  $C_i$  to the decision maker, and is generally assumed to be positive. The weights of the criteria are usually determined on a subjective basis. They represent the opinion of the decision maker or synthesize the opinions of a group of experts using a group decision technique, such as the Delphi Method (see Linstone and Turoff, 1975).

Various weighting techniques are used. Weights can be assigned directly, such that the assessor specifies the weight values for each criterion individually. These are then normalised by dividing each assigned weight by the sum of all weights.

$$w'_i = \frac{w_i}{\sum_{i=1}^n w_i}, \quad i = 1, \dots, m; \quad w_i > 0 \quad (5.4)$$

Note that  $\sum_{i=1}^n w'_i = 1; \quad \forall i \in F$

Weights can also be assigned using a rank ordering, where the assessor applies a rank order to the criteria. Normalised weights can then be calculated from these rankings using either the expected value method (5.5) or the rank sum method (5.6).

$$w'_i = \sum_{k=1}^{n+1-i} \frac{1}{n(n+1-k)}, \quad k = 1, \dots, n; \quad n \in \mathbb{Z}_{\geq 1} \quad (5.5)$$

$$w'_i = \frac{n+1-i}{\sum_{k=1}^n (n+1-k)}, \quad k = 1, \dots, n; \quad n \in \mathbb{Z}_{\geq 1} \quad (5.6)$$

Where  $w'_i$  is the normalised weight for criterion  $i$ ,  $k$  is the rank of the alternative, and  $n$  is the total number of alternatives being ranked.

Alternatively, a pairwise comparison technique can be used, as with the Analytic Hierarchy Process (see Saaty, 1980), where the assessor evaluates all the unique criterion pairs and assigns weights according to a predefined scale (such as the Likert Scale).

The final score for an alternative is then given by:

$$A_j = \sum_{i=1}^m w'_i \times r_{ij}, \quad i = 1, \dots, m; \quad j = 1, \dots, n \quad (5.7)$$

Where  $r_{ij} = f(a_{ij})$

## 5.2 Spatial Multiple Criteria Assessment (SMCA)

### 5.2.1 Overview

Although the development of multicriteria decision analysis has always been closely linked with advances in microcomputer technologies (see Power, 2003), since the early 1990s, the integration of GIS and multicriteria decision analysis has attracted significant interest as GIS gained in popularity.

Spatial decision problems often involve a large set of feasible alternatives and multiple, sometimes conflicting, evaluation criteria. The evaluation may be conducted by a range of individuals (decision-makers, managers, stakeholders and interest groups) who are characterised by unique, differing views on the relative importance of the various criteria. Many spatial decision problems are, therefore, suited to a GIS-based multicriteria decision analysis (Laaribi, 1996; Thill, 2000). GIS and MCDA are particularly well suited for combined analysis. GIS techniques and procedures have an important role to play in analyzing decision problems, whereas MCDA provides an established set of methodologies for structuring decision problems, and designing, evaluating and prioritizing alternative decisions.

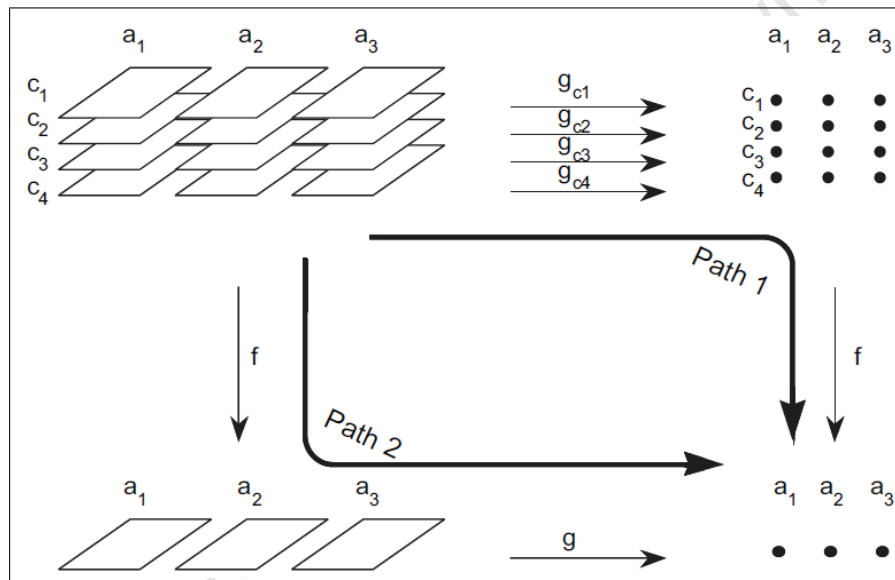
Some of the earliest work in the field was done by Diamond and Wright (1988), on the development of an integrated decision support system that addressed the multi-objective nature of the site-screening processes. Their system consisted of three distinct decision making technologies: a GIS; a rule-based system; and a multi-objective programming model that was linked together in the form of an integrated spatial information system. Site selection problems have, subsequently, become an important application of SMCA methodologies (see Carver, 1991; Eastman et al., 1995; Kontos et al., 2005; Moreno and Seigel, 1988).

Explicitly spatial criteria are required in many of the decision problems that have spatial characteristics. For example, in the context of a site search problem, site characteristics such as area, shape, proximity and compactness are often important factors in the decision (Brookes, 1997; Church et al., 2003). On the other hand, decision problems may involve criteria which are implicitly spatial. Criteria, such as the equity of income distribution, population density or the costs of solid waste disposal, require spatial inputs such as distance, proximity, accessibility, elevation or slope, in order to be used in the calculations (see Antoine, 1997; Macdonald, 1996). It should be noted that these two categories of criteria are not mutually exclusive, many studies involve both explicitly and implicitly spatial criteria.

Whereas in a non-spatial MCDA, attributes are compiled in a performance table, in a spatial decision problem, the information used to assess the criteria is described by a set of maps. Zucca et al. (2008)

identifies alternative strategies to process this information and derive an ordered ranking of the alternatives. They contend that the spatial decision problem can be visualised as either a ‘table of maps’, or a ‘map of tables’, with each conceptualisation taking a different analysis path, the results of which can then be transformed into a ranking of the alternatives.

According to their concept, the aggregation function can be split into two operations, aggregation in the spatial dimension, and aggregation in the ‘thematic’ (criterion) space. Following Path 1 in Figure 5.2, the spatial information is first aggregated to a non-spatial value for each criterion, separately. Then, traditional multiple criteria assessment (MCA) techniques can be used to derive the final utility for each alternative. Following Path 2 instead, first the theme maps are combined through MCA techniques to obtain a level of suitability map for each alternative, thus reflecting the spatial performance of the alternative. Each map can then be aggregated to a single non-spatial value. Following Path 2 it is possible to perform a multicriteria evaluation using both spatial criteria along with non-spatial criteria, and the assessment can be carried out without losing the spatial dimension.



**Figure 5.2:** Decision Paths in Spatial Decision Problems.

Source: Adapted from Zucca et al. (2008)

The value of GIS as a decision support tool was recognised early on in its development. Densham (1991) discusses the concept of a Spatial Decision Support System (SDSS), noting that they are specifically designed to support decision research processes for complex spatial problems. Using the example of the location of a bank branch, he notes that many of the questions typically asked around siting a branch, have a spatial component. The ability of a GIS to simply implement a set of algorithms to analyse spatial data greatly simplifies the data analysis process for decision makers.

Carver (1991) provides a good comparative assessment of the earliest spatial analysis work using GIS and more recent SMCA methods. The early methods used map overlay techniques to identify areas that simultaneously conformed to a number of factors. Although a useful application of the technology, this site screening method could not differentiate between the sites identified in terms of their relative suitability. In the paper Carver revisits an earlier such screening study by Openshaw et al. (1989) to identify candidate sites for a nuclear waste disposal facility in the UK, applying SMCA techniques to assess the utility of the technique in light of the recommendations made then. The work demonstrates

the power of being able to, not only identify candidate sites, but to be able to discern amongst those sites how their suitability varies according to each factor, and how the application of factor weighting alters the preference rankings. Along with Janssen and Rietveld (1990) and Aubert (1986), Carver's work is amongst the earliest forays into SMCA.

### 5.2.2 SMCA and Transportation

Since these earliest studies, SMCA has been applied to a wide range of spatial decision problems, most commonly, as discussed above, to site selection problems, but it is also increasingly, being applied to transportation problems. Since the decision problem being addressed in this thesis relates to transportation, it is useful to discuss some of the transportation problems that have already been addressed using this method.

A natural extension of the site selection procedures incorporating SMCA methods is the suitability analysis of corridors for linear infrastructure projects. To date, a fair amount of research has been carried out into the application of GIS for the environmental impact assessment of road projects. Li (1999) used a GIS-based weighted overlay method to evaluate the environmental impact of alternative alignments for a section of road from Nanjing City to Hefei City, in China. They used a simple overlay technique, but factored each layer with weights derived using the Analytical Hierarchy Process to account for the impact significance of the factors considered. Low impact road alignments were then identified by stringing together those polygons with the lowest overall cumulative impact.

Blaser et al. (2004) reported on the findings of a research project conducted for the Colorado Department of Transportation, describing the development and application of a GIS, remote sensing databases and analysis models for cumulative effects assessments (CEA) of growth associated with transportation infrastructure. They present four assessments using GIS, a habitat suitability study and a land use change analysis that demonstrates commonly used GIS overlay and distance techniques, a study linking a spatial database with commonly used flood design procedures to measure hydrologic impacts due to land use change, and a study that uses a range of techniques for development growth modelling. Specific applications for CEA are given in the land use and hydrologic studies. Brown and Affum (2002) describe a GIS-based environmental modelling system for assessing the environmental effects of road traffic plans. Their system utilises the capabilities of GIS to integrate the output from a transport planning activity with land use information to model the environmental impacts of different road traffic scenarios.

SMCA has also been used to solve routing problems for utility infrastructure. Rescia et al. (2006) conducted a study in which they made an assessment of the environmental impact of a pipeline project in Argentina. They constructed an environmental impact index and applied it to environmentally sensitive sections of three pre-defined alternative routes. They calculated the cumulative index scores for each route segment for the initial and post construction states. The difference between the initial and post construction state for each route formed the basis against which an alternative was selected.

Sharifi et al. (2006) describe the use of land use, environmental, engineering and social criteria to evaluate the performance of various railway networks for the Klang Valley in Malaysia. The study demonstrates SMCA's ability to incorporate a wide range of concerns (social, environmental, strategic development) into all phases of the planning process.

Yusof and Baban (2004) describe a least-cost pipeline routing exercise they conducted using SMCA in Malaysia. They used factors related to the land use and the land cover type to determine a set of 'friction factors', which they then mapped. A least-cost routing algorithm was then employed to extract suitable routes for the pipelines.

Bailey et al. (2005) reports on the use of a SMCA for transmission line routing. They note that the development of electricity grid infrastructure is often hampered, or even stalled completely, by public opposition to proposed alignments. Their methodology was somewhat similar to that of Yusof and Baban (2004), developing an 'impedance surface' to represent spatial costs from a MCA procedure that considers a range of relevant factors. Similar work was reported by Paulus et al. (2006) for telecommunications infrastructure and by Thomaidis and Mavrakis (2006) for gas pipeline routing.

This work highlights the power of using SMCA in routing problems, a common problem in infrastructure planning, but especially relevant for road infrastructure planning. The studies mentioned relied on the spatial allocation of generalised costs related to the infrastructure being planned, conceptualised as a 'friction' or an 'impedance' along the route. In Rescia et al. and Sharifi et al.'s work, these costs were then assigned to predefined alternatives, which were then assessed in terms of the cumulative costs along the routes. Bailey et al.; Paulus et al.; Yusof and Baban and Thomaidis and Mavrakis' work introduces the concept of using a least-cost routing algorithm to identify appropriate routes from within the data.

This approach is ably demonstrated by Keshkamat et al. (2009), who applied these concepts, along with the broad range of criteria used by Sharifi et al., to the identification of suitable alignments for the Via Baltica Highway in Poland. The method used improves upon previous work by Grossardt et al. (2001), to identify the best routing based upon spatial criteria. Grossardt et al. (2001) combined stakeholder priorities, such as economic development, connectivity, ecological factors, recreational areas, and so forth, to generate a continuous composite cost or cumulative impact surface using MCA processes. This surface is represented by a raster map in which every pixel corresponds to a weighted sum of the scores of the individual impedance elements. A cost-weighted distance algorithm was then used to identify the least cost path across this SMCA surface. The cost-distance function uses the impedance map to determine the least cost path between a designated origin and any other point on the map. The end result is a route which delineates the least cost path between the start and end points.

Keshkamat et al. argues that a weakness of this method is that it tries to find the path of least impedance regardless of the final length of the result or the presence of existing road segments. They maintain that such a route would not only be uneconomical to construct and maintain, but also inefficient in terms of vehicle-kilometres travelled. Their approach, although similar, employs tools from ArcGIS's (by Esri inc.) Network Analysis module to assign the cumulative impedance costs to links in the existing network using a line weighted mean. The least cost routing algorithm can now be employed to identify the lowest overall cost combination of existing segments in the network to define a final routing.

### 5.3 SMCA and Context

The methods developed by Grossardt et al. and Keshkamat et al. displays the possibilities of combining the spatial analysis routines and technologies of GIS, and the decision making methodologies of MCA to analyse and solve complex spatial problems. Returning to the decision problem identified at the beginning of this chapter: "*Given a set of information regarding a particular location and information regarding the various modes of transport, which mode is best suited to that location*", it becomes apparent that aspects of the methods discussed so far could be applied to solve this problem.

There are, however, significant differences between the work done by these authors and the requirements of the decision problem stated above. Firstly, previous work in infrastructure routing has always been on a large, regional or national scale. The decision problem, to be assessed in this research, calls for an analysis on a much more localised, neighbourhood or possibly, citywide scale. This has important implications, since contextual variations from point to point in a neighbourhood are much different

from those across entire provinces or regions. To capture these differences, data at significantly higher levels of disaggregation will be required for the assessment to be successful.

Also, the previous studies considered routing problems, where routes that were generated from the assessment, or that were selected beforehand, were the alternatives. All of the studies were designed for only a single mode of transport. In the case of this research, the route is fixed, and instead, the modes of transport are the alternatives. This is a critical distinction that raises an important issue.

The previous studies used standardised criterion maps of the various indicators to estimate the spatial variation in the suitability of the proposed infrastructure (see Table C.1 for an example of a criteria tree). The accumulation of these level of suitability map values was then used to derive routings using a least-cost algorithm. The indicators used in this research, however, would be used to describe the spatial context, and the accumulation of these contextual level of suitability map values would be used to inform modal suitability. There are, once again, key conceptual differences here.

Firstly, the hypothesis that modal priority can be informed by contextual suitability implies that, at any location, a ranking of modes can be derived based upon the contextual information at that location. The contextual suitability is, therefore, different for every mode at every location. However, this contextual suitability must be derived from a single data set. Therefore, the *interpretation* of how a mode relates to a criterion must differ for every mode.

It was already established in Chapter 4 that each criterion has different implications for every mode. However, this now means that every mode will require its own level of suitability map. Where previous research only required a single map output for any scenario, this evaluation will require multiple maps, that multiple being equal to the number of alternative modes being assessed. A modal priority ranking can then be defined by comparing the suitability scores for each mode at every location.

Since a distinct level of suitability map is required for each mode, a unique standardisation scheme has to be devised for each mode as well. In Table C.1, Keshkamat et al. lists the criteria categories (called 'themes' here), the individual criteria or indicators and an explanation of how they interpreted each criterion in the study. This singular interpretation of the costs or benefits attributed to each criterion, is applicable to all the scenarios, with these being differentiated from each other by varying the weights applied to the themes. This research, however, will require such a criteria tree for each alternative (each mode). Standardisation, which is used to interpret criteria data as either a cost or a benefit, and to normalise criteria scores to a common unit for arithmetic operations, now becomes the vehicle for expressing the relationship between the mode and the criterion. A cohesive standardisation framework must, therefore, be designed that allows for the relationships between modes and criteria to be expressed in an ordered fashion.

## 5.4 Standardisation Framework

Standardisation is the procedure used to interpret a criterion in terms of its implications for the alternatives being evaluated. Criteria can be interpreted as being either a cost or a benefit. Furthermore, criteria can be either categorical (such as for land use types) or continuous (such as for population density). In either case, the relationship to each mode must first be identified, and then a mathematical transformation can be applied in order to standardise the variable. To illustrate the approach taken, the case of land uses (a categorical variable) and household densities (a continuous variable) are considered.

### 5.4.1 Categorical Variables

Land uses have different characteristics in the types and volumes of traffic that they generate, the time of day and day of the week that peak volumes are generated, and the traffic needs specific to the land use. Consequently, when planning infrastructure to service any particular land use, these differences need to be considered and the design altered as required. By considering land use as an explicit variable in the criteria tree, these differences, and the costs or benefits of prioritising a particular mode as a result of them, can be captured.

These costs and benefits can vary according to the point of view taken in the analysis. For example, in an industrial area it is reasonable to expect high volumes of delivery vehicles and heavy vehicles. In fact, the businesses in these areas depend upon the ease of access provided to these vehicles. From this point of view, maximising the mobility of heavy vehicles (and, in fact, all motorised vehicles) is an important consideration. However, industrial areas, especially in South Africa, also see high levels of pedestrian traffic in the form of workers walking to and from work. The conflicts between pedestrian road users and vehicles factor as a significant cost from their perspective.

The question centres around the values that are imposed on the evaluation. These values can be translated into impacts through value statements. In the example of industrial areas, the following statement could be used to interpret the values from the perspective of NMT road users:

*“We want to maximise the safety of all road users.”*

A number of steps can be taken to maximise safety. As was seen in Chapter 3, limiting speeds and minimising potential conflict points are ways of improving safety. A further strategy is to afford increased protection to the most vulnerable road users. To express this value quantitatively, an appropriate indicator for safety must be identified. In this instance, the safety of each mode can be expressed qualitatively in terms of operating speeds. Since speed is an important strategy to improve safety, each mode can be listed according to its relative operating speed, and then standardise by giving the slowest mode the highest rank. This yields a qualitative ranking, as shown in Table 5.3.

**Table 5.3:** Safety maximisation qualitative ranking

	Car	Public Transport	Pedestrian	Bicycle	Freight
Relative Speed	+++++	+++	+	++	++++
Standardised Ranking	+	+++	+++++	++++	++
Standardised Score	0.00	0.50	1.00	0.75	0.25

From this perspective, ‘Pedestrian’ and ‘Bicycle’ are the slowest modes, and so score highest, whereas ‘Car’ and ‘Freight’ are the fastest modes, and so score the lowest. However, an alternative view could also be taken. Since an industrial area is being considered, the position could be taken that mobility for delivery vehicles is the most important consideration. The value statement could be expressed as:

*“We want to maximise the mobility of delivery vehicles.”*

In order to maximise mobility for delivery vehicles the highest priority must be given to speed, in which case the qualitative ranking will look as shown in Table 5.4, with the fastest mode receiving the highest score.

Clearly, with each land use, multiple value positions could be taken that would each yield different qualitative rankings. Furthermore, since the concerns around traffic varies between land uses (the

**Table 5.4:** Mobility maximisation qualitative ranking

	Car	Public Transport	Pedestrian	Bicycle	Freight
Relative Speed	+++++	+++	+	++	++++
Standardised Ranking	+++++	+++	+	++	++++
Standardised Score	1.00	0.50	0.00	0.25	0.75

concerns in a commercial district are different to that of a residential district), it is not possible to assume one value position for all land uses. Instead, the qualitative ranking must be individually defined for all land uses. The result can best be described as a value matrix. Each land use option is assessed from the value position that is chosen as being best suited to it. This yields an ordinal scale of benefits as seen in Table 5.5.

**Table 5.5:** Criteria, value statements, indicators and standardisations

Option	Value	Indicator	Car	Public Transport	Pedestrian	Bicycle	Freight
Residential	Safety for NMT	Lowest Speed	0.25	0.50	1.00	0.75	0.00
Commercial	Access for patrons	Highest Volume	0.50	0.75	1.00	0.25	0.00
Industrial	Mobility for vehicles	Highest Speed	0.75	0.25	0.50	0.00	1.00
Education	Safety for learners	Lowest Speed	0.25	0.50	1.00	0.75	0.00
Sports and Recreation	Access for spectators	Highest Volume	0.25	0.50	1.00	0.75	0.00
Vacant Land	Mobility for passersby	Highest Speed	1.00	0.75	0.00	0.25	0.50
Medical	Access for patients	Highest Volume	0.25	0.75	1.00	0.50	0.00
Office	Access for workers	Highest Volume	0.50	0.75	1.00	0.25	0.00

In Table 5.5, different value statements (to be read as: “We want to maximise [value] by conferring the highest rank to the mode with the [indicator]”) are used to express priorities varying between safety, access and mobility. In this case, only two indicators, the speed of the mode and the volumes by mode, are used to distinguish between the alternatives in terms of the value statement. Different impacts can be developed for different land use options despite the fact that the same indicator is used. The ordinal scale used for scoring is a result of the qualitative ranking used to interpret the value statement. The disadvantage of an ordinal scale is that the extent of preference is lost. Also, it is not possible to confer the same rank to different alternatives (or in this case, modes of transport).

Figures 5.3 and 5.4 show 1:100 000 extracts of the standardised land use score maps for Cape Town for cars and PT. In the figures, green areas are more suited to a mode than red areas. In this study, land

uses were differentiated by zoning type. Although particular land use types may vary somewhat within a given zone, for the purposes of developing the methodology, this variation is not significant. Further discussion on this topic is presented in Section 6.1.1 of this thesis.

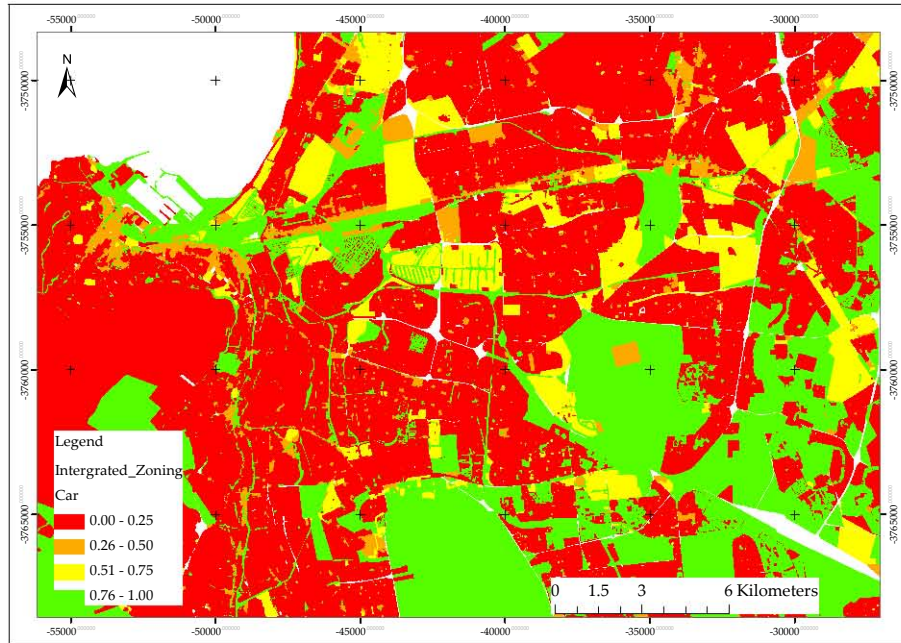


Figure 5.3: Extract of standardised land use scores for cars for Cape Town



Figure 5.4: Extract of standardised land use scores for PT for Cape Town

### 5.4.2 Continuous Variables

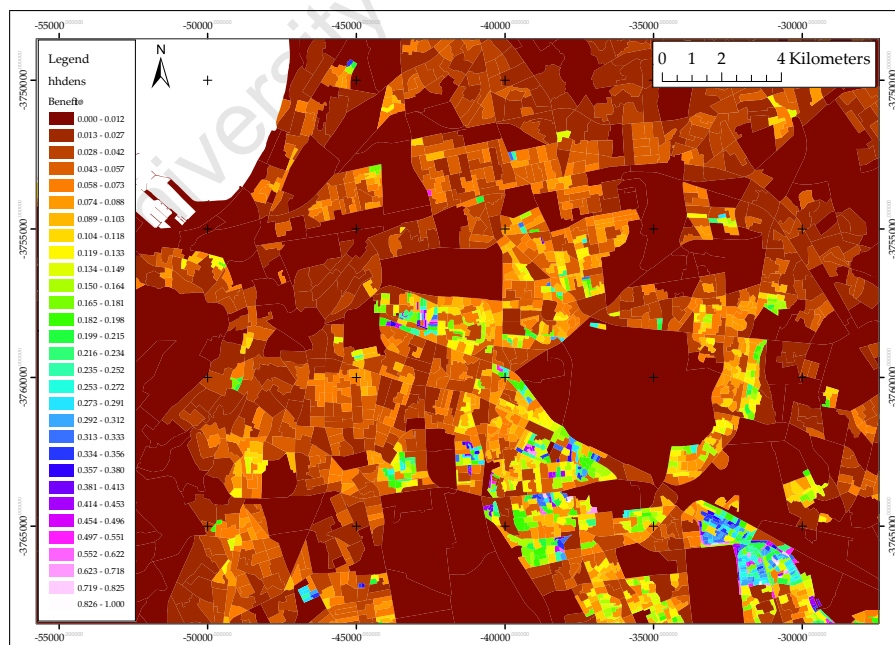
Higher density areas are better suited to a more intense activity mix, and can support better quality PT (City of Cape Town, 2009a). Density is often cited as an indicator of trip frequency and trip length (Chatman, 2008; Chen et al., 2007). Density is also an indicator of modal split, in that in less dense regions, with a higher uniformity of land uses, trips are more often made using motorised modes that are better suited to longer trips, whereas in high density areas, with a higher mix of land uses, trips can be shorter, and so better suited to PT and NMT modes (Kockelman, 1996; Limtanakool et al., 2006; Zhao et al., 2002).

Density is, therefore, important in defining the character of a location in relation to which mode should receive a higher priority. In low density residential environments, it is not unreasonable to assume that many trips will be made using private vehicles, especially when trip attractors, such as shops or work places, are far away, and when PT facilities are unavailable. So, although mobility may not be the overriding concern in a residential context, considering only the influence of density, the lower density gives justification to considering the needs of the private automobile as primary.

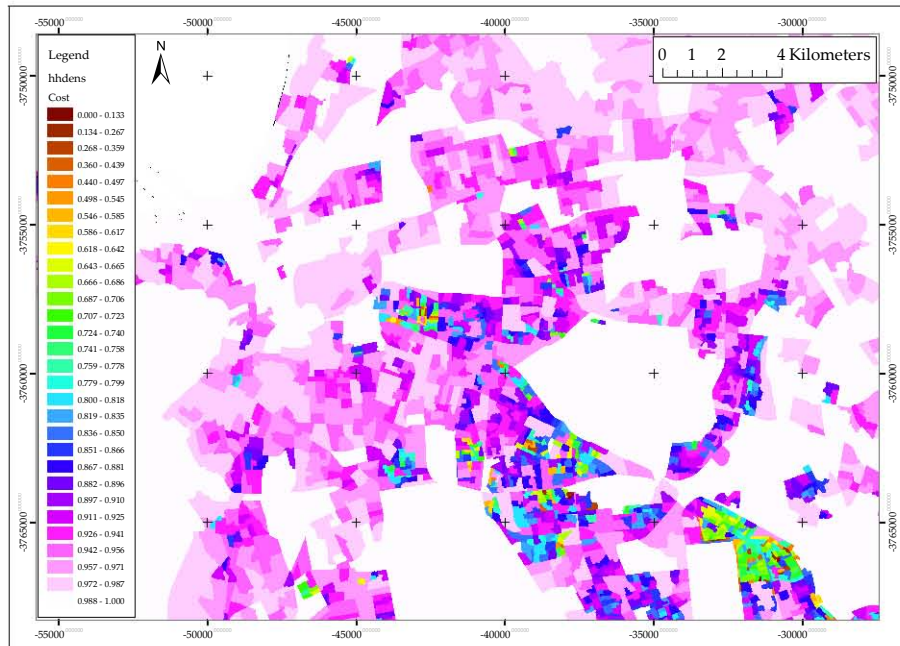
Returning to the value statements that were used to identify indicators for the criteria, if density is assumed to be a continuum of some range of values from high to low, then in higher density areas a higher preference should be given to NMT and PT and in lower density areas a higher preference should be given to PMT. The value statement can be stated as:

*“We want to maximise the mobility of the majority of road users”*

The relationship between density and modal priority can be modelled as a simple linear function. Each mode is modelled in terms of its modal priority given a certain density. After applying Equations (5.2) and (5.3) to household density data for Cape Town, Figures 5.5 and 5.6 illustrate how modal priorities shift with increasing densities.



**Figure 5.5:** Extract of benefit standardised household density scores for Cape Town



**Figure 5.6:** Extract of cost standardised household density scores for Cape Town

The same method was used to standardise remaining factors. Table 5.6 gives an overview of the criteria used and the standardisations that were applied to each criteria. The spatial cost/benefit standardisation approaches adopted for each criterion by mode are presented in Table 5.7.

Proximity maps are calculated by using a distance weighted score for each mode, depending on whether proximity to the feature is a cost or a benefit for that mode. Equations (5.8) and (5.9) are used to calculate the standardised score for any particular location.

For a benefit score:

$$a_m = \frac{d_i}{d_{\max}} \quad (5.8)$$

For a cost score:

$$a_m = 1 - \frac{d_i}{d_{\max}} \quad (5.9)$$

Where  $a_m$  is the score for the alternative,  $d_i$  is the distance to the nearest stop at point  $i$  and  $d_{\max}$  is the maximum distance from a PT stop for any location in the data set. Figure 5.7 shows an extract of the standardised proximity scores for Cape Town, where proximity to a PT stop is a cost for a mode. The black dots are the locations of PT stops (bus, taxi and train stations). Standardised maps for every mode of transport for all the criteria are provided in the Appendices.

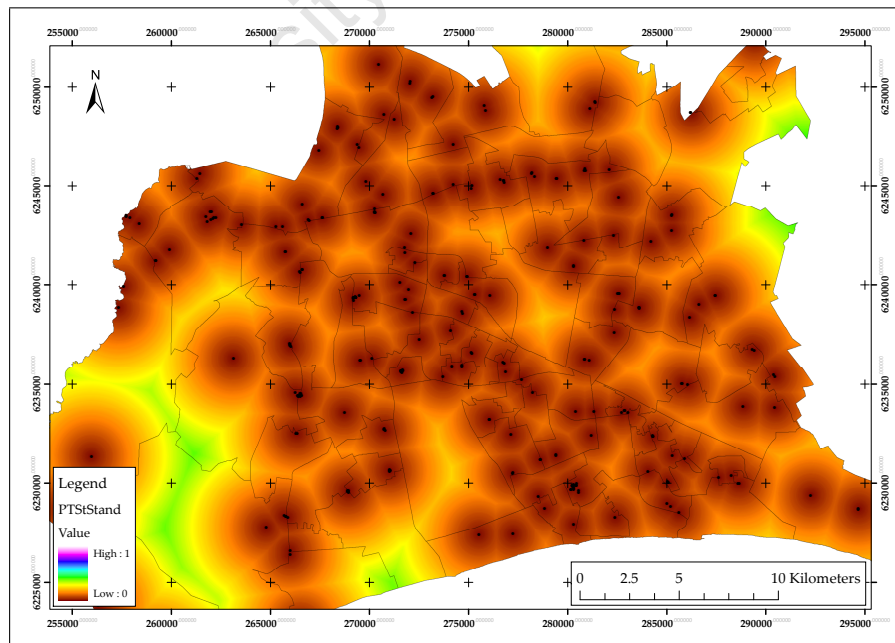
## 5.5 Aggregation

Once the standardised criteria maps have been created for every mode of transport, they can be used to evaluate the relative spatial performance of the various modes for that particular criteria. However, it is the accumulated scores for each mode that is required to assess the overall suitability of a mode

**Table 5.6:** Criteria, value statements and indicators

Criteria	Option	Value	Indicator
Land Use <sup>†</sup>	Varies	Varies	Varies
Household Density	Low	Mobility of majority	Highest Volume
Proportion of vulnerable road users	Low	Safety of vulnerable road users	Lowest Speed
Income	Low	Mobility of majority	Highest Volume
Employment	Low	Mobility of majority	Highest Volume
Proximity to heritage sites	No	Access	Lowest Speed
Proximity to wetlands	No	Minimise impact	Lowest Speed
Proximity to ecologically sensitive areas	No	Minimise impact	Lowest Speed
Public Transport demand	Low	Minimise impact	Highest Volume
Private Car demand	Low	Mobility of majority	Highest Speed
Proximity to public transport stop	Low	Minimise impact	Lowest Speed

<sup>†</sup> Land use inputs as per Table 5.5

**Figure 5.7:** Extract of cost standardised scores for proximity to PT stops in Cape Town

**Table 5.7:** SMCA criteria tree

Category	Criteria	Mode of Transport				
		Car	Public Transport	Pedestrian	Bicycle	Freight
Land Use	Land Use	Varies	Varies	Varies	Varies	Varies
	Household Density	–	+	+	+	–
Socio-Economic	Employment	–	+	+	+	–
	Proportion of vulnerable road users	–	+	+	+	–
Environmental	Income	+	–	–	–	+
	Proximity to heritage sites	–	+	+	+	–
	Proximity to wetlands	–	–	–	–	–
Transportation	Proximity to ecologically sensitive areas	–	–	–	–	–
	Public Transport demand	–	+	+	+	–
	Private Car demand	+	–	–	–	+
	Proximity to public transport stops	–	+	+	+	–

+ denotes a benefit criterion for the mode

– denotes a cost criterion for the mode

for a particular location. Referring to Figure 5.2, Sharifi et al. (2006) described two approaches to conducting a SMCA. The approach required for this analysis is that outlined by ‘Path 2’, where the map scores are aggregated spatially, thereby producing a spatial level of suitability map.

This process requires that all maps be converted to raster format. A raster consists of a matrix of cells (or pixels) organised into rows and columns (or a grid), where each cell contains a value representing information, such as temperature<sup>24</sup>. In a GIS, raster images are a particularly useful way of presenting continuously varying data, since the value of each cell can represent any quantity independently. Rasters are, therefore, commonly used to represent elevation, rainfall, temperature, concentration and population density.

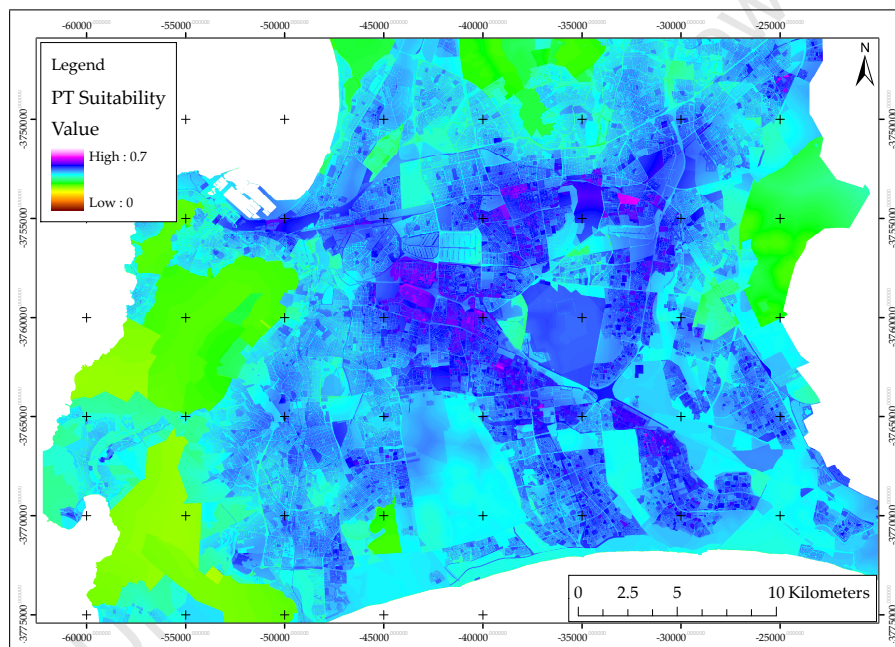
The nature of raster representations means that they are particularly useful for performing complex spatial analysis, statistics and data manipulation procedures. However, rasters do have some important limitations that are relevant to this analysis. The cellular nature of rasters means that there can be spatial inaccuracies in the representation due to the limits imposed by the raster cell dimensions - straight lines or edges can only be approximated by the jagged cell boundaries. Furthermore, raster data sets can be potentially very large. The resolution of the representation increases as the size of the cell decreases. Increased resolution, therefore, comes at the cost of increases in both disk space and processing time.

<sup>24</sup> <http://webhelp.esri.com>

Selecting a suitable cell size is, therefore, an important consideration. The size of the cells in the raster should not be so large as to lose information, since spatial variation in context could be very fine grained. However, if the cell sizes are too small, processing times can become unmanageable because of the scale of the analysis.

The cell size selected for the analysis is related to the smallest unit of data in the data set - in this case, a single parcel of land - taken as being, approximately, 10 m by 10 m. This is a useful scale to work at, since road reserves, themselves, are seldom narrower than 10 m, meaning that data on either side of the road will not necessarily become mixed.

The standardised criteria maps, now all at the same raster scale, can then be weighted by multiplying the value of each raster cell by the weight that has been defined for that criterion. These weighted maps can then be summed to produce a cumulative suitability score for every location. This score can then be normalised to a base of one. Figure 5.8 shows an extract of the results of such an aggregation procedure for PT, where all the criterion were equally weighted.



**Figure 5.8:** Extract of equally weighted level of suitability map for PT in Cape Town

In Figure 5.8, a higher score indicates a higher suitability for the mode at that location. The variation in the spatial suitability is quite apparent. By extension, therefore, the variation in the context is equally distinct, since context is being expressed in terms of modal suitability. The method requires that these level of suitability maps be calculated for every mode of transport (for every alternative).

At each location, therefore, there will be five suitability scores, one for each mode of transport. A ranking of these scores will give the relative priority of each mode at every location. To evaluate these scores, they need to be extracted and compared to each other in a spreadsheet. Since the analysis aims to identify a priority ranking for modes of transport along a road, the scores along the road's centreline must be extracted for each level of suitability map. The raster cells of each level of suitability map are 10 m by 10 m, so extracting values at 10 m centres along the centreline of the road would yield the highest possible information resolution for the analysis.

However, the score from each point only relates to the information that is directly beneath that point. Since this is almost invariably on the road surface itself, these values are not necessarily representative of the locations context. Figure 5.9 shows a 1:5 000 scale extract of the level of suitability map in Figure 5.8 for an arterial route in Cape Town. In the figure the white points are the score extract locations along the road at 10 m centres. In the figure it can clearly be seen that the point values would not capture the detail of the context along the route.

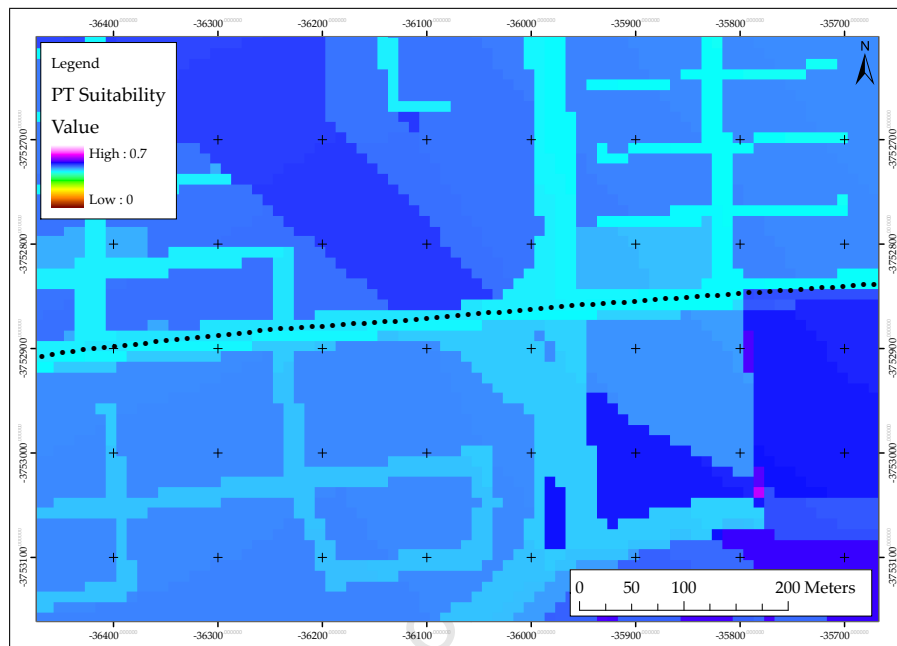


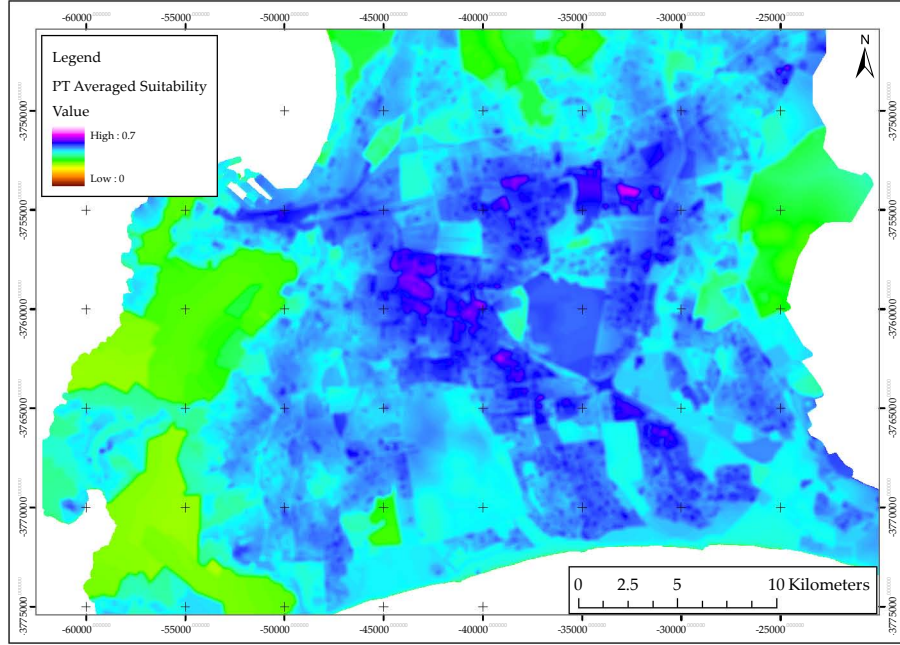
Figure 5.9: Example of point value extract locations along a road

### 5.5.1 Neighbourhood Averaging

To overcome this problem, an average of the scores from the area surrounding each point must be calculated. The area from which an average was calculated was taken as being a circle of radius 500 m from the centre of the cell. The radius was selected on the basis that 500 m is often cited as being a comfortable or reasonable walking distance (Cervero, 2010; Hossain et al., 2010; Shimizu et al., 2010). In urban areas, bus stops are often located at between 500 m and 1 000 m intervals, depending upon the intensity of activity along the route and the type of development being served (CSIR, 2000), and this distance is also often considered a reasonable estimate for the size of the catchment area for local businesses and primary schools. The sensitivity of the average score to the size of the influence area was also investigated. This is discussed in more detail at the end of this section.

Using a 500 m radius influence area, the scores used in the ranking would represent the average suitability of the mode for an area of, approximately, 78.54 hectares around each point. The score of each cell in the original level of suitability map is then recalculated as this average. The calculated raster image now constitutes an averaged representation of the spatial suitability of each mode across the study area. Point scores extracted along the centreline of the road from this average level of suitability map will now be more representative of the context surrounding the road. The results of this exercise are shown in Figure 5.10.

It is important to be aware of the assumptions made in creating this map. When averaging the scores



**Figure 5.10:** Extract of averaged level of suitability map for PT in Cape Town

across the map, the assumption is made that all points, within the area of influence surrounding the cell, affect the context of the cell's location equally. Although this may be a reasonable assumption in many cases, there may well be areas where this is not true. Physical barriers such as major roads, railway lines and canals, and geographic features such as rivers and parks, tend to diminish the extent of the interaction between areas on either side of them. Also, the further apart locations are from one another, the less influence they have on each other.

Modelling this diminishing influence factor, and applying it to the influence area around the cell, mitigates the barrier problem, as areas beyond the barriers will exert a smaller influence on the final cell score, and the problem of diminishing influence with distance is also dealt with. To test the effects of averaging the SMCA results in this way, the assumption was made that the influence that two locations have on each other decreases linearly with distance, and since a circle of 500 m radius was previously selected as the size of the catchment area for any cell, it was assumed that the contextual influence decreases to zero over this distance as well. As a result, the new averaged modal suitability score for any location on the map is given by Equation (5.10):

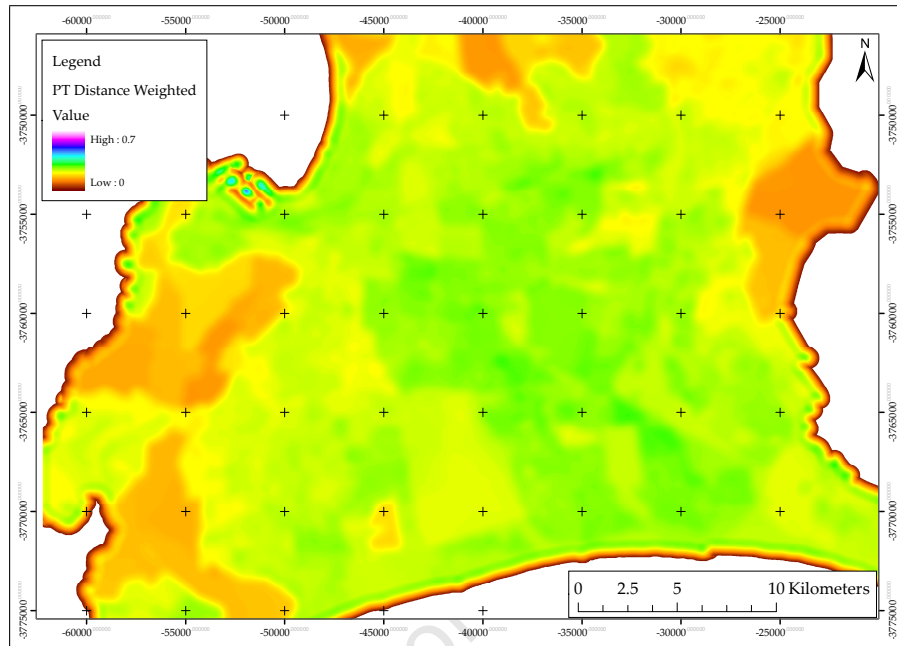
$$S_i = \frac{\sum_{j=1}^n \left( \frac{d_{ij}}{500} \times A_j \right)}{n} \quad (5.10)$$

Where  $S_i$  is the averaged, distance-weighted score for the mode at point  $i$ ,  $A_j$  is the accumulated criterion score for point  $j$ ,  $d_{ij}$  is the distance between point  $i$  and point  $j$  and  $n$  is the total number of cells in the averaging area. Note that  $A_j$  was calculated previously using:

$$A_j = \sum_{i=1}^m w'_i \times r_{ij} \quad (5.11)$$

The distance weighting,  $\frac{d_{ij}}{500}$ , applied to the cell value  $A_j$ , is given by the hypothetical height  $h_{xy}$  of a point  $x$  cells along the horizontal axis, and  $y$  cells along the vertical axis away from cell  $i$  on the raster grid, on a conical weighting surface of height 1. Figure C.1 gives a conceptual overview of the conical distance weighting scheme.

The results of this weighting method are presented in Figure 5.11.



**Figure 5.11:** Extract of averaged distance weighted level of suitability map for PT in Cape Town

An analysis of the differences between the two weighting approaches reveals the information presented in Table 5.8.

**Table 5.8:** Comparison of weighting scheme results

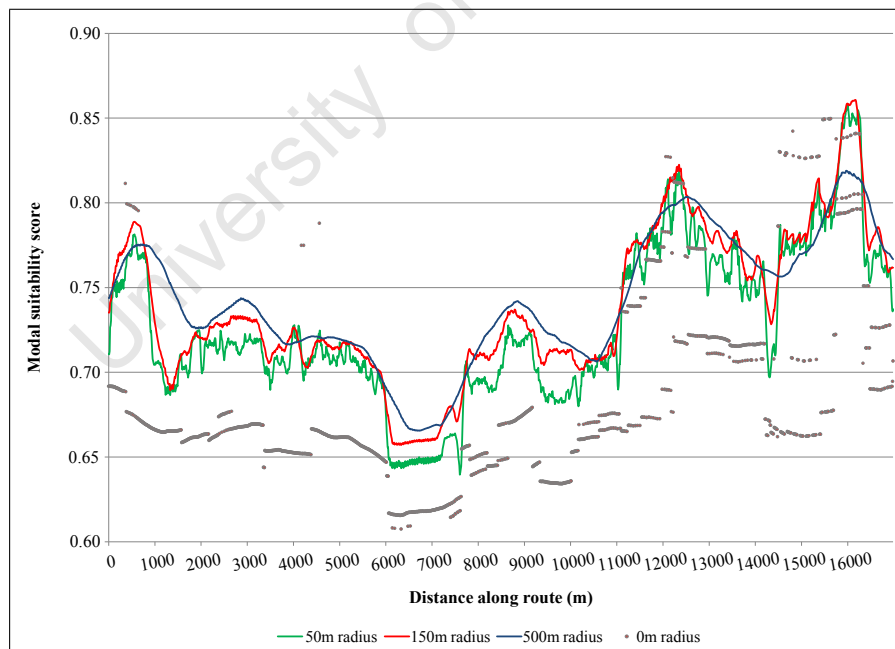
Statistic	SMCA Scores	Neighbourhood Averaging		Distance Weighted Averaging	
		Result	Difference	Result	Difference
Min	0.221	0.221	0.000%	0.000	-99.901%
Max	1.000	0.863	-13.697%	0.447	-55.272%
Range	0.779	0.642	-17.593%	0.447	-42.582%
Mean	0.538	0.536	-0.368%	0.175	-67.401%
Std Dev	0.161	0.156	-2.980%	0.058	-64.151%

A visual comparison of Figures 5.8, 5.10 and 5.11 indicates that applying a distance weighted averaging to the SMCA scores produces an output that differs significantly from the original values. From the statistics in Table 5.8, it becomes clear that distance weighted averaging does not yield representative modal suitability scores. The neighbourhood average along a straight line is given by  $(a_1 + a_2 + a_3 + a_4 + \dots + a_n)/n$ , whereas the distance weighted average along a straight line is given by

$(a_1(1 - d_1/500) + a_2(1 - d_2/500) + a_3(1 - d_3/500) + a_4(1 - d_4/500) + \dots + a_n(1 - d_n/500))/n$ . The distance weighted average score for any point is, therefore, quite different from the original value calculated from the SMCA. This has knock-on effects for the rest of the method when comparing mode suitability scores to each other.

The effect of both weighting methods is to ‘smooth out’ the results, somewhat similar to applying a low pass filter to the original spatial suitability raster. The neighbourhood averaging method produces a somewhat lower maximum score than the original data, but the minimum score is the same. The range of scores is, therefore, also slightly diminished but, importantly, the mean and standard deviation remains very similar. The distance weighted averaging process, however, has a dramatic effect on the SMCA scores, significantly reducing the minimum, maximum, mean and standard deviation of the score set. This makes using distance weighting unfeasible for the purposes of this analysis, since much of the original information is lost in the weighting process. It is, therefore, not possible, despite the advantages of doing so, to account for the effects of the distance of points in the influence area of a cell without distorting the SMCA results to such an extent that they become unusable.

This problem of data smoothing raises the issue of the sensitivity of the method to the size of the influence area. The figures presented in this section were all calculated using a 500 m influence area. As mentioned, there are good reasons for choosing this value - it relates to walking distances. However, there may be other, similarly appropriate influence area sizes that could be used that would yield better results. The issue stems from the problem of ‘noise’ in the data set, and how much noise constitutes superfluous information versus contextually important information. Figure 5.12 shows the extracted point scores from the original SMCA data (not averaged - represented as black dots), compared to the averaged data for influence areas of 50 m, 150 m and 500 m.



**Figure 5.12:** The effect of influence area size on the average suitability score

The figure demonstrates a practical reason for why it is necessary to average the raw SMCA suitability scores; there is simply too much noise in the data for any meaning to be taken from it. The scores along the road centreline are erratic, fluctuating wildly from point to point. However, averaging over just a

50 m radius (approximately the depth of one erf) dramatically improves the noise levels, revealing some of the contextual structure along the route. As more information is included in the influence area, a gradual smoothing of the profile starts occurring. The average scores start increasing, partly as a consequence of the road itself (zoned as a transport land use) having a lesser influence, but as other land uses and contextual factors start to exert more influence on the score. The effects of local contextual factors diminishes further as the radius is increased to 500 m, and local maxima and minima starts to even out.

This over simplification of the context of the route with the increase in influence area radius is undesirable, since the premise of the whole exercise is to identify local contextual variation. The ideal influence area size is, therefore, taken as being between 100 m and 200 m from the point of interest, or approximately one city block length. This is the size of the influence area that is best able to reflect the detail in the local context without including too much superfluous variation, or too much oversimplification of the SMCA data.

Prior to neighbourhood averaging, the standardised criterion rasters, created for use in the SMCA, must first undergo some processing. Specifically, cells with no data in them create 'holes' in the resultant level of suitability map. The input rasters are, therefore, all processed to remove these no-data cells by reclassifying these cells as having a zero value. No-data cells, thus, do not contribute to improving the final scores of any mode. After processing to remove no-data cells, neighbourhood averaging can be applied.

The suitability scores for the mode can then be extracted from the averaged level of suitability map by sampling the raster cell values beneath equidistant points along the road centreline, as shown in Figure 5.9. These points can be plotted for each mode on a chart as shown in Figure 5.12, and the suitability scores for each mode at every location can then be compared to identify a preference ranking for the modes along the route.

## 5.6 Résumé

The aim of this chapter was to order the criteria identified in Chapter 4 into an evaluative framework capable of producing measurable results. The chapter opened with a review of multiple criteria decision analysis. Some definitions of the field were discussed and the origins and history of the field was explored. Multiple criteria decision making was defined by Roy (1968) as “... a decision-aid and a mathematical tool allowing the comparison of different alternatives or scenarios according to many criteria, often conflicting, in order to guide the decision maker towards a judicious choice.” It was found that MCDA originated in the management sciences and developed from early work on non-linear programming.

The discussion then shifted to a sub-set of MCDA, namely Multiple Attribute Decision Making (MADM), which comprises those techniques designed to deal with a discrete set of alternatives. The various steps applied in MADM studies were discussed in some detail from a theoretical perspective. MADM methods, generally, involve the identification of a set of criteria and alternatives, which are ordered in a performance table. The scores of each alternative, with respect to each criteria, makes up the body of the table. Preferences, or weightings, may be applied to these scores using a range of techniques selected on the basis of whether the performance scores are qualitative or quantitative. The weighted scores can then be aggregated, and the summation of these scores represents the overall performance of each alternative.

A special application of MCDA (and MADM in particular), spatial multiple criteria analysis (SMCA)

was then discussed. This is the basis of the method developed for this research. SMCA was discussed in terms of its origins, and the methodological adaptations to MADM techniques used to conduct a SMCA was outlined. SMCA is used to analyse and develop solutions for decision problems with spatially varying attributes. In relation to MADM, spatial decision problems can be visualised as either a 'table of maps', or a 'map of tables', with each conceptualisation taking a different analysis path, the results of which can then be transformed into a ranking of the alternatives.

Either the spatial information is first aggregated to a non-spatial value for each criterion separately, whereafter traditional MCA techniques can be used to derive the final utility for each alternative or, instead, the theme maps are first combined through MCA techniques to obtain a level of suitability map for each alternative, thus reflecting the performance of the alternative across the space. Each map can then be aggregated to a single non-spatial value. The second method was identified as being most suited to this research, since it allows for the use of both spatial and non-spatial criteria, and it can be done without losing the spatial dimension.

The criteria maps used in the assessment have to be standardised to be able to perform the arithmetic operations used in the aggregation step on them, but also to express the criteria's relationship to the alternative (whether the criteria is a cost or a benefit in the context of the assessment). For standard SMCA studies, there is usually only one standardisation protocol developed, but the particular application of the method for this assessment requires that separate standardisation protocols be developed for each alternative. The method also develops five level of suitability maps, one for each mode of transport. The framework developed for the standardisation protocols relies on the identification of distinct value statements that are then translated into impact vectors by interpreting the value statements in terms of indicators related to the criterion, such as speeds or volumes.

In the next step, the standardised criterion maps must be converted into raster format as required. A raster cell size of 10 m by 10 m was selected based upon a balance of concerns around data resolution and data processing times. The rasterised maps can then be weighted. The weighted map scores can then be spatially aggregated using a map summation algorithm. The output is a spatial level of suitability map. This map is a representation of the spatial preference scores for that mode. A level of suitability map is generated for each mode.

The level of suitability maps must be reprocessed using a spatial averaging technique to draw in suitability information from areas beyond the road reserve into the analysis. A circular influence area of 150 m radius was identified as being a suitable size based upon the need to balance the amount of noise in the level of suitability map data and the tendency of the averaging process to oversimplify the contextual variation. This distance is also the approximate size of a city block.

The averaged level of suitability map scores along the road are then extracted at 10 m centres (since this is also the size of the raster cells) and can be plotted in a spreadsheet to compare the mode suitability scores and identify the modal ranking at every point along the road.

## Chapter 6

# Data Sources and Case Study Selection

One of the main aims of this thesis is to measure the contextual variation along roads, analyse this variation and then use this information to recommend infrastructural improvements to roads. In Chapters four and five of this thesis, the role of context in road planning, the factors that influence the context, and a method to measure and compare the context was identified. The remaining chapters of this thesis cover the measurement of contextual variation along major routes in Cape Town, the analysis of the contextual information and the development of a methodology to recommend infrastructural improvements, based upon the findings of this analysis.

This chapter starts by describing the data collection process, including a discussion of the data availability problems, as well as the limitations of the data in terms of the requirements of the methods used in the study. Three case study roads were used to develop, calibrate and test the method. A discussion on the rationale behind the selection of these routes is presented, followed by a description of the case study roads themselves.

### 6.1 Data Sources, Verification and Limitations

The methodology used in this research relies on the use of geocoded information to conduct the SMCA. Much of this information is collected by various branches of government for their own administrative and planning purposes, and is often made available to the public. Other than the need to verify the accuracy of the information collected, the need for manual data collection was minimal.

Table 6.1 lists the information used in the SMCA, the aggregation level at which the data was available, the number of unique features in the Cape Town metropolitan area for each data set and the source of the data set for each criteria.

The range of aggregation levels used in the data set is important, because this affects the confidence levels for the SMCA. The highest aggregation level used in the analysis was for income levels, with only 318 polygons in the metropolitan region of Cape Town, corresponding to the number of Sub-Place Areas used in the 2001 National Census for the Cape Town municipal area. The most disaggregated data sets used were those where the exact extent or location of each item was known (including zoning, heritage features, environmentally sensitive areas, wetlands and PT stops).

Including data at multiple levels of aggregation in the SMCA reduces the accuracy to which the cumulative scores can be calculated, thereby decreasing the confidence with which results can be reported. Having established that there may be much unknown variation within aggregated data sets, such as the income data set, considering that every household and, therefore, every erf would have

**Table 6.1:** Criteria, Aggregation Level, Number of Features and Data Sources

Criteria	Aggregation Level	Features	Data Source
Land use	Individual Erf	67198	City of Cape Town Corporate GIS Department
Household density	Small Area Level	3451	2001 National Census
Proportion of vulnerable road users	Small Area Level	3451	2001 National Census
Income*	Small Area Level	3451	2001 National Census
Employment	Sub-Place Level**	318	Regional Services Council
Proximity to heritage sites	Exact Extent <sup>†</sup>	N/A <sup>†</sup>	City of Cape Town Environmental Resource Management Department
Proximity to wetlands	Exact Extent	N/A	City of Cape Town Environmental Resource Management Department
Proximity to ecologically sensitive areas	Exact extent	N/A	City of Cape Town Environmental Resource Management Department
Public Transport demand	Travel Analysis Zone <sup>‡</sup>	842	City of Cape Town Transport, Roads & Major Projects Department
Private Car demand	Travel Analysis Zone	842	City of Cape Town Transport, Roads & Major Projects Department
Proximity to public transport stop	Exact Location	N/A	City of Cape Town Transport, Roads & Major Projects Department

\* Education level data was used as a proxy for income data (see Section 6.1.5).

\*\* In the 2001 National Census data, the Sub-Place is defined as the next spatial level up from the enumeration area in the place name hierarchy. In metropolitan Cape Town, Sub-Place boundaries roughly coincide with the planning suburb boundaries.

<sup>†</sup> Since the exact extent of these features are available, the number of features does not affect the analysis.

<sup>‡</sup> The travel analysis zones were generated by the city municipal authorities for traffic modelling purposes, and are more disaggregate than the census Sub-Place level, but more aggregated than the census Small Area Level.

a different income, an average value is already a simplification of the information. Averaging this information for larger areas decreases the reliability of this value even further.

The rasterization of the vector data sets carries these values over to the raster cells, and the SMCA process involves the summation of the individual raster cell scores. The cumulative score maps for each mode, therefore, exhibit the same amount of variation as the most disaggregated data sets. Also, since proximity maps contain continuously varying raster data, this variation carries through into the SMCA. Further variability is introduced by the neighbourhood averaging procedure. Therefore, despite the high aggregation levels of some of the data sets, the cumulative score maps exhibit significant variability from point to point. Since all the alternatives are evaluated using the same underlying data sets, the relative differences in the cumulative scores, at any point, is subject to the same data limitations. Since the aim of the analysis is to produce a ranking of modes, the relative differences between the alternatives is more important than the absolute scores.

The additional variation in the context that more disaggregated data in any one criterion would provide, is also somewhat diminished by the aggregation procedure. However, it is likely that a combination of weighting and increased disaggregation could lead to rank reversals at some locations. It is, therefore, important that the data used in the SMCA be as disaggregated as possible.

Every effort was made to collect the most recent data available in the city at the time to conduct the assessment. This has meant that the data sources used vary widely in the dates they were collected. Much of the data, especially socioeconomic and transportation data, is subject to variation with time, so that the results of the assessment may be time sensitive. This has no impact on the mechanics of the method, but it does affect the reliability of the results. It is not possible to quantify this impact without an estimate of the amount of variation each data set may undergo with time. The earliest data used in the study was the National Census data, which was collected during 2001.

A discussion of the methods used to construct, optimise and verify each data set is presented below.

### 6.1.1 Land Use

Land use can be described as the human use of land, or a description of the actual use of the land. Zoning is a land use planning device used by local governments to designate the permitted uses of a parcel of land. As mentioned in Section 5.4.1, zoning data was used as a proxy for land uses, since accurate land use data, at the extent required to conduct the analysis, is not available for Cape Town.

Land uses vary greatly, from schools to clinics to offices and shops; each particular land use has a different function and set of characteristics. In theory, the primary purpose of zoning is to segregate uses that are thought to be incompatible. Zoning bylaws regulate the kinds of activities that are acceptable on particular lots (such as open space, residential, agricultural, commercial or industrial), the densities at which those activities can be performed (from low-density housing such as single family homes, high-density or high-rise apartment buildings), the height of buildings, the amount of space that structures may occupy and the location of a building on the lot (setbacks).

The zoning data used in the study was obtained from the City of Cape Towns Integrated Zoning database compiled in 2009. Figure D.2 shows the distribution of zoning grouped into the categories unpopulated, residential, business, industrial and community facilities. Table D.1 shows the zoning categories, types and an explanation of each type for Cape Town (see City of Cape Town, 2007).

Since zoning is being used as a proxy for land uses, it was important to verify that the land uses match the zoning profiles described in Table D.1, since there is anecdotal evidence suggesting that this may not be the case. The method used to verify the land uses was to create a randomly selected points layer

overlaid onto the zoning layer. The zone description for the underlying polygon was then extracted to each point and the point layer exported to a KMZ file for use in Google Inc.’s Google Earth software, so that the land use at each point could be verified using satellite imagery and, where necessary, Google Inc.’s Streetview utility.

The number of sampling points created was calculated based upon the total number of unique zones in the population. A stratified sampling technique was not used because of the large differences in the numbers of land use categories versus the area that each category occupies. Stratified sampling would lead to very few, if any, sites from certain zones being selected, and the resultant sampling would be unlikely to satisfy the assessment requirements. Figure 6.1 shows the difference between the number of polygons per zoning category versus the percentage of the total area assigned to each zoning category.

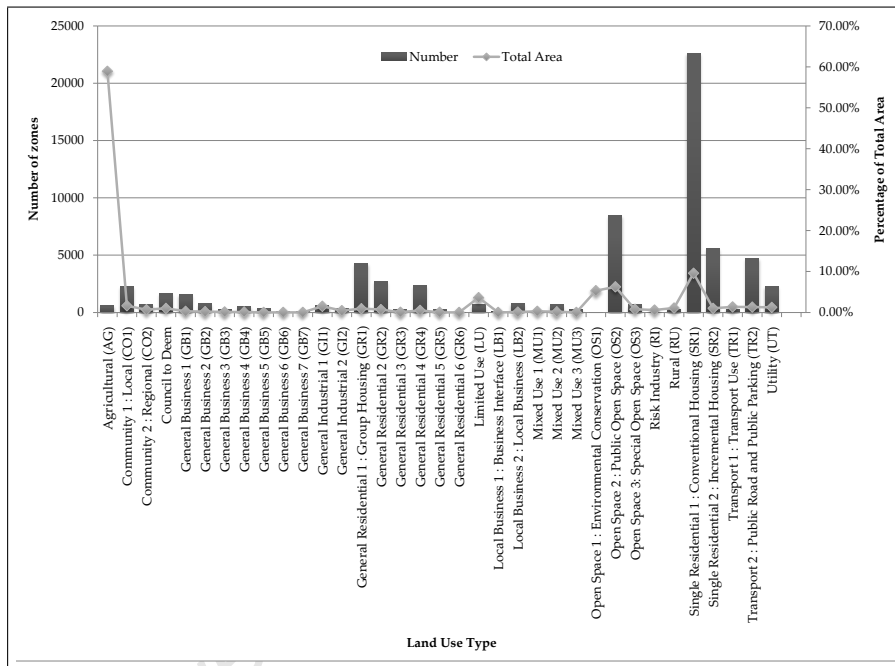


Figure 6.1: Number Of Polygons and the Percentage Area Per Zoning Category

Agricultural land accounts for, by far, the largest zoned area, more than 60%, but only a small fraction of the number of zoned sites, whereas by far the largest number of sites are zoned as Single Residential, although this zoning only accounts for around 10% of the land area. Since the number of unique items in the zoning layer is 67 198, for a 5% confidence interval at a 95% confidence level, 382 sample points were required. The listed point zoning’s were verified through visual inspection of the aerial photographs. The results are presented in Table 6.2. Based upon these results, it was accepted that the listed zoning is a reliable representation of the land use.

### 6.1.2 Household Density

The household density data used in the assessment was sourced from the 2001 National Census data published in 2003 (Statistics South Africa, 2003). The census was conducted by Statistics South Africa in October 2001. The data was made available at the Small Area Level. Figure D.3 shows the spatial distribution of households in Cape Town. Household density was calculated as the ratio of the number of households in a Small Area and the area of the polygon in hectares.

**Table 6.2:** Reliability of zoning labels

Description	Number	Verified
Agricultural (AG)	210	94%
Community 1 : Local (CO1)	8	88%
Community 2 : Regional (CO2)	3	100%
Council to Deem	1	N/A
General Business 2 (GB2)	1	100%
General Industrial 1 (GI1)	3	100%
General Residential 1 : Group Housing (GR1)	3	100%
General Residential 2 (GR2)	1	100%
General Residential 4 (GR4)	1	100%
Limited Use (LU)	20	100%
Local Business 2 : Local Business (LB2)	1	100%
Mixed Use 2 (MU2)	1	100%
Open Space 1 : Environmental Conservation (OS1)	17	100%
Open Space 2 : Public Open Space (OS2)	23	100%
Open Space 3: Special Open Space (OS3)	3	100%
Risk Industry (RI)	3	100%
Rural (RU)	5	100%
Single Residential 1 : Conventional Housing (SR1)	34	100%
Transport 1 : Transport Use (TR1)	5	100%
Transport 2 : Public Road and Public Parking (TR2)	3	67%
Utility (UT)	4	100%
Unknown	32	N/A

Two measures of density could have been used, either population density or household density, and both datasets were available from the census. In this instance, household density was selected because it is the more common metric used in trip generation estimation (Mathew and Rao, 2007). In terms of a description of the context, household density is also a more accurate descriptor of the built environment characteristics of the area. The population density data was used, instead, to normalise some of the other data sources.

### 6.1.3 Employment Density

Employment data was sourced from the Regional Service Council (RSC) Levy database for 2005. Now defunct, RSC levies were a tax based on an employees remuneration for Pay As You Earn (PAYE) purposes and RSC specifications. The data is provided per Sub Place area, of which there are 323 in total in the Cape metropolitan region. Employment density was calculated by taking the ratio of employment and the area of the Sub Place in hectares. Figure D.5 shows the employment density in Cape Town.

### 6.1.4 Vulnerable Road Users

The number of vulnerable road users was calculated from data sourced from the 2001 National Census data published in 2003 (Statistics South Africa, 2003). The data was made available at the Small Area

Level. Figure D.4 shows the spatial distribution of vulnerable road users in Cape Town. The value used in the assessment was calculated as the proportion of the population in a Small Area in the age categories below 14 years and above 65 years of age.

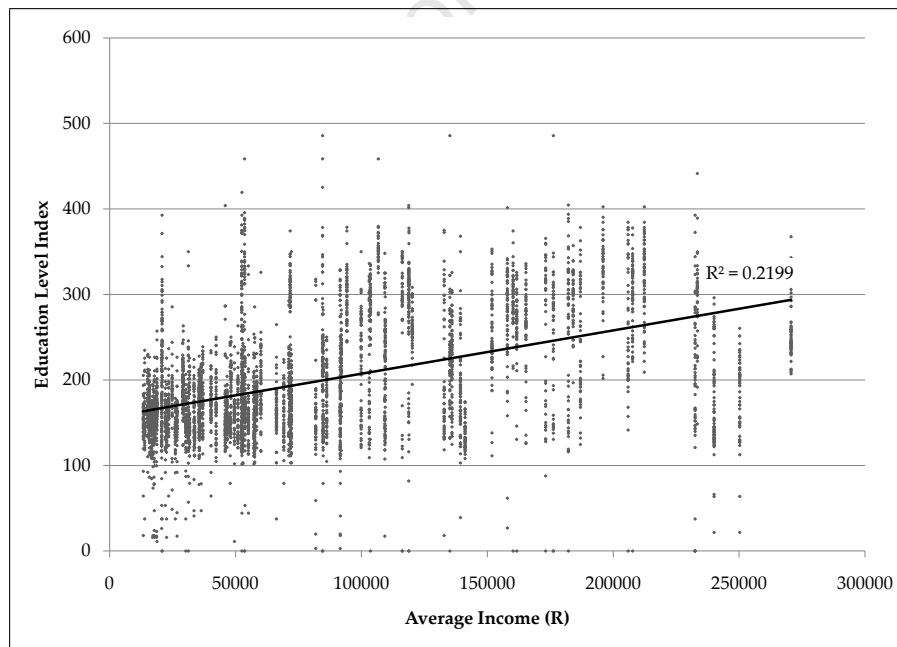
### 6.1.5 Income

The income data, used in the study, was drawn from the 2001 National Census data (Statistics South Africa, 2003). The data used was available at the electoral ward level, which, as discussed at the beginning of this section, is the most aggregate level of data used in the assessment. The income data is presented in the national census in terms of the number of people in each income category in each electoral ward. A single representative income value was calculated for each electoral ward using the Class Interval Arithmetic Mean,  $\bar{X}$  given by:

$$\bar{X} = \frac{\sum f \times X}{\sum f} \quad (6.1)$$

Where  $X$  is the midpoint of the interval, and  $f$  is the frequency within each interval.

Since the income data collected for this study is so highly aggregated compared to the other data sets, an alternative approach to represent income data was investigated, that of using education level as a proxy for income. Education level information is available at the Small Area Level from the census data. Education level can be a reasonable proxy data source for income, since people with higher education levels generally earn more than others. To test the viability of this approach, the relationship between income and education level was investigated. Figure 6.2 and Table 6.3 shows the results of the correlation analysis for the two data sets.



**Figure 6.2:** Income versus Education Level

The correlation between income and education level is statistically significant, and income has a strong positive correlation with education level ( $r = 0.491$ ,  $p < 0.001$ ). The value of  $R^2$ , the coefficient of determination, is 0.2199, meaning that despite the good correlation between the two data sets, education

**Table 6.3:** Pearson Correlation Coefficient for Income and Education Level

		<b>Average Income</b>	<b>Education level index</b>
<b>Average Income</b>	Pearson Correlation	1.000	0.491
	Sig. (2-tailed)		0.000
	N	5465.000	5437.000
<b>Education level index</b>	Pearson Correlation	0.491	1.000
	Sig. (2-tailed)	0.000	
	N	5437.000	5437.000

level can only account for 22% of the variation in incomes. The difference in aggregation level of the income data and the education level data possibly plays a large role in the uncertainty of the results. In addition, the incomes used in correlation test were calculated from a mean of the interval data presented in the census, and the score for education level is a weighted index. These values thus mask much of the detail within their areas, possibly weakening the correlation and increasing the variation between the data sets.

Considering all of these factors, it was decided to substitute the education data set for the income data set. The data used in the assessment, therefore, was the index of education level at Small Area Level. The data is presented in categories, or intervals, of age group and education level. An incremental weighting scheme was applied to the values in each category, and the summation thereof used as the index. This value was then normalised by the total population of the Small Area, and the normalised value was used in the SMCA. Normalising by population is critical, since Small Areas with high levels of education would otherwise rank the same as large areas with low levels of education. Table 6.4 shows the weighting scheme for the education level index.

### 6.1.6 Heritage Sites

The location of heritage sites, or sites designated as having some heritage or cultural value, was sourced from the City of Cape Towns Environmental Resource Management Department, who have been in the process of documenting and classifying this information for conservation purposes. The exact extent of each site or area has been geocoded, and some progress has been made to classify these sites in terms of their heritage value, but no final classification scheme had been accepted by the Department at the time of receiving the data from them. It was, therefore, decided to simply note the extent of the sites and accept them all as having equal heritage value. The proximity maps are, therefore, not weighted according to the relative heritage value of each site, although the method could be adapted to account for this. Figure D.6 shows the location of heritage features in Cape Town.

### 6.1.7 Ecologically Sensitive Areas

The location of environmentally sensitive areas, wetlands and waterbodies was sourced from the City of Cape Towns Environmental Resource Management Department, who have been in the process of documenting and classifying this information for conservation purposes. As with the heritage features database, the exact extent of each site or area has been geocoded, and the metadata accompanying the data mentions the studies used to compile the data, and the classification schemes used in them, but no

**Table 6.4:** Education level index weighting scheme

Age Group	Education Level	Weight
0 - 4	Not applicable	0
5 - 9	No schooling	0
	Some primary	1
10 - 14	No schooling	0
	Some primary	1
	Complete primary	2
	Some secondary	3
15 - 19	No schooling	0
	Some primary	1
	Complete primary	2
	Some secondary	3
	Grade 12	4
	Higher	5
20 +	No schooling	0
	Some primary	1
	Complete primary	2
	Some secondary	3
	Grade 12	4
	Higher	5

final classification scheme had been accepted by the Department at the time of receiving the data from them. As with the heritage data, it was decided to accept all the sites as being of equal significance.

The data sets for wetlands and environmentally sensitive areas are in separate files. These files were merged in a GIS and the merged data sets used in the assessment. The proximity maps, therefore, represent the proximity to both wetlands and environmentally sensitive areas. Figure D.7 shows the location of environmentally sensitive areas and wetlands in Cape Town.

### 6.1.8 Travel Demand

Data from the City of Cape Towns Origin Destination (OD) matrix for commuter trips in 2007 was used in the assessment to account for travel demand. The information was compiled by the City of Cape Towns Transport, Roads and Major Projects Department to assist with their transport planning operations. The data is provided at the Travel Analysis Zone (TAZ) level of aggregation, there being 842 TAZ's in the Cape metropolitan region. The data is for the morning commute period only, only includes work trips, and only includes the main mode to work.

The data is presented by mode of transport to and from each TAZ, including road based PT, rail based PT and car. The information used for the analysis is the demand for PT and for PMT per analysis zone. The origin and destination data was, therefore, combined to give an estimate of the total demand for travel in each area, by mode. Data for PT modes, including rail transport, was combined to give an overview of the demand for PT travel per TAZ. Rail was included in the analysis, because rail often

forms an important part of people's commuting trip chains, and access and egress modes to and from stations often include other PT modes and walking. The information was not normalised by area and, instead, the absolute number of origin and destination trips was used. Separate criterion maps were generated for PT and for PMT modes. Figure D.8 shows the travel demand data as the ration between the number of PT trips and car trips per TAZ for Cape Town.

### 6.1.9 Public Transport Stops

Spatial data on the location of PT stops around Cape Town was sourced from the City of Cape Towns Transport, Roads & Major Projects Department. The data is published in a point file format, and contains information for all the formal stops for road based and rail based PT. While it is appreciated that certain stops cater to larger numbers of travellers than others, no distinction was made in the assessment between stops based on patronage.

## 6.2 Case Study Selection

The Cape metropolitan region is a starkly polarised urban area. Wealthy suburbs and prosperous economic centres that offer rich opportunities of all kinds contrast with overcrowded, impoverished dormitory settlements on the periphery (Turok, 2001). Historically, Cape Town developed around the port established by the Dutch East India Company, as a resupply station for ships sailing between Europe and the East (Wilkinson, 2000).

The construction of a railway line, heading south from Table Bay towards Simons Town on False Bay, promoted the development of a series of suburban commuter villages around the small farming stations along its route. This was followed by the construction of a railway line to the interior, which passes through Bellville, leading to the development of a similar chain of commuter villages stretching eastwards from the central settled area. Additional lines were laid on intermittently to serve the growing African and Coloured 'townships' of the Cape Flats with major centres of employment in the city's historic core (Behrens and Wilkinson, 2003).

The large-scale grid of freeways and major arterials, which presently dominates the metropolitan road network, was largely planned during the 1960s (Morris, 1969), and constructed in phases through the 1970s and 1980s, although some of the arterial routes originated during the nineteenth century or earlier. The improved road infrastructure both promoted and developed in response to the growth of private car ownership in the city.

The settlement pattern of Cape Town is inextricably linked to the legacy of Apartheid spatial planning ideologies. The effects of the policies of racial segregation from that era are still largely unchanged, despite policy changes deracialising settlement planning and urban development. Racial segregation has been replaced by class divisions, but class divisions have remained mostly along racial lines. Historically African and Coloured suburbs are still populated by low-income African and Coloured communities, with poor access to employment opportunities, services and city amenities (Schensul and Heller, 2011). So, although the legislative and political structures of Apartheid were demolished nearly two decades ago, the geography it left behind has remained firmly in place - now bolstered by the economic realities of the housing market (Lucas, 2010; Naude, 2008; Rospabe and Selod, 2006). The radial arterials that head out from the city centre, therefore, generally pass through a range of communities, with vastly different circumstances.

The hypothesis proposed by this research, is that every location has a contextual setting that provides important information, useful for planning roads and, as this context varies along the route, the planning

of the road should vary to suit. In that sense, in developing the method used to measure the context of roads, case study roads were selected on the basis of the extent of variation in the factors used to describe the context. By necessity then, the case studies used to develop the method were all arterial roads, since these tend to be especially long, passing through a large range of contextual settings. They, therefore, provide the best opportunity to use a large, robust set of data to test various analysis strategies and calibrate the method.

As was shown in Table 3.4, all of the ten worst road sections regarding accidents involving pedestrians are either arterials or freeways. Further evidence is presented in Table 6.5, which shows the ten roads with the most accidents of all types in Cape Town for 2005 (City of Cape Town, 2005).

**Table 6.5:** Ten worst-known road sections for accidents in Cape Town

Location	Type	Fatal	Serious	Slight	Damage	EAN <sup>†</sup>
N1 - Cape Town	Freeway	31	45	399	1 643	3 293
N2 - Cape Town	Freeway	18	74	371	1 745	3 190
Lansdowne Rd - Lansdowne	Arterial	12	93	331	1 037	2 379
Voortrekker Rd - N. Suburbs	Arterial	5	37	297	1 516	2 304
N7 - Cape Town	Freeway	15	41	255	1 083	2 086
Koeberg Rd - Milnerton	Arterial	3	21	144	836	1 250
R300 - Delft	Freeway	18	21	106	317	1 057
Van Riebeeck Rd - Kuils River	Arterial	4	26	99	443	844
Modderdam - Ravensmead	Arterial	3	25	92	452	811
Klipfontein Rd - Cape Town	Arterial	3	22	87	423	758

<sup>†</sup> The following factors were used to calculate the EAN

Fatal accident: 25.4

Serious injury accident: 5.8

Slight injury accident: 1.5

Damage only accident: 1.0

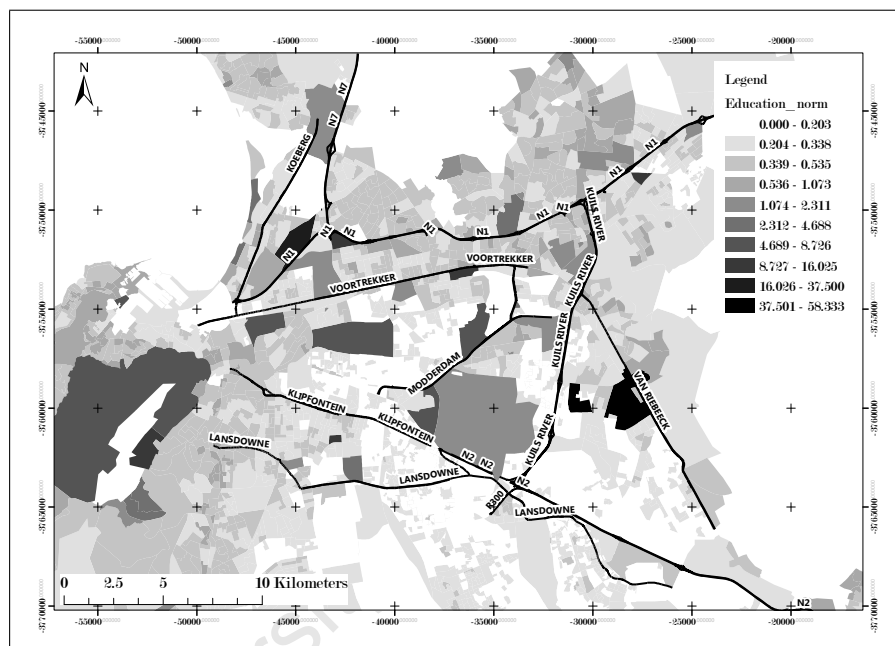
*Source:* Adapted from City of Cape Town (2005)

Many of the roads in Table 3.4 are also found in this list. Considering the arterials in these lists, Lansdowne Road and Voortrekker Road stand out as being particularly problematic, since they are the two most dangerous arterial roads on both lists. Ten out of twelve and three out of five fatalities, respectively, were listed as being pedestrians for Lansdowne and Voortrekker Roads. Such a high number of pedestrian accidents suggests a serious disconnect between the needs of the road users, and the infrastructure provided along these routes.

Figure 6.3 shows the arterial road network in Cape Town overlaid on a map of an index of the average education level in Cape Town at the Small Area Level<sup>25</sup>, calculated from data in the 2001 national census. The index was calculated by multiplying the number of people per Small Area in each category of education level, by a simple linear weighting that increases with higher categories of education level.

<sup>25</sup> The Small Area Level is a geographic area comprised of a combination of census enumeration areas, that when counted together have a population of 500 or more. This geographic area was devised by Statistics South Africa, who packaged and distributed the 2001 National Census data, to protect the privacy of those people living in enumeration areas with populations of less than 500 individuals.

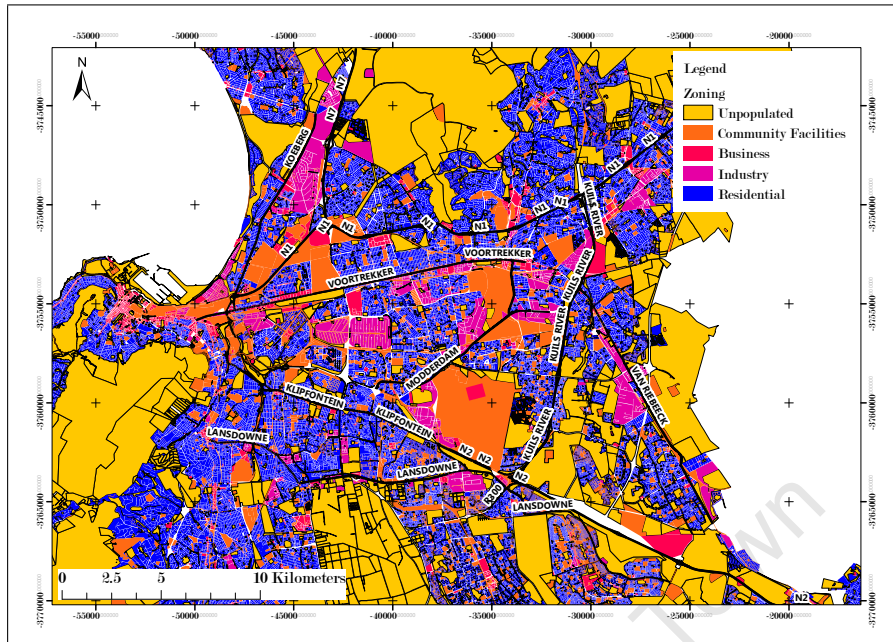
The sum of the weighted education data for each Small Area was then normalised by dividing by the population of the Small Area. This index gives a reasonable overview of the spatial distribution of education levels around Cape Town. Education level is a good proxy indicator of a range of transport related factors, such as income, car ownership, commuting mode, and data on education levels is readily available from the 2001 national census database at Small Area Level. Income data, which would be a better indicator overall, is not available at this level of aggregation. The best available income data is at the electoral ward level, which is a much more aggregated level of information and, therefore, not useful. A more in depth discussion around data availability, selection, optimization and formatting is presented in Section 6.1 of this chapter.



**Figure 6.3:** Ten worst-known roads for accidents in Cape Town and education level

In Figure 6.3, darker shaded areas indicate higher average levels of education. Lansdowne and Voortrekker Roads pass through areas with a wide variety of different education levels (and, by extension, income levels). Figure 6.4 shows the arterial road listed in Table 6.5 overlaid on a zoning map of Cape Town. As mentioned in Chapter 5, despite its limitations, zoning was used as a substitute for land uses. A small field verification study was undertaken to check the relationship between zoning and land use, and zonings were grouped into categories to minimise whatever variation there was between zoning and land uses. A description of this study is provided in Section 6.1 of this chapter.

Figure 6.4 shows that both Voortrekker and Lansdowne Roads pass through a wide variety of land use zones along their routes. Considering the range of incomes (estimated from the spatial distribution of education levels) and the wide variety of land uses, it is reasonable to expect the context along these routes to vary significantly. In addition, these routes are both important PT routes, and being arterial roads, tend to carry significant levels of freight vehicles as well. The combination of these factors, coupled with the high numbers of accidents, led to these two routes being selected for the development, testing and calibration of the methodology. The results produced by the method was then tested by analysing a third case study road, Koeberg Road, which shares many of the attributes of Voortrekker



**Figure 6.4:** Ten worst-known roads for accidents in Cape Town and land use zones

and Lansdowne Roads, and is also noted as being a particularly dangerous route for pedestrians and motorists alike.

## 6.3 Case Study Route Descriptions

A brief description of each route is presented in the following section.

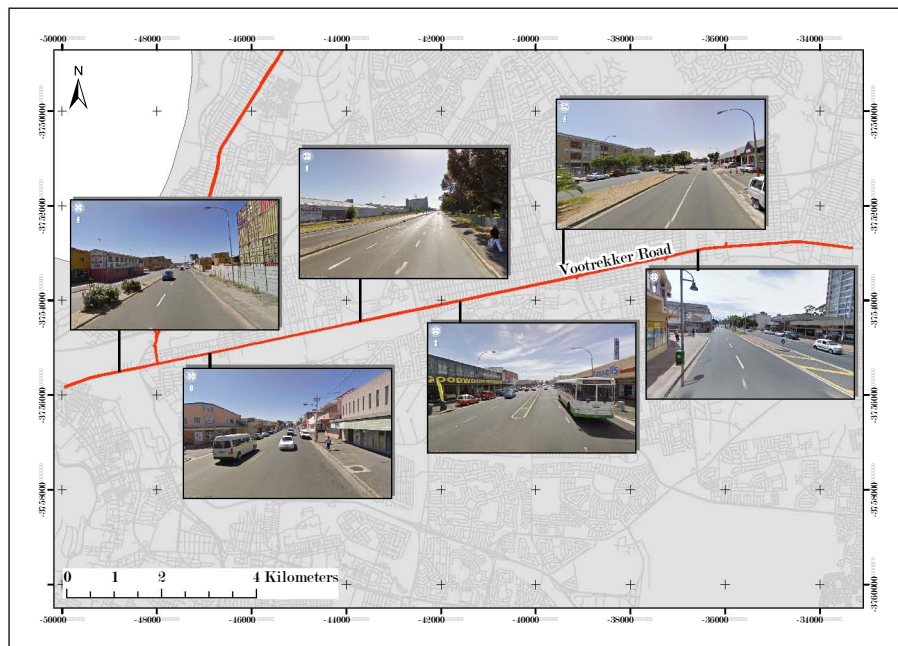
### 6.3.1 Voortrekker Road

Voortrekker Road (R102) is a major arterial linking the northern and eastern suburbs of Cape Town with the CBD. Voortrekker Road itself is approximately 17 km in length, running in an easterly direction between the districts of Salt River and Bellville. For the majority of its course (>60% of the length), the road comprises a four lane, dual carriageway roadway, with a narrow central median and is flanked by intermittent provision for parking space (sometimes in embayments) and sidewalks of varying width. The remainder of the route, generally, comprises a four lane undivided section with parking bays and sidewalks. The choice of cross-section is, mostly, related to the available reserve width, with wider reserves allowing for the preferred dual carriageway cross-section<sup>26</sup>. This uniformity of cross-section is striking given the range of contexts the road passes through. Figure 6.5 shows typical views at various points along the route of the road.

### 6.3.2 Lansdowne Road

Lansdowne Road (M24/M9) is one the main routes linking the southern suburbs of Cape Town, including many Coloured and African communities and the business districts of Claremont, Wynberg and Rondebosch. Starting in Claremont, it runs south east for over 27 km, traversing a very wide range

<sup>26</sup> Undivided four lane roads are avoided wherever possible since they are believed to be unsafe, being generally noted for having a higher incidence of head-on collisions.



**Figure 6.5:** Typical views at points along Voortrekker Road  
*Source:* Images sourced from Google Maps, accessed 15/04/2011



**Figure 6.6:** Typical views at points along Lansdowne Road  
*Source:* Images sourced from Google Maps, accessed 15/04/2011

of land uses and community types along its course, before terminating at the suburb of Macassar near False Bay.

For the first five kilometres of its course, Lansdowne Road passes through a mix of middle to upper middle income residential neighbourhoods interspersed with strip retail and local community facilities. This section of the route, generally, comprises a two lane undivided section, often with parallel parking and sidewalks on either side of the road. The road reserve in this section of the route is narrow, limiting the possibilities for additional vehicular capacity, although turning lanes are often provided at larger signalised intersections.

The route then passes along the northern boundary of the Phillipi agricultural zone and the southern boundary of the Crossroads industrial district. This section of the route is comprised of a four lane dual carriageway cross-section with a wide median, bicycle lanes (sometimes marked on the roadway, and sometimes sharing space with the sidewalk), and sidewalks on either side of the road. The road reserve in this area is very wide but, in places, has been taken over by informal housing built up to the outside edge of the sidewalks. This cross-section is maintained for approximately 7.5 km, before the route narrows to a two lane undivided section again, only splitting in the vicinity of major intersections.

Along the last half of the route it passes through a range of middle to low income residential neighbourhoods, interspersed with very high density informal settlements built right up to the road edge. This section of the route is characterised by extremely high volumes of pedestrian traffic and high levels of informal trading on the roadside. Figure 6.6 shows typical views at various points along the route of the road.

### 6.3.3 Koeberg Road

Koeberg Road is an important arterial route serving a range of residential, commercial and industrial activities along its route. It is, approximately, 11 km long and runs northwards from Maitland, passing through the suburbs of Brooklyn, Rugby, Tijgerhof, Milnerton and Montague Gardens, before terminating at the intersection with Blaauwberg Drive near Table View.

Koeberg Road has a variety of cross-sections along its route, ranging from three lanes undivided with sidewalks on either side at the start of the route, to four lane dual carriageway with two metre shoulders, no sidewalks and a service road in the vicinity of Montague Gardens. The function of the route also changes quite dramatically along the route. The route starts as being a mixed use arterial, even having residential accesses along it, eventually becoming a mobility only route with no direct access onto it, adjacent properties being served by a series of parallel serviced roads. Koeberg Road also accommodates a particularly high volume of freight vehicles, since it serves as an important access route for a number of large industrial areas located along the northern section of the route.

This mixture of land uses, income groups, and functionality makes this an interesting and challenging route to assess. Figure 6.7 shows typical views at various points along the route of the road.

## 6.4 Résumé

In this chapter, a description of the data collection process, the data sources, data availability and the data limitations in terms of the requirements of the methods used in the study was described. This was followed by a description of the case study roads selected for the development and testing of the method.

Primary data sources included various departments of the City of Cape Town Municipality and the 2001 National Census. Data used in the assessment were at different levels of aggregation. This has



**Figure 6.7:** Typical views at points along Koeberg Road  
 Source: Images sourced from Google Maps, accessed 15/04/2011

negative implications for the precision of the results. The data used in the assessment was also collected and compiled at different times in the past, which affects the reliability of the results, in that certain data sets may have outdated information in them.

A discussion of each data set and the methods used to optimise the data for analysis, followed. Zoning data sourced from the municipality was used instead of land use data. A verification study was conducted to investigate the variation between land use and zoning. Based upon these results, it was accepted that the listed zoning is a reliable representation of the land use.

Household density sourced from the national census was used in the assessment. Household density was calculated as the ratio of the number of households in a Small Area and the area of the polygon in hectares. The number of vulnerable road users was calculated from data sourced from the 2001 National Census. The number of vulnerable road users used in the assessment was calculated as the proportion of the population in a Small Area in the age categories below 14 years and above 65 years of age.

Education level data was used as a proxy for income data used in the assessment which was taken from the national census. The data is presented in categories, or intervals, of age group and education level. An incremental weighting scheme was applied to the values in each category, and the summation thereof used as the index. This value was then normalised by the total population of the Small Area, and the normalised value was used in the SMCA. The relationship between income and education level was investigated in a correlation analysis to test whether education level was a reasonable proxy data source for income. Income and education level was found to be positively correlated and the relationship was confirmed as being statistically significant ( $r = 0.491$ ,  $p < 0.001$ ).

Employment data was sourced from the RSC Levy database for 2005. Employment density was calculated by taking the ratio of employment and the area of the Sub Place in hectares.

The location of heritage sites, and the location of environmentally sensitive areas, wetlands and waterbodies was sourced from the City of Cape Towns Environmental Resource Management Department. Although some progress has been made to classify these sites in terms of their relative heritage or environmental value, no final classification scheme had been accepted by the Department at the time of receiving the data from them. It was, therefore, decided to simply note the extent of the sites and accept them all as having equal value.

Data from the City of Cape Towns OD matrix for commuter trips in 2007 was used in the assessment to account for travel demand. The data is for the morning commute period only, only includes work trips, and only includes the main mode to work. The OD data was combined, to give an estimate of the total demand for travel in each area, by mode. The information was not normalised by area and, instead, the absolute number of origin and destination trips was used. Spatial data on the location of PT stops around Cape Town was obtained from the same source.

Three case study roads were selected, Voortrekker Road between Salt River and Bellville, Lansdowne Road between Claremont and Macassar and Koeberg Road between Maitland and Table View. All the case study roads are arterial routes in Cape Town. The roads were selected based upon a combination of the road safety problems and the variety of land uses and community types along their courses.

All the roads selected pass through a wide variety of land use zones along their length. Considering the range of incomes (estimated from the spatial distribution of education levels) and the wide variety of land uses, it is reasonable to expect the context along these routes to vary significantly. In addition, these routes are both important PT routes, and being arterial roads, tend to carry significant levels of freight vehicles as well. The combination of these factors, coupled with the high numbers of accidents led to these two routes being selected for the development, testing and calibration of the methodology. The results produced by the method was then tested by analysing a third case study road, Koeberg Road, which shares many of the attributes of Voortrekker and Lansdowne Roads, and is also noted as being a particularly dangerous route for pedestrians and motorists alike.

## Chapter 7

# Case Study Results and Assessment

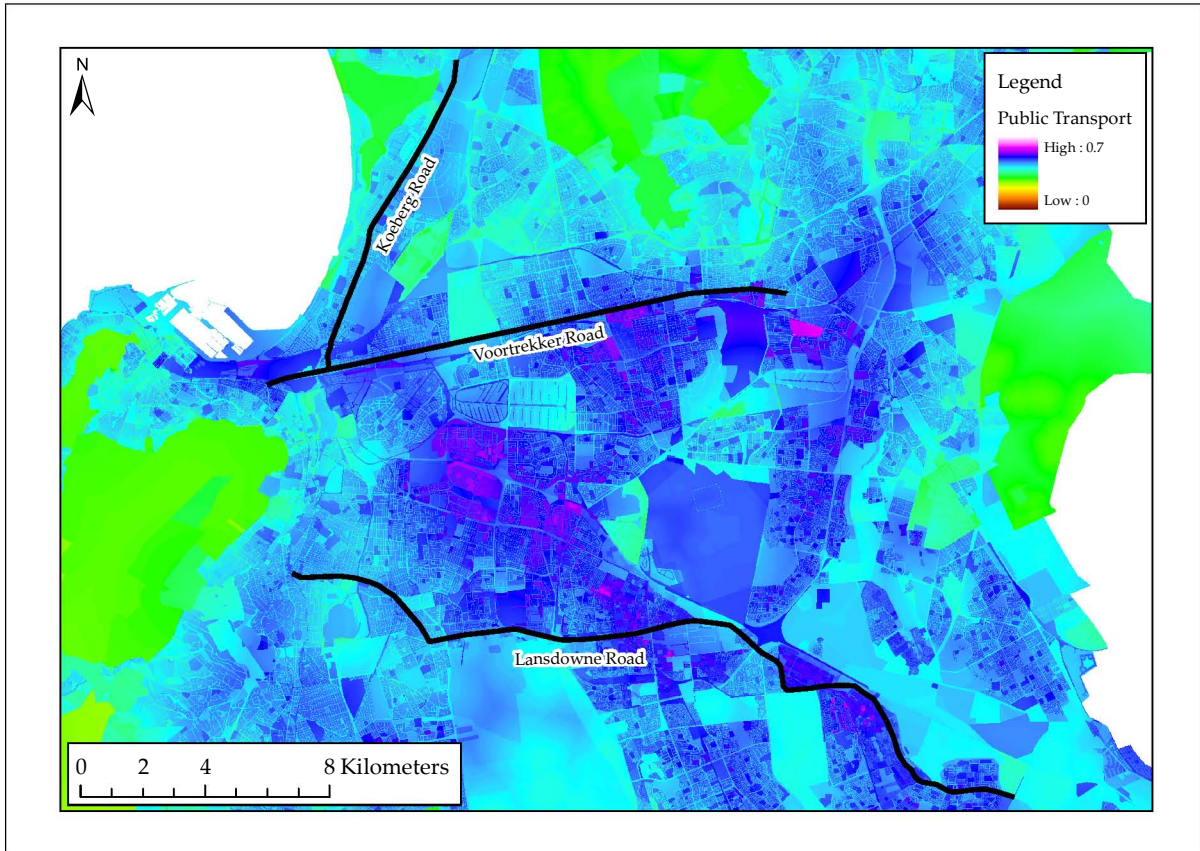
The case study roads and data used in the analysis were introduced and discussed in Chapter 6. In this chapter, the results of the SMCA conducted for the various modes are presented and analysed. As discussed in Chapter 5, citywide data sets were used to conduct the SMCA, thereby producing modal level of suitability maps for the whole city. The route analysis, however, is limited to the case study roads discussed in Chapter 6, for which suitability scores were extracted for comparison. The findings from this chapter contribute towards a better understanding of how context could be used to make planning recommendations.

### 7.1 Level of Suitability Rasters

The SMCA was conducted in Esri Inc.'s ArcMap GIS software using the methodologies discussed in Chapters 5 and 6. Five level of suitability maps were generated, one for each of the five modes of transport analysed along the three case study roads. Figure 7.1 shows the level of suitability map for PT generated from the SMCA before applying neighbourhood averaging. Figures E.1 to E.4 show the level of suitability maps for the other modes.

The visible light spectrum was selected for the map symbology because it highlights small variations in a narrow score range, whilst still allowing for a large total range of scores to be displayed. This is useful, since although most of the scores occur within a narrow range, there are important outlier areas along the route. The maximum scores achieved by any mode are all below 0.7. This was, therefore, selected as the maximum stretch value (purple to white). The minimum scores are nearer to zero (brown to black). Most regions score between 0.3 and 0.6 (green to blue). These colours are, therefore, the most prevalent across the city.

Figure 7.2 shows the frequency distribution of the suitability scores for each mode on a logarithmic scale. A logarithmic scale can show more detail across a wider range of scores than a linear scale. The figure shows that the suitability scores for pedestrian, bicycle and PT modes range between approximately 0.2 and 0.67, whereas the scores for freight and car range between 0 and 0.67. A large proportion of the suitability scores for bicycle and PT fall between 0.24 and 0.32, and between 0.36 and 0.47 respectively. The suitability scores for the pedestrian mode predominantly occur in one of three ranges, 0.24 to 0.30, 0.34 to 0.42 and 0.44 to 0.52. Freight and car modes have a small number of scores in the very low range. Figures E.2 and E.3 reveal that a large proportion of these scores fall in the waters near the harbour area. These lowest scores are irrelevant artifacts of the input data and can, therefore, safely be ignored. Notably, these modes also have areas where they are very highly suited.



**Figure 7.1:** Public Transport Suitability Scores from SMCA for Case Study Routes

Closer inspection of Figures E.2 and E.3 shows that at least some of these scores correspond to the Blackheath industrial area in the east of the city.

Figure 7.3 provides further insight into the suitability scores from the SMCA. The figure shows the cumulative frequency of suitability scores by mode and confirms that the majority of scores fall between 0.2 and 0.54. The figure also shows that, with the criteria used in this analysis, bicycle, pedestrian and PT modes have remarkably similar cumulative suitability scores for the study region, which is notably higher than that of freight and car. There are, however, still noticeable differences between the scores for each mode. The freight mode has a much more even distribution of scores than the car mode, although, in general, the car mode scores better than the freight mode. The various peak ranges of the other three modes are also clearly visible, bicycle and freight have two peak ranges and pedestrian has three.

These ranges indicate that there are typologies within the scores for each mode. In other words, some of the modes tend to score most frequently within specific suitability ranges. The suitability scores for the pedestrian mode occurs in three distinct ranges, suggesting three levels of suitability for the mode across the city. The suitability raster for pedestrians was reclassified into five ranges (as shown in Figure E.5) to investigate this further. The lowest scores were reclassified as 0, the first peak range of scores were reclassified as 1, the second peak range as 2, the third as 3 and the highest scores as 4. Figure 7.4 shows the spatial distribution of the reclassified suitability scores.

In Figure 7.4, the lowest scores, 0, are found on Robben Island, off the coast of Cape Town and,

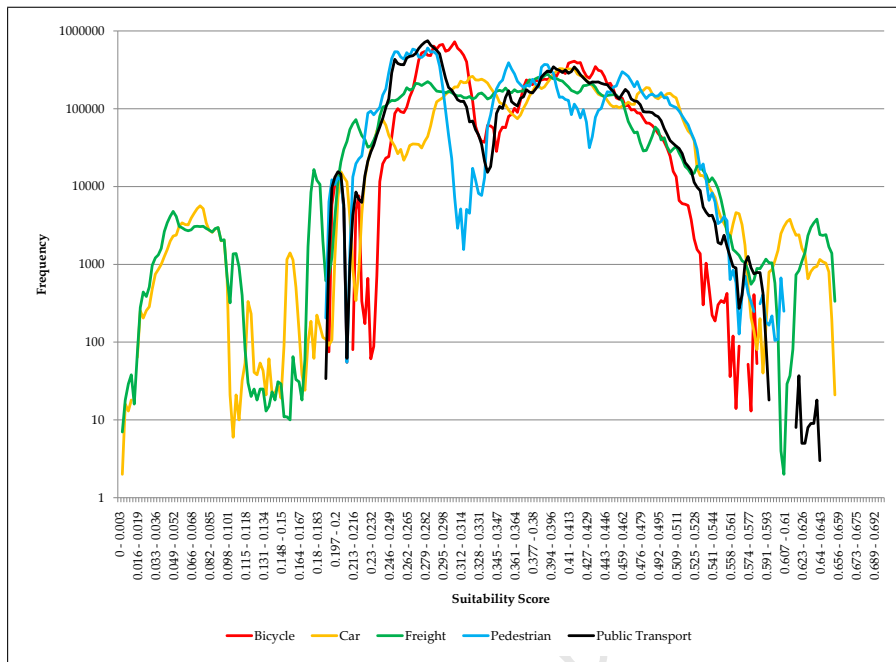


Figure 7.2: Frequency Distribution of Suitability Scores by Mode

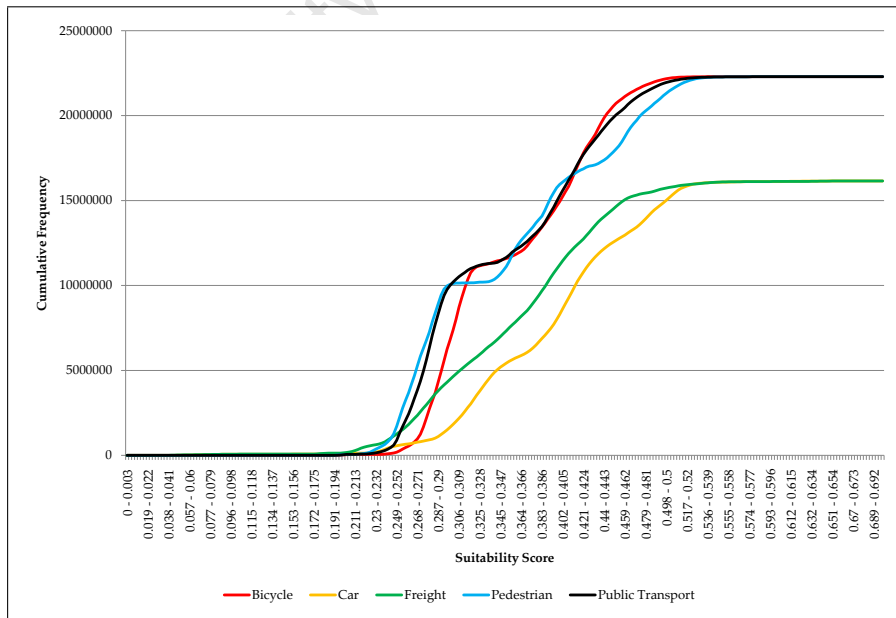
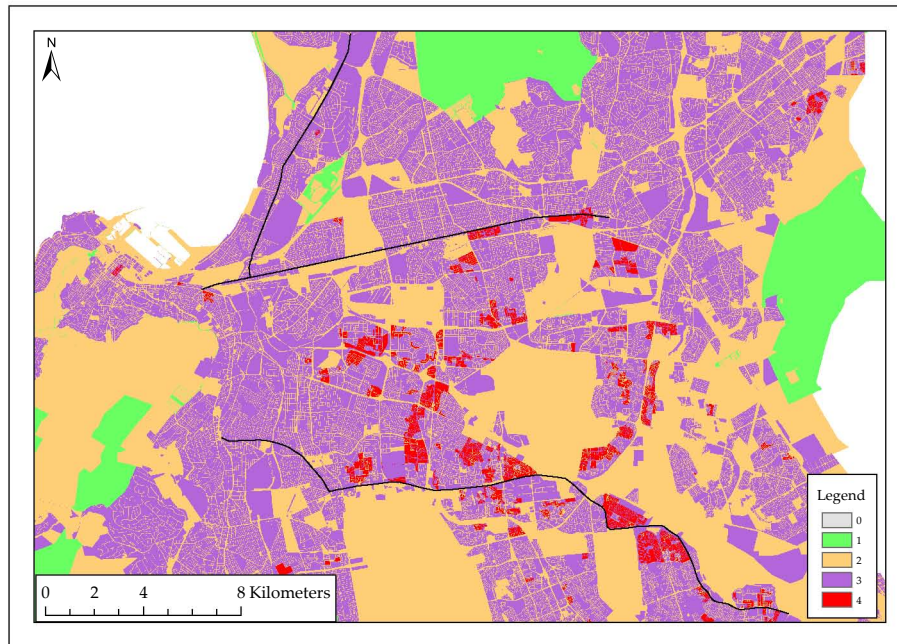


Figure 7.3: Cumulative Frequencies of Suitability Scores by Mode



**Figure 7.4:** Spatial Distribution of Reclassified SMCA Scores for Pedestrians

therefore, play no further part in the analysis. Scores in the ranges 1 and 2 occur only on sparsely populated areas, or on roads. The scores in ranges 3 or 4 are found in populated areas. These are, therefore, the scores that are of particular interest in this application. Plotting only the scores in range 3 (see Figure 7.5) shows that within this range, the scores are close to normally distributed. The score profiles for the other modes display similar attributes.

The next step in the method involves taking the average score of a circular area 150 m in radius around each cell. This will inevitably mean that there will be some blending of the results from range 2 into ranges 3 and 4, causing the slight lowering of results observed in the averaged scores. This lowering is not necessarily a reflection of the context, since at least some of the lowering effects are from the scores along the road itself (which does not, technically, play any role in the context<sup>27</sup>). This is not a concern, however, since the relative scores between modes will not be altered, as all the averaged scores will be affected in exactly the same way.

## 7.2 Extracted Raster Scores

The SMCA rasters were processed using neighbourhood averaging, and the averaged suitability scores along the case study roads were extracted to a spreadsheet (as discussed in Section 5.5) to conduct a more detailed analysis of the results. In this section the extracted raster scores are discussed by case study road.

### 7.2.1 Voortrekker Road

Figure 7.6 shows the suitability scores for Voortrekker Road.

The results for Voortrekker Road shows significant variation along the route. The route starts at

<sup>27</sup> Although it could be argued that this effect helps to account for the width of the road reserve, since areas with wider reserves will be affected to a greater extent.

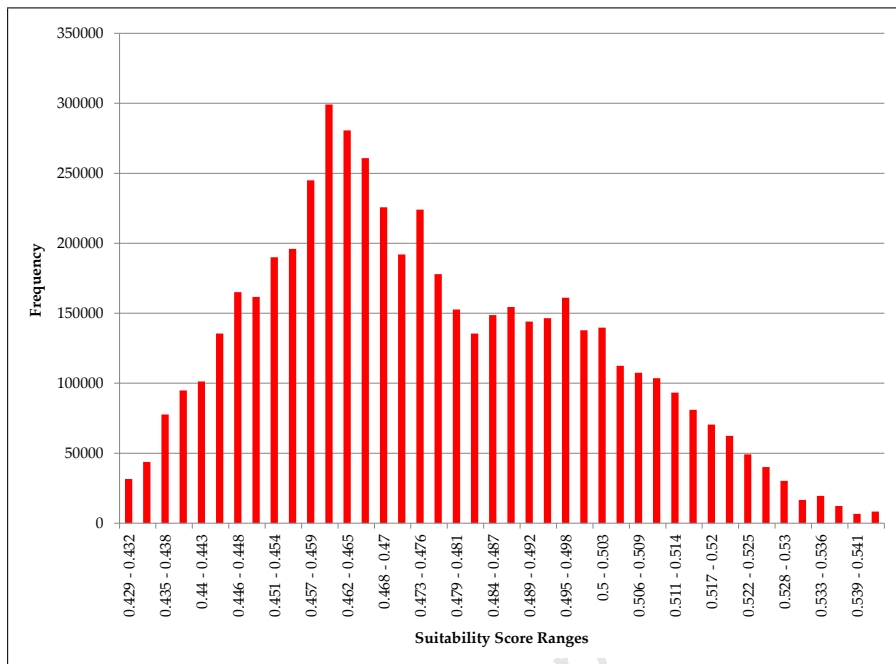


Figure 7.5: Frequency Distribution of Suitability Scores in Range 3 for Pedestrian Mode

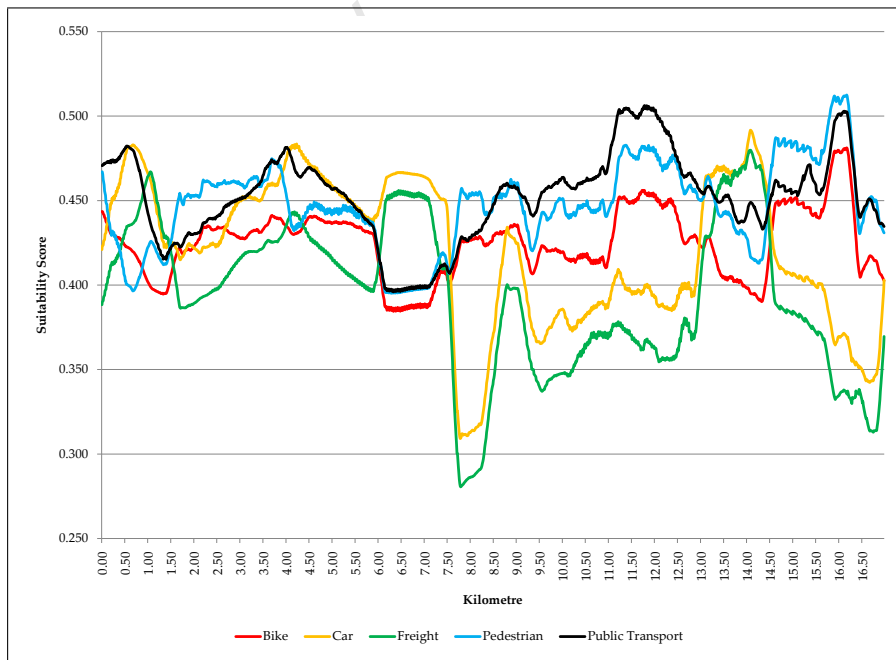
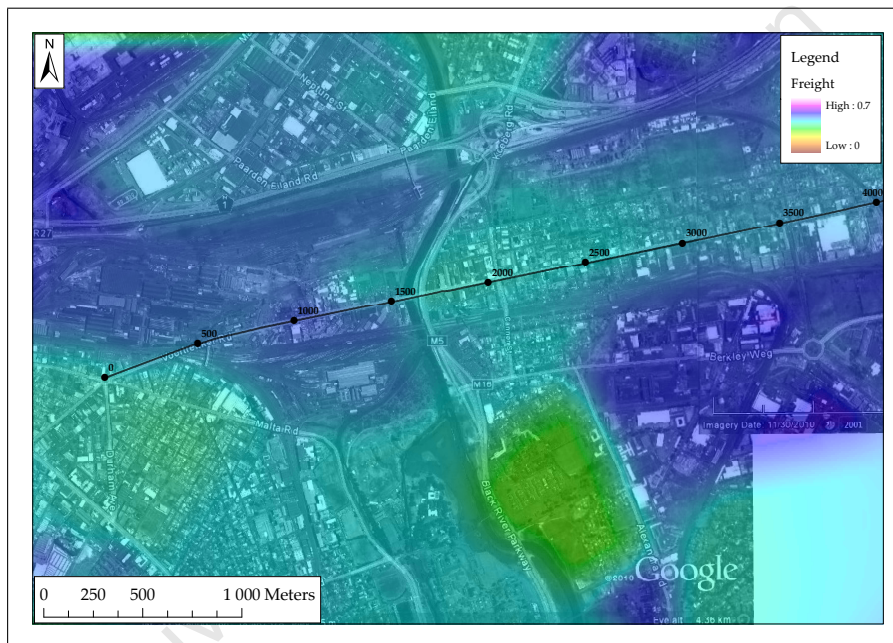


Figure 7.6: Suitability Scores along Voortrekker Road

Salt River circle, which is located in a commercial area, before moving through an industrial area, between kilometre 0.5 and 1.5. The route then passes through the commercial district of Maitland which, by kilometre 4, has become an industrial and warehousing district. Freight suitability will be used to demonstrate the analysis of the context from the extracted suitability scores. Near Salt River circle, freight is ranked poorly, whereas PT and pedestrian are ranked highly. The situation reverses in the industrial area between kilometre 0.5 and 1.5, thereafter, in Maitland's commercial district, the suitability score for the pedestrian mode increases dramatically again.

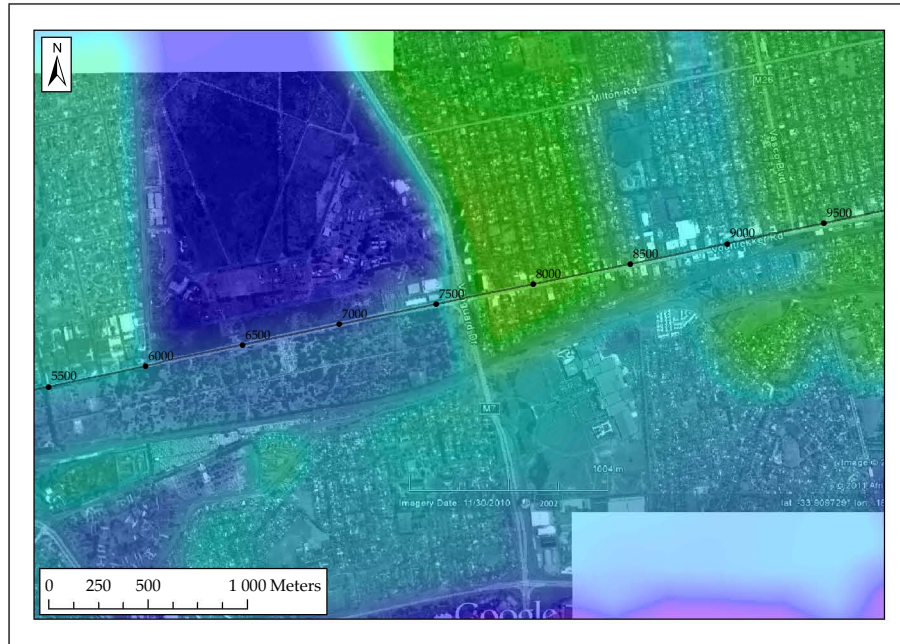
Figure 7.7 shows a satellite image of the area, overlaid by the averaged suitability raster for freight. The image provides some clarification of the scores along the route, giving some context to Figure 7.6. In the figure, the green areas are less suited to freight vehicles than the blue areas, with purple areas being most suitable for freight. The section of route between 500 m and 1 500 m, through an industrial area, is shown as being more suited to freight than the areas surrounding it.



**Figure 7.7:** Freight Suitability Raster on Voortrekker Road, Km. 0 - 4

Another distinctive area is between kilometres 6.0 and 9.0, where there is a dramatic reversal in priorities between freight and car and the other modes. Figure 7.8 shows that between kilometre 6.0 and 7.5, there is a large open area to the north of the road. Near kilometre 7.5 there is another industrial area. A large cemetery lies to the south along much of this section of the route. The route is then intersected by Vanguard Drive, a major arterial route linking the southern and northern areas of the city, and then enters a commercial area, passing by some residential tower blocks on the north eastern corner of the Vanguard Drive intersection. The combination of these factors lowers the suitability for freight vehicles in this area.

Between kilometres 9.5 and 13.0, the route maintains a relatively stable mode suitability rank order, PT being most suited, followed by pedestrian, bicycle, car and lastly, freight. This rank order reverses from kilometre 13.0 to 14.5, where the route passes by the Parow East industrial area. Thereafter, until the end of the route, the rank order changes to pedestrian, PT, bicycle, car and freight, in order of suitability. Table 7.1 contains various descriptive statistics for the suitability scores for the route.



**Figure 7.8:** Freight Suitability Raster on Voortrekker Road, Km. 5.5 - 9.5

**Table 7.1:** Descriptive Statistics for Voortrekker Road Suitability Scores

	Bike	Car	Freight	Pedestrian	Public Transport
N	1699	1699	1699	1699	1699
Mean	0.425	0.421	0.394	0.447	0.453
Std. Deviation	0.020	0.043	0.045	0.025	0.026
Skewness	0.032	-0.514	-0.291	-0.094	-0.180
Kurtosis	0.061	-0.499	-0.437	-0.070	-0.028
Minimum	0.384	0.309	0.281	0.395	0.396
Maximum	0.481	0.492	0.480	0.513	0.506
Range	0.097	0.183	0.199	0.117	0.110

The statistics provide some interesting insights into the performance of the various modes of transport along the route. There are 1 699 sample points along the route. On average, PT is the most well suited mode along the route, followed by pedestrian. Freight is, on average, the lowest scoring mode. The car and freight modes have the largest amount of variation in their data sets; on average double that of the other modes. This is an indication of the highly location specific suitability for these modes; they tend to only receive preference in areas where they are very highly suited, and everywhere else they score comparatively poorly.

The skewness value is a measure of the asymmetry of the frequency distribution of the mode suitability scores. All the modes, besides bicycle, have negative skewness values, indicating that these distributions are skewed to the right of their means, and bicycle is slightly skewed to the left. Car and freight have heavily skewed distributions, whereas bicycle and pedestrian are close to evenly distributed about their mean scores. Therefore, the most frequently occurring scores for car and freight tend to be higher than

the mean. This indicates that there are probably a small number of values having a negative influence on the mean, and that the mean itself may not be a reliable representation of the performance of these modes along the route. Referring back to Figure 7.6, the localised low scores around kilometre 8.0 and kilometre 16.5 could help explain the low mean scores for these modes.

The kurtosis value is a measure of the extent to which observations cluster around a central point. The bicycle mode has a slightly positive kurtosis, indicating that it is slightly more clustered than the normal distribution. The other modes have negative kurtosis, car and freight having the most dispersed frequency curves of all the modes. The range of scores for each mode is also revealing, with car and freight having a much larger range of scores than the other modes. The kurtosis and range values for these modes all point to the high levels of variation in the suitability of these modes along the route.

**Table 7.2:** Mode Suitability Correlation Matrix for Voortrekker Road

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
<b>Bike</b>	1.000	-0.305	-0.444	0.890	0.778
<b>Car</b>	-0.305	1.000	0.967	-0.442	-0.169
<b>Freight</b>	-0.444	0.967	1.000	-0.502	-0.269
<b>Pedestrian</b>	0.890	-0.442	-0.502	1.000	0.683
<b>Public Transport</b>	0.778	-0.169	-0.269	0.683	1.000

Table 7.2 lists the correlation coefficients for the suitability rankings of the modes along Voortrekker Road. The suitability of the bicycle mode is strongly positively correlated to pedestrian scores and the PT scores, indicating that, when these modes are well suited to an area, bicycle tends to be well suited as well. The bicycle mode has a moderate weak negative correlation to the car and freight modes, indicating that, although higher suitability for car and freight tends to coincide with lower suitability for bicycle, this relationship is not particularly strong.

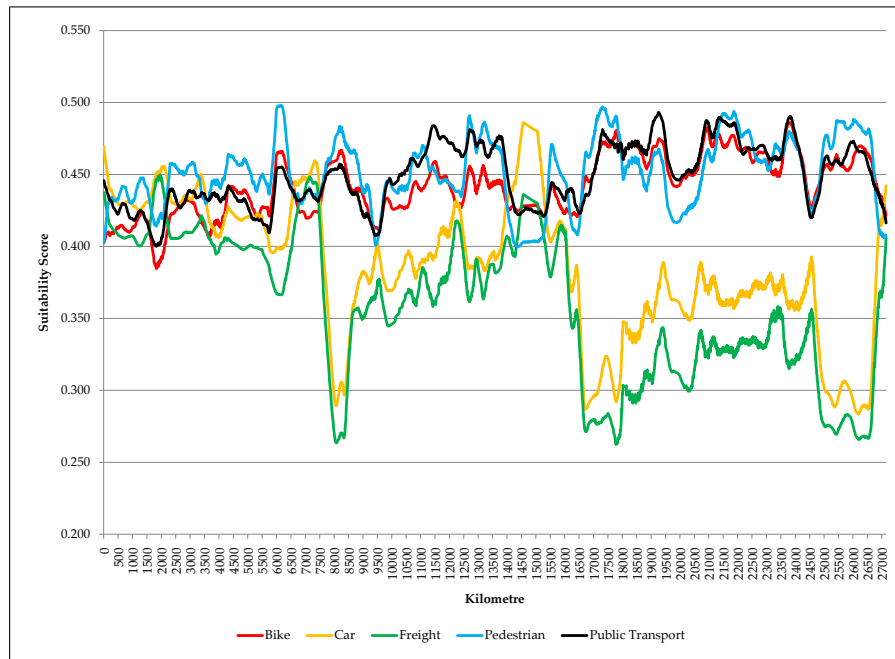
Car and freight suitability are very strongly positively correlated. From Figure 7.6 and Table 7.1, it can be seen that although the mean scores for these modes differ noticeably, their suitability performances tend to follow each other closely. When the suitability for the car mode increases, the suitability for the freight mode tends to increase as well, and so areas that are suitable for freight vehicles tend to also be suitable for cars, and vice versa.

Of all the modes, PT displays the weakest correlation to the other modes. PT has a high correlation with bicycle, but only has a weak negative correlation to car and freight, and has a medium to high correlation with the pedestrian mode.

The correlations between the performance scores is a result of the standardisation approaches adopted for the various data sets for the modes of transport. In the case of car and freight vehicles, all of the variations between the two modes is a result of differences in the standardised scores for land use, since the remaining data sets used in the SMCA for these modes were identical.

### 7.2.2 Lansdowne Road

Lansdowne Road is the longest case study road, being in excess of 27 km long. It passes through a wide range of land use types and income groups and could, therefore, be expected to show a large range of contextual variation along its route. Figure 7.9 shows the suitability scores for Lansdowne Road.



**Figure 7.9:** Suitability Scores along Lansdowne Road

There is significant variation in the suitability of the car mode and the freight mode along Lansdowne Road. These modes have approximately the same suitability levels as the other modes for the first seven kilometres, whereafter there is a dramatic localised drop in suitability, followed by a partial recovery from kilometre 9.0 to 14.0. In the vicinity of kilometre 15.0, freight vehicles have the highest suitability score of all the modes, but this is then followed by a significant, sustained reduction in suitability from kilometre 16.0 to kilometre 26.5. At the very end of the route, in the vicinity of kilometre 27.0, the vehicular mode suitability scores recover to similar levels as the other modes again.

The dramatic lowering of suitability scores for car and freight in the last third of the route is a result of the combination of high densities, low employment, high PT reliance and low incomes in this area. All of these criteria factor as spatial costs for the car and freight modes in the SMCA. The combination of these factors, therefore, has a strong negative influence on the suitability of the private vehicular modes in this area.

Throughout the route, NMT and PT maintain comparatively steady scores, suggesting that these modes tend to remain equally suited throughout the route. Considering the descriptives for the route, shown in Table 7.3, the effects are clear. All of the descriptives confirm that there is substantially more variation in the data sets for PMT, than for the other modes. Importantly, the mean scores for these modes are also significantly lower than for the other modes.

Table 7.4 lists the correlation coefficients for the suitability rankings of the modes along Lansdowne Road. The correlations between the mode suitability scores along Lansdowne Road are very similar to the other routes. This is, as mentioned, a result of the standardisation approach adopted for the various criteria.

**Table 7.3:** Descriptive Statistics for Lansdowne Road Suitability Scores

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
N	2716	2716	2716	2716	2716
Mean	0.442	0.385	0.358	0.451	0.449
Std. Deviation	0.021	0.048	0.050	0.023	0.021
Skewness	-0.064	-0.341	-0.247	-0.185	0.011
Kurtosis	-0.713	-0.458	-1.024	-0.588	-0.993
Minimum	0.385	0.283	0.262	0.400	0.400
Maximum	0.486	0.486	0.450	0.498	0.493
Range	0.102	0.203	0.187	0.098	0.093

**Table 7.4:** Mode Suitability Correlation Matrix for Lansdowne Road

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
Bike	1.000	-0.693	-0.763	0.715	0.848
Car	-0.693	1.000	0.970	-0.618	-0.520
Freight	-0.763	0.970	1.000	-0.540	-0.557
Pedestrian	0.715	-0.618	-0.540	1.000	0.669
Public Transport	0.848	-0.520	-0.557	0.669	1.000

### 7.2.3 Koeberg Road

Koeberg Road is the shortest case study road, being just over 10 km long. Figure 7.10 shows the suitability scores for Koeberg Road.

Koeberg Road has the least amount of variation of all the case study roads. This route is interesting to study because there are a large number of industrial areas along its course, as well as a large number of residential areas, with a wide range of incomes. One of the most noticeable features of the suitability scores along the route, is the very high score for freight vehicles in the first 1.5 km of the route. This is followed by a dramatic drop off in suitability at kilometre 2.0, and a similarly sharp recovery very shortly thereafter. To investigate this, Figure 7.11 shows an aerial overview of the suburbs along the route (shown in black), overlaid with the suitability raster for freight.

The first 1.5 km of the route passes by an oil refinery and a sewage treatment plant, explaining the high suitability for freight vehicles in this area. At kilometre 2.0 there is a community sports facility. This presence of this facility alters the context of the area, the effects of which can be seen in the sharp changes in the suitability of the motorised mode scores here. The sharp improvement thereafter, is a result of the high intensity industrial activity along the route for the next 3 km. This increase is then followed by a slow and unsteady decrease in the suitability of freight and car, and the inverse, an increase, for NMT and PT, as the context of the surrounding areas becomes more commercial and residential.

The next major feature along the route is a sharp decrease in the suitability of car and freight vehicles, in the area around kilometre 6.0. In this area the household density is sharply higher than the surrounding

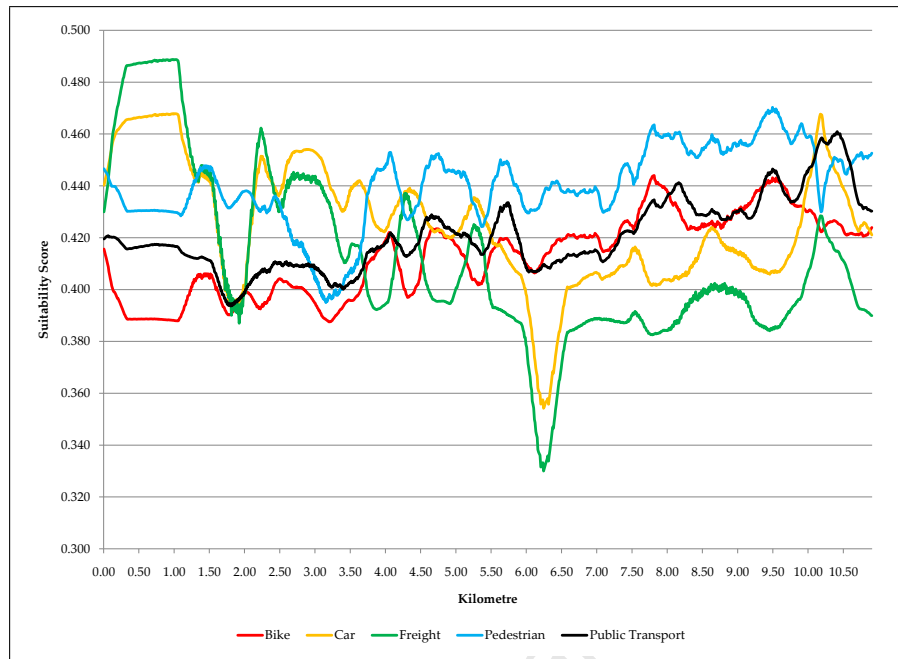


Figure 7.10: Suitability Scores along Koeberg Road

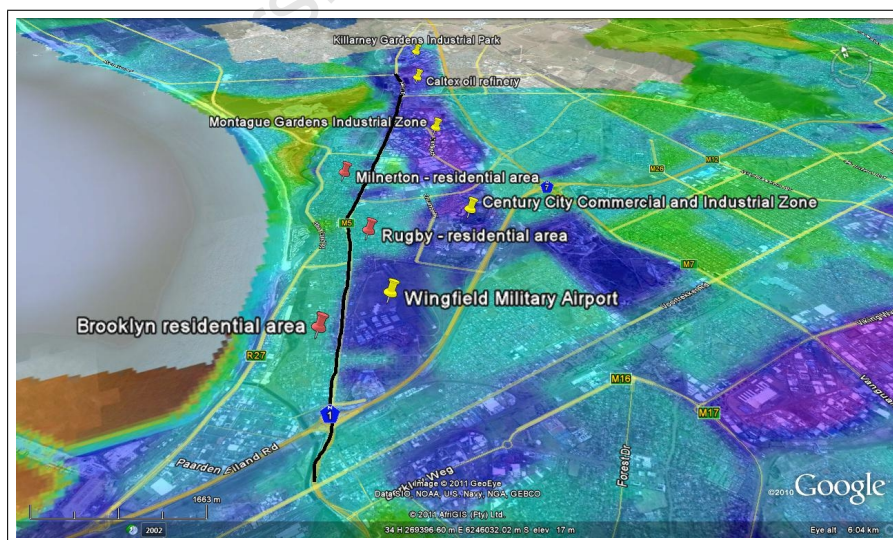


Figure 7.11: Aerial Overview of Freight Suitability along Koeberg Road

areas, there being a cluster of high rise residential tower blocks west of the road. The three kilometres beyond this area have a comparatively stable ranking order, where the modes have distinctly different average scores. In this area the route passes through the middle income residential neighbourhoods of Brooklyn and Rugby, and is flanked by shops and low rise mixed use buildings. NMT and PT modes score highest in this area. Near kilometre 10, where the route intersects with Voortrekker Road, the effects of the industrial areas of Paarden Island and Maitland changes the ranking orders again. Table 7.5 contains various descriptive statistics for the suitability scores along the route.

**Table 7.5:** Descriptive Statistics for Koeberg Road Suitability Scores

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
N	1091	1091	1091	1091	1091
Mean	0.413	0.426	0.412	0.440	0.421
Std. Deviation	0.015	0.023	0.033	0.015	0.014
Skewness	-0.026	-0.197	0.695	-0.705	0.700
Kurtosis	-1.060	0.080	0.356	0.641	0.247
Minimum	0.388	0.354	0.330	0.395	0.394
Maximum	0.444	0.468	0.489	0.470	0.461
Range	0.056	0.114	0.159	0.075	0.067

The descriptive statistics for the route show that, although the mean scores for the modes are very similar, the frequency distributions of the suitability scores are significantly different. The standard deviations for the NMT and PT modes are very small, indicating that the suitability of these modes is very stable along the route. Suitability for freight vehicles is the most varied amongst the modes along the route.

The kurtosis value for bicycle is very large and, being negative, it indicates that the frequency of scores for this mode is highly clustered, which is confirmed by the very small range of scores for the mode. The range of scores for freight vehicles is highest, with the majority of scores falling above the mean (because of the large positive skewness). This speaks to the duality of the suitability for this mode along the route. The mode is well suited in the first half of the route, but poorly suited in the second half. The range of scores is, therefore, high, but the mean score is very similar to the bicycle mode, which has a very stable suitability profile. Table 7.6 lists the correlation coefficients for the suitability rankings of the modes along Koeberg Road.

**Table 7.6:** Mode Suitability Correlation Matrix for Koeberg Road

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
Bike	1.000	-0.533	-0.692	0.826	0.738
Car	-0.533	1.000	0.922	-0.379	0.023
Freight	-0.692	0.922	1.000	-0.427	-0.199
Pedestrian	0.826	-0.379	-0.427	1.000	0.699
Public Transport	0.738	0.023	-0.199	0.699	1.000

As was seen with the data from the other studies, car and freight are highly positively correlated, as are

bicycle and pedestrian. Bicycle is also strongly positively correlated to PT. Moderate to high negative correlations are found between the public modes of bike, pedestrian and PT and the private modes of car and freight.

### 7.3 Weighting, Values and Scenario Planning

As discussed in Chapter 5, the cumulative scores in the SMCA are given by:

$$A_j = \sum_{i=1}^m w'_i \times r_{ij}, \quad i = 1, \dots, m; \quad j = 1, \dots, n \quad (7.1)$$

Where  $w'_i$  is the normalised weight for criterion  $i$ , and  $r_{ij}$  is the standardised score of alternative  $j$  for criterion  $i$ . In the analysis conducted in this chapter, the weights,  $w'_i$ , applied to the criterion scores were taken as being equal, that is, that all the criteria have an equal importance.

According to Yoe (2002), weighting is a technique for collecting data on the judgements of decision makers regarding the relative importance of a series of criteria. Hobbs and Meier (1994) notes that weights should be selected such that they are an accurate representation of the rate at which the decision maker is willing to trade-off one criterion for another. This notion of the weight as a measure of the trade-off value, implies that a utility gain in one criterion may be used to compensate for a loss in another criterion in proportion to the ratio of the weights of these two criteria (Keeney and Raiffa, 1993). Therefore, in terms of the assessment conducted in this research, the suitability of a particular mode to a particular location can be influenced using weighting to reflect the value judgements of the decision makers. An extension of this is that the analysis can be conducted such that a range of scenarios can be assessed and, consequently, a range of solutions generated.

Keshkamat et al. (2009) explored the use of weighting for scenario planning in their study on route selection for the Via Baltica project in Poland. They canvassed a range of stakeholders for inputs regarding weighting, and formulated four 'themes' from the criteria they selected, based upon these weights. Each theme (or weighting scheme) then generated an alternative route, ostensibly reflecting the priorities of the stakeholders who had contributed to the exercise.

Their approach presents an interesting opportunity for this research, since the concept of analysis themes is relevant to this research as well. However, implementing the concept in this research is not as straightforward. Firstly, the criteria used in this research was selected to give a measurable definition to the locational context. The spatial performance of each alternative (each mode) in relation to each criterion was assessed by defining the relationship between the alternative and the criterion, as either a spatial cost or benefit (see Tables 5.5 and 5.7). An identical increase in the value of all the alternatives for any criterion, at any one location would, therefore, have different effects on the relative performance of the alternatives for that location; some may improve their score, others will have their scores lowered.

For this same reason, the use of weights also presents a problem, since the standardisation process involves the application of value judgements to assess the performance of an alternative relative to a criterion (see Sections 5.4.1 and 5.4.2). The alternatives were each assigned a unique standardisation in accordance with the perceived relationship between the mode and the criteria. The influence of additional weighting on the final score could then be seen as being a sort of 'double weighting', since value judgements are being introduced at two points in the procedure. Since the standardisation

procedure already introduces these value judgements into the criterion scores, the addition of weighting will only serve to either reinforce, or weaken, the effects of the original value judgements on the alternatives.

Furthermore, since the SMCA scores are intended to be representations of a locations context, interpreted through the suitability of the various modes of transport to that context, weighting strategies only act so as to impose additional value judgements on this interpretation of the context. The values statements expressed in the standardisation framework (see Table 5.6), thus, become the distinguishing feature against which such weighting strategies are evaluated. Therefore, the stated values, such as '*maximise mobility*', or '*minimise impact*' would be the drivers of the weighting scheme selected, and not the individual criteria per se. However, following this approach would imply that the weights across all the criteria used in the evaluation will not sum to 1, meaning that some criteria could score more than their input values in the evaluation. An additional normalisation step must, therefore, be used to ensure that the weights applied sum to unity, as shown in Equation (7.2);

$$w_i'' = \frac{w_i'}{\sum_{i=1}^n w_i'} \quad (7.2)$$

where  $n$  is the total number of criteria and  $w_i'$  is calculated using either Equations (5.4), (5.5), (5.6) or the Analytical Hierarchy Process (see Saaty, 1980).

The land use variable was standardised separately, since the different modes each have different relationships to the various land uses. There are specific value statements for this variable, and so the method described here must be applied separately to each of its components. The individual land use component weights can then be applied to the standardised land use scores for inclusion in the analysis.

The weighting scheme would, therefore, take the form of Table 7.7. In the hypothetical weighting scheme shown, rank ordered weights were calculated using Equation (5.6), with the ranks selected arbitrarily. The values used are those described in Table 5.6.

Applying the same rank order as in Table 7.7, the land use variable is weighted, as shown in Table 7.8. In this way, the value statements identified are carried through to the land use variable, and are consistent with the weighting scheme as a whole.

The process described above develops a reasoned approach to the scenario evaluation described in Keshkamat et al. (2009), within the framework developed for this research. The scenarios developed here are a direct consequence of the values imposed on the evaluation (which are, in any event, reflected in the scenarios selected for evaluation by Keshkamat et al.), and the effects of favouring one perspective over another can be logically explored.

In terms of the methodology developed to define the context, then, the selection of weights is reduced to a mechanism with which to explore the implications of the values imposed during standardisation framework, and to develop scenarios that exploit the trade-offs between these values. An equally weighted scenario, then, is one where all the values imposed on the evaluation are held as being of equal importance.

## 7.4 Sensitivity to Weighting

In light of the discussion in Section 7.3, the results from the Voortrekker Road case study (Section 7.1) were explored under a range of scenarios. As mentioned in Section 7.3, the weights were applied to the

**Table 7.7:** A Hypothetical Value Driven Weighting Scheme

<b>Criteria</b>	<b>Value</b>	<b>Rank</b>	<b>Weight</b>
Land Use <sup>†</sup>	Varies	Varies	Varies
Household Density	Mobility of majority	1	0.143
Proportion of vulnerable road users	Safety of vulnerable road users	2	0.107
Income	Mobility of majority	1	0.143
Employment	Mobility of majority	1	0.143
Proximity to heritage sites	Access	4	0.036
Proximity to wetlands	Minimise impact	3	0.071
Proximity to ecologically sensitive areas	Minimise impact	3	0.071
Public Transport demand	Minimise impact	3	0.071
Private Car demand	Mobility of majority	1	0.143
Proximity to public transport stop	Minimise impact	3	0.071

<sup>†</sup> Land use inputs as per Table 5.5

**Table 7.8:** Hypothetical Value Driven Weighting Scheme for Land Uses

<b>Option</b>	<b>Value</b>	<b>Rank</b>	<b>Weight</b>
Residential	Safety	2	0.1429
Commercial	Access	3	0.0714
Industrial	Mobility	1	0.2143
Education	Safety	2	0.1429
Sports and Recreation	Access	3	0.0714
Vacant Land	Mobility	1	0.2143
Medical	Access	3	0.0714
Office	Access	3	0.0714

values expressed in the standardisation protocol. These values then formed the basis for the various scenarios generated. The values identified in Sections 5.4.1 and 5.4.2 are listed in Table 7.9, along with the rank afforded to each in the three scenarios investigated.

**Table 7.9:** Value Rankings for Scenarios for Voortrekker Road

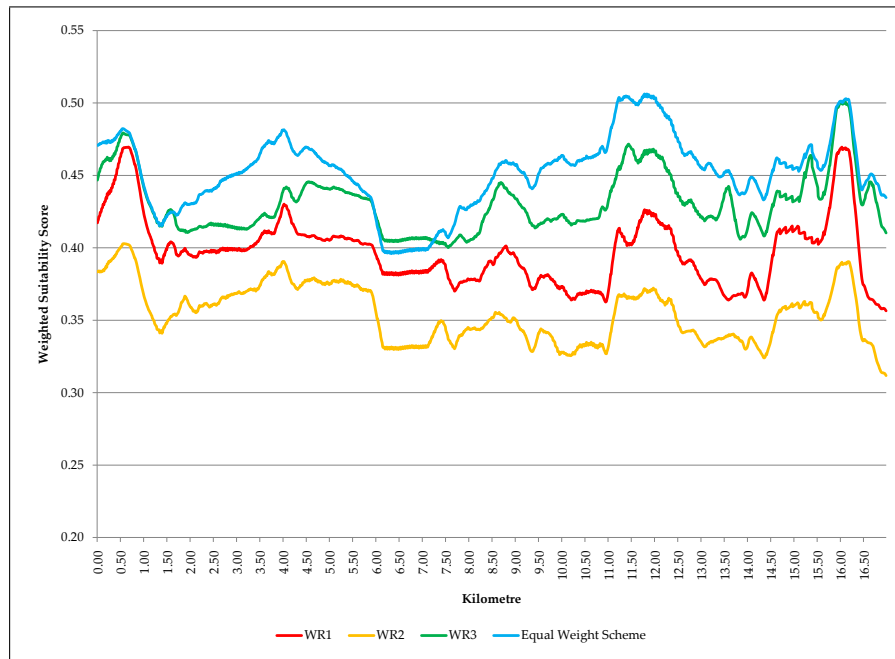
<b>Value</b>	<b>Rank Scheme 1</b>	<b>Rank Scheme 2</b>	<b>Rank Scheme 3</b>
Maximise Safety	2	1	4
Maximise Access	3	4	3
Maximise Mobility	1	2	2
Minimise Impact	4	3	1

**Table 7.10:** Weighting Schemes for Voortrekker Road

<b>Criterion</b>	<b>Value</b>	<b>Weight Scheme 1</b>	<b>Weight Scheme 2</b>	<b>Weight Scheme 3</b>
Household Density	Maximise Mobility	0.1429	0.1071	0.1071
Proportion of vulnerable road users	Maximise Safety	0.1071	0.1429	0.0357
Income	Maximise Mobility	0.1429	0.1071	0.1071
Employment	Maximise Mobility	0.1429	0.1071	0.1071
Proximity to heritage sites	Maximise Access	0.0714	0.0357	0.0714
Proximity to wetlands	Minimise Impact	0.0357	0.0714	0.1429
Proximity to ecologically sensitive areas	Minimise Impact	0.0357	0.0714	0.1429
Public Transport demand	Minimise Impact	0.0357	0.0714	0.1429
Private Car demand	Maximise Mobility	0.1429	0.1071	0.1071
Proximity to public transport stop	Minimise Impact	0.0357	0.0714	0.1429

The methodology used for the investigation was to extract a 400 m buffer around the centreline of Voortrekker Road, from the criteria rasters. This was done to minimise the processing time for the analysis and, since the neighbourhood averaging process only searches 150 m beyond each point, and the only points extracted are those along the route centreline, it has no impact on the final scores. The cell values in the extracted criteria maps were then factored by the required weight, and the weighted criteria rasters summed. The mode suitability scores were then extracted and exported to a spreadsheet for further analysis, as was done previously. The results of this exercise on the PT suitability scores are shown in Figure 7.12.

The weighted schemes produce significantly different results to the equally weighted results. The most notable difference is that the mean scores produced are substantially lower than the equally weighted



**Figure 7.12:** Comparison of Public Transport Suitability Using Different Weight Schemes

scores. The original suitability scores are, generally, greater than the weighted scores, except for a few locations, such as between kilometre 6.0 and 7.5, where the scores from weight scheme 3 are the best. Table 7.11 gives the mean suitability scores and the standard deviation of the suitability scores for the PT mode along Voortrekker Road.

**Table 7.11:** Descriptive Statistics for Weighted Analysis Results for Public Transport

	Mean	Std. Deviation
Public Transport	0.453	0.026
Weight Scheme 1	0.398	0.024
Weight Scheme 2	0.355	0.020
Weight Scheme 3	0.431	0.021

The mean scores for the weighted schemes are all lower than that of the equally weighted scheme. However, the standard deviations for the weighted schemes are lower than that of the equally weighted scheme. There is, thus, less variation within the weighted scheme data sets than in the equally weighted scheme. Table 7.12 shows the correlation between the weight scheme scores and the equally weighted scores.

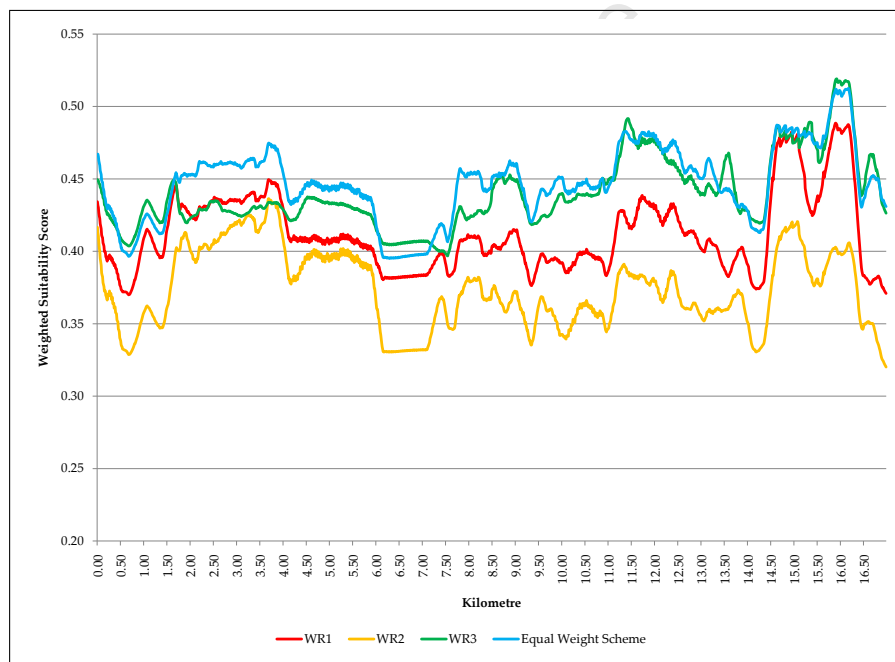
Although all the results are positively correlated to the original data set, weight schemes 1 and 2 have a lower correlation to the original results than weight scheme 3. In fact, the results from weight scheme 3 are more highly correlated to the original results than it is to weight scheme 1. The highest correlations are between weight schemes 1 and 2. A visual inspection of the results in Figure 7.12 shows that, the mean scores notwithstanding, the differences in the profiles themselves are more subtle than what

**Table 7.12:** Correlation Between Various Weight Analysis Results for Public Transport

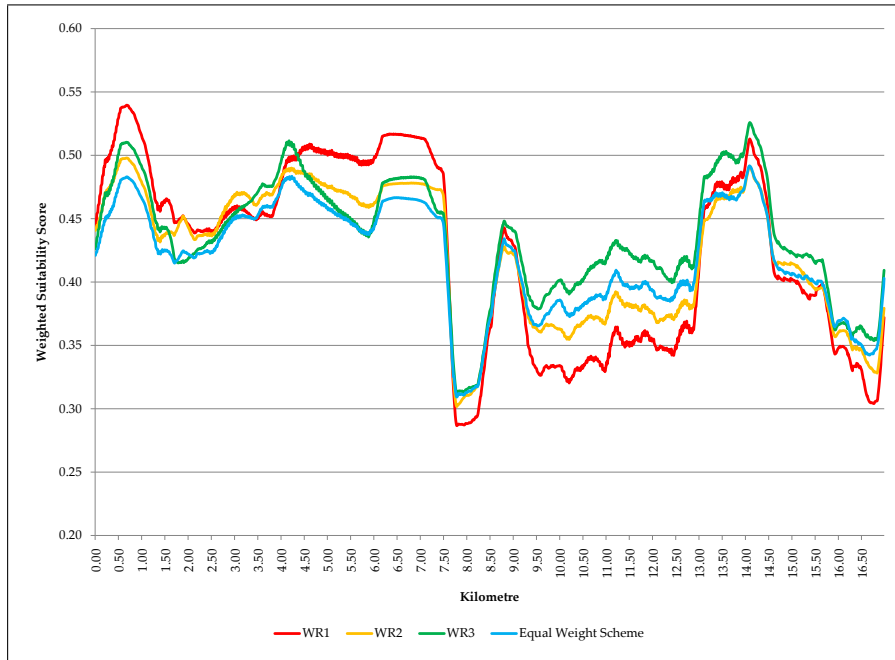
	Public Transport	Weight Scheme 1	Weight Scheme 2	Weight Scheme 3
Public Transport	1.000	0.529	0.541	0.796
Weight Scheme 1	0.529	1.000	0.898	0.782
Weight Scheme 2	0.541	0.898	1.000	0.665
Weight Scheme 3	0.796	0.782	0.665	1.000

Pearsons  $r$  statistic indicates. The underlying structure of the context can clearly be identified from all the results, and the various weighting schemes either accentuate or mask these features.

Similar results were found for the pedestrian mode. Weight scheme 2 delivered the lowest correlation to the unweighted data set, and weight scheme 3 delivered results that were most highly correlated to the original results. In Figure 7.13, the results from the different weighting schemes are compared to the equally weighted results for the pedestrian mode.

**Figure 7.13:** Comparison of Pedestrian Suitability Using Different Weight Schemes

In Figure 7.14, the results from the different weighting schemes are compared to the equally weighted results for the car mode. The weighting schemes have somewhat different effects than for pedestrian and PT. Whereas, on average, those modes performed worse under the weighting schemes, the car mode performs somewhat better. Under weight schemes 1, 2 and 3, the suitability is higher than with the equally weighted scenario, until kilometre 7.5. Thereafter, until kilometre 13.0, the suitability is lower under weight schemes 1 and 2, and better under weight scheme 3. This variation tends to cancel out, and so the mean scores for each weight scenario are not a good metric to convey the effects of the various weighting scenarios.



**Figure 7.14:** Comparison of Car Suitability Using Different Weight Schemes

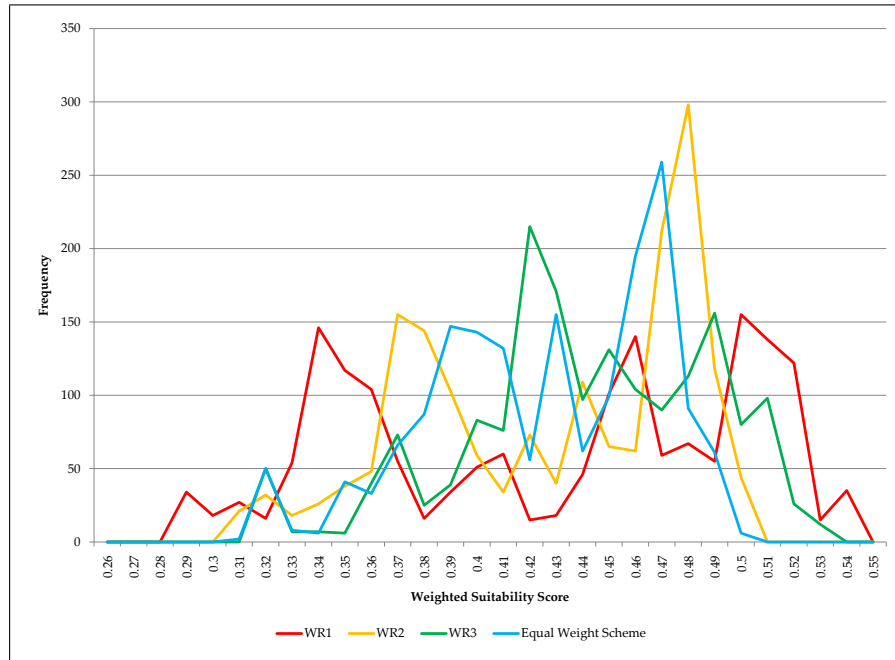
The descriptive statistics in Table 7.13 allow for a better analysis of the results. As mentioned, the mean scores are very similar, as are the standard deviations. The maximum and minimum values, as well as the range, provides some further clarity. Weight scheme 1 has the largest range, the lowest minimum and the largest maximum scores. This scheme tends to have the most dramatic effects on the scores for this mode.

**Table 7.13:** Descriptive Statistics for Weighted Analysis Results for Car

	Car	Weight Scheme 1	Weight Scheme 2	Weight Scheme 3
Mean	0.421	0.424	0.424	0.436
Std. Deviation	0.043	0.072	0.052	0.047
Skewness	-0.514	-0.225	-0.409	-0.457
Kurtosis	-0.499	-1.354	-1.121	-0.159
Range	0.183	0.253	0.196	0.214
Minimum	0.309	0.287	0.302	0.312
Maximum	0.492	0.540	0.498	0.526

Figure 7.15 shows the histograms for the various weighting scheme profiles. The histograms confirm the statistics in Table 7.13. Weight Scheme 1 has the largest range of scores, whereas the equally weighted data set has the lowest range. Weight scheme 2 produces results that are closest to the equally weighted results.

The effects of different weighting schemes on the results are significant. Figures E.6 to E.8 shows the suitability score results along Voortrekker Road for each weight scheme. The bulk of the variation



**Figure 7.15:** Comparison of Histograms for Different Weight Schemes Profiles for Car

in the results lies in the suitability of the car and freight modes, which vary, dramatically, along the route. The different weight schemes produce different rank orders along the routes. Some areas of the route have a different rank order under each scenario, whereas others always have the same rank order. Between kilometres 6.0 and 8.0 the rank order is very stable, whereas there is a lot of variation between weighting schemes in the rank orders between kilometres 9.0 and 13.0.

Areas of rank stability are important, because they have contextual features that make them highly suited to certain modes, irrespective of weighting. Areas of rank instability are also important, since, depending on the weight scheme used (from the decision makers perspective or priorities), the rank order in these areas may change. It is, therefore, important to explore how these changes occur, and which criteria are responsible for the reversal in rankings.

Weighting is, therefore, an important consideration in the analysis and results, under a range of weighting scenarios, should be used to explore the context as thoroughly as possible before making decisions about infrastructure provision.

## 7.5 Résumé

Suitability rasters were generated for the whole city, and an analysis of these results was conducted. The analysis of the results produced the descriptive statistics shown in Table 7.14.

Case studies were conducted on Voortrekker Road, Lansdowne Road and Koeberg Road in Cape Town. The results of the studies showed that mode suitability is highly variable along the routes, as was expected. With the criteria used in this SMCA, bicycle, pedestrian and PT modes have remarkably similar cumulative suitability scores, which is notably higher than that of freight and car. There are, however, still noticeable differences between the scores for each mode. The freight mode has a much more even distribution of scores than the car mode, although, in general, the car mode scores better than the freight mode.

**Table 7.14:** Descriptive Statistics for Case Study Routes

	<b>Bike</b>	<b>Car</b>	<b>Freight</b>	<b>Pedestrian</b>	<b>Public Transport</b>
<b>Voortrekker Road</b>					
<i>(N = 1 699)</i>					
Mean	0.425	0.421	0.394	0.447	0.453
Std. Deviation	0.020	0.043	0.045	0.025	0.026
Skewness	0.032	-0.514	-0.291	-0.094	-0.180
Kurtosis	0.061	-0.499	-0.437	-0.070	-0.028
Minimum	0.384	0.309	0.281	0.395	0.396
Maximum	0.481	0.492	0.480	0.513	0.506
Range	0.097	0.183	0.199	0.117	0.110
<b>Lansdowne Road</b>					
<i>(N = 2 716)</i>					
Mean	0.442	0.385	0.358	0.451	0.449
Std. Deviation	0.021	0.048	0.050	0.023	0.021
Skewness	-0.064	-0.341	-0.247	-0.185	0.011
Kurtosis	-0.713	-0.458	-1.024	-0.588	-0.993
Minimum	0.385	0.283	0.262	0.400	0.400
Maximum	0.486	0.486	0.450	0.498	0.493
Range	0.102	0.203	0.187	0.098	0.093
<b>Koeberg Road</b>					
<i>(N = 1 091)</i>					
Mean	0.413	0.426	0.412	0.440	0.421
Std. Deviation	0.015	0.023	0.033	0.015	0.014
Skewness	-0.026	-0.197	0.695	-0.705	0.700
Kurtosis	-1.060	0.080	0.356	0.641	0.247
Minimum	0.388	0.354	0.330	0.395	0.394
Maximum	0.444	0.468	0.489	0.470	0.461
Range	0.056	0.114	0.159	0.075	0.067

The results for Voortrekker Road shows significant variation along the course of the route. In terms of mean scores, on average, PT is the most suited mode along Voortrekker Road, followed by pedestrian. Freight is, on average, the lowest scoring mode. The car and freight modes have the largest amount of variation in their data sets, on average double that of the other modes.

Koeberg Road has the least amount of variation of all the case study roads. The mean suitability of freight along this route is the highest of all the case study routes, highlighting the importance of industry along the route.

Along Lansdowne Road the car and freight modes have approximately the same suitability levels as the other modes for the first 7 kilometres, whereafter there is a dramatic localised drop in suitability, followed by a partial recovery from kilometre 9.0 to 14.0. The dramatic lowering of suitability scores for car and freight in the last third of the route is a result of the combination of high densities, low employment, high PT reliance and low incomes in this area. All of these criteria factor as spatial costs for the car and freight modes. The combination of these factors, therefore, has a strong negative influence on the suitability of the private vehicular modes in this area. Throughout the route, pedestrian, bicycle and PT maintain comparatively steady scores, suggesting that these modes tend to remain equally suited throughout the route.

It was found that, in general, the suitability of the bicycle mode is most highly correlated to pedestrian scores and to the PT scores, indicating that when these modes are well suited in an area, bicycle tends to be well suited as well. Car and freight suitability was also found to be strongly positively correlated. Of all the modes, PT displays the weakest correlation with the other modes.

Weights should be selected such that they are an accurate representation of the rate at which the decision maker is willing to trade-off one criterion for another and, therefore, a utility gain in one criterion may be used to compensate for a loss in another criterion in proportion to the ratio of the weights of these two criteria. The suitability of a particular mode to a specific location can, therefore, be influenced using weighting to reflect the value judgements of the decision makers, and weighting can be used to explore various preference scenarios. Although weights have been used this way in other SMCA studies before, implementing this in this research presents some theoretical difficulties since, firstly, the spatial performance of each mode, in relation to each criterion, was assessed by defining the relationship between the mode and the criterion as either a spatial cost or a benefit and, therefore, an identical increase in the value of all the alternatives for any criterion at any one location would have different effects on the relative performance of the alternatives for that location.

The use of weights is, therefore, also somewhat problematic since the standardisation process already involves the application of value judgements to assess the performance of an alternative relative to a criterion. The influence of additional weighting on the final score would then be a sort of 'double weighting', since value judgements are being introduced at two points in the procedure. Since weighting strategies act to impose additional value judgements on this interpretation of the context, the values statements expressed in the standardisation framework, therefore, become the distinguishing feature against which such weighting strategies are evaluated. Therefore, the stated values would be the drivers of the weighting scheme selected, and not the individual criteria per se. Following this approach would imply that the weights across all the criteria used in the evaluation will not sum to 1, and so an additional normalisation step must be used to ensure that the weights used sum correctly. The land use variable was standardised separately, since the different modes each have different relationships to the various land uses.

The effects of three weighting schemes were tested on the suitability scores for Voortrekker Road. The methodology used for the investigation was to extract a 400 m buffer around the centreline of

Voortrekker Road from the criteria rasters. This was done to minimise the processing time for the analysis, and since the neighbourhood averaging process only searches 150 m beyond each point, and the only points extracted are those along the route centreline, has no impact on the final scores. The cell values in the extracted criteria maps were then factored by the required weight, and the weighted criteria rasters summed. The mode suitability scores were then extracted and exported to a spreadsheet for further analysis.

The effects of different weighting schemes on the results are significant. The bulk of the variation in the results lies in the suitability of the car and freight modes, which vary dramatically along the route. The different weight schemes do produce different rank orders along the routes in different places. Some areas of the route have a different rank order under each scenario, whereas others always have the same rank order. Between kilometres 6.0 and 8.0 the rank order is very stable, whereas there is a lot of variation between weighting schemes in the rank orders between kilometres 9.0 and 13.0.

Areas of rank stability were found to be important because they have contextual features that make them highly suited to certain modes, irrespective of weighting. Areas of rank instability are also important since, depending on the weight scheme used (from the decision makers perspective or priorities), the rank order in these areas may change. It is, therefore, important to explore how these changes occur, and which criteria are responsible for the reversal in rankings. Weighting is, therefore, an important consideration in the analysis, and results should be explored under a range of weighting scenarios to explore the context as thoroughly as possible before making decisions about infrastructure provision.



## Chapter 8

# Cluster Analysis

The data generated and analysed in Chapter 7 can provide useful insights into the spatial variation of the local context. The methodology used to generate this information, as described in Chapter 5, provides a means of quantifying the context and describing the implications thereof, in terms of the relative priorities between the modes of transport. Although the analysis of this data can provide useful information for the road planner, the variability within the output makes it necessary to simplify the results that are produced by clustering similar sections into groups that can then be analysed further and applied as proxies for the actual scores. These proxy results can be thought of as rank types, for which road design templates can be developed that would best suit the context. This chapter investigates clustering theories and methods and describes their application to the SMCA results. The clusters produced are briefly analysed, and their relationship to the context discussed.

### 8.1 Introduction to Cluster Analysis

Everitt (1994) defined cluster analysis as being “... a generic term used for a large number of techniques which attempt to determine whether or not a data set contains distinct groups, and, if so, to find the groups.” Jain (2010) defines the goal of clustering as discovering the natural grouping(s) of a set of observations. He notes that clustering is typically used in the following applications:

- To ascertain the underlying structure of a set of data in order to gain insight into data, generate hypotheses, detect anomalies, and identify salient features.
- To establish a natural classification by establishing the degree of similarity among forms in a data set.
- To simplify data through compression where clustering is employed as a method for organizing the data and summarising it through cluster prototypes.

All of these applications are relevant to the exercise being conducted in this research. The methodology used in this research rests upon the hypothesis that there is an underlying contextual structure along the route, and that this can be used to establish the priorities that should be afforded to each mode of transport in use along the route. In that sense, although certain sections of the route may be unique, it is likely that there will be locations that are spatially separate, but also contextually similar, and that these should, therefore, be treated similarly. The data generated displays a high ‘noise’ level that is undesirable in interpreting the results. Clustering the data could assist with eliminating such noise and simplify the planning process.

Jain et al. (1999) notes that there are two distinct fields of data clustering, supervised and unsupervised classification. In supervised classification there exists a collection of already labelled, or pre-classified, patterns of observations. The problem is to sort the remaining unclassified sets of observations into one of these pre-existing groupings. In the case of unsupervised classification, the problem is to group a collection of unlabelled patterns, or sets of observations, into meaningful clusters. Labels are, therefore, also associated with clustering, but these labels are now obtained solely from within the data set itself. This is the operation which needs to be performed on the data sets generated from the SMCA conducted in this research.

Han and Kamber (2006) identify two main clustering techniques, partitional and hierarchical. A partitional clustering algorithm constructs partitions in the data, such that each cluster optimises a pre-stated clustering criterion, such as the minimisation of the sum of the squared distances from the mean within each cluster. The popular K-means algorithm is an example of a partitional algorithm.

As the name states, a hierarchical algorithm creates a hierarchy amongst the sets of data points. Hierarchical algorithms can be either agglomerative or divisive. Agglomerative algorithms start with each object being a separate cluster itself, and successively merges groups according to a distance measure. The clustering stops when all objects are in a single group, or at any other point that the investigator chooses by analysing the cluster tree to find the most appropriate cluster groupings. Divisive algorithms start with one group containing all the observations and successively splits this group into smaller ones, until each object falls in a unique cluster, or as desired. Andritsos (2002) compiled a list of common algorithms, comparing their performance and limitations (Table F.1).

### 8.1.1 Partitional Algorithms

The K-means algorithm minimises the squared-error objective function to produce  $k$  clusters from a set of  $n$  objects.

$$E = \sum_{i=1}^k \sum_{\vec{p} \in C_i} |\vec{p} - \vec{m}_i|^2 \quad (8.1)$$

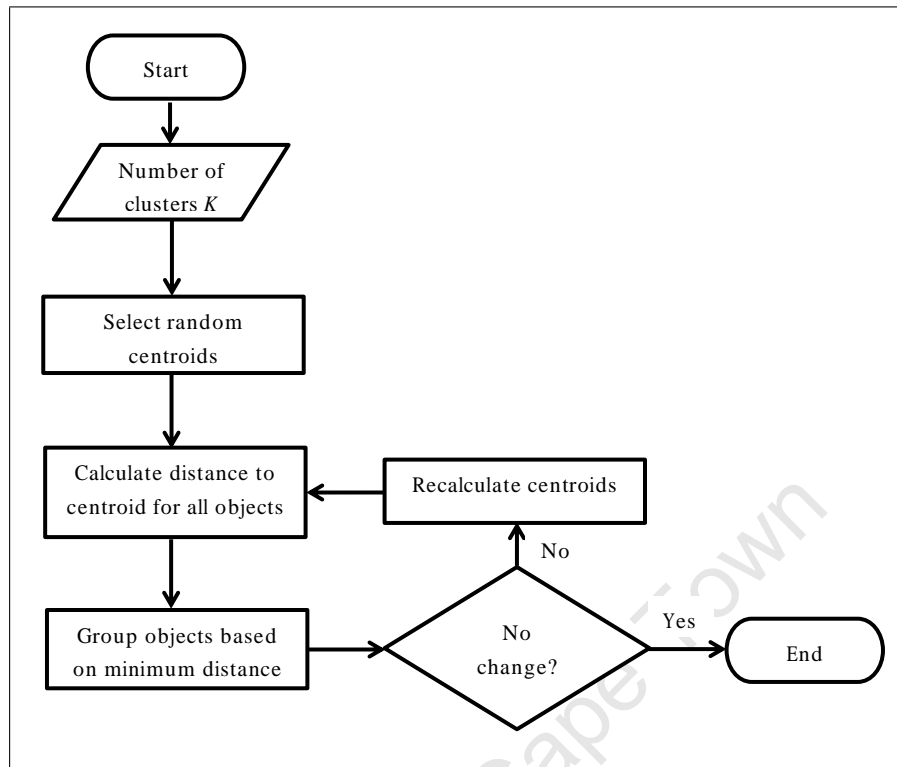
In Equation (8.1),  $C_i$  are the clusters,  $\vec{p}$  is a point vector in cluster  $C_i$  and  $\vec{m}_i$  is the mean of cluster  $C_i$ . The mean of a cluster is also given by a vector, which contains, for each attribute, the mean values of the data objects in this cluster, as shown in Equation (8.2).

$$\vec{m}_i = \frac{\sum_{i=1}^{n_{C_i}} \vec{p}}{n_{C_i}} \quad (8.2)$$

In Equation (8.2),  $n_{C_i}$  is the total number of individual points in cluster  $C_i$ .

The algorithm requires the user to input the required number of clusters and, as an output, the algorithm returns the means of every cluster  $C_i$  (Andritsos, 2002). The distance measure employed to calculate the difference between the point vector  $p_i$  and the cluster mean vector  $m_i$  is usually the Euclidean distance, although other measures are also occasionally used (Berkhin, 2006).

The steps performed by the K-means algorithm is summarised in Figure 8.1.



**Figure 8.1:** Steps performed by the K-means algorithm

In the first step, the user inputs the number of clusters  $k$  required. The algorithm then randomly selects  $k$  objects as initial cluster centroids. Each data object is then assigned to its closest cluster centroid, according to the distance calculated using whichever distance measure was selected. The centroids of each cluster are then recalculated using Equation (8.2). Steps two and three are repeated until the cluster centroids do not change (Andritsos, 2002).

In general, non-hierarchical partitional algorithms are very sensitive to the initial partition, which cannot be controlled by the user, since this is assigned randomly and, in general, the K-means algorithm performs poorly when random initial partitions are used (Andreas, 2005). However, a number of starting partitions can be used, also called replicates, that allows the algorithm to terminate at a local optimum. There are a number of variations on the K-means algorithm that attempt to address some of its shortcomings. These are briefly discussed in Appendix F.

The performance of the K-means algorithm can be increased when the results obtained from hierarchical methods are employed to constitute the initial partition (Fred and Jain, 2005). In this sense, hierarchical and non-hierarchical methods can be used as complementary clustering techniques.

### 8.1.2 Hierarchical Algorithms

A hierarchical clustering algorithm builds a cluster hierarchy, or a tree of clusters, also known as a dendrogram. Every cluster node contains child clusters, and sibling clusters partition the points covered by their common parent. This approach allows the investigator to explore the data at different levels of granularity. Hierarchical clustering methods are categorised into agglomerative and divisive approaches (Gower, 1967). An agglomerative clustering starts with one-point (singleton) clusters and recursively

merges two or more of the most similar clusters. A divisive clustering starts with a single cluster containing all data points and recursively splits the most appropriate cluster. The process continues until a stopping criterion (frequently, the requested number of clusters) is achieved. Advantages of hierarchical clustering include:

1. Flexibility regarding the level of granularity.
2. Ease of handling any form of similarity or distance measure.
3. Applicability to any attribute types.

Disadvantages of hierarchical clustering are related to:

1. Vagueness of termination criteria.
2. Most hierarchical algorithms do not revisit (intermediate) clusters once constructed.

Divisive algorithms are not widely used in practise. Although they can perform well in terms of computational speed and they scale well for large data sets, divisive clustering criterion are generally theoretically weak and inefficient, which leads to poor, unsatisfactory results (Andreas, 2005). It is computationally impossible to find division of a cluster into two new clusters that optimises a divisive clustering criterion (Garey and Johnson, 1979). Agglomerative algorithms tend to provide more precision at the bottom of the cluster tree, whereas divisive algorithms provide better precision at the top of the cluster tree. Divisive algorithms are, therefore, often used in applications where a small number of large clusters must be extracted from a large database, such as for genetics applications. Agglomerative algorithms are, on the other hand, by far more common in the literature (see Andreopoulos et al., 2009; Patras et al., 2011; Scheunders, 1997; Wu et al., 1999), and have been noted as producing more robust results than divisive algorithms, although they can be computationally expensive for very large data sets (> 2 000 objects) (Ng and Han, 1994).

Hierarchical algorithms function by calculating the dissimilarity between pairs of multidimensional objects which are represented as vectors. The dissimilarity indices are then compiled into a dissimilarity matrix. There are a large number of similarity measures available to calculate the index, although the simple Euclidean distance is the most commonly mentioned measure (Wu et al., 1999).

Once the distances between objects in the data set have been calculated, it can then be determined how objects in the data set should be grouped into clusters, using a linkage function. The linkage function uses the distance information in the dissimilarity matrix and links pairs of objects that are close together into clusters made up of two objects. The linkage function then links these newly formed clusters to each other, and to other objects, to create bigger clusters until all the objects in the original data set are linked together in a hierarchical tree. The linkage function uses distances to determine the order in which objects are clustered. The single linkage method is the most commonly used approach, using the shortest distances to pair objects. The results of the hierarchical cluster analysis are best represented as a dendrogram, which shows the relationships between clusters in the hierarchy of clusters (see Figure 8.3).

The dendrogram is a particularly useful visualisation, in that the investigator can ascertain the underlying structure of clusters in the data set, and visually determine the distance between clusters. The dissimilarity height can be used to determine appropriate points in the hierarchy to stop clustering. The number of clusters at this point is given by the number of vertical lines crossed. Aside from being able to cluster data objects this way, the dendrogram also provides a useful starting point for determining the number of clusters to use in other cluster methods, such as in the K-means method (see Section 8.1.1).

## 8.2 Clustering the SMCA Data

As discussed in Section 8.1, the various cluster algorithms have been developed to address the shortcomings of each method with respect to the data being clustered, theory, or computational complexity. In terms of computational complexity, the data sets being clustered in this analysis would have, up until recently, been considered large. Zhang et al. (1996) mention a feasible range of 2 000 datapoints with 3 dimensions as being the upper limit for reasonable computation times using hierarchical clustering in 1996. Improvements in computer speeds, however, means that the data sets for this research (Lansdowne Road has >2 700 points with 5 dimensions) can be processed in a matter of minutes<sup>28</sup>.

MATLAB<sup>®</sup> V7.10 by MathWorks Inc. was used to carry out the K-means and hierarchical cluster analysis, since it has a good range of customisable parameter settings. The Voortrekker Road case study will be used to investigate the clusters produced under different clustering methods, whereafter the results for each case study will be compared based on the preferred method.

### 8.2.1 Data Investigation

Before clustering the data, the relationships and distribution of datapoints was investigated. Since the data is 5 dimensional, it is not possible to view all the data simultaneously, only three modes can be viewed at once. A scatter plot matrix provides an overview of the mode score relationships for each case study road (Figures 8.2, F.2 and F.3).

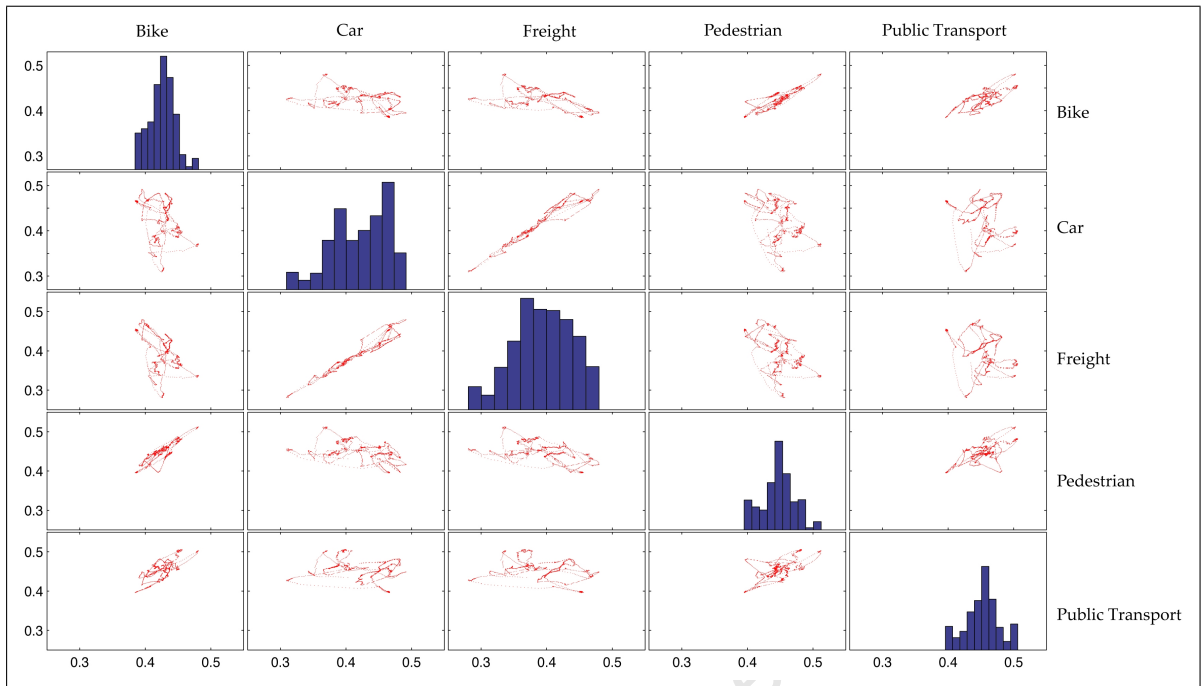
There is a surprising amount of similarity in the preference score relationships between modes across the three case study roads. Consider the relationships between bike and car (second column, first row). On Voortrekker Road, bike has a very narrow range of scores clustered around 0.43, whereas car has a very large range of scores, varying from 0.3 to 0.5 (this can be seen from the histograms for these two modes). The resultant scatter plot for these modes is a narrow, horizontally elongated strip of points. On Lansdowne Road, the pattern is more pronounced; since the route is longer, it has more sample points. On Koeberg Road, this basic shape is repeated again, albeit with sparser data points (since this is the shortest route). The remaining plots also display commonalities in the orientations of the primary axes and the location of the centroids for all the routes.

For some modes, scores tend to be tightly clustered (see pedestrian/PT and pedestrian/bike), whereas for others the scores are more dispersed (pedestrian/car, PT/car). These commonalities seem to indicate that there would be similarities between the clusters produced along the case study roads, and that there are, therefore, commonalities between the contexts found along each route.

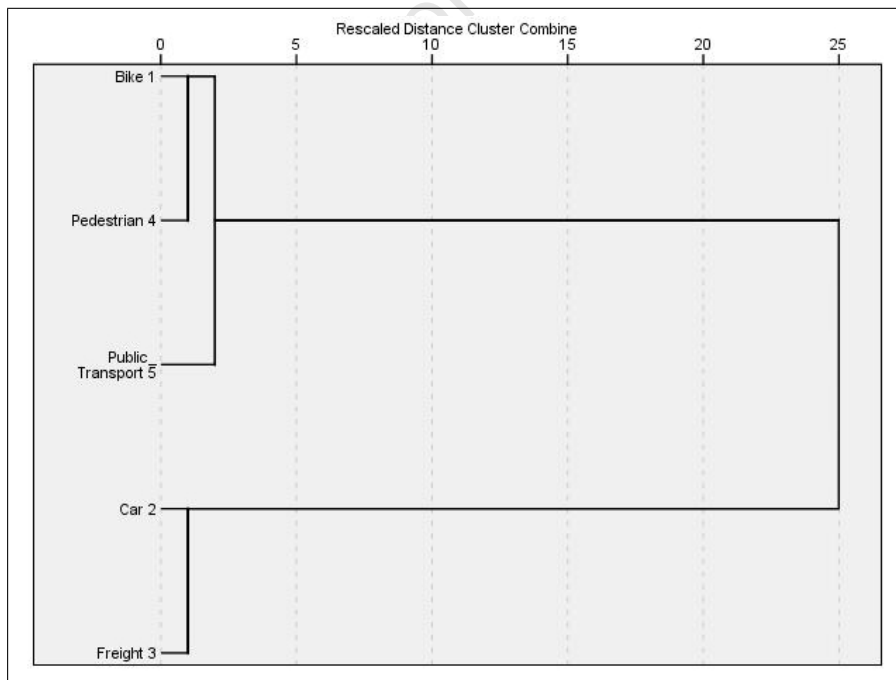
Certain mode groups also tend to have similar relationships. The relationships of car and freight (the vehicular modes) to the NMT modes tend to be similar, and vice versa. Pedestrian and bike tend to have similar relationships as well. This is evident for all the case study roads. These relationships are found in Figures 7.6, 7.9 and 7.10 as well. The general orientation of the primary axis of the points for any mode score comparison plot reveals the differences in the priority statements for those modes. When freight scores well, car tends to score well too, and the same can be said for pedestrian and bicycle. These relationships are, therefore, an artifact of the priority statements applied to the analysis.

An analysis of the SMCA data for Voortrekker Road reveals that there are two very distinct groupings in

<sup>28</sup> The actual time taken to process a hierarchical analysis is heavily dependent upon the distance measure being used. Euclidean distances can be calculated very rapidly, whereas more complicated measures, such as Pearson correlations, require much more time. Nonetheless, experimenting with the different approaches has shown that clustering, using all of the measures available, can be completed in less than 10 minutes per run on a modern computer.



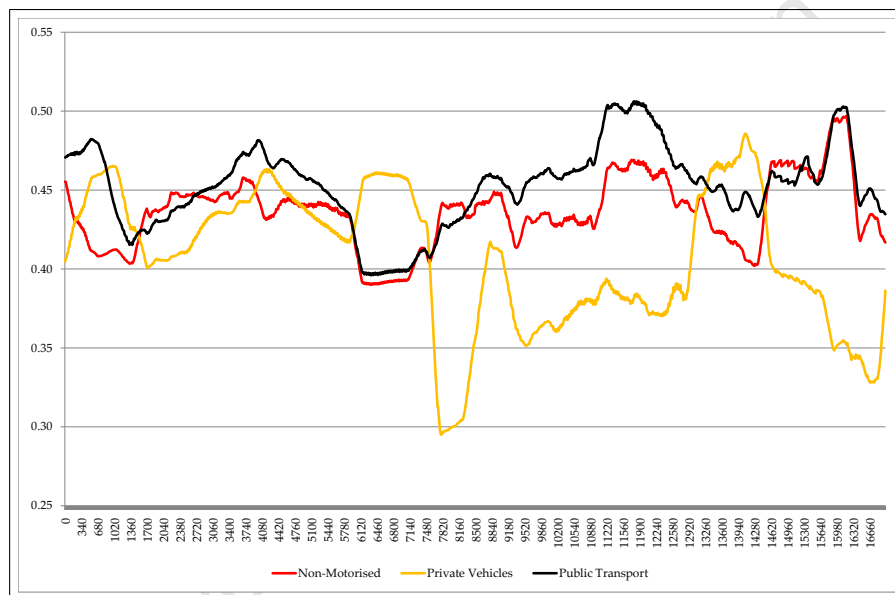
**Figure 8.2:** Mode score relationships for Voortrekker Road. Each mode’s preference scores are plotted against every other mode’s preference scores in a scatter plot, and the histograms for each mode is shown along the central diagonal.



**Figure 8.3:** Similarities between modes along Voortrekker Road. A dendrogram of the SMCA data for Voortrekker Road that has been clustered by mode using hierarchical clustering using the cosine distance<sup>29</sup> as the similarity measure, and between groups linkage. The lengths of the horizontal links are an indication of how different a mode is from the others.

the modes along the route (Figure 8.3). Freight and car form one group, and bicycle, PT and pedestrian make up the other. Within this second grouping, a subgroup, less distinct than the overarching groups, can be formed between pedestrian and bike. This correlation between modes presents a problem when clustering, since a strong clustering relies on there being distinct differences in the elements being clustered. The less differences there are, the less distinct are the clusters that are formed.

Figure 8.3 suggests that a simplified clustering could be conducted by merging freight and car, and pedestrian and bicycle, and leaving a third mode option of PT (Figure 8.4). The cluster analysis would, therefore, be looking for sections of the route with similar priorities between NMT, PT and PMT. The priorities of the individual modes in the mode groupings could then be determined from the average priority for each mode, at the location of the cluster. This may help to address the problem of correlation between mode priorities in the data set. A further benefit of this approach is that clusters can be readily visualised to determine whether they are logical or not.



**Figure 8.4:** Merged suitability scores along Voortrekker Road. The scores were calculated by taking the average score for pedestrian and bicycle as the score for NMT, the average for freight and car as the score for private vehicles, and leaving the original score for PT unchanged.

## 8.2.2 Selecting the Number of Clusters and Data Clustering

Choosing the final number of clusters to represent a data set is challenging. Each additional cluster may decrease the overall sum of cluster distances, but at the cost of increasing the complexity of the representation. Including too many clusters leads to a loss of information regarding the nature of key relationships in the data set. The Voortrekker Road dataset is used to demonstrate the problem, and the method used to overcome it.

<sup>29</sup> The cosine similarity is a measure of the cosine of the angle between two vectors (two data points). The cosine function equals 1 when the angle is 0, and is less than 1 when the angle is any other value. The cosine of the angle between two vectors is a measure of the extent to which two vectors are pointing in the same direction (Tan et al., 2005). The cosine similarity measure is given by:

$$\text{Similarity} = \cos \Theta = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

The K-means algorithm, described in Section 8.1.1, was carried out on the merged data set. K-means clustering uses an iterative algorithm that assigns objects to clusters, so that the sum of distances from each object to its cluster centroid, over all clusters, is minimised. It requires three inputs: the number of clusters, the distance measure used, and the linkage method.

There are a range of possible distance measures that can be used, and a range of linkage methods that can be selected. The ‘best’ combination amongst these depends upon the type of data used, and the distribution of this data in the data space. The data being clustered here is numerical, scalar data. That is, the data set consists of numbers upon which arithmetic operations can be sensibly carried out (as opposed to ordinal or nominal numerical data, such as ‘first’ and ‘second’ (ordinal) or ‘bus number 8’ (nominal)). Euclidean distances are, therefore, a sensible choice for the distance measure.

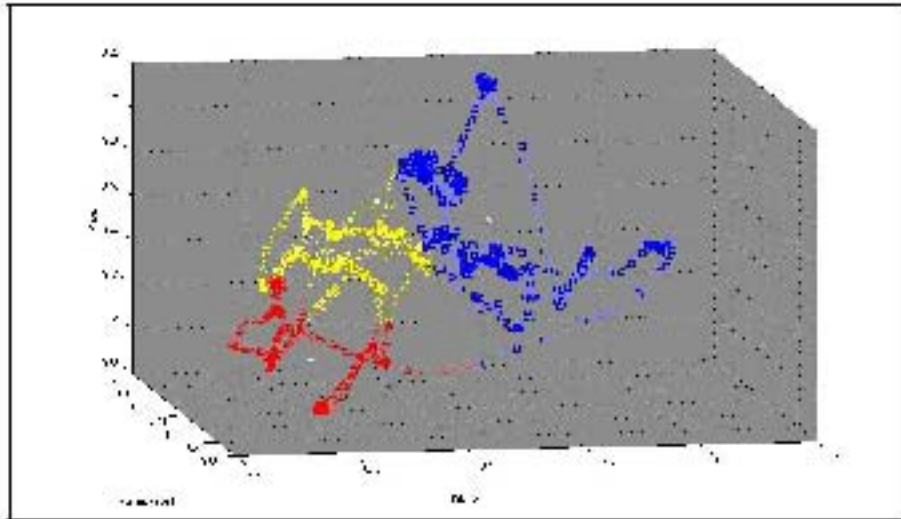
The linkage method refers to how the distance measure is applied to the clusters that are formed. K-means randomly creates  $k$  cluster centroids (where  $k$  is the number of clusters selected), and then assigns data points to the cluster, using the distance from the centroid given by the distance measure. The new cluster of two points (the randomly selected centroid and the first closest assigned point), now has a new centre of gravity, which then becomes the new cluster centroid. The linkage measure determines how the next, and consecutive, points are added. The shortest distance (where distance is measured using the distance measure chosen), as mentioned before, is the most common approach. For the purposes of this research, and considering the data type and the requirements of the clustering, it is not necessary to choose anything more complex than the shortest distance (single link method). This instructs the algorithm to search for the closest data point to the cluster centroid given the distance measure, which has been selected as the Euclidean distance.

As mentioned, the K-means algorithm starts by randomly assigning cluster centroids. Theoretically, this should not matter, because as the algorithm iterates, the centroids shift to their correct positions. However, in certain data sets the algorithm may encounter local minima, where the next shift in centroid position will result in an increase in the squared error objective function (Equation (8.1)), despite there being another configuration of centroids that would result in a lower state solution. To overcome this problem of ‘local minima’, the algorithm was carried out 10 times per analysis (10 replicates), each time selecting the solution with the lowest squared error objective value to ensure that the global minimum state was reached.

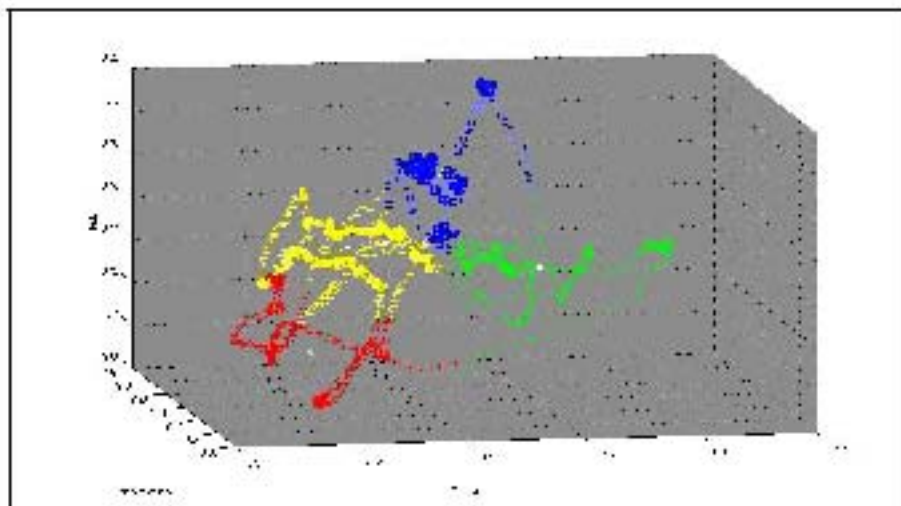
The analysis was run using the parameters described above, and selecting 3 clusters as a starting point (Figure 8.5). The clusters produced by the analysis can be classified into two sets; the blue cluster contains the points with lower private vehicle scores, and the red and yellow clusters contain points with higher private vehicle scores. Between the red and yellow clusters, the red cluster contains points with low PT and NMT scores, and the yellow cluster contains those points with moderate scores for NMT and PT. The blue cluster is much larger than the red and yellow clusters, suggesting that this cluster could be sub-divided further.

To test this, the algorithm was run again, this time using  $k = 4$  (Figure 8.6). When a fourth cluster was added, the original blue cluster split into two new clusters, consisting of points with high PT scores and high NMT scores (shown as blue squares), and points with lower PT and NMT scores (shown as green diamonds). Recall that the original blue cluster consisted of those points with lower private vehicle scores.

Although this new cluster configuration is more satisfying, the data set could, in theory, be divided into ever smaller clusters, with each additional cluster displaying ever smaller differences from its neighbours. Ideally, a cluster developed from the analysis should display enough differences in its configuration to warrant its separation from another cluster. Burnham (2004) notes: “. . . *model selection*



**Figure 8.5:** K-means clustering results for three cluster centres along Voortrekker Road. Each colour represents a different cluster. The blue cluster is much larger than the red and yellow clusters, suggesting that it could be subdivided further.



**Figure 8.6:** K-means clustering results with 4 cluster centres along Voortrekker Road. When a fourth cluster is added, the original blue cluster split into two new clusters.

*should be based on a well-justified criterion of what is the ‘best’ model, and that criterion should be based on a philosophy about models and model-based statistical inference, including the fact that the data are finite and ‘noisy’. The criterion must be estimable from the data for each fitted model, and the criterion must fit into a general statistical inference framework.”* Deciding what the correct number of clusters, that accurately reflects the structure of a data set, is a subject that has been the focus of a significant amount of study in the field of data mining (Fraley, 1998; Kothari, 1999; Milligan and Cooper, 1985).

Salvador and Chan (2004) list some of the most common methods of estimating the number of clusters in a data set. Cross-validation techniques, such as Monte Carlo cross-validation (Smyth, 1996), is used to assess how the results of an analysis will generalise to an independent data set. The data is randomly partitioned into two sub-sets, and the analysis is performed on one of these sub-sets (called a training sub-set). The results are validated by analysing how the clusters conform to the other sub-set (called the testing sub-set). To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the rounds.

An information theoretic statistic can also be used as a guide to model selection. Common statistics used in cluster analysis are the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC), sometimes called the Swartz Criterion (De Smith et al., 2007). These statistics penalise models with additional clusters by adding a term that is dependent on the number of clusters to the log-likelihood function. The AIC is defined as:

$$AIC = -2 \ln(L) + 2k \quad (8.3)$$

where  $k$  is the number of clusters used in the model and  $L$  is the likelihood function. The BIC, which places a greater weight on the number of parameters used, is defined as:

$$BIC = -2 \ln(L) + k \ln(n) \quad (8.4)$$

where  $n$  is the sample size,  $k$  is the number of clusters used in the model, and  $L$  is the likelihood function.

Locating the ‘knee’ of an error curve, in order to determine an appropriate number of clusters or segments, is another method that is commonly mentioned in the literature (Bezdek and Pal, 1998; Jain et al., 1999; Salvador and Chan, 2004; Vasko and Toivonen, 2002). The knee of a curve can be defined as the point of maximum curvature on the curve. The knee in a graph of the number of clusters versus an evaluation metric<sup>30</sup> can be used to determine the ‘correct’ number of clusters, since the mean squared error decreases monotonically as a function of  $k$ .

The K-means algorithm was run on the Voortrekker Road data set for all  $k$  between 2 and 100. The total sum of distances was plotted against the number of clusters,  $k$ . The curve given by  $y = 6.4454x^{-1.222}$  was fitted to the points, with the square of the correlation coefficient,  $R^2$ , being 0.9985. The curvature,  $\kappa$  of this fitted curve is given by:

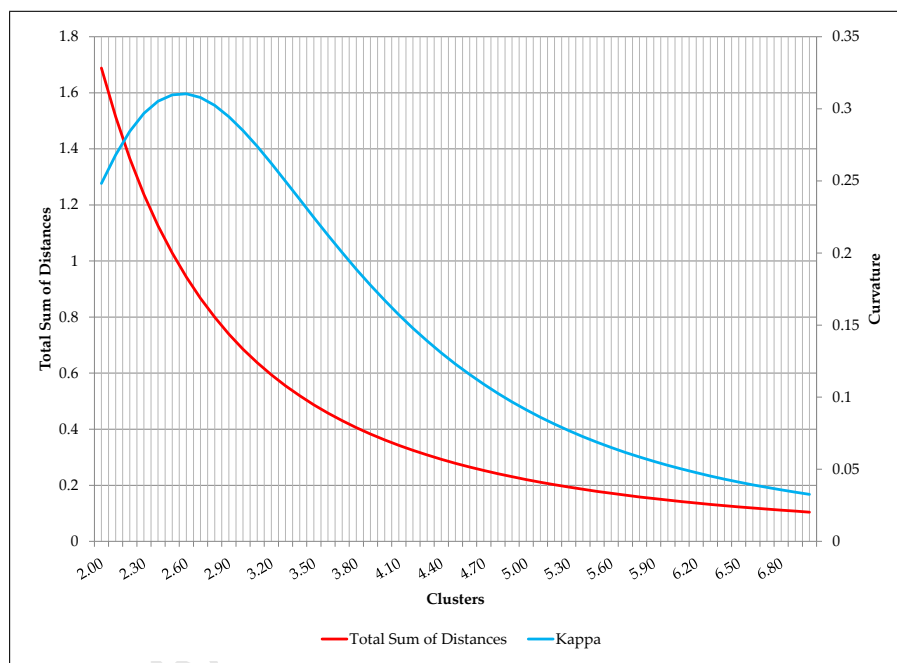
$$\kappa = \frac{|y''|}{(1 + y'^2)^{3/2}} \quad (8.5)$$

<sup>30</sup> Common metrics include the largest magnitude difference between two points, the largest ratio difference between two points or the first data point with a second derivative above some threshold value

which, given that  $y = 6.4454x^{-1.222}$ , is:

$$= \frac{17.5011}{\left(\frac{62.0358}{x^{4.444}} + 1\right)^{3/2}} \times x^{3.222} \quad (8.6)$$

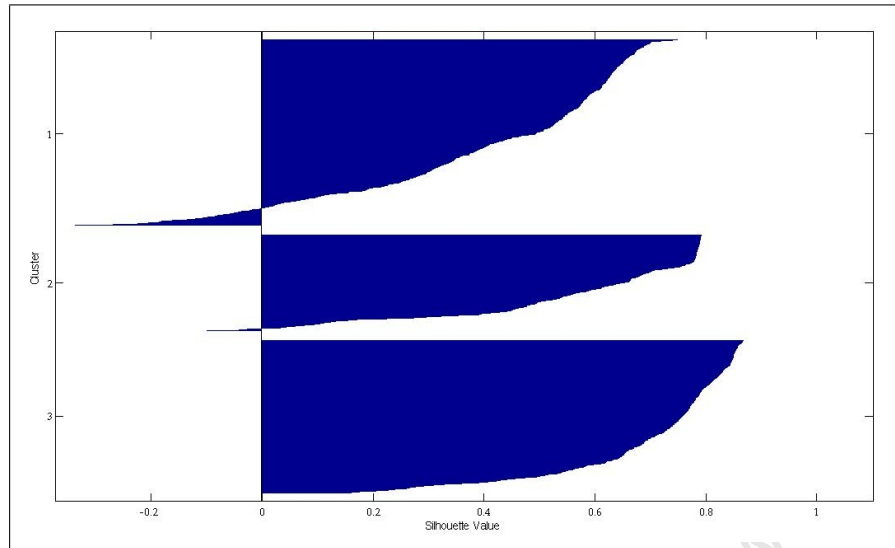
According to the definition, the knee is loosely defined being the point of maximum curvature of the curve, which occurs nearest  $k = 3$ , indicating that the ideal number of clusters is three (Figure 8.7). However, Salvador and Chan (2004) contend that, since the fitted curve is a smooth line, the knee of the curve is ambiguous, suggesting that the ideal number of clusters is, in fact, a range of values.



**Figure 8.7:** Total Sum of Distances Versus Curvature. The point of maximum curvature is nearest  $k = 3$ , suggesting that the ideal number of clusters is 3. However, since the curve  $\kappa$  is a smooth line that transitions evenly between large decreases in the sum of cluster distances and negligible decreases in the sum of cluster distances, the ideal number of clusters is, in fact, a range of values.

An alternative approach is to measure how close each point in one cluster is to points in the neighbouring clusters. This data can be represented in a silhouette plot, which plots the ratio of the average dissimilarity of a point to all other objects in its cluster, and the average dissimilarity of that point to all objects in the closest other cluster (Figure 8.8).

Assessing the silhouette plot for  $k = 3$ , there are quite a few points that have probably been incorrectly assigned to cluster 1, which also contains a large number of points with silhouette values nearer to 0 than to 1. These points are, therefore, not distinctly separate from clusters 2 and 3. Clusters 2 and 3 are better formed, although they both still contain a large number of points that are not distinctly separate from the other clusters. The mean silhouette value for all the clusters in the plot is an indication of the overall strength of the clustering. For this three cluster scenario the mean silhouette score is 0.5477, which is an indication that, overall, the clusters are not very distinct from each other.



**Figure 8.8:** Silhouette Plot for Three-Cluster Scenario for Voortrekker Road. The values plotted on the x-axis ranges from +1, indicating points that are very distant from neighbouring clusters, through 0, indicating points that are not distinctly in one cluster or another, to -1, indicating points that are probably assigned to the wrong cluster. The clusters themselves are shown along the y-axis.

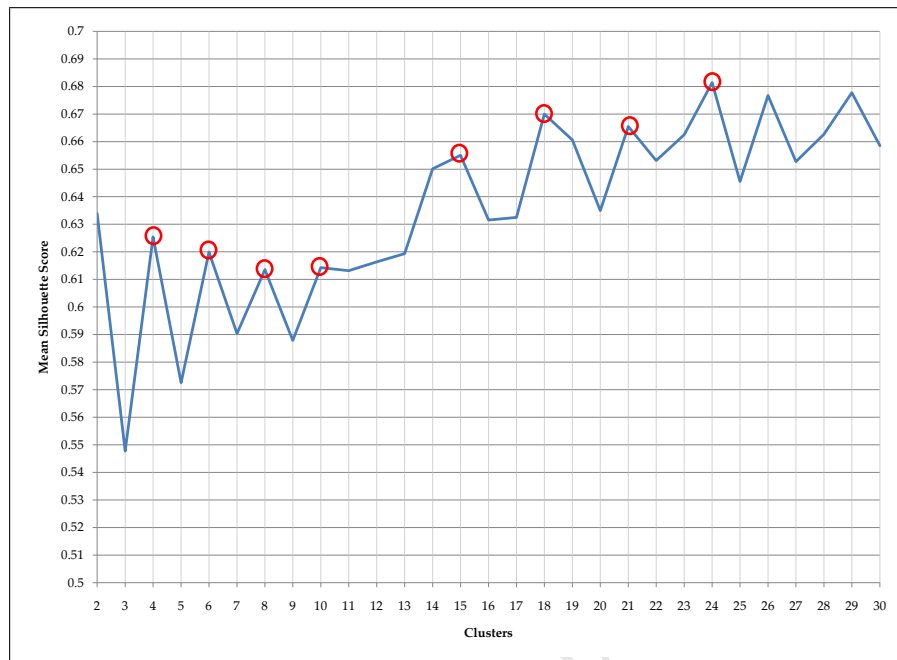
Increasing  $k$  to four produces an improvement over the three-cluster scenario (Figure F.4). This is confirmed by the mean silhouette score for this plot, which is calculated as 0.6254. However, there is still the question of how many clusters is the correct number. The K-means algorithm was run for 2 through 100 clusters, and the mean silhouette score was calculated for each iteration. The mean silhouette scores for each  $k$  was plotted against the number of clusters (Figure 8.9).

The mean silhouette scores reach a maximum at 24 clusters, indicating that the  $k = 24$  produces the most distinctive set of clusters. Although the clusters produced when setting  $k$  to 24 are, in this sense, well defined, the 24 clusters produced are not meaningfully distinct from each other. The smallest cluster contains only 36 points (equivalent to 360 m on the road), and the largest cluster only 134 points (or 1 340 m on the road). The differences between clusters becomes less significant the larger  $k$  becomes.

Peak mean silhouette values for  $2 \leq k \leq 24$  occur at 4, 6, 8, 10, 15, 18 and 21 clusters (highlighted with red rings). Any of these  $k$  will also produce a set of well defined cluster centres. The choice of which of these to select depends on the setting of a minimum threshold range for the distance between cluster centres (or the least acceptable difference between adjacent clusters). Table 8.1 shows the mean distance to the nearest centroid for all clusters in the data set for those  $k \leq 24$  with peak mean silhouette values. As expected, the distance to the closest centroid decreases as the number of clusters increases. This indicates that adjacent clusters become more similar as the number of clusters increases.

This threshold is best defined in terms of the range of scores along the road. Assuming a suitable threshold range of between 15% and 20% of the range of scores along Voortrekker Road (0.232), indicates that either 6, 8 or 10 clusters would best represent the natural pattern within the data set for Voortrekker Road.

To investigate this further, an agglomerative hierarchical clustering of the data set was done. The distance measure used was the Euclidean distance and the average linkage method was used to link

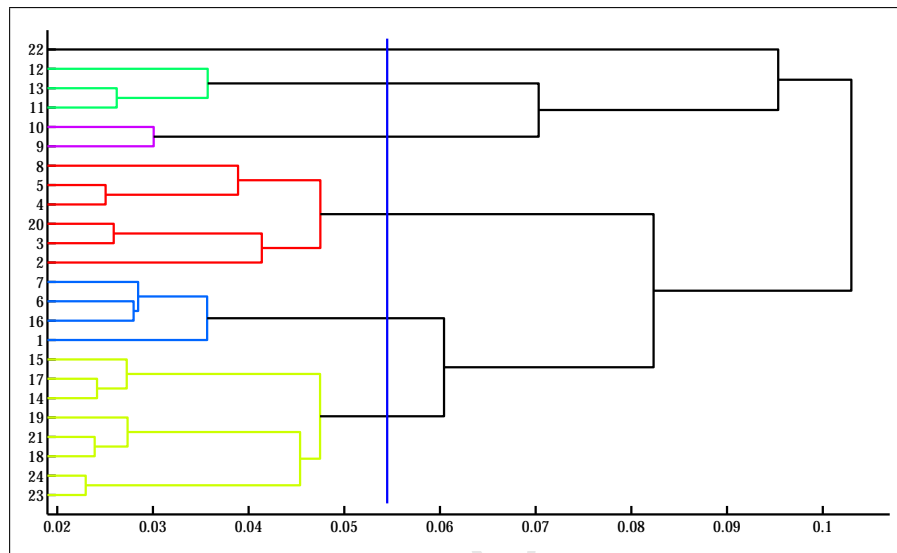


**Figure 8.9:** Mean Silhouette Score versus Number of Clusters (first 30 clusters shown).  $k = 24$  produces the most distinctive set of clusters for the Voortrekker Road data set.

**Table 8.1:** Mean distance to closest centroid for increasing values of  $k$  - Voortrekker Road. Suitable values are found for  $6 \leq k \leq 10$ .

Number of Clusters	Mean distance from centroid to closest centroid
4	0.0561
6	0.0452
8	0.0418
10	0.0363
15	0.0290
18	0.0280
21	0.0250
24	0.0223
<b>Range</b>	0.2320
<b>15% of Range</b>	0.0348
<b>20% of Range</b>	0.0464

clusters. Hierarchical clustering provides a useful perspective on the relationships between clusters in the data set, helping to reveal the structure of the clusters in the data. Figure 8.10 shows a dendrogram of the data for Voortrekker Road clustered using agglomerative hierarchical clustering using the squared Euclidean distance measure and average linkage measure. The dendrogram has been clipped at 24 clusters, since the mean silhouette scores show that cluster definition tends to decrease above this value. All partitions below this threshold are, therefore, collapsed into the 24 lowest groupings.



**Figure 8.10:** Dendrogram of Clusters along Voortrekker Road. The dendrogram shows that the correct number of clusters for the Voortrekker Road data set is 6.

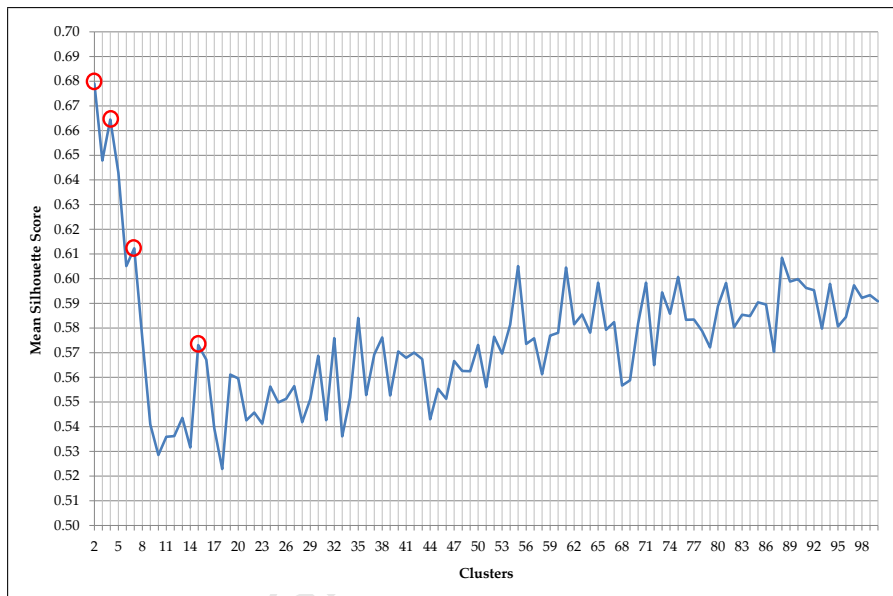
The dendrogram is interpreted by comparing the length of the horizontal ‘heights’ and looking for sustained ‘gaps’ between vertical links. The longer the height between links, the greater the dissimilarity between those clusters. From the figure it becomes clear that the correct number of clusters for the Voortrekker Road data set is 6. A cut-off line has been added on the 6 cluster solution. The 3 and 4 cluster solutions have already been ruled out for being too aggregated, as has the 5 cluster solution for producing a low mean silhouette score, and the 8 and 10 cluster solutions simply do not present a significant enough level of dissimilarity over that of the 6 cluster solution.

The procedure used to determine the correct number of clusters can, thus, be summarised as follows:

1. Merge highly correlated modal scores (s.t.  $\overline{NMT} = (Pedestrian + Bicycle)/2$ ;  $\overline{PMT} = (Car + Freight)/2$ ).
2. Compute K-means clustering mean silhouette scores for  $2 \leq k \leq 100$  clusters (K-means distance measure = Squared Euclidean, 10 replicates).
3. Identify maximum mean silhouette score, and local peak scores of lower cluster solutions.
4. For peak mean silhouette scores, compute the mean distance between adjacent centroids for all clusters.
5. Identify cluster solutions that satisfy the 15 - 20% of score range for case study threshold criterion.
6. Compute agglomerative hierarchical clustering (distance measure = Euclidean, average linkage method).

7. Evaluate dissimilarity for cluster levels that satisfy the threshold criterion.
8. Select  $k$  with maximum dissimilarity that satisfies the threshold criterion.

The procedure outlined above was repeated for the other case study data sets. The mean silhouette scores plotted against the number of clusters for the Lansdowne Road data set (Figure 8.11) shows that peak mean silhouette scores for this data set are located at 2, 4, 7 and 15 clusters. The next best peak values occur at 32 and 35 clusters, but it is unlikely that this many clusters would be meaningfully distinct from each other. Table 8.2 shows the mean distance to the nearest centroid for each cluster scenario for the Lansdowne Road data set. The data shows that the most suitable range of clusters is found for  $6 \leq k \leq 8$ . Considering the peak mean silhouette scores in Figure 8.11, this suggests that the ideal number of clusters for this data set is 7.



**Figure 8.11:** Mean Silhouette Score versus Number of Clusters - Lansdowne Road. Peak mean silhouette scores for this data set are found at 2, 4, 7 and 15 clusters

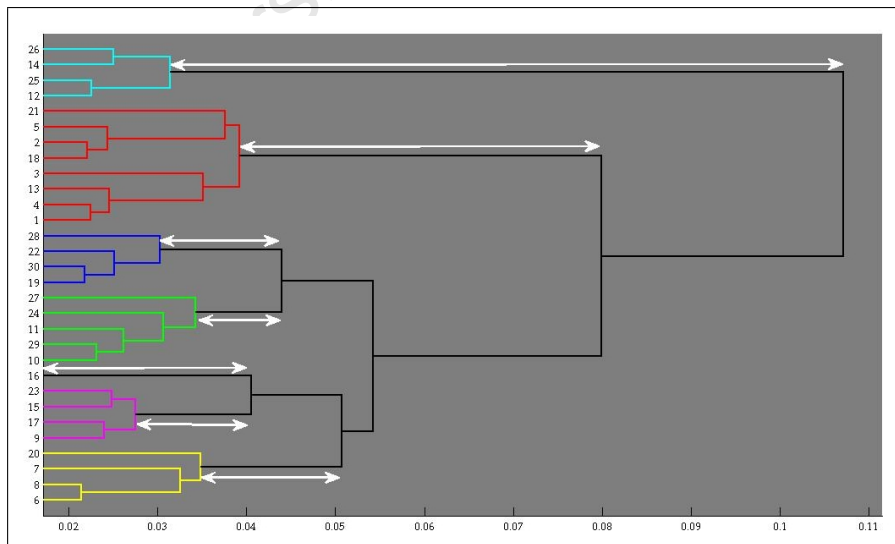
An agglomerative hierarchical clustering (distance measure = Euclidean, average linkage method) was conducted on the data set for Lansdowne Road. An assessment of the dendrogram for the Lansdowne Road data set (shown in Figure 8.12) confirms that the 7-cluster solution provides good differentiation between clusters. The white arrows show the dissimilarity height for the clusters, each of which are shown in a different colour.

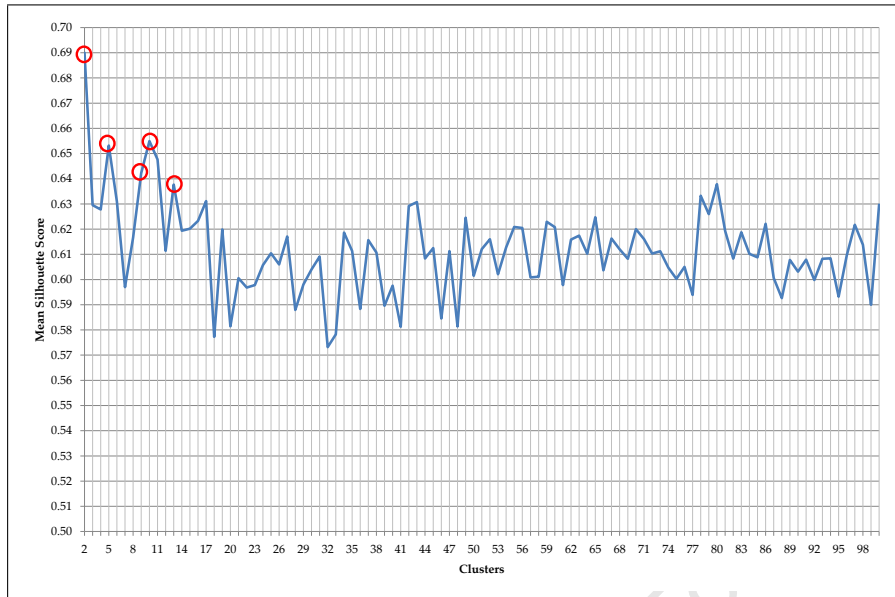
The same procedure was conducted for the Koeberg Road data set. The mean silhouette score plotted against the number of clusters for the Koeberg Road data set (Figure 8.13) shows that, in contrast to Lansdowne Road and Voortrekker Road, the lower  $k$  numbers perform better overall than the larger  $k$  numbers. As with the Lansdowne Road data set, the peak score is located at  $k = 2$ , but, it is unlikely that this aggregate clustering explains the variation in the data set satisfactorily. Instead, there are additional peak mean silhouette scores located at 5, 10 and 13 clusters. The next peak, at  $k = 17$ , is less than a peak further out at  $k = 80$ , and so is disregarded.

As with the previous case study roads, identifying which of these  $k$  numbers represents a reasonable representation of the data structure requires an analysis of the distances between clusters. Calculating the mean distances to the nearest centroids for all cluster centroids for all  $2 \leq k \leq 13$  shows that

**Table 8.2:** Mean Distance to Closest Centroid - Lansdowne Road. Suitable  $k$  are found when  $6 \leq k \leq 8$ 

Number of Clusters	Mean distance from centroid to closest centroid
2	0.0887
3	0.0693
4	0.0568
5	0.0493
6	0.0431
7	0.0405
8	0.0358
9	0.0315
10	0.0301
11	0.0285
12	0.0279
13	0.0282
14	0.0274
15	0.0268
<b>Range</b>	0.2360
<b>15% of Range</b>	0.0354
<b>20% of Range</b>	0.0472

**Figure 8.12:** Dendrogram of Clusters along Lansdowne Road. The dendrogram confirms that a 7-cluster solution provides good differentiation between clusters.



**Figure 8.13:** Mean Silhouette Score versus Number of Clusters - Koeberg Road. The peak scores are located at  $k = 2, 5, 10$  and  $13$  clusters. The next peak, at  $k = 17$ , is less than a peak further out at  $k = 80$ , and so is disregarded.

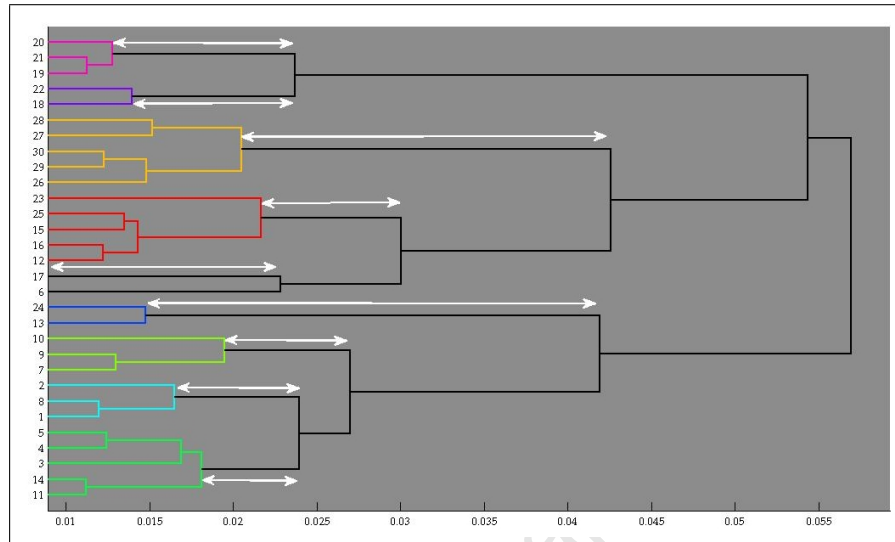
$7 \leq k \leq 10$  satisfies the threshold criteria. Since only  $k = 10$  falls within this range and is also a local peak in Figure 8.13, this must be the most suitable  $k$  for this data set.

**Table 8.3:** Mean Distance to Closest Centroid - Koeberg Road. Suitable  $k$  are found when  $7 \leq k \leq 10$ .

Number of Clusters	Mean distance from centroid to closest centroid
2	0.0526
3	0.0409
4	0.0371
5	0.0354
6	0.0340
7	0.0304
8	0.0259
9	0.0261
10	0.0251
11	0.0225
12	0.0212
13	0.0200
<b>Range</b>	0.1590
<b>15% of Range</b>	0.0239
<b>20% of Range</b>	0.0318

It is somewhat surprising that such a high  $k$ -value was identified for this data set considering that the

suitability scores along Koeberg Road (Figure 7.10) appeared to be the most uniform of all the case study roads. An agglomerative hierarchical clustering (distance measure = Euclidean, average linkage method) was conducted on the data set for Koeberg Road. The dendrogram for a 10 cluster solution along Koeberg Road (Figure 8.14) shows that, while a 5 cluster solution appears strong, the cluster dissimilarities for a 10 cluster solution are large, overall, hence the strong performance of this option.



**Figure 8.14:** Dendrogram of Clusters along Koeberg Road. The dendrogram for a 10 cluster solution shows that while a 5 cluster solution appears strong, the cluster dissimilarities for a 10 cluster solution are large overall.

To investigate these differences further, the cluster plots for both  $k = 5$  and  $k = 10$  were generated (Figure F.10). The cyan and pink clusters in the 5 cluster solution both span a very large range of scores, and these are split into 2 and 3 clusters respectively in the 10 cluster solution. The green cluster, on the other hand, appears well formed in the 5 cluster solution, but this is also split into 3 clusters in the 10 cluster solution. This suggests that an intermediate  $k$  value would produce a better clustering for this dataset. Considering both the mean silhouette scores and the threshold criteria, the remaining viable solutions are either  $k = 8$  or  $k = 9$ .

To investigate these solutions, the hierarchical clustering for both solutions was calculated, and the resulting dendrograms plotted (Figure F.11). As expected, the dendrogram shows how the 9 cluster solution is agglomerated to an 8 cluster solution by the linkage of the top 2 clusters (pink and purple) into a single cluster (red). Notice also that the black clusters from the 10 cluster solution dendrogram (Figure 8.14) has been linked in the 9 cluster solution. Given all of the above, the 8 cluster solution for Koeberg Road is a reasonable compromise and offers a good balance between cluster dissimilarity and information loss.

Having established the appropriate number of clusters for each data set, the K-means algorithm was run using these  $k$ -numbers on the original 5-dimensional data for each case study road. Since the complexity of the data set is now increased, a slight drop in the quality of the clustering was expected. The mean silhouette scores for the merged and original data was compared to ensure the drop in cluster quality was not greater than 10%. None of the case study roads exhibited a drop in the mean silhouette scores of greater than 10% (Table 8.4). The  $k$ -values identified in the cluster analysis for the merged data set were, therefore, accepted as being applicable to the original data sets as well.

**Table 8.4:** Comparison of silhouette score means for merged data sets and original data sets. Since none of the case study roads exhibited a drop in the mean silhouette scores of greater than 10%, the  $k$ -values identified in the cluster analysis for the merged data set were accepted as being applicable to the original data sets as well.

Case Study Route	Number Of Clusters	Mean Silhouette Score		Percentage Drop
		Merged Data Set	Original Data Set	
Voortrekker Road	6	0.6199	0.5756	7.15%
Lansdowne Road	7	0.6123	0.6089	0.55%
Koeberg Road	8	0.6166	0.6092	1.20%

The cluster centroids produced from each case study road are shown in Table 8.5.

**Table 8.5:** Cluster centroid details for case study roads

Case Study Route	Cluster	Bike	Car	Freight	Pedestrian	Public Transport
Voortrekker Road	1	0.4294	0.4305	0.3985	0.4436	0.4512
	2	0.4289	0.4646	0.4298	0.4644	0.4436
	3	0.4184	0.3817	0.3568	0.4564	0.4434
	4	0.4205	0.3301	0.3030	0.4350	0.4479
	5	0.3950	0.4611	0.4502	0.4178	0.4112
	6	0.4525	0.3936	0.3652	0.4824	0.4824
Lansdowne Road	1	0.4535	0.3498	0.3079	0.4406	0.4551
	2	0.4289	0.3816	0.3559	0.4331	0.4321
	3	0.4477	0.3943	0.3740	0.4668	0.4656
	4	0.4692	0.3684	0.3307	0.4694	0.4760
	5	0.4610	0.2994	0.2748	0.4789	0.4601
	6	0.4251	0.4251	0.4041	0.4439	0.4328
	7	0.4171	0.4575	0.4300	0.4188	0.4253
Koeberg Road	1	0.4022	0.4385	0.4349	0.4354	0.4120
	2	0.4205	0.4206	0.3966	0.4482	0.4256
	3	0.4115	0.4019	0.3885	0.4356	0.4094
	4	0.4358	0.4086	0.3888	0.4600	0.4355
	5	0.3952	0.4446	0.4302	0.4069	0.4055
	6	0.4257	0.4477	0.4146	0.4479	0.4540
	7	0.4126	0.3672	0.3447	0.4365	0.4088
	8	0.3903	0.4644	0.4825	0.4314	0.4164

### 8.3 Résumé

This chapter outlined the procedures used to cluster the data generated in Chapter 7. Clustering the data will assist with ascertaining the underlying structure of the data sets in order to gain insight into the data, evaluate the hypothesis, detect anomalies, and identify salient features. It will also help to establish a natural classification by establishing the degree of similarity between sections of the case study roads. It was used to simplify the data by compressing the original data into clusters that can be used to organise the data into route typologies.

The two most popular approaches to clustering are partitional and hierarchical. Partitional algorithms construct partitions in the data such that each cluster optimises a pre-stated clustering criterion. Hierarchical algorithms create a hierarchy amongst the sets of data points. Hierarchical algorithms can be either agglomerative or divisive. Agglomerative algorithms start with each object being a separate cluster itself, and successively merges groups according to a distance measure.

Although a large range of clustering techniques were identified, K-means clustering was used, since it is a well established method, and is highly customisable, and able to produce reliable results very quickly. The clusters produced were verified against the results of an agglomerative hierarchical clustering, using a range of techniques to identify the most appropriate number of clusters for each dataset.

An initial investigation of the data revealed that the modes could be split into three groups according to their suitability scores along the route; car and freight, and pedestrian and bike obtained similar scores, and could be combined into mode groups. PT formed the third group. These similarities amongst the modes are problematic when clustering the data since a strong clustering relies on there being distinct differences in the elements being clustered. The less differences there are, the less distinct are the clusters that are formed. The data was, therefore, first merged into three groups to identify the appropriate number of clusters for each case study. These cluster numbers were then used to cluster the original data set.

The clustering techniques used produced 6 clusters along Voortrekker Road, 7 clusters along Lansdowne Road, and 8 clusters along Koeberg Road.

## Chapter 9

# Infrastructure and Context

In Chapter 8, the SMCA outputs were clustered to identify the underlying context types along each case study road. The main hypothesis of this thesis is that a locations context can and should inform the planning of road infrastructure provided to serve that location. The context groups, or clusters, developed in Chapter 8, are differentiated from each other by variations in mode rankings and variations in mode suitability scores. In this chapter, these modal priority rankings and scores are used to assess infrastructural priorities at any location along the route.

### 9.1 Analysis of Route Clusters

The cluster means for each cluster within the data set describes the contextual characteristics of each section of the route in a much more compact form than can be achieved by simply analysing the raw SMCA data. There is, of course, some measure of information loss intrinsic to substituting the point scores with the cluster mean scores, since these will invariably differ. But one of the major benefits of clustering, is that it makes the contextual assessment of the route more manageable, whilst minimising information loss. Identifying clusters simplifies further analysis of the data in that much of the complexity, or 'noise', within the data set is eliminated.

As such, it is useful to compare the cluster means to the original SMCA scores. The cluster means are the 'coordinates' of the cluster centroid, the average suitability score for each mode in the cluster. The combination of those means, therefore, represents the average context for a given stretch of road, interpreted in terms of modal suitability. When compared to the original SMCA scores, the cluster means can be seen to be a good proxy for the original SMCA scores (Figures 9.1, G.3 and G.4).

Along Voortrekker Road, for the majority of the route, PT and pedestrian receive the highest priority. For the first half of the route, the private car is an important mode, although nearly always secondary to PT. Where modes scores are closely grouped, as in cluster 2, those modes can be said to be of similar suitability. Evaluating Figures 9.1 and 9.2 together, shows that clusters 1, 2, 3 and 6 are most common, and clusters 4 and 5 occur only in isolated areas. These areas are, therefore, somewhat contextually unique along the route, and may warrant more careful investigation to understand the specific needs here. In particular, whereas the steps between mode scores are, generally, incremental, between kilometre 7.0 and 8.0 there is a sharp reversal in priorities, suggesting that there is some kind of disjuncture in the context at this location, requiring careful consideration. Considering the remaining case studies, it is clear that the pedestrian, PT and, to a certain extent, the bicycle have received a high level of priority, with the car and freight modes only receiving preference at specific locations.



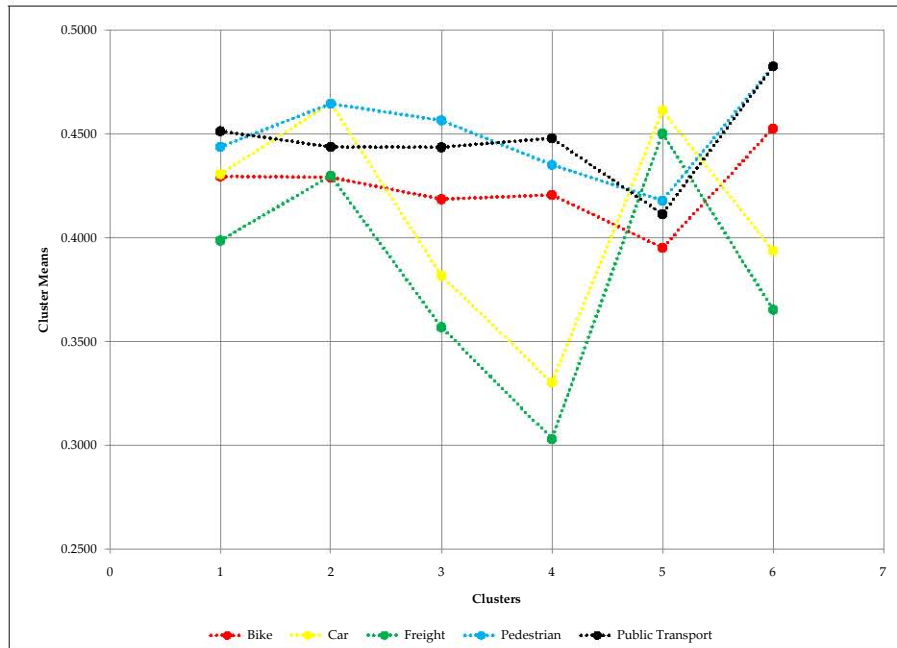
**Figure 9.1:** Comparison of Cluster Means and SMCA Scores for Voortrekker Road

Clustering the data also identifies sections of the route that are contextually similar. Clusters can be seen to be repeated within all of the data sets, which shows that there are contextually similar areas along the route that are spatially separate from each other. In the Voortrekker Road data set, cluster 5 is repeated at 3 points along the route, between kilometre 1.0 and 1.5, between kilometre 6.0 and 7.5 and again between kilometre 13.5 and 14.5. Similar repetition can be identified in the other case study data sets.

Clusters range from approximately 500 m to 3 000 m in length, and cluster means tend to change in a stepwise fashion. That is, cluster means do not change erratically, instead they increase and decrease in increments. This indicates that contexts also tend to change incrementally. This is significant, since continuity is an important consideration in road design. Efficient and safe route operations for all modes are dependent upon the predictability and continuity of the infrastructure.

Clusters are distinguished from each other by variations in rank and variation in score range (Figures 9.2, G.5 and G.6). These characteristics of clusters can form the basis of a description of each cluster in terms of its suitability for each mode. The clusters themselves, although a simplification of the SMCA outputs, are still representative of the values imposed on the analysis by the investigator at the early stages of the assessment. These values are, therefore, also represented in the descriptions of the clusters. Since these results were generated from an equally weighted assessment, they are only influenced by the standardisation framework used in the SMCA. It is entirely plausible, then, that an alternative standardisation approach would yield somewhat differing results. Similarly, the application of a weighting scheme would also influence the results. The sensitivity to weighting was explored in some detail in Section 7.4. Weighting was found to significantly affect the results of the analysis. However, since the weight schemes are developed from the values used in standardisation framework, weighting can only serve to reinforce or weaken the effect of these values on the analysis.

The cluster means are, therefore, just a simplified classification of the priority mixes generated given



**Figure 9.2:** Comparison of Cluster Means Voortrekker Road

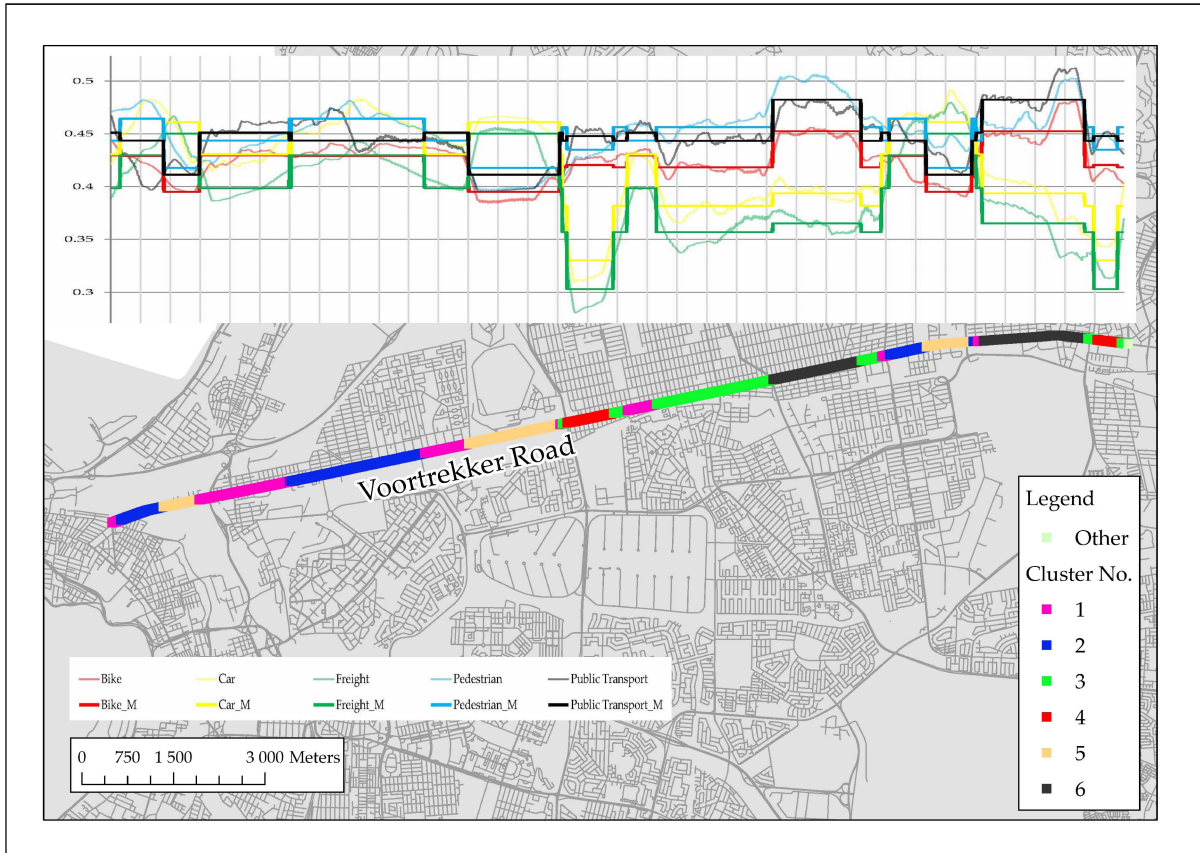
the standardisation framework employed, and the weighting scheme applied. Considering that both the standardisation framework and the weighting scheme are derived from the values imposed on the assessment, the clusters are, therefore, a perspective on these values interpreted in terms of modal priority. The infrastructure recommendations made using the cluster information will, therefore, also reflect the original assessment values.

## 9.2 Infrastructure Recommendations from Cluster Analysis

Using the rank order of the modes, the groupings of modes within the clusters, the score differences between individual modes or mode groupings and the range of scores within each group, it is possible to develop descriptive statements regarding the priority of each mode at each point along the route, with reference to the contextual setting for that location. Mapping the clusters along the route gives some perspective on how they relate to the surrounding infrastructure. In Figure 9.3 the clusters are mapped along Voortrekker Road, and the cluster means have been added to show how the individual mode priorities vary along the route.

The first half of the route is dominated by clusters 1, 2 and 5; those clusters where car and freight perform relatively well. Referring back to Figure 6.5, this corresponds to those sections of the road with higher levels of industrial activity. The second half of the route, with some exceptions, is dominated by clusters 3, 4 and 6, where car and freight perform poorly in relation to the other modes, corresponding to areas with high levels of residential, mixed use and commercial activities. It is tempting to think of these areas only in terms of their land uses, and attribute the mode suitabilities primarily to these. However, mode suitabilities are a function of a range of contextual inputs, and all of these inputs, having been equally weighted in this analysis, have contributed equally to the results produced here.

The method used to develop infrastructure recommendations relies on these descriptive statements of the mode suitability. As such, a formal mechanism is required to translate mode suitability into a statement



**Figure 9.3:** Comparison of Cluster Means and Cluster Location on Voortrekker Road

that can be interpreted in terms of infrastructure requirements. One of the simplest mechanisms is that of rank. If one mode is ranked higher than another then, by definition, it should receive higher priority, since it is better suited to that location. A second approach is that of score range or, specifically, relative score. Where one mode scores much higher than another, it should receive much higher priority than if it only scored slightly better than the other mode.

These mechanisms all rely on a comparison of cluster means to define mode priority and, consequently, infrastructure provision. To effect such a comparison, a standardised scale is required, against which cluster means can be ranked. Since cluster means are already scored on a scale between 1 and 0, an interval scale between these values is naturally suited to compare the cluster means. However, since the range of scores in all of the case studies is rather limited, it is necessary to rescale the cluster means to amplify their differences. To this end, a route maximum and minimum score is defined as being the highest and lowest score obtained by any mode, in any cluster along the route. The cluster means are mapped to a linear scale, with the route maximum and minimum being rescaled to 1 and 0 respectively, using the transformation in Equation (9.1):

$$x'_j = \frac{x_j - x_{jmin}}{x_{jmax} - x_{jmin}} \quad (9.1)$$

Where  $x_j$  is the cluster mean for mode  $j$ ,  $x'_j$  is the rescaled cluster mean value for mode  $j$ ,  $x_{jmax}$  is the route maximum score for mode  $j$ , and  $x_{jmin}$  is the route minimum score for mode  $j$ . The

transformation is linear, thereby retaining the proportional differences of the mode scores within each cluster, as well as between clusters.

The scale lends itself to a stratification of mode suitability. This stratification forms the basis for comparing mode operations, in any cluster. In turn, describing the operational characteristics for each mode in any given cluster forms the basis for identifying the types of infrastructure that best suits a location, given the context as identified by the analysis. The approach is somewhat similar to the traditional interpretation of the LOS, which, at its core, is also a descriptive statement of the operational conditions along a piece of infrastructure. Improving the LOS along any particular route, generally, has infrastructural implications, since LOS is a function of the infrastructure provided to accommodate traffic, and the characteristics of the traffic itself. To affect the LOS, one of these factors must be changed.

Similarly, suitability rankings have implications in terms of infrastructure provision, since a mode that is better suited to a particular location, in terms of that location's context, should be given better operational conditions than other modes. The question, then, centres around how best to describe the operational conditions for a mode in terms of its suitability ranking. The approach adopted was to reinterpret mode suitability in terms of three parameters: access to the segment, right of way or ease of movement within the segment and the level of independence afforded to the mode within the segment. This interpretation of mode priority develops a description of the level of operation for each mode, at all sections of the route (Table 9.1).

**Table 9.1:** Mode suitability in terms of operational permissions

<b>Suitability</b>	<b>Access</b>	<b>Priority</b>	<b>Independence</b>
Unsuitable	Restricted - physically prevented from accessing the road	Not applicable	None provided
Low	Access allowed, but movements physically restricted	Lowest priority afforded	Shared infrastructure, with minimal dedicated mode specific features
Medium	Partial access levels provided	Priority given according to need and in subservience to main mode	All needs catered for, mixture of shared and dedicated infrastructure
Highest	Highest access allowed	Priority of movement given for all circumstances	Dedicated infrastructure wherever practical, minimal interaction with other modes

## 9.3 Operational characteristics and Infrastructure

### 9.3.1 Access

In terms of the operational characteristics of a mode, in a particular section of the route, the defining characteristic of that section is to what extent a particular mode is afforded access to it. Modes are restricted access to certain areas for a range of reasons. In certain circumstances, it may be undesirable or unsafe for a mode to operate in an area at any time, or only at certain times of the day. Restrictions

such as this are common around the world. Obvious examples may include restrictions placed on pedestrian activity on high speed roads, or time of day restrictions on freight or delivery vehicles in residential areas. Some roads are designated for pedestrians only, others may allow for both pedestrians and cyclists.

Restrictions such as this also apply to the operations of modes along routes. In many areas, vehicles are allowed to access the area, but may not stop or park. Similarly, pedestrians are often afforded access to a route, but their crossing opportunities are limited to only certain controlled locations. Although access to an area is permitted, the mode is not allowed to operate unencumbered. A combination of infrastructural and regulatory measures are put in place to enforce these restrictions.

The location's context is often the deciding factor when determining where, when and to what extent a mode is afforded access to an area. Aspects such as the land uses and the most common activities associated with them, the efficiency of the area from a traffic operations standpoint, and interactions with other modes, commonly influence these decisions. In this sense, the use of the context, as defined in this research, is well suited as a basis for describing the level of access afforded to a mode at any location.

Modes that, in terms of the analysis, are poorly suited to the context at a location, should most probably be excluded from accessing that location. Conversely, modes that are well suited to the context at a location, should be afforded unencumbered access. Infrastructure should be planned so as to facilitate and enforce these various levels of restriction. There are a range of measures that could be used to enforce such restrictions.

Restrictions, as discussed, can be full or partial, and can affect many aspects of the operations of a mode. For motorised modes, speed restrictions are ubiquitous, and there are well documented measures that can be used to allow access to a location for a mode, but only at restricted speeds (see Section 3.4). Similarly, parking restrictions, and time of day based access restrictions, are common as well. Other measures include access control booms and bollards, that allow only certain types of vehicles, or certain drivers (such as residents), to access an area.

For pedestrians and cyclists, median barriers or fencing are commonly used to discourage midblock, uncontrolled crossing. Of course, not providing infrastructure such as sidewalks or cycling lanes, and prohibiting NMT modes from accessing an area, can also discourage access in areas where it is unsafe for these modes. Similarly, PT access to an area can be controlled by not allowing vehicles to stop for passengers, or limiting vehicle speeds. Of course, PT vehicles could also be prohibited from an area depending upon the routing.

### 9.3.2 Right of Way

The other defining characteristic of a route is the right of way, or priority of movement, allowed for the mode. Right of way is an aspect of traffic control that can have dramatic implications for the operations of a section of road. The determination of the right of way is important at conflict points along the route, since practice generally dictates that modes should be kept separate from each other. As such, at these conflict points, time needs to be allocated for each mode to use the space, in as efficient a manner as possible. Typically, right of way is determined by regulatory infrastructure, such as road markings, signage or traffic control devices, but these may be supported by infrastructure, such as vertical grade changes (as in the case of raised pedestrian crossings), or special coloured surfacing, to reinforce the right of way hierarchy.

If a mode is better suited to a location in terms of the contextual analysis and, consequently, receives

a higher priority in that location, this implies that it should be given priority of movement, or the right of way in that location, wherever possible. This dictate should, naturally, be tempered by safety considerations. But there are a range of methods that can be used to implement or enforce the right of way without sacrificing safety.

Traffic control devices, such as stop or yield signs, appropriately placed could enforce, from both a legal and a traffic control perspective, the right of way of the preferred mode. Remotely activated, or actuated, traffic signalling, as is sometimes used on Bus Rapid Transit (BRT) routes, are used to afford priority at signalised crossings. Whole sections of road can be redesigned to afford priority to a certain mode. Examples of such initiatives are woonerven in the Netherlands, and 30 km/h zones in Germany. These areas are noted for being specifically designed to create the impression that the automobile is the 'guest', and that the primary mode in these areas is NMT. A range of measures, going beyond simple traffic calming, is employed to create this impression, and enforce the modal hierarchy (see Figure B.1).

### 9.3.3 Independence of Operations

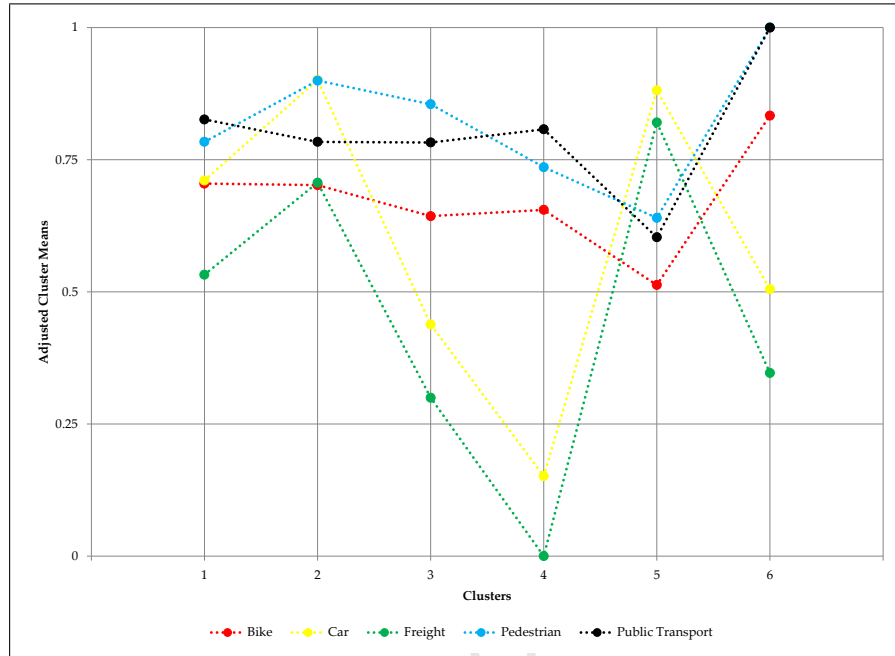
One of the guiding principles of multimodal road planning is that, as far as possible, it is best to keep modes separate from each other. To achieve this, separate infrastructure must be provided for all modes operating in the corridor. Although this is ideal, given constraints of space and cost, it is not always possible in practice, and may not always be desirable, given the operational aspects of that section of the route. The question of space allocation, of course, falls away, when space and cost are not constraints. However, these are often limiting factors, and allowing a mode to operate independently from the others in the corridor does afford that mode significant benefits. These benefits cannot be realised if the mode shares space with other modes. Space allocation which, in essence, translates to independence of operations is, therefore, the third way in which to afford one mode priority over another.

There are well established norms relating to the various infrastructural mechanisms that can be used to confer various levels of independence of operations to a mode. The *Guidelines for Human Settlement Planning and Design* (CSIR, 2000) lists a range of infrastructure classes for pedestrians and cyclists that are differentiated from each other by the extent to which they have to share space with other modes. The allocation of separate roadspace for PT vehicles is starting to be implemented throughout urban centres in South Africa, and abroad. The BRT systems being implemented around the country are noted for having physically delineated bus lanes as part of their trunk networks. On higher order arterial routes and freeways, freight vehicles are often limited to using a particular lane. None of the physical requirements for affording a particular mode priority, through the allocation of roadspace, is new, or untested in South Africa.

## 9.4 Interpreting Modal Suitability

Defining modal suitability in terms of operational parameters, provides the planner with some guidance as to what combination of infrastructural interventions or components would be appropriate in that particular context, whilst still allowing for innovation and flexibility in planning choices. This is important for two reasons, contexts are spatially and temporally fluid, and design circumstances may vary from place to place, even in areas with similar contexts. There are, thus, many possible design solutions for any given context that could function equally well, and only a careful analysis of each option, preceded by a thorough investigation into the reasons for the cluster means in any area, would lead to contextually appropriate infrastructure.

The adjusted cluster mean values for the Voortrekker Road data set (Figure 9.4), redistributed into the four categories described in Table 9.1, using Equation (9.1), are used to demonstrate the application of the method.



**Figure 9.4:** Adjusted Cluster Means for Voortrekker Road

Considering cluster number 1, all the modes are relatively well suited, but the pedestrian and PT modes are the best suited to the context. Accordingly, these modes should receive unrestrained access to the area, priority of movement over the other modes, and should have infrastructure dedicated for their exclusive use. This has a number of immediate important implications. Firstly, the pedestrian mode should receive full access to the area, implying that pedestrians should be allowed and expected to cross the road anywhere in the area. This would necessitate very low vehicle speeds. Secondly, there should also be dedicated PT infrastructure, which implies the provision of a dedicated lane for buses and taxis.

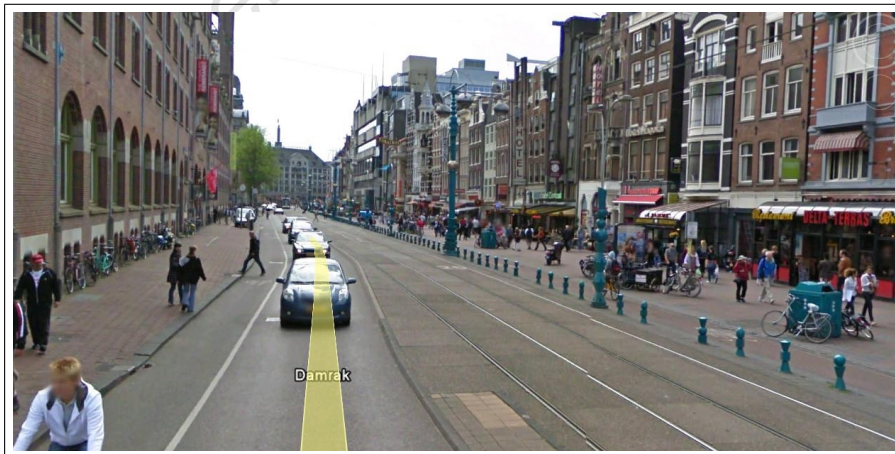
The other modes, car, bicycle and freight, all fall within the medium access range. This implies that these modes should be allowed access to the area, but with some restrictions on their operations. It may be appropriate to limit freight access to delivery vehicles only, and to not allow parking for private cars. Cyclists may be required to stay within a designated cycling lane, or may be required to dismount during certain times of day.

Figure 9.5 shows a typical street scene along Voortrekker Road in a cluster 1 area. This particular area is characterised by mixed land uses, including small retail stores, offices, and apartments along the road, and residential suburbs further away from the road. The road itself consists of two undivided lanes with on street parallel parking, and sidewalks on either side of the road.

In order to more closely comply with the requirements of Table 9.1, this section of the route might be reconfigured to more closely resemble the example in Figure 9.6. The allocation of space in this road, as well as the priority and independence of movement, quite closely matches what is suggested, given the context described by cluster 1. There is dedicated infrastructure for PT, and a large amount of space



**Figure 9.5:** Voortrekker Road in a cluster 1 area *Source:* Image sourced from Google Maps, accessed 15/04/2011



**Figure 9.6:** Damrakstraat, Amsterdam *Source:* Image sourced from Google Earth, accessed 15/04/2011

is given over to the pedestrian, who is free to cross wherever required. PMT movement is restricted to one lane, which, in this instance is one way only, and parking is not provided. Cyclists and cars share the same road surface, with bicycles not being allowed to cycle on the pedestrian only areas.

An immediate problem with this solution is that this section of Voortrekker Road is much narrower than the example shown in Figure 9.6. Compromises will, therefore, have to be made, but these should be made in deference to the mode ranking, and issues around road safety first. For example, it may not be possible to provide a dedicated bicycle lane. However, since speeds are likely to be low, and these modes have almost identical suitability scores, they could be made to share a lane. This demonstrates the need for flexibility in the approach since, often, compromise solutions will call for design innovation. Figure 9.7 shows the existing cross-section for this section of the route, and two proposals for context sensitive cross-sections, given the cluster mean scores for each mode, and the discussion above.

The two proposals highlight the roadspace allocation compromises that are required as a result of the road reserve constraints. In the existing section, approximately 70% of the roadspace is allocated to motorised modes (including PT). In both of the proposals presented, parking has been eliminated, and in Proposal 1, only one direction of flow has been allocated to PMT modes. This frees up significant amounts of space for the other modes, but could be argued to be impractical, and not truly representative of the suitability scores, since the Car mode does score comparatively well overall. Nonetheless, the space allocation has been changed to: 34% to PT, 16% to PMT and 46% to NMT. The remainder is made up by drainage and street furniture.

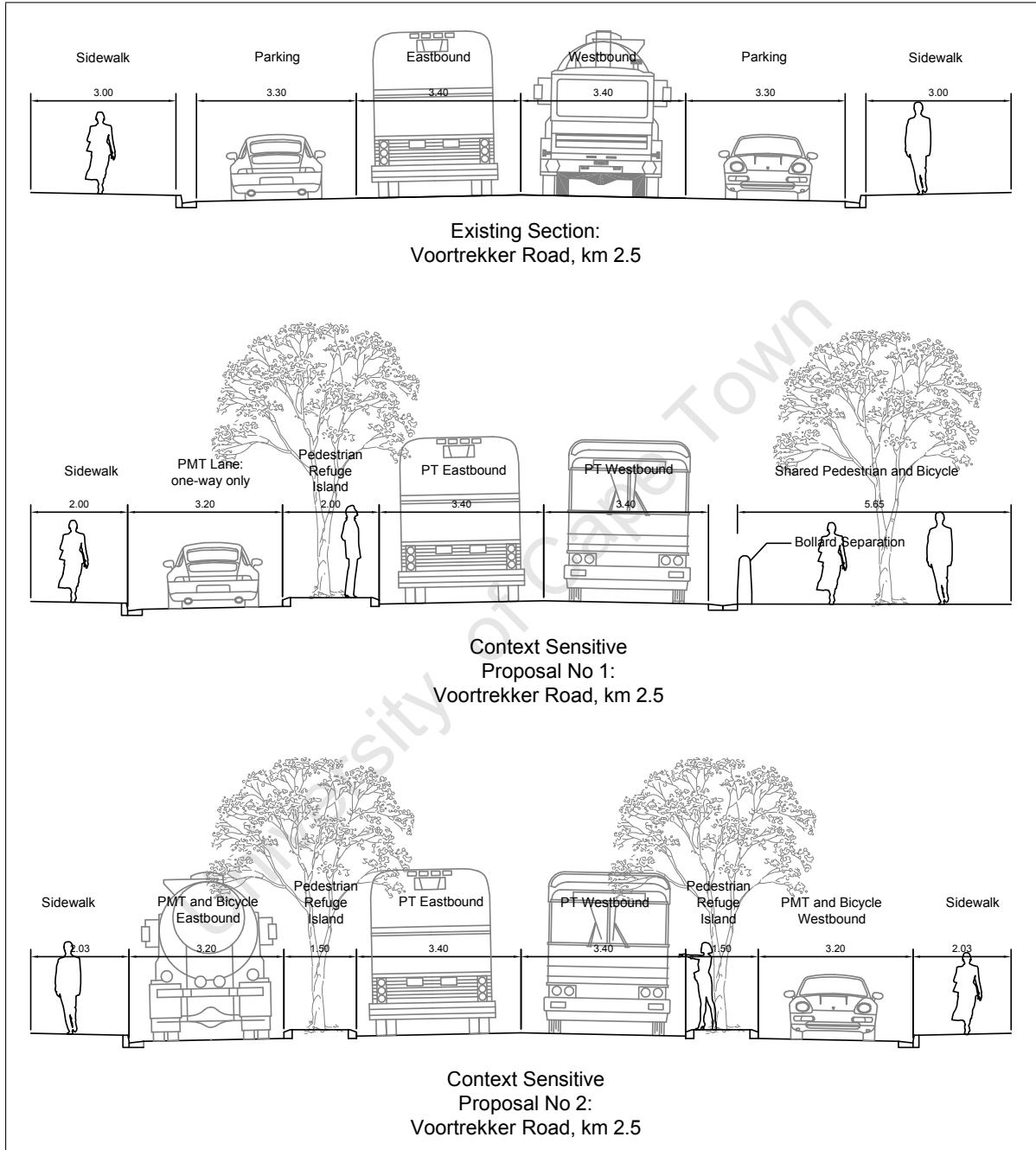
In Proposal 2, both directions of flow have been accommodated for all modes. To accommodate the extra PMT lane, some of the space for NMT had to be reassigned. However, the remaining NMT space is also now more fragmented, since refuge islands must be provided between PT and PMT lanes to assist with the unregulated crossing required by the context. The narrow PMT lanes can also be interspersed with raised humps to keep speeds down.

Another interesting example is that of cluster 4, which is found only in two, relatively short, stretches of the route. The cluster means suggest that PMT modes should be restricted from accessing these areas, which should only allow for PT and NMT modes. Of these modes PT is the preferred mode, with the NMT modes categorised as being of medium suitability (albeit that pedestrian scores towards the upper end of the scale). The implication is that provision should be made for dedicated PT facilities, possibly aligned along the median and shared pedestrian and bicycle facilities.

Cluster 4 occurs along the section of Voortrekker Road between kilometre 16.5 and kilometre 17.0. This area has already been subject to significant infrastructure improvements to improve its pedestrian friendliness, with frequent pedestrian crossings and pedestrian friendly street furniture and lighting having been installed in the past (Figure 9.8). This indicates that there is already an awareness of the special context in the area, although this analysis indicates that the interventions did not go far enough to address the problems.

Cluster 5 is unique in that it is the only cluster of the six generated for the Voortrekker Road dataset that has the PMT modes as being the most suitable to the context. Both PMT modes are highly suited to the context, whereas the remaining modes fall in the medium suitability category (with the bicycle mode being on the boundary of medium and poorly suited). In terms of the definitions outlined in Table 9.1, this means that the PMT modes must be afforded full access and priority of movement, and should have dedicated infrastructure. The remaining modes should receive partial access, with their movements or operations somewhat restricted, and can be made to share infrastructure.

Cluster 5 is found in three locations along the route. These are typically areas with either primarily



**Figure 9.7:** Existing Layout and Proposal for Context Sensitive Upgrade in a Cluster 1 area on Voortrekker Road



**Figure 9.8:** Voortrekker Road in a cluster 4 area *Source:* Image sourced from Google Maps, accessed 15/04/2011

industrial activities that attract lots of freight and delivery vehicles (Figure 9.9), or large undeveloped areas with very little activity of any kind (Figure 9.10).



**Figure 9.9:** Industrial area along Voortrekker Road designated as cluster 5. *Source:* Image sourced from Google Maps, accessed 15/04/2011

The infrastructure in these areas, particularly near the cemetery, could be said to already be quite appropriate given the context. The PMT modes are well catered for, there being two lanes in each direction, separated by a median. The addition of on-street parking could improve conditions in the industrial area, with space for parking taken from the very large sidewalks along the route. Dedicated turning lanes at intersections are not necessary, since there are already two lanes in each direction (although this would also depend upon turning movement volumes). PT stops should be provided here as well, given that it scores as medium suitability. However, this should be provided in embayments



**Figure 9.10:** Undeveloped area along Voortrekker Road designated as cluster 5. To the left is a large cemetery, and to the right is an undeveloped strip of marshland.  
*Source:* Image sourced from Google Maps, accessed 15/04/2011

only, so as not to interfere with the operations of the PMT modes. Bulbouts could be provided at intersections to improve pedestrian crossing LOS. Cycling facilities could be shared with pedestrian facilities beyond the on-street parking.

## 9.5 Résumé

This chapter outlined the approach proposed for making infrastructure recommendations from the clustered SMCA results. The approach relies upon an analysis of the cluster means, being the mean score for each mode in every cluster in the data set. The cluster means can be seen to be a good proxy for the original SMCA scores since the combination of these cluster means is a representation of the average context for a given stretch of road, interpreted in terms of modal suitability.

Clustering the data reveals which sections of the route are contextually similar to each other. Clusters can be seen to be repeated within all of the data sets, which shows that there are contextually similar areas along the route that are spatially separate from each other. These areas should, according to the hypothesis, have similar infrastructure, since the modal priorities, in terms of the context, are very similar.

Clusters range from approximately 500 m to 3 000 m in length and, in general, cluster means tend to change in a stepwise fashion, meaning that cluster mean scores do not change erratically, instead they increase and decrease in increments. This indicates that contexts also tend to change incrementally, and that modal priority tends to increase or decrease incrementally as well.

The method used to develop infrastructure recommendations relies on descriptive statements of the mode suitability. The mechanism used to translate mode suitability into a descriptive statement uses the rank of the mode, and the score of the mode in relation to the others in the cluster. The cluster means are mapped to a linear scale that sets the highest score of any mode in any cluster along the route (the route maximum) to 1 and the lowest score (the route minimum) to 0. The transformed cluster means

can then be classified into 4 score bands, highest suitability, medium suitability, low suitability and unsuitable.

For each suitability band, the appropriate operational limitations of the mode are described in terms of three parameters: the access afforded to the mode, the right of way or priority of movement afforded to the mode, and the level of independence of operations afforded to the mode. These operational characteristics are then used to determine the appropriate combination of infrastructural and regulatory interventions that would be most likely to produce the desired service levels.

## Chapter 10

# Conclusions and Recommendations

The high number of road deaths and, in particular, pedestrian deaths on South African roads remains a persistent problem, despite ongoing efforts to address it. It is also widely acknowledged that the transport system in South Africa remains highly inequitable and inefficient, hindering efforts aimed at the upliftment of the poor. The motivation for this research was to understand why, despite an improved and, arguably, facilitative legislative and policy environment, planning for road infrastructure provision in South Africa has not been able to make much headway in producing solutions to these problems. The hypothesis proposed was that planning and design practice has not fully embraced the ideals of policies, and that this plays a significant role in the poor quality and piecemeal nature of the infrastructure provided for the other road-based modes.

A review of local policies and the design manuals that guide road planning and design practices, suggested this hypothesis was correct, i.e. that policy and practice are often in conflict with each other. Acknowledging that current road planning and design practices are rooted in well established, scientifically sound methods and research, a need was, nonetheless, identified to incorporate the principles of sustainable transportation into the practice to produce a comprehensive, holistic approach to design. To this end, a problem statement was developed as:

*Each mode used in a road has its own specific characteristics and needs, which determine the design parameters for that mode. Also, each location in a network, or along a road, is defined by a set of contextual parameters that determine how, and by whom, the road is most often used. It is where the modal characteristics and the location specific factors, or the needs of a mode and the use of a location intersect, that an ideal planning solution can be found. Accordingly, certain modes are better suited to a particular set of contextual circumstances than others. Therefore, under a given mix of contextual circumstances, certain modes should be given a higher priority than the rest. This modal priority ranking can then be used to inform road planning recommendations, using established practice and guidelines, to inform decisions.*

The research outlined in this thesis investigated whether it is possible to rethink the road planning process, to incorporate a broader range of factors than is currently used, so that a more sustainable, contextually sensitive road facility can be planned. The method ultimately developed brings together techniques from a range of disciplines to demonstrate the importance of a locations context, and that this context can be quantified and used to infer the most appropriate mix of infrastructure for that location, given its context.

## 10.1 Answers to the Research Questions

Several research questions were initially posed to develop a response to the objectives set out for this research. These questions are addressed below:

1. *How are roads, currently, planned and designed, what are the reasons for this approach, and what are the pros and cons of the current approach to road planning?*

The history of transport planning was investigated, revealing the theoretical and historical underpinnings of the approach used in practice today. Whereas, during the nineteenth century, the primary concerns in street planning centered on issues of public health and well-being, the rise of the automobile through the twentieth century saw that focus shift towards traffic efficiency. The post-war era saw the rapid advancement of analytical methods of analysing infrastructure performance in relation to traffic efficiency, and the traditional role of the town planner, or architect, being usurped by the traffic engineer.

Road infrastructure, today, is planned using analytical methods that rely on forecasting future traffic efficiencies given a number of development scenarios. Planning horizons often span more than a decade, owing to the large capital costs and complexity of projects, despite the limitations of the forecasting methods. Analytical transport planning relies heavily on mathematical models that, inevitably, rely on assumptions and generalisations to simplify their complexity. Models are calibrated against empirical observations to minimise their predictive error.

In most instances, however, transport planning and, in particular, road design, is guided by a set of manuals and guidelines that outline best practices for a range of generic situations. These best practices, in turn, are based upon a combination of theoretical research and empirical and experiential evidence. There is, thus, much research effort behind the practices espoused in these documents, and they are generally uniformly accepted in the industry. Deviation from these guidelines, despite the guidelines themselves limiting their scope and calling for flexibility in their application, exposes planners and designers to professional risk, which could result in professional liability in the event of an error. As such, practice does not encourage innovation.

The focus on traffic efficiency has come to dominate transport planning, based on the premise that efficient motorised transport is an economic necessity. The concerns around the negative social and environmental impacts of motorisation, and the role of public transport (PT) and non-motorised transport (NMT) in promoting an efficient, balanced transport system, are not adequately addressed using the current approach to transport planning. Modes tend to be planned in isolation from each other, and this often results in PT and NMT being inadequately planned for, or not planned for at all. The continued improvement of service levels for private motorised transport (PMT) fuels the growth in the use of these modes and discourages the use of NMT and PT, thereby aggravating the situation further.

The domination of PMT has a wide range of negative effects on the transport system, and on the other urban and social systems with which it interacts. Inadequate provision for and investment in, PT disadvantages those who rely exclusively on this mode in terms of accessibility and quality of life. This, directly, impacts on the success of programmes aimed at improving the lives of the underprivileged and marginalised sectors of society, since it is primarily these sectors of society that rely on PT for their mobility. Transport planning, in its current form, can thus be said to be hampering the goals of the democratic revolution in South Africa.

Inadequate provision for NMT is another unfortunate consequence of the current planning paradigm. Since PT and NMT are intrinsically linked, and there are many people who are captive

to both modes, the effects of the poor provision and planning for these modes is self-reinforcing. Poor NMT services impacts on the overall quality of service for PT, and inadequate investment in PT services means that people have to walk or cycle further to access these services.

Poor planning for NMT also has road safety implications, in that people end up being forced to walk or cycle along routes that are not designed to accommodate them. The result is that the interactions between the more vulnerable NMT modes and PMT is not controlled or regulated and accidents occur, which often ends in death or serious injury for the pedestrian or the cyclist.

**2. *What are the legal, policy and regulatory frameworks that affect road planning practice, and to what extent do these influence the planning practice?***

Transport planning is guided by a range of policy and legislative instruments that address various aspects of the practice. The Moving South Africa (MSA) project was commissioned by the NDoT to, "...produce a data-driven program for strategic action that extends the short to medium-term policy formulation documented in the Transport White Paper into a long-term strategic formulation embodying the sets of trade-offs and choices necessary to realise the vision as set out in the White Paper." The first major revision of the legislation that governed transportation came in the form of the National Land Transport Transition Act of 2000, which has subsequently been superseded by the National Land Transport Act of 2009. The act furthers the visions and objectives of the 1996 White Paper and the 1999 MSA policy document, by placing a very strong emphasis on the importance of PT, and issues of social justice and the upliftment of the poor. The National Land Transport Strategic Framework (NLTSF) for the period 2006 to 2011 is a legal requirement in terms of Section 21 of the National Land Transport Transition Act and gives guidance on transport planning and land transport delivery by national government, provinces and municipalities.

The policy and legislative developments, since the dawn of democracy in South Africa, have put policy and practice at odds with each other in that they call for the prioritisation of PT, and they emphasise the importance of NMT, in particular walking, in the transport environment. Transport planning practice, on the other hand, appears ill-equipped to meet these policy needs, since much of the methods and tools used were optimised for a unimodal, automobile centric perspective on transport. Multimodal concerns are very poorly integrated into practice, and methods to address this have only, recently, begun to be developed and adopted.

As such, there remains a disjuncture between the ideals of policy and legislation, and the practicalities of planning transport infrastructure.

**3. *What is the role of context in the transport environment?***

In recent times, the role of the local context in developing transport infrastructure has begun to receive greater attention through initiatives such as Context Sensitive Design (CSD), the Complete Streets programme and the ARTISTS initiative. At a 2008 workshop called "*Thinking Beyond the Pavement: A National Workshop on Integrating Highway Development with Communities and the Environment*", the CSD process was defined as follows: "*Context sensitive design asks questions first about the need and purpose of the transportation project, and then equally addresses safety, mobility, and the preservation of scenic, aesthetic, historic, environmental, and other community values. Context sensitive design involves a collaborative, interdisciplinary approach, in which citizens are part of the design team.*"

The definition incorporates a range of environmental influences to construct a description of the context and, by implication, CSD. Importantly, it notes the primacy of identifying the needs and

purpose of the transportation project. This speaks directly to the function of the facility under consideration. Roads have been found to serve a whole range of urban functions, which are direct consequences of the very environmental aspects identified in the definition of CSD presented in the quotation.

This research has defined the context as being related to who is using the road and why. In this sense, the context relates to the people using the road and their characteristics, and the activities along the road and the characteristics of these activities. All of this information is interrelated. The activities being carried out, at a particular location, are related to the land uses at that location which is, in some sense, related to the characteristics of the people who frequent the location.

Nonetheless, the context is related to the transport environment, in that it plays a crucial role in determining the non-mobility related aspects of a road, collectively referred to as the 'access' function of a road. The mobility function of the road is related to the configuration of the broader road network, and the trip distribution patterns on that network.

**4. *What are the factors that best describe the context of a route and to what extent do these factors currently play a role in road planning practices?***

As mentioned, the context can be defined as being related to who is using the road and why they are there. This definition draws upon demographic information of the road users, and information on the activities being performed. This information is recognised as being critically important to the prediction of travel behaviour and, in particular, modal split.

The information that best describes the context can be divided into four categories. These include land use information, socioeconomic information, environmental information and transportation information. Variables were considered for these categories that are commonly cited in transportation literature. A particular emphasis was placed on information that impacts travel behaviour and, in particular, modal split, since modal priority is defined as the mechanism through which infrastructure choices are determined in relation to the context. Variables identified as being important include the land use type, population or household density, employment density, income, age, culturally and ecologically sensitive areas, the relative proportion of each mode using the route and the location of PT stops.

**5. *What would be an appropriate platform to incorporate context in the planning process?***

Prioritising modes, given a range of criteria, can be thought of as a decision problem. This problem can be stated as, "*Given a set of information regarding a particular location, and information regarding the various modes of transport, which mode is best suited to that location?*" Implicit in this statement is that modal suitability is assessed as the performance of each mode in relation to the characteristics of the location. Such a problem, where there are multiple alternatives, each to be assessed in relation to their performance in terms of multiple criteria, is termed a multiple criteria problem. The family of methods used to solve these problems is termed Multiple Criteria Analysis (MCA).

Although the decision problem being addressed in this research is a multiple criteria problem, it also has a distinct spatial element, in that the context, as it is defined here, varies spatially. As such, the method used must be able to conduct the evaluation spatially. The spatial application of MCA is called Spatial Multiple Criteria Assessment (SMCA). With SMCA, the spatial decision problem is visualised as a map of criteria tables. Each criteria is represented spatially with a map, and the criteria maps are combined through MCA techniques, to obtain a level of suitability map for each alternative, thus reflecting the spatial performance of each alternative. This method

makes it possible to perform a MCA, using both spatial criteria and non-spatial criteria, and the assessment can be carried out without losing the spatial dimension.

#### 6. *What are the data requirements for a context analysis?*

One of the key features of the context, as defined in this research, is that it varies spatially and, potentially, temporally as well. Furthermore, the description of the context, which is based upon who is using a location and why they are there, is made up of land use, socioeconomic, environmental and transportation information. The information, or data requirements, for a description of the context, therefore, requires the collection of spatial information from each of these categories of data.

The methodology used in this research, therefore, relies on the use of geocoded information to conduct the SMCA. Much of this information is already available from various branches of government who collect the information for their own administrative and planning purposes. Other than the need to verify the accuracy of the information collected, the need for manual data collection is, therefore, minimal, as long as such information is available.

Table 10.1 lists the information used in the SMCA and the source of the data for each criteria.

**Table 10.1:** Criteria and Data Sources

<b>Criteria</b>	<b>Data Source</b>
Land use	City of Cape Town Corporate GIS Department
Household density	2001 National Census
Proportion of vulnerable road users	2001 National Census
Income*	2001 National Census
Employment	Regional Services Council
Proximity to heritage sites	City of Cape Town Environmental Resource Management Department
Proximity to wetlands	City of Cape Town Environmental Resource Management Department
Proximity to ecologically sensitive areas	City of Cape Town Environmental Resource Management Department
Public Transport demand	City of Cape Town Transport, Roads & Major Projects Department
Private Car demand	City of Cape Town Transport, Roads & Major Projects Department
Proximity to public transport stops	City of Cape Town Transport, Roads & Major Projects Department

\* Education level data was used as a proxy for income data (see Section 6.1.5).

**7. Which corridors would be best suited for developing and testing the method?**

In developing the method, case study roads were selected on the basis of the extent of variation in the factors used to describe the context along the routes, and by selecting amongst those routes that were known to be, particularly, problematic regarding accidents. By necessity then, the case studies used to develop the method were all arterial roads, since these tend to be especially long, passing through a large range of contextual settings. They, therefore, provide the best opportunity to use a large, robust set of data to test various analysis strategies and calibrate the method.

Three case study roads were selected for the development, testing and the calibration of the method: Voortrekker Road, Lansdowne Road and Koeberg Road. As mentioned, these are arterial routes in the Cape Metropolitan Region. They are all longer than 10 km in length and have a variety of land uses and socioeconomic conditions along the routes. They are all, also, noted for being particularly dangerous from a road safety perspective (see Table 6.5).

**8. What are the crucial elements of the method that need to be developed and validated?**

The SMCA, itself, involves a range of processes that must be developed and calibrated. This includes the data normalisation and standardisation protocols, and the method used to weight the results. The results of the SMCA are raster maps of the suitability scores for each mode of transport, across the extent of the data sets used (in the case of this research, the city of Cape Town). This information must be processed further in order to extract meaning from it.

There are a number of processes that must be tested and calibrated to ensure that the results are representative of the context. Since the context is being interpreted in terms of modal suitability, the suitability scores along the case study routes must be extracted from the SMCA output rasters along the centreline of each case study route. Before this can be done, the scores need to be averaged to include those suitability scores further out from the centreline of the road. This is an image processing technique, and the parameters for the process must be calibrated to ensure that the results are representative of the context.

The processed SMCA results can then be extracted to a spreadsheet for further analysis. These results represent the point suitability scores along the route at a regular spacing interval defined by the resolution of the raster maps generated by the SMCA (in the case of this research, the resolution was 10 m). These point suitability results must now be clustered to identify areas along the route that are contextually similar, and to identify typical context 'regimes' or types along the route. Each of these context types can, in theory, be treated similarly with respect to the infrastructure provided for them.

The clustering methodology used is a separate procedure that must be defined, tested and calibrated to ensure that the clusters produced are acceptable. This involves, firstly, selecting a suitable clustering algorithm to apply, and setting the parameters for this algorithm, and secondly, identifying a suitable method for choosing the correct number of clusters, since most clustering algorithms will not automatically identify the 'best' number of clusters.

**9. How could context be used to make planning recommendations?**

Clustering the SMCA data reveals the underlying structure of the data sets, and allows further insight into the data, enabling the evaluation of hypotheses, the detection of anomalies, and the identification of salient features in the context. It also helps to create a classification of the context, by establishing the degree of similarity between sections of the route. It can be used to simplify the data, by compressing the original data into groups that arranges the data into route typologies.

The method used to develop infrastructure recommendations relies on descriptive statements of the mode suitability, based upon an analysis of the clustered data. The mechanism used to translate mode suitability into such a descriptive statement uses the rank of the mode, and the score of the mode, in relation to the others in the cluster. The cluster means are mapped to a linear scale, that sets the highest score of any mode in any cluster along the route (the route maximum) to 1, and the lowest score (the route minimum) to 0. The transformed cluster means can then be classified into score bands. In this research it is proposed that four score bands be used, namely, highest suitability, medium suitability, low suitability and unsuitable.

For each suitability band, the operational characteristics that are considered appropriate given the modes level of suitability to the context in that location are defined. The operational characteristics imposed on the mode in that section of the route, are described in terms of three parameters: the access afforded to the mode, the right of way or priority of movement afforded to the mode, and the level of independence of operations afforded to the mode. These operational characteristics are then used to determine the appropriate combination of infrastructural and regulatory interventions that would most likely produce the desired mix of service levels.

An overview of the methodology developed in the research is provided in Figure 10.1.

## 10.2 Reflection

The work conducted towards this thesis made a number of novel findings that contribute to knowledge, and have the potential to improve the status quo. The basis of the work was that there are a range of factors that can be used to describe the nature of the people using a road, and the activities they are involved in along it. This information should play a more direct role in the planning of these roads. The factors, collectively termed the context, have always, in some sense, been considered in transport planning, but have tended to be overshadowed by concerns around efficiency and cost, and there has not been a comprehensive framework within which the context could be evaluated, and its implications investigated. The way in which infrastructure interfaced with, or suited the context has, thus, always been left to the discretion or judgement of the engineer or planner of the facility. The development of a method to systematically evaluate the context, and interpret the results in terms of infrastructure is, therefore, the primary contribution of this research.

The research identified which factors could be used to describe the context of a location, and demonstrated that it is possible to quantify the context in terms of its effects on the suitability of the various modes of transport. Quantification has a number of advantages for planning and designing infrastructure. Being able to quantify the suitability of a mode of transport to a particular location, given the context, allows for the prioritisation of modes in terms of infrastructure provision, which can then be used as the basis for planning and design.

Quantification also has the benefit of facilitating accurate comparisons between different locations along the route. This is useful for planning and design in that the subtleties, in variation of the context along the route, are retained during the analysis, allowing for a fine grained tailoring of the required infrastructure. It also demonstrates that context varies spatially, that it is not static, and that for infrastructure to be contextually sensitive it must, therefore, vary to suit.

The method developed to assess the context employs a novel application of the principles of multiple criteria assessment to conduct the evaluation. Whereas SMCA had previously been used to compare the suitability of a number of sites for a development, or to identify routing alternatives in an analysis space, the application developed here assesses suitability of a number of alternatives in a constrained

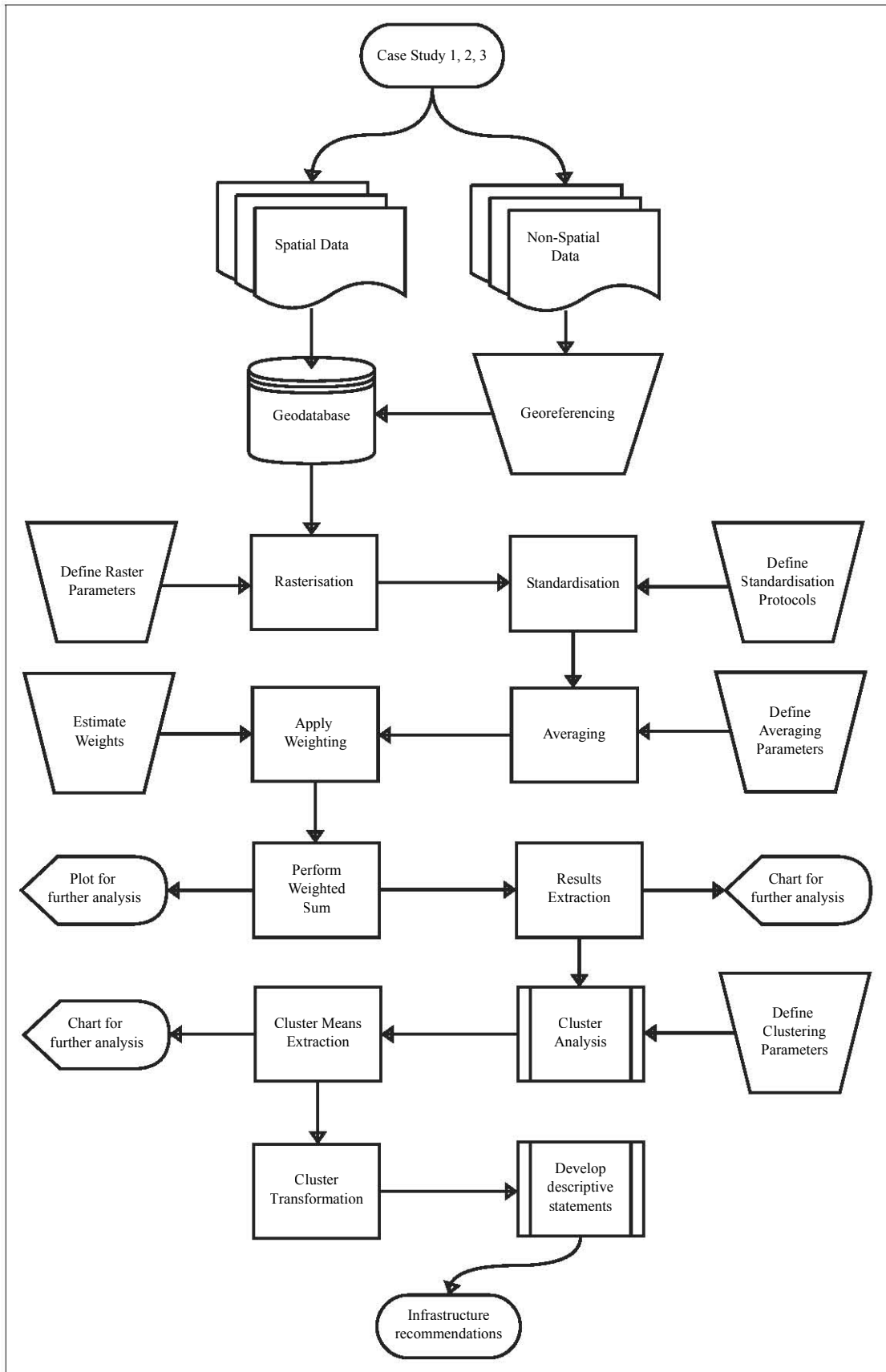


Figure 10.1: Overview of methodology

space and, uniquely, no alternative is necessarily taken as being the one correct solution. Instead, it may only be the correct solution at that location. Also, none of the other alternatives are abandoned if they do not rank highest. Instead, their relative suitability is used to determine their priority in terms of the operational aspects of the road being planned.

The standardisation framework used in the SMCA is also novel in that, for any one criterion, the alternatives being assessed cannot logically be standardised using the same standardisation method. A separate system of assessment values, therefore, needed to be employed to guide the standardisation used for each criterion, for the various alternative modes. This introduces an element of subjectivity into the analysis at an early stage, and complicates the application of weighting, using traditional weighting methodologies. To overcome this, the same system of assessment values was used to derive appropriate weights, thereby ensuring that a consistent set of values is used throughout the assessment.

The context, being an amalgam of a range of disparate factors, does not have any intrinsic meaning by itself. Instead, it is the implications of the context that has meaning and, therefore, context can only be understood in terms of its implications for other aspects of the facility. In this research, context is defined in terms of its implications for the suitability of the various modes of transport. The translation of contextual suitability into infrastructure recommendations, therefore, relied on defining the operations of the modes of transport on the road. This approach is somewhat similar to the definition of LOS, in that it also describes service levels in terms of operational characteristics and performances. However, in this instance, these descriptions are prospective, in that they define what should be, instead of retrospective, or defining what is.

The methodology developed in this research, therefore, successfully develops a definition of the context, demonstrates its importance, explores its characteristics and implications and translates these into descriptives that can be used to inform infrastructure provision. Although the research was conducted under the auspices of the *Cycling Academic Network* (CAN), the method developed treats all modes of transport equally, and is explicitly multimodal in nature from the outset. Since all modes receive equal treatment, the method is able to highlight any existing disparities in the provision of infrastructure for the various modes. During the course of the research, the use of the method to evaluate existing infrastructure was tested for just this purpose. Beukes and Zuidgeest (2010) compared the results of a multimodal LOS analysis to the contextual analysis results based on the methodology developed in this research for a short stretch of Lansdowne Road. The paper demonstrated one particular application of the methodology for road infrastructure planning.

### 10.3 Recommendations

The findings of this thesis indicate that it is important that the contextual setting be considered in road planning. Currently, a systematic analysis of the contextual variation along the route does not form part of road planning practice. In addition, current planning practice results in infrastructure provision that is biased towards the needs of PMT. Policy priorities are, therefore, not being met as effectively as they should be, which impacts negatively on the success of efforts to improve the lives of the underprivileged, and contributes to the road safety problems on the country's roads. Contextually sensitive planning could help to correct the disjuncture between the objectives of transport policy, and transport planning. It is, therefore, recommended that an analysis of the context be included in the current transport planning processes.

Road projects, typically, include a feasibility, or preliminary, investigation stage, the findings of which are then presented in a report. This is followed, after some deliberation between stakeholders, by a detailed design phase, during which construction drawings and contract documentation are produced.

The major decisions regarding the extent and focus of the project are, therefore, finalised before the detailed design phase and it is, therefore, during this initial stage of the project cycle that an analysis, such as the one developed in this research, would be most beneficial. It is recommended that planning authorities require that a contextual assessment be included as part of the initial project planning investigations.

Context studies could also be conducted prior to any project being initiated, as part of the overall strategic planning for an area. This information would be useful to investigate different planning scenarios, and could assist in identifying current deficiencies in the existing network. Context investigations could help provide an explanation for localised road safety problems, and could be used to identify interventions that may improve conditions in these areas. It is recommended that authorities conduct a context analysis similar to the type developed in this research, on their road networks.

Studies, such as that conducted for this research, are data intensive and, although geocoded data is expensive to generate, this data can be used to analyse a range of problems, and is very useful for planning, in general. One of the primary shortcomings of this research is that the data used was all at different levels of aggregation, and not all from the same reference year. This has negative implications for the precision with which this study was conducted. It is recommended that more effort is expended in geocoding data sets, and bringing existing data sets up to date.

The work conducted in this research also raises a number of opportunities for further research. Firstly, it is conceivable that, depending upon data availability, certain criteria may not be able to be included in the analysis. It would be beneficial to establish the 'core' data requirements for a reasonable description of the context. Moreover, the impacts of excluding certain information should be examined in more detail. Also useful would be a better understanding of the implications of including additional, non-core data in the analysis, and the development of guidelines for criteria selection that would outline these findings.

There is opportunity for developing a more streamlined process for the cluster analysis and, in particular, the selection of the correct number of clusters to use for the final results. This part of the methodology requires a detailed understanding of clustering theory, that may complicate the widespread adoption of the methodology. Data clustering is a constantly evolving, rapidly developing, field and there is definitely scope for a more comprehensive investigation of the different clustering techniques that have been developed.

All of the case studies investigated during the development of the methodology were urban arterial routes in Cape Town. The rationale for this decision was to select routes that have a large variety of contextual variation along their routes, but there is definitely opportunity for investigating the application of the method on non-arterial or even rural routes. It is suspected, but has not been verified, based upon this investigation of the SMCA results, that the recommendations for these types of roads will be in accordance with the findings for the case studies used. It could also be useful, for the same reasons, to investigate the results for other areas in South Africa and elsewhere.

There is a definite opportunity, beyond the scope of the work conducted for this thesis, to adapt the methodology for an area-wide analysis. Given that the SMCA was conducted across the city as a whole, such an adaptation will, most likely, involve clustering the SMCA results across the whole city. It is hoped that this work will be able to provide further insight into the contextual structure of the city as a whole, and urban areas in general, and open up opportunities for developing a generalised suite of context types.

# Bibliography

- Ahern, A., Hine, J. and Begley, E. (2010). Transport for the elderly - what happens in rural areas?, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–20.
- American Association of State Highway and Transportation Officials (AASHTO) (ed.) (2004). *A Policy on the Geometric Design of Highways and Streets*, 5 edn, American Association of State Highway and Transportation Officials (AASHTO), Washington, D.C.
- Anderson, M. D. (2002). *An Examination of Trip Generation and Travel Rate Indices for Alabama*, The University of Alabama in Huntsville, Huntsville, AL.
- Andreas, P. (2005). *Cluster Analysis: Overview, Algorithms, and Comparison*, Masters thesis, University of London.
- Andreopoulos, B., An, A., Wang, X. and Schroeder, M. (2009). A roadmap of clustering algorithms: finding a match for a biomedical application, *Briefings in Bioinformatics* **10**(3): 297–314.
- Andritsos, P. (2002). *Data Clustering Techniques*, PhD thesis, University of Toronto.
- Antoine, J. (1997). Multiple criteria land use analysis, *Applied Mathematics and Computation* **83**(2-3): 195–215.
- Appleyard, D. (1980). Livable Streets: Protected Neighborhoods?, *The Annals of the American Academy of Political and Social Science* **451**(1): 106–117.
- Aubert, P. J. (1986). The siting of nuclear power plants in The Netherlands: on the possible use of multicriteria decision methods in energy policy.
- Badoe, D. (2000). Transportation - land-use interaction: empirical findings in North America, and their implications for modeling, *Transportation Research Part D: Transport and Environment* **5**(4): 235–263.
- Bailey, K., Grossardt, T. and Jewell, W. (2005). Participatory routing of electric power transmission lines using the EP-AMIS GIS/Multicriteria Evaluation Methodology, *Proceedings of the 10th International Symposium on Information and Communication Technologies in Urban and Spatial Planning and Impacts of ICT on Physical Space (CORP 2005)*, Vol. 1, Wien, Austria, pp. 137–142.
- Banister, D. (1995). Transport and the environment: a review article, *The Town Planning Review* **66**(4): 453–458.
- Bates, J. (2000). History of demand modelling, in D. A. Hensher and K. J. Button (eds), *Handbook*

- of Transport Modelling*, Vol. 1, Pergamon Amsterdam, Amsterdam, The Netherlands, chapter 2, pp. 11–33.
- Behrens, R. (2005). Accommodating walking as a travel mode in South African cities: Towards improved neighbourhood movement network design practices, *Planning Practice and Research* **20**(2): 163–182.
- Behrens, R. (2010). Integrating paratransit in Bus Rapid Transit systems in South Africa: A review of progress, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 2010–2010.
- Behrens, R. and Wilkinson, P. (2003). Metropolitan transport planning in Cape Town, South Africa: A critical assessment of key difficulties.
- Beirao, G. and Sarsfieldcabral, J. (2007). Understanding attitudes towards public transport and private car: A qualitative study, *Transport Policy* **14**(6): 478–489.
- Belmont, D. M. (1954). Effect of shoulder width on accidents on two-lane tangents, *Highway Research Bulletin* **91**: 29–32.
- Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Applications to Travel Demand*, The MIT Press, Cambridge, Massachusetts, USA.
- Bendavid-Val, A. (1991). *Regional and Local Economic Analysis for Practitioners*, Praeger, Westport, Connecticut, USA.
- Berkhin, P. (2006). A survey of clustering data mining techniques, in J. Kogan, C. Nicholas and M. Teboulle (eds), *Grouping Multidimensional Data: Recent Advances in Clustering*, Springer, Berlin, Germany.
- Bester, C. J. and Makunje, J. A. (1998). The effect of rural road geometry on safety in Southern Africa, *International Symposium on Highway Geometric Design Practices*, Transportation Research Board, Boston, Massachusetts, pp. 1–10.
- Beukes, E. A. (2002). *The Effects of Traffic Calming Methods*, Undergraduate thesis, University of Cape Town.
- Beukes, E. A. and Zuidgeest, M. (2010). Alternative multimodal planning processes for the developing world - A case study in Cape Town, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–24.
- Bezdek, J. C. and Pal, N. R. (1998). Some new indexes of cluster validity, *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society* **28**(3): 301–15.
- Bhat, C. (1997). Work travel mode choice and number of non-work commute stops, *Transportation Research Part B: Methodological* **31**(1): 41–54.
- Birch, E. L. (1980). Radburn and the American Planning Movement, *Journal of the American Planning Association* **46**(4): 424–439.
- Black, A. (1990). The Chicago area transportation study: A case study of rational planning, *Journal of Planning Education and Research* **10**(1): 27.

- Black, C., Collins, A. and Snell, M. (2001). Encouraging walking: the case of journey-to-school trips in compact urban areas, *Urban Studies* **38**(7): 1121.
- Black, J. A., Paez, A. and Suthanaya, P. A. (2002). Sustainable Urban Transportation: Performance Indicators and Some Analytical Approaches, *Journal of Urban Planning and Development* **128**(4): 184.
- Blanco, H., Alberti, M., Forsyth, A., Krizek, K. J., Rodríguez, D. A., Talen, E. and Ellis, C. (2009). Hot, congested, crowded and diverse: Emerging research agendas in planning, *Progress in Planning* **71**(4): 153–205.
- Blaser, B., Liu, H., Mcdermott, D., Nuszdorfer, F., Phan, N. T., Vanchindorj, U., Johnson, L. and Wyckoff, J. (2004). GIS-Based cumulative effects assessment, *Technical Report CDOT-DTD-R-2004-6*, Colorado Department of Transportation - Research Branch, Denver, CO.
- BMJ Publishing Group (1906). The Motor Car Report, *The British Medical Journal* **2**(2380): 308–309.
- Bors, P., Dessauer, M., Bell, R., Wilkerson, R., Lee, J. and Strunk, S. L. (2009). The Active Living by Design national program: community initiatives and lessons learned, *American Journal of Preventative Medicine* **37**(6 Suppl 2): S313–21.
- Botha, G. (2009). Road traffic report 2008-2009, *28th Southern African Transport Conference (SATC 2009)*, Pretoria, South Africa.
- Bovy, P. H. L. and Van der Zijpp, N. J. (1999). Transportation Modelling, *Lecture notes {CTvk4800}*, Delft University of Technology, Delft, The Netherlands.
- Boyer, M. C. (1983). *Dreaming the rational city: The myth of American city planning*, The Massachusetts Institute of Technology, Cambridge, MA.
- Braza, M., Shoemaker, W. and Seeley, A. (2004). Neighborhood design and rates of walking and biking to elementary school in 34 California communities, *American Journal of Health Promotion* **19**(2): 128.
- Brilon, W. (2000). Traffic flow analysis beyond traditional methods, *Proceedings of the 4th International Symposium on Highway Capacity*, pp. 26–41.
- Brilon, W. and Blanke, H. (1993). Extensive traffic calming: results of the accident analyses in 6 model towns, *ITE 1993 Compendium of Technical Papers* pp. 119–23.
- Brookes, C. (1997). A parameterized region-growing programme for site allocation on raster suitability maps, *International Journal of Geographical Information Science* **11**(4): 375–396.
- Brown, A. and Affum, J. (2002). A GIS-based environmental modelling system for transportation planners, *Computers, Environment and Urban Systems* **26**(6): 577–590.
- Brysiewicz, P. (2001). Pedestrian road traffic collisions in South Africa, *Accident and Emergency Nursing* **9**(3): 194–197.
- Buchanan, C. (1961). Standards and values in motor age towns, *Journal of the Town Planning Institute* **47**: 320–329.
- Buchanan, C. (1963). *Traffic in Towns*, Penguin Books, London, UK.

- Budayan, C., Dikmen, I. and Birgonul, M. T. (2009). Comparing the performance of traditional cluster analysis, self-organizing maps and fuzzy C-means method for strategic grouping, *Expert Systems with Applications* **36**(9): 11772–11781.
- Bunn, F. (2003). Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis, *Injury Prevention* **9**(3): 200–204.
- Bunn, F., Collier, T., Frost, C., Ker, K., Roberts, I. and Wentz, R. (2003). Area-wide traffic calming for preventing traffic related injuries, *Cochrane Database of Systematic Reviews (Online)* . URL: <http://www.ncbi.nlm.nih.gov/pubmed/12535454>
- Bureau of Public Roads (1964). *Traffic Assignment Manual*, Department of Commerce, Urban Planning Division, Washington D.C., USA.
- Burnham, K. P. (2004). Multimodel inference: understanding AIC and BIC in model selection, *Sociological Methods & Research* **33**(2): 261–304.
- Cairns, S., Atkins, S. and Goodwin, P. (2002). Disappearing traffic? The story so far, *Municipal Engineer* **151**(1): 13–22.
- Caltrans (2005). *Main Streets: Flexibility in Design & Operations*, California Department of Transportation, Sacramento, California.
- Cameron, R. (1996). G3 F7 - An Expanded LOS Gradation System, *ITE Journal* **66**: 40–41.
- Carver, S. (1991). Integrating multi-criteria evaluation with geographical information systems, *International Journal of Geographical Information Science* **5**(3): 321–339.
- Castle, J. A. and Kamyra-Lukoda, G. E. (2006). *Review of International Road Safety Good Practice*, TRL Limited, London, UK.
- Cervero, R. (1989). *America's Suburban Centers: The Land-Use - Transportation Link*, Routledge, London, UK.
- Cervero, R. (1996). Mixed land-uses and commuting: evidence from the American housing survey, *Transportation Research Part A: General* **30**(5): 361–377.
- Cervero, R. (2001). Induced demand: An urban and metropolitan perspective, *Policy Forum: Working Together to Address Induced Demand*, Eno Transportation Foundation, Washington, D.C.
- Cervero, R. (2010). Ridership and Land Market Impacts of BRT Services: Experiences in Los Angeles, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal.
- Cervero, R. and Gorham, R. (1995). Commuting in Transit Versus Automobile Neighborhoods, *Journal of the American Planning Association* **61**(2): 210–225.
- Cervero, R. and Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods, *Transport Policy* **3**(3): 127–141.
- Chakhar, S. and Mousseau, V. (2007). Spatial multicriteria decision making, in S. Shehkar and H. Xiong (eds), *Encyclopedia of Geographic Information Science*, Springer-Verlag, New York.
- Charnes, A. and Cooper, W. (1961). *Management Models and Industrial Applications of Linear Programming*, Wiley, New York.

- Charnes, A., Cooper, W. and Ferguson, R. (1955). Optimal estimation of executive compensation by linear programming, *Management Science* **1**: 138–151.
- Chatman, D. G. (2008). Deconstructing development density: Quality, quantity and price effects on household non-work travel, *Transportation Research Part A: General* **42**: 1008–1030.
- Chen, C., Gong, H. and Paaswell, R. (2007). Role of the built environment on mode choice decisions: additional evidence on the impact of density, *Transportation* **35**(3): 285–299.
- Choocharukul, K. (2004). User perceptions and engineering definitions of highway level of service: an exploratory statistical comparison, *Transportation Research Part A: Policy and Practice* **38**: 677–689.
- Christie, R. (2001). The effectiveness of driver training as a road safety measure: An international review of the literature, *Road Safety Research, Policing and Education Conference*.
- Church, R., Gerrard, R., Gilpin, M. and Stine, P. (2003). Constructing cell-based habitat patches useful in conservation planning, *Annals of the Association of American Geographers* **93**(4): 814–827.
- City of Cape Town (2005). *Traffic Accident Report*, City of Cape Town, Cape Town, South Africa.
- City of Cape Town (2007). *City of Cape Town Zoning Scheme 4th Draft*, City of Cape Town, Cape Town, South Africa.
- City of Cape Town (2008). *Traffic Calming Policy*, City of Cape Town, Cape Town, South Africa.
- City of Cape Town (2009a). *Integrated Transport Plan for the City of Cape Town 2006 to 2011*, City of Cape Town, Cape Town, South Africa.
- City of Cape Town (2009b). *Integrated Transport Plan for the City of Cape Town 2006 to 2011 - Review and Update October 2009*, City of Cape Town, Cape Town, South Africa.
- Clark, B. (2003). Ebenezer Howard And The Marriage Of Town And Country: An Introduction to Howard's Garden Cities of To-morrow (Selections), *Organization & Environment* **16**(1): 87–97.
- Clark, D. (1998). Interdependent Urbanization in an Urban World: An Historical Overview, *The Geographical Journal* **164**(1): 85.
- Clifton, K. J. and Kreamer-Fulfs, K. (2007). An examination of the environmental attributes associated with pedestrian-vehicular crashes near public schools, *Accident Analysis & Prevention* **39**(4): 708–15.
- Cobian, R., Henderson, T., Mitra, S., Nuworsoo, C. and Sullivan, E. (2009). Vehicle Emissions and Level of Service Standards: Exploratory Analysis of the Effects of Traffic Flow on Vehicle Greenhouse Gas Emissions, *88th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- Committee of Land Transport Officials (COLTO) (1986). *Urban Transport Guidelines: Draft UTG1 - Guidelines for the Geometric Design of Urban Arterial Roads*, Council for Scientific and Industrial Research, Pretoria.
- Committee of Land Transport Officials (COLTO) (1999). *South African Road Safety Manual*, South African Department of Transport, Pretoria, South Africa.
- Committee of Urban Transport Authorities (1989). *Draft UTG 7: Geometric Design of Urban Local Residential Streets*, Department of Transport, Pretoria, South Africa.

- Cope, A. J. (1955). Traffic accident experience - before and after pavement widening, *Traffic Engineering* pp. 114–115.
- Council, F. and Stewart, J. R. (2000). Safety effects of the conversion of two-lane rural to four-lane rural roadways based on cross-sectional models, *79th Annual Meeting of the Transportation Research Board*, Transportation Research Board.
- Council for Scientific and Industrial Research (CSIR) (2000). *Guidelines for Human Settlement Planning and Design*, Department of Housing, Pretoria, South Africa.
- Crane, R. and Crepeau, R. (1998). Does neighborhood design influence travel?: A behavioral analysis of travel diary and GIS data, *Transportation Research Part D: Transport and Environment* **3**(4): 225–238.
- Crum, F. S. (1913). Street Traffic Accidents, *Publications of the American Statistical Association* **13**(103): 473–528.
- Curtis, C. (1995). Reducing the need to travel: strategic housing location and travel behaviour, *Reducing the Need to Travel: Some Thoughts on PPG13; A Collection of Papers* **1**(2): 29–47.
- Dart, K. O. and Mann, L. (1970). Relationship of rural highway geometry to accident rates in Louisiana, *Highway Research Record* **313**: 1–15.
- Datta, S. (2003). Comparisons and validation of statistical clustering techniques for microarray gene expression data, *Bioinformatics* **19**(4): 459–466.
- De Cerreño, A. L. C. and Pierson, I. (2004). *Context Sensitive Solutions in Large Central Cities*, Rudin Center for Transportation Policy & Management, New York.
- De Smith, M., Goodchild, M. and Longley, P. (2007). *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*, 3rd edn, Troubador Publishing.
- Deehr, R. C. and Shumann, A. (2009). Active Seattle: Achieving walkability in diverse neighborhoods, *American Journal of Preventative Medicine* **37**(6 Suppl 2): S403–11.
- Delbosc, A. and Currie, G. (2010). Exploring the separate and combined influences of transport disadvantage and social exclusion on well-being, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–13.
- Delucchi, M. A. (1998). The annualized social cost of motor-vehicle use in the U.S., 1990-1991: Summary of theory, data, methods, and results, *Technical Report UCD-ITS-RR-96-3 (1)*, Institute of Transportation Studies, University of California Davis, Davis, California.
- Densham, P. (1991). Spatial decision support systems, in D. Maguire, M. Goodchild and D. Rhind (eds), *Geographical Information Systems: Principles and Applications*, Vol. 1, Longman Scientific and Technical, London, pp. 403–412.
- Department of Environmental Affairs (1998). *National Environmental Management Act*, South African Government, Pretoria, South Africa.
- Diamond, J. T. and Wright, J. R. (1988). Design of an integrated spatial information system for multiobjective land-use planning, *Environment and Planning B* **15**(2): 205–214.

- Diana, M., Quadrifoglio, L. and Pronello, C. (2007). Emissions of demand responsive services as an alternative to conventional transit systems, *Transportation Research Part D: Transport and Environment* **12**(3): 183–188.
- Dijksterhuis, C., Brookhuis, K. a. and De Waard, D. (2010). Effects of steering demand on lane keeping behaviour, self-reports, and physiology. A simulator study, *Accident Analysis & Prevention* (2010): 8–15.
- Dimitrov, L. (2010). The effects of social exclusion and transport in South Africa, *29th Southern African Transport Conference (SATC 2010)*, Pretoria, South Africa, pp. 723–732.
- Dobbs, L. (2005). Wedded to the car: women, employment and the importance of private transport, *Transport Policy* **12**(3): 266–278.
- Dowling, R., Reinke, D., Flannery, A., Ryus, P., Vandehey, M., Petritsch, T., Landis, B., Roupail, N. and Bonneson, J. (2008). Multimodal Level of Service Analysis for Urban Streets, *Technical Report 616*, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Dunphy, R. and Fisher, K. (1996). Transportation, Congestion, and Density: New Insights, *Transportation Research Record: Journal of the Transportation Research Board* **1552**(1): 89–96.
- Duperrex, O., Bunn, F. and Roberts, I. (2002). Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials, *BMJ: British Medical Journal* **324**(7346): 1129–1129.
- Eastman, J. R., Jin, W., Kyem, P. A. K. and Toledano, J. (1995). Raster procedures for multi-criteria/multi-objective decisions, *Photogrammetric Engineering & Remote Sensing* **61**(5): 539–547.
- Eden, W. A. (1947). Studies in Urban Theory. II: Ebenezer Howard and the Garden City Movement, *The Town Planning Review* **19**(3): 123–143.
- Elvik, R. (2001). Area-wide urban traffic calming schemes: a meta-analysis of safety effects, *Accident Analysis & Prevention* **33**(3): 327–36.
- Elvik, R. (2010). A restatement of the case for speed limits, *Transport Policy* **17**(3): 196–204.
- Engel, U. (1990). The Effects of Implementation of 50 Km/H in Urban Areas, *Living and Moving in Cities*, Paris, France.
- Ester, M., Kriegel, H., Sander, J. and Xu, X. (1998). Clustering for mining in large spatial databases, *KI-Journal* **12**(1): 18–24.
- European Conference on Sustainable Cities and Towns (1994). Charter of European Cities and Towns Towards Sustainability, *European Conference on Sustainable Cities and Towns*, Aalborg, Denmark.
- Everitt, B. (1994). *A Handbook of Statistical Analyses Using S-Plus*, 1 edn, Chapman & Hall, London.
- Ewing, R. (1995). Beyond density, mode choice, and single-purpose trips, *Transportation Quarterly* **49**(4): 15–24.
- Ewing, R. H. (1999). *Traffic calming: State of the Practice*, Inst of Transportation Engrs.

- Ewing, R., Schroeder, W. and Greene, W. (2004). School Location and Student Travel Analysis of Factors Affecting Mode Choice, *Transportation Research Record: Journal of the Transportation Research Board* **1895**: 55–63.
- FHWA (1997). Flexibility in Highway Design, *Technical Report FHWA-PD-97-062*, Federal Highway Administration, Washington, D.C.
- FHWA (2007). What is CSS? **URL**: <http://www.fhwa.dot.gov/context/what.cfm>
- Figueira, J., Greco, S. and Ehrgott, M. (2005). *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer Verlag, London.
- Finch, D. J., Kompfner, P., Lockwood, C. R. and Maycock, G. (1994). Speed, Speed Limits and Crashes, *Technical Report PR 58*, Transport Research Laboratory (TRL), Crowthorne, Berkshire.
- Florida Department of Transportation (2009). *2009 Quality / Level of Service Handbook*, Florida Department of Transportation, Tallahassee, Florida.
- Forbes, G. (1999). Urban Roadway Classification, *Urban Street Symposium, Dallas, Texas*, Washington, DC.
- Fraley, C. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis, *The Computer Journal* **41**(8): 578–588.
- Frank, L. D. and Pivo, G. (1995). Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking, *Transportation Research Record: Journal of the Transportation Research Board* **1446**: 44–52.
- Fred, A. L. N. and Jain, A. K. (2005). Combining multiple clusterings using evidence accumulation., *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(6): 835–50.
- Fröhlich, P. (2003). Induced Traffic : Review of the explanatory models, *3rd Swiss Transport Research Conference, Arbeitsbericht Verkehrs- und Raumplanung*, Ascona, Switzerland.
- Frost, L. (1991). *The New Urban Frontier: Urbanisation and City Building in Australasia and the American West*, New South Wales University Press, Sydney, Australia.
- Gainewe, M. and Masangu, N. (2010). Factors leading to fatal crashes and fatalities on the South African roads: 2005-2009, *29th Southern African Transport Conference (SATC 2010)*, Pretoria, South Africa, pp. 2005–2009.
- Galante, F., Mauriello, F., Montella, A., Perneti, M., Aria, M. and D’Ambrosio, A. (2010). Traffic calming along rural highways crossing small urban communities: driving simulator experiment, *Accident Analysis & Prevention* **42**(6): 1585–94.
- Garber, N. J. and Kassebaum, E. A. (2008). Evaluation of Crash Rates and Causal Factors for High-Risk Locations on Rural and Urban Two-Lane Highways in Virginia, *Technical Report VTRC 09-R1*, Virginia Transportation Research Council, Charlottesville, Virginia.
- Garber, S. and Graham, J. (1990). The effects of the new 65 mile-per-hour speed limit on rural highway fatalities: a state-by-state analysis, *Accident Analysis & Prevention* **22**(2): 137–149.
- Gardner, P., Ivan, J. N. and Du, J. (2002). *Traffic Calming of State Highways: Application New England*, Massachusetts Institute of Technology, Boston, MA, USA.

- Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability, A Guide to the Theory of NP-Completeness.*, W. H Freeman, New York, NY.
- Garrard, J., Rose, G. and Lo, S. K. (2008). Promoting transportation cycling for women: the role of bicycle infrastructure, *Preventive Medicine* **46**(1): 55–9.
- Geraghty, A. B., Seifert, W., Preston, T., Holm, C. V., Duarte, T. H. and Farrar, S. M. (2009). Partnership moves community toward complete streets., *American Journal of Preventative Medicine* **37**(6 Suppl 2): S420–7.
- Glasgow, R. E. and King, D. K. (2009). Implications of Active Living by Design for broad adoption, successful implementation, and long-term sustainability, *American Journal of Preventative Medicine* **37**(6 Suppl 2): S450–2.
- Goldenbeld, C. and van Schagen, I. (2007). The credibility of speed limits on 80 km/h rural roads: The effects of road and person(ality) characteristics, *Accident Analysis & Prevention* **39**(6): 1121–30.
- Goldstine, R. (1991). Influence of road width on accident rates by traffic volume, *Transportation Research Record: Journal of the Transportation Research Board* **1318**: 64–69.
- Golob, T. and Hensher, D. (2007). The trip chaining activity of Sydney residents: A cross-section assessment by age group with a focus on seniors, *Journal of Transport Geography* **15**(4): 298–312.
- Goodwin, P. (1999). Transformation of transport policy in Great Britain, *Transportation Research Part A: Policy and Practice* **33**(7-8): 655–669.
- Gordon, P., Kumar, A. and Richardson, H. (1989). Gender Differences in Metropolitan Travel Behaviour, *Regional Studies* **23**(6): 499–510.
- Gower, J. (1967). A comparison of some methods of cluster analysis, *Biometrics* **23**(4): 623–637.
- Greaves, S. and Ellison, A. (2010). Personality, risk aversion and speeding: An empirical investigation, *12th World Conference for Transportation Research*, Vol. 2007, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–16.
- Grossardt, T., Bailey, K. and Brumm, J. (2001). Analytic minimum impedance surface: Geographic information system-based corridor planning methodology, *Transportation Research Record: Journal of the Transportation Research Board* **1768**: 224–232.
- Grosskopf, S. E., Labuschagne, F. J. J. and Moyana, H. (2010). Road safety audits: the way forward, *29th Southern African Transport Conference (SATC 2010)*, Pretoria, South Africa, pp. 576–587.
- Gudmundsson, H. (2001). *Research Notes No. 148: Indicators and performance measures for Transportation, Environment and Sustainability in North America*, Ministry of Environment and Energy, National Environmental Research Institute, Denmark.
- Hadi, M. A., Aruldas, J., Chow, L. F. and Wattleworth, J. A. (1995). Estimating safety effects of cross-section design for various highway types using negative binomial regression, *Transportation Research Record: Journal of the Transportation Research Board* **1500**: 169–177.
- Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques*, Morgan Kaufmann.
- Handy, S. L. (1992). Regional versus local accessibility: neo-traditional development and its implications for non-work travel, *Built Environment* **18**(4): 253–267.

- Handy, S. L. (1996a). Understanding the link between urban form and nonwork travel behavior, *Journal of Planning Education and Research* **15**(3): 183.
- Handy, S. L. (1996b). Urban Form and Pedestrian Choices: Study of Austin Neighborhoods, *Transportation Research Record: Journal of the Transportation Research Board* **1552**: 135–144.
- Handy, S. L., Cao, X. and Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California, *Transportation Research Part D: Transport and Environment* **10**: 427–444.
- Harwood, D. W. (1990). Effective utilization of street width on urban arterials, *Technical Report 330*, National Cooperative Highway Research Program, Washington, D.C.
- Hass-Klau, C. (1993). *The Pedestrian and City Traffic*, John Wiley & Sons Ltd, Oxford, UK.
- Hass-Klau, C., Nold, I. and Bocker, G. (1992). Civilised streets: A guide to traffic calming.
- Hebbert, M. (1999). A City in Good Shape: Town Planning and Public Health, *The Town Planning Review* **70**(4): 433–453.
- Hebbert, M. (2005). Engineering, Urbanism and the Struggle for Street Design, *Journal of Urban Design* **10**(1): 39–59.
- Heimbach C. L., C. P. A. and Chang, M. S. (1983). Some partial consequences of reduced traffic lane widths on urban arterials, *Transportation Research Record: Journal of the Transportation Research Board* **923**: 69–72.
- Highway Research Board (1965). *Highway Research Board Special Report 87: Highway Capacity Manual*, Transportation Research Board, Washington, D.C.
- Hillman, M., Henderson, I. and Whalley, A. (1973). *Personal Mobility and Transport Policy*, Political and Economic Planning, London, UK.
- Hjorthol, R. (2000). Same city - different options. An analysis of the work trips of married couples in the metropolitan area of Oslo, *Journal of Transport Geography* **8**(3): 213–220.
- Hobbs, B. and Meier, P. (1994). Multicriteria methods for resource planning: an experimental comparison, *IEEE Transactions on Power Systems* **9**(4): 1811–1817.
- Holtzclaw, J. (1990). Manhattanization versus Sprawl. How Density Impacts Auto Use Comparing Five Bay Area Communities, *Proceedings of the Eleventh International Pedestrian Conference*, pp. 99–106.
- Hossain, M. S., Hunt, J. and Wirasinghe, S. (2010). Impact of walking distance on attractiveness of transit use, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–21.
- Hummel, T., Mackie, A. and Wells, P. (2002). Traffic calming measures in built-up areas: Literature Review, *Technical Report PR/SE/622/02*, TRL Limited, London, UK.
- Iaych, K., Alexeev, V. and Latipov, O. (2009). *IRF World Road Statistics 2009*, International Road Federation, Geneva. **URL:** [www.irfnet.org](http://www.irfnet.org).
- Investec (Pty) Ltd. (2006). *SA Economic Research - Prospects 2006-2010: Consumption Expenditure Outlook.*, Investec (Pty) Ltd., Cape Town, South Africa.

- Jacobs, G. and Sayer, I. (1983). Road accidents in developing countries, *Accident Analysis & Prevention* **15**(5): 337–353.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means, *Pattern Recognition Letters* **31**(8): 651–666.
- Jain, A. K., Murty, M. N. and Flynn, P. J. (1999). Data clustering: a review, *ACM Computing Surveys* **31**(3): 264–323.
- Janssen, R. and Rietveld, P. (1990). Multicriteria analysis and geographical information systems: an application to agricultural land use in the Netherlands, *Geographical Information Systems for Urban and Regional Planning* pp. 129–139.
- Jobanputra, R. (2010). Quantifying the impact of infrastructure based traffic calming on road safety; A case study in Cape Town, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–25.
- Jones, R. (2004). Context-Sensitive Design: Will the Vision Overcome Liability Concerns?, *Transportation Research Record: Journal of the Transportation Research Board* **1890**(1): 5–15.
- Kane, L. and Del Mistro, R. (2003). Changes in transport planning policy: Changes in transport planning methodology?, *Transportation* **30**(2): 113–131.
- Keeney, R. L. and Raiffa, H. (1976). Decision analysis with multiple conflicting objectives.
- Keeney, R. and Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value trade-offs*, Cambridge University Press.
- Kemperman, A. D. A. M. (2010). Built environment and active travel behaviour of children: results of a bayesian belief network, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–16.
- Keshkamat, S., Looijen, J. and Zuidgeest, M. (2009). The formulation and evaluation of transport route planning alternatives: A spatial decision support system for the Via Baltica project, Poland, *Journal of Transport Geography* **17**(1): 54–64.
- Khalilian, M. and Mustapha, N. (2010). Data Stream Clustering: Challenges and Issues, *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. I, Hong Kong, pp. 17–20. URL: <http://arxiv.org/abs/1006.5261>
- Kim, K., Brunner, I. M. and Yamashita, E. (2008). Modeling violation of Hawaii’s crosswalk law, *Accident Analysis & Prevention* **40**(3): 894–904.
- Kitamura, R., Mokhtarian, P. L. and Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area, *Transportation* **24**(2): 125–158.
- Kleinemas, U. M. and Rudinger, G. (2010). Risk Profiles of Elderly Drivers with Accidents - Implications for Intervention and Policy, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–6.
- Kockelman, K. M. (1996). *Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from the San Francisco Bay Area*, Masters thesis, University of California, Berkeley.

- Kontos, T., Komilis, D. and Halvadakis, C. (2005). Siting MSW landfills with a spatial multiple criteria analysis methodology, *Waste Management* **25**(8): 818–832.
- Kothari, R. (1999). On finding the number of clusters, *Pattern Recognition Letters* **20**(4): 405–416.
- Kouri, C. (1999). *Wait for the Bus: How Low Country School Site Selection and Design Deter Walking to School and Contribute to Urban Sprawl*, South Carolina Coastal Conservation League, Duke University Terry Sanford Institute of Public Policy, Durham, N.C.
- Kuhn, H. W. and Tucker, A. W. (1951). Nonlinear programming, *Proceedings of 2nd Berkeley Symposium*, University of California Press, Berkeley, pp. 481–492.
- Laaribi, A. (1996). A spatial decision aid: A multicriterion evaluation approach, *Computers, Environment and Urban Systems* **20**(6): 351–366.
- Laplante, J. and McCann, B. (2008). Complete Streets: We Can Get There from Here, *ITE Journal* (May).
- Lave, C. and Elias, P. (1994). Did the 65 mph speed limit save lives?, *Accident Analysis & Prevention* **26**(1): 49–62.
- Leccese, M. and McCormick, K. (2000). *Charter of the New Urbanism*, McGraw-Hill Professional.
- Leden, L., Wikström, P. E., Gårder, P. and Rosander, P. (2006). Safety and accessibility effects of code modifications and traffic calming of an arterial road, *Accident Analysis & Prevention* **38**(3): 455–61.
- Levinson, D. M. and Kumar, A. (1994). Multi-modal trip distribution: structure and application, *Transportation Research Record: Journal of the Transportation Research Board* **1446**: 124–131.
- Levinson, D. M. and Kumar, A. (1997). Density and the journey to work, *Growth and Change* **28**(2): 147–72.
- Levinson, H. S. and Wynn, F. H. (1963). Effects of density on urban transportation requirements, *Highway Research Record* (2): 38–64.
- Li, X. (1999). GIS based map overlay method for comprehensive assessment of road environmental impact, *Transportation Research Part D: Transport and Environment* **4**(3): 147–158.
- Liabo, K. (2003). Can traffic calming measures achieve the Children's Fund objective of reducing inequalities in child health?, *Archives of Disease in Childhood* **88**(3): 235–236.
- Lillebye, E. (1996). Architectural and functional relationships in street planning: an historical view, *Landscape and Urban Planning* **35**(2-3): 85–105.
- Limtanakool, N., Dijst, M. and Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips, *Journal of Transport Geography* **14**(5): 327–341.
- Linstone, H. A. and Turoff, M. (1975). *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, Massachusetts.
- Litman, T. (2007). Developing Indicators for Comprehensive and Sustainable Transport Planning, *Transportation Research Record: Journal of the Transportation Research Board* **2017**(1): 10–15.
- Litman, T. and Burwell, D. (2006). Issues in sustainable transportation, *International Journal of Global Environmental Issues* **6**(4): 331.

- Lockwood, I. (1997). ITE Traffic Calming Definition, *ITE Journal* **July**: 22.
- Lucas, K. (2010). The role of transport in the social exclusion of low income populations in South Africa: A scoping study, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–21.
- Mabunda, M. M., Swart, L.-A. and Seedat, M. (2008). Magnitude and categories of pedestrian fatalities in South Africa., *Accident Analysis & Prevention* **40**(2): 586–93.
- Macaulay, J., Gunatillake, T., Tziotis, M., Fildes, B., Corben, B. and Newstead, S. (2004). On-Road Evaluation of Perceptual Countermeasures, *Technical Report CR 219*, Australian Transport Safety Bureau, Victoria, Australia.
- Macdonald, M. (1996). A multi-attribute spatial decision support system for solid waste planning, *Computers, Environment and Urban Systems* **20**(1): 1–17.
- Mackenzie, S., Seedat, M., Swart, L.-A. and Mabunda, M. (2008). Pedestrian injury in South Africa: Focusing intervention efforts on priority pedestrian groups and hazardous places, *Crime, Violence & Injury Prevention in South Africa*, Medical Research Council-University of South Africa, Tygerberg, South Africa.
- Mackett, R. L. (2010). Children's travel behaviour and its implications for their health, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–23.
- Maitra, B., Sikdar, P. K. and Dhingra, S. L. (1999). Modeling congestion on urban roads and assessing level of service, *Journal of Transportation Engineering* **125**(6): 508–514.
- Marshall, S. (2004a). Building on Buchanan: Evolving road hierarchy for today's streets-oriented design agenda, *Proceedings of the 2004 European Transport Conference*, Association for European Transport, London, pp. 1–16.
- Marshall, S. (2004b). D1.1: A First Theoretical Approach to Classification of Arterial Streets, *Arterial Streets Towards Sustainability*, University of Westminster, London, UK.
- Marshall, S. (2005). *Streets and Patterns*, Routledge.
- Martens, M., Comte, S. and Kaptein, N. (1997). Contract No RO-96-SC.202 - The Effects of Road Design on Speed Behaviour : A Literature Review, *Technical Report 2.3.1*, Transport RTD Programme of the 4th Framework Programme, European Commission.
- Mathew, T. V. and Rao, K. V. K. (2007). Trip generation, *Introduction to Transportation Engineering*, National Programme on Technology Enhanced Learning, Chennai, India.
- Matzopoulos, R., Myers, J. E. and Jobanputra, R. (2008). Road traffic injury: prioritising interventions., *South African Medical Journal* **98**(9): 692–696.
- McCann, B. and Rynne, S. (2010). Complete Streets: Best Policy and Implementation Practices, *Technical Report 559*, American Planning Association: Planning Advisory Service.
- McKnight, A. and Klein, T. M. (1990). Relationship of 65 mph limit to speeds and fatal accidents, *Transportation Research Record: Journal of the Transportation Research Board* **1**(1281): 71–77.
- McLean, J. (1997). Review of accidents and urban arterial cross-section treatments, *Technical Report 309*, ARRB Transport Research.

- McNally, M. (2000). The Four-Step Model, in D. Hensher and K. Button (eds), *Handbook of Transport Modelling Vol. 1*, Pergamon.
- Messenger, T. and Ewing, R. (1996). Transit-Oriented Development in the Sun Belt, *Transportation Research Record* **1552**(1): 145–153.
- Miller, E. and Ibrahim, A. (1998). Urban Form and Vehicular Travel: Some Empirical Findings, *Transportation Research Record* **1617**(1): 18–27.
- Milligan, G. and Cooper, M. (1985). An examination of procedures for determining the number of clusters in a data set, *Psychometrika* **50**(2): 159–179.
- Mohammed, S. and Labuschagne, F. (2008). Can draconian law enforcement solve the South African road safety crisis?, *27th Southern African Transport Conference (SATC 2008)*, Pretoria, South Africa, pp. 456–468.
- Mohan, D. (2008). Road safety challenges and associated policy issues, *27th Southern African Transport Conference (SATC 2008)*, Pretoria, South Africa.
- Mok, J.-H., Landphair, H. C. and Naderi, J. R. (2006). Landscape improvement impacts on roadside safety in Texas, *Landscape and Urban Planning* **78**(3): 263–274.
- Moore, V., Dolinis, J. and Woodward, A. (1995). Vehicle speed and risk of a severe crash, *Epidemiology* **6**(3): 258–262.
- Moreno, D. and Seigel, M. (1988). A GIS approach for corridor siting and environmental impact analysis, *GIS/LIS*, Vol. 88, pp. 507–514.
- Morris, E. (2007). From Horse Power to Horsepower, *Access* **30**(Spring): 2–9.
- Morris, S. S. (1969). Cape Town: Metropolis in the making: A decade of development, *The Town Planning Review* **40**(2): 102–118.
- Morrison, D., Petticrew, M. and Thomson, H. (2003). What are the most effective ways of improving population health through transport interventions? Evidence from systematic reviews, *Journal of Epidemiology and Community Health* **57**(5): 327.
- Moughtin, C. (1991). European city Part 2 : Relating and function, *The Town Planning Review* **62**(2): 153–199.
- Nakamura, H., Suzuki, K. and Ryu, S. (2000). Analysis of the Interrelationship Among Traffic Flow Conditions, Driving Behavior, and Degree of Driver's Satisfaction on Rural Motorways, *Proceedings of the Fourth International Symposium on Highway Capacity*, Transportation Research Board, Washington, D.C., pp. 42–52.
- Naude, W. (2008). Is there a spatial mismatch in South Africa's metropolitan labour market?, *Cities* **25**(5): 268–276.
- NCSC (2010). Complete streets: Complete Streets FAQ. Online; accessed 03 September 2010. **URL:** <http://www.completestreets.org>
- NDOT (1996). *White Paper On National Transport Policy*, South African National Department of Transport, Pretoria, South Africa.

- NDoT (1999). *Moving South Africa: Draft for Discussion*, South African National Department of Transport, Pretoria, South Africa.
- NDoT (2005a). *Key Results of The National Household Travel Survey*, South African National Department of Transport, Pretoria, South Africa.
- NDoT (2005b). *Technical Report: The First South African National Household Travel Survey 2003*, Department of Transport.
- NDoT (2006a). *National Land Transport Strategic Framework*, South African National Department of Transport, Pretoria, South Africa.
- NDoT (2006b). *National Road Safety Strategy*, South African National Department of Transport, Pretoria, South Africa.
- Neuman, T. R., Schwartz, M., Clark, L., Bednar, J., Forbes, D., Vomacka, D., Taggart, C., Glynn, M., Slack, K. and Abere, D. (2002). A guide to best practices for achieving context sensitive solutions, *Technical Report 480*, Transportation Research Board, Washington, D.C.
- Newman, P. and Kenworthy, J. (2000). The Ten Myths of Automobile Dependence, *World Transport Policy & Practice* **6**(1): 15–25.
- Newman, P., Kenworthy, J. and Vintila, P. (1995). Can we overcome automobile dependence? Physical planning in an age of urban cynicism, *Cities* **12**(1): 53–65.
- Newstead, S. and Mullan, N. (1996). *Evaluation of the Crash Effects of the Changes in Speed Zones in Victoria During 1993-1994*, Victoria Roads Corporation, Victoria, Australia.
- Ng, R. T. and Han, J. (1994). Efficient and effective clustering methods for spatial data mining, *Proceedings of the 20th VLDB Conference*, Santiago, Chile, pp. 144–155.
- NHTSA (1988). Fatal Accident Reporting System: 1987, *Technical Report DOT-HS-807-360*, National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington D.C.
- Nielsen, T. S. and Hovgesen, H. H. (2004). How do motorways shape commuting patterns? An evaluation based on time series., *European Transport Conference*.
- Nilsson, G. (1990). Reduction of the 110 km/h Speed Limit to 90 km/h During Summer 1989. Effects on Personal Injury Accidents, Injured and Speeds, *Technical Report VTI-358*, National Swedish Road & Traffic Research Institute, Linköping, Sweden.
- Noland, R. B. (2001). Relationships between highway capacity and induced vehicle travel, *Transportation Research Part A: Policy and Practice* **35**(1): 47–72.
- Norley, K. and Peters, A. (2010). Towards an understanding of the impact of deferring transit infrastructure implementation, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–23.
- Nyarirangwe, M. (2009). A causal-based multi-sectoral approach to the prevention of road traffic accidents and prioritization of response measures in South Africa, *28th Southern African Transport Conference (SATC 2009)*, Pretoria, South Africa.
- O’Flaherty, C. (1997). *Transport Planning and Traffic Engineering*, Butterworth-Heinemann.

- Ogle, J. H. (2005). *Quantitative Assessment of Driver Speeding Behavior using Instrumented Vehicles*, PhD thesis, Georgia Institute of Technology.
- Olszak, L., Goldbach, R. and Long, J. (2008). Do Context-Sensitive Solutions Really Work?, *Transportation Research Record: Journal of the Transportation Research Board* **2060**(-1): 107–115.
- Openshaw, S., Carver, S. and Fernie, J. (1989). *Britain's Nuclear Waste: Siting and Safety*, Belhaven Press, London, UK.
- Oppenheim, N. (1995). *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*, John Wiley & Sons, New-York, USA.
- Owens, S. (1995). From 'predict and provide' to 'predict and prevent'?: Pricing and planning in transport policy, *Transport Policy* **2**(1): 43–49.
- Pant, P., Adhami, J. and Niehaus, J. (1992). Effects of the 65-mph speed limit on traffic accidents in Ohio, *Transportation Research Record: Journal of the Transportation Research Board* **1**(1375): 53–60.
- Park, S. H., Kim, D.-K., Kho, S.-Y. and Rhee, S. (2010). Identifying hazardous locations based on severity scores of highway crashes, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–17.
- Parkany, E. (2006). Teaching Context-Sensitive Design to Undergraduates, *Transportation Research Record: Journal of the Transportation Research Board* **1984**(-1): 96–101.
- Parker, M. R. (1985). Synthesis of Speed Zoning Practices, *Technical Report FHWA/RD-85/096*, Federal Highway Administration, Washington, D.C.
- Parker, M. R. (1997). *Effects of Raising and Lowering Speed Limits on Selected Roadway Sections*, US Dept. of Transportation, Federal Highway Administration, Springfield, Va.
- Parsons Brinckerhoff Quade and Douglas Inc. (1996). Transit and Urban Form, *Technical Report 16*, Transit Cooperative Research Program, Washington D.C.
- Patras, A., Brunton, N. P., Downey, G., Rawson, A., Warriner, K. and Gernigon, G. (2011). Application of principal component and hierarchical cluster analysis to classify fruits and vegetables commonly consumed in Ireland based on in vitro antioxidant activity, *Journal of Food Composition and Analysis* **24**(2): 250–256.
- Pau, M. and Angius, S. (2001). Do speed bumps really decrease traffic speed? An Italian experience, *Accident Analysis & Prevention* **33**(5): 585–97.
- Paulus, G., Krch, M., Scholz, J. and Bachhiesl, P. (2006). Scenario-based spatial decision support for network infrastructure design, *VisASDS Workshop 2006* .
- Pearson, C. S. (1985). *Down to business: multinational corporations, the environment and development*, World Resources Institute, Washington, D.C.
- Peltola, H. (2000). Seasonally changing speed limits: Effects on speeds and accidents, *Transportation Research Record: Journal of the Transportation Research Board* **1734**(1): 46–51.
- Perry, C., Heydecker, W., Goodrich, E., Harrison, S., Adams, T., Bassett, E. and Whitten, R. (1929). *Neighborhood and Community Planning*, Regional Plan of New York and Its Environs.

- Peska, P. and Venter, C. (2009). A relook at residential trip generation variables, *28th Southern African Transport Conference (SATC 2009)*, Pretoria, South Africa, pp. 538–548.
- Peterson, J. A. (1979). The Impact of Sanitary Reform upon American Urban Planning, 1840-1890, *Journal of Social History* **13**(1): 83–103.
- Pfefer, R. (1999). Toward Reflecting Public Perception of Quality of Service in Planning, Designing, and Operating Highway Facilities, *Transportation Research Record: Journal of the Transportation Research Board* **1685**(1): 81–89.
- Plowden, S. (1972). *Towns Against Traffic*, Deutsch.
- Polk, M. (2003). Are women potentially more accommodating than men to a sustainable transportation system in Sweden?, *Transportation Research Part D: Transport and Environment* **8**(2): 75–95.
- Potts, I. B., Harwood, D. W. and Richard, K. R. (2007). Relationship of Lane Width to Safety on Urban and Suburban Arterials, *Transportation Research Record: Journal of the Transportation Research Board* **2023**(1): 63–82.
- Power, D. J. (2003). A brief history of decision support systems. **URL:** <http://dssresources.com/history/dsshistory.html>
- Pucher, J. and Dijkstra, L. (2003). Promoting safe walking and cycling to improve public health: lessons from The Netherlands and Germany, *American Journal of Public Health* **93**(9): 1509–16.
- Pucher, J., Dill, J. and Handy, S. L. (2010). Infrastructure, programs, and policies to increase bicycling: An international review, *Preventive Medicine* **50**(Suppl 1): S106–25.
- Pucher, J. and Renne, J. L. (2003). Socioeconomics of Urban Travel : Evidence from the 2001 NHTS, *Transportation Quarterly* **57**(3): 49–78.
- Pushkarev, B. and Zupan, J. (1977). *Public Transportation and Land Use Policy*, Indiana University Press.
- Reason, J. T. (1990). *Human Error*, Cambridge University Press.
- Republic of South Africa (2000). National Land Transport Transition Act.
- Republic of South Africa (2009). National Land Transport Act.
- Rescia, A. J., Astrada, E. N., Bono, J., Blasco, C. A., Meli, P. and Adámoli, J. M. (2006). Environmental analysis in the selection of alternative corridors in a long-distance linear project: a methodological proposal, *Journal of Environmental Management* **80**(3): 266–78.
- Retting, R. A., Ferguson, S. A. and McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes, *American Journal of Public Health* **93**(9): 1456–63.
- Richardson, B. (2005). Sustainable transport: analysis frameworks, *Journal of Transport Geography* **13**(1): 29–39.
- Richardson, T. (2001). The Pendulum Swings Again: In Search of New Transport Rationalities, *The Town Planning Review* **72**(3): 299–319.
- Rockey, J. (1983). From Vision to Reality: Victorian Ideal Cities and Model Towns in the Genesis of Ebenezer Howard's Garden City, *The Town Planning Review* **54**(1): 83–105.

- Romero, C. (1991). *Handbook of Critical Issues in Goal Programming*, Pergamon Press, Oxford.
- Root, A., Schintler, L. and Button, K. (2002). Women and Travel: The Sustainability Implications of Changing Roles', in W. Black and P. Nijkamp (eds), *Social Change and Sustainable Transport*, Indiana University Press, Bloomington.
- Rospabe, S. and Selod, H. (2006). Does city structure cause unemployment? The case of Cape Town, in H. Bhorat and R. Kanbur (eds), *Poverty and Policy in Post-Apartheid South Africa*, HSRC Press, Cape Town, South Africa.
- Roy, B. (1968). Classement et choix en présence de points de vue multiples (la méthode ELECTRE), *la Revue d'Informatique et de Recherche Opérationnelle (RIRO)* **8**: 57–75.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Dordrecht.
- Royal Commission on Environmental Pollution (1994). Transport and the Environment, *Technical Report 18*, Royal Commission on Environmental Pollution.
- Royal Commission on Environmental Pollution (1997). Transport and the Environment, *Technical Report 20*, Royal Commission on Environmental Pollution.
- RTMC (2009). *Road Traffic Report - 30 September 2009*, Road Traffic Management Corporation, Pretoria, South Africa.
- Russell, J. (1990). Traffic Calming and Town Planning, *The Town Planning Review* **61**(2): 3–6.
- Rust, F. C., Mccutcheon, R. T. and Coetzee, L. (2008). Transport Research: Quo Vadis?, *27th Southern African Transport Conference (SATC 2008)*, Vol. 0001, Pretoria, South Africa, pp. 50–67.
- Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, McGraw-Hill.
- Saaty, T. L. (2005). Making and validating complex decisions with the AHP/ANP, *Journal of Systems Science and Systems Engineering* **14**(1): 1–36.
- Sallis, J. E. and Glanz, K. (2011). The Role of Built Environments in Physical Activity, Eating, and Obesity in Childhood, *Obesity* **16**(1): 89–108.
- Salvador, S. and Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms, *16th IEEE International Conference on Tools with Artificial Intelligence*, IEEE Computer Society, pp. 576–584.
- SANRAL (2009). *Design Guidelines for Single Carriageway National Roads*, South African National Roads Agency Limited, Pretoria, South Africa.
- Schaeffer, K. and Sclar, E. (1980). *Access for All: Transportation and Urban Growth*, Columbia University Press.
- Scharping, F. K. (1994). 30 Km/h Speed Limited Zones in Hamburg. Speed Reduction Measures on Major Inner City Roads.
- Schensul, D. and Heller, P. (2011). Legacies, change and transformation in the post-apartheid city: Towards an urban sociological cartography, *International Journal of Urban and Regional Research* **35**(1): 78–109.

- Schermers, G. and Theyse, H. (1996). *National Guidelines for Traffic Calming*, CSIR, Pretoria, South Africa.
- Scheunders, P. (1997). A comparison of clustering algorithms applied to color image quantization, *Pattern Recognition Letters* **18**(11-13): 1379–1384.
- Schimek, P. (1996). Household Motor Vehicle Ownership and Use: How Much Does Residential Density Matter?, *Transportation Research Record: Journal of the Transportation Research Board* **1552**(1): 120–125.
- Scoones, I. (1998). *IDS Working Paper 72: Sustainable Rural Livelihoods : A Framework for Analysis*, Institute of Development Studies, University of Sussex, Brighton, UK.
- Sharifi, M. A., Boerboom, L., Shamsudin, K. B. and Veeramuthu, L. (2006). Spatial multiple criteria decision analysis in integrated planning for public transport and land use development study in Klang Valley, Malaysia, *ISPRS Technical Commission II Symposium, Vienna, Vienna*.
- Shimizu, T., Ojiro, F. and Phan, T. T. (2010). A study on survey method for definition of walking distance, *12th World Conference for Transportation Research*, Vol. 1998, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–10.
- Sliogeris, J. (1992). 110 Kilometre Per Hour Speed Limit: Evaluation of Road Safety Effects, *Technical Report GR 92-8*, Victoria Roads Corporation, Victoria, Australia.
- Smith, W. S. (1983). Mass Transport for High-Rise High-Density Living, *Journal of Transportation Engineering* **110**(6): 521.
- Smyth, P. (1996). Clustering using Monte Carlo cross-validation, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 126–133.
- Solesbury, W. and Townsend, A. (1970). Transportation Studies and British Planning Practice, *The Town Planning Review* **41**(1): 63–79.
- South African National Treasury (2009). *Provincial Budgets and Expenditure Review: 2005/06 - 2011/12*, South African National Treasury, Pretoria, South Africa.
- Southworth, M. and Ben-Joseph, E. (2003). *Streets and the Shaping of Towns and Cities*, Island Press, Washington, D.C.
- Stamatiadis, N., Adam, K., Don, H., Theodore, H. and Jerry, P. (2009). Quantifying the Benefits of Context Sensitive Solutions, *Technical Report 642*, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Stander, H., Kruger, P., Coetzee, J. and Lamprecht, T. (1995). *South African Trip Generation Rates*, 2nd edn, South African National Department of Transport, Pretoria, South Africa.
- Statistics South Africa (2003). Census in brief, *Technical Report 03-02-03 (2001)*, Statistics South Africa, Pretoria, South Africa.
- Statistics South Africa (2008). *Gross Domestic Product: Annual estimates: 1993 - 2007*, Statistics South Africa, Pretoria, South Africa.
- Stead, D. (2001). Relationships between land use, socioeconomic factors, and travel patterns in Britain, *Environment and Planning B: Planning and Design* **28**(4): 499–528.

- Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV) (2004). *SWOV Fact Sheet - Zone 30: Urban Residential Areas*, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, Leidschendam, The Netherlands.
- Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV) (2005). *Advancing Sustainable Safety*, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV), Leidschendam, The Netherlands.
- Streff, F. M., Wagenaar, A. C. and Schultz, R. H. (1990). Reductions in police-reported injuries associated with Michigan's safety belt law, *Journal of Safety Research* **21**(1): 9–18.
- Stuster, J., Coffman, Z. and Warren, D. (1998). FHWA-RD-98-154: Synthesis of safety research related to speed and speed management, *Technical Report 1*, Federal Highway Administration.
- Svensson, A. s. (ed.) (2004). *Arterial Streets for People: Guidance for Planners and Decision Makers when Reconstructing Arterial Streets.*, European Commission Fifth Framework Programme. **URL:** [www.tft.lth.se/artists](http://www.tft.lth.se/artists)
- Tan, P.-N., Steinbach, M. and Kumar, V. (2005). *Introduction to Data Mining*, Addison-Wesley, Boston, Massachusetts.
- Taylor, D. (2002). Highway Rules, *Urban Design Quarterly* (81).
- Tester, J. M., Rutherford, G. W., Wald, Z. and Rutherford, M. W. (2004). A matched case-control study evaluating the effectiveness of speed humps in reducing child pedestrian injuries, *American Journal of Public Health* **94**(4): 646–50.
- Thill, J. C. (2000). Geographic information systems for transportation in perspective, *Transportation Research Part C: Emerging Technologies* **8**(1-6): 3–12.
- Thomaidis, F. and Mavrakakis, D. (2006). Optimum route of the south transcontinental gas pipeline in SE Europe using AHP, *Journal of Multi-Criteria Decision Analysis* **14**(1-3): 77–88.
- Townsend, T., Lane, L. and Martin, J. (2005). Context-Sensitive Solutions Training Experience, *Transportation Research Record: Journal of the Transportation Research Board* **1904**: 67–74.
- Transportation Research Board (1994). *Special Report 209: Highway Capacity Manual*, 3rd edn, Transportation Research Board, Washington, D.C.
- Tripp, H. (1942). *Town Planning and Road Traffic*, E. Arnold & Co., London.
- Turok, I. (2001). Persistent polarisation post- apartheid? Progress towards urban integration in Cape Town, *Urban Studies* **38**(13): 2349–2377.
- Unal, A., Roupail, N. and Frey, H. (2003). Effect of arterial signalization and level of service on measured vehicle emissions, *Transportation Research Record: Journal of the Transportation Research Board* **1842**(1): 47–56.
- United States Code (1969). National Environmental Policy Act of 1969.
- Van Der Horst, A. (1983). A behavioural study in relation to bicycle traffic: Demonstration project on redesigning urban areas (Eindhoven and Rijswijk), *Technical Report IZF 1983 C-11*, TNO Human Factors Research Institute, Soesterberg, The Netherlands.

- Van Exel, N. and Rietveld, P. (2009). Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travellers on the main travel corridors to the city of Amsterdam, The Netherlands, *Transportation Research Part A: Policy and Practice* **43**(4): 374–385.
- Vasko, K. and Toivonen, H. (2002). Estimating the number of segments in time series data using permutation tests, *2002 IEEE International Conference on Data Mining*, IEEE Comput. Soc, pp. 466–473.
- Vernick, J. S., Li, G., Ogaitis, S., Mackenzie, E. J., Baker, S. P. and Gielen, A. C. (1999). Effects of High School Driver Education on Motor Vehicle Crashes, Violations, and Licensure, *American Journal of Preventative Medicine* **16**(1S).
- Visser, D., Steynberg, M., van Biljon, B. and Scheepers, J. (2003). GAUTRANS Guidelines for the provision of pedestrian and bicycle facilities on provincial roads in Gauteng, *22nd Southern African Transport Conference (SATC2003)*, Pretoria, South Africa.
- Voogd, H. (1982). Multicriteria evaluation with mixed qualitative and quantitative data, *Environment and Planning B: Planning and Design* **9**(2): 221–236.
- Vuchic, V. R. (1999). *Transportation for Livable Cities*, Center for Urban Policy Research.
- Walton, W. and Shaw, J. (2003). Applying the new appraisal approach to transport policy at the local level in the UK, *Journal of Transport Geography* **11**(1): 1–12.
- West, R. (1998). The Effect of Speed Cameras on Injuries from Road Accidents: The Technology Exists to Ensure All Road Traffic Is Regulated to Safe Speeds, *BMJ: British Medical Journal* **316**(7124): 5–6.
- WHO (2004). *World Report on Road Traffic Injury Prevention*, World Health Organization, Geneva, Switzerland.
- Wicking-Baird, M., De Villiers, M. and Dutkiewicz, R. (1997). Cape Town brown haze study, *Technical Report GEN 182*, Energy Research Institute, University of Cape Town.
- Wilkinson, P. (2000). City profile: Cape Town, *Cities* **17**(3): 195–205.
- Withers, M. M. and Bester, C. (2009). Trip generation rates for South African golf clubs and estates, *28th Southern African Transport Conference (SATC 2009)*, Pretoria, South Africa, pp. 549–559.
- Wohl, A. S. (1983). *Endangered Lives : Public health in Victorian Britain*, Dent, London.
- Wu, J. D., Milton, D. K., Hammond, S. K. and Spear, R. C. (1999). Hierarchical cluster analysis applied to workers' exposures in fiberglass insulation manufacturing, *The Annals of Occupational Hygiene* **43**(1): 43–55.
- Xianghai, M., Mei, L. and Zhiqiang, G. (2010). Research on identification of freeway black spots and accident induced factors, *12th World Conference for Transportation Research*, World Conference on Transport Research Society, Lisbon, Portugal, pp. 1–11.
- Yagar, S. (1984). Predicting speeds for rural 2-lane highways, *Transportation Research Part A: General* **18**(1): 61–70.
- Yoe, C. (2002). *Trade-off Analysis Planning and Procedures Guidebook*, U.S. Army Institute for Water Resources, Alexandria, VA.

- Yusof, K. and Baban, S. (2004). Least-cost pipeline path to the Langkawi Island, Malaysia using a geographical information system (GIS), *Proceedings of the 7th Annual International Map India Conference*, New Delhi, India.
- Zaidel, D. M. (2002). Contract No: RO-98-RS.3047 - The Impact of Enforcement on Accidents, *Technical Report 3 (WP2)*, Transport RTD Programme of the 4th Framework Programme, European Commission.
- Zein, S., Geddes, E., Hemsing, S. and Johnson, M. (1997). Safety Benefits of Traffic Calming, *Transportation Research Record: Journal of the Transportation Research Board* **1578**(1): 3–10.
- Zhang, M., Sun, Q., Chen, J. and Guo, J. (2008). Travel Behavior Analysis of the Females in Beijing, *Journal of Transportation Systems Engineering and Information Technology* **8**(2): 19–26.
- Zhang, T., Ramakrishnan, R. and Livny, M. (1996). BIRCH: An efficient data clustering method for very large databases, *SIGMOD 96*, Vol. 1, pp. 103–114.
- Zhao, F., Li, M.-t., Chow, L.-f., Gan, A., Shen, L. D. and Street, W. F. (2002). *NCTR 392-07, 416-03: FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access*, National Center For Transit Research (NCTR), Tampa, FL.
- Zucca, A., Sharifi, A. M. and Fabbri, A. G. (2008). Application of spatial multi-criteria analysis to site selection for a local park: a case study in the Bergamo Province, Italy, *Journal of Environmental Management* **88**(4): 752–69.
- Zuidegeest, M. H. P. (2005). *Sustainable Urban Transport Development: A Dynamic Optimisation Approach*, Phd thesis, University of Twente.

# Appendices

Appendix A: Planning, Policy and Practice	231
Appendix B: Transportation and Road Safety	233
Appendix C: Multiple Criteria Decision Making	235
Appendix D: Data Sources and Case Study Selection	237
Appendix E: Case Study Results and Assessment	247
Appendix F: Cluster Analysis	251
Appendix G: Infrastructure and Context	261

University of Cape Town



## Appendix A: Planning, Policy and Practice



Figure A.1: The 'Radburn Plan' for Radburn, New Jersey. *Source:* Adapted from Birch (1980)



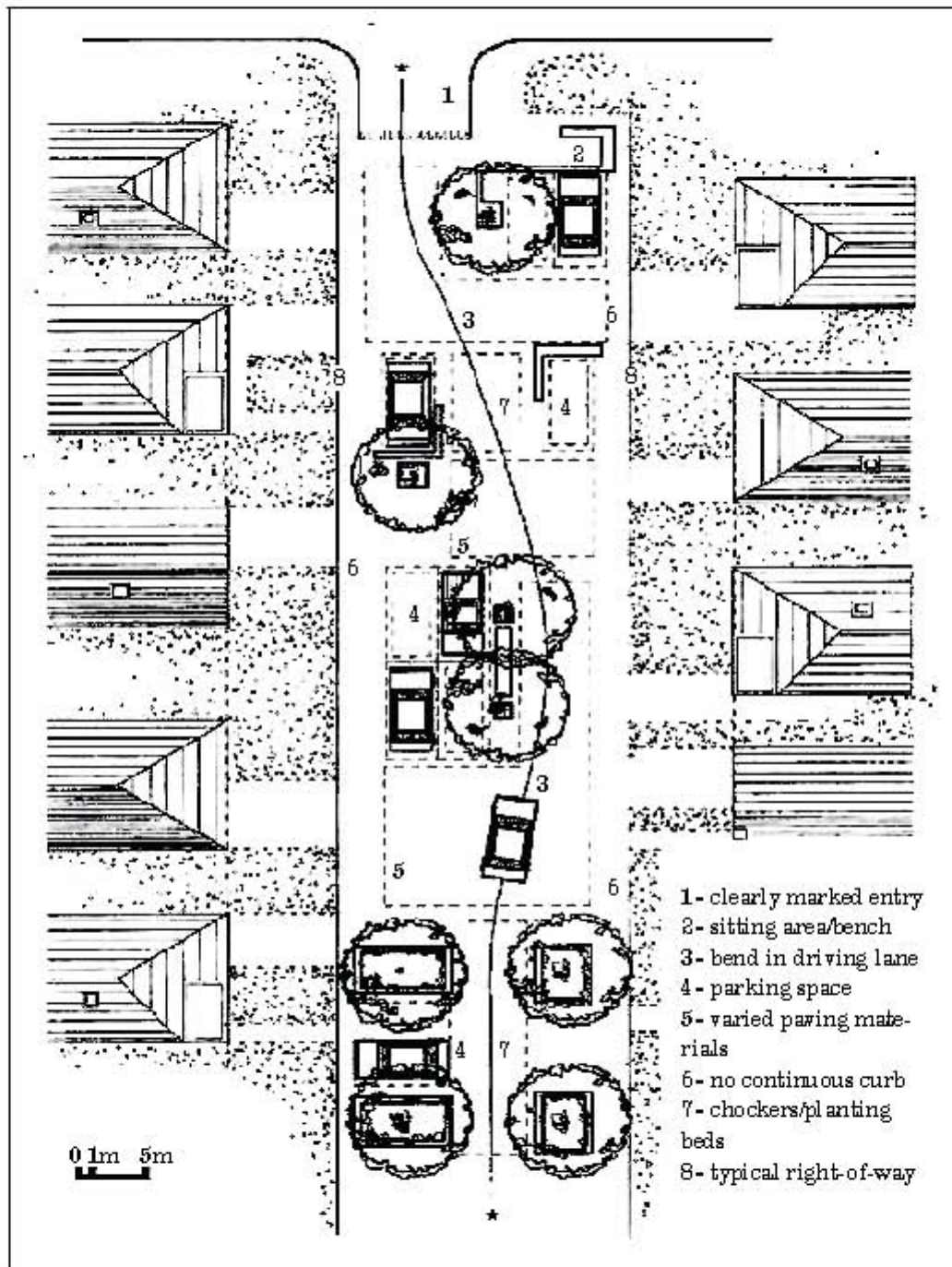
## Appendix B: Transportation and Road Safety

**Table B.1:** Operational Effects of Freeway Lane Widths. *Source:* Federal Highway Administration. (2011) Lane Width. Retrieved January 15, 2011, from <http://safety.fhwa.dot.gov.htm>

Lane width (m)	Reduction in Free-Flow Speed (km/h)
3.6	0
3.5	1
3.4	2.1
3.3	3.1
3.2	5.6
3.1	8.1
3.0	10.6

**Table B.2:** Operational Effects of Lane and Shoulder Width on Two-Lane Highways. *Source:* Federal Highway Administration. Retrieved January 15, 2011, from <http://safety.fhwa.dot.gov.htm>

Lane width (m)	Reduction in Free-Flow Speed (km/h)			
	Shoulder Width (m)			
	$\geq 0.0 < 0.6$	$\geq 0.6 < 1.2$	$\geq 1.2 < 8$	$\geq 1.8$
2.7 < 3.0	10.3	7.7	5.6	3.5
$\geq 3.0 < 3.3$	8.5	5.9	3.8	1.7
$\geq 3.3 < 3.6$	7.5	4.9	2.8	0.7
$\geq 3.6$	6.8	4.2	2.1	0



**Figure B.1:** A typical layout of a woonerf. Extensive use of vegetation, creating parking bays that jut out into the street and limiting the line of sight of the driver by creating sharp chicanes with the street furniture forced vehicles to slow down to walking pace in order to avoid all the obstacles. *Source:* Adapted from Southworth and Ben-Joseph (2003)

# Appendix C: Multiple Criteria Decision Making

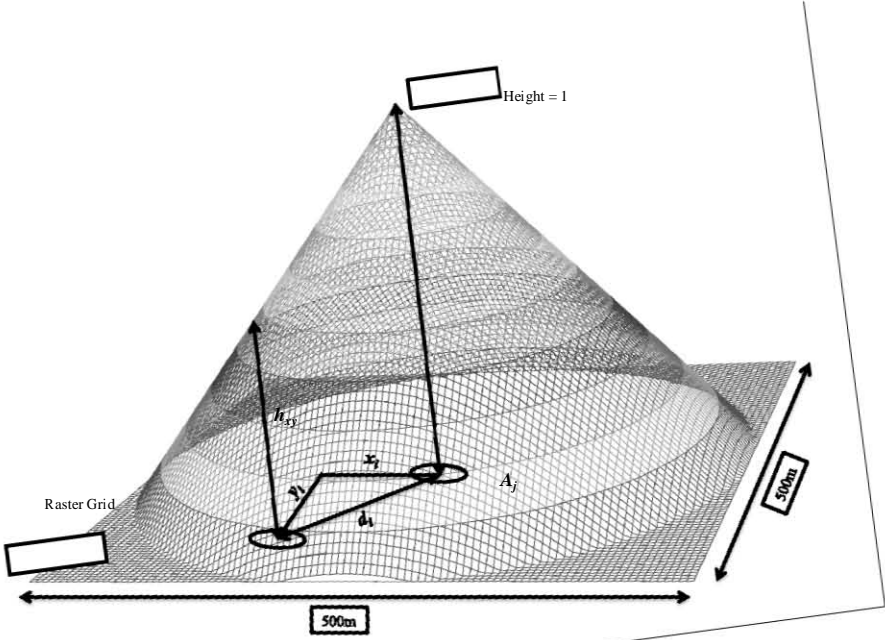
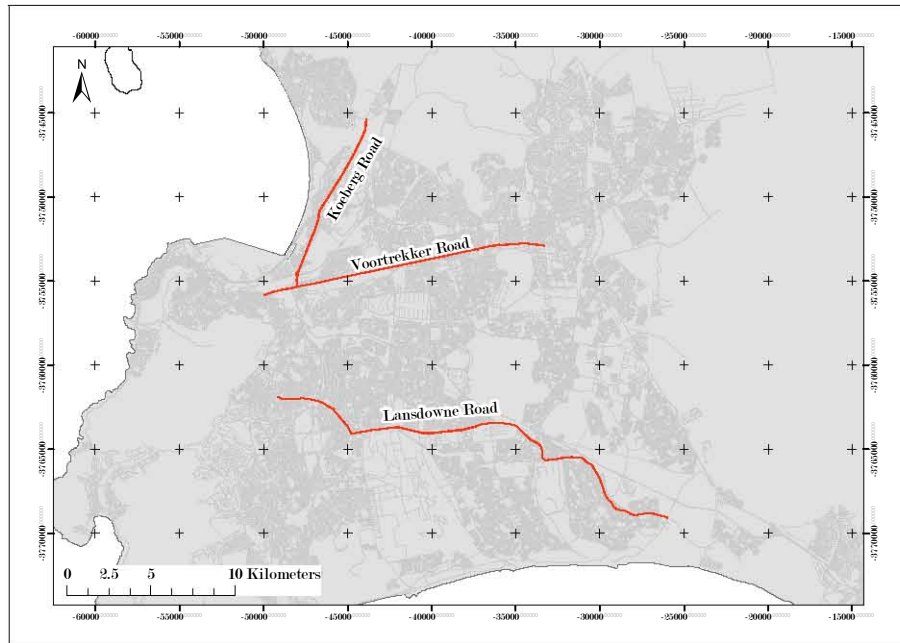


Figure C.1: Conical distance weighting

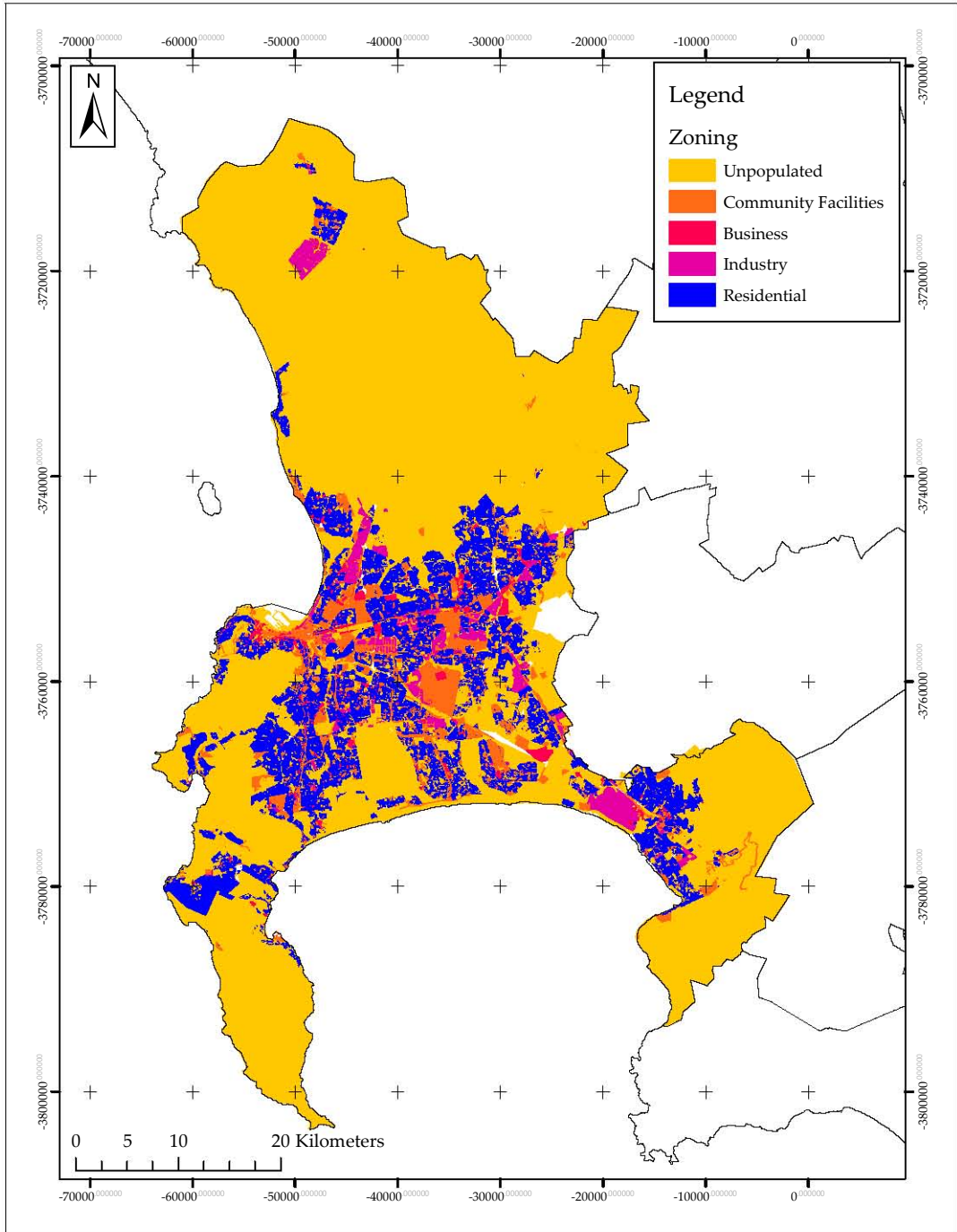
**Table C.1:** Themes, criteria and the explanation for use in the Via Baltica corridor study.  
 Source: Adapted from Keshkamat et al. (2009)

Theme	Criteria	Explanation
Transport efficiency	Proximity to existing rail network	<i>Spatial benefit.</i> The closer the expressway is built to an existing rail network, the better the future intermodality
	Proximity to the proposed Rail Baltica	<i>Spatial benefit.</i> The closer the expressway is built to the proposed rail route, the better the future intermodality
	Current traffic density	<i>Spatial benefit.</i> The higher the current traffic density, the more is the reason to upgrade the road
Ecology	Internationally protected natural areas (Natura 2000 sites)	<i>Spatial constraint.</i> Natura 2000 sites are strictly protected under EU regulations
	Nationally protected areas, such as National and Landscape Parks (and Reserves)	<i>Spatial cost.</i> May be passed through but at a high cost
	Forests and semi natural areas	<i>Spatial cost</i>
	Wetlands and peat bogs Water courses and lakes	<i>Spatial cost</i> <i>Spatial cost</i>
Social impact and safety	Proximity to urban areas	<i>Spatial benefit.</i> The closer the route is to an urban area, the greater the accessibility
	Risk of accidents in urban areas	<i>Spatial cost.</i> The closer the route is to an urban area, the greater are the incidences where resettlement of homes and establishments will be required
	Population served	<i>Spatial benefit.</i> The larger the population served, the more reasons to upgrade the road
	Hazardous areas	<i>Spatial cost.</i> The closer it is to a hazard prone area, the more will be the cost associated with providing safety features
Economic costs and benefits	Current agriculture land use	<i>Spatial cost.</i> Current livelihood
	Economic zones	<i>Spatial benefit.</i> The more the economic activity in the area, the more reasons to upgrade the road
	Best agricultural soils	<i>Spatial cost.</i> Potentially productive areas
	Current status of the road (Category of the road)	<i>Spatial benefit.</i> The higher the current category of the road, the lower will be the engineering cost of upgrading it
	Intersections with water bodies	<i>Spatial cost.</i> Bridges, viaducts, culverts etc involve the construction of expensive structures. Also, the longer the bridge, the higher the cost
	Intersections with secondary roads	<i>Spatial cost.</i> All intersections with secondary roads need to be upgraded. This involves the construction of expensive structures such as flyovers
	Problem soils for construction	<i>Spatial cost.</i> Soils like peat are prone to differential settlement and pose a potentially high construction cost and/or a high maintenance cost
Ancillary structures for urban areas	<i>Spatial cost.</i> The closer the route is to an urban area, the higher will be the engineering costs associated with building acoustic barriers, pedestrian subways and other ancillary structures	

## Appendix D: Data Sources and Case Study Selection



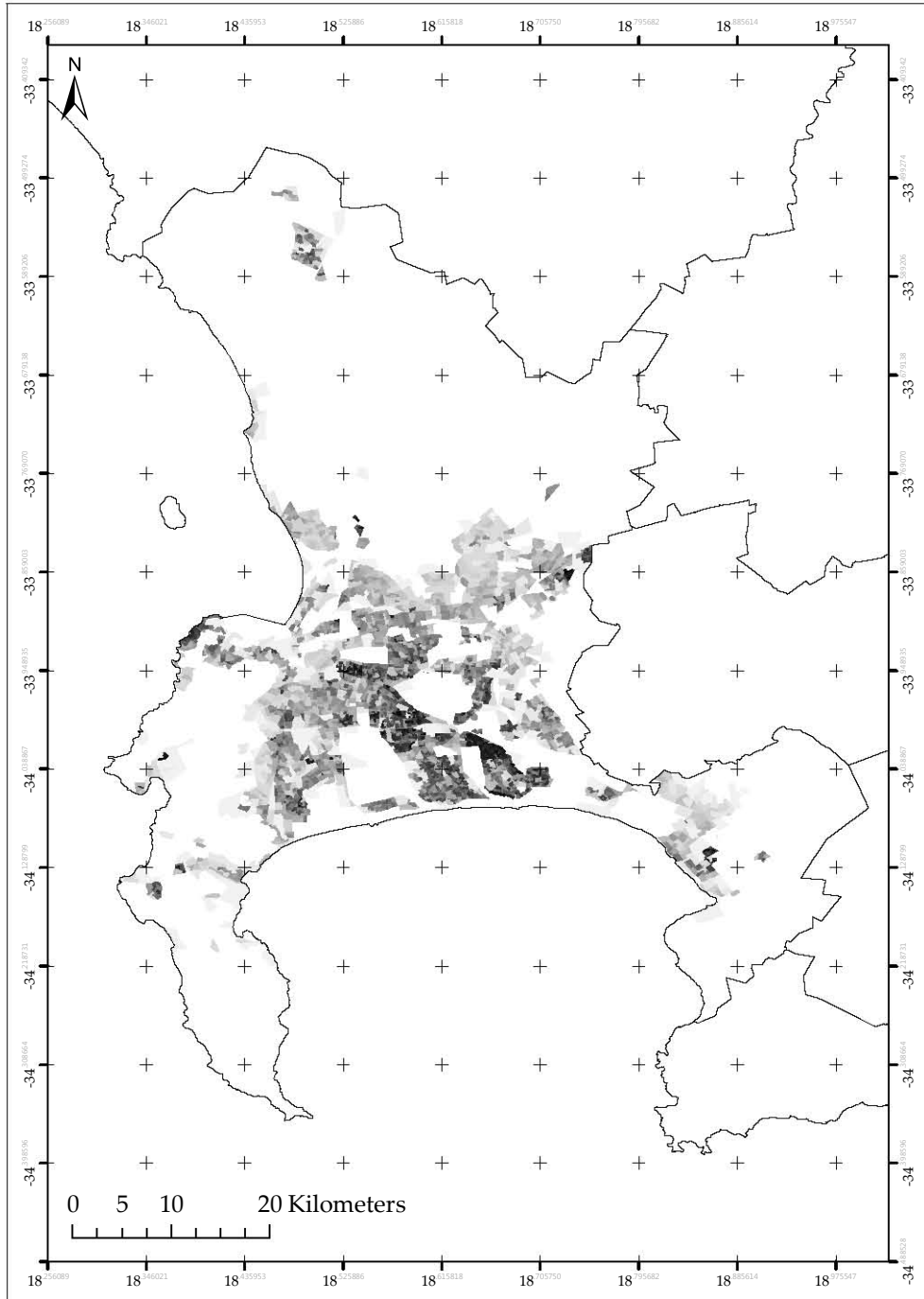
**Figure D.1:** Location of case study roads in Cape Town



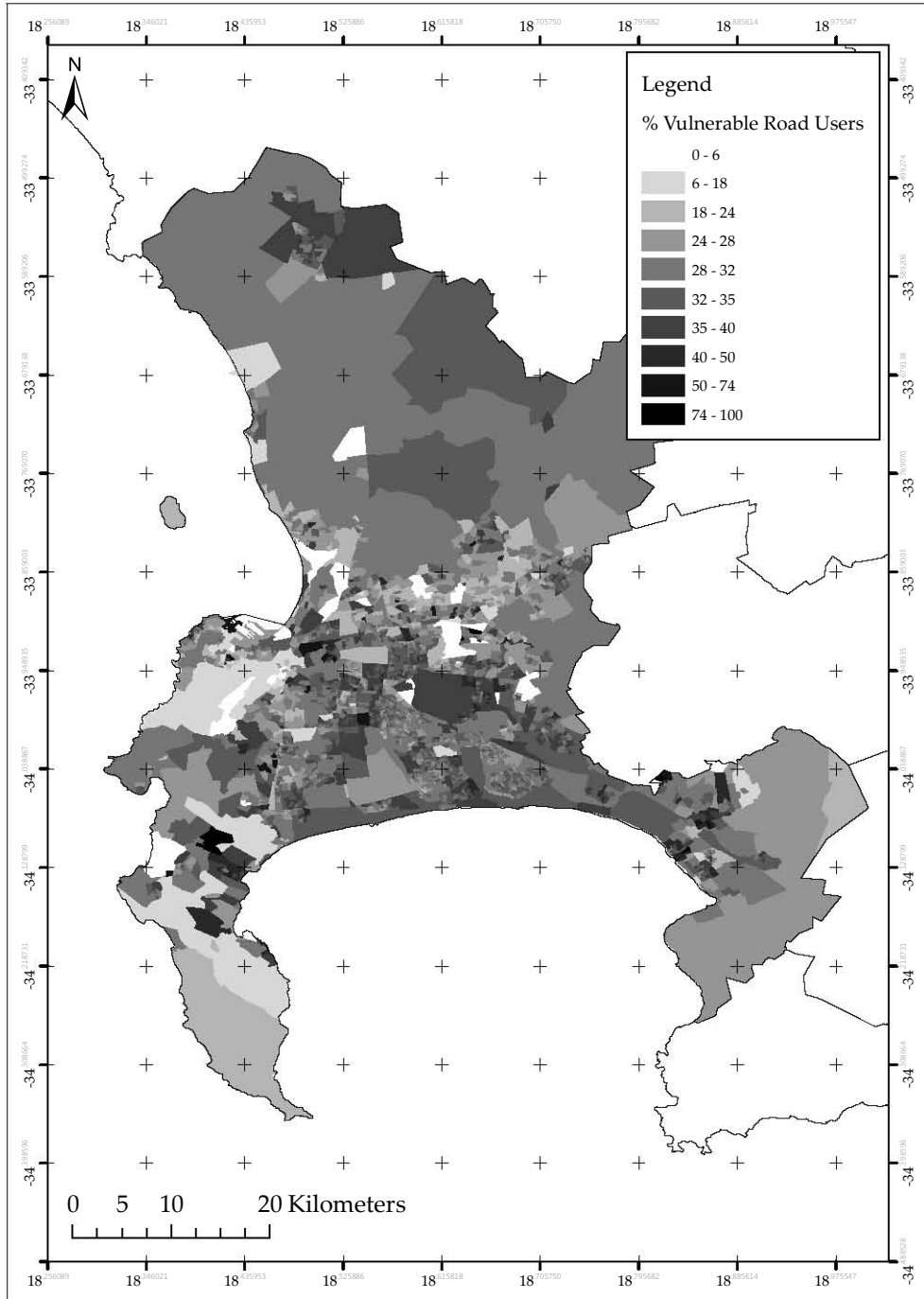
**Figure D.2:** Grouped Zoning Map for Cape Town

**Table D.1:** Zoning Categories and Base Zones for Cape Town.  
*Source:* Adapted from City of Cape Town (2007)

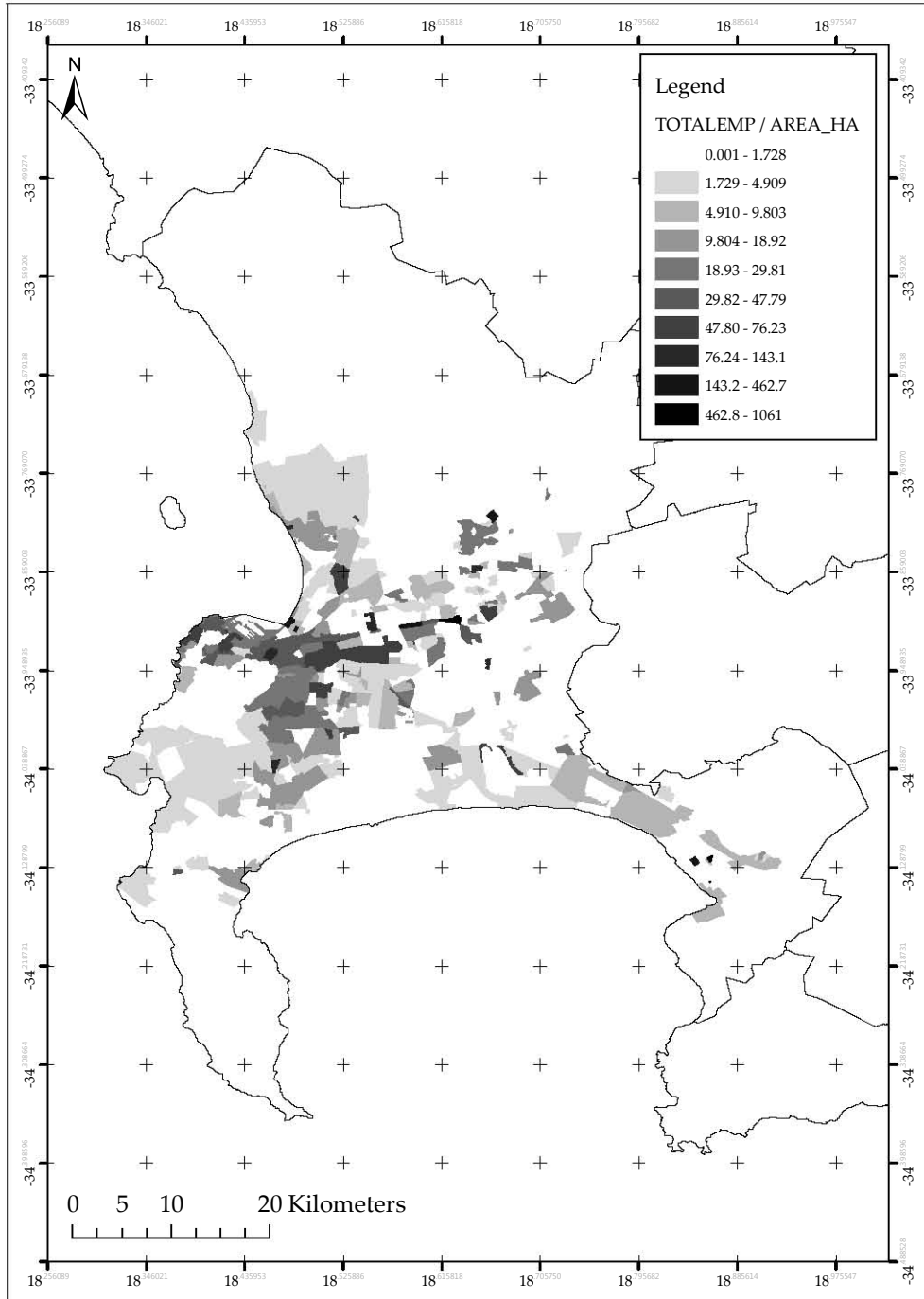
<b>Zoning Category</b>	<b>Base Zone</b>	<b>Purpose</b>
Single Residential Zones (SR)	SR1: Conventional Housing	Provides predominantly for single family dwelling houses.
	SR2: Incremental Housing	Provides for the upgrading of informal settlement to a formal settlement through incremental housing processes.
General Residential Zones (GR)	GR1: Group Housing	Medium density residential development where group housing is encouraged.
	GR2 to 6: Sub-Zones	Higher density residential development that includes blocks of flats. Different development rules apply to different sub-zones.
Community Zones (CO)	CO1: Local	To provide for local community facilities.
	CO2: Regional	To provide for the full range of community facilities. This can be local and regional in nature.
Local Business Zones (LB)	LB1: Intermediate Business	Provide buffer or interface between general business and lesser intensity zones.
	LB2: Local Business	Provide for low intensity commercial and mixed uses which serve local needs.
General Business and Mixed Zones (GB and MU)	GB1 to 7: Sub-Zones	Provide for general business and mixed uses of medium to high intensity. Different development rules apply to different sub-zones.
	MU1 to 3: Sub-Zones	Provide a mixture of business, industrial and residential development. Different development rules apply to different sub-zones.
Industrial Zones (GI and RI)	GI1 and 2: General Industry sub-zones	Provide for all industry except noxious industry. Different development rules apply to different sub-zones.
	RI: Risk Industry	Provides for industries which are noxious and carry a high risk.
Utility and Transport Zones (UT and TR)	UT: Utility Zone	Provide for utility services such as electrical substations and water reservoirs.
	TR1: Transport Use	Provides for transportation systems and transport undertakings which serve the public.
	TR2: Public Road and Public Parking	Provides for public streets and roads.
Open Space Zones (OS)	OS1: Environmental Conservation	Provides for the conservation of environmental resources.
	OS2: Public Open Space	Provides for active and passive recreational areas on public land
	OS3: Special Open Space	Provides for special reserved open spaces such as a golf course on a golf estate.
Agricultural, Rural and Limited Use Zones (AG, RU and LU)	AG: Agricultural Zone	Provides and protects agriculture on farms.
	RU: Rural Zone	Provides for smaller rural properties that may be used for agriculture and occupied as places of residence for people seeking a country lifestyle.
	LU: Limited Use Zone	Provides for a transitional mechanism to deal with land that was zoned as undetermined.



**Figure D.3:** Household Densities in Cape Town



**Figure D.4:** Percentage of Vulnerable Road Users



**Figure D.5:** Employment Densities in Cape Town

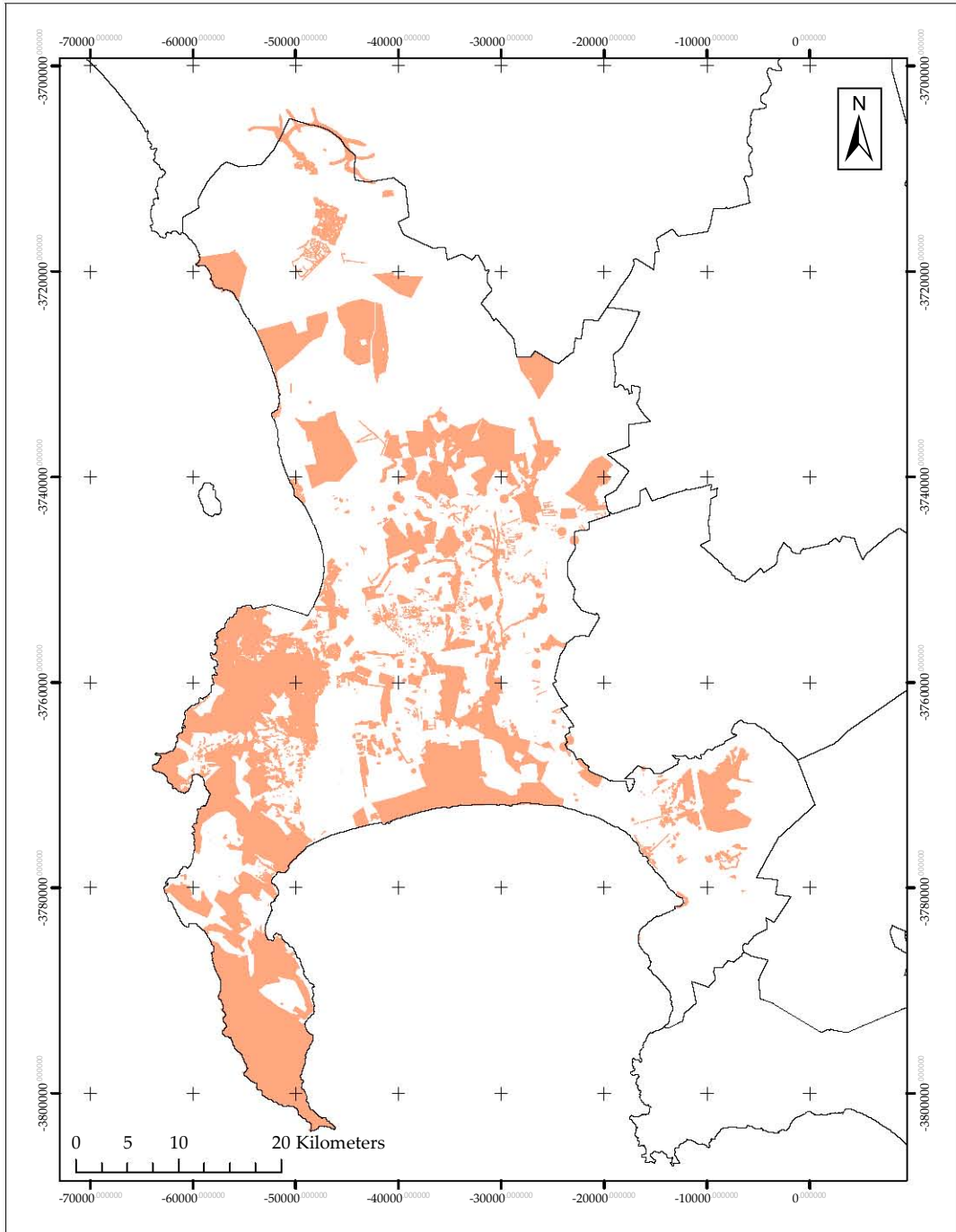
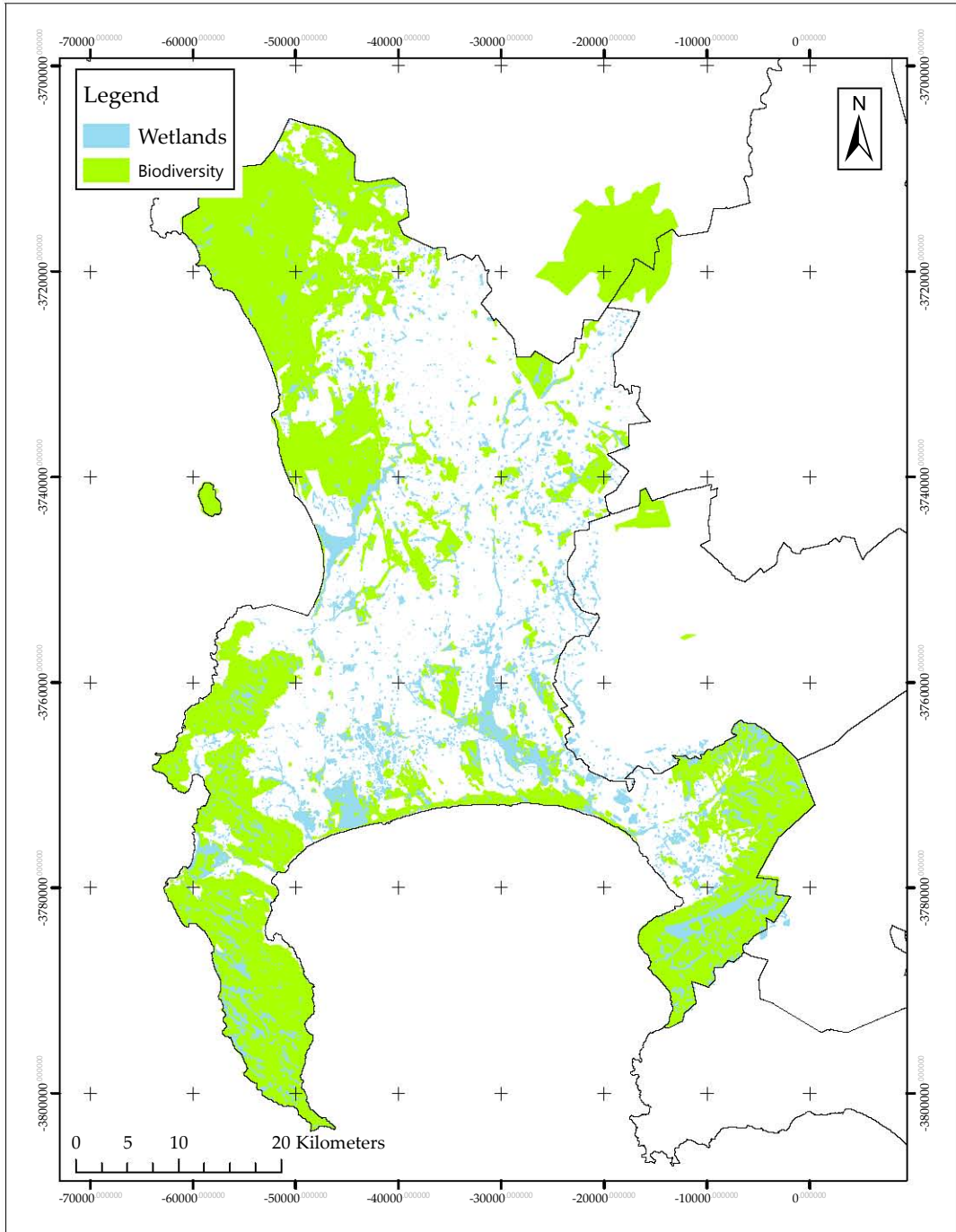
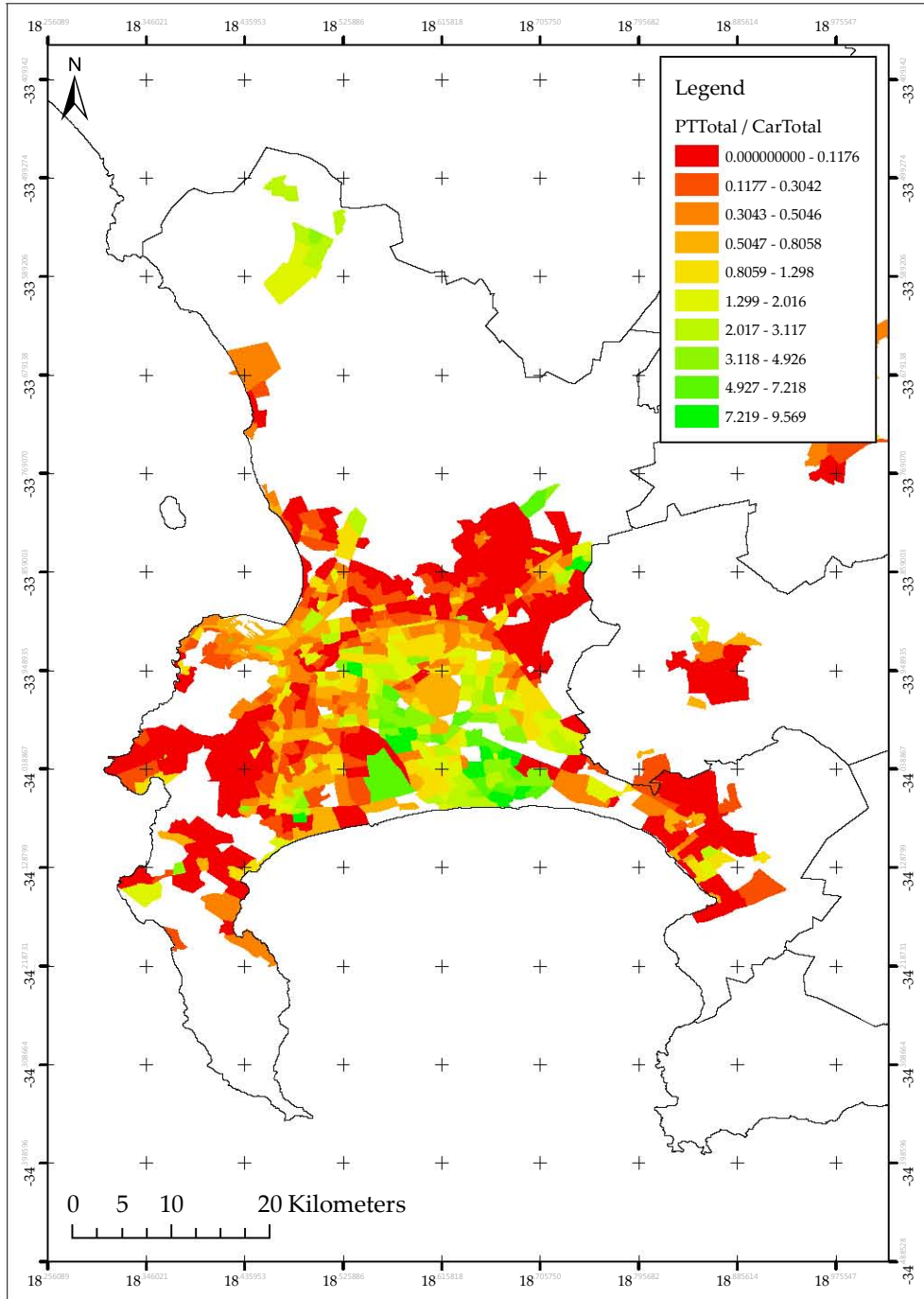


Figure D.6: Location of Heritage Features in Cape Town



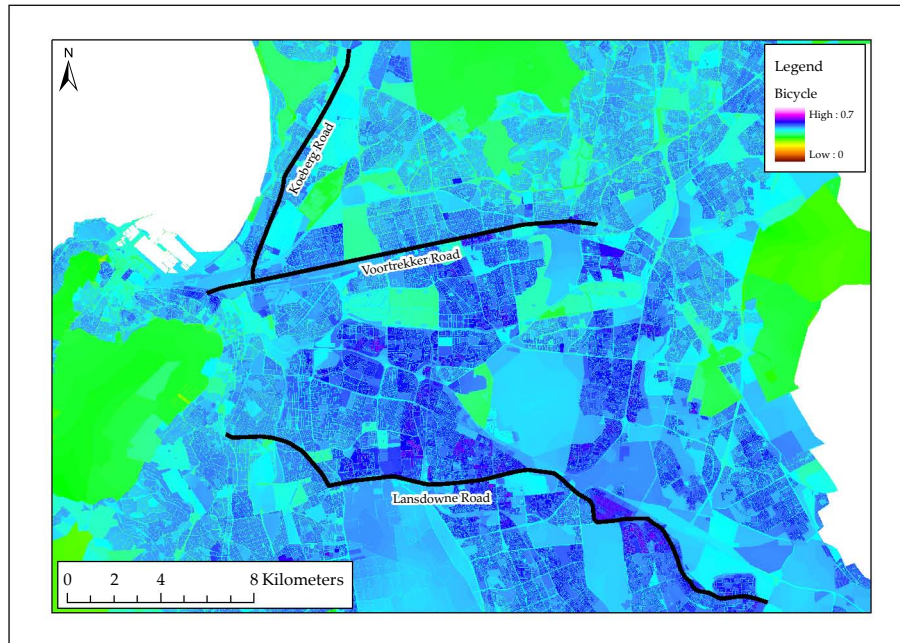
**Figure D.7:** Location of Wetlands and Environmentally Sensitive Areas in Cape Town



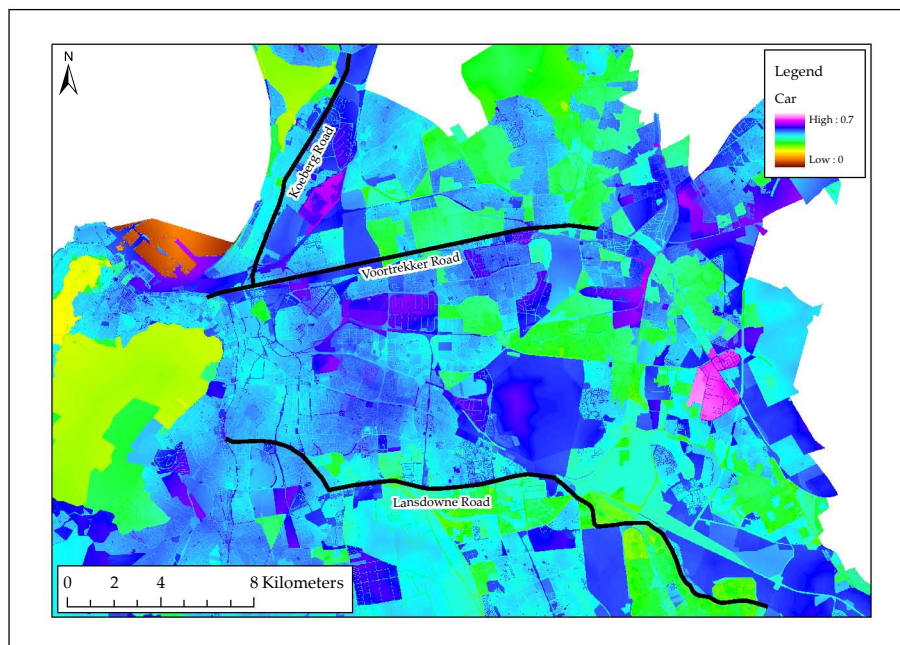
**Figure D.8:** Public Transport versus Car Trips per TAZ in Cape Town



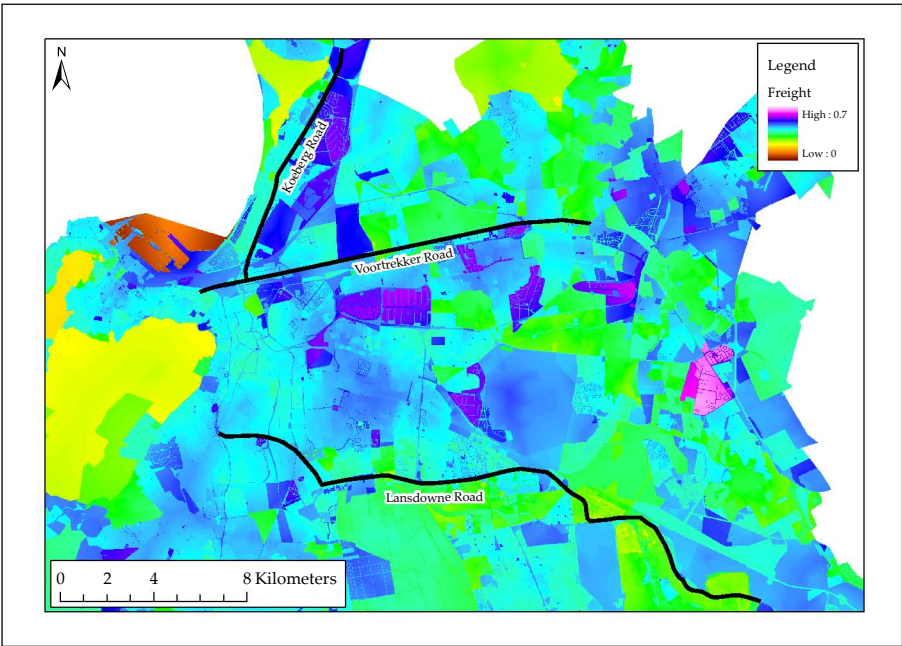
## Appendix E: Case Study Results and Assessment



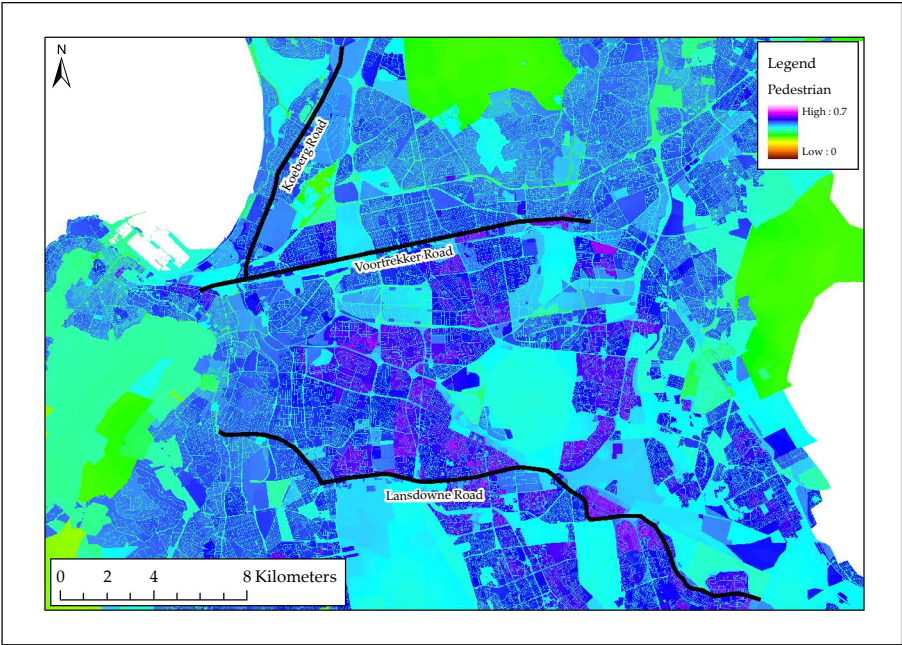
**Figure E.1:** Bicycle Suitability Scores from SMCA for Case Study Routes



**Figure E.2:** Car Suitability Scores from SMCA for Case Study Routes



**Figure E.3:** Freight Suitability Scores from SMCA for Case Study Routes



**Figure E.4:** Pedestrian Suitability Scores from SMCA for Case Study Routes

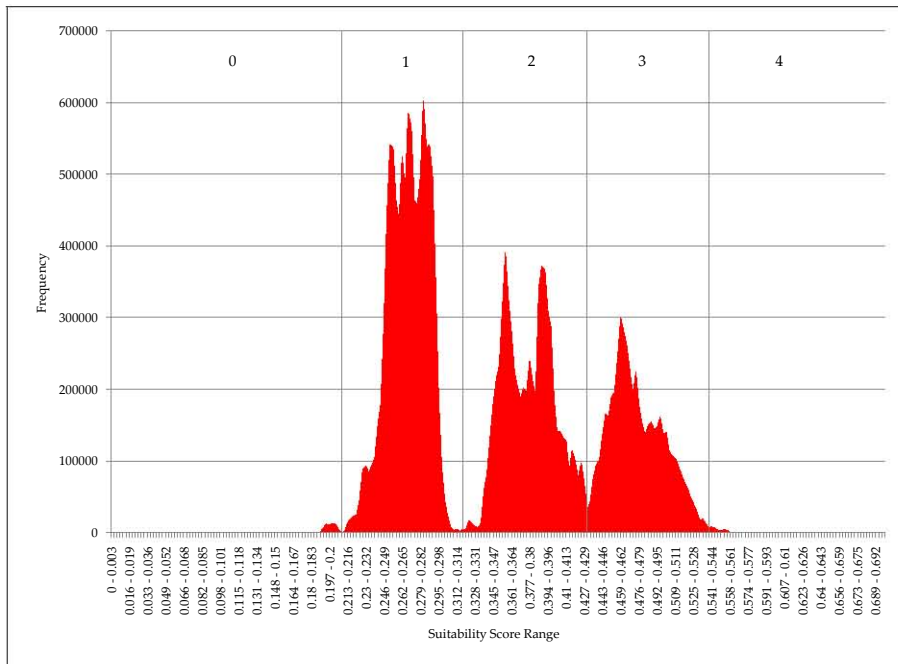


Figure E.5: Frequency Distribution of Suitability Scores for Pedestrian Mode

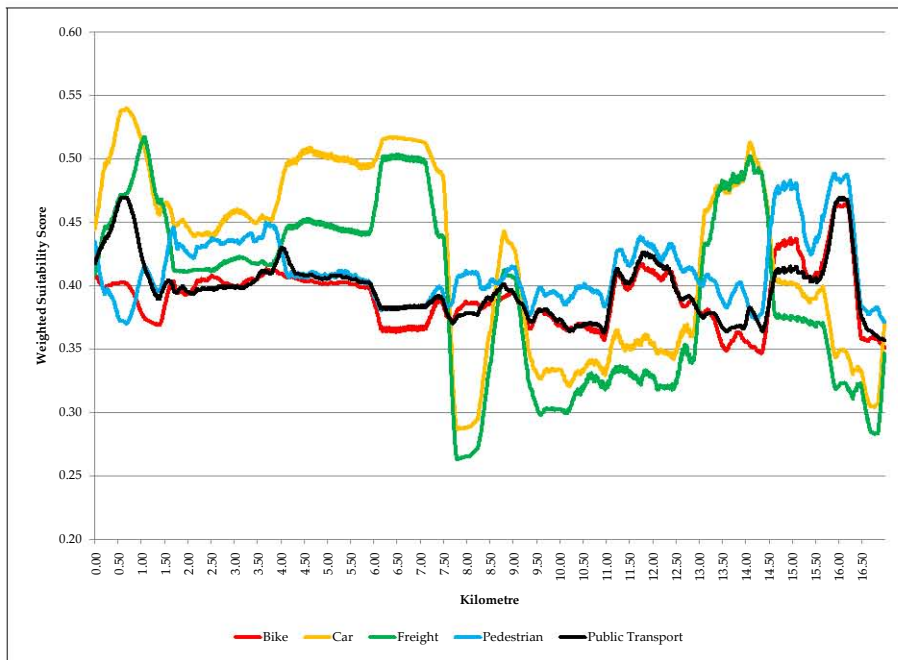
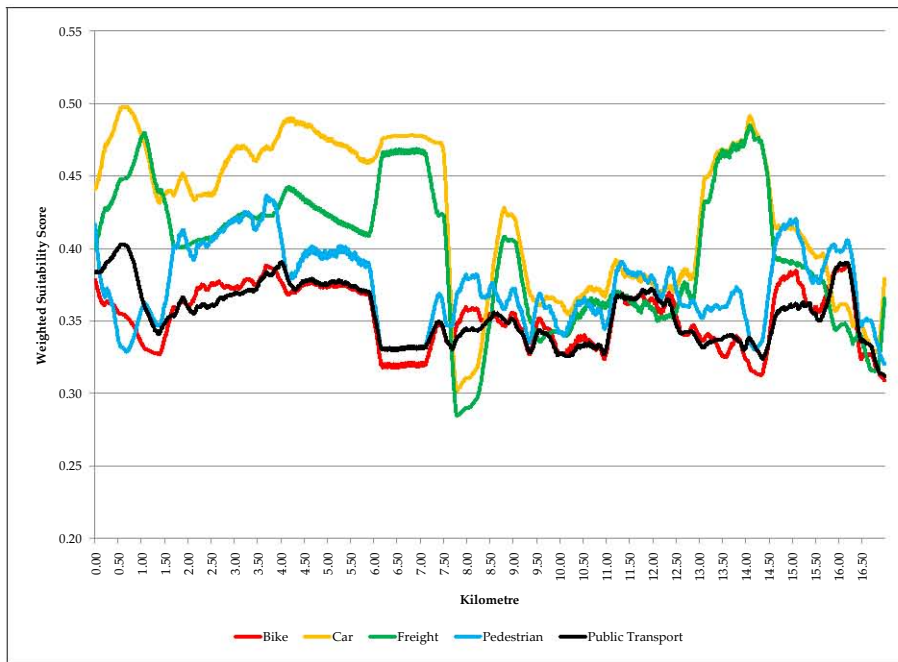
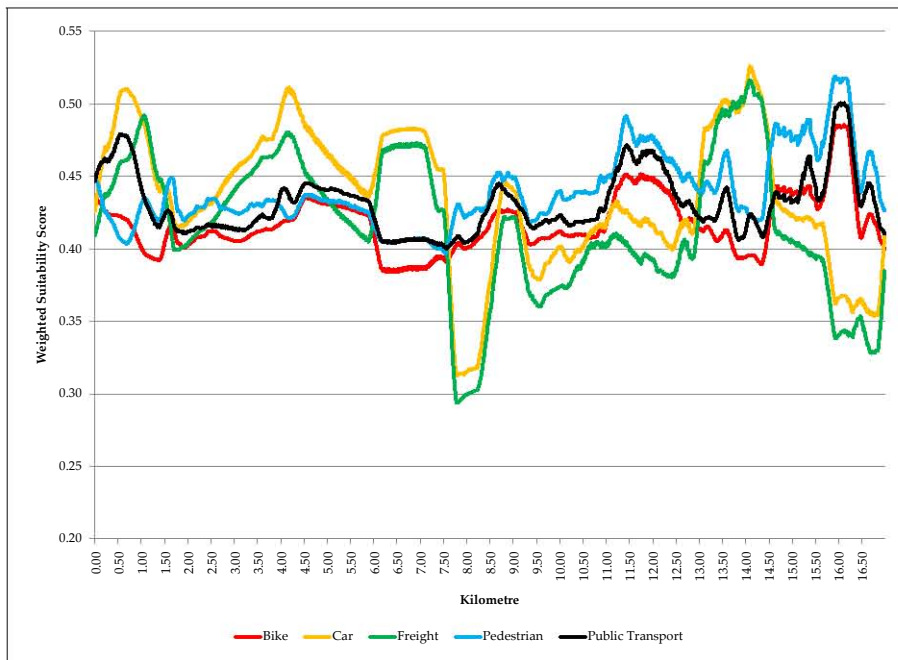


Figure E.6: Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 1



**Figure E.7:** Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 2



**Figure E.8:** Weighted Suitability Scores along Voortrekker Road under Weighting Scheme 3

## Appendix F: Cluster Analysis

### ALTERNATIVE CLUSTERING ALGORITHMS

The PAM, CLARA and CLARANS algorithms can be thought of as extensions or adaptations to K-means, targeting some of the shortcomings of the original algorithm, albeit at the expense of its simplicity, thereby incurring greater calculation costs. One of the common criticisms of K-means is its inherent sensitivity to outliers (see Budayan et al., 2009; Datta, 2003). Since K-means recalculates the cluster means after assigning data objects to clusters, if one of those objects is an outlier point, it will affect the recalculated cluster mean, perhaps skewing it such that it no longer accurately reflects the natural cluster patterns. PAM addresses this problem by representing each cluster centre by the medoid. The medoid is the most centrally located object in a cluster. The algorithm selects  $k$  medoids and tries to place objects into clusters based upon their eccentricity from the nearest medoid, all the while swapping medoids for non-medoids to improve the squared-error objective. Identifying the best medoids is a computationally intensive exercise, making the PAM algorithm costly for large  $n$  and  $k$  values (Andritsos, 2002).

CLARA and CLARANS attempts to get around this problem by breaking the data set into  $s$  samples, and then applying PAM to each sample sub-set. The best result achieved between each sub-set is then taken as the best overall result. This sampling technique may have quality issues if the samples that are initially selected are not representative of the overall database, and so it is recommended to run the algorithm a few times to ensure that the resultant clusters remain constant (Ester et al., 1998).

The K-means algorithm performs well on appropriately distributed (separated) and spherical-shaped groups of data. In case the two groups are close to each other, some of the objects in one group of data points might end up in different clusters, especially if one of the initial cluster representatives is close to the cluster boundaries. The algorithm does not perform well on clusters that are not circular or globular (so-called non-convex) due to the usage of Euclidean distance (Khalilian and Mustapha, 2010). As already mentioned, PAM tends to handle outliers better, since medoids are less influenced by outlier values than means.

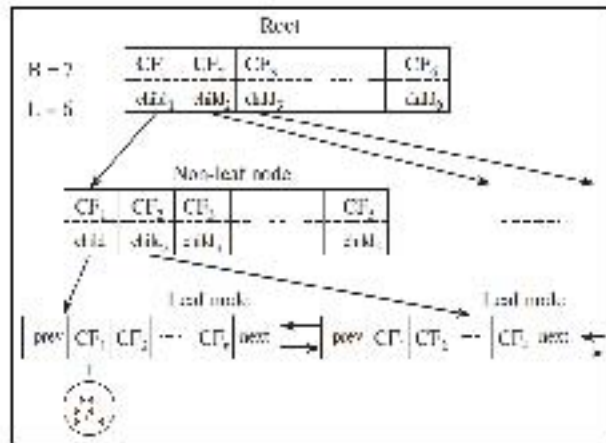
Since CLARA and CLARANS rely on the sampling of sub-sets of the database, their efficiency and effectiveness is highly dependent on the sample size used and its bias. A bias is present in a sample when the data objects in it have not been drawn with equal probabilities. As a result, their application tends to be restricted to numerical data of lower dimensionality, with inherently well separated, high density clusters (Andritsos, 2002).

### BIRCH Algorithm

Balanced Iterative Reduction and Clustering using Hierarchies (BIRCH) developed by Zhang et al. (1996) is suitable for large databases, improving upon the slow runtimes of other hierarchical algorithms. BIRCH relies on data squashing techniques that scans the data to compute summaries (called sufficient statistics) of the data. These summaries are then used instead of the original data for clustering. The summaries are called clustering features (CF), and BIRCH creates a data structure from these, called a CF-tree. The clustering feature is a three dimensional vector, defined as  $CF = (N, \vec{LS}, SS)$ , where  $N$  is the number of data points in the cluster,  $\vec{LS}$  is the linear sum of data points in the cluster ( $\sum_{i=1}^N \vec{X}_i$ ), and  $SS$  is the square sum of the  $N$  data points ( $\sum_{i=1}^N \vec{X}_i^2$ ) (Zhang et al., 1996).

A CF-tree is a multilevel summary of the data distribution. It consists of the root node, non-leaf nodes

and leaf nodes (Figure F.1). CF-trees are characterised by the branching factor,  $B$ , and the threshold,  $T$ . The branching factor controls the number of children for every non-leaf node, and the threshold controls the maximum distance between any pair of points or the diameter of the subclusters stored at the leaf nodes (Zhang et al., 1996).



**Figure F.1:** A CF-Tree used by the BIRCH algorithm. A CF-tree is a multilevel summary of the data distribution consisting of the root node, non-leaf nodes and leaf nodes. *Source:* Han and Kamber (2006)

The BIRCH algorithm follows the following procedure (Andritsos, 2002):

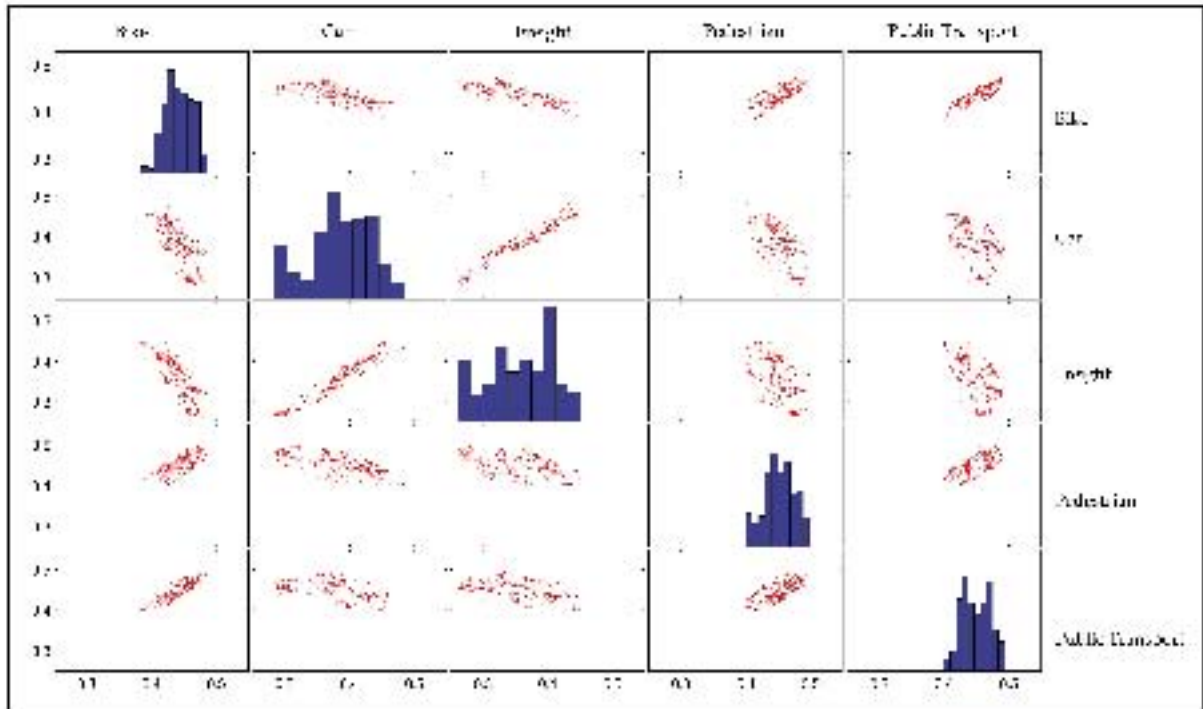
1. The data objects are loaded one by one and the initial CF-tree is constructed and an object is inserted into the closest leaf entry. If the diameter of the this sub-cluster becomes larger than the threshold value,  $T$ , the leaf node, and possibly its parent nodes (depending on whether the branching factor,  $B$ , is violated) are split. When the object is properly inserted in a leaf node, all nodes towards the root of the tree are updated with the new information.
2. If the CF-tree of stage 1 does not fit into memory, a smaller CF-tree is built. The size of a CF-tree is controlled by the parameter,  $T$ , and thus choosing a larger value for it will merge some sub-clusters making the tree smaller.
3. The leaf nodes of the CF-tree hold subcluster statistics. In this stage BIRCH uses these statistics to apply some clustering technique, e.g. k-means, and produce an initial clustering.
4. Redistribute the data objects using the centroids of the clusters discovered in step 3. This is an optional stage which requires an additional scan of the data set and re-assigns the objects to their closest centroids. Optionally, this phase also includes the labelling of the initial data and discarding of outliers.

The disadvantages of the method include a difficulty in finding arbitrary shaped clusters, since it uses the notion of a subcluster radius. Strongly elliptical, linear or concave cluster shapes are not as well detected. Since the algorithm uses the Euclidean distance, it only works well on well distributed numerical data. Furthermore, the parameter  $T$  affects the cluster sizes and thus their naturalness, forcing objects that should be in the same cluster to end up in different ones, while duplicate objects could be attracted by different clusters, if they are presented to the algorithm in different order. The algorithm should, therefore, be run a number of times on differently ordered data to ensure that the clusters are consistent.

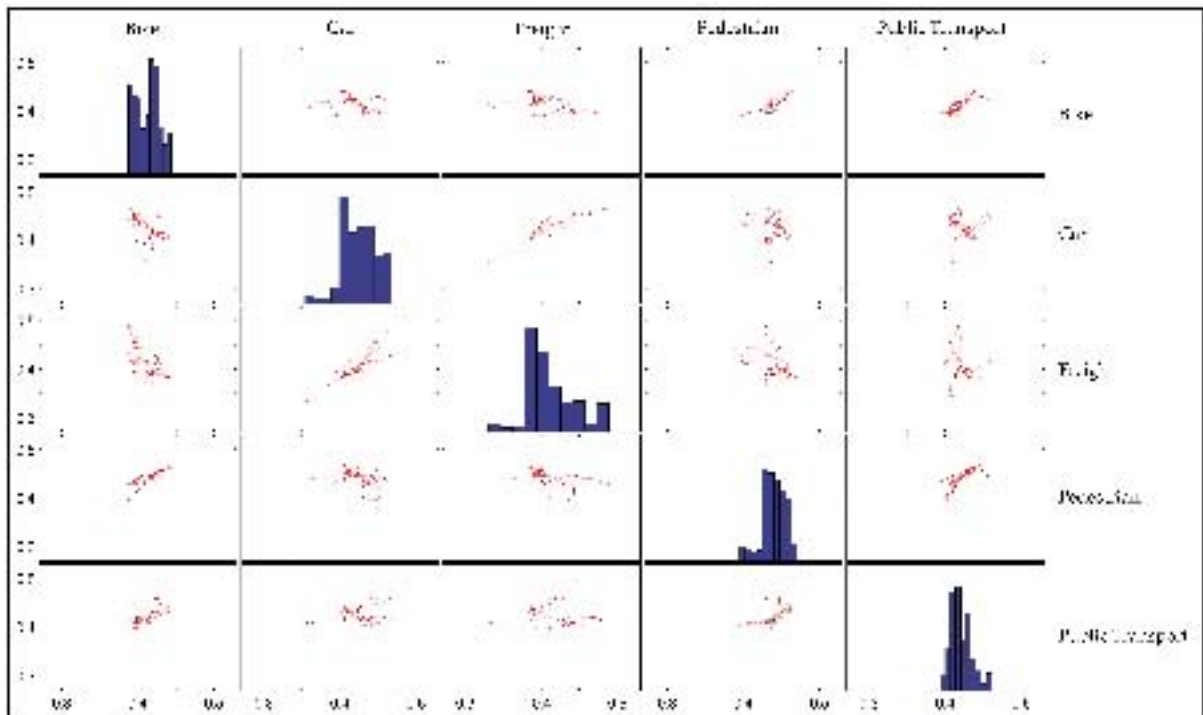
**Table F.1:** Properties of Various Clustering Algorithms. There are a large number of algorithms available that each take different approaches to clustering, improve on the shortcomings of other methods, or are optimised for different data types and different cluster needs.

*Source:* Adapted from Andritsos (2002)

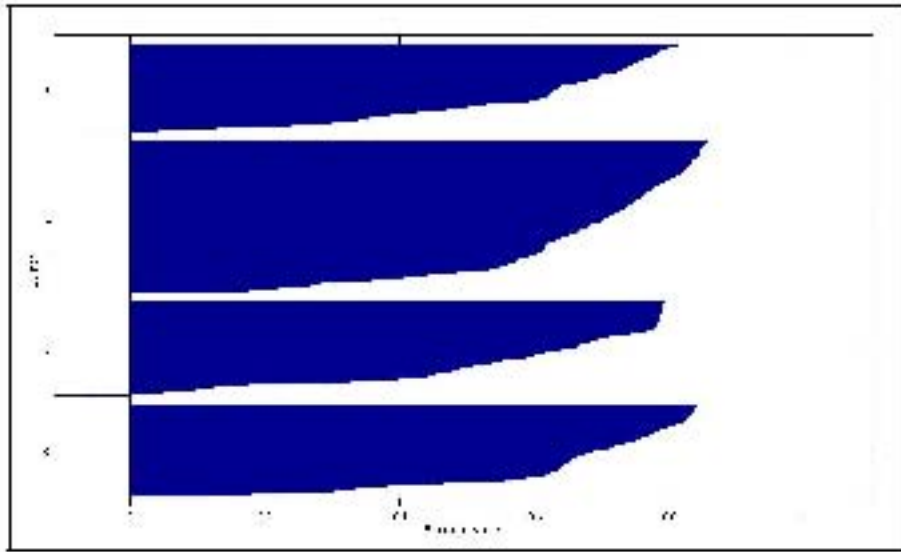
<b>Algorithm</b>	<b>Input Parameters</b>	<b>Optimised for</b>	<b>Cluster Structure</b>	<b>Outlier Handling</b>
<b>Partitional Methods</b>				
K-means	Number of Clusters	Separated Clusters	Spherical	No
PAM	Number of Clusters	Separated Clusters, Small Data Sets	Spherical	No
CLARA	Number of Clusters	Relatively Large Data Sets	Spherical	No
CLARANS	Number of Clusters, Maximum Number of Neighbours	Spatial Data Sets, Better Quality of Clusters than PAM and CLARA	Spherical	No
<b>Hierarchical Methods</b>				
BIRCH	Branching Factor, Diameter Threshold	Large Data Sets	Spherical	Yes
CURE	Number of Clusters, Number of Cluster Representatives	Arbitrary Shapes of Clusters, Relatively Large Data Sets	Arbitrary	Yes
<b>Density-Based Methods</b>				
DBSCAN	Radius of Clusters, Minimum Number of Points in Clusters	Arbitrary Shapes of Clusters, Large Data Sets	Arbitrary	Yes
DENCLUE	Radius of Clusters, Minimum Number of objects	Arbitrary Shapes of Clusters, Large Data Sets	Arbitrary	Yes
OPTICS	Radius of Clusters (min,max), Minimum Number of objects	Arbitrary Shapes of Clusters, Large Data Sets	Arbitrary	Yes
<b>Miscellaneous Methods</b>				
STING	Number of cells in lowest level, Number of objects in cell	Large Spatial Data Sets	Vertical and Horizontal Boundaries	Yes
WaveCluster	Number of Cells for each Dimension, Wavelet, Number of application of Transform	Arbitrary Shapes of Clusters, Large Data Sets	Arbitrary	Yes
CLIQUE	Size of the Grid , Minimum Number of Points within each Cell	High Dimensional Large Data Sets	Arbitrary	Yes
ScalableEM	Initial Gaussian Parameters, Convergence Limit	Large Data Sets with Approximately Uniform Distribution	Spherical	No (?)



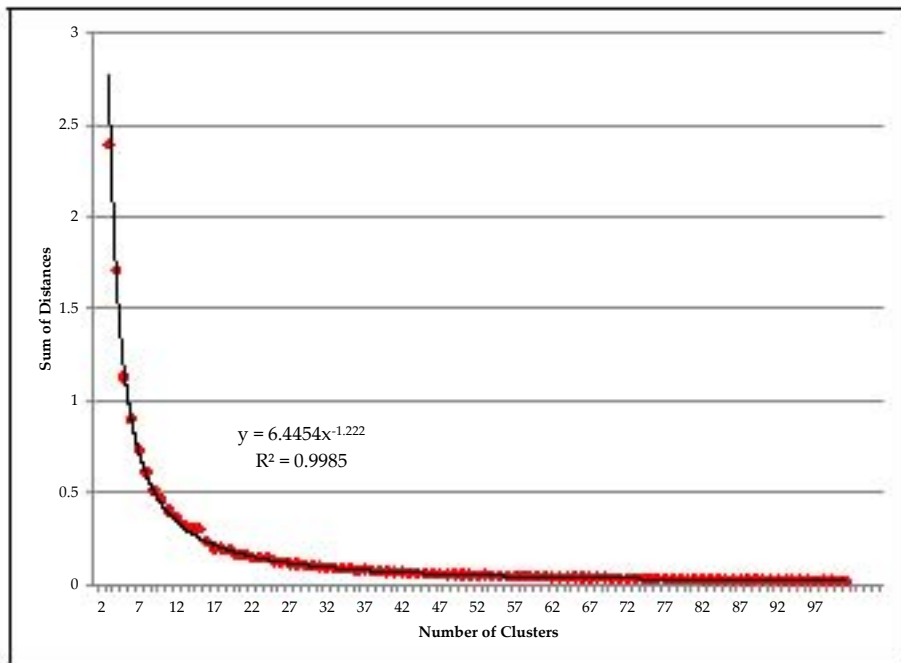
**Figure F.2:** Mode score relationships for Lansdowne Road. Each modes preference scores are plotted against every other modes preference scores in a scatter plot, and the histograms for each mode is shown along the central diagonal.



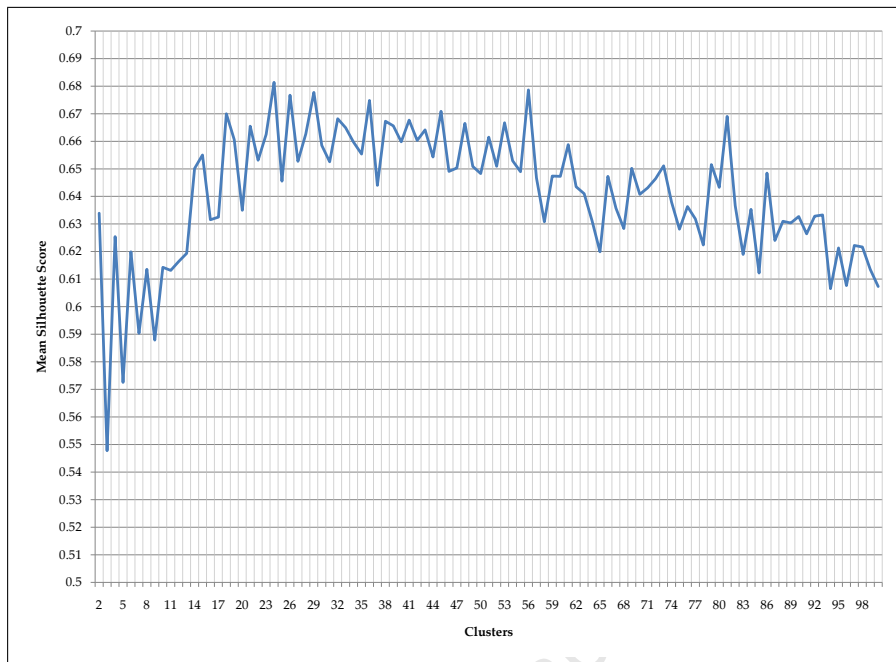
**Figure F.3:** Mode score relationships for Koeberg Road. Each modes preference scores are plotted against every other modes preference scores in a scatter plot, and the histograms for each mode is shown along the central diagonal.



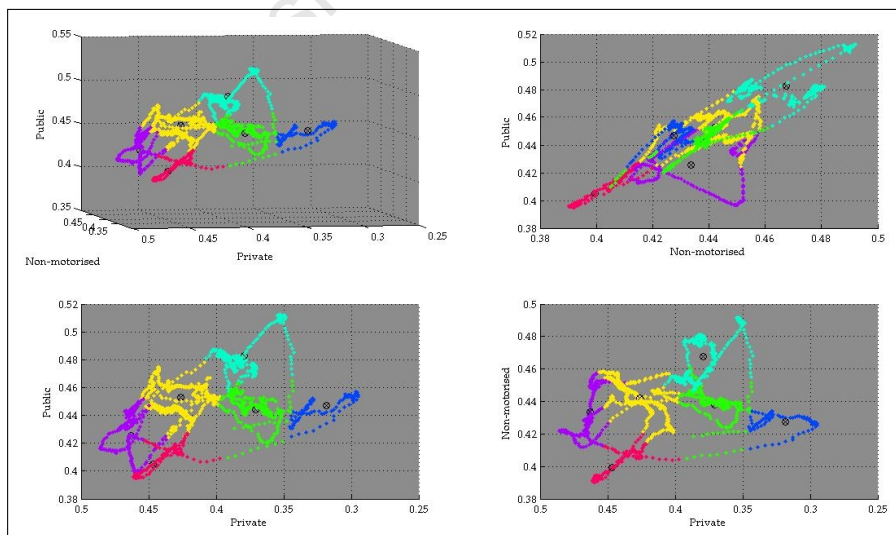
**Figure F.4:** Silhouette Plot for Four-Cluster Scenario for Voortrekker Road. The four-cluster scenario produces better formed clusters than for the three-cluster scenario.



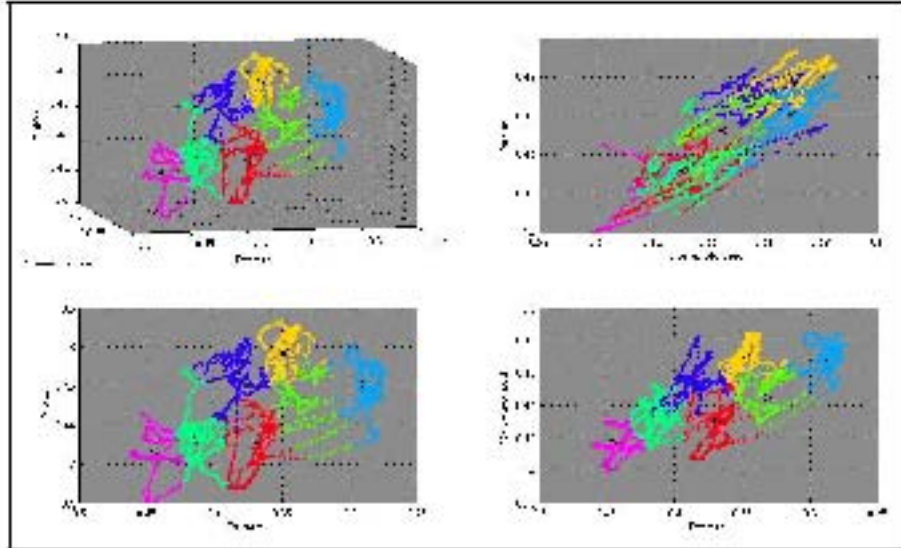
**Figure F.5:** Total Sum of Distances Versus Number of Clusters for K-means on Voortrekker Road. The point of maximum curvature is nearest  $k = 3$ , indicating that the ideal number of clusters is 3



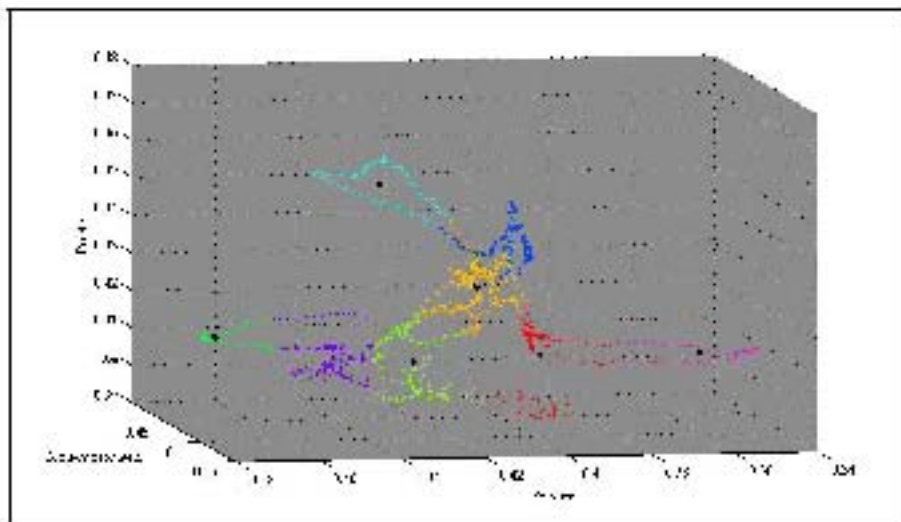
**Figure F.6:** Mean Silhouette Score versus Number of Clusters. For the Voortrekker Road data set,  $k = 24$  produces the most distinctive set of clusters.



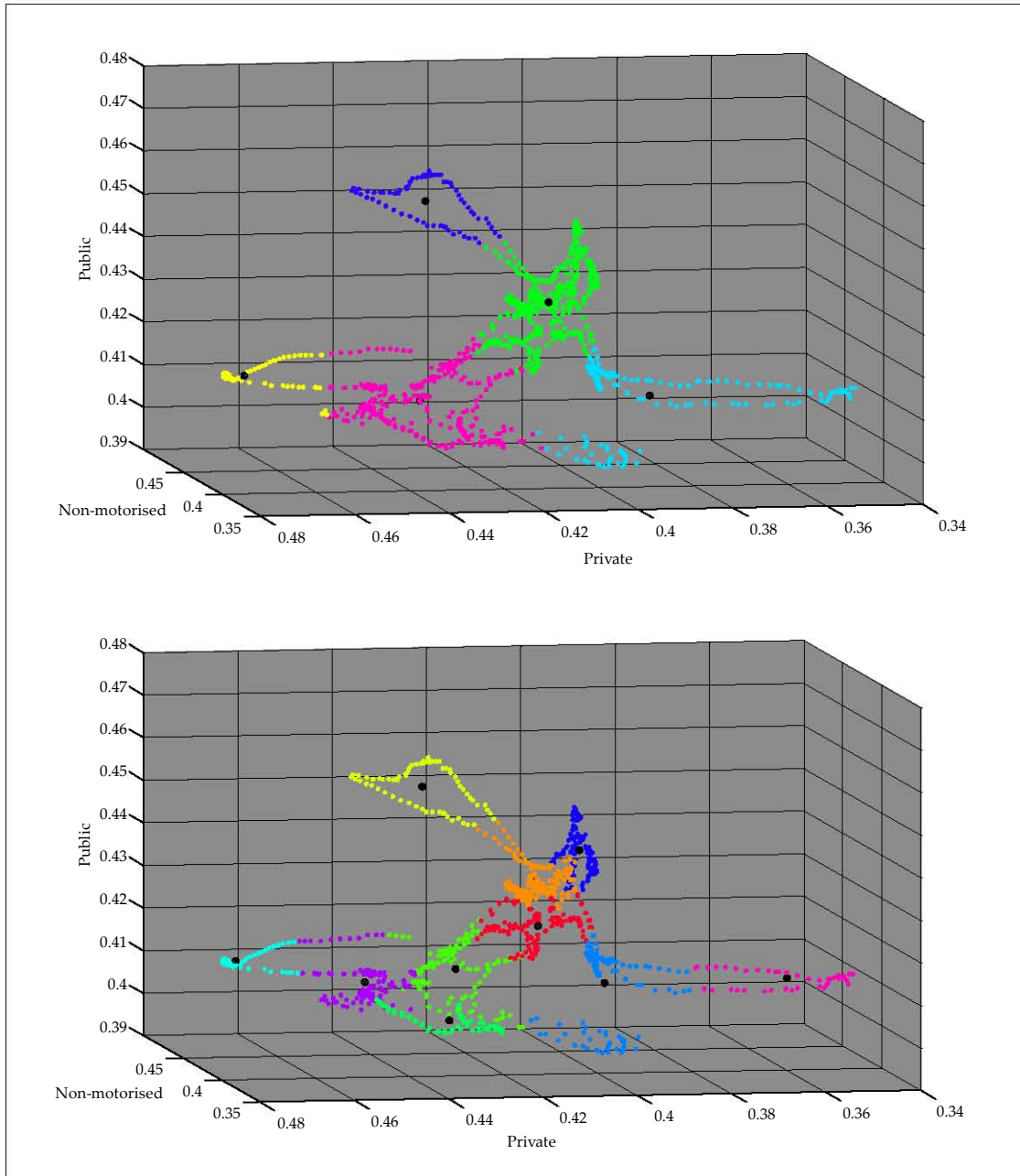
**Figure F.7:** Perspective views of k-means clustering results for 6 cluster centres along Voortrekker Road



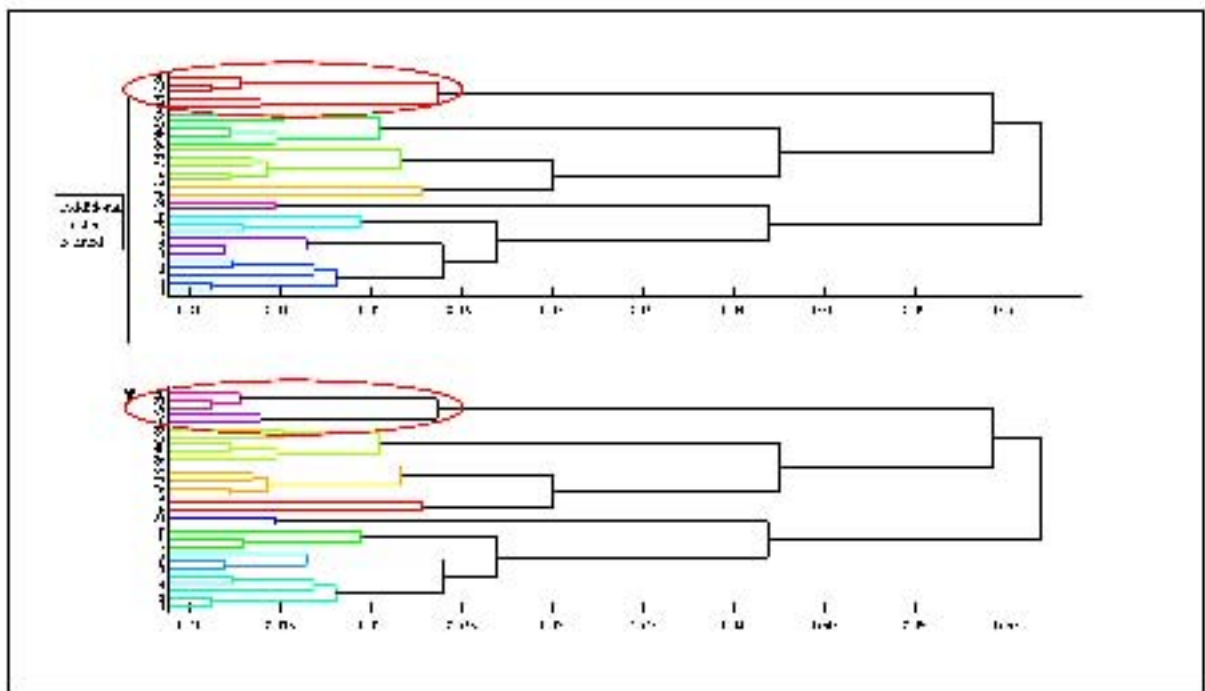
**Figure F.8:** Perspective views of k-means clustering results for 7 cluster centres along Lansdowne Road



**Figure F.9:** K-means clustering results for 8 cluster centres along Koeberg Road



**Figure F.10:** K-means clustering results for 5 (top) and 10 (bottom) cluster centres along Koeberg Road. The cyan and pink clusters in the 5 cluster solution are split into 2 and 3 clusters respectively in the 10 cluster solution. The green cluster, although well formed, is also split into 3 clusters in the 10 cluster solution. This suggests that an intermediate  $k$  value would produce a better clustering for this dataset.



**Figure F.11:** Dendrogram of 8 and 9 cluster solutions along Koeberg Road. The dendrogram shows how the 9 cluster solution is agglomerated to an 8 cluster solution by the linkage of the top two clusters (pink and purple) into a single cluster (red).



## Appendix G: Infrastructure and Context

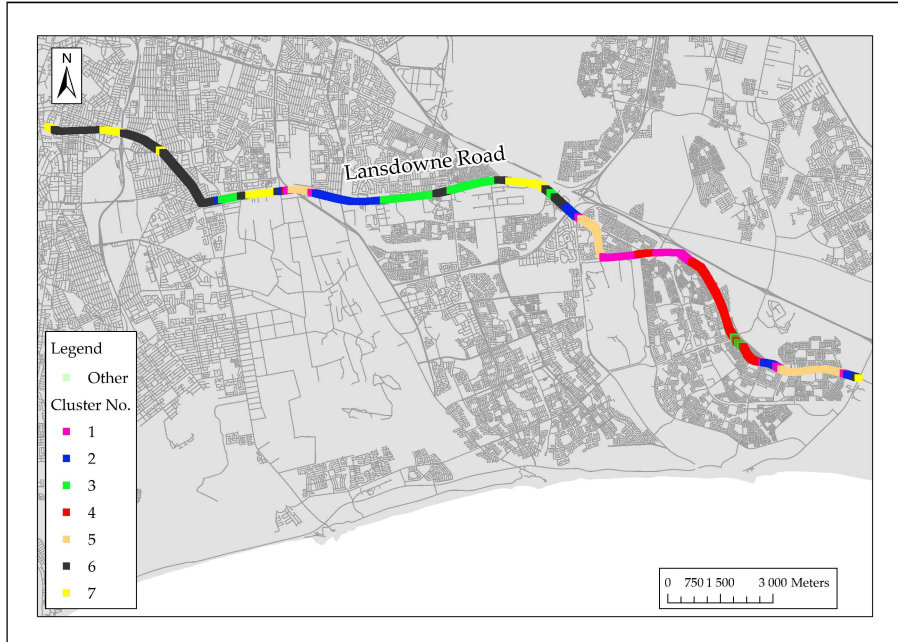


Figure G.1: Cluster locations along Lansdowne Road

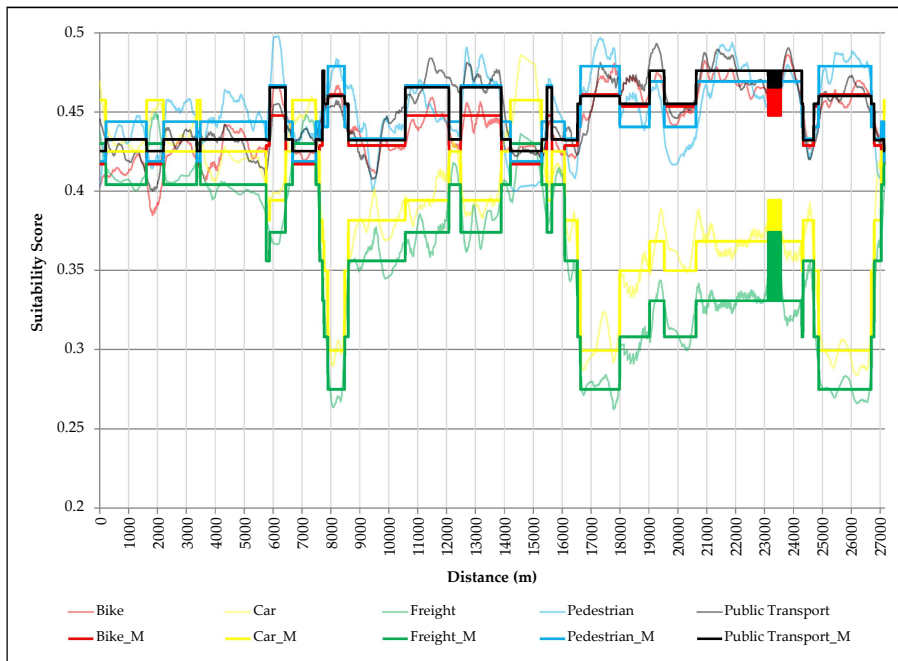
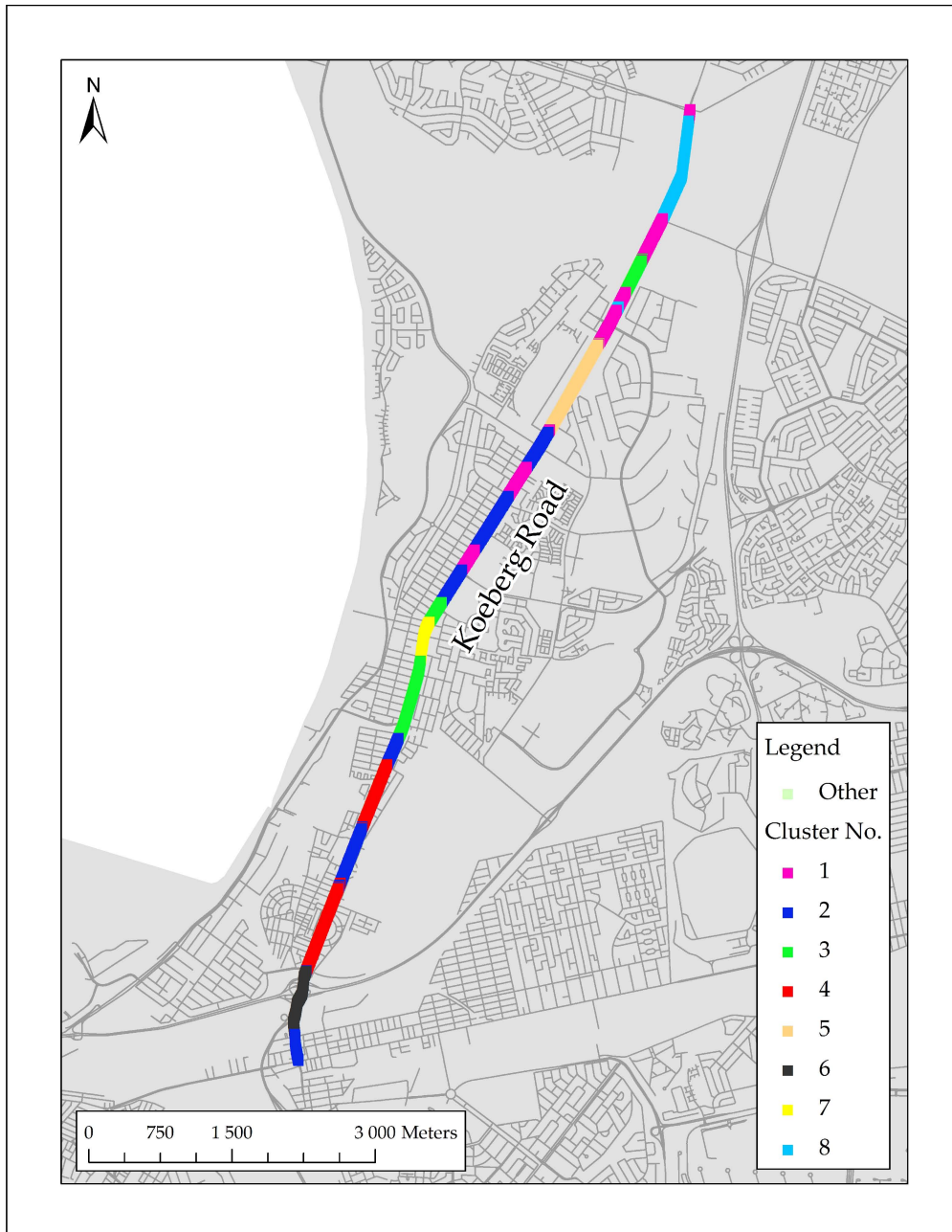
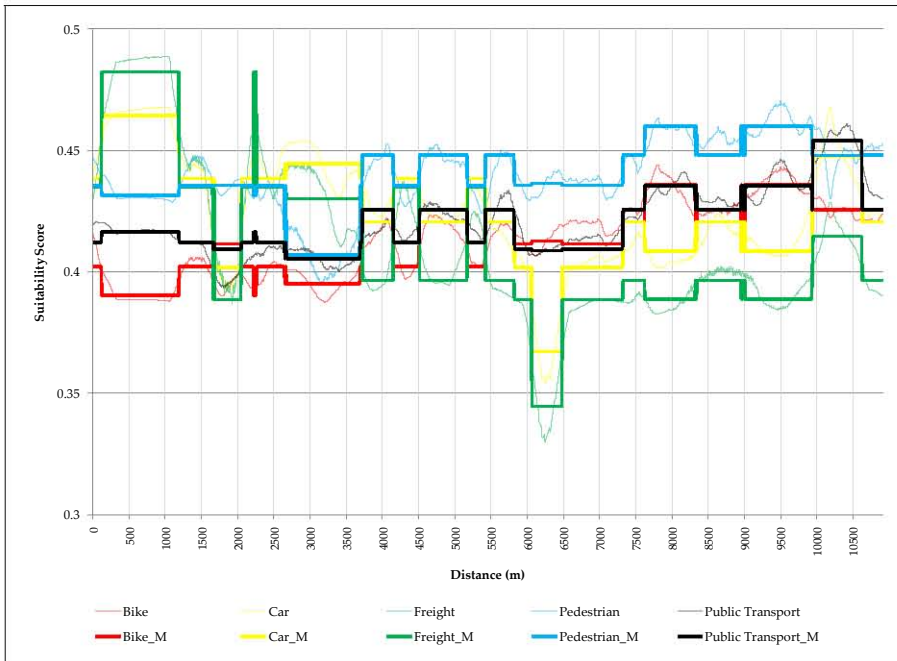


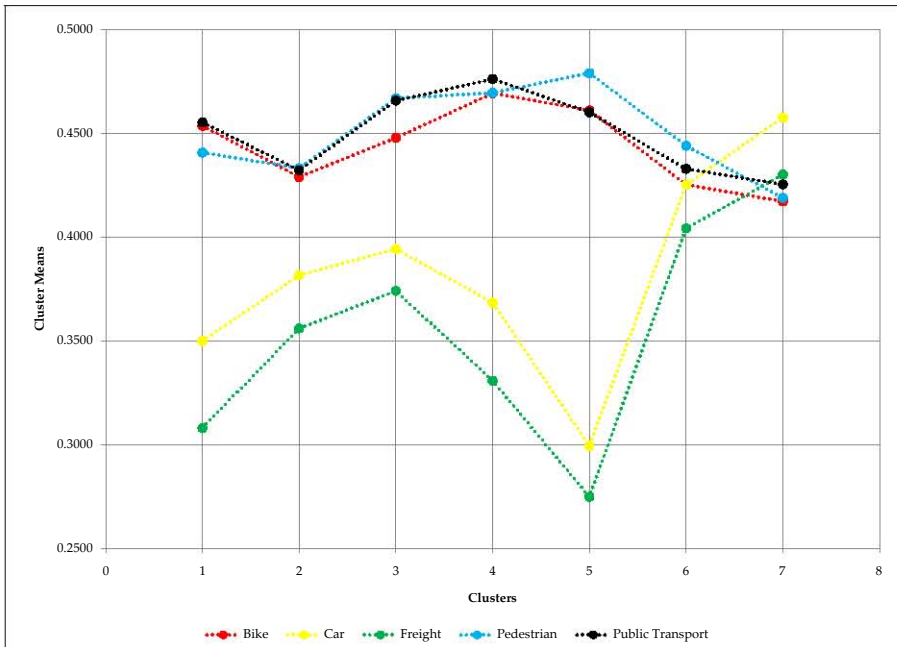
Figure G.2: Comparison of Cluster Means and SMCA Scores for Lansdowne Road



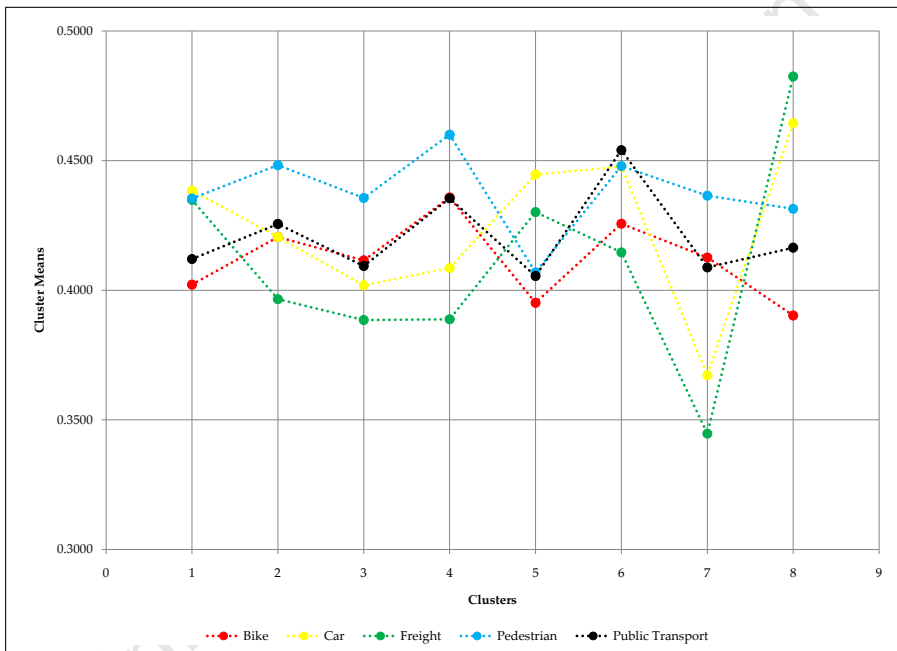
**Figure G.3:** Cluster locations along Koeberg Road



**Figure G.4:** Comparison of Cluster Means and SMCA Scores for Koeberg Road



**Figure G.5:** Comparison of Cluster Means for Lansdowne Road



**Figure G.6:** Comparison of Cluster Means for Koeberg Road