METHOD TO FILL PARTIALLY FILLED ORIGIN DESTINATION MATRIX IN SRI LANKAN CONTEXT

L. A. D. MENDIS

169182J

Master of Science Spatial Planning Management & Design

Department of Town & Country Planning
University of Moratuwa
Sri Lanka

November 2020

METHOD TO FILL PARTIALLY FILLED ORIGIN DESTINATION MATRIX IN SRI LANKAN CONTEXT

LIYANA ANUJA DHAKSHIMA MENDIS

169182J

Dissertation submitted in partial fulfillment of the requirements for the degree

Master of Science Spatial Planning, Management & Design

Department of Town & Country Planning

University of Moratuwa

Sri Lanka

November 2020

Contents

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books)

………………………………………..

Name of the student: L. Anuja D. Mendis

Registration Number: 169182J

Group: 2016/2018

Date: ………………….

CERTIFICATION

I certify herewith that L. Anuja Mendis, Index Number: 169182J of the Master of Spatial Planning Management & Design 2016/2018 Group has prepared this research project under my supervision.

………………………………… ……………………………

Signature of the Principal Supervisor

Signature of the Head of the Department of Town & Country Planning

Date: ………………………

Date: ………………………….

ABSTRACT

An Origin – Destination matrix is a vital aspect in the process of traffic planning. OD matrix table shows the trip production and trip attraction of the each zones. This zones should be identified with the certain physical boundaries. These zones will be called as Traffic Analysis Zones (TAZs) in a transport network.

The values of the origin destination matrix will be obtained by the surveys of the travel data. This will be empirical surveys and real observations. Since these data were obtained by the surveys, the result of the OD matrix may close to the reality. However if the number of TAZs is high, then it is not practical to carried out the surveys to obtain the data. These surveys are very costly and time consuming. Also since it is time consuming obtained survey data will be obsolete. Therefore, developing a partly filled OD matrix is only possible in this situation.

However, we cannot use partly filled OD matrix for decision making purposes in traffic planning. Then partly filled OD matrix is not essential. Then there is an essential requirement to fill the balance unfilled cells in the OD matrix before using it to the travel demand estimation.

In Sri Lankan context, there are 331 divisional secretariats with clear physical boundaries. Therefore divisional secretariat division can be identified as one traffic analysis zone. As a result, there are 109,561 OD pairs in OD matrix for whole Sri Lanka.

There are studies in the world context to solve these types of problem. Therefore it is usable to study those techniques which are used to solve this type of issues in the world wide by various scholars.

This study to find the relevant method to develop fully filled OD matrix with using partially filled OD matrix for entire Sri Lanka.

Finally intention of this research is to fill the trip generation and trip attraction for entire traffic analysis zones (i.e. all divisional secretariats) and find entire 109,951 trip interchange values for the origin destination matrix.

ACKNOWLEDGEMENT

I take this opportunity to thank my research supervisor Dr. P. C. P. De Silva, Department of Town & Country Planning, Faculty of Architecture, University of Moratuwa for his kind support and direction given to me throughout the project.

Also I wish to extend a similar gratitude to Mrs. Malini Herath, Course Coordinator of the Msc. Programme and Dr. Shanaka Kariyawasam for their clear guidance and direction for successfully completion of the master programme.

Also I would like to give my gratitude to Dr. Rangajeewa Rathnayaka to give us clear direction as a head of department of Town & Country Planning.

Further to that I would like to thank all government organizations who gave statistics and all customers and public to support to finish my research surveys successfully.

Finally, I am thankful to my parents, my wife and my family for their patience from the beginning of the master program.

CHAPTER 01

1 INTRODUCTION

1.1 Background

Nowadays, the world is growing fast, including Sri Lanka. On this process villages will change to cities and cities will be more congested. Meanwhile, transportation will be great challenge for the future development. If future development are planned without considering transportation, then entire country will be functioned ineffectively because there is close relationship between transport and land use. In transportation planning, one of the most essential requirement is Origin Destination Matrix (ODM) estimation for the purpose of arriving at the traffic pattern between several zones. (Bera & Krishna Rao, 2011). In transportation planning ODM is acting as a most crucial requirement. The origin- destination matrix is a matrix which is each cell represents the number of trips. In the matrix rows represents origins and column represents destinations.

Momentous transport planning is required for the developing countries like Sri Lanka when changes in the land use and economic state of affairs. Traffic pattern between various zones can be identified with the origin-destination matrix estimation.

As per the traditional methods of estimating origin destination matrix are through large scale sampled surveys like home interviews survey, road side interview, vehicle number plate survey, Telephone survey etc. These surveys are conducted once in 1-2 decades. However, according to the previous scholars in worldwide it is identified that these surveys are difficult to conduct in accurately due to financial constraints and time consuming. Also by the time the survey data collected and processed, the collected O-D data become obsolete. Therefore it is vital to update and revalidate of data set in order to conduct better transportation planning.

Usual methods for the ODM estimation are Roadside interviews, vehicle number plate survey, Telephone survey, Home interviews etc.

The accuracy of the estimation is depend on the quality of the available information, and how this data is combined and weighted from different sources. The usual methods for creating an OD matrix is based on sample size and counting locations. Conducting a complete OD survey manually for an entire city is not practical. Because collection of OD data in many counting locations are time consuming and costly. Due to that the accuracy of the ODM estimation also deducting.

With the new development trend and as a solution for the existing traffic congestion, in Sri Lanka highways, expressways, Bypasses, Flyovers are designing. With reference to this situation there is a need to modelling traffic demand in each towns in terms of trip distribution. Accuracy of the OD estimated increases with the number of counting locations. But with the resource limitation conducting an OD survey in every counting locations in related to the sample size and used to fill each and every cell of the OD matrix is not be possible. Using manual methods, there is a possibility to fill some (percentage, how much cells) cells by but not every cell.

Hence as a solution for that, in International context different scholars have developed ODM estimation methods such as gravity model (year), Growth factor method, intervening opportunity model (year), Entropy maximization model, Maximum likelihood estimation, Bayesian inference, Generalized least square estimator (years).

Throughout in this research it is focused on the ODM estimation method that can be used in Sri Lanka by minimizing these problems.

1.2 Research Problem

The origin destination matrix includes number of trips between a number of origins and destinations per time period. If the O-D matrix contains information on the present situation it is called “base year matrix” This matrix can be filled with all the available traffic and transport data. These data can be categorized as two types such as complete data and incomplete. Complete data means all the trips are served. However due to financial constraints and time constraints this process is never happened.

Accuracy of the household survey, road questionnaires are questionable. Also filling of entire matrix with above survey methods is impossible. Therefore it is mandatory to find a solution for complete a origin destination matrix.

In worldwide many researchers did studies in order to solve this problem in their context. Mathematical methods such as gravity model, entropy maximization, and multiple linear regression analysis are used for these studies.

In Sri Lankan context most of the traffic studies are conducted by the Department of Civil Engineering of University of Moratuwa. According to the sources available from Civil Department complete OD survey has been not conducted to the entire Sri Lanka.

1.3 Research Objective

The main objective is to develop a method to complete partially filled origin destination matrix in Sri Lankan context. Traditional methods of estimating origin destination matrix are obsolete in these days. Filling of some cells in O-D matrix can be done with traditional methods.

Therefore it is essential to find a method that can be used to fill the partly filled OD matrix in related to the Sri Lankan context. Many of researches have been done in order to find a method to fill the partly filled OD matrix in worldwide. However those are depends on other several factors such as geographical condition, economical condition and also thinking pattern of the people. Therefore it is mandatory to find a method to fill the partly filled OD matrix in Sri Lankan context since it is essential for transport planning.

If there is a way, then the remaining blank cells in the OD matrix can be filled without conducting surveys.

1.4 Sub Objectives

In order to complete this research it is necessary to set up other sub objectives as follows.

- To review existing methods and techniques for origin destination matrix estimation
- To select most suitable methods and techniques for origin destination matrix estimation for the Sri Lankan context
- To collect and estimate relevant socio economic data for developing equations
- To develop equations for trip production, trip attraction and trip interchange

- To validate the above equations using existing data

1.5 Frame Work

Origin Destination matrix includes every trip origins and destination within the closed system. These trip generation depends on the several factors of particular environment. That is varied depends on the country. A method calculated for a particular country is not compatible for other countries. Therefore there is a requirement to find a method for the Sri Lankan context. In order to do that it is necessary to find the factors by which has most dominate to the trip generation.

These factors are selected based on the assumptions. Some of these factors may be interrelated.

Surveyed data can be used for find out the value of the factors which are mentioned above. These values will be added to mathematical model and suppose to obtain best fit equation to use fill the black cells in partially filled OD matrix.

1.5.1 **Study Area**

There are 331 divisional secretariat division in all around the Sri Lanka. When we do the traffic planning for a particular area it is needed to get knowledge about the traffic generation from that area to the other whole area of the country and vice versa. Therefore, in order to get better result it is mandatory to select total number of DS divisions as a study area.

Therefore it was decided to consider traffic data from all 331 divisional secretariat divisions to conduct my research study.

1.5.2 **Limitation to the Study**

This research is focused on the estimation of partially filled origin destination matrix. Throughout this research, it identifies and evaluates the different modelling approaches which are used to estimate the OD matrix within the world. Ultimately, this research will find out the method which can be used to fill partially filled origin destination matrix.

In order to achieve that research aim and objectives are formulated. We can identified some limitations and shortcomings within this research process especially with the collection of data. Because of that used partially filled OD matrix to get the final outcome of this research is completed with the survey data which conducted in traditional ways. Therefore, sometimes those data is not fully accurate.

Also we have taken some assumptions such as dependent factors for trip generation and trip attractions. Even within the Sri Lanka also these factors can be changed in place to place. However we assume that those dependent factors are same in every place within the country. Therefore, there can be different types of dependent factors in place to place in the Sri Lanka.

As a result this will slightly affected to the final result of the study. However, I will try to get maximum number of dependent factors to my study thereby I will try to give maximum equity for my study.

1.6 Research Methodology

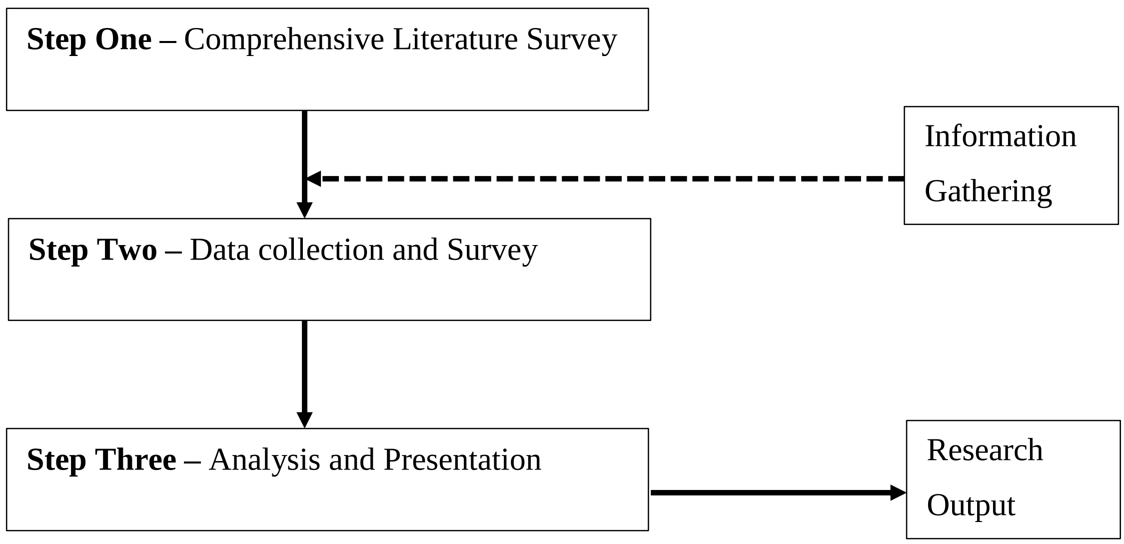


Figure 1.1

1.7 Structure of the Thesis

The Thesis was divided in to five main chapters.

1.7.1 Chapter one - Introduction

This chapter comprises the background of the problem, definition of the research problem, research objective, research questions, methodology and technical framework, data collection and sampling.

1.7.2 Chapter two – Literature Review

In this chapter, literature review explain what is origin destination survey, what is origin destination matrix and its importance in the transportation planning. Also this chapter describes the problems encounters with the estimation of OD matrix. Then it describes the methods which are used in international context as well as Sri Lankan context. After critically evaluation of the exiting method with the literature review, this chapter explains the advantages and disadvantages of current practices of the OD matrix estimation. Finally this chapter shows comparatively best model from the existing model which can be used in Sri Lankan context.

1.7.3 Chapter three – Research Design

This chapter explains details of methodology that are used in the research and its each stage. Also it is discussed that what are the data collection methods and what are the data analysis methods.

1.7.4 Chapter four – Analysis

This chapter comprises of an analysis that is used to continue this research, how the selected methods evaluates with the completed OD matrix by using Entropy Maximization and with the multiple linear regression. Also analysis will be done by soft wares such as Math labs. Finally, it shows the findings of the research.

1.7.5 Chapter Five – Conclusion

This chapter is the last chapter of this document which presents the conclusion of the main findings and the recommendations. Also it includes the study area of future research to be carried out.

1.8 Summary

The objective of this chapter is to provide an overview of background of the problem which focused and the research is about what and outline of the whole dissertation. On overview, the background of the problem was discussed in details while including the research questions of the study. The theoretical framework and research methodology were stated briefly. Villages are transformed into cities. Therefore transportation pattern will be changed. Therefore there is a close relationship between land use and transportation. Hence OD analysis will be very useful in order to overcome spatial issues pertaining to land use activities.

In next, Chapter two will study the literature about the method to do the OD estimation in worldwide and their merits and demerits.

CHAPTER 02

2 LITERATURE REVIEW

2.1 Introduction

Literature review will be helpful to find out all important elements which are used in many other research studies in relation to this research study. Research findings and suggestions on similar techniques those are used in many other countries and regions can be found by this literature review. Expectation on this literature review is to attain suitable measurements and eligible factors which are match to Sri Lankan context. Also it is expected to verify the applicability of the model which is selected by this research work. Further to that, some techniques which are combined to particular contextual factors will be examined in this literature review.

2.1.1 Origin- Destination Survey

Transportation professionals used Origin –Destination (OD) matrices as a most valuable tool. OD studies are conducted to understand the pattern of the movement of persons and goods in a particular area of interest during a particular period of time (Wang & S, 1997)

2.1.2 Origin- Destination Matrix

In transport analysis, OD matrix is an important study. Origin destination matrix is a two- dimensional table. Rows and columns of that table shows the trip attraction and trip production from one zone to other zone. Therefore we can get the information of the number of travelers who are travel between different zones of a region. (Kikuchi & Kronprasert, 2009)

2.2 Importance of the OD Matrix

It is necessary to know the correct number of trip generation and trip attraction from each zone to other zone in transport planning. According to the 2011 TRANS O-D Survey Report, it has mentioned that these OD surveys collect valuable data related to households, individuals and trips. On the other hand, information from the OD matrix gives the knowledge of the travel pattern and the characteristics of the travels. Also it could be identified the trends of the transportation in the future, identification of the transport model for a particular region. Further to that, we can input those data to the travel demand model development. Travel forecasting, planning and need of transportation infrastructure, monitoring of implemented transportation policies can be identified by the information of the OD matrix. It contain information about the spatial and temporal distribution of activities between different traffic zones in a determined study area. (Tornero, Martínez, & Castello)
Other than the above mentioned uses according to the Willumsen, L.G. (1978) Estimation of an OD Matrix from Traffic Counts – A Review, it mentioned number of uses such as,

- Transport demand modelling in towns
- Traffic management schemes design and assessment modeling in urban and rural areas
- Previous OD matrices updating
- Inter urban and rural transport demand modeling in sparsely populated areas.
- When other data is unavailable, unreliable or obsolete in most of the developing countries we can model transport demand with the information of OD matrices.

As per the traffic Engineering and transport planning, Dr. L.R.Kadiyali, specific uses of origin destination matrices can be listed down as follows.

i. In order to identify the requirement of the bypass, it is necessary to find the amount of by-passable travelers that enters a town.
ii. In transportation planning process, trip distribution models and trip generation models can be developed
iii. Requirement of the new highway system to the country and prediction of the new transportation infrastructure
iv. To provide the adequate parking facility and the plan for future.

Origin destination matrices can be estimated by the traditional techniques such as Home interviews, Motorists interview surveys, questionnaire survey by post cards, vehicle number plate surveys, surveys by telephone, GPS receiver, and Mail survey are used.

According to the RDA officer in charge of OD estimation, he mentioned that in Sri Lankan context, ODM estimate with the mostly used Home interview survey and Roadside interview survey.

While ODM estimation is one of the important requirement, there are various problems have arisen with adapted methods.

2.3 Problems of existing OD matrix estimation techniques

1. Counting location and Sample size problem

When the number of counting location increases accuracy of the OD estimation will be increased. (Bera & Krishna Rao, 2011). To conduct a complete or accurate OD survey in an entire city or region, number of counting locations and sample size with respect to the population in a particular region should be considered. Number of counting locations are decided on basis of the sample size. Therefore, when the sample size is being increased, also increase the number of counting locations. In manually conducting, an OD survey in more counting location is not practical because it consuming more time and cost. Therefore, with the resource limitation it may not be possible to conducting manual OD survey. With the counting location problem, it may cause to occur double counting errors and affected to the accuracy reduction of the dataset.

2. **Cost and Time**

Conducting many OD surveys in many counting locations with the high amount of sample size is consuming more time and high cost (Kikuchi & Kronprasert, 2009). In reality conducting OD surveys consumes millions of money. Therefore cost and time become as factor of reduction its accuracy.

3. **Reliability and accuracy**

Accuracy and reliability of the OD estimation is depend on the optimum traffic counting locations (Bera & Krishna Rao, 2011). The traffic counts collected provide much traffic information as possible. Data collection procedure of the OD estimation required manpower (Gan, Yang, & Wong, 2005). Resource limitation is also affected to the level of accuracy. Because of the fundamental issue of the OD surveys directly affected to the reliability and accuracy levels of the survey.

4. **Single time period problem**

These estimated OD values are static. Also it can be apply to only one observation time period. However theses estimations are required for a series of linked dynamic time periods. (Aerde, Rakha, & Paramahamsan, 2003)

2.4 Existing methods of ODM Estimation

2.4.1 Gravity model

Gravity model can be identified as one of the static Origin Destination Matrix estimation model. This model is used to explain the most of the travel behavior in the study area. (Willurnsen, 1978) According to the Traffic Engineering and Transport

Planning, this model explains that the interchange of trips between two zones is depends on the trip production of the particular zone, trip attraction of the particular zone and the distance between those two zones

General equation of the gravity model,

$$T_{ij} = \frac{KP_iA_j}{d_{ij}^n}$$

T_{ij} = Number of trips between zones i and j

P_i = Number of trips produced in zone i

A_j = Number of trips attracted to zone j

d_{ij} = Distance or time or cost of travelling between zone i and zone j

K = Constant which is independent from i and j

n = An exponential constant and value of this usually found between 1 and 3

The gravity models represent the travel costs occurred between two zones into the demand matrix estimation as per the Origin Destination Transportation Models.
According to the Origin Destination Transportation Models & Methods, it mentioned the gravity models introduce all travel costs into the estimation of the demand matrix. It shows the macroscopic relationship between places. So according to the notation of Origin Destination Transportation Models: Methods journal article, it mentioned,

It can be shown that the number of trips (T_{ij}) between zone i (Origin City) and zone j (Destination City) is proportional to the number people leaving from i zone (O_i) and number of people reaching to zone j (D_j). Also it is inversely proportional to the square of the (generalized) cost C_{ij} which is bear to travel between i and j

$$Tij = \propto \frac{OiDj}{C_{ij^2}}$$

In practical studies, formulation that is more flexible is used based on a deterrence function.

$$Tij = \propto OiDjf(Cij)$$

This deterrence function has calibration constant and with the good calibration accuracy of the model will be more.

However, there can be some limitation with the Gravity model when it is applying.

The gravity model compromises the sensitivity to the transport networks and the sensitivity to the levels of trip making ‘from’ and ‘to’ each zone. However, sometimes it can be act as unstable model and as a result sometimes resulting trip matrix may be quite different to the observed trip matrix even when it is calibrated. Therefore, it is mandatory to check the matrix produced against the observed matrix at a sector to sector level in order to verify that the model is working correctly. (Gupta & Shah, 2012 May)

According to the origin destination matrix estimation from traffic counts: the state of the art it mentioned that, some researchers like Robillard (1975), Hogberg (1976) used the Gravity model based approaches and some others (Tamin and Willumsen, 1989; Tamin et al, 2003) used Gravity-Opportunity (GO) based models for estimating ODM. These techniques require zonal data for calibrating the parameters of the demand models. The main drawback of the gravity model is that it cannot handle external-external trips (refer Willumsen, 1981)

These gravity type models misspecification errors can be occurred when use of a model for medium term planning according to the Estimation of an O-D Matrix from Traffic Counts – A Review report mentioned.

In inter urban areas most applications have been used in the gravity model as overall and this method can be used as suitable method for problems where each town represents a single zone. As well as this approach is not suitable for inner city areas because it involved short trip lengths.

Gravity type of trip making behavior is not likely to be a practical assumption (Willurnsen, 1978) In addition to that gravity models are basically depend on only the distance, employment and population parameters.

2.4.2 Entropy Maximization

The concept of Entropy is commonly used with the information theory. The transmission, processing, utilization and extraction of information are studied by the information theory. At the same time it explains the uncertainty should be manipulated, represented and qualified. Most fundamental quantify in information theory is entropy. The amount of uncertainty of unknown or random quantity is measured by the entropy in the information theory.

Entropy is a measure of randomness and also entropy can be identified as the average amount of the information from the event. The entropy concept is used to measure the uncertainty in a probability distribution. If someone have high uncertainty, there is a high entropy. When the uncertainty is low then there is a low entropy. Willumsen (1978) uses entropy maximization approach. The entropy-maximizing procedure analyzes the available information to obtain a unique probability distribution (Bera & Krishna Rao, 2011).

Concept of maximum disorder, the entropy maximization procedure finds the most likely configuration of elements within a constrained situation.

Estimation of the origin destination matrix can be derived from the entropy maximization concept in transport planning and it can be used as the model building tool. The model of entropy maximization can be identified as the one of the O-D matrix estimation techniques being commonly used in international context. This approach can be used in order to estimation of trip matrix and results in multi proportional problem. Entropy maximization is important because it minimize the commitments and model all that is known and assume nothing about what is unknown. It means using prior data and predict the unknown elements by seen the limited number of known things. It is linked with the probability distribution.

The concept of entropy is used to measure the uncertainty in a probability distribution. If someone have high uncertainty, there is a high entropy. When the uncertainty is low then there is a low entropy. Willumsen (1978) uses entropy maximization approach.

Another way it can be expressed as when number of unknowns is more than the number of available information is not sufficient to obtain the unique set of solutions. The approach will be determined the safest solution in the situation of many possible sets of solutions are exist. Then this approach will be called as maximum uncertainty or maximum entropy. It is dealt with uncertainty.

Entropy maximization uses minimum information and it reduces the commitment.

Entropy equation,

$$H = - \sum_{i=1}^{n} P_i \; \log_b(P_i)$$

H= Entropy Value

P_i=

n= number of different outcomes.

Entropy equation is derived,

$$I = -\sum_{i=1}^{n} (\mathrm{N} \times \mathrm{Pi}) \times \log_b(\mathrm{pi})$$

I = Total information from N occurrences

N = Number of occurrences

(N×Pi) = Approximated number that the certain result will come out in N occurrence.

Concept of entropy used set of parameters to estimate matrix such as Probability of travel from zone i and zone j, Total number of trips, Population, Total number of trip origins at I, Total number of trip destinations at j, human Choice behavior (Work, Adequate habitable place to live, Rent, Access to school, Hospitals, Gas connection, Local shopping facilities)

2.4.3 Growth factor method

This simple method updates an old OD matrix t using the actual number Oi of trips originating in each zone, and/or the actual number Dj of trips terminating in each zone. (Gupta & Shah, May 2012). This growth factor methods are developed with the assumptions of present travel patterns can be projected to the design year in the future. This will be created using certain growth factors. This model will be used with the number of trip produced and number of trip attracted by the certain zones as the input data.

There is an article called Trip Distribution Analysis at Urban Transportation Planning and as per that article, growth factor method will be used trip distribution matrix in the

base year as input data. In order to do so, we need to collect detailed information of the origin destination data. Also it is required to prepare a matrix with the data of trip distribution which are taken from the growth factor method.
As per the literature review, it is identified that there are many other growth factor methods which can be used in the transport modelling as follows.

- Uniform growth factor method
- Average growth factor method
- Frater Method
- Furness Method

- Uniform Growth Factor Method

This method is developed based on the assumption that the growth rate for the whole area is valid in order to predicting inter-zonal trips in the future. It is assumed that there is the uniform growth for the whole of the urban area is proportionally equal to the uniform rate for each of the traffic zones in the urban area. We used that assumption when we work on uniform factor method of trip distribution. (Prof. Dr Arasan)
If we have only available information is the general growth rate for the whole study area, then we can apply that uniform growth rate to each cell in the OD matrix with assumption that growth rate is uniform in each and every zones in the study area. (NPTEL, 2007).

$$Tij = t_{ij} \times E$$

T_{ij}= number of trips occurred between zone i and zone j in the design year
t_{ij}= number of trips occurred between zone i and zone j in the observed based year
E = Growth factor

- Average Growth Factor Method

In the Average Growth Factor Method, growth factor of a particular zone is calculated based on the average of growth factors calculated for both start and end of the trip at the origin and destination. This factor shows the average growth associated both with the origin and destination zones (L.R.Kadiyali, 2007)

$$T_{ij} = t_{ij}\left[\frac{E_j + E_i}{2}\right]$$

Tij = Number of trips from zone i to zone j in the future

t_{ij} = Number of trips from zone i to zone j in the present

E_i = Growth factor in zone i

E_j = Growth factor in zone j

- Fratar method

In this Fratar Method, the total number of trips for each and every zones will be distributed based on inter zonal movements as a first aproximation as per the relative attractiveness of each movement. After that, future trips will be estimated for a particular zone based on the movements happening in that zone in proportion to the existing trips between in that zone and other zone. Also it will proportion to the expected growth of each other zone.

$$T_{ij} = t_{ij} \times \frac{P_i}{p_i} \times \frac{A_j}{a_j} \times \frac{\sum^k t_{i-k}}{\sum^k \left[\frac{A_k}{a_k}\right] t_{i-k}}$$

T_{ij} = Number of trips from zone i to zone j in future

t_{ij} = Number of trips from zone i to zone j in present

P_i = Number of trips produced at zone i in future

p_i = Number of trips produced to zone j in present

A_i = Number of trips attracted to zone j in future

a_j = Number of trips attracted to zone j in present

K= Total number of zones

- Furness method

This Furness method estimates the future traffic at the originating zone and the terminating based on the origin growth factors and destination growth factors for each zone. In this method the traffic movements will be made to agree alternatively with the estimated future originated traffic in each zone and the estimated future terminated traffic in each zone, until both these two conditions are roughly satisfied. (L.R.Kadiyali, 2007)

Following advantages and disadvantages are relevant to the growth factor method.

Advantages

- This method is simple to understand.
- This method can be used to small areas.
- In this method observed trip pattern are preserved.
- This method is very useful in short term-planning.

Disadvantages

- In order to estimate the smaller zone to zone movements accurately, it is necessary to develop a present trip distribution matrix. Therefore large scale OD studies with large sampling sizes are required. It will be costly and time consuming.
- If there is an error in a specific zone. Then it will be magnified in zone to zone movements.
- There are no methods to evaluate the resistance to travel and it implies that resistance to travel will remain constant. This method neglects the effect of

changes in travel pattern which will create by the new transport infrastructure and creation of new transport network. (L.R.Kadiyali, 2007)

- Particularly weak in projecting small volumes of current traffic.
- Result of this method is highly depends on the observed trip pattern
- Unobserved trip pattern cannot be estimated by this method
- No consideration about the changes of travel cost
- This method is not suitable for policy studies such as introduction of a mode.

2.4.4 Intervening Opportunity Model

Opportunity models are based on the statistical theory of probability as the theoretical foundation (L.R.Kadiyali, 2007). This model is same as gravity model and it is used to distribute known or estimated attractions and productions and this model is a distribution model (Kaltenbach, 1972). After given the productions of all origin zones and the attractions of all recreation area, the model computes trip interchanges.
The concept of intervening opportunity model is the number of trips from an origin zone to a destination zone is directly proportional to the number of opportunities at the destination zone and inversely proportional to the number of intervening opportunities (Zhao, Chow, Min-Tang, Gan, & Shen, 2001 September). The assumption of this model is every time traveler use a shortest path to complete his trip. However, if the shortest trip does not satisfy his needs, the traveler must consider more distance destinations.

The model is based on probability and set of theories of transportation modelling which can be applied only to urban travel modelling. That is the probability that a particular trip from an origin zone will find a destination area is equal to the probability that an acceptable destination exists there times the probability that an acceptable destination has not been found elsewhere (Kaltenbach, 1972)
The model can be derived as,

$$T_{ij} = P_i \left[\frac{exp(-LV(j-1)) - exp(-LV(j))}{1 - \exp(-LV(J))}\right]$$

T_{ij} = Trips from zone i to zone j;

P_i = Trip produced in origin zone i;

L = the probability of accepting a destination opportunity;

$V_{(J)}$ = the total destination opportunities in all j destinations;

$V_{(j)}$= the total destination opportunities from origin zone i to the j^{th} ranked destination;

$V_{(j-1)}$= the total destination opportunities from origin zone i to the $(j\text{-}1)^{th}$ ranked destination.

The gravity model distributes trips with respect to distance or some other factors. However, in intervening opportunity model distributes trips according to a measure of opportunities. This method used to estimate outdoor recreational trip productions. When comparing the other models this model has less accuracy and it consumes high cost.

According to the Zhao, Chow, Min-Tang Li, Gan, & Shen, in their research report they have been mentioned that this model possibly not use in practice, due to following reasons,

- This is difficult to understand by practitioners since the theoretical base of this method is less well known.
- It is difficult to maintain the distance from the origin in practically. i.e. the n^{th} cell for origin i cell is not the destination, however the n^{th} destination away from i cell.
- The theoretical advantages and practical advantages between opportunity models and gravity model are not very much significant. As a result opportunity model will be replaced by gravity model.
- Inter zonal trips are underestimated by this model and as a result less number of intra zonal trips were gained by the intervening opportunity model.
- Inaccuracy can be happened in the attraction estimation.

2.4.5 Maximum Likelihood estimation

Likelihood can be identified as statistical approach. Also this is one of the oldest method and most important in estimation theory. Observed samples can be generated very often by the set of parameters from the Maximum Likelihood estimation. The likelihood of observing the target origin destination matrix and the observed traffic counts conditional on the true trip matrix are maximized by the Maximum Likelihood estimation in the ODM estimation. (Bera & Krishna Rao, 2011)

Probability of observing both the additional sampling survey results and the counted flows are maximized to create Maximum likelihood (ML) estimator. These two probabilities are independent in the maximum likelihood estimator under the usually acceptable assumptions. (Bera & Krishna Rao, 2011)

There are no any distribution assumptions in this method and that can be identified as and advantage of this method

2.4.6 Generalized Least Square

GLS is based on the minimization of the distances from an priori estimate t to the new OD matrix T. (Gupta & Shah, May 2012) It is mentioned that this method is combined with the model prediction, traffic counts and target trip matrix within a single framework as per the state of art. According to Bell, errors between the values to be estimated and the observed values are minimized by using Generalized Least Square subject to a set of constraints.

Considering to the prediction of the true trip matrix using a toy-network scenario, this method can be identified as outperform the Maximum Entropy Maximization approach.

2.4.7 Bayesian Inference

Bayesian Inference is valid for origin destination estimation problems and for estimating turning flows at intersections as per the state of art article. Apart from the priori trip matrix it can be potentially balanced the information of link traffic counts with many other sources of information. When the least prior estimated are equally distributed among OD pairs, this method extends the Maximum Entropy/Minimum Information criterion being equal.

Advantages

- The value of the various sources of information neatly distributed well
- proportional assignment approach

Disadvantages

- The relationship between OD flows and link flows should not be linear.
- Congestion phenomena and route choice are strongly correlated.

2.5 Summary

In this literature review, it was found that there are some methods to fill the partially filled origin destination matrix in other countries across the world. The researcher found that those methods depends on environment and other factors of the relevant countries. They have finalized with different equations relevant to conditions of their countries.

Also it is noted that most of the methods are derived based on the assumptions in relating to their context. Therefore, it is required to take correct assumptions in order to make best method to fill partially filled origin destination matrix in Sri Lankan context.

The literature shows that there is need to find this type of methods in worldwide in order to develop their business and transportation planning.

Model	**Description**	**Advantages**	**Limitations**
Gravity Model	• Gravity model is capable of explaining most of the travel behavior in the study area • It has suggested that the interaction between two locations declines with increasing (Distance, time and cost) between them	• Gravity model most applications have been used in inter urban areas • This method is suitable for problems where each town represents a single zone • Ensure more practical zone	• Only consider about the distance factor • Required zonal data for calibrating the parameters of the demand models • When calibrated the resulting trip matrix may be quite different from the observed trip matrix from which it was derived

			• It cannot handle with accuracy external- external trips • Require more data and have to make greater effort to calibration • Misspecification errors disqualify the use of a model for medium term planning • This approach is not suitable for inner city areas
Entropy Maximization Approach	• Model all that is known and assume nothing about what is unknown • Using prior data • Entropy equation, $H = -\sum_{i=1}^{n} P_i \ \log_b(P_i)$ H = Entropy Value P_i = n = Number of different outcomes	• Used as model building tool in transportation • Minimize the commitments and model all that is known and assume nothing about what is unknown	• Based on only probability value

Growth factor method	**Uniform growth factor method** • It assumes that the growth rate for the whole area is valid for predicting future inter-zonal trips $Tij = t_{ij} \times E$ T_{ij}= number of trips occurred between zone i and zone j in the design year t_{ij}= number of trips occurred between zone i and zone j in the observed based year E = aGrowth factor **Average growth factor method** • Growth factor for each zone is calculated based in the average of growth factors calculated for both ends of the trip • This factor represents the average growth associated both with the origin and destination zones.	• Simple to understand • Preserve observed trip pattern • Useful in short term planning • Easy to balance origin and destination trips at any zone	• Present trip distribution matrix has to be obtained first ,for which large scale OD studies with high sampling size are neede3d to estimate the smaller zone to zone movements accurately • The error in origin data collected on specific zone to zone movements are magnified • None of the methods provides a measure of the resistance to travel and all imply that resistance to travel will remain constant. They neglect the effect of changes in travel by the construction of new facilities and new network. • Depends heavily on the observed trip pattern • It cannot explain unobserved trips

	$$T_{ij} = t_{ij}\left[\frac{E_j + E_i}{2}\right]$$ Tij = Number of trips from zone i to zone j in the future t_{ij} = Number of trips from zone i to zone j in the present E_i = Growth factor in zone i E_j = Growth factor in zone j **Fratar Method** • The future trips estimated for any zone would be distributed to the movements involving that zone in proportion to the exiting trips between it and each other zone and in proportion to the expected growth of each other zone		• Do not consider changes in travel cost • Not suitable for policy studies line introduction of a mode • Does not reflect changes in the frictions between zones • Does not reflect changes in the network

	$T_{ij} = t_{ij} \times \frac{P_i}{p_i} \times \frac{A_j}{a_j} \times \frac{\sum^k t_{i-k}}{\sum^k \left[\frac{A_k}{a_k}\right] t_{i-k}}$ T_{ij} = Number of trips from zone i to zone j in future t_{ij} = Number of trips from zone i to zone j in present P_i = Number of trips produced at zone i in future p_i = Number of trips produced to zone j in present A_i = Number of trips attracted to zone j in future a_j = Number of trips attracted to zone j in present K= Total number of zones		

Intervening Opportunity Model	• Model concept is the number of trips from an origin zone to a destination zone is directly proportional to the number of opportunities at the destination zone and inversely proportional to the number of intervening opportunities • Based on the probability equation $T_{ij} = P_i \left[\frac{exp\left(-LV(j-1)\right) - exp\left(-LV(j)\right)}{1-\exp\left(-LV(J)\right)}\right]$ T_{ij} = Trips from azone i to zone j; P_i = Trip produced in aorigin zone i; L = the probability of accepting a destination opportunity;	• Used as alternative model of Gravity Model	• The theoretical base is less well known and probably more difficult to understand by practitioners. • The idea of matrices with destination ranked by distance from the origin is more difficult to handle in practice • The theoretical and practical advantages of opportunities models over gravity models are not so significant as to warrant their replacing gravity models • There is a lack of suitable software to calibrate and use them • These models are underestimated the intra zonal trips and much smaller number of intra zonal trips were gained from the intervening opportunity model

	$V_{(J)}$ = the total destination opportunities in all j destinations; $V_{(j)}$= the total destination opportunities from origin zone i to the j^{th} ranked destination; $V_{(j-1)}$= the total destination opportunities from origin zone i to the $(j-1)^{th}$ ranked destination.		• Inaccuracy may be occur in the estimation of attraction.

Problems and the methodology of the study carried out will be discussed in the Chapter three.

CHAPTER 03

3 METHODOLOGY

3.1 Introduction

This chapter presents the designing of structured questionnaires survey and theoretical approach of using a statistical tool for analysis of the data. In order to develop a method to fill partially filled OD matrix, it is required to obtain type of sources to make trip generation and trip attraction. Therefore primary data will be collected from the various organization such as department of Civil in university of Moratuwa, Road Development Authority and Statistics department by conducting surveys.

This research will focused to entire country. Therefore the researcher has targeted to obtain various information from whole 331 divisional secretariats in Sri Lanka.

Also this chapter describes how to collect the data and analysis the data. Further to that this chapter describes different types of data analysis techniques.

The analysis will be done considering the all the information and assumptions which are taken to this study are accurate. Therefore produced output is subject to the accuracy of the data collection.

Available information from the various organizations which are mentioned in earlier are obtained from the various survey techniques. Therefore we have to accept the level of accuracy of those survey technique during this research.

3.2 Methodology Flow

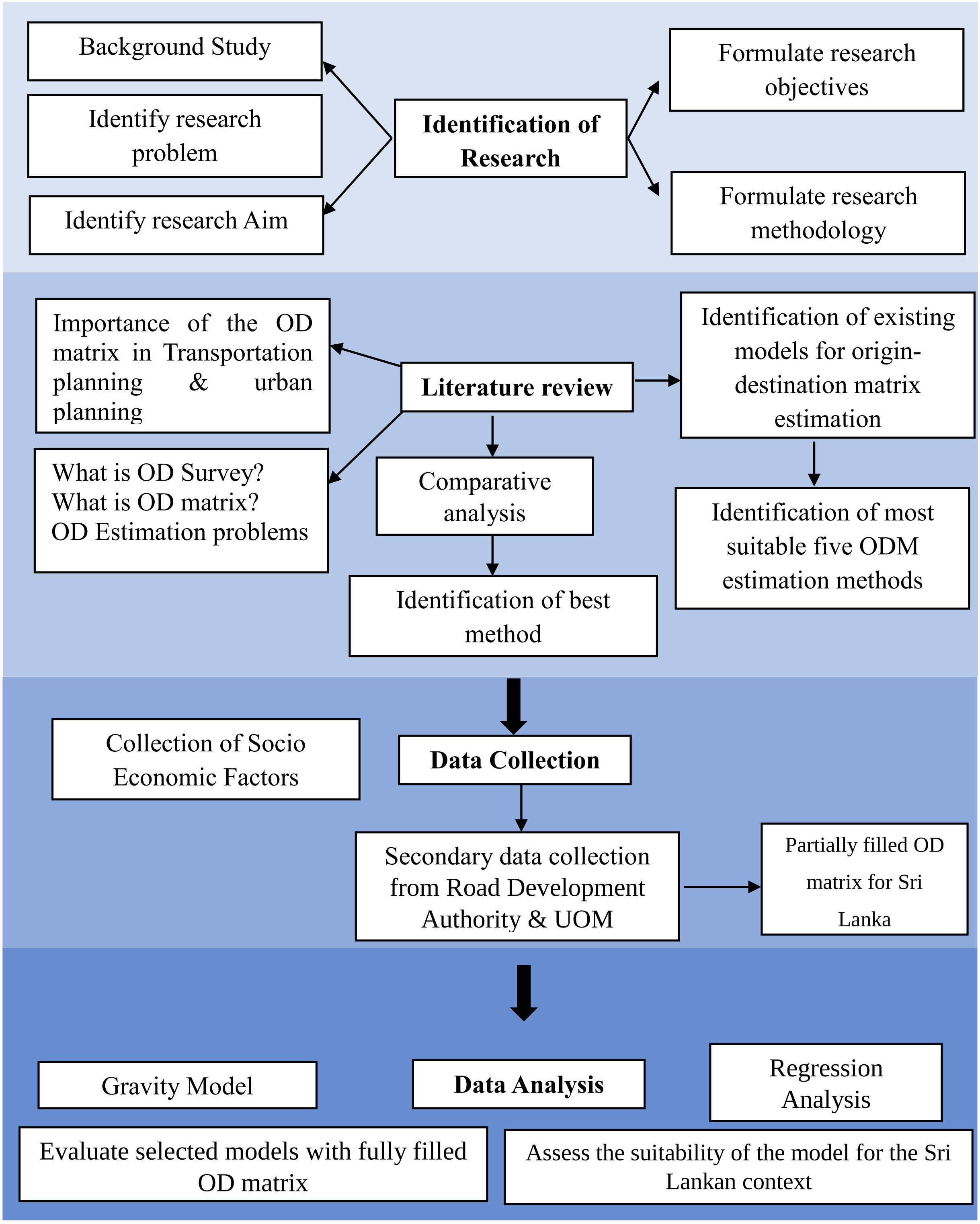


Figure 3.1

3.3 Research Statement

There are number of ways to conducting the origin destination matrix for a particular area. Currently origin destination surveys have been conducted by the researchers who are doing traffic related studies. As per the study they can select traffic analyzing zone as they wish.

Civil Department of University of Moratuwa has conducted Origin Destination matrix as taking districts as traffic analyzing zones. There are 625 cells in that origin destination matrix. It was noticed that at least 62 OD surveys has to be done in order to fill that origin destination matrix. Therefore it is noted that approximately 10% of the OD pairs to be done as OD surveys in order to fill that particular OD matrix.

When we consider about the macro level analysis for the entire Sri Lanka, it is required to consider divisional secretariat division as traffic analyzing zone. When we consider divisional secretariat as traffic analyzing zone, OD matrix comprise of 109,561 cells. Therefore according to the previous example 10,956 OD surveys to be done in order to fill the said OD matrix. Generally it is difficult to conduct more than 10,000 practically. Conducting more than 10,000 OD surveys is not practicable due to time and financial constraints. Therefore it is must to develop a method to fill partially filled origin destination matrix for Sri Lankan context.

3.4 Research Objectives

There are 331 divisional secretariat within the Sri Lanka. One divisional secretariat can be identified as one Traffic Analyzing Zone (TAZ). Therefore there are 331 TAZs. The scope of the present work is confined to prediction of future trip interchanges using several models for the entire Sri Lanka taking the DS divisions as TAZs. Therefore following objectives shall be derived.

- Identify the existing methods and techniques for origin destination estimation by reviewing index journals in order to identify the trip production and trip attraction of the each TAZs.

- Collection of relevant socio economic data with regards to the each TAZs
- Find the best fit equation model for develop trip production and trip attraction separately for each TAZs
- Development of a trip interchange model based on the gravity model

There are 331 district secretariat divisions within the Sri Lanka. Since four main objectives can be derived in this research study is to determine a method to fill partially filled origin destination matrix in Sri Lankan context.

The objectives of the research can be listed down as follows;

1. To review existing methods and techniques for origin destination matrix estimation
2. To select most suitable methods and techniques for origin destination matrix estimation for the Sri Lankan context
3. To collect and estimate relevant socio economic data for developing equations
4. To develop equations for trip production, trip attraction and trip interchange
5. To validate the above equations using existing data

3.5 Theoretical Framework

Research framework was developed with reading several previous studies done by the researchers about the origin destination matrices. Different approaches are used different methodologies in respective of their countries and other environment aspects. Most of those approaches are used several mathematical techniques such as entropy maximization, T method, Gravity model, Multiple Linear Regression Analysis etc. Then, comparative analysis is done in order to identify the advantages and disadvantages of each techniques used by several other researchers.

Content Analysis is done for identify the parameters used for several other techniques in OD matrix. With the literature review it can be identified that what type of parameters are mostly used in the similar techniques in worldwide.

3.6 Study Area

For a particular area attracts trips from all other area of the country. Also trips can be generated to all other area of the country from that particular area. Therefore it is necessary to study whole country in order to obtain better result from this research study. There are 331 number of divisional secretariats within Sri Lanka. We can take one divisional secretariat as a particular region of the trip matrix. Therefore data and information from all 331 divisional secretariat are obtained for this research analysis.

3.7 Limitation to the Study

In any research study accuracy level is depends on the accuracy level of the input data. Complete origin destination matrix has been obtained for the analysis of this research study. That complete OD matrix is obtained from Road Development Authority or Civil Department of University of Moratuwa. That matrix is completed with the traditional survey methods. Therefore accuracy of the said matrix is questionable.
Also during the research study there are parameters and assumptions to be considered. Accuracy of those parameters and assumptions can be changed according to the area even within the same country. Also available partially filled origin destination matrix is not included with transportation by train. That also can be considered as limitation of this study.

3.8 Research Methodology

As a first step review of literature to be done. In order to do this comprehensive review, publications such as Journals, Magazines, Books, Reports and News Papers were examined. In addition, internet facilities were used and many relevant literature were downloaded. It can be studied about methodologies and techniques which are used in several other countries in order to solve this type of problems in origin destination matrix. It can be identified the advantages and disadvantages of the existing models in the worldwide and their suitability for the Sri Lankan context.

Secondly, completed origin destination matrix is collected from Civil Department of University of Moratuwa. Also socio economic parameters are collected from several other organizations such as Land Use & Policy Planning Department and Valuation Department etc.

Then data analysis is started. In order to do the data analysis following mathematical techniques are used.

a. Bivariate Correlation Analysis
b. Multiple Linear Regression Analysis
c. T –Testing
d. Gravity Model
e.

First of all ten socio economic factors will be selected for the analysis and values of those socio economic factors were collected for 331 divisional secretariat divisions. Then we got the trip production values for some divisional secretariat from the partially filled origin destination matrix. After that correlation analysis will be done for the trip production and values of socio economic factors its other mathematical factors such as log value, ln value, sin value and cos value etc. With that analysis we can select the best five socio economic factors which are mostly link to the trip production. Then we can use those five socio economic factor for develop model for trip production. It can be developed 31 different models using those five socio economic factors. Then part of the known trip production values are taken to develop the trip production model. In order to develop trip production model SPSS software will be used. R value, R squared value, adjusted R squared value and standard error of the estimate will be calculated through the SPSS software.

Then the derived model with highest R value and R squared value is selected as the model equation for trip production.

Same process described as above will be followed for develop a model for trip attraction.

Then after developing models for trip production and trip attraction, validation to be done for both derived models. In order to do that we will use overall model fit and goodness of fit. Also t test to be done as a validation process. With the results of the validation process it can be confirm the derived model equation for trip production and trip attraction.

Then it can be filled trip production values and trip attraction values for 331 divisional secretariats. Then next task in to fill trip interchange values of partially filled origin destination matrix. In order to do that gravity model will be used. As per the available partially filled origin destination matrix 37% of the trip interchange values are filled. Therefore balance 63% trip interchange values to be filled.

In this analysis, it can be ascertained that the level of influence to the OD matrix from the selected parameters. So, correlation between these parameters is developed. Also from these parameters it can be identified dependent and independent variables.

Finally, the equation can be derived and it will be tested with known data in order to validation.

3.9 Research Design

This research was designed with the mathematical analysis techniques such as Bivariate Correlation Analysis, Multiple linear regression analysis, T Test and Entropy maximization. Also parameters are selected according to facts collected from literature review and personal experience. This research approach used is based on a mix of qualitative and quantitative methods with higher portion of quantitative methods. In this type of studies researcher has to depend exclusively on mathematical equations and their outputs.

However during the research study it is needed to identify the correct independent variables and dependent variables. Since primary data are collected through the traditional survey methods we have to do some mathematical techniques to obtain better results from this research study. This research is done with the completed OD matrix which obtained from Civil Department of University of Moratuwa.

3.10 Sample Selection

The sampling selection has been carried out by focusing limitations of conducting OD survey manually in the entire country. In practically OD survey can be done for complete limited number of cells. It is possible to fill about 40% - 50% percentage of cells. Therefore, this study is to identify, using partly filled OD matrix, can it used to fill other cells in the matrix. To examine that situation half of the data have removed from the actual OD matrix. Completed 60 sample are selected on the basis of highest OD demand for each zones.

CHAPTER 04

4 ANALYSIS AND DATA COLLECTION

4.1 Introduction

The main objective of this chapter is to analyze the collected data and the present with the findings. This chapter consists of all the analysis which are done throughout the research. This chapter describes in each stage flow to analyze the data and their ultimate results (findings). Application of the multiple linear regression is one of the important analysis techniques used in this research. Also gravity model is used for the finding of this study. As well, this interprets its results way to derived research output with the developed formula and method to estimate partly filled OD matric with reference to the developed formula.

4.2 Study Area

For this study, entire Sri Lanka was taken as the study area. Total Sri Lanka has area of 65,610 square kilometers. Basically, it is divided into 9 provinces due to many aspects of administratively and planning. Then, there are 25 districts lie under these 9 provinces. These districts are further subdivided into 331 secretariat and 14,008 Grama Niladhari divisions. Transportation of Sri Lanka is mainly based in the road network which is centered on city of Colombo which is commercial capital of Sri Lanka. There are approximately 12000 km of the road transportation which includes both A class and B class roads and approximately 250km of expressways. Also there is a railway network which handles a small fraction of the total transportation of the entire country. Although, there are some domestic airline services available in Sri Lanka and it is negligible fraction.

In my study it is only considered about road transportation. In this study, we considered a DS division as Traffic Analysis Zone (TAZ) to develop the origin destination matrix for the entire Sri Lanka.

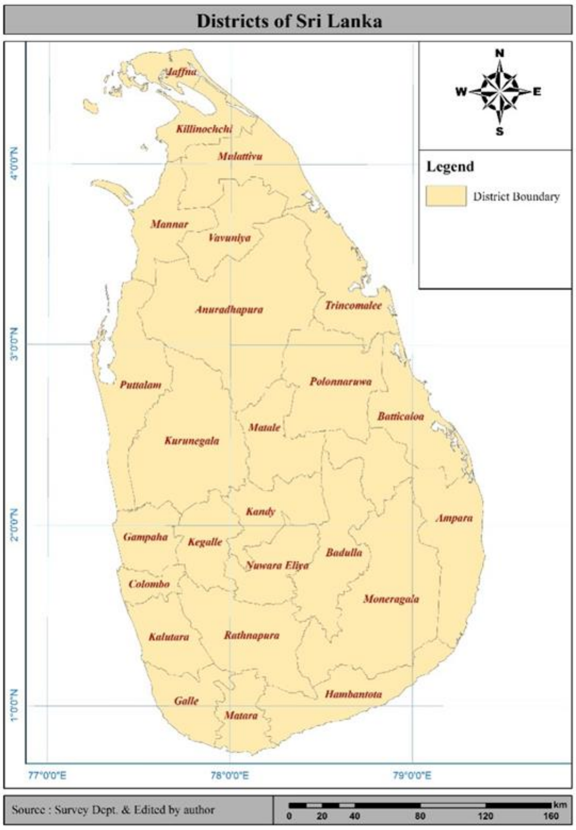
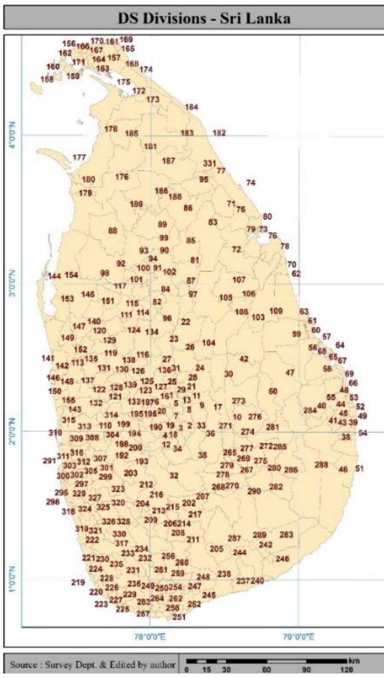


Figure 4.1: District of Sri Lanka



No	DSD	56	Eravur Pattu	108	Thamankaduwa	164	Nallur	218	Akmeemana	275	Passara
1	Akurana	57	Eravur Town	109	Welikanda	165	Vadamaradchi North	219	Ambalangoda	276	Rideemaliyadda
2	Delthota	58	Kattankudy	110	Alawwa	166	Valikamam South West (Sandilipay)	220	Baddegama	277	Soranathota
3	Doluwa	59	Koralai Pattu West	111	Ambanpola			221	Balapitiya	278	Uva Paranagama
4	Ganga Ihala Korale	60	Koralai Pattu	112	Bamunakotuwa	167	Valikamam South (Uduvil)	222	Benthota	279	Welimada
5	Harispattuwa	61	Koralai Pattu Central	113	Bingiriya			223	Bope Poddala	280	Badalkumbura
6	Hatharaliyadda	62	Koralai Pattu North	114	Ehetuwewa	168	Vadamaradchi East	224	Elpitiya	281	Bibile
7	Kandy Four Gravets & Gangawata	63	Koralai Pattu South	115	Galgamuwa	169	Vadamaradchi South west (Karaveddy)	225	Galle Four Gravets	282	Buttala
		64	Manmunai North	116	Ganewatta			226	Gonapeenuwala	283	Katharagama
8	Kundasale	65	Manmunai Pattu	117	Giribawa	170	Valikamam North	227	Habaraduwa	284	Madulla
9	Medadumbara	66	Manmunai South & Eruvil Pattu	118	Ibbagamuwa	171	Island South (Velanai)	228	Hikkaduwa	285	Medagama
10	Minipe			119	Kobeigane	172	Kandavalai	229	Imaduwa	286	Moneragala
11	Panvila	67	Manmunai South West	120	Kotavehera	173	Karachchi	230	Karandeniya	287	Sevanagala
12	Pasbage Korale			121	Kuliyapitiya East	174	Pachchilaipalli	231	Nagoda	288	Siyambalanduwa
13	Pathadumbara	68	Manmunai West	122	Kuliyapitiya West	175	Poonakary	232	Neluwa	289	Thanamalvila
14	Pathahewaheta	69	Porativu Pattu	123	Kurunegala	176	Madhu	233	Niyagama	290	Wellawaya
15	Poojapitiya	70	Veruga	124	Maho	177	Mannar Town	234	Thawalama	291	Colombo
16	Thumpane	71	Gomarankadawala	125	Mallawapitiya	178	Manthai West	235	Welivitiya Divithura	292	Dehiwala
17	Udadumbara	72	Kantalai	126	Maspotha	179	Musalai	236	Yakkalamulla	293	Homagama
18	Udapalatha	73	Kinniya	127	Mawathagama	180	Nanaddan	237	Ambalantota	294	Kaduwela
19	Udunuwara	74	Kuchchaveli	128	Narammala	181	Manthai East	238	Angunakolapelessa	295	Kesbewa
20	Yatinuwara	75	Morawewa	129	Nikaweratiya	182	Maritimepattu	239	Beliatta	296	Kolonnawa
21	Ambanganga Korale	76	Muttur	130	Panduwasnuwara	183	Oddusuddan	240	Hambantota	297	Maharagama
22	Dambulla	77	Padavi Sri Pura	131	Panduwasnuwara East	184	Puthukudiyiruppu	241	Katuwana	298	Moratuwa
23	Galewela	78	Seruvila			185	Thunukkai	242	Lunugamvehera	299	Padukka
24	Laggala Pallegama	79	Thambalagamuwa	132	Pannala	186	Vavuniya	243	Okewela	300	Ratmalana
25	Matale	80	Trincomalee Town & Gravets	133	Polgahawela	187	Vavuniya North	244	Sooriyawewa	301	Hanwella
26	Naula			134	Polpithigama	188	Vavuniya South	245	Tangalle	302	Sri Jayawardanapura Kotte
27	Pallepola	81	Galenbindunuwewa	135	Rasnayakapura	189	Vengalacheddikulam	246	Thissamaharama	303	Thimbirigasyaya
28	Rattota	82	Galnewa	136	Rideegama	190	Aranayaka	247	Walasmulla	304	Attanagalla
29	Ukuwela	83	Horowpothana	137	Udubaddawa	191	Bulathkohupitiya	248	Weeraketiya	305	Biyagama
30	Wilgamuwa	84	Ipalogama	138	Wariyapola	192	Dehiowita	249	Akuressa	306	Divulapitiya
31	Yatawatta	85	Kahatagasdigiliya	139	Weerambugedara	193	Deraniyagala	250	Athuraliya	307	Dompe
32	Ambagamuwa	86	Kebithigollewa	140	Anamaduwa	194	Galigamuwa	251	Devinuwara	308	Gampaha
33	Hanguranketha	87	Kekirawa	141	Arachchikattuwa	195	Kegalle	252	Dickwella	309	Ja?Ela
34	Kothmale	88	Mahavilachchiya	142	Chilaw	196	Mawanella	253	Hakmana	310	Katana
35	Nuwara Eliya	89	Medawachchiya	143	Dankotuwa	197	Rambukkana	254	Kamburupitiya	311	Kelaniya
36	Walapane	90	Mihinthale	144	Kalpitiya	198	Ruwanwella	255	Kirinda Puhulwella	312	Mahara
37	Addalachchenai	91	Nachchadoowa	145	Karuwalagaswewa	199	Warakapola	256	Kotapola	313	Minuwangoda
38	Akkaraipattu	92	Nochchiyagama	146	Madampe	200	Yatiyanthota	257	Malimbada	314	Mirigama
39	Alayadiwembu	93	Nuwaragam Palatha Central	147	Mahakumbukkadawala	201	Ayagama	258	Matara Four Gravets	315	Negombo
40	Ampara			148	Mahawewa	202	Balangoda	259	Mulatiyana	316	Wattala
41	Damana	94	Nuwaragam Palatha East	149	Mundel	203	Eheliyagoda	260	Pasgoda	317	Agalawatta
42	Dehiattakandiya			150	Nattandiya	204	Elapatha	261	Pitabeddara	318	Bandaragama
43	Irakkamam	95	Padaviya	151	Nawagattegama	205	Embilipitiya	262	Thihagoda	319	Beruwala
44	Kalmunai	96	Palagala	152	Pallama	206	Godakawela	263	Weligama	320	Bulathsinhala
45	Karaitheevu	97	Palugaswewa	153	Puttalam	207	Imbulpe	264	Welipitiya	321	Dodangoda
46	Lahugala	98	Rajanganaya	154	Vanathavilluwa	208	Kahawatta	265	Badulla	322	Horana
47	Mahaoya	99	Rambewa	155	Wennappuwa	209	Kalawana	266	Bandarawela	323	Ingiriya
48	Navithanveli	100	Thalawa	156	Valikamam West	210	Kiriella	267	Ella	324	Kalutara
49	Ninthavur	101	Thambuttegama	157	Thenmaradchi	211	Kolonna	268	Haldummulla	325	Madurawala
50	Padiyathalawa	102	Thirappane	158	Delft	212	Kuruvita	269	Hali?Ela	326	Mathugama
51	Pothuvil	103	Dimbulagala	159	Jaffna	213	Nivithigala	270	Haputale	327	Millaniya
52	Sainthamaruthu	104	Elahera	160	Karainagar	214	Opanayaka	271	Kandaketiya	328	Palindanuwara
53	Sammanthurai	105	Hingurakgoda	161	Kalmunai Tamil Division	215	Pelmadulla	272	Lunugala	329	Panadura
54	Thirukkovil	106	Lankapura	162	Island North (Kayts)	216	Ratnapura	273	Mahiyanganaya	330	Walallavita
55	Uhana	107	Medirigiriya	163	Valikamam East	217	Weligepola	274	Meegahakiwula	331	[illegible]

Figure 4.2: DSD Map for Sri Lanka

Since there are 331 TAZs it should be 109,561 cells in the proposed origin destination matrix for entire Sri Lanka. Therefore it is difficult to fill all of these 109,561 cells in manual counting.

As per the previous studies done by the department of Civil Engineering in university of Moratuwa, it is identified that there are 397 counting locations within the entire Sri Lanka in order to eliminate double counting problems and leaky screen line problems. Also it is identified as there are should be 62 counting location by which covers more than 40% of the OD matrix. Therefore partially filled OD matrix which is covered approximately 40 % of the entire OD pairs prepared by the Department of University of Moratuwa is taken to this study. These OD surveys were carried out in year 2015. Therefore trip production and trip attraction of 40% of the TAZs can be taken to this study.

With the available literature review which were conducted in many other countries in the world have concluded that there is a close relationship in between trip attraction and trip production with socio economic variables with related to a particular region. Therefore in this study, assumption was made that trip attraction and trip production are depends on the socio economic variables of the Sri Lanka.

4.3 Collection of Socio Economic Data

Travel demand forecasting process is the important part of the transportation system planning morphology. There are four stages in this planning morphology as follows.

- Trip Generation
- Trip Attraction
- Modal Split
- Traffic Assignment

As per the literature review of this study it was identified that there is a relationship between trip attraction and trip production of a particular region with socio economic variables of that particular region. These socio economic variables depends on the urban activity system of that region. Therefore, in this study basic assumption of the travel demand forecasting process is that there is a stable relationship between transport demand and urban activity system. Therefore it is mandatory to obtain an understanding of the characteristics of the travel in order to transport needs and necessary solutions for that particular area.

Trip attraction and trip production of a particular area in the region can be effectively forecasted in association with socio economic data. As per the previous studies done in other countries it is identifies that trip production is closely associated with socio economic variables such as household structure, income of the person and vehicle ownership. Similarly, trip attraction is closely associated with socio economic variables such as commercial and industrial services, schools, land values etc. All socio economic variables obtained for this study can be summarized as follows.

- Population
- Household
- Car ownership
- Birth Rate
- Commercial Area
- Number of employment
- Total Income
- Residential Area
- Death Rate
- Number of Schools

I have obtained data of above mentioned socio economic factors in various ways as follows.

Population – data of population in divisional secretariat wise is obtained from the department of census and statistics of Sri Lanka. Latest information of the population in Sri Lanka is available in year 2012. Therefore I have to find the population of year 2018 because available partially filled OD matrix is for year 2018. Therefore in order to find the value of population in year 2018 I used the growth factor of 1.048.

Household – I could find the details if household for year 2012 from the department of census and statistical of Sri Lanka. Since data is required for year 2018, I used growth rate of 1.048.

Car ownership – details of the car ownership for year 2018 is obtained from the department of Motor Traffic Sri Lanka.

Birth Rate – Birth rate in DS division wise is available for year 2015 as per the department of census and statistics Sri Lanka. Therefore I find the value of the birth rate for year 2018 by using growth factor of 1.048.

Commercial Area – I have to find the commercial area in divisional secretariat wise for the year 2018. Land use GIS file for entire Sri Lanka is obtained from the Urban Development Authority. Values for commercial area in DS division wise are obtained from the attributes table.

Number of employment – details of the number of employment DS division wise for year 2018 are obtained from the department of census and statistics Sri Lanka.

Total Income – Total income for the divisional secretariat for 2018 are obtained from the department of census and statistics Sri Lanka.

Residential Area – I have to find the commercial area in divisional secretariat wise for the year 2018. Land use GIS file for entire Sri Lanka is obtained from the Urban Development Authority. Values for residential area in DS division wise are obtained from the attributes table.

Death Rate – Birth rate in DS division wise is available for year 2015 as per the department of census and statistics Sri Lanka. Therefore I find the value of the birth rate for year 2018 by using growth factor of 1.048

Number of Schools – List of the schools in divisional secretariat wise in year 2018 is collected from the Ministry of Education.

Above 10 number of socio economic factors were selected based on the literature review. It could be identified that there are close relationship between transportation and above mentioned socio economic factors as per the research journals. For example normally population is high for a particular region, then trip generation should be higher than other area in which low population. Also when number of schools are high in a particular region then that area has higher trip attraction than other area in which low number of schools.

Likewise all the necessary socio economic factors for the entire 331 divisional secretariats are collected and tabled.

Then it is required to find the relationship between the collected socio economic variables and the transport pattern. In order to find out that there are various mathematical techniques and it is required to evaluation properly.

Data type	Available years	Prediction method
Population	2012	Population growth rate
No. of Households	2012	Calculated growth rate
Car Ownership	2018	Calculated growth rate
Birth Rate	2015	Calculated growth rate
Commercial Area	2018	Not applicable
Number of Employment	2018	A survey was carried out.
Total Income	2018	A survey was carried out.
Residential Area	2018	Calculated growth rate
Death Rate	2015	Calculated growth rate
Number of Schools	2018	Calculated growth rate

Table 4:1 Available Socio Economic Factors

In order to identify the actual relationship between trip attraction and trip production, there are several mathematical models which can be used. However, in this study Regression based mathematical modeling is used to future traffic production.

For future predictions, each of these independent variables should be predicted first using appropriate growth rates.

Many scholars in worldwide used several mathematical models to evaluate transport models and transport and land use models etc. The results of mathematical models are no more universally valid than those of empirical studies although mathematical models of dynamic behavior are also based on empirical surveys and observations.

There are various techniques used in several scholars in around the world as mathematical models to solve this problem. Regression Analysis is one of those methods.

4.4 Collection of partially filled Origin Destination Matrix

Partially filled origin destination matrix is collected from the Civil Department of University of Moratuwa. That Origin Destination matrix is prepared for the year 2018. Most of the cells in that OD matrix are not filled. (It is annexed as Appendices I). Approximately 25% of the total cells are filled in that partially filled origin destination matrix. As per that OD matrix it can be observed that there are trip production values for 80 number of traffic analyzing zones and trip attraction values for 84 number of traffic analyzing zones are available. Trip attraction values and trip production values for other traffic analyzing zones not accurate because most of the trip inter change values are not available.

There are 40,486 cells which represents trip interchange values filled in the available partially filled origin destination matrix. That means 37% of the total cells are filled in partially filled origin destination matrix. There are 331 divisional secretariats in this origin destination matrix. As per the standard it should be 25 number of values of trip production and trip attraction values at least in the origin destination matrix. However as per the available partially filled origin destination matrix, there are 69 and 62 traffic analyzing zones which have values for trip production and trip attraction respectively.

4.5 Regression Analysis and its importance

With the introduction of the computer, the field of statistics has dominantly changed. Many mathematical models can be used with the computer algorithms and computer applications.

The regression technique is the statistical methods to fit the relationship between one variable and one or more other variables. In this study there are many socio economic variables and therefore regression technique is the most suitable technique to solve this problems. Regression technique is used to predict what is likely to happen in the next quarter, year or even further into the future. That is the main advantage of the regression technique. If the equation will develop with existing data, then values of the dependent variables can be predicted. There are few of essential features of this technique as follows.

1. This techniques shows significant relationship between independent and dependent variables
2. This technique shows the strength of impact independent variables on dependent variables
3. This equation is derived is purely empirical in nature
4. This technique is based on the premises that the regression coefficients initially established will still remain unchanged in the future and can be used in the regression equation for predicting the future travel.

Most widely known model of regression technique is linear regression technique. In this technique, it can be seen that independent variables can be continuous or discrete while the dependent variable is continuous. The nature of the regression line is linear. In many cases independent variable is identified as ‘Y’ and dependent variable is identified as ‘X’. Linear regression establishes a relationship between the dependent variable(Y) and one or more independent variables(X) using a best fit straight line. This line is known as the regression line. Linear regression analysis consists of following three stages.

1. Analyzing the correlation and directionality of the data
2. Eliminate the multicollinearity
3. Estimating the model
4. Evaluating the validity and usefulness of the model

There are so many methods it can be used to develop a method for trip production and trip attraction. As per the literature review there are some methods are studied. However it is necessary to check the suitability before using those methods in Sri Lankan context.

The entropy-maximizing procedure analyzes the available information to obtain a unique probability distribution. In Sri Lankan context trip production and trip attraction for some DS divisions are unknown. Therefore if we use entropy maximization the accuracy of the results will be lower.

In growth factor model it is necessary to having previous data of the particular zones in order to predict the data of that particular zone. Although there are four methods of growth factor methods, all methods use that growth factor. Since in this case the data for some DS divisions are not available, it cannot be used growth factor method for Sri Lankan context.

As per the Intervening opportunity model, number of trips from one zone to other zone is proportionately to the number of intervening opportunity of the both zones. As per the literature review it is identified that this method is difficult to handle in practice and result of this model are inaccurate. Therefore this method is not suitable for the Sri Lankan context.

4.6 Gravity Model

According to the Scholar Wilson, the most probable trip distribution is same as the gravity model distribution. There are three categories of gravity model named as doubly constrained, singly constrained and unconstrained. The unconstrained gravity model makes an estimated trip distribution matrix which only matches the total number of trip on the network. However, these models are relatively low degree of accuracy. Also singly constrained gravity model also produces as estimated trip distribution matrix but only matches either trip production (production constraint) or trip attraction (attraction constraint) while doubly constrained gravity model accounts both. Out of these three models, doubly constrained gravity model can be identified as the best model since that is implementing in developing countries for forecasting trip distribution in the future which is very difficult task and complicated process. (Herijano and Thorpe, 2005). Scholar Wilson is developed the doubly constrained model by maximization of an entropy maximization.

Although the resultant base matric more closely reflects the observed behavior, Inability to relate any policy variables to these factors can be seen as one of the major drawbacks of this model. Hence, it is difficult to assess their validity in the future. It is able to estimate the trips between each origin and destination zone subject to given the trip productions, trip attractions and the accessibility between regions/zones. The doubly constraint gravity model will be used to perform the trip distribution.

4.7 Formulation of Models

As initial step, it is required to find the relationship between trip production and trip attraction with the social economic data. Social economic data for every 331 divisional secretariat divisions are available (it is annexed as appendices 3). Therefore the prime objective is to build two regression analysis based models for trip production and trip attraction. In order to do this we used to SPSS computer software.

Here, the prime objective was to build two regression analysis based models for trip production and trip attraction with the socio-economic data. Initially, the correlation between trip production and each socio-economic data was examined. The correlations derived with socio-economic data are as follows. Also significance value for each socio economic will be examined here.

By correlation analysis, parameters, which have a higher correlation and significance level with the trip production, were selected for further analysis. Initially we selected 69 traffic analyzing zones with socio economic factors in order to evaluate using the SPSS software (It is annexed as Appendices 4)

In order to do the correlation analysis we used SPSS software. Before that table was prepared for each socio economic factors with the values of ln (X), $\log_{10}$ (X), Sin X, Cos X, X^2, X^3, and e^X for values of socio economic factors. Separate tables were prepared for the each socio economic factors. (Those are annexed as appendices 5 to 14) Then data of the all tables were transferred to the SPSS software.

Then in the SPSS bivariate correlation analysis was done for the each socio factors and trip production values for traffic analyzing zones. Pearson correlation and two tailed significance analysis were carried out for the above data. Likewise same process was carried out for the tip attraction and the socio economic factors. Results from the analysis was tabled to one table to understand easily.

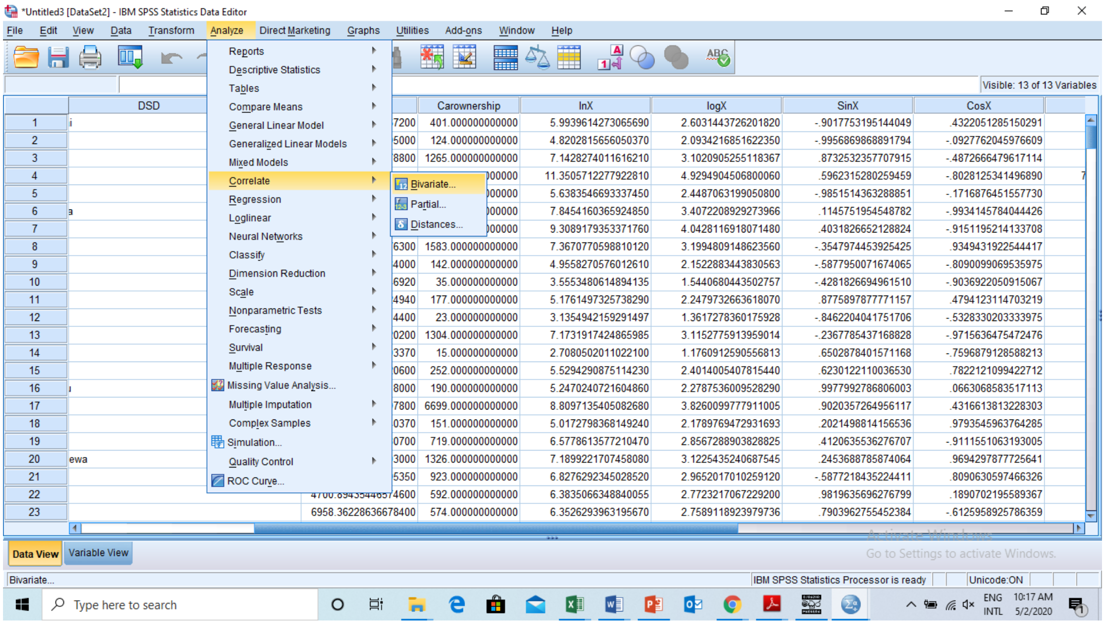
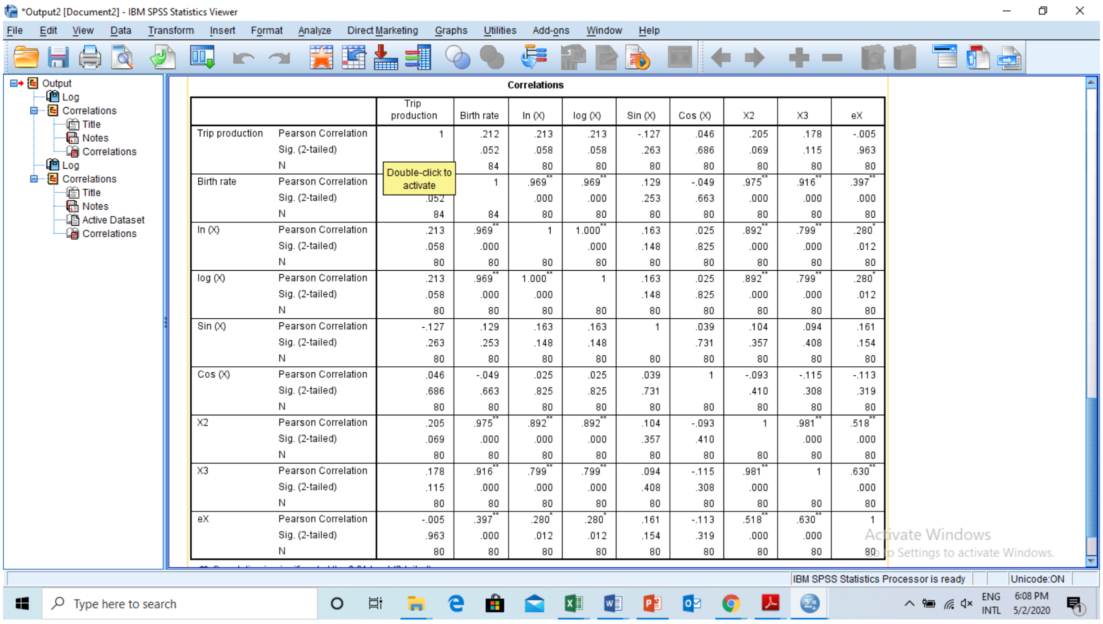


Figure 4.3: Analysis from SPSS Software



Correlations

		Trip production	Birth rate	ln (X)	log (X)	Sin (X)	Cos (X)	X2	X3	eX
Trip production	Pearson Correlation	1	.212	.213	.213	-.127	.046	.205	.178	-.005
	Sig. (2-tailed)		.052	.058	.058	.263	.686	.069	.115	.963
	N		84	80	80	80	80	80	80	80
Birth rate	Pearson Correlation	[illegible]	1	.969**	.969**	.129	-.049	.975**	.916**	.397**
	Sig. (2-tailed)	.052		.000	.000	.253	.663	.000	.000	.000
	N	84	84	80	80	80	80	80	80	80
ln (X)	Pearson Correlation	.213	.969**	1	1.000**	.163	.025	.892**	.799**	.280*
	Sig. (2-tailed)	.058	.000		.000	.148	.825	.000	.000	.012
	N	80	80	80	80	80	80	80	80	80
log (X)	Pearson Correlation	.213	.969**	1.000**	1	.163	.025	.892**	.799**	.280*
	Sig. (2-tailed)	.058	.000	.000		.148	.825	.000	.000	.012
	N	80	80	80	80	80	80	80	80	80
Sin (X)	Pearson Correlation	-.127	.129	.163	.163	1	.039	.104	.094	.161
	Sig. (2-tailed)	.263	.253	.148	.148		.731	.357	.408	.154
	N	80	80	80	80	80	80	80	80	80
Cos (X)	Pearson Correlation	.046	-.049	.025	.025	.039	1	-.093	-.115	-.113
	Sig. (2-tailed)	.686	.663	.825	.825	.731		.410	.308	.319
	N	80	80	80	80	80	80	80	80	80
X2	Pearson Correlation	.205	.975**	.892**	.892**	.104	-.093	1	.981**	.518**
	Sig. (2-tailed)	.069	.000	.000	.000	.357	.410		.000	.000
	N	80	80	80	80	80	80	80	80	80
X3	Pearson Correlation	.178	.916**	.799**	.799**	.094	-.115	.981**	1	.630**
	Sig. (2-tailed)	.115	.000	.000	.000	.408	.308	.000		.000
	N	80	80	80	80	80	80	80	80	80
eX	Pearson Correlation	-.005	.397**	.280*	.280*	.161	-.113	.518**	.630**	1
	Sig. (2-tailed)	.963	.000	.012	.012	.154	.319	.000	.000	
	N	80	80	80	80	80	80	80	80	80

Figure 4.4: Results for 'Birth Rate' Socio Economic Factor

Comparison of correlations for the socio economic factors for trip production can be derived as following table.

Trip Production		**X**	**Ln (X)**	**Log_{10} (X)**	**Sin (X)**	**Cos (X)**	**X^2**	**X^3**	**e^x**
Population	Pearson Correlation	.837	.827	.827	-.031	-.044	.753	.701	.000
	Sig. (2-tailed)	.000	.000	.000	.594	.438	.000	.000	1.000
Number of House holds	Pearson Correlation	.789	.728	.725	.668	-.668	.624	.522	.000
	Sig. (2-tailed)	.000	.000	.000	.085	.194	.000	.000	1.000
Car Ownership	Pearson Correlation	.684	.661	.561	-.520	.607	.588	.408	-.007
	Sig. (2-tailed)	.000	.000	.000	.007	.002	.000	.000	.012
Birth Rate	Pearson Correlation	.317	.293	.293	-.162	.014	.186	.298	.301
	Sig. (2-tailed)	.000	.000	.000	.000	.298	.000	.000	.000
Commercial Area	Pearson Correlation	.422	.381	.381	-.365	.238	.178	.385	.000

	Sig. (2-tailed)	.000	.000	.000	.004	.005	.011	.000	1.000
Number of Employment	Pearson Correlation	.440	.354	.354	.014	-.140	.328	.380	a
	Sig. (2-tailed)	.000	.000	.000	.380	.011	.000	.000	.
Total Income	Pearson Correlation	.852	.788	.788	.000	-.024	.527	.495	a
	Sig. (2-tailed)	.000	.000	.000	.095	.072	.000	.000	.
Residential Area	Pearson Correlation	.725	.634	.634	-.024	.022	.730	.691	.000
	Sig. (2-tailed)	.000	.000	.000	.647	.682	.000	.000	1.000
Death Rate	Pearson Correlation	.156	.096	.091	-.041	-.072	.149	.104	a
	Sig. (2-tailed)	.000	.000	.000	.009	.002	.000	.000	.
Number of Schools	Pearson Correlation	.228	.271	.271	-.153	.049	.003	-.020	a
	Sig. (2-tailed)	.000	.000	.000	.007	.389	.976	.727	.

Table 4:2 Comparison of corelations for Trip Production

X value of the above table refers the pure value of the socio economic factor for each traffic analyzing zone. Others are ln value, log value, sin value, cos value, square value and triple value of the socio economic factor of each traffic analyzing zone.

It is outstanding that significant value of all correlations above are less than 0.01. That means those relationships are really exist. These selected parameters were again tested for the multicollinearity effects firstly.

According to the value of the correlation of the attributes we can summarized the first five socio economic factors as follows.

No	Name if the Attribute	Correlation	Significance
1	Total Income	0.852	0.000

2	Population	0.837	0.000
3	Number of House holds	0.789	0.000
4	Residential Area	0.725	0.000
5	Car Ownership	0.684	0.000

Table 4:3 HIgh correlated attributes with Trip Production

In order to develop model for trip production above five numbers of socio economic factors were considered. Then we took different combinations of the above socio economic variables for the model calibration for trip production. 31 different combinations can be made with the above said five socio economic factors. Also 69 number of trip production values for 69 traffic analyzing zones can be obtained as per the available partially filled origin destination matrix. From that 69 number of trip production values 40 number of trip production values are taken to the development of the model equation. Other 29 number of trip production values are used for model validation. Therefore the said 40 number of trip production values and 31 different combinations which are made from above mentioned five number of socio economic factors are input to the SPSS computer software. Then 31 different model equations are derived (annexed as appendices 15). R value, R squared value and standard error of the estimates can be derived by the SPSS software. Out of 31 derived model equations 10 model equations can be shown as follows.

Model No	Input Parameter	Model Equation	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	Total Income	4.472(Total Income) - 58902.47	0.840	0.706	0.7048	49356.9810
2	Total Income No of Household	2.498(Total Income) + 0.765(No of Household) - 58449.378	0.858	0.736	0.7186	48562.14695
3	Total Income No of Household	1.561 x 10(-5)(Total Income) + 1.328(No of	0.862	0.744	0.7205	49468.5103

	Population	Household) + 0.585 (Population) -48519.74				
4	Total Income No of Household Population Car Ownership	1.792 x 10(-5)(Total Income) + 2.086(No of Household) + 0.469 (Population) - 1.006(Car ownership) -38857.26	0.878	0.753	0.7286	51856.896
5	Total Income Population Car Ownership Residential Area	1.486 x 10(-5)(Total Income) + 0.458(Population) - 0.864(Car Ownership) +1.855(Residential Area) – 24865.85	0.854	0.766	0.7415	50864.963
6	Total Income Population No of Household Residential Area	1.953 x 10(-5)(Total Income) + 0.691(Population) + 5.862(No of Household) - 105 (Residential Area) – 32856.15	0.846	0.716	0.7002	49516.468
7	Total Income No of Household Residential Area Car Ownership	1.463 x 10(-5)(Total Income) + 0.497(No of Household) - 0.042(Residential Area) +2562 (Car Ownership) – 15856.26	0.758	0.574	0.5382	50895.149
8	Population No of Household Residential Area Car Ownership	85.45(Population) + 0.298(No of Household) + 10.46(Residential Area) – 152.12 (Car Ownership) – 24625.83	0.788	0.621	0.5968	49516.286
9	Total Income Population Car Ownership	1.324x10(-5)(Total Income)+0.385(Population) – 1.956(Car Ownership) – 21647.51	0.818	0.669	0.6125	51628.185
10	Population Residential Area Car Ownership	105.85(Population) + 0.385(Residential Area) – 5.049(Car Ownership)– 19524.63	0.828	0.686	0.6452	50149.859

Table 4:4 Calibrated Equations for Trip Production

After obtaining calibrated equations, R value, R squared value, adjusted R squared value and standard error of the estimate are calculated for each calibrated equations separately.

R value, R squared value, adjusted R squared value and standard error of the estimated are critically evaluate after the analysis. High R value means correlation is high with independent variables and dependent variable. Therefore we have to choose highest R value model equation as most suitable trip production model. The model with highest R value and R squared value is selected as the best fit trip production model.

Trip Production Model

Trip Production = 1.792 X $10^{(-5)}$ (Total Income) + 2.086 (No of Household) + 0.469 (Population) - 1.006 (Car Ownership) -38857.26

As per the above derived equation trip production is directly function to the total income, number of employment, population and car ownership of the particular traffic analysis zone other than other socio economic factors of that traffic analysis zone. Therefore, if we know those mentioned three factors for a particular traffic analysis zone we can predict the trip production of a particular traffic analysis zone.

4.7.1 Comprehensive Validation of the Trip Production Model

4.7.1.1 Overall Model Fit

The overall model fit is considered to find the variance of the dependent variables. If value of the R^2 is 0.766 which means 76.6% of the variance of dependent variables can be explained by the model. Overall model fit is tabulated in following table.

R	R Square	Adjusted R Square	Standard Error of the Estimate
0.8750	0.766	0.7415	50864.963

Table 4:5 Overrall Model Fit for Trip Production

4.7.1.2 Goodness of Fit

Paired wise T test is carried out for checking the goodness of fit of this model. In order to that comparison is carried out derived model equation with the available balance 30 number of trip production values. Here it can be checked the gap between actual Origin destination matrix and estimated origin destination matrix. The average result of the paired wise T test should be less than 0.05. According to the T test result can be illustrated as follows.

Mean	Variation	Mean	Variation	Correlation	hmd	df	p	t
1.185	1.04	1.215	1.008	0.9420	0	24	0.124	-0.94

Table 4:6 Results of paired wise T test analysis for Trip Production

P value of the paired wise t test analysis should be higher than 0.05 according to the analysis. If so, it would accept the null hypothesis. As per the above table p value is 0.124. This means that there is a 12.4% of probability to get null hypothesis truth. Also value of the t value is -0.94 which is close to zero means there is no high level of difference between two data distribution.

Similarly, correlation analysis was carried out for the trip attraction with the socio economic variables.

In order to do the correlation analysis we used SPSS software. Before that table was prepared for each socio economic factors with the values of ln (X), log10 (X), Sin X, Cos X, X2, X3, and eX for values of socio economic factors. Separate tables were prepared for the each socio economic factors. (Those are annexed as appendices 16 to 25) Then data of the all tables were transferred to the SPSS software.

Then we have values for total income, number of households, population and car ownership for entire 331 divisional secretariats. Also we have derived trip production model. Then we could calculate trip production values for each traffic analyzing zone. We have known trip production values for 69 number of trip analyzing zones.

Then we calculated trip production values by the derived model for those 69 number of trip analyzing zones. After that we could check the accuracy of the derived model by conducting t test to goodness to fit.

As a results of the t test it is noted that derived trip production model is fit for the Sri Lankan context.

Then in the SPSS bivariate correlation analysis was done for the each socio factors and trip attraction values for traffic analyzing zones.

Pearson correlation values and significance values which are obtained from analysis done by SPSS software can be summarized as following table. The socio economic variables which have higher correlation and significance level with the trip production, were selected for further analysis.

Trip Attraction	**X**	**Ln (X)**	**Log_{10} (X)**	**Sin (X)**	**Cos (X)**	**X^2**	**X^3**	**e^x (X)**

Population	Pearson Correlation	.821	.809	.792	-.753	-.052	.740	.785	.000
	Sig. (2-tailed)	.000	.000	.000	.677	.349	.000	.000	1.000
Number of Household	Pearson Correlation	.752	.658	.658	.594	-.665	.549	.601	.000
	Sig. (2-tailed)	.000	.000	.000	.087	.242	.000	.000	1.000
Car Ownership	Pearson Correlation	.387	.339	.339	-.024	.009	.262	.186	-.020
	Sig. (2-tailed)	.000	.000	.000	.658	.875	.000	.001	.724
Birth Rate	Pearson Correlation	.251	.185	.185	-.216	.008	.177	.131	.101
	Sig. (2-tailed)	.000	.000	.000	.000	.943	.000	.000	.000
Commercial Area	Pearson Correlation	.808	.789	.769	-.757	.737	.669	.179	.000
	Sig. (2-tailed)	.000	.004	.004	.301	.502	.210	.000	1.000
Number of Employment	Pearson Correlation	.782	.747	.747	.623	-.633	.530	.673	a
	Sig. (2-tailed)	.000	.000	.000	.004	.002	.000	.000	.
Total Income	Pearson Correlation	.685	.650	.650	-.456	-.621	.535	.518	1.000
	Sig. (2-tailed)	.000	.000	.000	.001	.006	.000	.000	.
Residential Area	Pearson Correlation	.332	.533	.415	-.026	.019	.720	.681	.000
	Sig. (2-tailed)	.000	.000	.000	.637	.734	.000	.000	1.000
Death Rate	Pearson Correlation	.399	.289	.289	.248	-.075	.418	.392	a
	Sig. (2-tailed)	.000	.000	.000	.562	.188	.000	.000	.
Number of Schools	Pearson Correlation	.744	.678	.678	-.555	.655	.007	.522	a
	Sig. (2-tailed)	.000	.000	.000	.005	.436	.898	.687	.

Table 4:7 Comparison of Correlations for Trip Attraction

X value of the above table refers the pure value of the socio economic factor for each traffic analyzing zone. Others are ln value, log value, sin value, cos value, square value and triple value of the socio economic factor of each traffic analyzing zone.

Then it is identified that highest Pearson correlation values are derived for the pure value of the socio economic factors. Therefore it can be concluded that trip attraction values are correlated with the pure values of the socio economic factors.

It is outstanding that significant value of all correlations above are less than 0.01. That means those relationships are really exist. These selected parameters were again tested for the multicollinearity effects firstly.

According to the value of the correlation of the attributes we can summarized the first ten socio economic factors as follows.

No	Name of the Parameter	Correlation	Significance
1	Population	0.821	0.000
2	Commercial Area	0.808	0.000
3	Number of Employment	0.782	0.000
4	Number of House hold	0.752	0.000
5	Number of Schools	0.744	0.000

Table 4:8 High Correlated Attributes with Trip Attraction

Then above five numbers of socio economic factors were taken for the preparation of the model for trip attraction. Also we took pure values of the above mentioned socio economic factors for the model development.

Then we took different combinations of the above socio economic variables for the model calibration for trip attraction. 31 different combinations can be made with the above said five socio economic factors. Also there 62 trip attraction values for 62 traffic analyzing zones can be obtained from the available partially filled origin destination matrix. Out of those 62 trip attraction values, 40 trip attraction values are taken to the development of model equation for trip attraction and balance 22 number of trip attraction values are taken to the derived model validation. Then the said 40 trip attraction values and all 31 combinations are input to the SPSS computer software. Then 31 different model equations are derived. R value, R squared value and standard error of the estimates can be derived by the SPSS software (it is annexed as appendices 26). Out of 31 derived model equations 10 model equations can be shown as follows.

Model No	**Input Parameter**	**Model Equation**	**R**	**R Square**	**Adjusted R^2**	**Std. Error of the Estimate**
1	Population	0.481(Population) - 26847.915	0.848	0.719	0.717	51264.126
2	Population No of House hold Commercial Area No of Employment	2.96(Population) +0.124 (No of House hold) + 9.548 (Commercial Area) + 8542(No of Employment) - 15985.14	0.851	0.724	0.722	0.46785.451
3	Population No of House hold No of Employment No of Schools	2.589(Population) + 0.984(No of House hold) +7.495(No of Employment) – 10.85(No of Schools) - 49456.878	0.862	0.743	0.740	50156.159
4	Population No of House hold Commercial Area No of Schools	1.85(Population) + 0.844(No of House hold) + 2.52 (Commercial Area) - 3450(No of Schools) - 84582.48	0.875	0.781	0.780	48652.44
5	Population Commercial Area No of Employment No of Schools	1.245(Population) + 15.486(Commercial Area) + 0.586(No of Employment) – 78.146(No of Schools) – 50864.26	0.884	0.765	0.762	35894.15
6	No of House hold	1.859(No of House hold) +	0.852	0.726	0.764	46215.25

	Commercial Area No of Employment No of Schools	2.855(Commercial Area) + 6.89(No of Employment) – 1.198(No of Schools) – 19846.18				
7	Population No of House hold Commercial Area	0.465(Population) + 1.845(No of House hold) +1.795(Commercial Area) -51856.45	0.844	0.712	0.710	48652.74
8	No of House hold Commercial Area No of Employment	5.489(No of House hold) + 0.654(Commercial Area) + 1489(No of Employment) - 34582.15	0.826	0.682	0.680	40561.62
9	Commercial Area No of Employment No of Schools	49.985(Commercial Area) + 14.594(No of Employment) + 8.567(No of Schools) – 48532.62	0.871	0.759	0.755	51349.47
10	Population Commercial Area No of Schools	1.493(Population) + 1.482(Commercial Area) + 14.658 (No of Schools) – 28154.68	0.868	0.753	0.750	48514.19

Table 4:9 Calibrated Equations for Trip Attraction

R value, R squared value, adjusted R squared value and standard error of the estimated are critically evaluate after the analysis. High R value means correlation is high with independent variables and dependent variable. Therefore we have to choose highest R value model equation as most suitable trip attraction model.

After obtaining calibrated equations, R value, R squared value, adjusted R squared value and standard error of the estimate are calculated for each calibrated equations separately. The model with highest R value and R squared value is selected as the best fit trip attraction model.

Trip Attraction Model

Trip Attraction = 1.245 (Population) + 15.486 (Commercial Area) + 0.586 (Number of employment) – 78.146 (Number of Schools) – 50864.26

According to the above equation model, trip attraction for a particular traffic analysis zone is directly functioned by the population, commercial area, number of employment and number of schools in a particular traffic analysis zone. Therefore, with the value of those mentioned four socio economic factors of a particular TAZ we can predict the trip attraction of that traffic analysis zone.

4.7.2 **Comprehensive validation of the trip attraction model**

4.7.2.1 Overall model fit

The overall model fit is considered to find the variance of the dependent variables. If value of the R^2 is 0.954 which means 95.4% of the variance of dependent variables can be explained by the model. Overall model fit is tabulated in following table.

R	R Square	Adjusted R Square	Standard Error of the Estimate
0.884	0.781	0.780	48652.44

Table 4:10 Overall Model Fit for Trip Attraction

4.7.2.2 Goodness of Fit

Paired wise T test is carried out for checking the goodness of fit of this model. Here it can be checked the gap between actual Origin destination matrix and estimated origin destination matrix. The average result of the paired wise T test should be less than 0.05. According to the T test result can be illustrated as follows.

Mean	Variation	Mean	Variation	Correlation	hmd	df	p	t
74.35	4485.42	56.48	3168.28	0.85	0	24	0.218	1.850

Table 4:11 Result of paired wise T test analysis for Trip Attraction

P value of the paired wise T test analysis should be higher than 0.05 according to the analysis. If so, it would accept the null hypothesis. As per the above table p value is 0.218. This means that there is a 21.8% of probability to get null hypothesis truth. Also value of the t value is 1.850 which is close to zero means there is no high level of difference between two data distribution.

Then we have values for population, commercial area, number of employment and number of schools for entire 331 divisional secretariats. Also we have derived trip attraction model. Then we could calculate trip attraction values for each traffic analyzing zone. We have known trip attraction values for 62 number of trip analyzing zones.

Then we calculated trip attraction values by the derived model for those 62 number of trip analyzing zones. After that we could check the accuracy of the derived model by conducting t test to goodness to fit.

As a results of the t test it is noted that derived trip attraction model is fit for the Sri Lankan context.

It can be found the trip attraction and trip production for the other Traffic Analysis Zones of the Origin Destination Matrix. Then we have origin destination matrix with the values of trip attraction and trip production of whole origin destination matrix.

During the SPSS analysis multicollinearity check was done for the dependent variables. In SPSS first we choose analyze tab and then select correlate. Then bivariate was choose. Then window called bivariate correlation was opened. Then independent variables have been selected and click OK button. Then calculation is done by SPSS. Then it can be seen that Pearson correlation is not very high. Then clicked analyze tab and went to the regression under that category. After that opened linear regression tab by clicking linear under regression tab. Then chose independent variables and dependent variable separately. Then clicked statistically and chose collinearity diagnosis and finally clicked OK button. Then SPSS software calculated multicollinearity. It could be calculated of tolerance and VIF. As per the outcomes tolerance values are less than 1 and VIF is less than 10. Therefore we can argued that our independent variables are actually independent. Also we checked condition index in collinearity diagnostics table. All values are less than 15 means it could be concluded that our independent variables are actually independent to each other.

However, trip interchange values of the origin destination matrix are not evaluated. Therefore, next step is to find these trip interchange values.

4.8 Trip Interchange Model

After developing models for trip production and trip attraction, it could be calculated for trip production values and trip attraction values for the entire 331 divisional secretariat divisions. Then from that calculation final column and final raw can be filled.

Then next task is to fill the cells of the partially filled origin destination matrix. As per the available partially filled origin destination matrix 40,486 cells are filled. Therefore there are balance 69,075 cells are unfilled. Then we have to fill those cells which represents trip interchange values.

There were many models found out by many scholars in order to solve this problems. Entropy maximization, gravity model, Furness model etc. each models have pros and crones. Entropy maximization and gravity model are the mostly used by the various studies.

Although entropy maximization is widely used in the other studies it will obtain results based on the unknown factors. Also if there are number of unknown is more than the number if available information then result will be not accurate. This method is based on the uncertainty.

Therefore gravity model is based on the number of trips production and attraction in a particular zones and the distance between those two zones. The general gravity model can be illustrated as follows.

$$T_{ij} = \frac{KP_iA_j}{d_{ij}^n}$$

T_{ij} = Number of trips between zones i and j

P_i = Number of trips produced in zone i

A_j = Number of trips attracted to zone j

d_{ij} = Distance or time or cost of travelling between zone i and zone j

K = Constant which is independent from i and j

n = An exponential constant and value of this usually found between 1 and 3

k = Total number of zones.

It can be examined that this gravity model developed based on the travel cost between two zones because this model is based on the distance between those two zones. Travel cost is depends on the distance between two particular zones.

Then, this general gravity model is applied to find the interchange values between two zones. However, results from the general gravity zone have significant difference from the O-D surveys values. Therefore importance of the new version of the gravity model was arisen. As a result new version gravity model called doubly constrained gravity model was developed. Also this new model is called as production-attraction constrained gravity model. Therefore this doubly constrained gravity model is used to find the values of the interchange. Then interchange values obtained from the doubly constrained gravity model and values from O-D surveys were evaluated. Then it was found that those tow values were matched remarkably. Therefore this production-attraction constrained gravity model was selected. This model equation can be illustrated as follows.

$$T_{ij} = \frac{A_i B_j O_i D_j}{d_{ij}^2}$$

T_{ij}= No. of trips from origin i to destination j

O_i= No. of trips beginning from zone i

D_j= No. of trips ending at destination j

A_i= Origin specific correlation factor

B_j= Destination specific correlation factor

d_{ij}= Distance between origin i and destination j

Then all the balance trip interchange values of the origin destination matrix were calculated using the above model equation. However several iteration processes have been done for obtain best fit values for trip interchange.

Finally we have fully completed origin destination matrix for entire Sri Lanka.

4.8.1 Goodness of Fit

Then T test was carried out to find the goodness of fit of the derived gravity model. This T test analysis was done with using 50 of sample data. Then results of the T test can be illustrated as following table.

Mean	**Variation**	**Mean**	**Variation**	**Correlation**	**hmd**	**df**	**p**	**t**
1.5844	0.276	1.582	0.284	0.971	0	50	0.278	-1.05

Table 4:12 The results of T test Analysis

According to the above results it shows the p value is very higher than 0.05 and the difference is less between two data distribution. Therefore this result concludes that there is high probability to get null hypothesis truth. The value of t value of all factors are very close to zero. It shows there is no high level of significance variance between two variables.

Finally we could be able to obtain complete origin destination matrix for the entire Sri Lanka. (It is annexed as appendices 27)

CHAPTER 05

5 CONCLUSION AND RECOMMENDATION

5.1 Introduction

This research investigates to find a method to estimate a fully completed origin destination matrix from the partially filled origin destination matrix. Therefore, this chapter concludes all the findings that gained from the from the research. In addition to that, limitations of the study have been surfaced and recommendation for future researches are also given in this chapter.

5.2 Background

Origin destination matrix is one of the most crucial requirement of the traffic planning. According to the literature review, it was identified that most scholars have mentioned that traditional methods are important to estimate the origin destination matrix. Different countries used various methods to solve the problems incorporated with the OD matrix in related to their environmental context. It was found by the literature review that various methods such as entropy maximization, gravity model. Regression analysis, growth factor methods used in other countries according to their country.

However, there are no such literature on Sri Lankan context under this model.

However, some relevant organizations have tried to develop origin destination matrix for Sri Lankan context with doing manual surveys such as Road Development Authority, Civil Department of University of Moratuwa. However they couldn't develop fully filled origin destination matrix to entire Sri Lanka due to lack of finance and time consuming. Therefore this research helps to develop fully filled origin destination matrix which can used for traffic planning purposes.

This research found two separate equation model to find the trip production and trip attraction for each traffic analysis zone. In this research a particular divisional secretariat is considered as traffic analysis zone (TAZ). Therefore, there are 331 TAZs in the entire Sri Lanka. After that method was develop to find the trip interchange values in the origin destination matrix. There are 109,561 trip interchange values in the origin destination matrix which developed to the entire Sri Lanka. Also finally these methods are validated with the suitable techniques and finally developed the fully filled origin destination matrix for entire Sri Lanka.

5.3 Findings & Usefulness for planners

In order to find the model for trip production and trip attraction for the OD matrix, regression analysis was used. SPSS software used for this purpose. As per the literature review, it was found that trip production and trip attraction for a particular area depends on the socio factor of that particular area or the zone. Therefore in order to find the relationship between those we used to SPSS software.

Using linear regression analysis, model equations for trip production and trip attraction were found as follows.

Trip Production = $1.486 \times 10^{(-4)}$ (Total Income) + 0.458 (Population) – 0.864 (Car Ownership) + 1.855 (Rental Value) -24865.85

Trip Attraction = 1.245 (Population) + 15.486 (Commercial Area) + 0.586 (Number of employment) – 78.146 (Number of Schools) – 50864.26

With the above findings it is identified that trip production and trip attraction were closely related to two different set of socio economic factors. As a result trip production of a particular traffic analysis zone is depends on the total income, total population and car ownership of that zone. Similarly, trip attraction is depends on the land value, number of occupants, inter emigration and commercial area of that zone.

As per the above outcome it could be argued that trip production for a particular divisional secretariat will be increased when the income of the people of that divisional secretariat are high. Also we can say that income of the people is positive correlated to the trip production.

As same as commercial area is highly trip attracted. If the commercial area is more in a particular traffic analyzing zone, then trip attraction will be higher in that particular TAZ.

It was able to find the total trip production and trip attraction of a particular traffic analysis zone from that trip production model and trip attraction model. Then next task was to find the trip interchange values of the origin destination matrix. In order to find that doubly constrained or production attraction constrained gravity model was selected. Every cell of the origin destination matrix were calculated with that doubly constrained gravity model.

After elaborating all three models, validation process was carried out for all these three models. In the validation process all three models have good t value and therefore these three models have very high confidence for applying for actual scenarios.

Furthermore, these models can be used by the transport planners in order take their decision in long term transport options strategies. Since these models are developed for entire Sri Lanka, these models can be used for national decision making as well as regional decision making. Benefits of the finding of this research can be summarized as follows.

1. Understand preferred routes and trip trends to determine which regions are best for a targeted advertisement or a new retail location.
2. This can be used for preparation of the development plan for a particular area

by which how to change the land use pattern in order to enhance traffic management

3. This can be used by the government agencies in order to allocate budget for transportation in future.
4. It can be decide how to plan public transportation for a particular area with the results of accurate origin destination matrix'
5. With these derived trip production model and trip attraction model we can find the trip generation and trip attraction for divisional secretariat in the future without doing OD surveys. With the values of the socio economic factors which influence to trip production and trip attraction it can be found the trip generation and trip attraction for a particular divisional secretariat division.
6. This results can be used to decide highway construction, interchange location, number of lanes of the highway
7. This matrix can be used to decide to relocation of government organization so that convenient to public transportation

5.4 Limitation of the Study

This study was carried out based on the partially filled origin destination matrix which was prepared with the conventional survey techniques. Therefore, trip production model, trip attraction model and trip interchange model were checked with the exiting data of OD matrix in which data was obtained from the said conventional survey technique. Therefore result of this study is directly based on the actual survey data of the partially filled origin destination matrix.

Therefore if accuracy level of the partially filled OD matrix high means accuracy of the result of this study is high and vice versa.

5.5 Recommendation for future research

This research study was carried out based on the various assumptions regarding the existing transportation of the Sri Lanka. These assumptions were obtained for this study due to time limitation and data availability. Therefore this research study and the results of this study can be used for the future direction in order to get best opinion regarding the transportation behavior of the Sri Lanka.

Good transportation and railway transportation were not included in this study. Therefore it can be continue this study further with the railway transportation with the result of this study. Also further researches can be done in order to enhance the accuracy of the research.

References

[1] J. W. Hutchinson, "Plastic buckling," Advances in applied mechanics, vol. 14, pp. 67-144, 1974.

[2] N. R. S. M. Subha Nair, "TRIP MATRIX ESTIMATION FROM TRAFFIC COUNTS USING CUBE," International Journal of Research in Engineering and Technology, 2013.

[3] F. J. P. J. W. X. L. R. R. B. Yang, "Dynamic Origin-Destination Travel Demand Estimation using Location Based Social Networking Data," in 92nd Transportation Research Board Meeting, 2012.

[4] K. V. K. R. Sharminda Bera, "Estimation of origin-destination matrix from traffic counts: the state of the art," European Transport, no. 49, pp. 3-23 , 2011.

[5] S. S. K. G. J. K. Navya S V., "TRIP GENERATION MODEL FOR THE CORE AREA OF THIRUVANANTHAPURAM CITY," International Journal of Innovative Research in Science, Engineering and Technology, vol. 2, no. 1, 2013.

[6] M. Wegener, "Transport Geography and Spatial Systems," in OVERVIEW OF LAND-USE TRANSPORT MODELS, Dortmund, Germany, Spiekermann & Wegener Urban Regional Research, 2004, pp. 127-146.

[7] T. L. Paez, "Introduction to Model Validation," Validation and Uncertainty Quantification Department.

[8] R. Borck, M. Pflüger and M. Wrede, "A Simple Theory of Industry Location and Residence Choice," 2007.

[9] Goodbody Economic Consultants,Ballsbridge Park, Ballsbridge, Dublin, "Transport and Regional Development".

[10] INTERIM ADVICE NOTE (IAN 36/01), "THE USE AND APPLICATION OF MICROSIMULATION TRAFFIC MODELS," June 2001.

[11] A. C. Jordaan, B. E. Drost and M. A. Makgata, "Land Value as a Function of Distance from the CBD: The Case of the Eastern Suburbs of Pretoria," SAJEMS NS7, 2004.

[12] R. M. Pendyala and C. R. Bhat, "Emerging Issues in Travel Behavior Analysis," 2004.

[13] J. Schaeffer, A. Christian and D. Helms, "Transportation".

[14] Science Applications International Corporation, "Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land Use Petterns," September 2000.

[15] P. Tisato and C. Mayer, "Local vs Inter-Regional Travel: A Comparison of Two Regions of Adelaide," in 28th Australasian Transport Research Forum.

[16] U.S.Department of the Interior, U.S.Geological Survey, "Analyzing Land Use Change In Urban Environments," USGS, 1999.

[17] Torgil Abrahamsso, “Estimation of Origin-Destination Matrices Using Traffic Counts – A Literature Survey”, Interim Report