Towards Integrated Land Use and Transport Modelling: Evaluating Accuracy of the Four Step Transport Model- the Case of Istanbul, Turkey

James Njiraini Gachanja

March, 2010
Towards Integrated Land Use and Transport Modelling: Evaluating Accuracy of the Four Step Transport Model- the Case of Istanbul, Turkey

by

James Njiraini Gachanja

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Urban Planning and Management.

Thesis Assessment Board

Chairman: Prof.Dr.Ir. M.F.A.M. van Maarseveen
External Examiner: Prof.Dr.Gurcan Buyuksalih
First Supervisor: Dr. Sherif Amer
Second Supervisor: Dr. R.V. Sliuzas
Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.
‘The acid test for any model is how well it corresponds with reality’

Michael Pacione, 2005.

In memory of my mother and father
Abstract

Urban land use and transportation are known to interact and influence each other yet it is common practice for planners in both domains to prepare plans without due consideration for this interaction. Integrated land use and transport models are therefore important because they enable policy makers to foresee and evaluate the effects of transport and urban plans jointly thus providing solutions to common planning problems. The four step model (FSM) is identified as a weak point in integrating land use and transport modelling because it does not deal with land use explicitly. Another concern is that the sequential approach of the FSM creates a situation where no feedback is provided from one step to the other leading to inaccurate outputs. This research focuses on evaluating the (FSM) in order to assess the factors that influence its accuracy.

The City of Istanbul in Turkey is used as the case study. The forecast period for this research is between 1996 and 2006. The methodology involves comparing the FSM daily traffic against the actual observed daily traffic derived from RTMS records. Using the 1997 Istanbul transport model developed by ITU and IMM as the basis for analysis, forecasts for population, students and employment as well as the planned road network are validated against actual data for 2006. A total of six FSM runs are then undertaken using different inputs, namely: projected data, actual data, disaggregated TAZs, actual highway network, new model parameters and finally a feedback mechanism. The results from these six model runs are assessed for accuracy by using the %RMSE statistic. The effect size of using each different input is then measured by means of the F-test and omega squared ($\omega^2$).

Results show that the population and employment forecasts are overestimated while those for students are underestimated. The planned road network significantly differs from the actual road network. The %RMSE accuracy assessment results for the FSM under different inputs is as follows: projected data (204.43%); actual data (195.02%); disaggregated TAZs (181.81%); actual highway (138.78%); new parameters (100.92%); and feedback (170.75%). In terms of effect size, the use of new parameters has the most significant effect ($\omega^2=.549$) followed by application of a feedback mechanism ($\omega^2=.501$). The use of disaggregation, actual highway and projected data has insignificant effect on the FSM.

The land use and socio economic projections account for the greatest inaccuracy in the FSM while the use of new parameters and feedback improves the accuracy of the FSM significantly. A change in model inputs leads to a change in model outputs above and beyond the effect of extraneous factors, hence the variation in model outputs due to model inputs used is significant. The study recommends improving the quality of input data, application of feedback and estimation of the trip generation parameters at different time steps in the forecasting horizon based on anticipated changes in travel behaviour.

Keywords: land use and transportation modelling; four step model; accuracy assessment.
Acknowledgements

All glory and honour be to God the Father Almighty creator of heaven and earth; thank you Lord for blessing me. Thank you Netherlands, and thank you NUFFIC and ITC for giving me a chance to pursue an MSc degree, I have achieved personal and professional growth. I thank the entire ITC staff for making my stay here in Enschede a wonderfully enriching experience. I also thank my employer Mr. Geoffrey Njoroge of Eco Plan Kenya for supporting me in the pursuit of this endeavour.

I extend a special token of appreciation to my supervisors Dr. Sherif Amer and Dr. Richard Sliuzas for guiding me through the course of this study, your advice provided clarity and direction, thank you. I also extend much gratitude to Dr. Ir. Zuidgeest for the free consultancy on transport modelling, to Ir. Frans van Den Bosch for ensuring that the technical tools at the heart of this study were available and to Dr. J. Flacke and Dr.J.A.Martinez for support in the field work exercise. To the rest of the UPM members of staff, thank you for your support and input during the entire course of my studies, God bless you.

I am grateful to my UPM classmates, thank you for the time we spent together; it has been great knowing you all, God bless you. Special thanks and appreciation to the UPM Istanbul research group comprising Ms. Dubovyk, Ms. Blakelock Mr. Gavu, Mr. Devkota and Mr. Kowe for an excellent and successful field work exercise in the greatest city in the world, Istanbul.

I would like to thank the entire team at IMP in Istanbul for their support and undivided attention during the field work exercise. Thank you Prof. Dr. Ibrahim Baz, Prof. Dr. Gurcan Buyuksalih, Mr. Ulas Akin, Ms. Guley, Mr. Karadeniz, Dr.Darcin Akin, Ms. Serap Sengul, Mr. Orhan Atkas, Mr. Selahattin Taspinar and all others not mentioned here. I also extend deep gratitude to Prof. Dr. Haluk Gurcek of ITU and Mr. Orhan Demir of PlanOfis for providing insight and data on the 1997-2010 Istanbul Transport Model. Many thanks also to Mr. Batuhan Altun and Ismail Adar from the TCC in Istanbul for providing data and information crucial to this research. To all of you who made my experience in Istanbul a success, teşekkür ederim.

Thank you my fellow Kenyans, ladies and gentlemen of goodwill, you made my stay in Enschede feel like I was at home in Gatundu. Asanteni sana na Mungu awabariki, tena sana. Thegio global momma.

Thank you Ms. Bo, for everything in total, God bless you and reward you forever everywhere you may go from Enschede to Kandara.

And thank you Kuria, Wanjiru, Wangui, Kamata and Githire for your support and for making it possible for me to pursue my MSc degree. I promise to apply the wisdom from my study experience in advancing the interests of Gacunji, kazi iendelee. God bless you.

James Njiraini Gachanja
Enschede, March 2010
# Table of contents

1. Introduction ......................................................................................................................... 1
   1.1. Integrating Land Use and Transport Planning .............................................................. 1
   1.2. Integrating Land Use and Transport Models ............................................................... 1
   1.3. The Four Step Model .................................................................................................... 1
   1.4. Research Problem ........................................................................................................ 2
   1.5. Research Objectives and Questions ............................................................................. 2
   1.6. Justification .................................................................................................................. 3
   1.7. Research Design .......................................................................................................... 4
   1.8. Limitations of the Research ....................................................................................... 5
   1.9. Thesis Structure .......................................................................................................... 5
2. Land Use and Transport: Planning and Modelling ................................................................. 6
   2.1. Planning and Decision Support ................................................................................... 6
   2.2. Land use Planning ....................................................................................................... 6
   2.3. Land use-Transport Interaction .................................................................................. 7
   2.4. Transport planning ...................................................................................................... 8
   2.5. Models in Transport Planning .................................................................................... 8
      2.5.1. Fundamentals of Transport Modelling ................................................................. 9
      2.5.2. Historical development ....................................................................................... 10
   2.6. The Four Step Model (FSM) ....................................................................................... 10
      2.6.1. Trip Generation - *How many trips will there be?* ........................................... 11
      2.6.2. Trip Distribution - *Where do the trips go?* .................................................... 13
      2.6.3. Models of mode split/choice-*How will people travel?* .................................. 14
      2.6.4. Traffic Assignment - *What routes will be used?* ............................................ 15
   2.7. Methodological concerns ............................................................................................. 16
      2.7.1. Trip Generation ................................................................................................. 16
      2.7.2. Trip Distribution ............................................................................................... 16
      2.7.3. Mode Choice/Modal split ................................................................................... 17
      2.7.4. Traffic Assignment ............................................................................................ 17
   2.8. Summary of the FSM .................................................................................................. 17
   2.9. Empirical studies on Accuracy of the FSM ................................................................ 18
   2.10. The Conceptual Framework ....................................................................................... 22
3. Background to the Study Area .......................................................................................... 24
6.9.  Research question 9- Effect size of manipulating FSM using different inputs ...............66
6.10.  Comparison of the spatial distribution of Percent Error Case study links ..................67
7.  Summary of Findings ........................................................................................................69
  7.1.  Summary ......................................................................................................................69
  7.2.  Limitations ....................................................................................................................71
8.  Conclusion and Recommendation ......................................................................................72
  8.1.  Conclusion ....................................................................................................................72
  8.2.  Recommendations .........................................................................................................73
References ................................................................................................................................75
Appendices ..........................................................................................................................78
## List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Research Design</td>
<td>4</td>
</tr>
<tr>
<td>2-1</td>
<td>Concept of land-use-transport interaction</td>
<td>7</td>
</tr>
<tr>
<td>2-2</td>
<td>Supply–demand equilibrium</td>
<td>9</td>
</tr>
<tr>
<td>2-3</td>
<td>The Four Step Transport Model (FSM)</td>
<td>11</td>
</tr>
<tr>
<td>2-4</td>
<td>Conceptual Framework</td>
<td>23</td>
</tr>
<tr>
<td>3-1</td>
<td>Location map of Istanbul depicting the 39 districts</td>
<td>24</td>
</tr>
<tr>
<td>3-2</td>
<td>Population growth trend in Istanbul</td>
<td>25</td>
</tr>
<tr>
<td>3-3</td>
<td>Land Cover Map-Istanbul 1995</td>
<td>25</td>
</tr>
<tr>
<td>3-4</td>
<td>Land Cover Map-Istanbul 2005</td>
<td>25</td>
</tr>
<tr>
<td>3-5</td>
<td>Land Cover Change Map 1995-2005</td>
<td>26</td>
</tr>
<tr>
<td>4-1</td>
<td>Traffic sensor (RTMS)</td>
<td>34</td>
</tr>
<tr>
<td>4-2</td>
<td>Displays how RTMS micro-waves are beamed on the road network</td>
<td>34</td>
</tr>
<tr>
<td>4-3</td>
<td>Traffic intensity map and traffic video capture</td>
<td>35</td>
</tr>
<tr>
<td>5-1</td>
<td>Methodology for research question 1</td>
<td>39</td>
</tr>
<tr>
<td>5-2</td>
<td>Methodology for research question 2</td>
<td>40</td>
</tr>
<tr>
<td>5-3</td>
<td>The BPR function</td>
<td>46</td>
</tr>
<tr>
<td>5-4</td>
<td>Research Methodology</td>
<td>48</td>
</tr>
<tr>
<td>5-5</td>
<td>Location of case study RTMS and links</td>
<td>50</td>
</tr>
<tr>
<td>6-1</td>
<td>Planned Highway network 1996</td>
<td>53</td>
</tr>
<tr>
<td>6-2</td>
<td>Highway network 2006</td>
<td>54</td>
</tr>
<tr>
<td>6-3</td>
<td>Road Density Map for planned highway network 1996-2010</td>
<td>54</td>
</tr>
<tr>
<td>6-4</td>
<td>Road Density Map for actual highway 2006</td>
<td>55</td>
</tr>
<tr>
<td>6-5</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 3</td>
<td>56</td>
</tr>
<tr>
<td>6-6</td>
<td>Percent error for individual links-research question 3</td>
<td>56</td>
</tr>
<tr>
<td>6-7</td>
<td>Spatial distribution of case study links and land cover change -1995 to 2005</td>
<td>57</td>
</tr>
<tr>
<td>6-8</td>
<td>RTMS vs TAZ Forecast PE-Population</td>
<td>57</td>
</tr>
<tr>
<td>6-9</td>
<td>RTMS vs TAZ Forecast PE-Employment</td>
<td>57</td>
</tr>
<tr>
<td>6-10</td>
<td>RTMS vs TAZ Forecast PE-Students</td>
<td>58</td>
</tr>
<tr>
<td>6-11</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 4</td>
<td>58</td>
</tr>
<tr>
<td>6-12</td>
<td>Percent error for individual links-research question 4</td>
<td>59</td>
</tr>
<tr>
<td>6-13</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 5</td>
<td>60</td>
</tr>
<tr>
<td>6-14</td>
<td>Percent error for individual links-research question 5</td>
<td>60</td>
</tr>
<tr>
<td>6-15</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 6</td>
<td>61</td>
</tr>
<tr>
<td>6-16</td>
<td>Percent error for individual links-research question 6</td>
<td>62</td>
</tr>
<tr>
<td>6-17</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 7</td>
<td>64</td>
</tr>
<tr>
<td>6-18</td>
<td>Percent error for individual links-research question 7</td>
<td>64</td>
</tr>
<tr>
<td>6-19</td>
<td>Scatter plot: Model ADT vs Observed ADT-research question 8</td>
<td>65</td>
</tr>
<tr>
<td>6-20</td>
<td>Percent error for individual links-research question 8</td>
<td>65</td>
</tr>
<tr>
<td>6-21</td>
<td>Spatial distribution of case study link percent error for different FSM runs</td>
<td>68</td>
</tr>
</tbody>
</table>
List of tables

Table 1-1: Research objectives and questions..................................................................................3
Table 3-1: Trip rates in Istanbul for the years 1996 and 2006 .................................................................28
Table 3-2: Comparison of the 1997-2010 and 2007-2023 Transport Master Plans for Istanbul........29
Table 4-1: Data acquired during field work in Istanbul ........................................................................32
Table 5-1: Employment and un-employment rates by Income group –Istanbul, 2006. ..................41
Table 5-2: Trip production rates classified by trip purpose and employment status..........................41
Table 5-3: Trip attraction rates classified by trip purpose ....................................................................42
Table 5-4: Variables used for trip production and attraction in the 2007-2023 Transport Model ..........43
Table 5-5: Average occupancy rates and PCU for Vehicles, Istanbul .................................................45
Table 5-6: Contents of RTMS data in MS excel format .......................................................................49
Table 6-1: Accuracy assessment for forecast vs actual land use and socio economic inputs ..........52
Table 6-2: RTMS vs TAZ Forecast PE ...............................................................................................58
Table 6-3: Population, employment and student numbers for Istanbul between 1996 and 2006 .......63
Table 6-4: Results of within-subject contrast statistics .......................................................................67
Table 7-1: Summary FSM accuracy assessment ..................................................................................69
## List of acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>Average Daily Traffic</td>
</tr>
<tr>
<td>AE</td>
<td>Absolute Error</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>FSM</td>
<td>Four Step Model</td>
</tr>
<tr>
<td>IMM</td>
<td>Istanbul Metropolitan Municipality</td>
</tr>
<tr>
<td>IMP</td>
<td>Istanbul Metropolitan Planning and Urban Design Centre</td>
</tr>
<tr>
<td>ITC</td>
<td>International Institute for Geo-Information Science and Earth Observation</td>
</tr>
<tr>
<td>ITU</td>
<td>Istanbul Technical University</td>
</tr>
<tr>
<td>JICA</td>
<td>Japanese International Cooperation Agency</td>
</tr>
<tr>
<td>KMZ</td>
<td>Keyhole Markup Language (zipped) files</td>
</tr>
<tr>
<td>LUT</td>
<td>Land use-transport interaction</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td>MPE</td>
<td>Mean Percent Error</td>
</tr>
<tr>
<td>O-D</td>
<td>Origin to Destination</td>
</tr>
<tr>
<td>PE</td>
<td>Percent Error</td>
</tr>
<tr>
<td>PCU</td>
<td>Passenger Car Units</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>%RMSE</td>
<td>Percent Root Mean Square Error</td>
</tr>
<tr>
<td>RTMS</td>
<td>Remote Traffic Microwave Sensor</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>TAZ</td>
<td>Traffic Analysis Zone</td>
</tr>
<tr>
<td>TCC</td>
<td>Traffic Control Center</td>
</tr>
<tr>
<td>V</td>
<td>All vehicle count</td>
</tr>
<tr>
<td>VL</td>
<td>Long vehicle count</td>
</tr>
</tbody>
</table>
1. Introduction

This chapter gives a general presentation of the main topics, scope and components of the thesis including the research problem, research objectives, research questions and justification. It discusses the theme of the thesis which revolves around integrated land use and transport modelling. The focus of this research is on transport modelling with emphasis on evaluating the accuracy of the Four Step Model (FSM).

1.1. Integrating Land Use and Transport Planning

Urban land use and transportation are known to interact and influence each other yet it is common practice for urban planners to prepare plans without regard for transportation while transportation planners design and implement transportation strategies without consideration of urban development (Henscher et al., 2004). As a result the interventions from these separate plans and strategies are often sub optimal or conflicting. For example, the findings from a study by Lefèvre (2009) on integrated land use and transport planning indicate that the savings obtained from the integration of transport and land use policies are much more important than the savings obtained from a transport investment alone. It was observed that energy consumption and emissions of green house gases would only increase by 9% if the land use and transport policies were integrated whereas they would increase by between 50-70% if integration was not implemented. Furthermore, the chances of conflicting land use and transport plans and projects are significantly reduced where shared land use-transport (LUT) visions and concepts are in place (Te Brömmelstroet and Bertolini, 2008).

1.2. Integrating Land Use and Transport Models

One of the key barriers to integration of land-use and transportation planning is the lack of a “common language”, that is, tools, instruments and indicators that can support planners from both domains in developing integrated land-use and transportation strategies (Te Brömmelstroet and Bertolini, 2008). Despite this barrier, there have been efforts by urban researchers to formalize the relationship between land use and transport using mathematical, statistical and logical models capable of predicting changes to transportation and land use systems as the result of policy measures in both fields (Iacono et al., 2008).

Since the focus of this research is on the transport component of the integrated land use and transport modelling theme, the discussion that follows introduces the FSM.

1.3. The Four Step Model

In seeking to model travel demand and predict the effect of transportation system changes on the transportation network, transport planners have relied on the FSM (Hensher and Button, 2000). This model sequentially comprises four sub-steps, namely: trip generation, trip distribution, mode choice,
and traffic assignment. The model predicts future travel demand and its implications on the transportation system. Inputs into the model include existing and planned land use and socio economic variables as well as the planned road network. Through a sequence of computations the model is able to predict the future volume of traffic on each transport network link as well as provide other indicators such as total number of trips, vehicle speeds, travel times and congestion per link among others (Hensher and Button, 2000; Bates, 2000; McNally, 2000). Transportation planners use these indicators to evaluate the impact of proposed policies, plans or projects on the transport system performance. On this basis, the FSM is seen as an important planning and decision support system.

1.4. Research Problem

In section (1.2) above it was observed that there have been efforts to formalise the relationship between land use and transport using models. However, the FSM has been identified as weak point in these efforts because it does not explicitly consider land use effects and depends upon land use methods and models to make demographic and socio-economic projections. Land use data with base year figures for population of different types, employment activities, shopping spaces, educational and recreational facilities and location choices are used in the model to estimate the total number of trips originating by and attracted to a zone of the study area (trip generation). This data is used as exogenous input and is not changed during the modelling process.

Another concern is that the sequential approach of the FSM creates a situation where no feedback is provided from one step to another. This means that results of the final step (assignment) of the FSM do not have an influence on the previous steps (trip generation, distribution and modal split). This leads to inconsistency between the inputs and the outputs of the model and hence inaccuracy in representation of reality. The FSM in its traditional form does not really capture the interaction between land use and transport (Bates, 2000; McNally, 2000; Tillema, 2004).

This research evaluates the application of the traditional FSM and assesses factors that influence its accuracy in forecasting travel demand using the case of Istanbul, Turkey. The purpose is to demonstrate the weakness of the FSM and make suggestions for its improvement based on evidence.

The research is important because it highlights the factors that need to be considered when using the FSM and how considerations for the interaction between land use and transport can be included in the model.

1.5. Research Objectives and Questions

The research objectives and questions that guided this research are presented in table 1-1 below. They are based on the research problem presented in section 1.4 above. The general objective is to evaluate the FSM and assess its accuracy. The city of Istanbul in Turkey is used as the case for this study.
Table 1-1: Research objectives and questions

<table>
<thead>
<tr>
<th>General Objective</th>
<th>Specific objectives</th>
<th>Research Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>To evaluate the FSM and assess its accuracy using the case of Istanbul Turkey.</td>
<td>1. To assess the validity of FSM inputs</td>
<td>i. How do projected land use and socio economic projections for 2006 match with the actual situation in 2006?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. What differences or similarities exist between the planned and actual highway networks for 2006?</td>
</tr>
<tr>
<td></td>
<td>2. To assess the accuracy of the FSM outputs based on different inputs.</td>
<td>iii. How do the FSM traffic volumes based on projected land use and socio economic data compare with observed traffic volumes on selected links?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. Do the FSM traffic volumes based on actual 2006 land use and socio economic data have a better match with observed traffic volumes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v. Do the FSM traffic volumes based on the actual highway network in 2006 have a better match with observed traffic volumes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vi. Do FSM traffic volumes based on 2006 parameters have a better match with observed traffic volumes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vii. Do FSM traffic volumes based on disaggregated input data have a better match with observed traffic volumes?</td>
</tr>
<tr>
<td></td>
<td>3. To determine the effect size of different FSM inputs on model outputs.</td>
<td>viii. Does the implementation of feedback lead to a better match between the FSM and observed traffic volumes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ix. Based on the previous questions iii-viii above. Which FSM input has the most significant effect on the model outputs?</td>
</tr>
</tbody>
</table>

1.6. Justification

No one can ever know the future but we can often better prepare for it. Planners require techniques and tools to improve the accuracy of their predictions and to support them in dealing with the problems of risk and uncertainty. Van Delden et.al., (2009) note that plans and policies should be designed to incorporate and work with uncertainty rather than ignore it.

Barredo and Demicheli (2003) observe that estimating future socioeconomic and environmental impacts of existing spatial plans and policies on urban development and the consideration of alternative planning and policy scenarios for impact minimisation are of particular interest for urban planners. They further note that there are complex rules at work that make difficult the forecasting of urban dynamics. Given the complexity of urban dynamics, Te Brömmelstroet and Bertolini (2008) suggest that focus should be on developing a common land use transport language.

De Bok (2009) holds that integrated land use and transport models enable policy makers to foresee and evaluate the effects of transport and urban plans hence enabling the solution of common planning problems. Geurs and van Wee (2004) argue that the use of land use and transportation interaction...
models is necessary because of the inclusion of feedback mechanisms between land-use, travel demand and accessibility. However, they note that not many evaluation studies of the accessibility impacts of land-use and transport projects are based on such models.

This research is important because it will demonstrate the weaknesses of the traditional FSM making suggestions for its improvement and showing how it can be best applied within the framework of integrated land use and transportation modelling.

1.7. Research Design

The research comprises three phases as shown in figure 1-1. The pre-field work phase relies largely on literature and supervised discussions to formulate the research problem and define the aim and objectives of the study. It also comprises formulation of research questions and establishment of data requirements. The second phase involves field work where the required data was collected together with information pertinent to the research. The research matrix that guided the fieldwork is presented at the end of this report as appendix 1.

The post field work phase involved data analysis where the FSM was implemented and evaluated for accuracy.

![Figure 1-1: Research Design](image-url)
1.8. Limitations of the Research

A number of challenges and limitations were encountered which had an effect on the overall outputs. The major limitations experienced during field work included language barriers and difficulty in accessing historical data for the year 1997 because the organisational set up at IMP had changed since then and the individuals who were privy to the data and information had changed stations. Limitations during the post-field work phase included difficulties in data conversion, technical difficulties in digital transfer of observed traffic data from the TCC due to immense size, challenges in accessing and applying the FSM within the available transport modelling software packages and finally methodological limitations were experienced in data preparation and analysis mostly on account of gaps in the historical data for 1997.

1.9. Thesis Structure

The thesis is divided into seven chapters. Chapter 1 deals with the introduction highlighting the research problem, objectives and questions. It also addresses the research design. Chapter 2 presents the literature review with a focus on the FSM mentioning its workings and historical development as well as empirical studies on it.

Chapter 3 looks into the data collection methodology and approach. It highlights the activities during field work as well as the sources and nature of data.

Chapter 4 presents the study area giving a brief introduction on the background and process of land use and transport development in Istanbul.

Chapter 5 presents the methodology for data analysis used in this study. It provides a step by step account of how the data was analysed in reference to each research question.

Chapter 6 presents the results of data analysis for each research question. A discussion to interpret each result is further provided.

Chapter 7 presents a summary of the findings based on a synthesis of the results and discussion.

Chapter 8 deals with the conclusions and recommendations of the study. These are based on the research objectives and conceptual framework.
2. Land Use and Transport: Planning and Modelling

This chapter presents the literature review and gives the theoretical and empirical context of this research. It focuses mainly on the traditional four step model (FSM) highlighting its fundamental elements, its historical development, the methodological issues of concern as well as areas for improvement. The purpose is to provide insight into the model and its workings. This section also briefly reviews land and transportation. The final part of this section delves into empirical studies that have been conducted in relation on the FSM focusing on accuracy assessment and model validation. The literature review provides the rationale, scope, and methodology to be used in the analysis stage of this research. A conceptual framework to guide the empirical scope of the study is also presented.

2.1. Planning and Decision Support

Planning is concerned with the future, Sharifi and Rodrigues (2002) define planning as a continuous process which involves decision, or choices about alternative ways of using available resources, with the aim of achieving particular goals at some time in future. Couclesis (2004) mentions that the three basic questions of planning – what may be, what should be and what could be require serious thought about the future. This means that planners require methodologies and techniques to help them anticipate the future and plan ahead.

Planning support is defined as all the professional help in the form of dedicated information, knowledge and instruments that people actively involved within formal spatial-planning practices use to improve quality, increase ease and speed of performance in carrying out their planning tasks and activities (Geertman, 2006; Batty, 2007).

The following section briefly presents the concept of land use and land use planning. It is not exhaustive because the focus of this report is on the FSM but it serves to show the reader how the element of land use is dealt with in urban planning.

2.2. Land use Planning

The term land use is used broadly to signify the human activities on land consisting of modifications of the natural environment into a built or non-built environment. There are three general land uses in urban areas namely residential, commercial and industrial. Other land uses include specialized functions like national security uses for the military as well as government functions (Pacione, 2005). Land-use planners play a key role in guiding the process of community growth and development by helping a community achieve its vision for the future and overseeing the land development process (Schock, 2000).

Various scholars have described how land use is organized in space, how it develops and how it grows over time using theories and models as explained in Pacione (2005). Most models of land use often refer to a central business district which is the core of the city and acts as the focal point for other land uses (Pacione, 2005). The study of urban morphogenesis or town plan analysis has sought to advance
from description and classification of urban forms to analysis of the causal forces underlying changes in pattern of urban land (Pacione, 2005). Conzen (1960) as cited in Pacione (2005) divided the urban landscape into three main elements of namely, town plan, building forms and land use, noting that land use is most susceptible to change. Land-use models often incorporate a variety of land use categories as inputs and thereby can account for different sub-classifications of urban and nonurban land use and these models seek to project changes in land use (Schock, 2000).

2.3. Land use-Transport Interaction

Land use and transportation systems interact and the link to this interaction is through the concept of accessibility which is defined as the extent to which the land-use and transport systems enable groups of individuals to reach activities or destinations by means of a combination of transport modes (Geurs and van Wee, 2004). Figure 2-1 below illustrates the interaction between land use –activities- transport and accessibility. From this it can be seen that the interaction between land use, accessibility and transport is cyclical (Geurs and van Wee, 2004; Tillema, 2004).

![Figure 2-1: Concept of land-use-transport interaction (source: Tillema, 2004)](image)

The transportation system has an impact on accessibility and accessibility in turn influences the land use system for instance by having an impact on location choice. Land use generates activities such as housing, working, shopping and leisure. The need for individuals to participate in these activities and the spatial distribution of these activities generates the need to travel which has an impact on travel demand leading to impacts on the transport system. It should be noted that the reverse also holds as true, where land use has an impact on accessibility and this leads to changes in the transport system(Geurs and van Wee, 2004; Tillema, 2004).

Based on the concept of accessibility it is observed that the spatial structure of land use can influence the modes of transport used. For instance the convenience of private cars for those able to afford them and the extent to which retailing, concentrations of job opportunities and new residential developments encourage their use suggests that it will be difficult to shift car users to other transport modes (Pacione, 2005). Consequently land use and urban development influences transport, Pacione (2005) mentions that most cities have been transformed by the growth in the number of road vehicles.

It is noted that there are four components of accessibility namely; transport, land use, temporal and individual. Accessibility measures are of prime importance to land use and transport planners because they act as indicators of the impact of land-use developments, transport developments and policy plans on the functioning of the society (Geurs and van Wee, 2004).
In summary land use planners try to forecast future demographic and socio-economic patterns, and land use changes. They make use of a variety of methods and models such as comprehensive plans, use of experts, allocation rules, statistical methods, and regional economic models among others. Travel demand forecasting models provide forecasts of future travel patterns based on projections of future demographic and socio-economic patterns derived from land use methods and models.

The following section focuses on the field of transport planning; it provides a general description of transport planning and mentions its objectives and role.

2.4. Transport planning

According to Beimborn (1995) transportation planning is a process that develops information to help make decisions on the future development and management of transportation systems, especially in urban areas. The author notes that it involves the determination of the need for new or expanded facilities, their location, their capacity and the management of their demand with a forecast period of 15 to 25 years into the future.

The role of transport planning is well described by Ortúzar and Willumsen (2001) as that of ensuring the satisfaction of certain demand for person and goods movements with different trip purposes, at different times of day and the year, using various modes, given a transport system with a certain operating capacity. They note that the transport system itself can be seen as made of infrastructure, a management and control strategy and a set of transport modes and their operators.

In summary, literature suggests that the main objectives of transport planning involve addressing traffic congestion, travel growth, land use-transportation coordination and environmental compliance issues among others (Beimborn, 1995).

The following section presents a review of transport modelling, it begins with a general overview of models and how they function, and then proceeds to present the history and fundamentals of travel demand modelling. The FSM is also discussed in considerable amount of detail to provide an understanding of its components and functionality.

2.5. Models in Transport Planning

Models are the basic tool of analysis for planners working in the fields of transportation and land use forecasting. Over the course of many years, professional spatial planners have used a plethora of methods and tools to support their various planning activities (Geertman, 2006). Models continue to represent an important means of testing theories and developing knowledge about the behaviour of urban systems and in this way provide support for planning and decision making (Iacono et.al., 2008).

One of the key attributes of models is the fact that they simplify and abstract the reality to make the decision making process easier (Waddell and Ulfarsson, 2004). Within the field of transport, modelling focuses on the ways in which one can simplify and abstract important relationships underlying the provision and use of transport and it deals with the study of the behaviour of individuals (Hensher and Button, 2000). It has been observed that transport models may be complex
but this complexity is traded off by the value of these models in offering a ‘common ground’ for discussing policy (Ortúzar and Willumsen, 2001).

In practice, transport demand forecasting models are used to help in planning transport infrastructure and in anticipating exogenous changes in travel demand patterns (Fox et al., 2003). It is interesting to note that those models that have exerted the greatest influence have been elegant and simple (Hensher and Button, 2000). Models are developed for different reasons and are intended for the consideration of different users. Hensher and Button (2000) note that academicians focus on developing transport models that explore the technical efficiency of a transport system while policy makers are much more interested in the impacts of various transport actions on different societal groups, political modelling is perceived in the sense that decision makers often adopt models because they offer a framework similar to their political ideology.

The importance of models cannot be underscored further than by the fact that future transportation plans are based on what the models say will happen rather than on what individual people may think will happen (Beimborn, 1995). It appears that people are more confident in making decisions that are backed by these models. However, over reliance on models might be handicapped in the sense that good predictions require good models yet in some cases pure chance can lead to fairly accurate predictions (Hensher and Button, 2000).

Transport modelling draws some comparisons to economic theory in terms of its conceptualisation of demand and supply. This is described albeit briefly in the following section, the indicated references are recommended for a more in-depth insight.

2.5.1. Fundamentals of Transport Modelling

The application of the concepts of demand and supply in transport economics require that they be defined with more care than is generally the case in economic theory. This is because travel is described as a derived demand due to the fact that it is demanded as a consequence of the desire to partake in activities in different locations (Bates, 2000; Ortúzar and Willumsen, 2001). Furthermore, the distinctive characteristic about transport supply is that it is a service and not a good (Ortúzar and Willumsen, 2001).

Fig 2-2 below displays the interactive relationship between demand and supply; demand is a function of cost which is plotted on the vertical axis. The rationale of this is that travelling will cost money and involves expenditure of time. Hence in transport economics cost is considered in terms of generalised cost which is defined as a combination of cost and time. Time is converted to money units by means of the “value of travel time savings” (Bates, 2000).

![Figure 2-2: Supply –demand equilibrium (Bates, 2000).](image-url)
The focus of the supply relationship in transport has been on the non-monetary items and on time in particular because many of the issues of demand with which transport analysts are concerned with affect performance of the transport system (Bates, 2000). “The supply function reflects the response of the transport system to given level of demand. If the estimated demand were loaded onto the system, the supply effect would be the deterioration in highway speeds, as traffic volumes rise” (Bates, 2000).

Ortúzar and Willumsen (2001) note that congestion is one of the most important features of transport supply which arises when demand levels approach the capacity of a facility and the time required to use it (travel through it) increases well above the average under low demand conditions.

The actual demand predicted to arise as a result of a transport strategy is the outcome of the equilibrium process. The actual volume of travel must be where the two curves cross and this is known as the equilibrium point (Bates, 2000). For more on demand and supply in transport modelling see Bates (2000) and Ortúzar and Willumsen (2001).

The following section briefly introduces the history of the FSM. A more detailed account of the history and development of the FSM can be obtained from the indicated citations.

### 2.5.2. Historical development

The fundamentals of transport modelling can be traced back to the Detroit and Chicago transportation studies in USA in the 1950s (Bates, 2000). The last 40 years has seen the development and application of a large number of statistical and mathematical procedures directed towards improving the understanding of the behaviour of agents who make decisions that impact the transport system (Hensher and Button, 2000).

Bates (2000) notes that transport models have evolved from many disciplines; most notably economics, psychology, geography, sociology, and statistics. He notes that their initial focus was on estimating peak demand for transport services and predicting what provision should be made for this peak. The FSM was originally designed for the analysis of urban highway investment although substantial improvements have now made it usable in the public transport field (Bates, 2000).

The following section delves into the FSM discussing how it works and highlighting the main elements of its composition.

### 2.6. The Four Step Model (FSM)

Transport planners seek to model travel demand and predict the effect of transportation system changes in terms of traffic or ridership for a particular link, route or entire transportation network. Over the years they have relied on the traditional four step transport model (FSM) to forecast travel demand (figure 2-3 below). This model sequentially comprises four sub-steps: trip generation, trip distribution, modal split/choice, and traffic assignment.

The initial step, *trip generation*, makes use of land use and socioeconomic data such as population, household size, and household income to determine the number of trips produced by and attracted to analysis zones. The second step referred to as *trip distribution* determines the spatial distributions of trips that are generated in the first step. The third step, modal split, shares the trips into different modes
of transport such as private car, bus, train and walking. The fourth and final step, traffic assignment, allocates the trips in different modes to the transportation network (Levinson and Kumar, 1994; Tillema, 2004).

The FSM can be viewed in two stages, the first stage relates to characteristics of the traveller and the land use –activity system in which travel demand is measured. The second stage loads this demand onto the transportation network (McNally, 2000).

The following sections give detailed reviews of each of the four models that comprise the FSM. The rationale, assumptions and formulations are outlined. The presentation follows the sequential nature of the FSM as applied in practice.

2.6.1. Trip Generation- How many trips will there be?

The first step in travel forecasting is trip generation. In order to undertake this step, the study area is delineated in to special geographic units called traffic analysis zones (TAZs) from which transportation planning data is derived. Information from land use, population and economic forecasts is used to estimate how many trips will be made to and from each zone (Bates, 2000; Beimborn, 1995; Wells, 1975). The quantities of the trips produced and attracted to each zone are taken as symmetrical meaning that the total trips attracted should equal the total trips produced (Bates, 2000).

This step translates the four stage model from activity-based to trip-based and separates each trip into a production and an attraction; it should be noted that when determining trip generation, network performance measures such as congestion do not influence the frequency of travel which is a weakness of the model (McNally, 2000). The trip generation stage essentially defines total travel in the region and the remaining steps are effectively share models sharing the trips to different areas, modes of transport and links of the transportation network (McNally, 2000).
It should be understood that the nature of travel is such that, the basic requirement for travel is produced at one end of the trip, typically the home, and is then attracted to a particular zone which will meet the purpose of the journey (Bates, 2000). In terms of the level of modelling, trips can be modelled at the zonal, household, or person level, with household level models most common for trip productions and zonal level models most common for trip attractions (McNally, 2000). In practice it has been easier to develop models for the productions than for the attractions since they are taken as well defined, but the attractions are merely an indication of the relative attractiveness of different zones (Bates, 2000).

When modelling trip production we need to define the time periods to which the travel relates, the set of modes included, the distinction of person type, the distinction of trip purpose and the different levels of household car ownership (Bates, 2000). The choice of explanatory variables to be used in estimating the trip production model is constrained according to model significance, policy sensitivity and forecast ability (McNally, 2000). It is common practice to separate the trips by purpose where the trip purposes normally used are; home based work trips (work trips that begin or end at home), home based shopping trips, home based other trips, school trips, non-home based trips (trips that neither begin or end at home), truck trips and taxi trips (Beimborn, 1995).

According to Bates (2000) the model of trip production is written as shown in equation 2.1 below:

$$T_i[k] = f(X^k[C^k_i,..])$$  \hspace{1cm} (2.1)

Where \((k)\) is a segmentation of the population (typically a combination of journey purpose and person/household characteristics), \((i)\) is the origin, \((X^k)\) is a vector of characteristics for segmentation \((k)\) and \((C^k_i)\) is the composite cost of travelling from the origin. Calibration of the model is carried out using information from local area or national household surveys (Bates, 2000; Wells, 1975).

The two main methods of computing trip generation are zonal regression analysis and category analysis. The latter was proposed as an improvement to zonal regression as it recognised the value of identifying different categories of household and having different trip rates for each category and provides a reasonably accurate measure of trip frequency (Bates, 2000; McNally, 2000; Wells, 1975). Regression analysis is however used for determination of trip attractions (Wells, 1975).

When modelling, Bates (2000) notes that variation in persons between children, employed adults, non-employed adults and persons under retirement age contributes to variation of trip rates by purpose and this should be recognised by the model developer. In terms of data acquisition for model development, it should be noted that the use of household and person characteristics means that there are taxing data requirements which leads to the reliance on other techniques such as predictions of car ownership (Bates, 2000).

Trip production models have proved reasonably stable over time and attempts to introduce the effects of accessibility on the level of trip making have been unsuccessful, hence most models rely on the external effects such as land use and socio-economic conditions for estimation of trips produced in the study area (Bates, 2000).
Upon the computation of the trips produced and attracted in each of the TAZs, the next step is to establish spatial distribution of these trips. The following section presents the model of trip distribution. It should be noted that the trips in this stage are in the form of person trips.

### 2.6.2. Trip Distribution - *Where do the trips go?*

The process of trip distribution links the trip ends (productions and attractions) from the trip generation step to form an origin-destination matrix (McNally 2000; Bates 2000; Beimborn, 1995). This step is used to represent the process of destination choice (Beimborn, 1995). Bates (2000) notes that the trip distribution modelling process should be viewed as that of reproducing a matrix of movements in which the number of trips in the cell of the matrix are related to the characteristics of the production zone \((i)\); the characteristics of the attraction zone \((j)\); and the characteristics of the “separation” or “cost” of travel, between zones \((i)\) and \((j)\).

According to Wells (1975) “the principle which is applied in trip distribution is that trips will be made to a given zone in direct proportion to its relative attractiveness and in some form of decreasing proportion to the separation of the attracting zone from the originating zone.”

The principle is expressed as shown in equation 2.2 below (Wells, 1975):

\[
T_{ij} = \frac{G_i A_j}{dn}
\]  

(2.2)

*Where:*

- \(T_{ij}\) = trips from zone \(i\) to zone \(j\)
- \(G_i\) = total number of trips generated to zone \(i\)
- \(A_j\) = total number of trips attracted to zone \(j\)
- \(d\) = distance or other measure of spatial separation
- \(n\) = a constant (usually assumed to be between 1 and 2.5)

The relationship is more usually expressed as equation 2.3 below (Wells, 1975):

\[
T_{ij} = k G_i A_j f(C_{ij})
\]  

(2.3)

*Where:*

- \(T_{ij}\), \(G_i\), \(A_j\) - are as before in equation (2.2) and
- \(f(C_{ij})\) = is a function of the separation of the zones \((i)\) and \((j)\) (a deterrence function/trip decay function which is based on the generalized cost of the journey from \((i)\) to \((j)\)).
- \(k\) = is a constant.

A matrix of inter zonal trips \((T_{ij})\) can be completed when \((f)\) and \((k)\) are determined and the trips as well as \((C_{ij})\) are applied (Wells, 1975). The model is commonly referred to as the “gravity” model with reference to the Newtonian law of gravitational attraction (Bates, 2000).
In a model where both origins and destinations are assumed known the problem is essentially confined to the estimation of a suitable $f(C_{ij})$. “The earliest forms of the model used zonal population or employment weights for $(G_i)$ and $(A_j)$ and simple forms for $f(C_{ij})$ based on distance” (Bates, 2000).

There are three types of decay functions namely, the power function, the exponential function and the gamma function, see Wells (1975) and Bates (2000) for more on these functional forms. The most commonly used measure of deterrence is the inter-zonal generalised cost. According to Wells (1975) this is the cost that the traveller attaches to travel from point A to B. This may include driving cost, car operating cost, parking charges and waiting time among others. The estimated function is assumed to capture underlying travel behaviour and to be stable in the future to allow its use in forecasting (McNally, 2000).

Wells (1975) observes that the generalised cost for a specific mode can be expressed as equation 2.4:

$$C_{ij} = a_1 t_{ij} + a_2 e_{ij} + a_3 d_{ij} + p_j (+\delta)$$

(2.4)

Where:

$(C_{ij})$ - generalised cost

$t_{ij}$ - Travel time from i to j

$e_{ij}$ - Excess time (access, waiting, etc)

$d_{ij}$ - Distance from i to j

$p_j$ - Terminal cost at destination end of trip(j)

$\delta$ - The inherent modal handicap – (also referred to as fudge factor representing such immeasurable factors as comfort and convenience).

$a_1, a_2, a_3$ - Constants representing the values of the components

The distribution model has been reinterpreted in terms of discrete choice theory and it has been mentioned that it is appropriate to assume that the model is a destination choice model, distributing a known total of trip productions from each zone among the attraction zones (Bates, 2000).

The next step after trip distribution is modal split; a discussion on modal split is presented in the following section.

2.6.3. Models of mode split/choice-How will people travel?

Once the origin and destination of trips has been established in matrix form, the next step is to convert the person trips into vehicle trips by applying the modal split model. At this stage of the FSM trips between a given origin and destination are split into the different modes of transport such as trips using transit, trips by car pool or as automobile passengers and trips by automobile drivers (Beimborn, 1995; Wells, 1975). Calculations are conducted that compare the attractiveness of travel by different modes to determine their relative usage (Beimborn, 1995). Mode choice effectively factors the trip tables from trip distribution to produce mode-specific trip tables (McNally, 2000).

According to Ortúzar and Willumsen (2001) the choice of transport is probably one of the most important classic models in transport planning because of the key role played by public transport in
policy making. It also has implications on the general efficiency with which we can travel, the amount of urban space devoted to transport function, and whether a range of choices is available to travellers. Mode choice is influenced by the characteristics of the trip maker, the characteristics of the journey and the characteristics of the transport facility (Ortúzar and Willumsen, 2001).

According to Bates (2000) the model of mode choice/split may be written as equation 2.5:

\[
P_m[ij : k] = f(C^k_{ij}, C^k_{ij(m)})
\]  

\(Where:\)

- \(k\) - A segmentation of the population;
- \(i\) and \(j\) - The origin and destination of the journey;
- \(m\) - Is the mode

\(P_m[ij : k]\) - Is the proportion of all travellers of type \(k\) moving between \((i)\) and \((j)\) who use mode \(m\);

- \(C^k_{ijm}\) - is the associated cost and \{\(m\)\} is the set of modes being considered.

The chief sources of variation in the models used in practice are the number and type of modes actually distinguished and the detail of the “generalised cost” \(C^k_{ijm}\) (Bates, 2000).

While formulating the model, assumptions have to made that certain types of households are captive to public transport and hence mode choice tends to be confined to predicting the proportion of persons assumed to have access to a car but who actually use public transport (McNally, 2000; Bates, 2000; Wells, 1975).

Bates (2000) notes that an S-shaped curve where the probability of choosing the mode vanishes when its costs are greatly in excess of the costs of the other mode but which allows reasonable sensitivity when the costs are comparable is required. In practice, there is a preference for the logit model and the most common model estimated is the nested logit model (McNally, 2000). It is more preferable than the multinomial logit model because the modes may be inherently similar in some cases (Bates, 2000).

The following section presents the trip assignment step of the FSM. This is the final step of the model where the vehicle trips from the modal split model are assigned to the transportation network.

### 2.6.4 Traffic Assignment - What routes will be used?

Traffic assignment is the step that assigns trips to the specific path that they use to travel from their origin to their destination once they have been split into highway and transit trips (Beimborn, 1995). Modal O-D trip matrices which are built by the application of average car and bus occupancy rates relating to trip purposes are loaded on the modal networks. It should be noted that public transport network is separated from the highway network while undertaking the assignment (Bates, 2000; McNally, 2000; Wells, 1975).

According to Bates (2000) assignment has a number of separate processes, namely, choice of route(or path) for each \((i – j)\) combination; aggregating \((i-j)\) flows on the links of the chosen paths; dealing with supply –side effects (capacity restraint) as a result of the volume of link flows relative to capacity; and
obtaining the resulting cost for each (i-j) combination. He further adds that the produced travel matrices are on annual average day basis and assignment tries to relate flows on links to sensible definitions of “capacity”.

There are different methods for assignment, namely, the all or nothing assignment, capacity restraint assignment and multi route assignment. The latter two techniques take account of the unlikelihood of all trips using the single least-cost route while the multi-route method assumes that the driver does not know which is the least cost route and depends on random selection of links (Wells, 1975). The capacity restraint technique depends on the relationship between the volume of traffic and the speed at which traffic can move on a link. It is mentioned that the best practice is converging on equilibrium methods of traffic assignment (Bates, 2000; Wells, 1975).

It is important to note that supply effects will change the costs of travel and that the costs output from the assignment process are inconsistent with those used to drive the distribution and modal split models and this is one of the shortcomings of the FSM (Bates, 2000). The author adds that assignment deals with managing the interface between demand (a pair of zones) and supply (at the network level). The traffic assignment stage produces traffic volumes on links, the output of this stage is what is mainly applied in evaluation of policies and some of the key effects considered are congestion, accidents, travel times and air pollution emissions (Bates, 2000).

The following section looks at the methodological concerns of the FSM showing the weaknesses that contribute to its inaccuracy.

2.7. Methodological concerns

The main methodological concerns for each sub step of the FSM are presented briefly. It should be noted that this review is not exhaustive of all concerns. For a more detailed description of the methodological weaknesses of the FSM, please refer to the indicated citations.

2.7.1. Trip Generation

There are a number of methodological concerns attributed to the trip generation step of the FSM. For instance the interdependency in trip making is not considered yet decisions of one household member are dependent on others. Another concern is that there are limited trip purposes used in the trip generation model and combinations of trips are ignored. Furthermore there are feedback as well as cause and effect problems relating to how the FSM calculates trips as a function of factors that in turn could depend on how many trips there are (Beimborn, 1995). All these factors amongst others not highlighted here lead to errors that are propagated through to the other stages of the FSM and contribute to its overall inaccuracy.

2.7.2. Trip Distribution

The main methodological concerns of the trip distribution step of the FSM according to Beimborn (1995) include the use of constant trip times; the use of automobile travel times to represent ‘distance’; the limited effect of social-economic-cultural factors and the lack of feedback due to the sequential nature of the FSM. These problems reduce the accuracy of the model (Ortúzar and Willumsen, 2001).
To add on to the above mentioned weaknesses, Bates (2000) mentions that the deterrence function used at this step of the FSM poses a problem because it attempts to explain a large amount of variation using a very small number of parameters which means that the matrix produced is not sufficiently realistic to carry forward to the remaining stages of the model. He proposes the formulation of the trip distribution problem in terms of destination choice, use of additional parameters, use of observed matrices or the introduction of a number of constants to as solutions to the problem. He also adds that the distribution model is a major weakness in the FSM.

2.7.3. Mode Choice/Modal split

It has been observed that the main methodological concerns of the modal choice/ split model are that choice is only affected by time and cost characteristics. More so, there are omitted factors such as crime, safety and security. Personal factors affecting modal choice are not generally taken into account and there is usually no treatment of walk or cycle modes. Another concern is that access times are simplified and the model applies constant weights where the importance of time, cost and convenience is assumed to remain constant for a given trip purpose yet in reality this may not necessarily be the case (Beimborn, 1995).

2.7.4. Traffic Assignment

Some of the main methodological concerns of the assignment stage are that assignment methods generally focus on link travel times ignoring or placing less emphasis on intersection delays. Capacities are often over simplified neglecting to allow for such things as heavy vehicle movements or highway geometry. Intra-zonal travel is ignored and the zone-network system is a simplification of reality where some links may not be included in the network. Times of day variations are not modelled while the models are unable to represent how travellers often cope with congestion by changing the time they make their trips. Emphasis is on peak hour travel meaning the duration of congestion beyond the peak hour is not determined. Finally variations in travel by time of year or day of the week are usually not considered (Beimborn, 1995). These weaknesses together with the errors from the previous steps are compounded to contribute to the overall inaccuracy of the FSM outputs.

The next section presents a summary of the FSM giving an account of the key issues regarding its formulation, logic and merits.

2.8. Summary of the FSM

The foregoing review of the FSM suggests that there are various factors that affect the models reliability and applicability. To add on to these, the model is not run iteratively to achieve equilibrium because of the heavy computational burden (Bates, 2000).

There is a tautological character to the conventional four-step travel demand modelling process manifested by the fact that the trip distribution and mode split stages depend on estimates of interzonal travel time yet estimates of travel time are not available until after completion of traffic assignment (Deakin et.al., 1993).

Furthermore there is debate about the correct ordering of the mode choice and distribution models and additional stages such as choice of day have been used as well as adequate treatment of supply
feedbacks with the most common feedback being to mode and destination choice (Bates, 2000; McNally, 2000). Bates (2000) adds that the weakest point of the models is their static conception in terms of a lack of ‘knock-on’ effects between successive periods.

Besides these criticisms it has been noted that the four stage model can be used to investigate reasonably precise locations of infrastructure in terms of the impacts on accessibility between specific zones, this is due to the level of detail provided by the networks (Bates, 2000).

Appendix 2 presents the logic of the transportation study process showing the data, procedures and outputs of the transportation study indicating the flow and logic of the steps followed. The following section presents empirical studies that have been conducted on the FSM. The focus is on the validation of model outputs, accuracy assessment and sensitivity analysis.

2.9. Empirical studies on Accuracy of the FSM

The following section presents empirical studies undertaken to validate the FSM outputs. It includes examples of accuracy assessment, sensitivity analysis and comparisons of model outputs against actual observations. These provide the foundation to develop a conceptual framework for this research.

A report by Zhao et al., (2005) investigated the performance of the four step model by measuring the effect of input data on the accuracy of traffic volumes projected for Florida in the USA. Different input data, namely, updated zonal data; the actual highway network and two digit coding system were used to run the model. The modelled traffic volumes were compared with the field collected traffic volumes using the %RMSE. The study found that the zonal data had been underestimated and that the use of updated zonal data improved the accuracy of the model. The traffic volumes based on projected zonal data gave a %RMSE value of 0.2896 while the model results based on the updated zonal data gave a value of 0.2833. It was also found that not all the committed highway improvement projects had been undertaken.

Model results based on the updated traffic network gave a %RMSE of 0.2610 meaning that the use of an updated traffic network had a more significant effect on model accuracy. The use of the two-digit coding system improved model performance giving an %RMSE of 0.2520 and was identified to be the key factor in achieving model accuracy. The study found that the three factors, zonal data, the traffic network and the two-digit zoning system, were not independent meaning that the presence of one factor enhanced the effect of the other.

A similar study by Parthasarathi and Levinson (2008) tested for the presence of inaccuracy in roadway traffic forecasts and identified the reasons for this. The analysis involved looking at the input assumptions (roadway network, socio-economic forecasts and trip rates) that went into creating the forecasts. Errors in the socio-economic inputs that feed into the model, the inability to incorporate shifts in trip generation/travel behaviour and differences between the assumed highway network and the actual in-place network were seen as reasons for forecast inaccuracy.
The study found that there was a trend of overestimation in demographic forecasts where the inaccuracy ratio was greater than 1.0. It was also found that the inability of the travel demand models to incorporate fundamental shifts in travel behaviour such as changes in trip length, trips per capita, trips per household, auto occupancy and persons per household could be an important reason for inaccuracy in traffic forecasts. Similarly it was observed that differences between the assumed networks and in-place networks contributed to inaccuracy in project forecasts. Other factors seen to affect model accuracy included the number of years between the report year and forecast year, the highway type, highway functional classification and roadway direction. The researchers also note that the long-term nature of the forecasting process makes it difficult to anticipate changes and control for errors.

Horowitz and Emslie (1978) undertook a study that compared the measured and forecast traffic volumes on interstate highways. They compared the forecasts of 1975 average daily traffic (ADT) on 78 interstate highway segments with measurements of 1975 ADT on the same segments. The study revealed that the 1968 and 1972 forecasts over-estimated the 1975 ADT by 24% and 21% respectively. They identified errors in traffic assignment techniques and residual effects of the 1973-74 gasoline shortage as the main reasons for the forecasting bias.

Zhao and Kockelman (2002) investigated the stability of transport demand model outputs by quantifying the variability in model inputs. They used inputs such as zonal socioeconomic data and trip generation rates and simulated the propagation of their variation through demand models based on a 25–zone network. Their study found that uncertainty compounds itself over a series of models. Errors in prediction from the earlier stages of the multi stage models such as trip generation were found to amplify across later stages. The simulation results from the study suggest that the trip assignment equilibrium technique may reduce the overall uncertainty.

The study also involved a sensitivity analysis to identify which model inputs are key contributors to uncertainty in model output. It was found that the parameter that had the strongest correlation with link flows is the trip generation rate. The overall outputs were seen to be sensitive to the demographic inputs which in this study were the number of households and employments. The study also made use of regression analysis to identify the most important contributors to overall uncertainty. From this it was found that the major contributors to variation in flow estimates are the parameters from trip generation step and total employment input levels per TAZ.

Bonsall et.al., (1977) undertook research that involved running the transport model several times to investigate extremes of policy and carried out sensitivity analysis on the models. The transport model was calibrated for a coarsely zoned West Yorkshire Region. The model runs tested variation in parameters representing policy alternatives, variations in parameters which are normally regarded as fixed exogenous inputs and variations in minor modifications of model form.

In particular the different type of variables tested included land use inputs, network characteristics, parking fees and trip rates. The indicators of the impacts of policies and parameter values being tested included those related to network usage, trips and trip ends, expenditures and costs, accessibility and consumer surplus measures of benefits. The methodology involved taking one set of policies and parameter values which seemed to be the most likely future (MLF) and plotting all variations against that basis. They undertook analysis on single parameter, two parameter and multi parameter tests.
The MLF values were compared with the base year values to identify the differences. From the results the authors conclude that the method involving execution of a large number of model runs and the presentation of results by comparison with the most likely future is a useful approach. The method was able to show those indicators which should be used with particular caution in relation to a particular model design.

It was found that the mean speed of private transport in the study area was extremely sensitive to the method of assignment used while other indicators were relatively insensitive to this. In other cases it was found that a greater range of indicators were sensitive to model parameters such as the deterrence parameters and the value of time. It was found that the FSM gives an apparently sound “quantified common-sense” view of transport system characteristics in relation to possible policy changes. However, it was observed that for policy changes of great magnitude, the model’s predictions were rarely sensible because certain aspects of behaviour, such as car occupancy levels, are beyond its predictive scope. Consequently, it was observed that system changes relate to variables which are outside the control of policy makers and that the system characteristics are relatively insensitive to variables which can be changed as a matter of policy.

Giuliano (1984) undertook research on the 1960 Los Angeles Regional Transport Study which forecasted transportation needs for the year 1980 and investigated the Atlanta rapid transit system (MARTA) demonstrating some of the pitfalls of travel forecasting. From the Los Angeles study, it was observed that regional population was overestimated by 28 % while regional, employment was underestimated by 13%.

Comparison of the expected and actual trip generation showed large differences in some categories. It was found that trip generation had been underestimated by 21% using an area of 10 000 dwelling units, of which 13.8 % was due to car ownership error, 2.0% due to dwelling unit error and 5.7% due to trip generation rate error. The trip distribution and modal split components of the transport model were not assessed due to lack of data. The traffic assignment outputs were measured in terms of Average Daily Traffic (ADT). It was found that there were cases of overestimation and underestimation; this was explained by the fact that the forecasts represented travel desires given the land use and population assumptions and did not take into account the capacity constraints. Similarly the errors in ADT were seen to be explained by the fact that the road network was not built as planned and that some roads were at capacity. The trip rate and the vehicle miles of travel per capita were seen to have increased more than anticipated.

The Atlanta rapid transit system investigation was based on system ridership estimates of 1962, 1971-A and 1972-B and the 1980 actual observations. It was found that the actual ridership was almost exactly equal to the 1971 system average estimates and about 15% less than the other estimates. It was observed that the overall system productivity would decline due to population and employment decentralisation. The study drew the conclusion that in order to improve the forecasting, transportation must be considered in the context of land use. It was also concluded that the importance of economic and demographic trends that affect travel behaviour must be recognised. The study recommended that expectations in transport forecasting should be put into perspective with the aim of being flexibility.

According to Niles and Nelson (2001) major differences can be observed when actual model outcomes are compared with past forecasts. In their paper titled ‘Identifying uncertainties in forecasts of travel
demand’, they review different transport forecast studies and conclude that there are many forces at work shaping the urban system and increasing its complexity and the uncertainty of forecasts of travel. They found that the aggregate uncertainty attributed to the planning horizon of 20 or more years may lead to highly problematic estimates of travel demand and that it is important to build an accounting of uncertainty and risk needs into the planning process. The paper identified the following sources of uncertainty in travel demand modelling: uncertainty in model design and structure; transportation network uncertainty; demographic and behavioural uncertainty and uncertainty resulting from social/political bias.

Krishnamurthy and Kockelman (2003) investigated the propagation of uncertainty in outputs of a standard integrated model of transportation and land use. Model predictions of residence and work locations were used as inputs to a travel demand model and the resulting travel times were fed forward into the future period’s land use models. Monte Carlo sampling of 200 scenarios were used to accommodate covariance in inputs. The study analysed the variances in land use and travel predictions over time and as a function of input values. The results found that output variations were most sensitive to the exponent of the link performance function, the split of trips between peak and off-peak and several trip generation and attraction rates. The study also found that central point estimates of key model outputs were likely to fall 38% to 50% below or above the mean value. It was concluded that such substantial variation was due solely to standard model parameter and input uncertainties, it was also noted that uncertainty about future and human behaviour also exists and will add further variation.

Wegmann and Everett (2008) in their report titled Minimum Travel Demand Model Calibration and Validation Guidelines for State of Tennessee; mention that model calibration adjusts parameter values until the predicted travel matches the observed travel for the base year while model validation tests the ability of the model to predict future behaviour. They add that a model can be considered validated if the model output and the independent data are in acceptable agreement. The authors identify reasonableness checks and sensitivity analysis as two types of validation checks. Reasonableness checks are defined as tests that evaluate the models in terms of acceptable levels of error and ability to perform according to theoretical as well as logical expectations. Sensitivity tests check the responses of the model to transportation, socioeconomic or political changes and are viewed as the elasticity of a variable.

The authors mention that calibration and validation of travel demand models is essential to accurately model current and future travel for metropolitan areas. They add that there are a number of evaluations and reasonableness checks that can be performed to enhance the travel demand model’s forecasting ability and that these should be performed after each step of the four-step modelling process. They recommend that evaluation and reasonableness checks should be applied during the process of calibrating each individual step of the FSM and this should be followed by overall model validation.

The report outlines the various checks that are undertaken for the different stages of the model, for example, the report mentions that for the traffic assignment stage, the statistics used to make comparisons between model outputs and actual situations include; absolute difference in volumes, percentage difference in volumes, average error, average percent error, standard deviation, R Square, Root Mean Square Error (RMSE) and correlation coefficient. The report also gives the desired values
for these validation measures. It is mentioned that traffic counts on 10% or more of the region wide highway segments being analysed should be used in the validation.

In summary, there are four main factors that account for most of the inaccuracy of the FSM outputs and these are as follows; first, prediction or forecast errors of the land use and socio economic variables used as inputs to the preliminary stages of the FSM; second, the difference between the planned road networks and the actual road network; third, errors in estimating the parameters and rates used in the different steps of the FSM; fourth, the model design/set up and techniques used for running the model.

The FSM was identified to be sensitive to these factors with the observation that socio economic attributes and the network used account for the greatest elasticity of the model outputs. Some of the errors contributing to the overall inaccuracy of the FSM can be explained by the fact that it is difficult to predict human behaviour especially given long forecasting period and the unforeseeable socio-economic, political and environmental dynamics that may occur.

The techniques and procedures that were used to undertake model validation, sensitivity analysis and reasonableness checks were clearly defined in the reviewed studies. They provided the empirical foundation to develop the conceptual framework on which this research is based. The following section discusses the conceptual framework.

2.10. The Conceptual Framework

This section introduces the reader to the conceptual framework which provides the empirical scope for this thesis. It is derived from elaborate literature review and findings from related studies. The conceptual framework is abstract in nature by virtue of the fact that it presents a general overview of the building blocks of this research.

The concept holds that there are different inputs and modelling techniques that are applied in the FSM which affect the accuracy of model outputs. These factors are categorised into three groups namely; physical and socio economic, behavioural and finally model set up and run. Figure 2-4 below captures the conceptual framework; it is explained in the sections below.

The basic physical and socioeconomic factors are for example the transport network and the spatial distribution of land use and socio economic attributes in TAZs. These affect model accuracy depending on how they are used and computed for application in the FSM. As seen in figure 2.10 below, these factors also influence the observed traffic volumes on the road network.

Behavioural factors include such aspects as the trip making behaviour, the trip rates, travel patterns and general human behaviour that is captured in the FSM, the parameter values used in different steps of the model fall under this category. The literature shows that changes in behavioural factors affect model accuracy. As seen in figure 2-4 below, these factors also influence the observed traffic volumes on the road network.
Factors associated with model structure include the sequence in which the steps of the FSM are undertaken as well as the techniques for each of the four steps. Examples within this category include the choice between linear regressions or cross classification analysis for the trip generation step and the technique of applying feedback between trip distribution and assignment using model iteration. These factors do not have an effect on the observed traffic volumes on road networks.

In the data analysis, chapter 5 of this report, each of the above mentioned factors are tested separately by running the FSM six different times in order to establish their effect on model accuracy.

The figure below shows the conceptual framework

![Conceptual Framework](image)

**Figure 2-4: Conceptual Framework**

The conceptual framework was operationalised using the methodology presented in the data analysis section of this report. The following section introduces the study area.
3. Background to the Study Area

This section introduces the study area and it entails a discussion about the land use, transport and socio economic situation in Istanbul. Part of this section is based on information obtained from key informants during the field work while majority of the details presented are based on the Urban Mobility in Istanbul Report prepared by Gurcek and Demir (2008) for Plan Bleu.

3.1. Administrative Structure

The province of Istanbul is divided into 39 districts; some of the districts are totally urbanized while others distant from the city centre comprise both urban and rural settlements (villages). All districts have a district centre, a local municipality and an elected mayor. The city has a metropolitan administration and it is under the jurisdiction of Istanbul Metropolitan Municipality (IMM) which has three main organs: The Metropolitan mayor, the Metropolitan council and the Metropolitan executive committee. Local authorities within Istanbul comprise: Municipalities, special provincial administrations and village administrations (Gurcek and Demir, 2008).

3.2. Socio economic background

Istanbul is the biggest metropolitan city in Turkey with a population of 12.6 million in 2007. The city’s population has grown rapidly since 1950. The population growth rates have increased over the years, for the period 1935 – 1950 the rate of growth was 1.87% while for the period 2000 - 2007 it was 3.30 %. With a total area of 5343Km² the average population density in Istanbul has increased from 1,067 inhabitants per km² in 1985 to 2,333 inhabitants per km² in 2007. Figure 3-2 below captures the population growth trend for Istanbul.
Between 1990 and 2004, Istanbul produced 21% to 22.7% of Turkey’s total annual GDP. Total employment in Istanbul was estimated as 3,862,821 in 2006. Average employment density in 2006 was estimated at 717 employees per km2. However, zonal densities are as high as 33,797 employees per km2 in the European side of the city.

### 3.3. Land use in Istanbul

The trend in urbanization in Istanbul has been between the European and Anatolian sides and is mainly influenced by its historical spatial patterns, topography and geographic (physical) thresholds. In terms of topography, Istanbul has a unique urban form with hilly terrain divided by the strait of Bosporus linking with Marmara Sea. Istanbul has developed as a linear city; the maps figure 3-3 and 3-4 below depict the land cover for the years 1995 and 2005 while figure 3-5 illustrates the growth areas of human settlements between 1995 and 2005. The land cover maps are based on LANDSAT satellite images and classification undertaken by IMP. The maps indicate that majority of the developments have occurred around the core urban areas and more towards the west. In 1995 the land cover of human settlements had an area of 71,475 Ha, while in 2005 it covered an area of 85,239 Ha. However, further research should be undertaken to determine and quantify the growth areas.

---

**Figure 3-2: Population growth trend in Istanbul**  
(Source: Gercek and Demir, 2008)

---

**Figure 3-3: Land Cover Map-Istanbul 1995**  
(Source: IMP)

**Figure 3-4: Land Cover Map-Istanbul 2005**  
(Source: IMP)
According to sources at Istanbul Metropolitan Planning and Urban Design Centre (IMP), land use planning in Istanbul is carried out in different scales namely 1:25000, 1:50000, 1:100000 with feedback between these plans. The first land use plan in Istanbul was prepared in 1937 by Henry Prost, a French architect/planner. He proposed to decentralize the industry from the historic peninsula to outside of the old city walls with a new plan. Since then there have been a number of land use planning initiatives that have guided the development of the city. The first 1/50000 scale land use master plan of Istanbul Metropolitan Area was approved in 1980. By the implementation of that plan decentralization begun and mainly the manufacturing industry moved to outside the residential areas, the city macro-form became more linear in east-west direction than it was before.

IMP prepared a 1:50,000 scale zoning master plan in 1995 with one of the three main strategies of this plan being to achieve growth of the urban macro-form in a linear and multi-centred form, but with a degree of hierarchical ranking. This was replaced by The Istanbul Provincial Environmental Plan - scale 1:100,000 which was prepared in 2006 and revised in 2009 with the objectives to improve the quality of life, ensure sustainability of natural assets and improve the standard of living.

In summary as a result of rapid and extensive growth, the macro form of the city has changed from a single centred one to a multi-centred one with a number of sub-centres. Land use maps over time show a dense, polycentric urban structure extending in linear fashion, with a major drop in density in the outlying districts (Vallouis, 2008). The agricultural and natural areas have decreased as the lands allocated to housing, industry, commercial functions and transportation infrastructure have increased since 1940s. Land use and transport interaction in Istanbul is manifested in the fact that the first impact

Figure 3-5: Land Cover Change Map 1995-2005 (source: IMP)
of the Bosphorus Bridge was on the distribution of the population between two sides of the city. Nearly 80% of the population was living on the European Side of Istanbul in the year 1965. This ratio has decreased down to 76% in 5 years and to 73% in 10 years.

3.4. Transportation in Istanbul

Urban transportation in Istanbul is characterised by a road-based policy focusing on providing more road capacity to accommodate the rapidly increasing number of motor vehicles in Istanbul. It has been noted that additional capacity provided by these road investments facilitated a rapid growth in car use and created “induced traffic” as a result of the changes in the land-use and activity patterns. This is manifested by the fact that 90% of the private and public passenger trips are made by road vehicles in Istanbul (Gercek and Demir, 2008).

Constraints to the transportation system in Istanbul can be attributed to the separated geographical condition between the Asian and European side as well as the insufficient road network capacity and narrow access roads. The urban road network and village roads in Istanbul are not sufficient and this leads to traffic problems such as congestion.

The European and Asian sides of Istanbul are connected by two highway bridges that cross over the Bosphorus Strait. The first Bosporus crossing was constructed in 1973 while the second bridge commenced operations in 1988 and proposals are at an advanced stage to construct a third bridge. Roads in Istanbul are classified into 3 categories: freeways, arterial roads and other roads.

The public transportation system in Istanbul has been unable to keep pace with the rapid growth and changing urban structure. However, it has been assessed that the public transport system is relatively good and well managed with a variety of road public transport modes including buses and taxis. The rail system has a total length of 137 km of which most provides a low level of service and further development is impeded by the hilly topographical nature of the city. Sea transport systems in Istanbul include: the fast ferry, seabus, intercity passenger ferry, intercity vehicle ferry and the Mavi Marmara passenger ferry.

3.5. Travel Demand and Mobility Trends

Trip rates in Istanbul have increased over time as shown in the table 3-1 below, it can be seen that the total trip rate has increased from 1.54 in 1996 to 1.74 in 2006, while the motorised trip rate has decreased from 1.00 in 1996 to 0.88 in 2006(Gercek and Demir, 2008). The authors add that the changes in motorised trip rates could be explained by the suppressed urban travel demand by motor vehicles due to increasing traffic congestion and travel time on the urban road network. The increase in total trip rate can be attributed to improved economic performance where the GDP growth rate rose from 8.0 percent in 1996 to 20.3 percent in 2006. According to the 2007-2023 Istanbul Transport Masterplan Study, the increase in the GDP translates to increased household incomes which in turn lead to increased vehicle ownership and hence increased motorised households. The trip rate is seen to increase with an increase in motorised households.
This assertion is supported by the fact that the number of automobiles registered in Istanbul has increased dramatically from 200,000 in 1980 to 1.7 million in 2007. It is recommended that further analyses considering other factors such as changes in travel distance and travel time, travel-money budgets of households, land-use pattern among others should be undertaken to better understand the changes in the urban mobility pattern.

Table: 3-1: Trip rates in Istanbul for the years 1996 and 2006

<table>
<thead>
<tr>
<th></th>
<th>1996</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Old perimeter</td>
<td>New Perimeter</td>
</tr>
<tr>
<td>Total trip rate</td>
<td>1.54</td>
<td>1.79</td>
</tr>
<tr>
<td>Motorised trip rate</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td>Walk trips(%)</td>
<td>35.0</td>
<td>50.8</td>
</tr>
</tbody>
</table>

(Source: Gercek and Demir, 2008).

In regard to trip distribution there is a closer and stronger social and economic relation between the pair of OD districts and movements are restricted by the Bosphorus crossing. The trend has been that the average travel distance has dropped whilst the average motorised travel time has increased by 20% in the last decade. The average travel distance for all purposes and all modes in Istanbul in 2006 was estimated at 7.2 km and the average distance of motorized (vehicle) and walk trips was estimated at 11.2 km and 3.6 km respectively. The average travel time for motorised trips was 41 minutes in 1996 while in 2006 it was 49 minutes. It has been observed that about 21 million trips are made daily in the Istanbul metropolitan area of which 49.3% are walk trips (Gercek and Demir, 2008).

3.6. Travel Demand Modelling and Transport Planning in Istanbul.

Istanbul has witnessed four different transportation master plans; these were for the years 1985, 1987, 1997 and most recently the 2007-2023. This study focuses on the 1997-2010 and the 2007-2023 plans. The objectives of the 1997-2010 Istanbul Transport Master Plan were: towards a more integrated transportation system and a sustainable city for social, economic and urban development as a whole. This plan was prepared by the Istanbul Technical University (ITU) in conjunction with the Istanbul Metropolitan Municipality (IMM), (Alpkokin and Hayashi, 2003). The major investment suggested, was the tube tunnel railway system through the Bosporus with major rail improvements within both sides of the city.

The methodology of the 1997 transportation plan was the conventional four-step transport model. Two software packages, EMME/2 and TRANPLAN were used for calibrating the model and forecasting the future. A household survey with the sampling percent of 0.42% was conducted and necessary data of existing demographics, economics, and transportation system was collected. As the next step, a four step conventional transportation model consisting of 6,423 transport links and 250 zones, was calibrated. Finally future proposals of land use decisions and transport infrastructure projects were tested by the model (Alpkokin and Hayashi, 2003).

In 2007 “The Study on Integrated Urban Transportation Master plan for Istanbul Metropolitan area” was launched and carried out by IMM and JICA. It made use of the TRANSCAD software to calibrate
the four step transport demand model consisting of 451 zones. The table below presents a comparison of the 1997 and 2007 Transport master plans, it is observed that majority of the variables have increased since 1997 and only the statistics for none home based trips went down.

Table 3-2: Comparison of the 1997-2010 and 2007-2023 Transport Master Plans for Istanbul

<table>
<thead>
<tr>
<th>Working</th>
<th>1997</th>
<th>2007</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study field (Ha)</td>
<td>154,733</td>
<td>539,000</td>
<td>71%</td>
</tr>
<tr>
<td>Population in study field</td>
<td>9,057,747</td>
<td>12,007,000</td>
<td>25%</td>
</tr>
<tr>
<td>Employment</td>
<td>2,532,211</td>
<td>3,957,336</td>
<td>36%</td>
</tr>
<tr>
<td>No. of registered private car</td>
<td>889,342</td>
<td>1,522,521</td>
<td>42%</td>
</tr>
<tr>
<td>Average trip distance (minutes by motorised vehicle)</td>
<td>41</td>
<td>45.8</td>
<td>10%</td>
</tr>
<tr>
<td>Home based work (HBW)</td>
<td>43</td>
<td>48.26</td>
<td>11%</td>
</tr>
<tr>
<td>Home based school (HBS)</td>
<td>37.4</td>
<td>41.52</td>
<td>10%</td>
</tr>
<tr>
<td>Home based other (HBO)</td>
<td>42</td>
<td>44.64</td>
<td>6%</td>
</tr>
<tr>
<td>Non-home based (NHB)</td>
<td>34</td>
<td>40.73</td>
<td>17%</td>
</tr>
<tr>
<td>Number of motorised trips by purpose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home based work (HBW)</td>
<td>4,981,761</td>
<td>5,623,964</td>
<td>11%</td>
</tr>
<tr>
<td>Home based school (HBS)</td>
<td>1,313,373</td>
<td>1,550,353</td>
<td>15%</td>
</tr>
<tr>
<td>Home based other (HBO)</td>
<td>1,657,568</td>
<td>2,579,805</td>
<td>36%</td>
</tr>
<tr>
<td>Non-home based (NHB)</td>
<td>1,105,045</td>
<td>848,137</td>
<td>-30%</td>
</tr>
<tr>
<td>Total No. of trips by motorised vehicles</td>
<td>9,057,747</td>
<td>10,602,258</td>
<td>15%</td>
</tr>
<tr>
<td>Modal split %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private Transport</td>
<td>40</td>
<td>29</td>
<td>-38%</td>
</tr>
<tr>
<td>Public Transport</td>
<td>60</td>
<td>71</td>
<td>15%</td>
</tr>
</tbody>
</table>

(Source: Gercek and Demir, 2008)

3.7. Summary

Istanbul is a dynamic city that has undergone changes in land use and transportation between 1997 and 2007; the city has witnessed growth in this period driven largely by socio economic, political, physical and environmental factors. There has been debate about sustainable development issues related to mobility, urban development and transport in Istanbul. The need to involve new tools at decision-making level particularly by taking external costs into account is recommended (Vallouis, 2008).
4. **Research Methodology**

This chapter discusses the field work exercise in Istanbul. It highlights the activities undertaken as well as the data and information collected. The procedures for data preparation are also elaborated.

4.1. **Data needs**

In order to operationalize the conceptual framework and address the objectives of this research various data needed to be collected. The data required for this study was identified after extensive literature review and consultation with the research supervisors, due consideration was observed for feasibility given the time and technical constraints. The data needs for each step of the FSM and the data wish list used to guide the field work exercise are presented as appendix 3 and 4.

Land use and socio economic data needs included the traffic analysis zones (TAZs) used in the 1997-2010 and 2007-2023 Istanbul transport models, these would provide one of the main inputs for the model. Some of the required attributes were population, students, employment, income, car ownership among others.

The transport system data sought included the multi-modal transport network comprising road, rail and sea. Differentiation between highway and public transport was required. The networks would have the prerequisite attributes needed to run the FSM including mainly the travel speed and capacity.

Model calibration and validation data required included the parameters, coefficients, formulae and assumptions to run the FSM model. Due to the time constraints, this data was sourced from calibration reports as opposed to the standard practice of primary data collection through travel and household surveys. Validation data required for the study comprised the traffic counts captured by Remote Traffic Microwave Sensors (RTMS) which showed the actual traffic volumes on the roads.

The identification of data needs provided the basis for planning the data collection exercise. A justification for the use of a case study approach and the rational for the selection of Istanbul as the case study area is provided in the following section.

4.2. **Justification and Rational for Istanbul Case Study**

A case study approach was selected because this would give the research an element of novelty and realism. It was the intention that the research should be based on real life situations and actual data rather than hypothetical situations. By this the research would have scientific rigour and the outputs could be more reasonable and applicable to other context. This would also enable future work to be undertaken to replicate, compare or validate this work.

It is true that this research can be undertaken in almost any location in the world but Istanbul was selected as the case study location based on a number of factors. Firstly, there have been transportation master plan studies conducted in 1985, 1987, 1997 and the most recent being the 2007-2023 transport master plan. This provided the research with ample historical and current data to meet the objectives. Secondly, Istanbul is a highly populated dynamic and vibrant city that poses numerous challenges for
transport planning, its location astride two continents, Europe and Asia, as well the geographic phenomena of the Bosporus strait provide an interesting situation for travel demand modelling. Lastly, ITC and the IMP entered into an MOU where it was agreed that research work would be supported for the city of Istanbul.

Upon the selection of the case study area, the next step was to undertake the field work. The field work and data collection approach is presented in the following section.

4.3. Field Work Phase

A 21 day fieldwork exercise was carried out in Istanbul for the purpose of data collection. The research team was based at the Istanbul Metropolitan Planning and Urban Design Centre (IMP) offices.

The field work was justified by the fact that the required data could not be acquired by any other available means. Given the time constraints and nature of the study, data was collected from mainly secondary sources. Together with this, key informant interviews were also conducted with selected experts in the field of transportation and land use planning. Field observation was also carried out with the purpose of familiarisation with the dynamics of the study area and to confirm some of the secondary sources of data. The data collection approach and methodology are described in the section below.

4.3.1. Field work station- Istanbul Metropolitan Authority (IMP)

IMP is a consultancy body initiated by the Mayor of Istanbul in December 2004. The idea was to have an organizational form that cuts bureaucratic lines for solving problems and capitalising on opportunities to get the results. It is structured on the basis of problem solving and innovation in changing environments. IMP provides technical support and consultancy for the metropolitan municipality and other public agencies in the fields of advanced technology applications, cultural heritage management, urban design and regeneration, urban and regional planning among others.

Its core functions include research, planning, project design, survey and coordination. Other functions include generation and dissemination of knowledge by extensive research reports, GIS based maps, innovative 3D surveys, and master plans for different scales, plan reports and executive summaries, brochures, bulletins, almanacs, presentations, workshops, competitions, exhibition and organization of meetings.

The core mission of IMP is to support the preparation of supra scale plans in accordance with the strategic planning approach projected to the year 2023 for the metropolitan municipality area of jurisdiction which covers the Province of Istanbul. It organizes all kinds of technical studies for this purpose simultaneously contributing to the project development process in different ways. For more about IMP please refer to the IMM website following the link provided here: (www.ibb.gov.tr/en-US/Pages/Haber.aspx?NewsID=153).

4.3.2. Memorandum of Understanding between ITC and IMP

A memorandum of understanding was reached between IMP and ITC. This provided the framework for cooperation in research between the two institutions. It spelt out the roles and responsibilities of
each party and the scope and nature of cooperation. The field work activities were undertaken with reference to this MOU. A copy of the MOU is attached as appendix 5 of this report.

4.3.3. Working environment

Six workstations complete with desk top computers, internet connection, intranet connection, printing facilities and a telephone were provided. An English speaking contact person from IMP was assigned to supervise and assist the research team, this proved very helpful for coordination of field work activities, familiarisation of working environment and overcoming language barriers. The contact also helped to organise meetings with key informants.

4.3.4. Secondary Data Collection

Secondary data was collected from various sources mainly within IMP. The transportation department provided majority of the data which included the land use and socio economic data, the transport network data, the transport study report and model calibration report for the 2007-2023 Istanbul Transportation Master plan as well as the 1997-2010 Istanbul Transportation Master plan. The 1997-2010 transport model data was provided from sources within Istanbul Technical University and from Planofis.

Data on existing traffic conditions was acquired from the traffic control centre (TCC) and it comprised traffic volumes on links monitored by centre.

The table 4-1 below provides details on the secondary data that was collected; it gives a description of the data, the attributes, spatial scale, temporal scale, format and source.

Table 4-1: Data acquired during field work in Istanbul

<table>
<thead>
<tr>
<th>Data description</th>
<th>Attributes</th>
<th>Spatial Scale</th>
<th>Temporal scale</th>
<th>Format</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport networks; Highway and Public Transport(roads, bus routes, rail and sea)</td>
<td>Length, hierarchy, capacity, speeds, traffic flow data.</td>
<td>metropolita n level</td>
<td>1997/07</td>
<td>shape file and emme/2</td>
<td>IMP, ITU</td>
</tr>
<tr>
<td>TAZ and Districts layer</td>
<td>Land use and socio economic data(population, employment, school students)</td>
<td>metropolita n level</td>
<td>1997/07</td>
<td>shape file</td>
<td>IMP, Plan ofis, ITU</td>
</tr>
</tbody>
</table>
Further to the secondary data collection, meetings were held with key informants in order to get information and knowledge regarding land use and transportation in Istanbul, details of the key informant meetings are provided in the following section.

### 4.3.5. Meetings with Key Informants

A series of open discussions were held with relevant informants in the field of land use and transportation planning in Istanbul. The informants were identified through literature review and recommendation by contact persons. The overall purpose of the interviews was to provide tacit knowledge that would aid in the successful accomplishment of the study. A summary of the discussions held with the identified experts is presented below.

The main contact person at IMP was instrumental in introducing the research team to the working environment. He made a presentation about IMP and the history and culture of planning in Istanbul. He also organised a session where the research team made brief presentations on their respective topics of research to the staff at IMP. This helped to identify the relevant contact persons within the various departments at IMP as well as identify data availability and source. From these sessions the research team was able to assess the feasibility of undertaking the research with suggestions for modifications made by the IMP officials. The research team was also introduced to the chief advisor and co-ordinator for IMM who guaranteed approval for the research undertaking.

The research team met with one of the informants who had been involved in the preparation of the land use plans for Istanbul. The land use planning process in Istanbul and its historical development was discussed as well as the main issues of concern and challenges. From this it was established that the land use planning department provided the inputs that were used by the transportation department in the preparation of the Transport Master plan. Literature and data on land use planning was provided as well as contacts to other key informants.
In regards to transportation planning the research team met with key informants from this field at IMP and ITU. Information was gathered relating to the nature, practice and details of transportation planning in Istanbul. The methodology and preparation of the 2007-2023 and 1997-2010 Istanbul Transportation Master Plans was discussed and crucial information on calibration, computation and technical aspects related to application of software packages used in travel demand modelling in Istanbul was obtained, it was established that the plans were prepared using the FSM. Data used in the preparation of these plans was also made available to the research team.

4.3.6. The Traffic Control Centre (TCC)

The research team made a visit to the TCC in Istanbul, the TCC makes use of traffic cameras, radar sensors and electronic violation detection Systems (EDS) to survey traffic in Istanbul and obtain data from which traffic information is disseminated to motorists by use of intelligent transportation applications. Real time traffic information is broadcast to radios and television channels by a call centre 7 days 24 hours continuously, it is also presented on the web and on mobile phones to public. For more on the TCC refer to the website provided: (http://tkm.ibb.gov.tr/en-EN/index.aspx).

Figure 4-1: Traffic sensor on the left and depiction of the captured data on the right. Sensors can measure vehicle speed, vehicle number, vehicle classification, intensity of lanes (upto 8 lanes) and queue length.

Figure 4-2: Displays how RTMS micro-waves are beamed on the road network.

The research team met key informants at the TCC who provided a short description of how the TCC functions and described the system of sensors used. It was revealed how these sensors record and relay data into the TCC database for dissemination to the public via the TCC website in the form of traffic intensity maps which depict the congestion levels and travel speeds on monitored links as shown in figure 4-1 and 4-2 above. They also provided insight into the traffic problems and transport related issues facing Istanbul. Finally they provided the RTMS data which was used in this study for validation of the FSM outputs.
4.4. Field Observation

Field observation was undertaken in order to familiarise research team with the study area and to counter check information from the secondary sources and key informant interviews on transportation issues such as congestion. The aspects observed were the traffic conditions on links identified as having high congestion levels, the use of the public transport system namely rail, sea and bus to check the scheduling, frequency, speed among other observations. The rational of the field observation was that the information would be used in the modelling process and could explain inaccuracy.

Information on traffic flow from the TCC is published in the internet and it displays a traffic intensity map showing the traffic conditions in terms travel speed and congestion based on data captured by the RTMS. The information displayed on the TCC website was cross checked for accuracy during the field observation to see if it was congruent to what was on the ground. An example of this is illustrated figure 4-3 below, it was established that the traffic intensity maps gave an accurate account of the observed traffic conditions meaning that on this basis, and not considering other factors, the data from TCC sensors can be relied for validation of the FSM outputs and accuracy assessment.

![Traffic Intensity Map](image1)

![Video Capture](image2)

Figure 4-3: On the upper right is the traffic intensity map depicting congestion levels on roads based on travel speed. The displayed travel time and graph is for the link shown by the black arrow and indicates an average travel time of 10km/h between 18:04 and 19:00 hrs. On the lower left is a video capture of the traffic flow on the same link shown in the intensity map, taken at 18:51 hrs

In summary, the field work exercise was successful in obtaining valuable information and data. Due to the nature and scope of this research, majority of the data acquired was secondary in nature. In hindsight the contacts made and key informant interviews proved to be useful in the data analysis stage of the research especially during execution of the FSM.

Some of the data collected required further processing in order to meet the specifications of the research. The main focus was on making the data ready for input into the FSM for analysis. The following section presents briefly some of the data preparation procedures undertaken.
4.5. Data Processing, Validation and Reasonableness Check

The data was assembled and put into a geodatabase using ESRI ArcGIS 9.3 software application. The pre-analysis processing required the change of format of some of the data into shapefile format in preparation for the analysis. For example the highway network for 1996 was in the TRANPLAN format and was changed into shapefile using CUBE transport planning software. Similarly the model results showing the projected volumes on links for the 1997-2010 transport models were in TRANPLAN format and were opened and deciphered using CUBE.

Furthermore, some of the data collected for example the attributes on the shapefiles were in Turkish and hence there was need to translate them into English for ease of reference. Translation was carried out during the field work and was undertaken by two research assistants hired for the study.

The data was validated and checked for reasonableness using different approaches. The key approach used was to cross check the land use and social economic data against other sources of data mainly census records where aggregated totals were compared (Wegmann and Everett, 2008 and Zhao et.al., 2005). The dataset contained reasonable figures for population, employment and students.

The 1997 TAZs (250 in number) lacked land use and socio economic attributes of population, employment and students. These attributes were contained in the districts layer (32 in number). Hence, there was need to assign attributes to the TAZs before undertaking the FSM. The approach followed is related to the work by McCray, et.al (2008) and its implementation, which was done in ArcGIS 9.3 is shown in the section below:

**Data processing 1: Adding socio economic and land use attributes into the TAZs shape files for 1997 derived from the districts layer.**

**Step1: Computing the population, students and employment density in the districts layer using the equation 4.1,4.2 and 4.3 below:**

The input for this data was the districts shapefile obtained during field work, it contained attributes for population, employment and students aggregated into 32 districts polygons covering the entire Istanbul Metropolitan area.

\[
\text{Population Density} = \frac{\text{Population}}{\text{District Shape area}} \quad (4.1)
\]

\[
\text{Employment Density} = \frac{\text{Employment}}{\text{District Shape area}} \quad (4.2)
\]

\[
\text{Students Density} = \frac{\text{Students}}{\text{Districts Shape area}} \quad (4.3)
\]

The figures for population, employment and students are the absolute values while the shape area is in km².
**Towards Integrated Land Use and Transport Modelling: Evaluating Accuracy of the Four Step Transport Model - The Case of Istanbul, Turkey**

Step 2: Overlay the TAZs layer with the districts layer using the intersect option in ArcGIS. This increases the number of rows in the new overlay layer.

The input for this step was the districts layer with necessary attributes from the step above and the TAZ polygon layer.

Step 4: Recalculate the population, students and employment in new fields using the formula 4.4, 4.5 and 4.6 below:

\[
\text{Absolute Population} = \text{Population Density} \times \text{TAZ Shape area} \quad (4.4)
\]

\[
\text{Absolute Employment} = \text{Employment Density} \times \text{TAZ Shape area} \quad (4.5)
\]

\[
\text{Absolute Students} = \text{Students Density} \times \text{TAZ Shape area} \quad (4.6)
\]

Step 5: Summarise the attribute table using the unique ID of the TAZ layer and sum the population, students and employment attributes. The summary reduces the records to the number of TAZs.

Step 6: The summary table from the step above is then joined with the original TAZ layer and thus creating the attributes in the TAZs.

It is important to note that a major weakness of this approach is the assumption that population, employment and students are uniformly distributed throughout the district. Hence these attributes are recalculated in step 4 above using the density as a factor of redistribution within the TAZs that intersect with the district layer. The variation in the values will depend on shape area.

However, the results of this technique were checked and found to be reasonable. When visualised in ArcGIS together with the land cover maps, it is seen that the TAZs with highest numbers of population, employment and students also had the greatest concentration of human settlements land cover. It is recommended that a better approach would be to consider the magnitude of land use per TAZ as criteria for calculating population, employment and students.

Data processing 2: Aggregating the 451 TAZs of the year 2006 into 250 TAZs as used in 1996

The TAZs for the year 2006 contain the land use and socio economic data used in this thesis as the actual / observed data. They comprised 451 TAZs, while the 1996 TAZs were 250 in number. In order to achieve uniform analysis areas to be used in this thesis, the 2006 TAZs were aggregated into 250 with necessary attributes. The steps followed to achieve this are similar to those used in data processing 1 shown previously.
5. Data Analysis

This section presents the methodology used in data analysis in order to answer the research questions. The overriding idea was to compare FSM traffic volumes against the reality (observed traffic volumes) using statistical methods. Research question 1-2 assess the validity of input data with analysis between actual and modelled data. Research questions 3-8 mainly focus on the FSM and provide an accuracy assessment of the assigned traffic volume based on different inputs. Research question 9 measures the effect size of using different inputs and model specifications. Figure 5-6 at the end of this section provides a summary of the main scope of analysis that guided this study. The FSM is implemented focusing on the private car and assignment of traffic to the highway network on 24hr basis. The results and discussion ensuing from the analysis phase are presented in chapter 6.

5.1. Research question 1: Comparing Forecast and Actual TAZ Data

Rational and Justification
Data analysis in this part involved a comparison between the actual and forecasted land-use and socio economic data in the TAZs as used in the FSM. The rational applied was to determine the difference between forecasted and actual data in order to establish how these differences affect the FSM accuracy. A similar approach is used in the work by Wegmann and Everett (2008) and Zhao et.al., (2005).

Inputs
The actual data is based on the TAZs for the 2007-2023 Istanbul Transport Master Plan study conducted by IMM and JICA in which surveys were carried out within Istanbul to establish base line information for the study (JICA and IMM, 2007). The forecast data is based on the TAZs used in 1997-2010 Istanbul Transport Master Plan.

Methodology, Tools and Techniques
The TAZ data from the 1997-2010 Transport Master Plan was forecasted to 2006 to allow for comparison with actual data. In preparation for this analysis, the TAZs for the year 2006 were aggregated to match the year 1996 TAZs to allow for uniform spatial areas of analysis. This was done by using overlay functions in ArcGIS 9.3 as indicated in the previous section dealing with data processing 2. The attributes compared were the population, employment and students in terms of absolute numbers.

The statistical comparison entailed computing the absolute error (AE), percentage error (PE), mean absolute error (MAE), mean percentage error (MPE), root mean square error (RMSE) and percent root mean square error (%RMSE) between forecast and actual data using equations 5.1-5.6 shown below. It should be noted that for this analysis, the outlier values were excluded from the comparative analysis. For an example of previous use of this approach the reader is referred to Zhao et.al., (2005); McCray et.al., (2008) and Wegmann and Everett (2008). The statistical procedures were carried out using Microsoft Excel, below are the equations used:

\[ AE = \text{forecasts} - \text{actual data} \]  

(5.1)
Towards Integrated Land Use and Transport Modelling: Evaluating Accuracy of the Four Step Transport Model - The Case of Istanbul, Turkey

\[ PE = \frac{\text{Forecasts} - \text{Actual data}}{\text{Actual data}} \times 100 \]  

(5.2)

\[ MAE = \frac{\sum (\text{Forecasts} - \text{Actual data})}{n} \]  

(5.3)

\[ MPE = \frac{\sum (\text{Forecasts} - \text{Actual data}) \times 100}{n} \]  

(5.4)

\[ \text{RMSE} = \frac{\sum (\text{Model}_j - \text{Count}_j)^2}{(\text{Number of Counts} - 1)} \]  

(5.5)

Where:

\( \text{Model}_j \) - is the Forecast values for TAZ \( j \)

\( \text{Count}_j \) – is the Actual values for TAZ \( j \)

\( \text{Number of Counts} \) – is the number of TAZs

\[ \% \text{RMSE} = \frac{\text{RMSE}}{\sum_{i=1}^{n} X_i \, / \, n} \]  

(5.6)

Where:

\( \text{RMSE} \) - is the RMSE from equation (5.5) step above.

\( X_i \) - is the actual values for TAZ \( i \) (Hence, all the actual values are summed up).

\( n \) – is the number of TAZs.

The flow chart below summarises the methodology used to address research question 1.

![Flow chart](image)

Figure 5-1: Methodology for research question 1
5.2. Research Question 2: Comparing the 1996 and 2006 Highway networks

Rational and Justification
Data analysis for research question 2 involved comparing the planned highway networks based on the 1997-2010 Istanbul Transport Master Plan with the transport network for the year 2006. The rational was to determine whether the transport networks were developed as planned and to establish the effects of applying planned and actual networks. A similar approach is used in the work by Zhao et.al., (2005).

Inputs
The inputs for this part of the analysis included the planned highway network for the year 1996; the second input used was the actual highway network in 2006.

Methodology, Tools and Techniques
Using ArcGIS 9.3, both highway road networks were compared in terms of the total lengths and total number of links by summarising the attribute data. A visual inspection of the highway networks was also undertaken to establish differences. Finally road density maps for both highway networks were prepared in ArcGIS in order to make comparisons. The flow chart below captures the methodology used in addressing research question 2.

![Flowchart for research question 2]

Figure 5-2: Methodology for research question 2

5.3. Research Question 3-9: The FSM

This section is a prelude to the analysis of the remaining research questions which were operationalized using Flowmap7.3 professional software. It presents the steps followed to build and execute the FSM in Flowmap software.

About Flowmap
Flowmap is a program for geographical analyses. It is specialized in displaying interaction data, like commuting and migration flows, interaction analysis like accessibility analysis, network analysis, and interaction modelling (De Jong et.al., 2009).
Methodology, Tools and Techniques
The FSM was executed using a combination of Flow Map 7.3 (Professional version), ArcGIS 9.3, OmniTRANS version 5.1, SPSS and MS Excel. The procedures and calculations applied are outlined in the steps below.

Step 1: Trip Generation
Trip generation using the 1996 TAZs was computed based on the 1997-2010 Istanbul Transport Master Plan data and model specifications. Trip purposes were classified into Home Based Work trips (HBW), Home Based School trips (HBS), Home Based Others (HBO), and Other trips (OT). The productions and attractions for 1996 were computed using the cross-classification method based on the employment rates and trip purpose.

Table 5-1 below shows the employment and unemployment rates for Istanbul in 1996 which were based on surveys conducted for the 1997-2010 Istanbul Transportation Master Plan. These were used to quantify the number of employed and unemployed population in the TAZs.

Table 5-1: Employment and un-employment rates by income group –Istanbul, 2006.

<table>
<thead>
<tr>
<th>Income Group</th>
<th>Sample population (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employed</td>
<td>Unemployed</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>8.6</td>
<td>13.43</td>
<td>22.04</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>27.98</td>
<td>40.10</td>
<td>68.08</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>4.32</td>
<td>5.57</td>
<td>9.88</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>40.90</td>
<td>59.10</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

(Source: 1997-2010 Istanbul Transportation Master Plan)

The number of employed and unemployed population was derived by multiplying the rates with the population in the TAZs.

Upon computing the level of employed and unemployed in the TAZs the next step was to compute the trip productions. This was based on the trip production rates by purpose as shown in table 5-2 below.

Table 5-2: Trip production rates classified by trip purpose and employment status

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>Employed trip production rate</th>
<th>Unemployed trip production Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home-Work</td>
<td>1.26</td>
<td>0.06</td>
</tr>
<tr>
<td>Home-School</td>
<td>0</td>
<td>0.27</td>
</tr>
<tr>
<td>Home-Other</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Others</td>
<td>0.28</td>
<td>0.11</td>
</tr>
</tbody>
</table>

(Source 1997-2010 Istanbul Transportation Master Plan)

The equations 5.7-5.10 were used to compute productions per TAZ as shown below:
The trip attractions were computed based on the trip attraction rates classified in terms of trip purpose as shown in the table 5-3 below:

**Table 5-3: Trip attraction rates classified by trip purpose**

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>Rate</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home-Work</td>
<td>2.07</td>
<td>Employment</td>
</tr>
<tr>
<td>Home-School</td>
<td>0.92</td>
<td>Students</td>
</tr>
<tr>
<td>Home-Other</td>
<td>0.67</td>
<td>Employment</td>
</tr>
<tr>
<td>Others</td>
<td>0.43</td>
<td>Employment</td>
</tr>
</tbody>
</table>

(Source: 1997-2010 Istanbul Transportation Master Plan)

Equations 5.11-5.14 were used to compute trip attractions per TAZ as shown below:

\[
\text{Home–work } \text{Attraction} = 2.07 \times \text{No.employed population} + 0.06 \times \text{No. unemployed population} \tag{5.7}
\]

\[
\text{Home–school } \text{Attraction} = 0 \times \text{No.employed population} + 0.27 \times \text{No. unemployed population} \tag{5.8}
\]

\[
\text{Home–other } \text{Attraction} = 0.22 \times \text{No.employed population} + 0.22 \times \text{No. unemployed population} \tag{5.9}
\]

\[
\text{Others } \text{Attraction} = 0.28 \times \text{No.employed population} + 0.11 \times \text{No. unemployed population} \tag{5.10}
\]

Trip generation using the 2006 TAZs was computed based on the 2007-2023 Istanbul Transport Master Plan data and model specifications. Trip purposes were classified into Home Based Work trips (HBW), Home Based School trips (HBS), Home Based Other (HBO) and None Home Based (NHB). Trip production and attraction models for HBW and HBS were forecasted using trip rates while HBO and NHB models were developed using linear regression analysis. Table 5-4 shows the variable list used for production and attraction.
Table 5-4: Variables used for trip production and attraction in the 2007-2023 Istanbul Transport Model

<table>
<thead>
<tr>
<th>Trip purpose</th>
<th>Production</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>Workers in home</td>
<td>Employment in working place</td>
</tr>
<tr>
<td>HBS</td>
<td>Students in home</td>
<td>Students in school place</td>
</tr>
<tr>
<td>HBO</td>
<td>Population, Average income, Workers in home</td>
<td>Population, Student in school, Employment in working place</td>
</tr>
<tr>
<td>NHB</td>
<td>Students in school, Employment in working place</td>
<td>Population, Student in school, Employment in working place</td>
</tr>
</tbody>
</table>

(Source: IMM and JICA-2007-2023 Istanbul Transport Master Plan Study)

The specifications for the model are presented in equation 5.15-5.23 below:

\[ HBW \text{ production} = \text{Net trip rate}(1.94) \times \text{Working ratio}(0.88) \times \text{Number of workers(Home place) by zone} \] (5.15)

\[ HBW \text{ Attraction} = \text{Net trip rate}(1.94) \times \text{Working ratio}(0.88) \times \text{Number of employment(Working place) by zone} \] (5.16)

\[ HBS \text{ production} = \text{Net trip rate}(2.02) \times \text{Studying ratio}(0.87) \times \text{Number of students(Home place) by zone} \] (5.17)

\[ HBS \text{ Attraction} = \text{Net trip rate}(2.02) \times \text{Studying ratio}(0.87) \times \text{Number of students(School place) by zone} \] (5.18)

\[ HBO \text{ production} = -748.943 + (0.425737 \times \text{Population}) + (1.163230 \times \text{average income}) + (0.558952 \times \text{workers}) \] (5.20)

\[ HBO \text{ Attraction} = 1.063040 + (0.44342 \times \text{Population}) + (0.290994 \times \text{Students at school}) + (0.257524 \times \text{Employment}) \] (5.21)

\[ NHB \text{ production} = 511.73 + (0.089921 \times \text{Students at school}) + (0.296634 \times \text{Employment}) \] (5.22)

\[ NHB \text{ Attraction} = 464.118 + (0.015623 \times \text{Population}) + (0.03754 \times \text{Students at school}) + (0.286727 \times \text{Employment}) \] (5.23)
The following section presents the methodology and approach used for the trip distribution step of the FSM.

**Step 2: Trip Distribution**

Trip distribution was carried out using Flow Map 7.0 software. The shape file of the TAZs with the calculated productions and attractions was converted into Flowmap format with productions and attractions being the main attributes of focus. The highway network was also converted to Flowmap format.

In this step, the task was to match the origins and destinations to develop an O-D trip table which is a matrix that displays the number of trips going from each origin to each destination. The doubly constrained gravity model was used for trip distribution where the productions and the attractions were used to fit the distribution function. Flow map allows introduction of different types of decay functions namely, the exponential, power and the Tanner function. The mean trip length (MTL) and beta values were also used to produce a trip proportional fit.

In Flow Map a distance table from each zone to each other zone using the network has to be created to before execution the gravity model. The network distance table was created using the set up shown in appendix 6.

The origins and destinations in Flow Map are depicted as the centroids of zones. The intrazonal distance, which is the average distance to go from the centroid to the network, was calculated on the basis of equation 5.24 shown below:

\[ C_i = 0.667 \sqrt{\left( \frac{S_i}{\pi} \right)} \]  

\[ (5.24) \]

Where \( C_i \) the intrazonal distance of the area \((i)\) and \( S_i \) the surface (SIZE) of zone \((i)\)

In Flow Map this was calculated using equation 5.25 below:

**Intrazonal Distance** = \( 0.667 \times \left( \frac{\text{SIZE}}{\#p_1} \right)^{0.5} \)  

\[ (5.25) \]

Having developed the network distance table the next step was to actually perform the doubly constrained gravity model. The origin constraint applied was the production field from the TAZs and the destination constraint was the attraction field. The distance decay function was set to exponential with the assumption that the propensity to travel decreases proportionately with distance and this decrease is quick. The choice of function was based on the 1997-2010 Transport Master plan. The beta values and the MTL used were based on the estimated values during calibration. The convergence criterion for all cases was set to 1.0%. The formula for trip distribution is shown earlier in section 2.6.2.

The following section presents the methodology used for modal split. This was undertaken using the flows generated from the gravity model output.

**Step 3: Modal Split**

Modal split was carried out in Windows MS Excel, the Flowmap flow-file in data base format (DBF) from the trip distribution step was used to share the person trips into different modes. According to the
1997-2010 Istanbul Transport Master Plan, the share of private transport was 40% and that of public transport was 60%. While in the year 2007, the share of private transport was 29% while that of public transport was 71%. The average occupancy rates for the different vehicle types and passenger car units according to the plan are shown in the Table 5-5 below, these rates were used to convert the person trips to car trips using equations 5.26 and 5.27 shown below:

\[
\text{Vehicle trips}(1996) = \frac{(\text{Score} \times (40/100))}{1.9} \\
\text{Vehicle trips}(2006) = \frac{(\text{Score} \times (29/100))}{1.57}
\]

Where: Score is the person trips in the Flowmap DBF flow file

<table>
<thead>
<tr>
<th>Year</th>
<th>Average Occupancy</th>
<th>PCU</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1.90</td>
<td>1.0</td>
</tr>
<tr>
<td>2006</td>
<td>1.57</td>
<td>1.0</td>
</tr>
</tbody>
</table>

(Source: 2007-2023 Istanbul Transportation Master Plan Study)

The following section presents the methodology applied for the traffic assignment step of the FSM. It relies on the outputs derived from the modal split step.

**Step 4: Assignment – Flowmap to OmniTRANS**

Upon completing the modal split, the following step was to assign the vehicle trips to the transport network. The private car trips were assigned to the highway network using OmniTRANS version 5.1 which is a transport modelling software application.

The shift from Flowmap to OmniTRANS was done because Flowmap uses the All or Nothing assignment technique which is weak as it does not consider the effect of congestion and thus does not represent the reality of travel behaviour appropriately (Bates, 2000; Wells, 1975). The assignment method used was the Volume Averaging assignment which is an equilibrium assignment method that takes into account congestion effect and a random component of the route costs.

Assumptions of this modelling technique are:

- Each trip-maker chooses his/her route through the network which minimizes his/her individual travel time
- Equilibrium is achieved when every trip-maker is using the best route possible, given prevailing congestion levels
- No user can switch routes and improve his/her travel time, and so no user will switch voluntarily
- For each origin-destination pair of zones, all used routes have equal perceived travel times, and no unused route has a lower perceived travel time.

Inputs required to perform an equilibrium assignment in OmniTRANS include the O-D matrix with traffic flows and the highway network comprising links, nodes and TAZ centroid. The O-D matrix
from Flowmap with traffic flows was converted to matrix form in Ms Excel using pivot table function and imported into OmniTRANS.

The highway network was imported into OmniTRANS via a conversion function that converts ArcGIS shape files into OmniTRANS network files with the main attributes being speed, capacity and cost attributes of distance and travel time. The link speed used in this analysis was derived from speed limits in Istanbul because the network acquired during fieldwork lacked this attribute. Based on the speed limit regulations in Istanbul, it was assumed that links close to areas with urban development had a speed of 50Km/h and those outside urban development areas had a speed of 100 Km/h. This was a compromise that should be considered as a limitation of the study.

The assignment parameters were specified within the job function in OmniTRANS using Ruby programming language; this included defining the congestion effect function, the number of iterations and the epsilon. The Bureau of Public Roads (BPR) congestion effect function developed by the American Bureau of Public Roads was used in this analysis. This function defines the relationship between travel time, volume and capacity as shown in the equation 5.28 below.

\[ T = T_0 \left(1 + \alpha \frac{V}{Q}\right)^\beta \]  
(5.28)

Where:  
\( T \) - travel time  
\( V \) - volume  
\( Q \) - capacity  
\( T_0 \) - free-flow travel time  
\( \alpha \) and \( \beta \) - alpha and beta function coefficients which can be set in the job

As explained in the OmniTRANS manual, if \( \alpha = 0.5 \), delays will occur if the link volume is approaching full capacity (main highways). If \( \alpha = 2.0 \), significant delays will occur well before full capacity is reached (residential roads) as shown in the figure 5-3 below, from previous experience, the parameter \( \beta \) is usually set to 4.0.
An example of the job specification for the assignment using Ruby is attached as appendix 7 to this report. The next section presents the set up of the FSM to address specific research questions of this research.

5.4. Variation of Inputs in the FSM Based on Research Questions

In order to address specific research questions, the set up and inputs of the FSM were varied for each model run. There were six research questions that required application of the FSM and for each the model was run and the results recorded separately. The combination of variables used is depicted in figure 5-4 below that depicts the general research methodology.

The variable under analysis for research question 3 was the projected TAZ data, for research question 4 the variable under analysis was actual TAZ data, research question 5 measured the effect of the road network, and research question 6 addressed the use of different parameters while research question 7 measured the effect of disaggregating the TAZs.

Research question 8 measured the effect of using a feedback mechanism between traffic assignment and trip distribution on FSM accuracy. The variables used to analyse feedback were actual data for 2006 in 250 TAZs, the 1996 planned highway network, year 1996 model parameters and a feedback mechanism based on travel time. In order to achieve the feedback mechanism, the FSM was run for a number of two iterations. The initial iteration was conducted as shown above and made use of the free flow link travel time as the impedance unit for the network distance matrix. The second iteration adjusted the free flow link travel times to factor in the congestion effect derived from the assignment step in the first iteration. The BPR travel time function, equation (5.28) in the previous section, was used to estimate the effect of congestion on each link. The network distance matrix was created using the new travel times computed.

The following section captures the second phase of the analysis which focuses on measuring the accuracy of the FSM outputs (traffic volumes) by comparing them to the actual traffic volumes captured by the RTMS for specific road links.
Figure 5-4: Research Methodology

KEY:

- The input boxes marked with a darker outline indicate the new variable to be analysed
5.5. Comparing the FSM traffic assignment volumes with the RTMS traffic volumes (ADT)

In order to evaluate the FSM accuracy, the model traffic volumes were compared to the actual traffic volumes using statistical measures. The data from the RTMSs containing actual traffic volumes was in raw form and needed to be prepared before analysis could take place. The methodology to achieve this is shown below.

**Step 1: Identifying the RTMSs and Specifying the Spatial Reference**

The RTMS point data were delivered in KMZ format and visualised using Google earth and ESRI ArcGIS 9.3. The point data did not have a spatial reference and thus could not be directly related spatially with the highway network. Hence, a new point feature shapefile was created in ArcGIS and the X and Y coordinate values from the KMZ file were used to capture the RTMS points. Attributes of RTMS_ID were added and a visual check done to confirm accuracy.

**Step 2: Computing the Average Daily Traffic (ADT) from the RTMS Data**

The RTMS data in MS excel files consisted of a number of attributes listed in the table 5-5 below:

<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>No. of Columns*</th>
<th>No. of Rows(Approx)</th>
<th>No. days</th>
<th>Temporal Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTMS_ID</td>
<td>1</td>
<td>900,000</td>
<td>90</td>
<td>5 minutes</td>
</tr>
<tr>
<td>VL-long vehicle count</td>
<td>8</td>
<td>900,000</td>
<td>90</td>
<td>5 minutes</td>
</tr>
<tr>
<td>V-All vehicle count</td>
<td>8</td>
<td>900,000</td>
<td>90</td>
<td>5 minutes</td>
</tr>
<tr>
<td>O-Occupancy</td>
<td>8</td>
<td>900,000</td>
<td>90</td>
<td>5 minutes</td>
</tr>
<tr>
<td>S-Speed</td>
<td>8</td>
<td>900,000</td>
<td>90</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>

*No. columns is based on the fact that each RTMS records data for 8 lanes.

The ADT was computed from this data set using SPSS software based on equation 5.29 below:

\[
ADT = \frac{Volumes}{No. of Days}
\]

(5.29)

Where: Volumes is the number of small vehicles.

The volumes of small cars were computed by subtracting the long vehicle counts (VL) from the all vehicle count (V). This was based on the assumption that VL comprised public transport vehicles and freight vehicles such as trucks.

The RTMS data with the ADT was converted to DBF format and joined with the RTMS shape file using the join function in ArcGIS.
Step 3: Relating the RTMS point features to the Highway Network

The new RTMS point features were displayed in ArcGIS together with the highway network. A total of 31 RTMSs were used and each was associated with a single highway network link using select functions in ArcGIS and visual analysis. This was followed by keying in the RTMS_ID and ADT into the associated highway way network link.

Figure 5-5: Location of case study RTMS and links

Step 4: Statistical Analysis of the FSM traffic volumes and RTMS traffic volumes (ADT)

The FSM traffic volumes and RTMS traffic volumes were analysed using MS Excel. The outputs from this stage are presented as the final results indicating FSM accuracy. RMSE and %RMSE statistics were used as shown in the previous section, equation (5.5) and (5.6) with the only difference being the use of links instead of TAZs for the formula.

5.6. Research question 9- Effect Size

Research question 9 measured the effect size of the different FSM manipulations from the previous research question (3-8). In order to determine the effect size of the various FSM variables analysed in this research, one way repeated measures ANOVA was computed in SPSS to determine the F-ratio and effect size. This statistic compares the size of the variation due to the experimental manipulations with the size of the variation due to random factors (Field, 2005). The percent error on the 31 case study links based on FSM traffic assignment results under actual data input were used as the control against which the percent error of the other 5 model runs were measured. The rational for this was based on the fact that actual data had been used to set up the FSM for 5 out of the total 6 model runs in this study.
The $F$-ratio and effect size of FSM manipulation were derived from the within-subjects contrasts in this manner. SPSS reported the $F$-ratio, refer to Field, (2005), but the effect size had to be calculated.

In order to compute the effect size ($\omega^2$), the equation (5.30) provided by Field (2005) was used.

$$ r = \frac{F(1, df_R)}{\sqrt{F(1, df_R) + df_R}} $$  \hspace{1cm} (5.30)

Where:-

- $r$ is the effect size ($\omega^2$ omega squared)
- $F$ is the $F$-ratio
- $df_R$ is the degrees of freedom for the residuals
6. Results and Discussion

This chapter presents the results of the data analysis which are based on the application of the methodology presented in the previous chapter. Results are reported for each research question and are followed by a short discussion that interprets the outcomes. In the end there is a summary discussion that compares the results from the different analyses applied.

6.1. Research Question 1 - Projected vs actual land use and socioeconomic data for the year 2006

The data analysis in this step involved comparing the projected and the actual land use and socioeconomic variables for the year 2006. The statistics indicate that population and employment were overestimated while the student projections were under estimated. The percent error for employment is the greatest with a value of 20 % while that of students was -14% and population 6%. The averages for the TAZs give a %RMSE for population as 134%; for Employment at 135% and for students at 125%. Table 6-1 presents the results from the accuracy assessment.

Table 6-1: Accuracy assessment for forecast vs actual land use and socio economic inputs: 1995-2005

<table>
<thead>
<tr>
<th>Study area</th>
<th>Forecast</th>
<th>Actual</th>
<th>AE</th>
<th>PE %</th>
<th>MAE</th>
<th>MPE %</th>
<th>RMSE</th>
<th>%RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>12841086</td>
<td>11784774</td>
<td>1056312</td>
<td>9</td>
<td>4737</td>
<td>323</td>
<td>71039</td>
<td>134%</td>
</tr>
<tr>
<td>Employment</td>
<td>4634820</td>
<td>3877205</td>
<td>757615</td>
<td>20</td>
<td>3397</td>
<td>275</td>
<td>23516</td>
<td>135%</td>
</tr>
<tr>
<td>Students</td>
<td>2189254</td>
<td>2532613</td>
<td>-343359</td>
<td>-14</td>
<td>-1540</td>
<td>270</td>
<td>14250</td>
<td>125%</td>
</tr>
</tbody>
</table>

Discussion

The outcome of the error assessment suggests high inaccuracy between the forecast and actual figures for population, employment and students. The values for PE appear small because they are based on aggregate values, the %RMSE for all the variables is high at a value of +130% and this is a better measure because it takes into account the error in all TAZs. The values for the PE and MPE show a logical discrepancy because when calculating MPE the error for individual TAZs cancel out but in the PE they do not.

One of the reasons explaining the error between forecast and actual values could be due to the fact that the land use development did not take place as planned between 1995 and 2005. It should also be noted that errors in land use and socio-economic projections are unavoidable (Hanson and Giuliano, 2004). Several unforeseen events could have occurred between the 10 year forecasting period and this could explain the error reported because human beings cannot predict what will happen in the future precisely. These results tally with the findings in the work by Zhao et.al., (2005) and Parthasarathi and Levinson (2008).

When the land cover change map for human settlements between 1995 and 2005 is overlaid with the TAZ map visualising the forecasting error it is seen that the zones with high error are not necessarily also the zones with highest land cover growth. This assertion holds true for population, employment and students when based on visual analysis.
The preparation of the zonal land use and socio-economic data in ArcGIS could yet be another reason explaining the error reported. The overlay functions used as explained in section 4.6 could result in error because of loss of some of the polygons after intersecting the district layer and TAZ layer leading to loss of spatial information related to land use and socio-economic variables. Error in forecasting the land use and socio-economic input data leads to inaccuracy in the FSM outputs measured in terms of ADT, this assertion is tested in section 6.3.

6.2. Research Question 2- Planned vs actual highway network for the year 2006

The planned highway network in 1996 was compared with the 2006 highway network in order to establish whether the network was developed as planned. The reason for doing this was because the highway network would be used in the FSM and could be a source for model inaccuracy. The maps below, figure (6-2) and (6-3), present the highway networks for 1996 and 2006, a visual comparison indicates that they are different and the reason for this is explained in the discussion below.
On further analysis, the road density maps indicate general similarities in that the distribution of TAZ road density across the study area is occurs in the same pattern for both networks when visualised in ArcGIS, the high densities and low densities occur in the same locations as shown in figure (6-3) and (6-4). However, the statistics show differences with the planned highway network having a total of 5617 links amounting to 4601 Km of highway network, a minimum road density of 58Km² and a maximum road density of 9538Km² while the actual road network in the year 2006 had a total 15333 links amounting to 6711 Km of highway network, a minimum density of 16 km² and a maximum of 11002Km².

Figure 6-2: Highway network 2006

Figure 6-3: Road Density Map for planned highway network 1996-2010
Discussion
The results indicate that there is a difference between the planned and the actual highway network. The main reason explaining this is the manner in which the networks were modelled and prepared for use in the FSM. The planned highway network is greatly simplified and the geometry does not reflect the network on the ground, from key informant interviews it was determined that the simplification and generalisation of the network was done in order to make it easier to implement in the FSM. The actual net for 2006 was modelled differently and resembles the real network on the ground; a visual inspection with google earth images shows a good match. Abstraction of the planned highway network could account largely for the differences witnessed.

Another reason explaining the difference between the observed and actual network is the fact that the highway network from 1997-2010 was not developed as planned. A link by link analysis of the proposed and committed road networks as per the 1997-2010 Istanbul Transport Master plan has not been undertaken in order to substantiate these claims, the reason being that the available data set did not enable such analysis to be undertaken. Observations from this analysis lead to the conclusion that the highway network used in the FSM could lead to inaccuracy in forecasting ADT; this notion is tested in section 6.5.

6.3. Research Question 3- Projected land use and socioeconomic data inputs for 2006

To address this research question the FSM was run based on the projected population, employment and students for 1996-2006 and the highway network and parameters used were for as per 1997-2010 master plan. The general observation was that the FSM overestimates the ADT by 68.43% PE using projected data inputs. The RMSE reported was 112,947 and the %RMSE was 204.43% indicating high level of inaccuracy. The scatter plot, figure 6-5 below shows that the model ADT and Observed (RTMS) ADT generally disagree. It also reveals the presence of outliers and majority of the observations cluster around 0-60,000 ADT.
The percent error for individual links is presented in figure 6-6 showing the variation in error for the analysed links. It is observed that there are 9 links with a percent error greater than 100% with link No. 62 having the greatest over estimated error of approximately 500%. The under estimated links all fall between 0 to -100% error.

Discussion

The results in general indicate a considerable level of inaccuracy of the FSM outputs and this can be explained by a number of factors. First, the projected land use and socio economic data used was seen to have error as outlined in section 6.1, it could be that this error is propagated through the FSM and hence leading to the ADT error as observed by Zhao and Kockelman (2002). In order to test this assertion, the FSM was run using different land use and socio economic input data while the other variables were held constant, the results are shown in section 6.4.

In terms of the distribution of error on individual case study links it is observed that the error ranges between -100% and 500%. Links No. 27, 29, 62 and 81 reported the highest PE above 300% and from further analysis it was established that they are not located in the areas that experienced growth in human settlement between 1995 and 2005 as shown in figure 6-7 below. Hence the error reported may not necessarily be related to unexpected development of human settlements but could be due to other factors such as road type and number of lanes.
When the percent error of the case study links is compared to the error in the land and socio economic forecasts, it is observed that the links which have the highest PE are not in the TAZs with the highest forecast PE (figures 6-8, 6-9, 6-10 and table 6-2). The PE on links does not correspond with the PE of forecasts in the TAZs, however the links analysed are too few to make a firm conclusion. The error from the analysed links might be generated from other sections of the highway network.

Figure 6-7: Spatial distribution of case study links and land cover change -1995 to 2005

Figure 6-8: RTMS vs TAZ Forecast PE-Population

Figure 6-9: RTMS vs TAZ Forecast PE-Employment
Table 6-2: RTMS vs TAZ Forecast PE

<table>
<thead>
<tr>
<th>RTMS ID</th>
<th>RTMS PE %</th>
<th>Percent error Population %</th>
<th>Percent error Employment %</th>
<th>Percent error Student %</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>342</td>
<td>69</td>
<td>-1</td>
<td>39</td>
</tr>
<tr>
<td>29</td>
<td>426</td>
<td>54</td>
<td>-52</td>
<td>21</td>
</tr>
<tr>
<td>60</td>
<td>269</td>
<td>115</td>
<td>155</td>
<td>199</td>
</tr>
<tr>
<td>62</td>
<td>507</td>
<td>-9</td>
<td>-31</td>
<td>-22</td>
</tr>
<tr>
<td>81</td>
<td>453</td>
<td>-9</td>
<td>-31</td>
<td>-22</td>
</tr>
</tbody>
</table>

Figure 6-10: RTMS vs TAZ Forecast PE-Students

6.4. Research Question 4- Actual land use and socio economic data inputs for 2006

To address this research question the FSM was run based on the actual population, employment and students for 2006 and the highway network and parameters used were for as per 1997-2010 master plan. The general observation was that the FSM overestimates the ADT by 68.45% PE using actual data inputs. The RMSE is 107,749 and the %RMSE was 195.02% indicating high level of inaccuracy. The scatter plot, figure 6-11 below shows that the model ADT and observed (RTMS) ADT generally disagree; it also reveals the presence of outliers and majority of the observations cluster around 0-60,000 ADT.

Figure 6-11: Scatter plot: Model ADT vs Observed ADT-Actual Data Input

The percent error for individual links is presented in figure 6-12 showing the variation in error for the analysed links. It is observed that there are 9 links with a percent error greater than 100% with link No. 81 having the greatest over estimated error of approximately +500%. The under estimated links all fall between 0 to -100% error with no outliers reported.
Discussion

The results from this analysis also indicate that the FSM underestimates the ADT even when actual data for 2006 is used in the model. This could mean that the actual land use and socio economic data used was not accurate or that there are other factors explaining the model inaccuracy. However the reasonableness checks and validation of the actual land use and socio economic data presented in section 4.6 showed that the data was a good representation of the situation in 2006 as it was based on official records and surveys as per the 2007-2023 Istanbul Transport Master Plan. Hence the error could be explained by the other factors such as the highway network, the model parameters, the model set up and errors arising due to relating the RTMS data to the case study links. Studies by Zhao and Kockelman, (2002), Horowitz and Emslie, (1978), Parthasarathi and Levinson, (2008), Bonsall et al., (1977) mention these factors.

In general the FSM performed better when run using actual data than when using projected data which means that emphasis should be placed on achieving accurate forecasts. The %RMSE for actual data was 195.02% and that of projected data was 204.43% showing a difference of approximately 9%. These results differ from the observation made in the work by Zhao et.al.,(2005) because the application of actual data did not improve the %RMSE by a large magnitude. The distribution of error among the individual case study links had a moderately similar pattern for actual and projected data as shown in figures 6-6 and 6-8.

In this case links No. 62 and 81 reported the highest PE above 300% and they are not located in the areas that experienced growth in human settlement between 1995 and 2005 as shown in figure 6-7. This demonstrates that other factors other than the land use and socio economic variables affect model accuracy. However, these links are part of the International highway network and are located near the first Bosporous bridge this could mean that they experience a lot of traffic and hence are more prone to error.
6.5. Research Question 5- Year 2006 highway network

To address this research question the FSM was run based on the same inputs as for research question four but using the highway network for 2006. The general observation was that the FSM overestimates the ADT by 49.63% PE using the highway network for year 2006. The RMSE was 76,677 and the %RMSE was 138.78% indicating high level of inaccuracy. The scatter plot, figure 6-13 below shows that the model ADT and Observed (RTMS) ADT generally disagree; it also reveals the presence of outliers and this time the observations seem to be evenly distributed.

![Scatter Plot: Model ADT Vs Observed ADT - Highway network 2006](image)

Figure 6-13: Scatter plot: Model ADT vs Observed ADT-research question 5

The percent error for individual links is presented in figure 6-14 showing the variation in error for the analysed links. It is observed that there are 10 links with a percent error greater than 100% with link No. 21 having the greatest over estimated error of approximately 300%, no outliers are observed. The under estimated links all fall between 0 to -100% error with no outliers reported.

![Percent Error for Individual Links](image)

Figure 6-14: Percent error for individual links-research question 5

**Discussion**

Model results from this analysis suggest that the FSM reports high inaccuracy when using the actual highway network but the results are better than when using the planned highway network, all other variables held constant. It was seen that the actual network reported a %RMSE of 138.78% compared to 195.02% for the planned network which marks a difference of 56.245. This is a big difference
meaning that the highway network used in the FSM could affect the accuracy results greatly. The reason for error in this analysis could be due the modelling technique used to represent the network as explained in 6-2.

The error reported in this analysis could largely be explained by the fact that the networks were prepared differently and also by the fact the perhaps the planned network consisting of committed and proposed projects was not developed as planned. A similar observation can be found in the work by Giuliano (1984), Zhao et.al., (2005) and Parthasarathi and Levinson (2008). The percent error on individual links had a range between 300% and -100% and no outliers were observed, this means that the distribution of error among the case study links was more uniform with the actual network than with the planned network.

In this case links No. 2, 20 and 58 reported the highest PE and it was observed that link No. 2 and 20 are in close proximity to areas that witnessed growth in human settlements between 1995 and 2005, as can be seen in figure 6-7 in section 6.3. This means that the FSM land use and socio economic inputs may fail to capture the dynamics of development and urban growth due to their static conceptualisation. The inputs used did not capture the urban growth witnessed in areas surrounding link No. 2 and 20.

On another note it is observed that these links with high error form part of the international highway network with Link No. 2 being at close proximity to the second Bosporus crossing (Fatih Sultan Mehmet bridge), it is likely that errors can be linked to this aspect but exactly how is not clear.

6.6. Research Question 6- Year 2006 model parameters

To address this research question the FSM was run based on the same inputs as for research question four but using the model parameters for the year 2006. The general observation was that the FSM underestimates the ADT by -38.40% PE using model parameters for year 2006. The RMSE was 55,757 and the %RMSE was 100.92% indicating high level of inaccuracy. The scatter plot, figure 6-15 below, shows that the model ADT and observed (RTMS) ADT generally disagree; it also reveals the presence of outliers and majority of the observations cluster around 0-60,000 ADT.

![Scatter Plot: Model ADT vs Observed ADT-Year 2006 parameters](image-url)

Figure 6-15: Scatter plot: Model ADT vs Observed ADT-research question 6
The percent error for individual links is presented in figure 6-16 showing the variation in error for the analysed links. It is observed that there are 3 links with a percent error greater than 100% with link No. 62 an outlier having the greatest over estimated error of approximately 350%. The under estimated links all fall between 0 to -100% error with no outliers reported.

![Percent Error for Individual Links-Year 2006 Parameters](image)

Figure 6-16: Percent error for individual links-research question 6

**Discussion**

The results from this analysis indicate that the %RMSE is better than that reported in the previous section which means that the parameters used have an effect on the FSM accuracy. This analysis made use of trip generation parameters for the year 2006 and implemented regression analysis for this step of the FSM. The trips rates used and coefficients for the regression model are shown in section 5.3. It is observed that the trip rate changed between 1996 and 2006 and this could be a major factor affecting the accuracy of the model, the error could have arisen because it is difficult to predict the trip rate accurately as this depends on the socio economic and land use situation at the specific point in time. A similar observation can be found in the work by Niles and Nelson (2001), Krishnamurthy and Kockelman (2003).

The 1996 FSM model developed by IMM and ITU made use of cross classification analysis for trip generation, it is not clear exactly how the different trip generation methods and variables affect the model accuracy and it is recommended that further research should test this.

Trip rates in Istanbul have increased over time from 1.54 in 1996 to 1.74 in 2006, while the motorised trip rate has decreased from 1.00 in 1996 to 0.88 in 2006. An explanation of the changes in mobility trends in Istanbul can be found in section 3.5 of this report and in Gercek and Demir, (2008). It is expected that changes in the socio economic variables shown in table 6-3 below would have a direct consequence in increasing trip generation assuming all other factors in the study area are held constant, that is, assuming that there are no polices and transport strategies aimed at reducing travel demand.
Table 6-3: Comparison of population, employment and student numbers for Istanbul between 1996 and 2006

<table>
<thead>
<tr>
<th></th>
<th>Year 1996</th>
<th>Year 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>9,349,965</td>
<td>11,766,554</td>
</tr>
<tr>
<td>Employment</td>
<td>2,505,000</td>
<td>3,677,000</td>
</tr>
<tr>
<td>Students</td>
<td>1,462,617</td>
<td>2,566,554</td>
</tr>
</tbody>
</table>

(Source: 2007-2023 Istanbul Transport Master Plan Study)

Land use factors and the urban structure also play a key role in influencing trip rates. With examples shown in the literature review section 2.3 of this report, it is noted for example that a compact land use structure in which the residential areas are in close proximity to the working and shopping areas is likely to reduce the travel demand and hence reduce the trip rates. According to the 2007-2023 Istanbul Transport Master Plan Study, the city has developed through expansion of urban areas in three major directions namely, towards the west, towards the east and towards the north from the old urban centre. This has lead to a linear east-west development which has likely increased the travel distances and the travel demand.

As discussed in the literature review, section 2.3, infrastructure development and transportation strategies are known to have an effect on travel demand. In the case of Istanbul, it was mentioned by one of the key informants involved in this study that the construction of the Bosporus bridges increased the rate of trips made between the European and Anatolian sides of Istanbul. Although not yet studied, it is assumed that the increase in road, rail and sea transport capacity has lead to induced travel demand.

In summary, this brief explanation of the effect of changes in FSM parameters is not conclusive and lacks scientific rigour. However it gives an indication of some of the variables that can explain the changes in FSM parameters especially for trip generation. Further studies should be undertaken to explain the factors that lead to changes in FSM parameters for trip generation as well as the trip distribution, modal split and assignment steps of the model because they are important factors in the FSM accuracy.

6.7. Research Question 7- Disaggregated TAZs (451)

To address this research question the FSM was run based on the same inputs as for research question five but using 451 TAZs as opposed to 250. The general observation was that the FSM overestimates the ADT by 92.16% PE using model parameters for year 2006. The RMSE was 100,451 and the %RMSE was 181.81% indicating high level of inaccuracy. The scatter plot, figure 6-17, below shows that the model ADT and observed (RTMS) ADT generally disagree; it also reveals the presence of outliers and the observations are fairly well spread.
The percent error for individual links is presented in figure 6-18 showing the variation in error for the analysed links. It is observed that there are 11 links with a percent error greater than 100% with link No. 2 having the greatest over estimated error of approximately 300%. The under estimated links all fall between 0 to -100% error with no outliers reported.

Discussion

Results from this analysis can be compared with those from research question five because the model inputs are the same and the only variation being that in this research question 451 TAZs are used instead of 250. The results for the 451 disaggregated TAZs show a % RMSE of 181.81 while those of 250 TAZs show a %RMSE of 138.78% marking a difference of 43.03% due to disaggregation. The results are interesting because it was expected that disaggregation should lead to better accuracy of the FSM yet in this case it lead to less accurate results. This analysis was based on 31 RTMSs and these could account for the unexpected results, it is suggested that further analysis based on the entire highway network could better explain the effect of disaggregated TAZs. Links No. 2, 20 and 58 register the highest error and as mentioned previously in section 6.5, they form part of the international highway network and share the same linear spatial location.
Disaggregation affects trip distribution and also traffic assignment because the TAZ centroids are connected to the nearest highway network link. Thus the nearest highway link connected to the TAZ centroids and assigned with traffic change if the total number of TAZs is different. The individual case study links show a uniform distribution of percent error with no major outliers, this could mean that the error is spread through the network but this could be on account of the equilibrium assignment used.

6.8. Research Question 8- Feedback analysis

To address this research question the FSM was run based on the same inputs as for research question four but implementing a feedback mechanism between traffic assignment and trip distribution. The general observation was that the FSM overestimates the ADT by 53.18% PE using a feedback mechanism. The RMSE was 94,340 and the %RMSE was 170.75% indicating high level of inaccuracy. The scatter plot, figure 6-19 below, shows that the model ADT and observed (RTMS) ADT generally disagree; it also reveals the presence of outliers and majority of the observations cluster around 0-60,000 ADT.

The percent error for individual links is presented in figure 6-20 showing the variation in error for the analysed links. It is observed that there are 8 No. links with a percent error greater than 100% with link No. 81 an outlier having the greatest over estimated error of approximately 500%. The underestimated links all fall between 0 to -100% error with no outliers reported.
Discussion
The FSM accuracy assessment results presented above suggest that the use of a feedback mechanism between the assignment step and the trip distribution step by means of applying the BPR congestion function on travel time improves the performance of the model. The %RMSE for the FSM with feedback was 170.75% while that for a similar model set up without feedback was 195.02% indicating an improvement of 24.27%. The performance for the individual links was almost similar with links No. 62 and 81 reporting the highest error (PE) of approximately 400% and 500% respectively. The explanation for this is similar to that discussed in section 6.4.

These results can be interpreted to mean that the use a feedback mechanism in the trip distribution step of the FSM is more logical and represents reality better, this because in the traditional form of the model, trip distribution relies on free flow travel time which fails to capture the realistic traffic situation on the ground in which travel time on links is affected by the volume of traffic in relation to the capacity of the links. The estimates of actual travel time are not available until after completion of traffic assignment (Deakin et al., 1993), hence the need for iterations.

The number of iterations between assignment and trip distribution could also have an effect on model accuracy, this research made use of only one iteration, further investigation should be undertaken to determine the effect using more iterations.

6.9. Research question 9- Effect size of manipulating FSM using different inputs
The F-ratio was used to determine the effect size of using different FSM inputs. It compares the size of the variation due to the experimental manipulations with the size of the variation due to random factors. If the variance due to model inputs is big relative to variation due to random factors the value for the F-ratio is big and we can conclude that the observed results are unlikely to have arisen by chance. If the F-ratio is greater than 1, we can say that experimental manipulation had some effect above and beyond the effect of extraneous factors (Field, 2005).

The Mauchly’s test indicated that the assumption of sphericity had been violated (χ²(5)=184.730,p<.05); therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ε=.32. The results indicated that the PE on case study links was significantly affected by the model inputs used F(1.600, 47.995)=4.058, p<.05, above and beyond extraneous factors. In order to determine the effect size of the individual FSM input, tests of within-subject contrasts were undertaken and the results are summarised in table 6-4 below: Appendix 8 gives the full SPSS output.

The results show that using new parameters and applying feedback to the FSM has the greatest effect on model performance, the within-subject contrasts revealed an F-ratio of 12.962 and 10.073 respectively with p<.05 indicating that there was significant difference between FSM results when these variables were used, they also give large effects of .549 and .501 respectively. Manipulation of the FSM by using disaggregated TAZs, actual highway network and projected data also leads to differences in the model results but these differences are not significant with values of p>.05 and effect sizes of .142, .062 and .142 respectively. These results differ with the observations made in the work
by Zhao et al., (2005) where the effect of using the actual highway network was the most significant but are similar to the observations by Zhao and Kockelman (2002) where the trip generation parameters were seen to have the most significant effect.

Table 6-4: Results of within-subject contrast statistics

<table>
<thead>
<tr>
<th>Manipulation Variable</th>
<th>F-ratio</th>
<th>Sig.*</th>
<th>( \omega^2 ) **</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. New parameters</td>
<td>12.692</td>
<td>.001</td>
<td>.549</td>
</tr>
<tr>
<td>2. Feedback</td>
<td>10.073</td>
<td>.003</td>
<td>.501</td>
</tr>
<tr>
<td>3. Disaggregation</td>
<td>.617</td>
<td>.438</td>
<td>.142</td>
</tr>
<tr>
<td>4. Actual highway</td>
<td>.116</td>
<td>.736</td>
<td>.062</td>
</tr>
<tr>
<td>5. Projected data</td>
<td>.036</td>
<td>.851</td>
<td>.035</td>
</tr>
</tbody>
</table>

*Critical (P) value = .05
** \( \omega^2 \) -Omega squared (effect size)

6.10. Comparison of the spatial distribution of percent error on case study links

The spatial distribution of percent error in case study links under different FSM model runs is presented below (figure 6-21). The negative errors are indicated in red representing underestimation while colour green is used to represent over-estimation. The size of the point symbols signifies the magnitude of the percent error.

The spatial distribution of percent error varies depending on the inputs of the FSM. For all of the model runs, the largest positive percent error follows a sequence from sensor ID 29 through to 28, 27, 50, 51, 15, 60, 80, 81 and 62. These links comprise the O-1 motorway that connects the European and Asian sides of Istanbul through the Bosporus Bridge on the south; this motorway is older and is mostly used for inner-city traffic. However this observation is different when the model is run using the actual highway network.

It is also noted that the links with lowest and negative percent error follow a sequence from sensor ID 58 through to 20, 13, 2, 72, 82 and 63. These links form the O-2 motorway which crosses the Bosporus via the Fatih Sultan Mehmet bridge on the north, this motorway is more recent and mostly used by intercity and intercontinental traffic. However this distribution of error changes when the model is run using the actual highway network.

It is therefore likely that the spatial distribution of percent error on the highway links is influenced by the highway network used in the model or by the fact that the model over estimates inner-city traffic flows.
Figure 6-21: Spatial distribution of case study link percent error for different FSM runs
7. Summary of Findings

This section briefly presents a summary of the main findings from the data analysis results and discussion. The results have been compared and a synthesis provided that addresses the specific objectives of this research. Limitations experienced while undertaking the research are explained.

7.1. Summary

The results from analysis undertaken in this study indicate that the FSM has a %RMSE of between 100-200% depending on the inputs used. There are cases both of overestimation and underestimation of the ADT. Results of the 1997-2010 Istanbul Transport model developed by IMM and ITU using TRANPLAN/EMME/2 software packages were assessed for accuracy using the same methodology as used in this study. It was observed that the IMM/ITU model underestimates the traffic volumes for the assessed links but is generally more accurate than the results from this research as shown in table 7-1 below.

Table 7-1 shows that the use of new parameters in the FSM gives the best results of 100.92%RMSE followed by, the highway network, feedback, disaggregated TAZs, actual data and projected data. The highest error is 204.43%RMSE due to application of projected data.

<table>
<thead>
<tr>
<th>FSM Manipulation</th>
<th>PE (%)</th>
<th>RMSE</th>
<th>%RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. New Parameters 2006</td>
<td>-38.40</td>
<td>55,757</td>
<td>100.92</td>
</tr>
<tr>
<td>2. Actual Highway 2006</td>
<td>49.63</td>
<td>76,677</td>
<td>138.78</td>
</tr>
<tr>
<td>3. Feedback</td>
<td>53.18</td>
<td>94,340</td>
<td>170.75</td>
</tr>
<tr>
<td>4. Disaggregate data</td>
<td>92.16</td>
<td>100,451</td>
<td>181.81</td>
</tr>
<tr>
<td>5. Actual Data</td>
<td>68.45</td>
<td>107,749</td>
<td>195.02</td>
</tr>
<tr>
<td>6. Projected Data</td>
<td>68.43</td>
<td>112,947</td>
<td>204.43</td>
</tr>
<tr>
<td>1997-2010:IMM/ITU model*</td>
<td>-61.99</td>
<td>19,770</td>
<td>68.67</td>
</tr>
</tbody>
</table>

(*Based on 1997-2010 Istanbul Transportation Master plan)

The main findings of this research are summarised as follows:

Land use and socio economic forecasts v.s Actual data

Land use and socio economic forecasts of population, employment and students for the year 2006 did not match with the actual data in 2006. This can be attributed to the fact that it is difficult to predict the future with accuracy especially given a long-term horizon such as 10 years. Furthermore the forecasts are based on anticipated future conditions which might not have materialised between 1996 and 2006. The error could also have been due to the inaccuracy of the data preparation process and overlay procedures used in this research.
The use of projected/forecast land use and socio economic data in the FSM reported the greatest error and the main explanation for this is that the input data used was with error. The use of actual data in the FSM leads to a model improvement of 9.41%RMSE which means that the forecast error propagates through the FSM but it does not have a significant effect on final assignment results.

**Planned highway network vs Actual highway network**

The planned and actual highway network were inconsistent and the reason for this was that the planned network was simplified while the geometry of the actual network was a better geographical representation of the on the ground network. Another reason for the inconsistency between the networks could be that the planned highway network consisting of committed and proposed road construction projects was not developed accordingly; however, this was not tested in this research.

Application of the actual highway network improves FSM model results by 56.24% RMSE which means that the network used in the FSM affects model results but it is also noted that the effect is not significant.

**New model parameters 2006**

The use of FSM model parameters for the year 2006 improves the results by 95.10% RMSE. This means that accurate estimation of parameters has the greatest effect on accuracy of the FSM outputs. In this case parameters for trip generation were tested and this means that the use of accurate land use and socio economic forecasts may not necessarily translate into better FSM results if the parameters used are unrealistic to the forecasted period.

**Disaggregation of TAZs from 250 to 451**

The application of disaggregated TAZs from 250 to 451 leads to a decline in model performance with a difference of -43.03%RMSE but the effect was not significant. This could either mean that disaggregation leads to greater inaccuracy or alternatively that the other model runs using higher aggregation levels of 250 TAZs gave better but unrealistic results. The latter is more likely to be true based on the theory that disaggregation should result in better model accuracy. The use of more RTMS observations may affect these results.

**Feedback mechanism**

Implementation of a feedback mechanism in the FSM between traffic assignment and trip distribution improved model performance by 24.27%RMSE and the effect was significant. This means that the sequential nature of the FSM leads to inaccuracy in model results and that inclusion of a congestion effect on travel times is a better representation of reality.

**Spatial distribution of error in the case study links**

It was established that spatial distribution of FSM error followed a regular pattern along a sequence of links forming a route; this indicates that the error is systematic rather than random. This is confirmed by the observation that the spatial distribution of error changed when a different highway network was used as seen in figure 6.21.

These findings are interesting but they should be appreciated in consideration to the fact that this study was faced by a number of limitations which are presented in the section below.
7.2. Limitations

There were a number of limitations encountered while undertaking this research which could influence the findings significantly. The main limitation experienced during field work was the limited access to data for the 1997-2010 Istanbul Transport Master plan.

Another major limitation was in regard to the data acquired, although majority of the data needs were met, some crucial data was not available in the desired form. Furthermore, the acquired data was assumed to be accurate with minor validation checks undertaken. It would have been interesting to make use of the existing land use as of 1996 and 2006 in the study but such data was not forthcoming, hence, land cover was used as an alternative. Data preparation was another limitation where the necessary computer software to convert data formats into usable form had to be acquired, studied and installed for use in this research.

The methodology used in this research also had some limitations, for instance the method used to assign land use and socio economic attributes to the TAZs for the year 1997 was not a good representation of the reality as explained in section 4.6. In order to deal with this limitation, the spatial distribution of population, students and employment in TAZs was compared against the land cover map and it was seen that areas with high values for these variables were also the areas with high concentration of human settlement land cover.

The method of computing RTMS ADT and associating with the highway network as discussed in section 5.5 might have been a source of error in this research. The use of 31 sensors and links for this study is another major limitation, the sample is small and may not give a true account of the accuracy of the FSM. The reason for the application of such a small sample size was due to the great volume of the data captured by each RTMS and the immense capacity demands with over 900,000 records for each RTMS.

A combination of Flowmap, ArcGIS, OmniTRANS, MS Excel and SPSS were used to operationalise the FSM. It is likely that in using this computer software applications some errors could have been committed or factors omitted. For example while working in Flowmap, the doubly constrained gravity model for trip distribution was set up using the model estimated parameters for MTL and beta values, the MTL for Istanbul of approximately 18Km was not applied in this research. Some errors might have also occurred while importing the highway network links, nodes, centroids and TAZ shape files into OmniTRANS in order to perform traffic assignment. The use of speed limit for the links free-flow speed was yet another limitation and potential source of error.

The focus of this study was on private transport and 24hr traffic, the effect of other modes and time of day considerations were not analysed. These omissions may have had an influence on the results.
8. Conclusion and Recommendation

This section presents the main conclusions of the research based on the objectives established at the onset; the purpose is to reflect on the salient features of the study. Recommendations are also provided based on the findings and limitations with the aim of addressing the research problem.

8.1. Conclusion

The theme of this research was integrated land use and transportation modelling. Focus was placed on transport modelling with the aim of evaluating the application of the FSM to assess the factors that influence its accuracy using the case of Istanbul, Turkey. According to literature it was shown that there were different inputs and modelling techniques used in the FSM that acted as factors affecting the accuracy of model outputs. The conceptual framework developed in this study categorised these factors into three groups namely, basic physical, land use and socio economic factors (disaggregation of TAZs, road network, population, employment and students), behavioural factors (model parameters for trip rates) and model structure factors (feedback mechanism).

The forecast period for this research was between 1996 and 2006. The methodology involved comparing the forecasted FSM daily traffic against the observed daily traffic derived from RTMS records. The 1997-2010 Istanbul transport model developed by ITU and IMM was used as the basis for analysis. Forecasts for population, students and employment and also the planned road network were validated against actual data for 2006. A total of six FSM runs were undertaken using different inputs, namely projected data, actual data, disaggregated TAZs, actual highway network, new model parameters and finally a feedback mechanism. The results from these six model runs were assessed for accuracy by using the %RMSE statistic. The effect size of using each different input was measured by means of the F-test and omega squared ($\omega^2$).

The population and employment forecasts had been overestimated while those for students had been underestimated when compared against actual data. The planned road network was also significantly different from the actual road network. The %RMSE accuracy assessment results for the FSM under different inputs was as follows; projected data 204.43%; actual data 195.02%; disaggregated TAZs 181.81%; actual highway 138.78%; new parameters 100.92%; and feedback 170.75%. In terms of effect size, the use of new parameters had the most significant effect ($\omega^2=.549$) followed by application of a feedback mechanism ($\omega^2=.501$). The use of projected data, disaggregation, actual data and the actual highway had insignificant effect on the FSM.

It can be concluded that the land use and socio economic projections were inaccurate because they did not match with the actual data. They account for the greatest source of inaccuracy within the FSM when measured in terms of %RMSE because when other inputs are used in the model, the %RMSE improves significantly. Even with the use of actual land use and socio economic data, the %RMSE was still the worst and changed insignificantly. However, it should be noted that in this study only population, employment and students were used and there were limitations in the data preparation.
The application of new model parameters however leads to the best results and most significant effect on the FSM with a %RMSE of 100%. This means that between 1996 and 2006 the travel behaviour and trip generation characteristics in Istanbul changed significantly but due to the static and sequential nature of the FSM, the trip generation parameters for the 10 year forecast period did not change and still reflected the 1996 situation. Hence, it seems that predicting the future trip generation parameters together with forecasts of land use and socio economic variables would lead to improvement of FSM accuracy. The component of land use could thus be made more implicit in the FSM based on how the parameters are used.

It can be concluded that the method in which the transportation network is modelled has a significant influence on the accuracy achieved. Results based on the actual highway network for 2006 were better than those based on the planned highway network in 1996. The reason for this is that the 2006 network was a better representation of reality than the 1996 network which was abstract.

The inclusion of feedback also significantly improves the FSM accuracy. It means that if the trip distribution step is undertaken with consideration for the congestion effect based on the assignment results, then the trips (productions and attractions) will be distributed across the TAZs based on more realistic impedance factors.

Disaggregation of the TAZs does not lead to improved FSM accuracy and neither is the effect significant because when 250 TAZs were used the %RMSE was better than when 450 TAZs were used. This was surprising because it was expected that disaggregation would lead to better accuracy.

In summary, despite the inaccuracy reported by the FSM it is still applicable for use because based on the F-test, the results obtained were on account of model manipulation above and beyond the effect of extraneous factors. The behavioural factors were seen to be the most important factors to be considered in the FSM followed by the model structure and finally the physical, land use and socio economic factors.

There were a number of limitations observed in this research which should be considered when interpreting the findings and appreciating the conclusions. The next section presents recommendations aimed at addressing the limitations observed in this research.

8.2. Recommendations

In line with the mentioned findings, limitations and conclusions the following recommendations are made:

The projection of future land use and socio economic variables in the TAZs should be made more accurate because errors from these projections are propagated through the FSM leading to error in the ADT reported by the model.

Estimation of the trip generation parameters in the FSM should be improved. The process should be dynamic in order to capture the anticipated changes in trip making behaviour over the planning horizon. Furthermore, the use of cross classification and regression analysis in the trip generation step should be assessed to determine the effect on FSM accuracy.
The FSM structure should include components of feedback between the different steps of the model where the congestion effect based on the BPR function is applied on travel time and speed. In this study iteration between traffic assignment and trip distribution was used leading to more accurate results.

Further research should use the logit model of modal split based on the 2007-2023 IMP/JICA Transport Master plan and multiple modes of transportation to determine the effect of modal split estimates on the FSM accuracy.

Developing on this study, further research should be undertaken where the temporal aspect of the FSM is disaggregated to AM and PM peak hours as opposed to 24hr traffic used herein. The RTMS data acquired from this study provides opportunity for further research into dynamic travel demand modelling because it provides real time traffic data at 5-minute intervals daily for a period between 2006 and 2009.

The migration from Flowmap to OmniTRANS in this study could have resulted in error at the assignment step. It is recommended that further research should be undertaken using a single transport modelling software application. OmniTRANS provides different methods for traffic assignment; the method used in this study was the volume averaging equilibrium assignment method in which a uniform stochastic factor was applied. Further research should consider assessing the effect of the alternative methods of traffic assignment on FSM accuracy.

This study used a total of 31 RTMS observations of ADT for a period of 3 months; it is recommended that more RTMS observations having a wider coverage of the study area should be used in further research to improve the validity of the accuracy assessment.

Further research should be conducted to explore the effect of changes in land use and socio economic variables on the FSM. The analysis should involve quantifying the changes in for example residential, commercial, and industrial land use in the TAZs based on time steps within the planning horizon. Calculation of the population, employment and students in the TAZs should be based on this analysis of land use change. This means that a dynamic model of land use would need to be developed to compute the changing land use variables over the planning horizon. The outputs from this model would then be used in the FSM in an iterative process. The concept of land use-transport interaction as shown in section 2.3, figure 2-1 and explained by Tillema, (2004) is recommended for application.

The need for the FSM could be replaced by use historical and real time RTMS data from the monitored links. If the traffic flows are captured for a considerable period of time such as 5 years, the traffic flow behaviour and patterns can be established. These patterns can then be projected into the future and thus estimates of future travel demand established based on the assumption that people will keep travelling in the same way that they have been in the past. Scenarios can be used to make these projections realistic. If new links are added to the transport network, a formula can be derived to estimate how the traffic flows will be redistributed to make use of the new link.
TOWARDS INTEGRATED LAND USE AND TRANSPORT MODELLING: EVALUATING ACCURACY OF THE FOUR STEP TRANSPORT MODEL - THE CASE OF ISTANBUL, TURKEY

References


IMM., ITU. (1997). Istanbul transportation master plan, final report (Turkish version), Istanbul, Turkey.


Appendices

Appendix A: Research matrix ................................................................. 79
Appendix B: Logic of FSM ................................................................. 80
Appendix C: Data needs ................................................................. 81
Appendix D: Data wish list ................................................................. 82
Appendix E: MOU between IMP and ITC .................................................. 83
Appendix F: O-D matrix set up in flow map ................................................ 84
Appendix G: OmniTRANS job description in Ruby – assignment ................................................ 84
Appendix H: Results for F-test in SPPSS .................................................. 85
### Appendix A: Research matrix

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data/Information Required</th>
<th>Data Sources</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How far does projected land use in 1997 match with the observed land use situation in 2006?</td>
<td>Reports and statistics on transport master plan studies 1997-2010 and 2007-2023</td>
<td>Secondary</td>
<td>Literature review, Key informant interview, Comparative analysis</td>
</tr>
<tr>
<td>3. How do the model outputs based on projected land use for the 1997-2010 model match with observed traffic volumes on selected links?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary/secondary</td>
<td>Four step Transport model run, Comparative analysis, visual analysis</td>
</tr>
<tr>
<td>4. Do the model outputs based on actual 2006 land use inputs have a better/worse match with reality on selected links?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary/secondary</td>
<td>Four step Transport model, Comparative analysis, visual analysis</td>
</tr>
<tr>
<td>5. Do the model outputs based on 2006 transport network have a better/worse match with reality on selected links?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary and secondary</td>
<td>Four step Transport model, Comparative analysis</td>
</tr>
<tr>
<td>6. Do the model outputs based on 2006 parameters have a better/worse match with reality on selected links?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary and secondary</td>
<td>Four step model, Comparative analysis</td>
</tr>
<tr>
<td>7. Does the implementation of disaggregation in the FSM lead to better accuracy?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary</td>
<td>Four step model, Comparative analysis</td>
</tr>
<tr>
<td>8. How does the implementation of feedback affect the convergence of model outputs and reality?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary/secondary</td>
<td>Literature review, Comparison of Outputs from research question 3-8</td>
</tr>
<tr>
<td>9. Which variation of model input set up has the greatest effect on outputs?</td>
<td>Model outputs, Observed traffic volumes</td>
<td>Primary and secondary</td>
<td>Iteration in four step model, Comparative/statistical analysis</td>
</tr>
</tbody>
</table>
Appendix B: Logic of FSM

Source: Wells, 1975)
## Appendix C: Data needs

<table>
<thead>
<tr>
<th>Step</th>
<th>Model</th>
<th>Inputs</th>
<th>Output</th>
<th>Information /Data needs</th>
</tr>
</thead>
</table>
| Trip generation       | o Multiple regression models                                          | Socio economic and land use data such as population, employment, students, car ownership etc, presented in TAZs Parameters for model Trip rates | Prediction of the number of trips each household makes as a function of its characteristics and the number of trips attracted to a zone as a function of the zone’s characteristics. | - Model : 1997-2010  
- Land use and socio economic data 2005  
- Calibration report showing parameters, formulae and methodology for 1997-2010 and 2007-2023 transport plans |
|                       | o Cross-classification                                                |                                                                        |                                                                        |                                                                                          |
|                       | o Trip rate analysis                                                 |                                                                        |                                                                        |                                                                                          |
| Trip distribution     | o Gravity model with generalized cost function                       | Zonal Productions and attractions from the step above Gravity Model parameters | Prediction of origin-destination flows, i.e, the linking of the trip ends predicted by the trip generation model together to form trip interchanges or flows( | - Calibration report showing parameters, formulae and methodology for 1997-2010 and 2007-2023 transport plans |
|                       | o Growth-factor update                                               |                                                                        |                                                                        |                                                                                          |
| Modal split           | o Direct-demand model, Trip-interchange models, Logit-models         | O-D matrix from the step above Model parameters based on travel behaviour; modal split ratios | Prediction of the percentages of travel demand that will use each of the available modes between each origin destination path(vehicle trips) | - Calibration report showing parameters, formulae and methodology for 1997-2010 and 2007-2023 transport plans |
| Assignment            | o All-or-nothing User-equilibrium modeling Social-equilibrium modeling | O-D matrix by mode form above ;Multi modal ;transport network;Vehicle occupancy rates | Assignment of the modal O/D flows on specific routes of travel through the respective modal network(using shortest paths) | - Calibration report and 2007-2023 transport plans Traffic counts for validation. |
## Appendix D: Data wish list

<table>
<thead>
<tr>
<th>Data description</th>
<th>Attributes</th>
<th>Scale</th>
<th>Time period</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public and road transport networks (i.e. for roads, bus routes, rail, tram, ferry etc)</td>
<td>Hierarchy, capacity, schedules/ frequencies of use, travel speeds/time, Annual Average Daily Traffic, traffic flow data, Origin-Destination data, bus/train/metro/tram stops. Present network; Base network (present network+ committed projects under construction or bidding and design after approval; MAXIMUM network = base network plus new projects (new projects i.e planned by imm but not approved, planned or opposed by organisations, projects planned in the study); Master Plan network= maximum network + alternative network)</td>
<td>metropolitan level</td>
<td>1997/2007 base year and projected</td>
<td>shape file or Transcad</td>
</tr>
<tr>
<td>Map of proposed new transport networks</td>
<td>All modes</td>
<td>metropolitan level</td>
<td>1997 and 2007</td>
<td>shape file or Transcad</td>
</tr>
<tr>
<td>TAZ</td>
<td>With land use and socio economic data(population, employment, school students)</td>
<td>metropolitan level</td>
<td>1997/07 (base year and projected)</td>
<td>shape file or Transcad</td>
</tr>
<tr>
<td>Transport master plan IMM/ITU &amp; IMM/JICA</td>
<td>Reports (english)</td>
<td>metropolitan level</td>
<td>1997 and 2007</td>
<td>Text</td>
</tr>
<tr>
<td>Calibration Report</td>
<td>Parameters to use in 4 steps, formulae, methodology, (english)</td>
<td>metropolitan level</td>
<td>1997 and 2007</td>
<td>Text</td>
</tr>
<tr>
<td>Transport surveys and analysis</td>
<td>Travel behavior, demand and system characteristics, vehicle ownership, trip production rate</td>
<td>metropolitan level</td>
<td>1997 &amp; 2007</td>
<td>Text</td>
</tr>
<tr>
<td>Transport policies</td>
<td>Strategies, objectives</td>
<td>metropolitan level</td>
<td>1997 and 2007 latest</td>
<td>Text</td>
</tr>
<tr>
<td>Existing Traffic conditions</td>
<td>Travel times, volumes, congestion on networks</td>
<td>metropolitan level</td>
<td>latest (2009)</td>
<td>shapefiles or transcad</td>
</tr>
<tr>
<td>Land use master plan</td>
<td>Reports and GIS data</td>
<td>metropolitan level</td>
<td>1995-2006-2009</td>
<td>shapefile</td>
</tr>
<tr>
<td>Existing land use</td>
<td>Reports and GIS data</td>
<td>metropolitan level and 1:25000</td>
<td>1995-2006-2009</td>
<td>shapefile</td>
</tr>
<tr>
<td>Census</td>
<td>GIS data and reports</td>
<td>neighborhood and districts</td>
<td>latest</td>
<td>shape files</td>
</tr>
</tbody>
</table>
Appendix E: MOU between IMP and ITC
Appendix F: O-D matrix set up in flow map

- Origins= TAZ based on research question
- Destinations= TAZ based on research question
- Transport network = 1996 and 2006 based on research question
- Distinction between direction= YES
- Impedance (out) Attribute= TIME 1 or TIME AB (free flow time)
- Impedance (home) Attribute= TIME 2 or TIME BA (free flow time)
- Impedance Unit = Seconds
- Connection method = Lines
- Access Attribute= Full
- Shortest Distance (in impedance units) = 0
- Conversion factor (map units to impedance units) = 0.0

Appendix G: OmniTRANS job description in Ruby – assignment

```ruby
# Perform stochastic equilibrium assignment
FMTURI=[1,10,10,1,3,1]
Write("performing stochastic VA assignment...")
assign=OtTraffic.new
assign.assignMethod=VOLUMEAVERAGING|UNIFORM
assign.load=[1,10,10,1,3,1]
assign.bprFerType=[[1], [0.5, 4.0]]
assign.spread=0.05
assign.seed=-3
assign.iterations=50
assign.epsilon=0.001
assign.saveIterations=true
assign.execute
```
Appendix H: Results for $F$-test in SPPSS

Mauchly’s Test of Sphericity

<table>
<thead>
<tr>
<th>Within Subjects Effect</th>
<th>Mauchly's W</th>
<th>Approx. Chi-Square</th>
<th>df</th>
<th>Sig.</th>
<th>Epsilon&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Greenhouse-Geisser</th>
<th>Huynh-Feldt</th>
<th>Lower-bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectsize1</td>
<td>.001</td>
<td>184.730</td>
<td>14</td>
<td>.000</td>
<td>.320</td>
<td>.335</td>
<td>.200</td>
<td></td>
</tr>
</tbody>
</table>

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept
   Within Subjects Design: Effectsize1

Tests of Within-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectsize1</td>
<td>Sphericity Assumed</td>
<td>261709.607</td>
<td>5</td>
<td>52341.921</td>
<td>4.058</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>261709.607</td>
<td>1.600</td>
<td>163585.360</td>
<td>4.058</td>
<td>.032</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>261709.607</td>
<td>1.676</td>
<td>156163.741</td>
<td>4.058</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>261709.607</td>
<td>1.000</td>
<td>261709.607</td>
<td>4.058</td>
<td>.053</td>
</tr>
<tr>
<td>Error(Effectsize1)</td>
<td>Sphericity Assumed</td>
<td>1934858.746</td>
<td>150</td>
<td>12899.058</td>
<td>4.058</td>
<td>.053</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>1934858.746</td>
<td>47.995</td>
<td>40313.711</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>1934858.746</td>
<td>50.276</td>
<td>38484.739</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>1934858.746</td>
<td>30.000</td>
<td>64495.292</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>