# AN APPROACH TO GEOGRAPHIC INFORMATION SYSTEM (GIS) MODELING AND DETERMINATION OF THE QUICKEST ROUTE FOR VEGETABLE DELIVERY



# INSTITUTE OF GRADUATE STUDIES UNIVERSITY OF MALAYA KUALA LUMPUR

2016

# AN APPROACH TO GEOGRAPHIC INFORMATION SYSTEM (GIS) MODELING AND DETERMINATION OF THE QUICKEST ROUTE FOR VEGETABLE DELIVERY

# MOHAMMAD ABOUSAEIDI

# THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# INSTITUTE OF GRADUATE STUDIES UNIVERSITY OF MALAYA KUALA LUMPUR

2016

# UNIVERSITY OF MALAYA ORIGINAL LITERARY WORK DECLARATION

Name of Candidate:

(I.C/Passport No:

)

Registration/Matric No:

Name of Degree:

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study:

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature

Date:

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name:

Designation:

#### ABSTRACT

This study involves the adoption of the Geographic Information System (GIS) modelling approach to determine the quickest routes for delivery of fresh vegetable. The distribution and delivery services of fresh vegetables are a major concern for producers of fresh produce especially concerning transportation. This research focuses on the Sime Darby Company in Malaysia which is producing fresh vegetables. The company is spending a significant part of its budget for a high standard of transportation and delivery services from the depot to hypermarkets. This research develops a spatial database management system based on significant parameters to solve the distribution problem. The optimisation of the model is based on several significant parameters that show the fastest routes for delivery. The main contribution of the developed model is its ability to capture data from different fields to solve distribution problem. This research integrates social data such as population and spatial data such as land use in developing the spatial database management system. Another contribution is to connect this information with digital road networks to select the fastest delivery route. The first objective of this thesis is about to identify the key factors that influence on the selection of the best routes for the distribution purpose. The second one is to develop the Spatial Database Management System and the third one is to develop and propose a GIS model for distribution problem. This PhD thesis presented the implementation of statistical analysis using a linear regression model to identify the key factors affecting driving time. It has been developed several scenarios in order to give the set of alternative fast route in order to distribute fresh vegetables. Among all of the factors in transportation and delivery of fresh product, the focus of this study is on examining the characteristic of transportation such as traffic density, drive time, population, land use and the location of hypermarkets in street network within the study area. The selected routes are chosen in several phases, based on the distance, driving time, average speed, population density, land use, car volume and time-frames as impedances.

The final output of this thesis is producing different maps integrated with road networks for decision makers to deliver their fresh vegetables by using the most efficient and fastest routes to reach their intended destinations.

university

#### ABSTRAK

Kajian ini melibatkan penggunaan pendekatan model Sistem Maklumat Geografi (GIS) untuk menentukan laluan yang paling cepat bagi penghantaran sayur-sayuran segar. Perkhidmatan pengedaran dan penghantaran sayur-sayuran segar adalah salah satu masalah yang membawa kebimbangan kepada pengeluar produk segar terutamanya yang melibatkan pengangkutan. Kajian ini tertumpu kepada syarikat Sime Darby Malaysia yang menghasilkan pelbagai jenis sayur-sayuran segar. Syarikat ini telah menggunakan peruntukan yang besar bagi mencapai standard yang tinggi dalam perkhidmatan pengangkutan dan penghantaran dari depot ke pasar raya utama untuk permintaan pengguna.

Kajian ini dapat membangunkan satu sistem pengurusan pangkalan data spatial berdasarkan parameter yang penting untuk menyelesaikan masalah pengagihan. Pengoptimuman model ini adalah berdasarkan kepada beberapa parameter penting yang menunjukkan laluan terpantas untuk tujuan penghantaran.

Sumbangan utama model yang dibangunkan berbanding model yang sedia ada adalah menyediakan laluan yang cekap berdasarkan masa perjalanan yang dikaitkan dengan keupayaannya menanggapi pelbagai data dari pelbagai kawasan bagi memudahkan pengedaran. Kepelbagaian data ini termasuklah dari aspek sosial dan ruang serta bidang lalu lintas dan pengangkutan. Perkara baru mengenai penyelidikan ini adalah pengintegrasian data sosial dan data spatial seperti penggunaan tanah dalam membangunkan sistem pengurusan pangkalan data spatial penduduk. Selain daripada itu, kajian ini turut menyumbang dari aspek pengintegrasian beberapa jenis maklumat dengan data rangkaian jalanraya digital untuk menyelesaikan masalah pengagihan bagi memilih laluan yang cepat untuk tujuan penghantaran.

Kajian ini menunjukkan pembolehubah penting yang mempunyai pengaruh yang paling tinggi masa pemanduan bagi mengatasi masalah pengagihan dengan cara memilih

laluan cekap. Kajian ini dijalankan untuk menentukan semua kemungkinan kesan pembolehubah terpilih untuk pengedar yang membuat sebarang keputusan mengenai perkhidmatan penghantaran. Objektif pertama kajian ini adalah untuk mengenalpasti faktor-faktor utama yang mempengaruhi pemilihan laluan yang terbaik untuk tujuan pengagihan. Yang kedua adalah untuk membangunkan Sistem Pengurusan Pangkalan Data Spatial dan yang ketiga adalah untuk membangunkan dan mencadangkan model GIS bagi masalah pengedaran dalam menentukan laluan penghantaran terpantas.

Kajian PhD ini telah menggunakan analisis statistik dengan menggunakan model regresi linear bagi mengenal pasti faktor-faktor utama yang mempengaruhi masa pemanduan. Ia telah membangunkan beberapa senario untuk memberi set laluan pantas alternatif bagi pengedaran sayur-sayuran segar. Fokus kajian ini adalah untuk meneliti ciri-ciri pengangkutan seperti kesesakan lalu lintas, masa pemanduan, kepadatan penduduk, penggunaan tanah dan lokasi pasar raya besar di rangkaian jalan di dalam kawasan kajian. Laluan dipilih dalam beberapa fasa, berdasarkan jarak, masa, kelajuan purata, kepadatan penduduk, guna tanah, jumlah kenderaan dan jangka masa pemanduan.

Dapatan akhir kajian ini ialah menghasilkan peta yang berbeza dan berintegrasi dengan rangkaian jalan raya bagi membolehkan pengedar menghantar sayur-sayuran segar mereka dengan menggunakan laluan paling berkesan dan paling cepat ke destinasi.

#### ACKNOWLEDGEMENTS

All praise be to Allah the Most Compassionate the Most Merciful for his favour so that I may complete this research.

I express my heartfelt thanks to my supervisor, Associate Professor Dr. Rosmadi Fauzi, for his keen constructive comments, generous support of knowledge and his critical review of successive drafts chapter by chapter. I also express my heartfelt thanks to cosupervisor, Associate Professor Datin. Dr. Rusnah Muhamad, for his advice, guidance, encouragement and providing academic assistance during my study.

I would like to thank Sime Darby Bhd for their generous support, guidance and valuable references of market places in Kuala Lumpur and Selangor.

These few words do not express my deep love and gratitude to my father Ahmad Abousaeidi, my mother Batul Madani, my lovely spouse Sholeh Rostami and my sister Zohreh Abousaeidi for their prayers, understanding, moral support and encouragement.

I acknowledge my classmates in the Geography Department, UM student, and my friends in the Main Library for their encouragement.

Finally, I would like to thank all those who may not have been mentioned here, but have contributed to this study in one way or another.

# TABLE OF CONTENTS

| ABSTRACT III  |
|---|
| ABSTRAKV  |
| ACKNOWLEDGEMENTSVII   |
| TABLE OF CONTENTS VIII  |
| LIST OF FIGURESXIV  |
| LIST OF TABLESXVII  |
| CHAPTER 1: INTRODUCTION1  |
| 1.1 An overview1  |
| 1.2 Background of the study1  |
| 1.2.1 The importance of fresh vegetables in this study              |
| 1.2.2 The influence factors on transportation and delivery services |
| 1.3 Research problem  |
| 1.4 Research objectives   |
| 1.4.1 The main aim6   |
| 1.4.2 The objectives  |
| 1.5 Scope of the study  |
| 1.6 Motivation and contribution of this research                    |
| 1.6.1 The contribution of the GIS model in this research            |
| 1.7 Organisation of the thesis                                      |
| 1.7.1 Chapter 1: Introduction                                       |
| 1.7.2 Chapter 2: Literature review11                                |
| 1.7.3 Chapter 3: Research methodology11                             |
| 1.7.4 Chapter 4: Data preparation11                                 |
| 1.7.5 Chapter 5: Results of the model and discussions               |
| 1.7.6 Chapter 6: Conclusion and recommendations12                   |

| CHA | APTER 2:  | LITERATURE REVIEW  | 13 |
|-----|-----------|--|----|
| 2.1 | Introduct | ion  | 13 |
| 2.2 | Distribut | ion planning   | 13 |
|     | 2.2.1     | Transportation management                                      | 14 |
|     | 2.2.2     | Transportation issue in Malaysia                               | 15 |
|     | 2.2.3     | Fuel consumption on transportation cost                        | 17 |
| 2.3 | Vehicle r | routing problem (VRP)  | 18 |
|     | 2.3.1     | Prior studies on VRP   | 19 |
|     | 2.3.2     | GIS application regarding some influential factors and issues  | 23 |
|     | 2.3.3     | Parameters regarding transportation and delivery of goods      | 24 |
| 2.4 | Geograph  | hic information system (GIS)                                   | 26 |
|     | 2.4.1     | Transportation geography                                       | 28 |
|     | 2.4.2     | Geographic information system for transport                    | 29 |
| 2.5 | Fresh veg | getables   | 30 |
|     | 2.5.1     | Distribution planning for fresh vegetables                     | 31 |
|     | 2.5.2     | Distribution planning for fresh vegetables in Malaysia         | 32 |
|     | 2.5.3     | Prior research on transportation and delivery of fresh produce | 34 |
|     | 2.5.4     | Malaysian's consumer behaviour                                 | 35 |
|     | 2.5.5     | Drive time modelling approach based on several parameters      | 37 |
|     | 2.5.6     | Application of the dijkstra model                              | 38 |
| 2.6 | Route pla | anning system  | 39 |
|     | 2.6.1     | Route planning system based on GIS                             | 39 |
|     | 2.6.2     | Route choice for travelling                                    | 40 |
| 2.7 | Regressio | on model   | 43 |
|     | 2.7.1     | Application of regression model                                | 43 |
|     | 2.7.2     | The influence of regression analysis on GIS practices          | 46 |

|      | 2.7.3     | Vulnerability analysis                               | 47 |
|------|-----------|--|----|
|      | 2.7.4     | Solid waste collection using GIS network analyst     | 48 |
| 2.8  | GIS mod   | els  | 49 |
|      | 2.8.1     | The role of GIS in modelling                         | 49 |
|      | 2.8.2     | Type of models                                       | 50 |
|      | 2.8.3     | Spatial data models                                  | 51 |
| 2.9  | Distribut | ion as a DSS   | 52 |
|      | 2.9.1     | Components of DSS                                    | 54 |
|      | 2.9.2     | GIS as a DSS   | 54 |
| 2.10 | Network   | analysis in GIS                                      | 55 |
|      | 2.10.1    | Network analyst model                                | 56 |
| 2.11 | Database  | management system (DBMS)                             | 56 |
| 2.12 | Marketin  | g factors  | 60 |
|      | 2.12.1    | Development of Carrefour                             | 60 |
|      | 2.12.2    | Changing the name of Carrefour and Jusco to AEON BIG | 61 |
|      | 2.12.3    | Population issue in Malaysia                         | 62 |
| 2.13 | Conclusi  | on   | 63 |
| CHA  | APTER 3:  | RESEARCH METHODOLOGY                                 | 64 |
| 3.1  | Introduct | ion  | 64 |
| 3.2  | Duration  | and location of research                             | 64 |
|      | 3.2.1     | Location of study                                    | 68 |
| 3.3  | Material  | and tools  | 71 |
|      | 3.3.1     | Data collection                                      | 71 |
|      | 3.3.2     | Source of the data                                   | 74 |
|      | 3.3.3     | Tool requirements                                    | 75 |
|      | 3.3.4     | Methodology  | 77 |

| 3.4 | Database   | Management System                                   |     |
|-----|------------|---|-----|
| 3.5 | Conceptu   | al framework  | 82  |
| 3.6 | Regressi   | on model  | 85  |
| 3.7 | Conclusi   | on  | 87  |
| CHA | APTER 4:   | DATA PREPARATION                                    |     |
| 4.1 | Introduct  | ion   |     |
| 4.2 | Data prej  | paration  |     |
|     | 4.2.1      | Adding school zones                                 |     |
| 4.3 | GIS regr   | ession model  |     |
|     | 4.3.1      | Data observations                                   |     |
|     | 4.3.2      | Table of abundance distribution                     | 110 |
|     | 4.3.3      | Estimation of correlation                           | 114 |
| 4.4 | Statistica | l calculation and prediction results                | 116 |
|     | 4.4.1      | Durbin-Watson Test (D-W)                            | 116 |
| 4.5 | Creating   | the network analyst from dataset                    | 117 |
|     | 4.5.1      | Directions  |     |
| 4.6 | Conclusi   | on  | 122 |
| CHA | APTER 5:   | RESULTS OF THE MODELS AND DISCUSSIONS               | 124 |
| 5.1 | Introduct  | ion   | 124 |
| 5.2 | Function   | s in network analyst                                |     |
|     | 5.2.1      | Finding the best route                              | 126 |
| 5.3 | Names o    | f hypermarkets and their locations on the map       | 126 |
| 5.4 | Street da  | ta processing                                       | 129 |
|     | 5.4.1      | Restrictions  | 130 |
| 5.5 | Scenario   | S   | 130 |
|     | 5.5.1      | Phase one: Modelling of routes in terms of distance | 130 |

|     | 5.5.2      | Phase two: Impedance modelling of optimal routes in terms of drive<br>time        |
|-----|------------|---|
|     | 5.5.3      | Phase three: Impedance modelling of the primary time pattern—7 to 10 135          |
|     | 5.5.4      | Phase four: The impedance modelling of the secondary time pattern—<br>10 to 12 am |
|     | 5.5.5      | Phase five: the population as an impedance for calculating driving time           |
|     | 5.5.6      | Phase six: Car volume as the impedance for selecting the best routes 145          |
|     | 5.5.7      | Phase seven: The land use modelling to determine the optimal routes               |
| 5.6 | Statistica | data evaluation and analysis162   |
|     | 5.6.1      | Equation estimation   |
|     | 5.6.2      | Substituted coefficients  |
|     | 5.6.3      | Final substituted coefficients166   |
|     | 5.6.4      | Coefficients evaluation obtained from final regression model167                   |
|     | 5.6.5      | Research philosophy   |
| 5.7 | The final  | map result of the fast routes   |
| 5.8 | Statistica | l results   |
| 5.9 | Conclusio  | on171   |
| CHA | APTER 6:   | CONCLUSION AND RECOMMENDATIONS  |
| 6.1 | Introduct  | ion173  |
| 6.2 | Research   | findings173   |
| 6.3 | Outcome    | of the research   |
|     | 6.3.1      | Key influential factors on route selection  |
|     | 6.3.2      | Spatial Database Management System176   |
|     | 6.3.3      | The contribution of the GIS model in this research                                |
| 6.4 | The signi  | ficance of this research177   |

| APP | APPENDICES                      |     |  |
|-----|---------------------------------|-----|--|
| REF | ERENCES                         |     |  |
| 6.7 | Conclusion                      |     |  |
| 6.6 | Suggestions for future research | 179 |  |
| 6.5 | Limitations of the study        | 178 |  |

university

# LIST OF FIGURES

| Figure 1.1: Conceptual Framework9  |
|--|
| Figure 2.1: GIS organization (Bolstad, 2012)   |
| Figure 2.2: The Decision-Making/Modelling Process (Turban, 1993)53   |
| Figure 2.3: The Architecture of a GIS (Martin, 1991)   |
| Figure 2.4: GIS database design stages (Healey, 1991)  |
| Figure 3.1(continued) The Study Location (Data Source: From land use and cartographic department in KL)  |
| Figure 3.2: Distribution locations on the map (Source: The Malaysia map and Selangor area from the Land use and cartographic department in Kuala Lumpur) |
| Figure 3.3: Truck for the delivery of fresh vegetables (Sime Darby website)70  |
| Figure 3.4: Map of Hypermarket locations (Source: GIS and Cartographic Lab, UM) .70  |
| Figure 3.5: The Architecture of ArcGIS(Esri website)76   |
| Figure 3.6: Application developments with GIS79  |
| Figure 3.7: Development of Database Management System  |
| Figure 3.8: The conceptual framework of capturing data for analysis  |
| Figure 3.9: Variables selected for regression analysis   |
| Figure 4.1: Combination of all Road Layers90   |
| Figure 4.2: File dataset   |
| Figure 4.3: Map of Finalized Dataset   |
| Figure 4.4: The total number of road Networks  |
| Figure 4.5: Attribute Table of the Distance  |
| Figure 4.6: Attribute Table of Average Speed for Road Networks   |
| Figure 4.7: Map of Average Speed for Roads97   |
| Figure 4.8: Attribute Table of Average Speed for Two Time Patterns   |
| Figure 4.9: Map of Average Speed for Roads in Time Pattern 7-10 100  |
| Figure 4.10: Map of Average Speed for Road Networks in Time Pattern 10-12 101  |
| Figure 4.11: Attribute of drive time for time frames   |
| Figure 4.12: Field calculation for drive time  |
| Figure 4.13: Attribute table for Drive time  |
| Figure 4.14: Attribute Table for Population and Land Use   |
| Figure 4.15: Map of Schools' Locations   |

| Figure 4.16: Attribute Table for One-Way Roads and Car Volume107                                    |
|---|
| Figure 4.17: Distribution Cost Forecasting (SIMAFORE)   |
| Figure 4.18: Table of Abundance Distribution for Average Speed at7am to 10am111                     |
| Figure 4.19: Table of Abundance Distribution for Average Speed at10am to 12am111                    |
| Figure 4.20: Table of Abundance Distribution for Distance   |
| Figure 4.21: Table of Abundance Distribution for Population   |
| Figure 4.22: Table of Abundance Distribution for Car Volume   |
| Figure 4.23: Table of Abundance Distribution for Two-Ways Directions                                |
| Figure 4.24: Table of Abundance Distribution for School Area114                                     |
| Figure 4.25: Connectivity Example (Esri, ArcMap,2014)118  |
| Figure 4.26: Creating Network Analysis  |
| Figure 4.27: Attributes for the Network Dataset   |
| Figure 4.28: Summary of Network Dataset Establishment   |
| Figure 5.1: The Conceptual Framework of The Final Map   |
| Figure 5.2: The Location of Hypermarkets (Cartographic and GIS Lab, UM)128                          |
| Figure 5.3: Location of Starting Point (Sime Darby Bhd)   |
| Figure 5.4: Routes Selection Based On Distance  |
| Figure 5.5: Total Time and Distance for Routes Based Distance and One-Way Road 132                  |
| Figure 5.6: Total Time and Distance with Consideration of One-Way Roads                             |
| Figure 5.7: Best Routs Based on Drive Time  |
| Figure 5-8: Total Time and Distance by Consideration of Average Speed for Drive<br>Time             |
| Figure 5.9: Total Time and Distance Based on Average Speed Without Consideration<br>of One-Way Road |
| Figure 5.10: Fast Routes for Delivery in the 7 to 10Period  |
| Figure 5.11: Total Distance and Time for Delivery in the 7 to 10 Period                             |
| Figure 5.12: Total Distance and Time for Delivery Without Consideration of One-<br>Way              |
| Figure 5.13: Best Routes for Delivery in the 10 to 12 Period  |
| Figure 5.14: Total Time and Distance for Delivery in Time 10 to 12                                  |
| Figure 5.15: Total Time and Distance Without Consideration of One-Way Roads 140                     |
| Figure 5.16: Population Density on the Map  |
| Figure 5.17: Classification of Population   |
| Figure 5-18: Result of Fast Routes for Delivery Based on Population Impedance 144                   |

| Figure 5.19: Total Distance and Time for Delivery Based on Population Factor   | 145 |
|--|-----|
| Figure 5.20: Map of Average Speed for All the Roads in KL                      | 147 |
| Figure 5.21: Best Routes Based on Car Volume                                   | 148 |
| Figure 5.22: Total Time and Distance for Delivery Based on Car Volume          | 148 |
| Figure 5.23: Map of Open Spaces and Industrial Areas in Kuala Lumpur           | 150 |
| Figure 5.24: Land Use Map  | 151 |
| Figure 5.25: Location of Residential Areas                                     | 152 |
| Figure 5.26: Map of Road Combination and Residential Zones                     | 153 |
| Figure 5.27: Buffering Zone Map  | 154 |
| Figure 5.28: Map of Residential Area Selection                                 | 155 |
| Figure 5.29: Combination of Roads with Land Use                                | 156 |
| Figure 5.30: Result Obtained for Selecting Routes Based on Land Use Impedance  | 158 |
| Figure 5.31: Result of Total Distance and Time for Routes in Delivery Services | 158 |
| Figure 5.32: The Differences for Road Direction in Two Time Frames             | 162 |
| Figure 5.33: The final Map of Fast Routes Based on All Variables               | 169 |
|  |     |

# LIST OF TABLES

| Table 2.1: Variables adopted in GIS application  | 24 |
|--|----|
| Table 2.2: Variables for route choice behaviour  | 42 |
| Table 2.3: Variables in different regression analysis                                      | 44 |
| Table 2.4: Variables involved in the Regression model                                      | 45 |
| Table 2.5: Variables Adopted In Current Research   | 46 |
| Table 2.6: Variables in Vulnerability Analysis   | 47 |
| Table 3.1: The data needed for selecting the optimal route                                 | 73 |
| Table 3.2: Non-Spatial Data collection   | 75 |
| Table 3.3: The Spatial data  | 75 |
| Table 3.4: ArcGIS Extension (ESRI, 2012)   | 77 |
| Table 4.1: Information Regarding Data Abundance 11   | 10 |
| Table 4.2: Probability Value Result from Pearson Correlation 11                            | 15 |
| Table 4.3: Statistical Test 11   | 17 |
| Table 5.1: The Information Chart for Hypermarkets as Destination for Delivery      Purpose | 26 |
| Table 5.2: Classification of Population 14   | 42 |
| Table 5.3: Comparison between Population and Drive Time 14                                 | 45 |
| Table 5.4: Land Use Classification and Description About Classes in Attribute Table 15     | 57 |
| Table 5.5: Comparison Result of Distance and Drive Time 15                                 | 59 |
| Table 5.6: Comparison Results of Time Pattern of7 to 10 Driving Time                       | 60 |
| Table 5.7: Comparison results of time pattern of 10 to 12 and driving time                 | 60 |
| Table 5.8: Compression Results of Time Pattern of 7 to 10 and 10 to 12      10             | 61 |
| Table 5.9: First Round of Regression Results   | 64 |
| Table 5.10: 10:Second Round of Statistical Regression Results    16                        | 65 |
| Table 5.11: Final Results of Regression Model 16   |    |
|  | 66 |

## **CHAPTER 1: INTRODUCTION**

#### 1.1 An overview

Fresh vegetables are vulnerable to immediate contamination. They require quick delivery to optimise their freshness. An efficient delivery system will decrease the chances of incurring high traffic density and reduces delivery times. Sime Darby Bhd is a major supplier and distributor of fresh vegetables throughout Malaysia.

Distribution problems are a daily challenge (Belenguer, Benavent, & Martínez, 2005). This is a critical issue especially for companies involved in the delivery of goods, where their income strongly depends on how distribution problems are solved. Therefore, it is crucial for such companies to be able to perform the tasks as efficiently as possible. The use of GIS technology is widely defined by various scholars. One example would be in the area of distribution planning (Kim, 2014). The focus of this study is on application of network analysis in determining the fastest routes to distribute fresh vegetables to selected hyper markets in Kuala Lumpur and Selangor.

This section presents a brief introduction of the thesis. The next section describes of the background of the study, followed by the problem statement. The following section discusses the aims and objectives of the study. Section 1.5 describes the scope of the study. Section 1.6 highlights the motivation and contribution of this research.

### 1.2 Background of the study

GIS is an indispensable software tool in delivery and distribution of fresh vegetables planning. In the delivery of fresh produces, GIS can manage the data on the road network and solve the problems of delivery services of fresh vegetables such as identifying the fastest rout (Silva, 2016).According to Silva (2016), the aim is to determine the pattern of the road networks and define transport routes using GIS technology. She used the Network Analyst extension to find the shortest and fastest routes. Xue (2016) developed a model based on optimisation route for collection of municipal solid waste using the ant colony optimisation based multi-objective routing model. He considered drive time, accident probability and population based on mathematical equation and the optimal routes are proposed in diagrams. Sanjeevi (2016) studied solid waste transportation using ArcGIS software to reduce cost and identify the shortest distance. He developed a spatial database system for route optimisation.

Parsafard (2015) conducted a research for fuel dispatch between a bulk terminal and gas station as a distribution centre. The shortest path model based on multi-criteria objective was proposed to solve a routing problem.

Kuby (2014) presented an online mapping tool to find the feasible path in a road network to give the vehicle driving range and station locations. His thesis applied Dijkstra's shortest path algorithm to find the shortest feasible path. The efficiency of the shortest distance depends on several factors that can be calculated by ArcGIS software. For example, the shortest distance is not always the best choice due to several other parameters that effect drive time.

The model for travel time was developed by Salonen (2013) who considered congestion when comparing between the car and bus to obtain the best result for travel time.

The present research will develop a spatial database management system based on parameters to solve the distribution problem. The optimisation of the model is based on several significant parameters that show the fastest routes for delivery. The other advantages of this study are to produce a map and to consider different time patterns in the morning. The contribution of this research is that the land use characteristics as spatial data has a significant impact on drive time. ArcGIS software determines the most efficient and effective routes based on minimum drive time and proposes the final map with directions and travelling time.

#### **1.2.1** The importance of fresh vegetables in this study

Fresh vegetable in agribusiness requires efficient transportation (Khan et al., 2014). It is necessary to implement a good delivery services for distribution of fresh vegetables to obtain an efficient system that reduces the delivery cost and identifies the best route for delivery.

According to Pimenta (2010), the quality of the final fresh produce depends on the transportation and storage of the product. The product must be transported as fast as possible from its production area to consumer markets due to climate conditions. Highly perishable products such as fresh vegetables and fruits must also arrive early at the distribution centres.

## 1.2.2 The influence factors on transportation and delivery services

Malaysia's growing urban population and increasing household incomes have led to a rise in car ownership. According to Ariffin, Khairi &Zaharib (2013), the most developed and fastest growing region in the country is the Klang Valley that covers an area of about 2,826 square kilometres. It comprises more than 30 rapidly growing new towns that include Kuala Lumpur.

Kuala Lumpur is known as one of the most congested cities in Southeast Asia and therefore to find the suitable routes for delivery of fresh produce such as fresh vegetables is critical (Dissanayake et al., 2012). Kuala Lumpur is known as a metropolitan city (Dissanayake et al., 2012). KL has undergone extensive growth and modernisation due to economic development that started in 1987 and as a result, new industrial zones have appeared in and around KL especially in the cities in Klang Valley.

The road network is significant part of route planning which can be modelled by GIS technology (Li, 2008). Zhu (2015) studied the route choice model and mentioned that the quality of model estimation is sensitive to the appropriateness of the consideration set, and most people did not choose the shortest path. According to Fernandez (2014),

developing a model of road networks is useful because ArcGIS can calculate the shortest route between two locations lined to the road system. The model estimates an average speed for every section on the network according to regional location.

To solve the distribution problem, social data like population and spatial data such as land use characteristics combined with traffic data and integrated into the proposed model to help determine the most efficient routes. The parameters influencing drive time obtained from statistical analysis are added to the database. The development of a spatial database can decide the fastest routes based on parameters obtained from the statistical analysis.

#### **1.3 Research problem**

Because of traffic condition and the perishable nature of the fresh produce such as vegetables, selecting the best routes while reducing the distribution cost is imperative. Moreover, perishables usually have a shortlife period, so the timely delivery of fresh food significantly affects its costs (Hsu, Hung & Li, 2007). However, little attention has been given to transportation in most e areas within KL (Tey, 2003).

The distribution and delivery services of fresh vegetables are a concern for producers of fresh produce. Sime Darby Bhd is spending a significant part of its budget for high standard transportation and delivery services from the depot to hypermarkets.

The application of GIS modelling to solve distribution problems has become an important area of GIS (Ismail, 2014). This research developed a GIS model to address the problem of distribution in determining the fastest routes for fresh vegetables.

According to Tatomir, Rothkrantz & Suson (2009), in the case of traffic accidents along the chosen route, drivers can be delayed for hours, which cause a loss in income; making designates traffic congestion as a high priority problem. The construction of new highways in some areas is sometimes impossible due to ecological, political and financial factors. In the short-term, the only solution is to use the road network in an optimal and proper way, along with finding the fastest routes to avoid traffic congestion (Tatomir, Rothkrantz & Suson, 2009).

Sime Darby's business capacity has increased substantially. This caused the manager to realise the necessity for changing the way in which the routes are selected to deliver their fresh vegetables to hypermarkets due to the tropical climate of Southeast Asia, optimising distribution planning systems, especially for perishable goods such as fresh vegetables, is of utmost importance. Malaysia is a country with mostly invariable high temperatures throughout the year (Badgie, 2012). Thus, the optimisation of routing systems in Malaysia continues to draw extensive attention.

According to the existing models developed by ArcGIS software for the road networks, the distribution problem is based on distance and average speed in the determination of efficient routes. Parameter other than traffic are not considered in solving the distribution problem. For example, the volume of cars, population density as well as land use are not considered to solve the distribution problem. The current study uses ArcGIS to calculate the fastest routes based on the parameters that effect on drive time.

Popular navigation services such as Waze are used by drivers both to plan out routes and optimally navigate drive time. Waze collects traffic information in areas of interested and so can take real-time traffic conditions which can be calculated by individual drivers when computing optimal route recommendation. The disadvantage of Waze is that it can recommend a route to shorten travel time of each individual driver. While the system can work if there is only one user on any road at any given time with the speed limit of that particular car. However, it may not be possible for each user to avoid traffic without creating congestion on the clearer roads and its recommendation might lead to longer aggregate routes (Vasserman, 2015). Real-time traffic data, on the other hand, is rarely available and almost never network-complete (Brosi, 2014). In conclusion, the gap between those existing traffic models is that they developed the model based on traffic data. In the current model, the social data is also considered in distribution planning.

### **1.4 Research objectives**

### 1.4.1 The main aim

The main aim of this research is the application of the GIS modelling approach by using the network analysis tool to solve the problem of finding the fastest routes for the delivery of fresh vegetables. The selected routes are based on minimum drive time according to the variables derived from the model. This study highlights the variables which have the most influence on drive time to solve the distribution problem to select the efficient routes. This paper highlights the finding of an impact assessment study.

## 1.4.2 The objectives

The following objectives will be accomplished in this study.

- **1.** To identify the key factors that influence selecting the fastest routes for the distribution of fresh vegetables.
- 2. To develop a Spatial Database Management System (SDMS) for analysing the road networks data.
- **3.** To propose a proper GIS model for the distribution problem and to determine the fastest delivery routes.

### **1.5** Scope of the study

This study seeks to develop a model that can help producers of fresh vegetables to manage their delivery of fresh vegetables. This model uses the complete road network in Kuala Lumpur and some areas in Selangor. Landuse classification was considered as a parameter for decision makers to select the fast routes. It is assumed that landuse will affect driving time. The model will use the spatial data that is commonly used in GIS. This study is of interest to researchers in the area of transportation planning and delivery services of the fresh produce. Also, decision makers can easily decide in determining the most efficient routes for the delivery of fresh vegetables. The scope of this study is on GIS application modelling approach in solving the routeing problem. The regression model is applied to determine the parameters that affect route selection with respect to the fastest delivery of fresh vegetables.

#### **1.6** Motivation and contribution of this research

GIS is adopted to determine the fastest route based on travel time and the shortest distance between the selected hypermarkets. Transportation is a fastest growing field for the application of GIS technology (Memon, 2005). It is an ideal information management and tool for analysing the spatial data within transportation. In this study, GIS technology provides the capability of spatial data and network system to represent real data in producing different maps.

Previous researches have been conducted to solve the distribution problem. They have proposed different methods in transportation planning and management system in various fields of transportation, such as civil engineering. Some researchinvolves the distribution of foodstuffs, while others consider the distribution of dangerous goods based on selecting the safer routes due to the avoidance of potential disasters. Network analysis has also been discussed several times in the field of computer science and engineering.

For this research, the concern is fresh vegetables as a perishable product. Other relevant researches concern the same issue and address different methods to solve the problem. However, in most of their research, they usually apply the mathematical methods and algorithm to approach the problem of finding the optimal routes. The weakness for all of the researches mentioned in this study is the fact that most of them employed mathematical algorithms to solve the problem of finding the optimal route while neglecting the spatial data, which results in the output of their respective model being absent of performances on route maps.

Considering the importance of the spatial issue in a distribution planning system, the main objective of some research is to analyse the result obtained from different approaches of a spatial regression model (Lopes, 2014). Hence for the purpose of this research and to obtain the main objective a multi-linear regression model were developed for solving the routeing problem of delivery of fresh vegetables.

This thesis focuses on the delivery of fresh vegetables to multiple hypermarkets in the area around Kuala Lumpur. The final decision to on delivery and the places as the destination of the fresh produce will be decided upon by the manager, based on time brackets. The manager will then use a pre-programmed delivery system, who will then determine the fastest delivery routes based on GIS modelling and final location of the hypermarkets. GIS combines digital maps with traditional databases and provides visual representations of information. Visualisation helps people approach the problems in the dimensions of space and time, with intuitive maps, instead of dimensionally restricted data tables and graphs. The overall conceptual framework of this thesis is detailed in Figure 1.1.



**Figure 1.1: Conceptual Framework** 

#### **1.6.1** The contribution of the GIS model in this research

The advantage of this new method is the consideration based on time patterns within two periods of time for calculating the fast route. This is more accurate, and the decision makers can rely on the results obtained from GIS model. Also, the average speed of the vehicle in the vicinity of the school zone has to be considered. For the mentioned roads it has to be considered that the average speed of the car should not exceed more than 30 kilometres per hour.

The main contribution of the developed model as compared to the existing models on route efficiency based on travel time relates to its capability in capturing data from different fields to solve the distribution problem. This diversity includes the data from social and spatial aspects as well as traffic and transportation fields. The novelty is to integrate the social data such as population and spatial data such as land use in developing the spatial database management system. The contribution is to connect this information with digital road networks to solve the distribution problem to select the most efficient and effective route for delivery purpose. The final output of this thesis is producing different maps integrated with road networks for decision makers to deliver their fresh vegetables.

## 1.7 Organisation of the thesis

This thesis will be organised into six chapters.

#### 1.7.1 Chapter 1: Introduction

The chapter provides an overview of the research. It describes the background of the study detailing the objective and scope of the study. It also provides justification and motivation for the research and contribution of the new model as well as the organisation of the thesis.

#### **1.7.2 Chapter 2: Literature review**

Chapter 2 reviews the literature and gives a brief explanation of various concepts of distribution planning, vehicle routeing problem, GIS and its applications. Also some explanation of the decision support system and its component. Previous research related to this study are discussed along with the different models used. After reviewing the different models, this study will develop a new model to solve the problem of routeing and find the decision for optimal distribution of fresh vegetables. It then talks about the spatial data and non-spatial data and explains how it will connect to the network GIS software for further analyses. It introduces different models in GIS and gives a brief explanation of each of them and the application of the mentioned model. It continues to describe the regression model and illustrates the previous works related to the application of the regression model in different fields of study.

## 1.7.3 Chapter 3: Research methodology

Chapter 3 examines the duration and location of this research. It identifies the tools and the data needed for analyses. It then discusses the data collection and the source of the data. The chapter then evaluates the data analysis and the steps for preparing the data to solve the distribution problems. The parameters and analysing the data are explained in the form of a flowchart.

#### **1.7.4 Chapter 4: Data preparation**

This chapter considers the data analysis, results and findings to evaluate and compare with each other for the best result based on the optimal cost. Data preparation for developing the system and data design to complete the attribute table for network analyses will be done to allow the GIS software to proceed for selecting the best roads. Types of the data and the format which is suitable for this study are also described. From the statistical point of view, different charts and figures are displayed. Creating a new street network dataset is described in various graphs and figures in ArcGIS software. This chapter extended the definition of GIS models and defined its use for this research.

## 1.7.5 Chapter 5: Results of the model and discussions

This chapter presents the data input and analyses the spatial data to obtain the best result for distribution. The final decision for selecting the fast routes is based on a calculation of minimising the drive time for delivery according to the company decision. Real-time traffic information for all of the major roads, especially congested roads in urban areas, is information necessary to produce a dynamic route planning system. The selected routes are chosen by the decision maker in several phases, based on the distance, driving time, average speed, population density, landuse, car volume and different time frames as impedances. The final GIS model has been used with traffic and transportation models for the purpose of various transportation system analyses.

#### **1.7.6 Chapter 6: Conclusion and recommendations**

This chapter presents a summary of the main findings and the conclusions drawn from the research. It provides the key findings of the research and possible routes for the delivery of fresh vegetables. The GIS model used for this study has identified the key factors for finding the optimal path in terms of time. This research demonstrated how a decision maker should choose the best orders to deliver fresh vegetables based on criteria and constraints in this research. GIS can both model transportation networks and integrate the association of network characteristics directly into a database. Measures of drive time arebased on average speed, time patterns indifferent times of the day, residential area, population, the total car volume and distance. Finally, it outlines the limitation of the research and includes suggestions for future studies.

#### **CHAPTER 2: LITERATURE REVIEW**

#### 2.1 Introduction

This chapter discusses the various concepts of distribution planning, vehicle routing problem and application of GIS in the distribution of fresh vegetables. GIS and transportation are discussed. This section is about the combination of GIS regarding the information data in transporting network and other demographic characteristics to aid in the decision-making process. At the same time, this section will explain how GIS is used to predict future congestion and to solve the problems of finding the fastest routes for the distribution of fresh produces to avoid traffic congestion and reducing time consumption. This section gives a brief explanation of the models and defines different types of models and selects the model which is most appropriate for this research.

Previous works will be discussed in this chapter. A brief explanation of GIS and different types of models in GIS areexplained. Previous works related to different GIS modelling is listed in this chapter. Various methods will be proposed for developing and solving the vehicle routing problem with previous research. Section 2.7discusses the modelling process as a conceptual design to better understanding the process of decision making and finally deciding on the fastest path based on objective requirements. Section 2.12 identifies the key factors of markets and gives a brief explanation of the concept of hypermarkets.

#### 2.2 Distribution planning

GIS technology is enhancing the status of geography in business and delivery service planning since geography holds the key to further enhancement of information management system. Most utilities implement a Distribution Management System (DMS) to increase service reliability, improve customer responses, reduce operational costs (mostly labour), and meet regulatory requirements. GIS-based routing solutions have proven to be one of the most popular and cost-effective GIS applications of recent years. Typically, these systems find the lowest cost path to a series of nodes on a linear network against a series of constraints (Tram, 1999).

GIS combines the computer system and spatial data linked to the map features. GIS can be used in many ways in transportation planning and management system, including route planning, traffic monitoring, etc (Memon, 2005). The main aim of the selected company as a case study for this research is to reduce the total distribution cost when improving the efficiency of delivery services in selecting the optimal routes. Most companies are interested in determining the best route, especially in the context of time reduction and the best cost/benefit ratio. This problem concerns the distribution planning system and delivery of perishable foods, such as fresh vegetables.

## 2.2.1 Transportation management

According to Memon (2005), transportation systems move people and goods. They consist of streets, highways, railroads and mass transit. Transportation systems involve issues related to land use, demographics, economics and environment. Transportation is one of the fastest growing fields in which GIS software is used. One reason is that GIS is the ideal information management and analysis tool for many aspects of the transportation industry including both the public and the private sectors.

Using GIS in the field of transportation opens up a wide range of possible applications, as diverse as the field of transportation itself. Whether these are cars and trucks along a road, trains along a track and ships across the sea and airplanes in the sky, all applications have one thing in common. They are the objects that move along a path in space. A GIS can provide a valuable tool for managing these objects in a spatially referenced context, viewing the paths as a transportation network (Memon, 2005)

A primary route is the route that a commuter tends to use during specific times. This group of variables tries to identify the effect of the driving experience of the drivers on the commuter's decision-making process (Li, Guensler & Ogle, 2005).

In other words, transportation is the process by which vehicles and other commodities moves from one geographic location to another and is the essential part of geographic phenomena and key factor for GIS(Nayati, 2008).

According to Tey (2003), the necessity for transport planning, in fact, is largely selfevident. During the day and night, people are engaged in a variety of activities which including working, shopping and delivery purposes. Hence, to take part in these activities people frequently need to travel between their origin and destinations for some distance, and they want to reduce their drive time as much as they can. A faster transportation system will reduce the delivery time but increase the transportation cost (Pimenta, 2010).

This study developed GIS model to solve the distribution problem of finding an efficient route to reducing the total drive time for the delivery of fresh vegetables.

### 2.2.2 Transportation issue in Malaysia

Transportation and delivery services will play a central role in the logistics system of corporations competing in the global market (Zailani, 2014). The need for effective and efficient transportation system are crucial for people involving in delivery services, hence overcoming the issue of traffic congestion. The transportation system in Malaysia including delivery services also faces various challenges, hence cannot offer the quality of services desired by companies involved in distribution planning for fresh produce (Khalid, 2014). This highlights the need for efficient delivery services to solve the distribution problem.

According to the research conducted by Indati (2013), the transportation sector is the second most energy consuming sector, accounting for 40% of total energy consumption after industrial sector in Malaysia (Indati, 2013). In this regard, the transportation sector

will be among the first sectors that need to be addressed to achieve the goal of solving distribution problems (Indati, 2013).

To achieve the goal of having efficient delivery routes as a primary objective, a GIS model will solve the distribution problem and find the fastest routes among the road networks. It will connect all the necessary data for decision makers to deliver fresh vegetables.

Traffic congestion is a common problem in Klang Valley and increases the drive time (Hashim, 2013). Also, the limited number of parking spaces in the city centre aggravates this situation.

Little attention has been givento parking and transportation issues in most cities in Malaysia. Such issues are important for future policy makers (Tey, 2003). Therefore, in this study, Sime Darby, which is the biggest supplier of fresh vegetables in Malaysia, is examined as a case study.

The private car is a much more convenient form of transportation than public transport in Malaysia, especially in the areas with low density when the land use pattern reflects the majority of the areas in Klang Valley (Ariffin & Zahari, 2013).

Experience tells that building more traffic capacity is not always the answer to solvingthe transport problem. In many cases in Malaysia especially in the big cities such as Klang Valley, new highways are jammed as soon as they are opened to traffic. For this reason, a transportation planning system is required to find the quickest delivery routes for fresh produce (Tey, 2003).

The transportation of perishable products is more complicated because of products deteriorating over time. Perishable products are affected by temperature variations and humidity or transportation time. Hence, it is very important that transportation time, handling, storage and other requirements are well planned maintain the products' characteristics when received by the consumers (Pimenta, 2010).

16

#### 2.2.3 Fuel consumption on transportation cost

Transportation costs such as fuel cost is significant component of total transportation cost for companies in the movements of fresh produce (Pimenta, 2010).Transportation cost (labour cost of the drivers, fuel cost, and cost of the trucks) are significant component of the total cost (Pimenta, 2010).Transportation costs are more important when the delivery of fresh produces involved, and special handling is required. For such a company, the transportation cost is about 25% of the total cost.

The price of fossil fuel has been on the rise over the past decade. Vehicle fuel consumption is approximately 30% greater under heavy congestion conditions, especially during peak hours, with the result being longer delay time (Shokri, Chu, Mokhtarian, Rahmat, & Ismail, 2009). According to this study, they found that availability of pre-trip information to the driver enhances the driver to better select the best routes to avoid traffic congestion and allows the commuters to make better-informed transit choices and save drive time (Noonan & Shearer, 1998). Also, the decision maker can improve the network traffic performance and manage the transportation plan as efficiently as possible (Shokri et al., 2009).

The price of fossil fuel has increased from 18 US\$(00)/bbl in 1990, to 80 US\$(00)/bbl in 2007. Although the fossil fuel prices continue to rise at the present, this trend is wrought with uncertainty in the near future(Rout et al., 2008).

According to the statements above for fuel condition, energy saving and technology adoption aroused to solve the vehicle routing problem and improve the distribution planning system, taking into account the optimal routes for reducing the overall costs.

In considering fuel consumption when selecting the best routes for travelling, research was conducted to find the best travelling path regarding fuel consumption as an important element in the process of improving energy intensity of the transport sector (Shokri et al., 2009). This work tries to understand the effect of different levels of traffic on the delay time for finding the best routes in static and dynamic traffic networks by considering drive time and fuel cost. The programme approach in basic visual guides the users to the best routes after the user picks the origin and destination. It also shows the total distance in programme application created in this study. It is concluded that fuel consumption is often higher in a congested area but shorter route. Sometimes a longer route maybe less congested and is more fuel economy (Shokri et al., 2009).

The mentioned research is inadequate in the sense of applying spatial dataon network analysis for calculating the fastest routes. The car volume for each road network is not considered as a variable for transportation management.

#### 2.3 Vehicle routing problem (VRP)

Vehicle routing was first formally studied as a mathematical research problem in the late 1950s (Dantzig & Ramser, 1959). The work of these researchers regarding truck dispatching and routing helped to formalise the problem that is today known simply as the VRP (Rice, 2005). One of the duties of municipal in every country is to provide a variety of services that relates to moving people and goods to the benefit of the community. The majority of these problems can be modelled as VRP. In its most basic form, the VRP consists of designing the cheapest distribution pattern to service a number of geographically scattered customers with known demand from a single depot, using identical, capacitated vehicles travelling at a constant speed (Bräysy, Dullaert, & Nakari, 2009).

The VRP consists of defining an optimal set of routes for vehicles to serve a set of customers. There are many variants of the problem, but the common objective is to minimise the total travel time (Rice, 2005).

The other definition for VRP which is traditional VRP is a set of routes originating from a single depot should be selected to visit all the required customers and finally the total distribution cost is minimised. Each customer has a given demand, and the capacity
of the vehicle is fixed. The problem can be extended in other ways such as to consider the time window to visit each customer on time and second to minimise the number of routes to be visited by customers (Andersson,2008).

In different routing applications, the given vehicles may take more than one route during a specific time. This happens when the customer demands increase with respect to vehicle capacity. In this case, one vehicle may be sent to more than one route hence reducing the number of vehicles in order to reduce overall cost (Battarra, Monaci, & Vigo, 2009).

In this study, we consider the problem of finding the efficient and effective routes for the delivery of fresh vegetables to hypermarkets from a single fresh produce company by using the spatial database management system and the output of the study will propose different maps showing the routes for decision makers to choose and minimise the drive time.

## 2.3.1 Prior studies on VRP

Several studies have addressed the distribution of fresh produces, related to factors, issues and VRP.

Osvald & Stirn (2008) conducted research on a vehicle routing algorithm for the distribution of fresh vegetables. They focused on the distribution of fresh vegetables in which the perishability represents a critical factor. This particular problem was formulated as a vehicle routing problem with timewindows and time-dependent travel times (VRPTWTD) where the travel times between two locations depends on both the distance and on the delivery time of the day. To minimise the overall distribution cost, the objective function must additionally consider the loss of quality of the load. To solve the problem, Osvald & Stirn (2008) used a mathematical algorithm and applying the heuristic method based on tabular search. The limitation of the mentioned research is that the author used the mathematical equation instead of spatial data for finding the best

routes. In addition, the time pattern for selecting optimal routes is not applied. The study is not based on GIS modelling to visualise the results of optimal routes for delivery purposes. The application of Network Analyst is not appropriate for this research to approach the problem.

In the field of VRP, several researchers have created a computer package to design dispatching routes in the meat industry named as RutaRep. This study presents a computer program that has been developed to design the dispatching routes of a medium-sized meat company in Spain. The objectives of this research are to improve the quality of delivery services to customers by minimising the lateness and also minimise the total travel distance which is travelled by the vehicle in one route (Belenguer, Benavent & Martínez, 2005). For this research, a number of heuristic algorithms were implemented to solve the routing problem. The GIS application is not applied for pathfinding and modelling.

Tarantilis & Kiranoudis (2001) analysed the distribution of fresh milk. They formulated the problem as a heterogeneous fixed fleet VRP. It means that a VRP of vehicles has different capacities. The objective of this research is route scheduling to find a set of routes that minimise the total cost of delivering fresh milk from the single factory to supermarkets. This study is based on mathematical algorithm with technical engineering background to solve the specific problem. The GIS modelling is not applied to illustrate the optimal routes for the best distribution of fresh produce. The GIS modelling does not identify the factors that have the most influence on distribution activities.

Tarantilis & Kiranoudis (2002) also presented a real life distribution problem of fresh meat in an area of the city of Athens. They adopted a new stochastic search metaheuristic algorithm belonging to the class of thresholdaccepting algorithms for solving VRP. The objective is to frame a set of routes that minimise the total travelling by vehicle and vehicle operating cost for distribution.

20

Research has been conducted on efficiently and scheduling home care, transportation of the elderly, and home meal delivery services (Bräysy et al., 2009). The objective of this research is to solve routing problem to explore the cost-saving of the optimal route and to improve the efficiency of arousing problem in the future. There is research in the dairy industry for milk collection (Butler, Herlihy & Keenan, 2005). Adopting Information Technology (IT) in this research will facilitate the data collection, manipulation and building the DSS to support logistics management in the milk collection sector. The main objective is to reduce the transportation cost in order to improve the price per litter that the dairy company can offer to frames. The research focuses on GISbased DSS allows the scheduler to interact with the optimisation algorithm for selecting the best routes for the collection of milk from the farmers. The transportation and delivery cost including labour and fuel cost for this company is approximately 20% to 25% of the total cost.

The problem with all of the mentioned researches is that mathematical algorithms were employed to solve the problem of finding the optimal route while neglecting the spatial data as well as traffic data for selecting the routes, which depends on drive time and other parameters in GIS modelling. The social data such as population is not part of their calculation in determining the fastest routes. The land use as a geographical aspect is not included in their model to solve the distribution problem of finding an efficient delivery route for their product.

On the other hand, one research has proposed a decision support system (DSS) and employed a metaheuristic algorithm called Bone Route, for solving the open VRP (OVRP) (Tarantilis, Diakoulaki & Kiranoudis, 2004). The algorithm applied in Tarantilis' study determined the optimal set of customers (Nodes), with the information of the known distance between them. Another research presented an application service provider, to be used for central food markets; which coordinates and disseminates tasks

and related information for solving the VRP. For solving the problem of Vehicle Routing, the metaheuristic technique was used (Prindezis, Kiranoudis & Marinos-Kouris, 2003). Another study has extended a VRP for the randomness of delivery process of fresh produce. The objective was to minimise the transportation, inventory, energy costs and penalty costs for violating time-windows. The model is adopted by 'stochastic VRP with time-windows' (SVRPTW) to obtain optimal delivery routes (Hsu, Hung, & Li, 2007). There is research on the exact algorithm for the solution of the basic version of VRP where the vehicle capacity is the only constraints to be considered. The objective is to minimise the total cost to serve all of the customers based on mathematical equation (Toth & Vigo, 2002). A group of researchers proposed a nonlinear mathematical model with vehicle routing and time-windows for perishable food. The quality of fresh produce will decrease rapidly once they are produced and will decay during the delivery process. The income of suppliers depends on the quality and condition of the produce when received. So timely production and improving the delivery service by choosing the optimal routes significantly affects the supplier's income. The weakness of the research mentioned above is that using mathematical equation as well as heuristic algorithm while improving the delivery services and finds an efficient transportation system is the main goal (Chen, Hsueh, & Chang, 2009). The GIS model is not used to find the best route for the transportation of fresh produce and the Network Analyst tool based on best routes is not considered for analysing the spatial data.

The reviewed researches use VRP which cannot handle the social data such as population. The significance of the current model is the capability to select efficient routes based on population density. It has been observed from the statistical analysis that the population density has a significant effect on drive time in the delivery process. Furthermore, the land use as a geographical parameter is not considered when calculating the minimum drive time for transportation planning. The spatial database management system with two categories of data was developed to solve the distribution problem to determine the most efficient delivery routes.

## 2.3.2 GIS application regarding some influential factors and issues

On transportation safety management, a group of researchers proposed the GIS technology to find the best set of routes for nuclear waste transportation. They considered the drive time factor for vehicles along the routes. They calculated the distance divided by the average speed of the vehicle to find the best drive time in ArcView software (Chen, Wang & Lin, 2008) The concentration of this research is more focused on safety transportation and environmental aspect in transportation of dangerous goods. The time pattern for the best delivery approach is not considered as the part of developing the model. The car volume is not considered as a factor which effects on drive time for transportation purposes.

In the field of VRP a research proposed a system as a spatial DSS that geocodes and maps customer's locations to the vehicle (Tarantilis et al., 2004). The role of GIS the background information system of DSS and the tasks are manipulating and organising all the large values of spatial data through the database system and performing a cartographical role related to VRP. Moreover, one research hasconducted a study on the application of GIS in the transportation field. The objective of this thesis is the application of GIS in transportation to find the optimal road among asset of routes. The optimal routes are based on a complex road network in the aspects of time, length, speed,etc.(Memon, 2005). A project was developed that focused on determining the best route between two destinations based on distance and travel time. The goal is the application of GIS to determine the quickest and fastest route between two locations. This study is based on vector-based model and Arc Viwe with Network Analyst that would be the best GIS software solution in this case (Echols, 2003). The weakness of the research above is that any GIS model was considered in their work to identify the key factors that affect the best

delivery of fresh produce. The study focused on the quickest routes for transportation regarding the aspect of time, length and drove time. For the current research, two categories of spatial and non-spatial data were collected to develop the database management system for solving the distribution problem.

A prominent researcher has proposed a Decision Support System (DSS) that enables distributors to approach intra-city VRPs with time-windows using the appropriate computational method that contains information about traffic and spatial data to achieve the goal (Ioannou, Kritikos, & Prastacos, 2002). In the field of GIS, research was proposed based on DSS to assist fire managers in determining the fastest and safest access routes from the firefighting headquarters to fire areas. The travel time variable was computed based on road length and average speed of the fire truck which is depending on the road types and road status (Akay, Wing, Sivrikaya & Sakar, 2012). The mentioned study was inadequate in the sense that to solve the distribution problem to find the efficient routes, only the algorithms approach and traffic data wereapplied.

| Author<br>Variables | Wen Chen (2008) | Akay (2011) | Echols (2003) |
|---------------------|-----------------|-------------|---------------|
| Drive Time          |                 |             | $\checkmark$  |
| Average Speed       |                 |             |               |
| Distance            |                 |             |               |
| Road Length         |                 |             |               |
| Type of the Road    |                 |             |               |
| Road Status         |                 |             |               |
| School Zone         |                 |             |               |
| Traffic Light       |                 |             |               |
| Rush Hour Traffic   |                 |             |               |

Table 2.1: Variables adopted in GIS application

## 2.3.3 Parameters regarding transportation and delivery of goods

A thesis titled Transportation Optimisation Model of Palm Oil Products for Northern Peninsular Malaysia sought to determine the optimal distance for the selected set of routes, and minimise the transportation cost. Shumsudin (2008) developed an integer mathematical programming model to solve the transportation problem. The GIS application was not applied. The inadequacy of this study is that ArcGIS based on Network Analyst is not used to approach the transportation problem.

Faulin (2003) studied the implementation of the mixed algorithm (MIXALG) procedure in routes management in agribusiness. The applied algorithm helps solve some VRPs related to canning companies. The aim of MIXALG is to optimise the supply transportation cost, and the distribution expenses of a canning company (Faulin, 2003). A multi-objective GIS for the best route selection of nuclear waste transport as dangerous goods was proposed. The objective is to minimise the travel time for selecting the best routes for nuclear wastes transportation. One of the factors dealt with in this research is the risk probability of conducting a system to avoid problems during transportation. The point is that it has been focused on the safety of transportation for nuclear waste (Chen, Wang & Lin, 2008).

In industrial engineering, researchwas conducted with the objective of determining the optimal routes and schedules for newspaper delivery to minimise the deliver cost, while the total delivery time decreases. For solving the routing problem, a practical approach was used based on the heuristic algorithm and digital map (Song, Lee &Kim, 2002). In economic analysis, a researcher used the stated preference method for determining the value of time in traffic situations taking into consideration a parallel road network. The purpose of this research is to find the social and economic impact of the various transport policies (Avelar, 2008).

One research sought policies that would enhance the efficiency of the transportation supply, with particular concern in improving environmental, socioeconomic and budget factors. The objective is to minimise the cost of externalities based on road traffic. For the implementation of this strategy, a set of the mathematical model has to be taken into consideration (Arampatzis, Kiranoudis, Scaloubacas & Assimacopoulos, 2004). The limitation of this study is that the GIS application modelling is not applied for practical approach to solving the routing problem. However, the application of network analysis is not applied to improve the efficiency of transportation for this research.

# 2.4 Geographic information system (GIS)

GIS is designed for collecting, storing, manipulating, assembling, analysing, displaying, and managing of geographic spatial information or an object and phenomena where geographic location is an important characteristic or critical to analysis (Aronoff, 1989a). This technology has developed so rapidly it now accepted as an essential tool for the effective use of geographic information. GIS provides the decision maker with a powerful set of tools for the manipulation and analysis of spatial information. GIS technology has been traditionally associated with physical infrastructures and proximity analysis (Aronoff, 1989b).

Two researchers well tell the history of GIS development. They describe the development of GIS from the first application in Canada in the 1960s to the present day, highlighting significant developments such as the arrival of supper-mini computers and the incorporation of relational database management system technology in the 1980s (Coppock & Rhind, 1991). Much of the technical advance has been made in the United States, especially with regard to the development of proprietary systems (Clarke, 1991).

In some studies, GIS works as a decision support tool. Data are collected, entered and organised into a database management system, and further analysis is performed to help the decision makers. The spatial analysis in GIS sometimes requires more data for more accurate results. The results go through several iterations such as collection, organisation, analysis and assessment steps before a final decision reached. Figure 2.1 illustrates how GIS will be integrated into the decision-makingprocess of organisation. Efficient use of GIS depends on a set of protocols and integrationinto the data collection, analysis, decision and action loop of organisation (Bolstad, 2012).



Figure 2.1: GIS organization (Bolstad, 2012)

GIS technologies primarily provide the capability of spatial and network database management as well as technology management. Computer-based GIS integrates data from diverse disciplines and various formats to generate useful information about physical features of an area such as street networks (Memon, 2005).

GIS has been instrumental in addressing some of our most pressing social problems. It can save billions of dollars in the delivery of governmental and commercial goods and services. GIS helps in the day-to-day management of many natural and man-made resources such as transportation networks. We need GIS due to increased population and consumption that have reached a level such that many resources including air and land are placing substantial limits on human action.

The potential of GIS is best illustrated through an examination of what it can offer the geographer or planner, whether public or private sector, regarding spatial analytical functions (Clarke, 1991). One of the potential areas of GIS applications is the market and business application. Markets and businesses require detailed information on consumption patterns and the geographical location. It is also important to know how to

distribute the fresh goods through those supermarkets which are linked with the network information of roads (NCGIA, 1990). Using the GIS technology wisely can help the decision makers to manage their goods to required locations in the least amount of time and lowest transportation cost.

## 2.4.1 Transportation geography

Transportation geography seeks to understand the movement of people and goods between locations and the spatial organisation of the systems that facilitate these exchanges. Furthermore, the use of GIS to tackle questions in transportation geography is widespread, and so-called geographic information systems for transportation (GIS-T) have become a fruitful area of research and application.

According to Horner and Downs (2007), technology has greatly influenced the way in which transportation geographers conduct research and also has presented new issues for them to study. Today, GISs are widely used to address problems in transport geography. This is because the functionality of GIS technology has vastly improved over the past decade, allowing a broader range of transport research questions to be considered. At the same time, computing technology has become more affordable, and the proliferation of the commercial Internet has provided GIS users with a convenient and inexpensive way in which to exchange information. Finally, the current environment can be characterised as data rich because a substantial number of government and private firms use GIS in their work and create spatial databases that are shared regularly and freely among transport researchers throughout the world.

Transport geography emphasises on spatial dimensions affecting cost components. These are called frictions in transporting geography. The present trend of increasing transport costs because of higher fuel prices will strengthen the possibilities for using other ways for optimal distribution (Henstra, 2006). The most significant reason for applying the network analysis and route planning in transportation is that businesses are interested in determining the best route to save cost and time and ultimately achieve the best cost-benefit ratio (Memon, 2005). This research focuses on hypermarkets for delivery purposes and determining the suitable routes to reach the final selected destinations. The aim of this research is to conduct a model based on GIS for selecting the best distribution routes for the delivery of fresh vegetables. The consideration of this research is to evaluate the statistical analysis to obtain the most significant variables affecting drive time.

# 2.4.2 Geographic information system for transport

The term GIS-T, which stands for GIS for transportation, emerged in the 1990s. GIS and transportation have been developed as two independent systems originated from the two disciplines geography and transportation. The range of models that needs to be employed has expanded rapidly, and the integration of transport models and technologies such as GIS has become a major requirement in any process of transport planning (Hensher & Button, 2000).

Transport problems are getting more and more dynamic due to changes in our complex social, economic, and physical world. Decisions must now take account of social, economic and environmental pressure, the interaction between them. GIS are proving to be effective in integrating the data needed to support transport modelling and data management (Hensher & Button, 2000).

Spatial relations among points, lines and polygons represent a street network. Streets are presented as lines between points, whereas blocks are defined by a boundary composed of lines. In addition, one of the most attractive features of a GIS is the possibility of attaching the addresses of streets so that the street can be traversed in the direction desired. This is a basic component in the application of a GIS to solve different routing problems on the transportation network. In the transportation context, three classes of GIS models are relevant as following (Goodchild, 2000).

**Field models,** or representation of the continuous variation of a phenomenon over space. Terrain elevation is as an example for this kind of model.

**Discrete models,** according to this model, a discrete entity is like (points, lines or polygons) populate space. Highway rest areas, toll barriers, urbanised areas may use this model.

**Network models,** to represent topologically connected linear entities (such as roads, rail lines, or airlines).

Considering all three models which may be useful in the field of transportation. The network model plays the most prominent role amongthese three application domains because of the single, and multi-modal infrastructure network is the fundamental of supporting the passengers and freight movement (Thill, 2000b).

Given the spatial data, GIS also helps determine the best route for delivery. GIS can also aid in the collection of data. "A major limitation in the use of transportation models is the difficulty of obtaining adequate data for them. GIS are a means of addressing this need"(Wong et al., 1997).

The consideration of this research is on developing a GIS model to analysis the spatial data to solve the routing problem in road networks. The data includes all spatial and non-spatial data to support the distribution problem. A database management system was developed to facilitate the data and propose the fastest routes based on the variables that affect drive time.

## 2.5 Fresh vegetables

Because of the perishable nature of the products, spending time for delivery plays a significant role in the supply management according to the optimal route. It may affect a perishable product in one or both of the following aspects: 1) Physical decay, which

affects the effective supply of the product 2) Value deterioration, which affects the demand for the product (Xiaolin, 2006).

Freshness is as one of the most influential variables influencing consumer demand and their decision to purchase fresh produce. Freshness is also one factor that attracts the consumers to buy the product (Chamhuri & Batt, 2013).

From the moment of harvest, agricultural products have a limited life because of loss of quality during the period between harvest and consumption; even optimal conditions are used during distribution (Sloof, Tijskens, & Wilkinson, 1996).

Fresh produce often deteriorates due to extended travel times and frequent stops to serve customers, during the delivery process. It is difficult to effectively manage cold chain distribution and ensure maximum freshness during hot or humid weather. In addition, perishable food usually has a short shelf life; thus, timely delivery of perishable food not only significantly affects the delivery operator's costs, but also the revenues of retailers (Hsu et al., 2007). Temperature control and modification of atmosphere are two important factors in the prolonging the shelf life of fresh produces (Arvanitoyannis & Vaitsi, 2007).

All activities related to delivery services depend on the level of services of traffic and transportation system available in the zoning areas (Tey, 2003). Quality is considered a huge issue when talking about fruits, vegetables and flowers (Pimenta, 2010).

One of the most important concerns regarding fresh vegetables is preserving their nutritional value during transportation. Fresh vegetables are an example of perishable goods; so selecting the fastest and optimal set of routes for delivery is the most significant factor (Osvald & Stirn, 2008).

## 2.5.1 Distribution planning for fresh vegetables

The following statements outline the distribution planning for fresh vegetables for the purpose of quick delivery of fresh produce.

- Identify the major industry supplier involved in the fresh produce system.
- Identify the global food system of fresh produce items from the main location through distribution to the hypermarket.
- Describe the hypermarket in different location and try to have the optimal distribution for the end users.

Nowadays, because of competition between marketplaces, retail buyers are increasing their purchases from the point of deliveries, which provide the value-added services demanded by increasingly sophisticated consumers (Clarke et al. 1991). Fresh produce with a higher consumer demand has higher revenue for the hypermarkets as well as the lifetime value of economic activities (Bohari, Rainis, & Marimuthu, 2012).

The market demand for the fresh produce depends on its freshness level as well as the selling price of the distributor. The manufacturer has to determine his wholesale price based on its impact on the order quantity of the distributor, whereas the distributor has to determine his order quantity and selling price based on the price of the manufacturer, the likely loss during the transportation, the freshness level of the product, and the possible market demand.

The transportation of perishable products ismore complicated because of products deteriorating over time. Perishable products are affected by temperature variations and humidity or transportation time. Hence, it is critical that transportation time, handling, storage and other requirements are well planned to maintain the products' characteristics when received by the consumers (Pimenta, 2010).

## 2.5.2 Distribution planning for fresh vegetables in Malaysia

Fresh food has the potential to be a strategy of sustaining the hypermarket lifetime value and to increase the profit as well. Most of the hypermarket use fresh fruits as their strategy for consumers to revisit and repurchase in their stores. Fresh fruit is the main advertisement strategy for them to keep their regular customers. Fresh fruits and vegetable sales in Malaysia are doing remarkably well despite the economic downturn (Bohari et al., 2012).

According to Kamil (1996), Malaysia produces a wide variety of vegetables, of which 50 species are grown commercially. In Peninsular Malaysia, the area under vegetables increased sharply from 7735 ha in 1983 to about 25,974 ha in 1994, which resulted in a production increase from 145,000 tonnes to 399,000 tonnes in that period. Fruit vegetables recorded the biggest increase in both cultivated area and production.

In Malaysia, the hypermarket business is the top most segment of the retail business. In the rapidly changing Malaysia hypermarkets business, a retailer's fresh produces critical success factor. Fresh food produce is critical for store visits and shopping frequency that attract consumers to repeat shopping activities as well as the opportunity for retailers to engage with their customers (Bohari, Rainis & Marimuthu, 2012).

Kamil (2006) stated that more vegetables are consumed in urban areas rather than in rural areas. Additional demand for fresh vegetable is expected in Malaysia's rural areas as urbanisation proceeds at a faster rate through the 21<sup>st</sup>century. In most real-world distribution problems, it is important to consider the fluctuations of the travel time in the solution tithe problem. Travel time plays an important role in the distribution of the perishable goods since its fluctuations may extend the time that the goods spend on the vehicles.

The National Agricultural Policy (1992-2010) outlines a strategic plan for the vegetable industry to attain a targeted self-sufficiency level of 125% by 2010. Production is expected to increase by 9.3% and 7.1% per annum in 1991-2000 and 2001-2010, respectively, to reach 2.7 million tonnes by 2010. Demand for quality vegetables is expected to rise as a result of consumers' greater awareness about health and nutrition (Kamil, 2006).

Fresh vegetables are sensitive produce in the agricultural sector a. The consumer demand depends on how fresh vegetables are delivered to hypermarkets. The quality of the produce is important for the consumer. In Sime Darby, the vegetables are cultivated in a netted greenhouse on a five-hectare farm in Seremban, Negeri Sembilan, including temperate lettuces, lettuces, subtropical leafy and chillies.

## 2.5.3 Prior research on transportation and delivery of fresh produce

The reviews of extant literature reveal that various studies have been conducted in the past on the distribution of fresh produce. In this case, there is research on the distribution of fresh vegetables. Solving the problem required these of a mathematical algorithm and applying the heuristic method based on tabular search. This method is a mathematical equation (Osvald & Stirn, 2008). It does not require any spatial data such as road network data, and the calculation is not based on vector-based data. The transportation variables such as car volume is not capable of this research.

In the transportation of fresh vegetables and fruits produce, a researcher has conducted research with consideration of quality and extend of the road infrastructure. He focused on the time-distance relationship in determining whether the produce reaches the market fresh and in marketable quality (Dannenberg, Kunze, & Nduru, 2011). SpatialAnalystwas usedfor working with the raster dataset. The analysis was based on the different velocities of three kinds of the road in the region. This study did not consider the time pattern for transportation. The vector data is not applied finding the suitable routes for transportation. The weakness of this study is that the car volume for road networks is not considered as a factor to estimate the best delivery routes of fresh vegetables.

Belenguer et al. (2005) conducted research on dispatching routes in the meat industry. The data analysis is not based on vector-based data. There is a lack of spatial data and GIS model to solve the transportation problem. This research uses the mathematical algorithm in solving the routing and dispatching problem in the meat industry instead of spatial data such as road network analysis.

Tarantilis and Kiranoudis (2001) analysed the distribution of fresh milk. Tarantilis and Kiranoudis (2002) also solves the distribution problem on fresh meat. To solve the routing problem, this study is based on a mathematical algorithm with technical engineering to solve the specific problem. The GIS model is not applicable for the research above. The study lacks spatial data for analysing the data.

Bräysy et al. (2009) researched the efficiently scheduling home meal delivery. Butler et al. (2005) havea research on the dairy industry, which is mainly in the food sector and milk collection. This paper considers how a GIS-based DSS allows a scheduler to interact with the optimisation algorithm to plan milk collection routes. The weakness of this research is that spatial data such as road networks are not applied for solving the transportation problem. The traffic data is not used for predicting the drive time for selecting the optimal routes.

## 2.5.4 Malaysian's consumer behaviour

Malaysia, as in many Asian countries, with rapid socioeconomic development, has increased in the standard of living and changes in lifestyle. This has changed the structure of agrion production and in meeting the demand by society. Rezai (2011) indicates that the factors influencing the consumer behaviour are the quality of production and the price (Rezai, Mohamed, & Shamsudin, 2011).

Malaysian consumer lifestyle has changed due to rising income and education levels. Hypermarkets are attracting consumers with their one-stop and all under oneroof concepts. Malaysia's population is young where 35% of the population is below the age of 15, 61% are in the working age category, and 3.7% are above 65. The median age of population at 22 years reflects a continuing young population. This group is globally influenced and within Malaysia prefer higher value products including fresh food. There is also an increase in contracts between farmers and supermarkets (Shamsudin & Selamat, 2005).

The increasing household income can be directly translated as stronger purchasing power, suggesting that the Malay population could be considered the leading market growth sector in Malaysia (Sheng et al., 2008&Rittgers, 2013). Malaysians spend a high percentage of their household income on fresh produce, groceries and personal care items. For this reason, it ranks third out of ten major economies in the Asia-Pacific Region (Shahnaei, 2012).

Regardingapplication of GIS in proposing a model based on traffic congestion with respect to the time, Alivand, Alesheikh, and Malek (2008)proposed a business model. This model will help the traffic manager to have real-time traffic data in different times and locations in order to find the optimal path. It also helps the user to obtain the best routes in urban trips using GIS analysis based on network analysis. In this research, for calculating the optimal path between two destinations, a new vision based on partitioning space-time for solving shortest path problem is proposed. This vision uses heuristic methods in addition to optimisation techniques. The weakness of this study is that there is no spatial data for analysing the data. It lacks additional parameters to estimate the exact drive time for delivery services. The GIS model and data evaluation are not appliedin this research.

Leonard and Oliveira (2000) have research on accessibility in highway capacity manual. His research is the measure of accessibility by calculating the speed and improves the accuracy of capacity in regional travel demand forecasting model. The weakness of this study is that the GIS application is not used to calculate the optimal path between the origin and destination.

Winyoopradist and Siangsuebchart (1999) developed a model on Network Analyst based on vehicle speed at different times of the day. The network analysis tool has a capability that calculates the shortest travelling time between two locations by specifying the starting time. They focused on the speed pattern of some roads in the different time of the day and day of a week. Ichoua, Gendreau, & Potvin (2003)developed a model based on time-dependent travel speeds. The proposed model is performed in a static and dynamic setting using a parallel taboo search heuristic. The speed of the car is a major factorfor determining and selecting the fastest routes as a required goal. The weakness of this study is that the spatial databased on the road network is not used for analysing and evaluation of final data.

Cairns (1998) researched GIS to find the best travel path to food shopping centres. He considered the issue on how to reduce travel for food shopping by using TransCad. The objective was to investigate new alternative ways in which food shopping might be organised in order to reduce the traffic. In this case, the spatial data for calculating the optimal routes is not proposed. The GIS model to identify the key factors is not considered for determining the drive time in transportation planning. The traffic issue is not discussed in delivery and finding the optimal set of routes.

## 2.5.5 Drive time modelling approach based on several parameters

Statistical reports show that there are 184 cars per 1000 people in KL hence 45 for Hong Kong, 106 in Singapore and 153 in Bangkok. In Kuala Lumpur, the modal share of public transport is about 20% of total trips. In contrast, in Hong Kong and Singapore and Tokyo the modal share is about 70% and in Bangkok, Manila and Jakarta the modal share of public transport is between 40 and 60% of total person trips. Kuala Lumpur has a higher per capita than Bangkok, Jakarta and Manila but a lower modal share of public transport (Schwarcz, 2003). Therefore, an efficient transportation system will decrease the traffic density in the main road in central of KL and the drive time for the vehicle will decrease consequently. Tatomir, Rothkrantz, and Suson (2009)researched computing the shortest path while being able to have a good prediction model of expected congestion on dynamic changes in the network caused by accidents. The GIS model is not applicable for the research above. The weakness is that the spatial data such as road network is not used for calculating optimal routes for delivery services.

Lam (1997) presents a heuristic approach to routing bus transit by considering spatial analysis techniques with analytical network functionality of GIS. The focus is to solve the routing problem by finding an optimal path with either the shortest travelling time or distance by applying the Dijkstra's algorithm used in Arc/Info and transCAD. The development of bus network is based on bus spaces and transit demand that is affected by land use, population and employment distribution as well as government policies on land development and public transportation. The average speed for road network is considered for estimating travelling time. The weakness of the above study is that the time pattern approach is not considered for drive time calculation in different periods. The other point is that the car volume factor is not an impedance to predict real-time transportation.

# 2.5.6 Application of the dijkstra model

The definition of Dijkstra's algorithm is to find the shortest path from x to y. The practical approach of this method is the weight or length of a path that is calculated as the sum of the weights of the edges in the path. In the initial stage, it chooses the first minimum edge, store this value and adds the next minimum value from the next edge it selects. It is also known as the single source shortest path problem. It computes the length of the shortest path from the source to each of the remaining vertices in the graph.

Leonard and Oliveira (2000) implemented Dijkstra's Algorithm to compute the shortest travel times between any single origin-destination pair. The primary highways and interstate are the only data including for this research to simplify the network and speed computation. This study demonstrated the application of proposed area-wide service measures within a real-time environment.

Alivand et al. (2008) also adopted Dijkstra's Algorithm in his research. He proposed the spatial analysis of finding the optimal path between two specific locations in a network considering the changing of traffic congestion continuously. Several features have been evaluated in this study and anew method based on partitioning space-time was presented. Wang et al. (2006) also adopted Dijkstra's Algorithm to develop an efficient system to compute the shortest path in real road networks.

This algorithm has limitations in terms of many factors affecting drive time to solve the distribution problem. This equation only calculates the shortest path among two nodes by consideration only mathematical algorithm. Another parameter such population density as a social data and land use as a geographical datais not considered as the part of determining the efficient routes. It this research the GIS regression model was applied to combine the spatial and non-spatial data to evaluate the most significant variables effect on drive time.

## 2.6 Route planning system

Computing optimal routes in road networks is one the most useful applications of algorithms to accelerate route planning queries. The following research is based on route planning system.

## 2.6.1 Route planning system based on GIS

Gohari (2010) developed a model for analysing the shortest path and closest facilities. He considers several factors such as speed limit of a car in the road network. There is no GIS model to approach the problem of path finding for delivery services. The traffic volume for road networks is not considered in this research. It requires more traffic data of different times of the day. For this reason, the calculation of pathfinding is not efficient. The parameters applied for this research are speed limit of vehicle, distance and drive time.

Sadeghi-Niaraki, Varshosaz, Kim, and Jung (2011) worked on the road network in the GIS for network analysis. Sensitivity analysis is used for model evaluation to finding the shortest route. The selection of variables as impedance modelling is based on the characteristics of road network including human factor. The weakness of this study is that GIS software does not evaluate the spatial data. The drive time in different times of the day is not considered in this study.

Papinski (2011) developed and applied route choice analysis. This GIS-based analysis generates variables describing route characteristics. The contribution of the model is the capability of generating routes based on land use characteristics and population density. The model was developed based on significant variables influence drive time to solve the distribution problem of finding fast routes for fresh produce delivery.

# 2.6.2 Route choice for travelling

In this section, the following researches consider the information for the purpose of expressing the driver route choice behaviour with multiple sources of traffic information. The experts realised that the focus of transportation had been shifted to the application of new technologies to achieve more efficient use of transportation systems. According to Li (2005), the ultimate route choice decision is the result of many factors such as travellers' socioeconomic characteristics such as age, gender, income, personalities, habits and in other hand trip characteristics such as time and location, flexibility in arrival time, availability of alternative routes and traffic conditions.

The following studies focus on route choice behaviour for selecting the suitable route in travelling from the point of origin to the destination. In this regard, Selten, Chmura, Pitz, Kube, and Schreckenberg (2007) researched the route's choice behaviour. The two variables considered in this research are travel time and its reliability in taking the suitable road. Li (2005) examined the morning commute route choice behaviour of some drivers in order to describe how these drivers tend to behave in selecting the suitable routes based on spatial data collection. Variables used in this research are travel time, distance and the average travel speed. Cascetta, Russo, Viola, and Vitetta (2002) examined different criterion in order to find the suitable path. The mentioned criteria in this study are 1) freeflow travel time; 2) physical distance (minimum); 3) congestedtravel time. This paper proposed a model based on simulating route perception by drivers. Lam & Small (2001) studied the value of time in their research. It has been evaluated and confirmed that the actual traveller behaviour of the drivers isbased on the data in the real pricing scenario. The observation is on the people who face a choice between two parallel routes, one free and congested, while the other with a time-varying toll.

Chen, Chang, and Tzeng (2001) conducted research on developing the individual behaviour based on mechanisms for determining the influential criteria for driver's route choice decision making. The possible objectives that most drivers consider before selecting routes include the minimal travel cost, fastest driving speed, optimal safety and comfort, lowest risk factors, the most familiar route, and the ones having the least number of stops. Hato, Taniguchi, Sugie, Kuwahara, and Morita (1999) created a model for route choice, based on the information for the purpose of expressing the driver's behaviour with multiple sources of traffic information for route selection. The variables mentioned in this study are traffic time, queue time as traffic information, and distance as route characteristic. Table 2.2 identifies the key factors in route choice behaviour by each author.

| Authors                            | Selten | Li     | Cascetta | Lam and | Chen   | Hato   |
|------------------------------------|--------|--------|----------|---------|--------|--------|
| Variables                          | (2007) | (2005) | (2002)   | (2002)  | (2001) | (1999) |
| Travel Time                        | ×      | ×      | ×        | ×       | ×      | ×      |
| Travel time reliability            | ×      |        |          |         | ×      |        |
| Travel distance                    |        | ×      | ×        |         | ×      | ×      |
| Time of day                        |        |        |          | ×       |        |        |
| Trip purpose                       |        |        |          |         | ~0     | ×      |
| Route familiarity                  |        |        |          |         | ×      |        |
| Congestion                         |        |        | ×        |         |        | ×      |
| Level of service                   |        |        | ×        |         |        |        |
| Safety                             |        |        |          |         | ×      |        |
| Volume                             |        |        |          |         | ×      |        |
| Directness                         |        |        |          |         | ×      |        |
| Number of Intersections complexity |        | ×      |          |         | ×      |        |
| Habit information                  | 0      |        |          |         | ×      |        |
| Comfort                            |        |        |          |         | ×      |        |
| Average speed                      |        | ×      |          |         |        |        |

In the research mentioned in the above table, the selection of the routes is based on the driver's behaviour for selecting the best routes for travelling based on human factors. They concentrate on sight selection areas to take the right routes for travelling. The GIS model is not applied to determine the variables effects on drive time. Spatial data are not used for calculating the distance and shortest time for selected routes.

### 2.7 Regression model

The regression model allows the users to examine and explore spatial relationships and assist in explaining the factors behind observed spatial patterns. According to Gkolias and Boile (2001), the linear regression models are mainly vehicle-based methods, and can be considered rather simple and straightforward for problem-solving in the transportation industry. In other words, regression is used to specify the nature of the relation between two variables. For this reason, to determine a GIS model based on regression analysis requires considering the parameters as independent and dependent variables.

A regression model uses map overlay operations in a GIS to combine all the independent variables needed for the analysis. According to the objective of this thesis, and the nature of the variables in GIS, the linear regression model is to be used, due to the fact that dependent and independents variables are all numeric in nature. For this study, different layers in the GIS are used, such as road networks, traffic volume, land use and population. The main distribution for this research is that average speed in different time in the morning is evaluated and tested in regression analysis.

## 2.7.1 Application of regression model

Gkolias and Boile (2001) applied a GIS-based tool for a transportation planner to estimate truck volumes on sections of highways system. They used the ArcView software which is the old version of spatial analysis and GIS application. In this study, the lack of traffic data for different times of the day and road network information is evident. The independent variables considered in regression mode are number of employees, sales volume, the number of establishment for different standard industrial classification (SIC) categories.

Using regression model in traffic management, Sofia, Nithyaa & Arulraj (2013) analysed the problem of optimal routes associated with transportation. Variables taken

into consideration were impedance for intersections, type of road and speed. Minimising the distance and drive time at the same time is the main goal of this research.

Çela, Shiode and Lipovac (2013) used regression analysis to find the most significant variables related to the road condition, time of the day and main causes for the high rate of accidents.

Hawas (2013) studied the bus routing problem to estimate bus route and network travel time using micro-simulation. The research is based on predicting the effectiveness of bus route design using traffic indicators. The indicators are average network traffic intensity, posted speed, route length, the frequency of bus operation, and average passenger loading. Regression models are used to predict both route and overall network travel time.

|                                   | Gkolias               | Sofia        | Çela         | Havas  |
|-----------------------------------|-----------------------|--------------|--------------|--------|
| Authors                           | (2001)                | (2013)       | (2012)       | (2013) |
| Variables                         | <b>O</b> <sup>*</sup> |              |              |        |
| Number of employees               |                       |              |              |        |
| Sales volume                      | V                     |              |              |        |
| Number of establishment           |                       |              |              |        |
| Type of road                      |                       | V            |              |        |
| Intersections                     |                       | $\checkmark$ |              |        |
| Speed                             |                       | $\checkmark$ |              |        |
| Road Condition                    |                       |              |              |        |
| Time of the day                   |                       |              |              |        |
| Main cause to affect the accident |                       |              | $\checkmark$ |        |
| Traffic Intensity                 |                       |              |              |        |
| Road Length                       |                       |              |              |        |
| Bus operation                     |                       |              |              |        |
| Average passenger loading         |                       |              |              |        |

Table 2.3: Variables in different regression analysis

Another research based on regression model shows the capabilities of GIS can enhance the understanding and information to be used in marketing planning and produces various statistical graphs and diagrams. The linear regression model was applied to study the market segmentation that describes the selection of groups of people who will be most receptive to a service or product. The assumptions in this research are the speed limit for each road, there is no any prohibited turns at the nodes, the time delay of the turns at the nodes is negligible, and traffic flow is smooth (Musyoka, Mutyauvyu, Kiema, Karanja, & Siriba, 2007).

| Author<br>Variables                                | Musyoka (2007) |
|--|----------------|
| Speed limit for Road Networks                      | $\checkmark$   |
| There are no prohibited turns at the nodes         | N              |
| The time delay of turns at the nodes is negligible | $\checkmark$   |
| Traffic flow is smooth                             |                |

Table 2.4: Variables involved in the Regression model

The mathematical model has been used for predicting sales for market zones in this research. The objective is to ensure the optimal distances in delivering the beverage and to also minimise fuel expenditure.

The inadequacies present in the mentioned researches are that the GIS model is not applied to find the fast roads for delivery services. The traffic data is not used for analysing the best roads within the market location. The analysis did not finalise the best routes regarding different time frames to make an optimal model. The application of network analysis for this study determines the destinations within the particular travel time mentioned in the research. Statistical results show that the car volume for each road network has a significant effect on drive time so it is necessary to consider this factor when selecting the best routes based on drive time.

The variables adopted for this research are detailed in Table 2.5.

| Variables                            | Sub—Variables     |
|--------------------------------------|-------------------|
| Speed of a car within the 7 to 10 am |                   |
| Speed of a car within 10 to 12 am    |                   |
| The length for each Road Network     |                   |
| Average speed of a selected car      |                   |
| One-way road restriction             |                   |
| The volume of traffic                |                   |
| Population for each road networks    |                   |
| Land use                             | Residential area  |
|                                      | Industrial region |
|                                      | open space        |
| School Zone Area                     |                   |

**Table 2.5: Variables Adopted In Current Research** 

This thesis develops a GIS model including socio, spatial and traffic data. The model is designed to integrate all the data as a spatial database management system. From regression analysis, the variables that affect the drive time are obtained to determine the fastest routes.

# 2.7.2 The influence of regression analysis on GIS practices

From previous researches, it is concluded that some factors are not considered to select the fast routes when regression analysis is taken into consideration. The most influential factor with a significant value on drive time is the distance for all the road networks. It included that the residential characteristics of the used route have a significant effect on drive time. Although the statistical results reveal that the distance is the most significant coefficient on drive time, it is not always the best choice for selecting the faster routes for delivery.

A regression model approach carried out for this research tested the variables that have the most influence on drive time. One of the most powerful uses of regression analysis is to predict the values of the dependent variable, sometime for a future period. By using the regression model the relationship between the variables and drive time will be specified. The relationship between GIS and spatial analyses is that both are concerned with the quantitative location of important features as well as the attribute tables of those features. The advantage of the current research is that using the regression model in GIS will assist to evaluate the variables and select the most influential variables in transportation.

# 2.7.3 Vulnerability analysis

Jenelius (2010) analysed road network vulnerability to assess social and economic consequences of disruption caused by events on the road networks. He proposed the combination of supply-side (link redundancy, network scale, road density, population density) and demand-side (user travel time, traffic load) indicators and combined them in statistical regression models. This study focused on vulnerability analysis and the impact on social and economic consequences. Jenelius studied the effect of sudden increases in travel times and cancelled trips following road link closures.

The spatial data is not used for data analysis. This study focuses on the road transport system especially on disruption of road networks. The vulnerability analysis is used to refer to the study of the potential degradation of the infrastructure and their impacts on society. A spatial database was not developed to approach the pathfinding in Network Analyst. The time pattern in the morning and afternoon was also not considered.

| Main-Variables | Sub- Variables     |  |
|----------------|--------------------|--|
|                | population density |  |
| Supply-Side    | road density       |  |
|                | network scale      |  |
|                | link redundancy    |  |
| Demand-side    | user travel time   |  |
|                | traffic load       |  |

**Table 2.6: Variables in Vulnerability Analysis** 

Liu (2008) proposed the framework to assess the reliability of using the fire incident reporting system database on studying the emergency response. The framework is based on alarm time and arrival time records in database system that can be used to estimate the emergency response route selection. To determine the influence of factors in actual driving behaviour a linear regression model was applied for this kind of analysis. The weakness of this study is that the speed limit for different road network as a variable in calculating drive time in selecting the fast routes is not considered. The GIS software for selecting the short path is not part of this research. The time pattern in the morning for travelling was also not part of the study.

## 2.7.4 Solid waste collection using GIS network analyst

Bhambulkar (2011) conducted research to find the best routes. The weakness of this research is that the GIS Model is not used for this research. The different time in a day is not important for such research and traffic data is not used.

Ghose, Dikshit, & Sharma (2006)optimised the routing system for the collection of solid waste. The parameters considered in the model are the information on population density, waste capacity, road network, type of road, storage bins and collection vehicles. Based on population density and width of the roads, different size bins is proposed to be placed in different parts of the studied area.

In The GIS technology helps to analyse the map layer to finalise the overlay maps to identify the best site selection for landfill of hazardous waste management(Sharifi et al., 2009). Land use is a critical factor for selecting the site selection for landfills for waste disposal. The shortcoming of this study is that the Network Analyst is not applied to calculate the optimal routes, and the road networks are not mentioned.

#### 2.8 GIS models

Description of a model is a simplified representation of a phenomenon or a system, and GIS modelling is the application of a GIS and its functionalities in creating a model with geospatial data. GIS models can be categorised into four general types: binary, index, regression, and process models (Chang, 2008).

### 2.8.1 The role of GIS in modelling

GIS can help the modelling process in several ways. First, a GIS is a tool that can process, display, and integrate different data sources including maps, digital elevation models (DEMs), GPS (global positioning system) data, images, and tables. These data are needed for the implementation, calibration and validation of the model. A GIS can function as a database management tool and, at the same time, is useful for modellingrelated tasks such as exploratory data analysis and data visualisation(Chang, 2008).

Second, models built with a GIS can be vector-based or raster based. Choosing the appropriate model for conducting a research it depends on the nature of the model, data sources, and the computing algorithm. A raster-based model is preferred if the spatial phenomenon to be modelled varies continuously over the space such as soil erosion and snow accumulation. A raster-based model is also preferred if satellite images constitute a major portion of the input data, or if the modelling involves intense and complex computations. However, raster-based models are not counselled for studies of travel demand, for example, because travel demand modelling requires the use of a topology based road network (Chang, 2008).

Vector-based models are recommended for spatial phenomena that involve exact locations and shapes such as road network (Chang, 2008). The purpose of this study is implementing a GIS model for the routing approach. Tofulfil the defined objectives of this thesis, it is more suitable to apply a vector-based model for this research. Applying

the network analysing tool to define the road networks requires using the vector-based model for final processing.

## 2.8.2 Type of models

### 2.8.2.1 Binary model

A binary model uses logical expressions to select spatial features from a composite feature layer or multiple rasters. The output of a binary model is in binary format: 1 (true) for spatial features that meet the selection criteria and 0 (false) for features that do not (Chang, 2008). The model above is not suitable for selecting in this research because based on the objective of this thesis it has to apply the road network data in GIS software to find the fastest roads.

## 2.8.2.2 Index model

An index model calculates the index value for each unit area and produces a ranked map based on the index values. An index model is similar to a binary model in that both involve multi-criteria evaluation, and both depend on overlay operations for data processing. However, an index model produces for each unit area an index value rather than a simple yes or no (Chang, 2008). To mention about the application of the index model, the index models are commonly used for suitability analysis and vulnerability analysis.

## 2.8.2.3 Regression models

A regression model relates a dependent variable to a number of independent variables in an equation, which can then be used for prediction or estimation of the events in any conducted research (Rogerson, 2001). Like an index model, a regression model can use overlay operations in a GIS to combine variables needed for the analysis. There are two types of regression model: linear regression and logistic regression (Chang, 2008). The most appropriate model for this research is the linear regression model based. It has developed the road network data to develop a model for best delivery of fresh vegetables based on time.

A multiple linear regression models is defined by:

$$y = a + b1x1 + b2x2 + \dots + bnxn$$

Where y is the dependent variable, xi is the independent variable i and b1,.....,bn are the regression coefficients. All variables in the equation are numeric variables. They can also be the transformation of some variables. According to the specifications of this type of model and fulfilling the objectives of the current research, this model represents the best delivery routes based on time (Chang, 2008).

## 2.8.2.4 Process model

A process model integrates existing knowledge about the environmental processes in the real world into a set of relationships and equations for quantifying the processes (Beck et al., 1996). Environmental models are typically process models because they must deal with the interaction of many variables including physical variables such as climate, topography, vegetation, and soils as well as cultural variables such as land management (Brimicombe, 2009).

### 2.8.3 Spatial data models

Spatial data has been traditionally stored and presented in the form of a paper map. Two basic types of spatial data model have been developed for storing geospatial data digitally. These are known as vector and raster models. These methods are used to reduce geographic phenomena to forms that can be coded in computer databases (Longley, Goodchild, Maguire, & Rhind, 2005).

## 2.8.3.1 Application of spatial data

GIS offers spatial modelling options that account for the influence of various factors, such as distances between the main location of producing the fresh vegetables, market places and type of roads. GIS modelling requires its provision to run the model while identifying the least-cost path generated through an exhaustive and iterative process for vehicle routing between origins and destinations.

The relationship between GIS and spatial analyses is that both of them are concerned with the quantitative location of important features as well as the attribute tables of those features. GIS user may also decide what features are important and what is important to them. In many situations, the GIS is the only way to solve spatially-related problems (Bolstad, 2013).

The traditions of GIS are strongly based on the maps, and even today it is common to introduce GIS through the ideas of representing of contents of maps in computer databases. The mapis based on ideas such as layers; projection, generalisation, and symbolization are still prevalent in GIS and account for a large proportion of the capabilities in GIS (Goodchild, 2007). There is no doubt how a technology built essentially for handling maps can be adapted to the needs of dynamic simulation modelling, and finally, GIS technology is the optimum platform of related modelling (Goodchild, 2000).

ArcGIS is a software in GIS to store the spatial data and other socioeconomic data to identify the hypermarkets and develop the distribution routes for this study. ArcGIS has the capability to explore, query, and analyse the data geographically. ArcGIS comes with ready-to-use data, which can be used to create various maps and to solve problems by uncovering and analysing trends and patterns (Panwhar, Pitt & Anderson, 2000).

### 2.9 Distribution as a DSS

A DSS can be a personal system for an individual manager or a shared system for a group of managers with similar jobs. The system may be customised to the specific need of managingit or it may be assembled from commercially available components. In any case, the manager-users are the only people who can evaluate any DSS. DSS is a movement to extend computer-based decision aiding to those areas in which the global, general systems of information management system was not successful. Accordingly, the DSS seeks to build flexible, quickly changed, tailored decision-aiding systems for particular problem areas (Turban, 1993).

To better understand the modelling process it is advisable to follow the decisionmaking process, which involves three major phases (Turban, 1993):

1) Intelligence 2) Design 3) Choice and Implementation of Solution

The modelling process starts with the intelligence phase where reality is examined, and the problem is identified and defined. In the design phase, a model that represents the system is constructed. The choice phase includes a proposed solution of the model. A conceptual picture of the modelling process is displayed in Figure 2.2.



Figure 2.2: The Decision-Making/Modelling Process (Turban, 1993)

For this research, from the view of intelligence phase, the data collection can be divided into two groups, the spatial data and non-spatial data. Problem identification concerns finding the optimal routes for distribution of fresh vegetables as perishable goods and preserving the nutrition value to gain more benefit from consumers. Companies also attempt to reduce their costs based on total distance travelled and the length of drive time, which can also reduce the loss of quality. The model was implemented using the ArcView GIS software, after collecting and gathering all the spatial data and inserting relevant attributes to the system for the final decision.

## 2.9.1 Components of DSS

1. Data Management: Includes the database, which contains relevant data for calculating the variables. This research considers spatial data and non-spatial data. The spatial data consists of information about the land features such as roads and hypermarket location. The non-spatial data includes information about the route such as distance, speed of the vehicle and drive time.

2. Model Management: The software package used for network analysis in this research is ArcView GIS software, which can analyse the best route features and makes the decision for the related purposes.

**3.** Communication Subsystem: The user interface consists of the optimal route map, including the calculation based on travel expenses and length of drive time.

## 2.9.2 GIS as a DSS

GIS technology offers extremely significant power in transport modelling. The spread of GIS use facilitates the efficient and portable spatial data storage, updating and processing. In addition, a GIS system facilitates model accessibility, database maintenance and updating, and cartographic display of model results. This cangreatly
enhance the role of the transport model as a decision support system (DSS) in transportation planning and policy development. The interface of network based equilibrium models with a GIS platform offers substantial potential of modelling transportation planning, analysis and control (Bohari et al., 2012)

According to Rich & Davis (2010), concepts and patterns are much easier to realise when complex information is communicated through geographic visualisation. On the other hand, the value of visualisation as it relates to decision making refers to the potential for better and rapid intellectual processing of information when starting with a complex especially richly expressive visual image (Rich & Davis, 2010).

# 2.10 Network analysis in GIS

Specific areas of GIS application could determine optimal and shortest paths such as routinganalysis or allocating resources based on demand and capacity. Much of this analysis is based on a set of connected linear features – network – that form a framework through which resources flow. This framework could be visualised as a road network through which vehicles move (Mukund& Rao 1992).

Network applications in a GIS are oriented towards planning, administering and operational management of resource facilities. Some of the applications are:

- **A**) Traffic routing for transportation planners, facility managers for efficient routing of facility movement and so on.
- **B**) Facilities management for planning demand-capacity rations for resources and optimal allocation of resources based on demand and capacity (Samet, 1990).
- C) Best routes based on the shortest distance or shortest time between destinations can be found by using Network Analyst. Thebes routes can be found during peak hour by changing weighting factors for route calculation (Memon, 2005).

Each network consists of different elements, each of which could be associated with an attribute defining the characteristics of the elements. The different elements in network are:

A) Lines B) Turns C) Stops D) Facility points E) Blocks (Mukund et al., 1992).

According to Moon et al. (2000), GIS can manage network data more comprehensively than a transportation model, and can be linked to transportation models by specific software. Many GISs support analysis in networks which are systems in interconnected lines representing in vector data. In the real world, networks consist of roads system and water supply and in all of them, there is one thing in common which transports moveable resources. The most common objective in routing across networks is to minimise the cost of the route. Cost can be defined and measured in many ways but is frequently assumed to be a function of distance, time, or impedance in crossing the network.

## 2.10.1 Network analyst model

According to Bhambulkar (2011), ArcGIS Network Analyst is a powerful extension that provides network based on spatial analysis including routing, travel directions, closest facilities and so on. It enables users to dynamically model realistic network conditions, including one-way restriction, speed limit, height restrictions, and traffic condition at the different time of the day. Users with Network Analyst extension can:

- Find efficient travel routes
- Generate travel directions
- Moreover, find a service area around the site

# 2.11 Database management system (DBMS)

Database management system is comprised of a set of programmes that provide facilities to manipulate and maintain the data in the database. Database can be described as an integrated collection of data and or information in a structured format. There are several advantages of a database management system:

- 1) Centralised control
- 2) Data can be shared efficiently
- 3) Data Independence
- 4) Implementation of new database application is easier
- 5) Direct user access and
- 6) Redundancy can be controlled(Healey, 1991).

Figure 2.3pertains to the data input and the data output in GIS. After inputting and collecting the non-spatial data, it will be stored in an attributes table in ArcView GIS software. With a combination of the road network layer as a spatial data and the non-spatial data which can be retrieved later, the data will be analysed, and the result will be the output of the alternative scenarios for solving the problem using the DSS.



Figure 2.3: The Architecture of a GIS (Martin, 1991)

The three major phases of a database are the conceptual, logical and physical designs.

*Conceptual Design* defines the application needs and the objectives of the database. This includes identification of spatial and non-spatial elements and identifies how entities will be represented in the database.

*Logical Design* refers to the components involved in the software. A database management system usually determines logical design. It gives details about how entities are related to each other.

*Physical Design* refers to the assessment of load, disk space requirement, memory requirement, access, and speed requirement of the GIS. Figure 2.4 shows the three major components of database design:



Figure 2.4: GIS database design stages (Healey, 1991)

#### 2.12 Marketing factors

The market for fresh and perishable produce especially fresh vegetables is increasingly guided by demands for freshness (Pimento, 2010). The hypermarket business is the top retail business in Malaysia contributing to national growth. International hypermarkets contributed 33.6% growth during 2001-2005 while local hypermarkets contributed 16.7% in the same period (Bohari et al., 2012).

A hypermarket is a concept of supermarket and department store together under oneroof. However, different retailers have different marketing strategies to attract customers. The best examples of leading hypermarkets in Malaysia are Tesco and Carrefour. Tesco is an international hypermarket from the United Kingdom. With a huge growth of profit, Tesco decided to look into the expansion of its international market to maintain its position. Malaysia was selected as the next market of entry, and the first Tesco Hypermarket in Malaysia was opened in Puchong in May 2002 and followed with the second Tesco which opened in October in Malacca. Tesco carries a total of 86,000 lines of products including more than 1,300 Tesco branded items (Hasliza, 2012).

Globalisation of the retail food system has impacted on the distribution and marketing of fresh food. In most developing countries including Malaysia, traditional retail markets are being replaced by supermarkets and hypermarkets (Chamhuri, 2013).

### 2.12.1 Development of Carrefour

Carrefour was created in 1959, which has grown from a single supermarket in France to more than 11,080 stores in 29 countries. Operations range from supermarkets stores and a variety of other outlets to convenience makes it one of the largest retailers in the world. Carrefour began its expansion to foreign countries in 1969 and the first hypermarket established abroad was in Belgium. In the mid-1970, Carrefour began its expansion to international market outside of Europe. The first store established in Brazil and the first supermarket was established in Argentina in 1982. Carrefour marched to Asia in 1989, and the first supermarket was established in Taiwan (Hasliza, 213).

This kind of hypermarket was well received by French consumers and has become one of the marketing models in retail commerce. This new concept of a supermarket delivered consumers with an amazing experience of shopping. They can buy whatever they need in one place. Compared with traditional retail store the retail price is 10-50% lower in Carrefour. Due to the low price, the annual sales of Carrefour increased by 50% from 1965 to 1971 (Hasliza, 2013).

Carrefour offers a different kind of products such as grocery product, clothing, consumer goods and household appliances besides these commodity products. It can satisfy all of the customers' need for only one trip. It offers healthy products such as organic products for health-conscious customers. In its operation, Carrefour adheres to the concept of "fresh quality product at discounted price and direct from producer to customer"(Kamath & Godin, 2001).

In Malaysia, Giant has become synonymous with everyday low prices, big variety and great value. This hasbeen underscored by the Shoppers Trend Survey, which showed that Giant was perceived as the cheapest place, in Malaysia, to shop for everyday groceries.

# 2.12.2 Changing the name of Carrefour and Jusco to AEON BIG

In November 2012, AEON acquired the operation of Carrefour Malaysia. All the current Carrefour hypermarkets in Malaysia will be fully re-branded into AEON Bing in the near future. Also, Carrefour has changed its name to AEON Big. This company is a Japanese leading company, which manages many other hypermarket chains in Malaysia. All of the current Carrefour hypermarkets and supermarkets in Malaysia will be fully rebranded into "ÆON Big" in the near future. The acquisition of Carrefour Malaysia will make ÆON as the second largest retailer in the nation (Hasliza, 2013).

'JUSCO' and now 'AEON' are well established among Malaysians as well as foreigners, especially due to its association with the international AEON group of companies. At all times, in every market, AEON's activities are guided by its unchanging 'Customer First' philosophy. Its aim is to surpass expectations by combining excellent products with unique personal services that enhance the shopping experience. In March 2012 all the Jusco stores in Malaysia under the AEON umbrella in Japan officially changed their names to AEON (Hasliza, 2013).

### 2.12.3 Population issue in Malaysia

According to the year 2000 statistics, 4,791,000 people who were accounted for about 21% of the nation's population lived in Klang Valley. Among this population, 84% of the households are car owners (Ariffin & Zahari, 2013). Like many other developing countries, Malaysia's growing urban population and increasing household incomes have led to a rise in car ownership. According to Ariffin & Zahari (2013), the most developed and fastest growing region in the country is the Klang Valley that covers an area of about 2,826 square kilometres. It comprises more than 30 rapidly growing new towns that include Kuala Lumpur.

According to Dissanayake, Kurauchi, Morikawa and Ohashi (2012), Kuala Lumpur is one of the most congested cities in Southeast Asia and therefore finding suitable routes for delivery of fresh produce such as fresh vegetables is critical. Kuala Lumpur is known also as a metropolitan city in Asia. It has undergone extensive growth and modernisation since 1987 and as a result, new industrial zones have appeared in and around KL especially in the Klang Valley.

Kuala Lumpur is the largest city in Malaysia with a population of around 1.6 million. As highlighted earlier the road networks in Kuala Lumpur and Klang Valley has reached the maximum capacity, and the changes in consumer behaviour demanding for more quality and nutritious food, thus giving rise for an improved delivery of fresh vegetables to marketplaces to meet consumers' demand.

In Malaysia, the car ownership is growing rapidly. Growth in the city has expanded the aspect of population, facilities and infrastructures including traffic and transportation network (Tey, 2003). This thesis selects the population as a social parameter. It has to evaluate the statistical approach to find the impact of this variable on drive time.

# 2.13 Conclusion

This chapter discusses the transportation system and the application of GIS. A brief explanation about GIS and its influence in the distribution of fresh vegetables was discussed. Different types of models were described following the explanation of the application of the model in GIS. The study extends the database management system, and the component of DBMS was proposed. The application of DSS on distribution was explained, and the flowchart for decision making and the modelling process was developed. The aspects that influence the marketing factors were explained. Therefore, in this part, the brief introduction about Carrefour has been mentioned, and the expansion as a huge hypermarket was highlighted. GIS is an excellent example of representing spatial variation in the initial and boundary conditions of the model and final output. GIS includes numerous tools for pre-processing and transforming data for use in modelling including data management, format conversion, projection change resampling and other facts that would be needed to assemble the data for dynamic simulation.

### **CHAPTER 3: RESEARCH METHODOLOGY**

### 3.1 Introduction

This chapter describes the research methodology of the study. It provides an overview and discussion on how the research is designed and the choice of methods adopted for gathering and analysing data. Section 3.2 gives a brief explanation about the location of the study and the main activities of the company regarding the destinations for distribution of fresh vegetables. This section also gives a brief explanation about hypermarkets studied in this research and also defines their locations in the digital map. The types of trucks for the delivery of fresh vegetables have been illustrated as a figure in this section. Section 3.3 discusses the material and tools which consist of a data including source of the data. This section is focused on using the GIS software to handle the spatial data in the database system and the process to adopt the network dataset for further analysing. This section also describes ArcGIS software as a tool and identifies the extensions using ArcGIS as well as its application in different fields. Section 3.5 extends the methodology process and how to achieve the goal of this research. The database management system and the framework of study as a conceptual design are applied to solve the distribution problem.

# 3.2 Duration and location of research

This research wasconducted from December 2009 to January 2015. To develop the GIS model, vehicle speed data were collected from the road traffic and monitoring department to determine the average speed of the vehicle in one year. This data can be updated by the GIS model for each year. The database management system that was created to solve the distribution problem can be updated every year. The new information regarding the average speed of the vehicle from the department of transportation and traffic and can be updated in the database for further analysis. This study will cover Selangor including

Kuala Lumpur and Klang Valley. The area is located approximately in the latitude of 3.1425 and longitude of 101.6501. Figure 3.1 is the map showing the study area for this research.

Dissanayake et al. (2012) clarified that Kuala Lumpur is known as one of the most congested cities in Southeast Asia and therefore finding inefficient and effective route for delivery of fresh produce is an essential concern. There is a trend of rapid motorisation in many Asian countries, and thus congestion is an important issue. In many of those cities, the travel time for work trips and other purposes is about 45 minutes and even more. Kuala Lumpur currently has traffic congestion problem and the statistics shows that the problem is going to get worst. There are as many as 2 million vehicles on the streets of KL every day. Traffic of KL is also heading to a direction with more traffic in road network (Schwarcz, 2003).



Figure 3.1(continued) The Study Location (Data Source: From land use and cartographic department in KL)



Figure 3.1: The Study Location (Data Source: From land use and cartographic department in KL)



Figure 3.2: Distribution locations on the map (Source: The Malaysia map and Selangor area from the Land use and cartographic department in Kuala Lumpur)

The pink colour in Figure 3.2 is the location for producing the fresh vegetables. The main goal of this study is to improve the delivery services by selecting the least drive time routes. The transportation will startform this point to the city centre. The blue colour is Kuala Lumpur and the main destination of fresh vegetables. Most of the hypermarkets are located in this area. The map in Figure 3.2 was designed and developed by ArcGIS software.

# 3.2.1 Location of study

The study area for this research is a place for producing fresh vegetables. The name of the place is Sime Aerogreen Technology Sdn Bhd, which is located in Seremban, Negeri Sembilan. This place is one of the most important places for producing fresh vegetables in Malaysia and from this place distributes into many different hypermarkets in Selangor. This research will focus on hypermarket places as destinations for delivery and to know the suitable routes for distribution.

With its rich background in the plantation, Sime Darby Bhd has ventured into Aeroponic Technology since 1998. The technology has been developed into a state-ofthe-art for commercial production of premium temperate vegetables in Seremban under Sime Aerogreen Technology SdnBhd, which was incorporated in 1997. The 5-hectare farm is located in Seremban a one hour drive from Kuala Lumpur city, and the products are marketed under the brand name "Sime Fresh" which are available leading hypermarkets and supermarkets in the Klang Valley, Penang, Melaka and Johor Bahru. Sime Fresh is a fully Malaysian company.

The company was incorporated in 1997 and commenced operation in 1998 with the initial capacity of 900 troughs. In 2005, the top platform (Green house A) was refurbished, and the trough's capacity had been increased from 900 units to the present 1216 units. In Aeroponics, the plant roots are suspended in the air. The availability of oxygen-rich aeration lacing around the root-zone enables fast and healthy plant growth. To grow temperate crops, chillers are used to 'trick' the plants into 'thinking' that they are being grown in a cold climate by cooling the root-zones of the plants as the nutrients and moisture are sprayed onto the roots.

This research focuses on the four main hypermarkets located in Selangor. The main aim of this research is on the determination of the fastest route between the Sime Fresh and the four main hypermarket destinations. The names of these hypermarkets are Jusco (AEON), Giant, Carrefour (AEON), and Tesco.

For the purpose of transportation, this agency uses several trucks in order to deliver the fresh vegetables to the different locations. The capacity of each truck is fixed. Figure 3.3is an example of a lorry vehicle for the purpose of delivery services.



Figure 3.3: Truck for the delivery of fresh vegetables (Sime Darby website)

Figure 3.4 shows the locations of the markets that the sampled companies plan to distribute their fresh vegetables. The total number of markets is 31. The objective is to improve the delivery services by achieving the least drive time to reach the marketplaces for keeping better quality produce. The starting point is in Seremban which is about 100 kilometres from KL.



Figure 3.4: Map of Hypermarket locations (Source: GIS and Cartographic Lab, UM)

#### **3.3 Material and tools**

Database development is focusedon two categories, 1) Spatial data component. It consists of maps to visualise them in different format as a map and which have been prepared by the interpretation of remote sensing data. Some examples of the maps are road network, geological, land use, village. Many of these maps are available in an analogue form, and recently some other map's information is available directly in digital format (Nielsen, 1994). The aim of this research is to analyse the spatial data in identifying the best routes for delivery purpose. To achieve this goal, the digital format of the map must install in related software. All the information regarding the road network is available in the digital format for better interpretation in ArcGIS software. The land use data is also available in digital format for better interpretation for selecting the appropriate roads. The last decision is based on map development for choosing the right paths for the delivery of fresh vegetables.

2) Non-spatial data component. It consists of attributes as complementary related to the spatial data. Attributes describe what is at a point, what about along a line or in a polygon. The attribute of socio-economic category could be demographic data or traffic volume data of roads in a city. The non-spatial data are mainly available in tabular records based on different tables in analogue format and need to be converted into digital format for incorporation and evaluation in GIS (Nielsen, 1994). The travelling speed of the vehicle as well the drive time are examples of non-spatial data for this research.

# 3.3.1 Data collection

According to the research objectives and requirements of this thesis, the data that must be collected and support the network analysis to find the fast routes. Table 3.1 illustrates how the data is gathered from different sources and the flow of data into the GIS. The first column is the information about the spatial data such as road network that is including in base map. For this case, all available data for routes are required for solving the problem for distribution of vegetables and analysing the data. Then it is necessary to know the location of each market on the map to see how far from each other and finally compute the total distance from the main point (Sime Darby).

From the view of the Base Map, the road network is the main part of spatial data which plays an important role in applying in the software for selecting the routes. The source of the data is from a private company producing different kinds of maps for various purposes. From the aspect of a land use map layer, the only important points are the locations of hypermarkets, and it is necessary to know the distance between them. The land use data, which was gathered for data analysis, was collected from the GIS Lab Faculty of Geography in the shapefile format.

The non-spatial data is divided into two different groups. The first is the overall distribution cost, and the other one is the length of drive time which contains several parameters. Distribution cost considers the mileage of the truck during the routes plus the number of vehicles to deliver the fresh vegetables.

The length of drive time contains several parameters such as the specific travel distance which the vehicle travelled between two locations. The next is the speed of the vehicle and the speed limit for each road network. The average speed of the vehicle is the consideration of one-year speed of the vehicles along the road. The school area is a zone where the movement should be slow so drivers have to reduce their travel speed in this specific area. The other parameter to be considering in the model is the time of the day spent travelling. In the peak hours, the traffic is more congested and increases the travelling time. The variable of time in the study is classified into twotime patterns, one from 7 am to 10 am, and the other from 10 am to 12 am. Data gathered from individual time frames will be separately transferred to the ArcGIS attribute table. To develop a GIS regression model regarding the variables the drive time is considered as the dependent variable and the average speed in two time frames is part of the dependent variables.

The land use is another important parameter that has a strong relation with time consumed for the distribution of fresh vegetables. The land use has several subparameters including residential area, commercial area and other private building. The population is a parameter that affects travel time which is the number of the people living in a specific area or the density of people in a location.

From the spatial data which is the location of each market, we need the name and description of each market in the map to distinguish between the marketplaces. The start point for the purpose of delivery services is the main agriculture land to produce fresh vegetables and from their it is distributed to all hypermarkets in KL city as mentioned in the previous map.

| No. | Type of Data  |   |
|-----|---|---|
|     | Spatial Data (Feature)  | Non – Spatial Data (Theme)  |
| 1.  | Base Map<br>• Road Network<br>• Location of the study   | <ul> <li>Travel Expense:</li> <li>Mileage</li> <li>Car volume</li> <li>Length of Drive Time:</li> <li>The specific travel distance between the certain places</li> <li>The average speed along the routes and also considering speed limit</li> <li>Drive time</li> <li>Time patterns for delivery including average speed in the 7 to 10 am and 10 to 12 am</li> </ul> |
| 2.  | Land use Map<br>Market location<br>Residential Area<br>Commercial Area<br>Open Space<br>School Zone<br>Population | The marketplace for the specific location   |

 Table 3.1: The data needed for selecting the optimal route

#### **3.3.2** Source of the data

Data for a GIS comes from different sources – mainly characterised by the type of data, which include following statements.

### (a) Available map data:

The most important source of data for a GIS application is an available map. Remote sensing data, images are another source of data. Data can be "collected" through digitising, scanning, interactive entry, etc. Producing different kinds of maps based on road networks for better decision and interpretation of delivery services is the main contribution of this research.

# (b) Tabular data:

Attribute data for a GIS are mainly tabular data collected by sampling. Most common of the attribute data are the census records, which represent the population, occupation structure, and development levels (O'Brien, 1992). The study has considered about the spatial data to find the shortest distribution routes for fresh vegetables and to minimise the travel expenses in order to reduce the total distribution cost.

For collecting and gathering the non-spatial data, it has been referred to the source of the data that is SimeAerogreen Technology SdnBhd which is called SIME FRESH. In addition, the information regarding the average speed for the road network in Kuala Lumpur was collected from the Department of Transportation and Traffic Management under Dewan Bandaraya Kuala Lumpur (DBKL). The data regarding the speed of the vehicle is the average speed of the vehicle for one year. For collecting data regarding car volume, the information was gathered from the same department (DBKL). For gathering the data regarding the timeframes, the data was collected from the same department.

| No. | Non – Spatial data   | Source  |
|-----|--|---|
| 1   | Destination, the number and name of<br>hypermarkets studied in this research for<br>distribution | SimeAerogreen Technology SdnBhd<br>(SIME FRESH)                                   |
| 2   | Delivery Times   | SimeAerogreen Technology SdnBhd<br>(SIME FRESH)                                   |
| 3   | Receiving Time Window  | SimeAerogreen Technology SdnBhd<br>(SIME FRESH)                                   |
| 4   | Truck size for each Delivery time  | SimeAerogreen Technology SdnBhd<br>(SIME FRESH)                                   |
| 5   | The speed limit during the way   | Department of Transportation and traffic management under (DBKL)                  |
| 6   | Car Volume   | Department of Transportation and traffic management under (DBKL)                  |
| 7   | Name of State where hypermarket located  | SimeAerogreen Technology SdnBhd<br>(SIME FRESH)                                   |
| 8   | Time pattern 7 to 10 am  | Department of Transportation and<br>traffic monitoring management under<br>(DBKL) |
| 9   | Time Pattern 10 to 12 am   | Department of Transportation and<br>traffic monitoring management under<br>(DBKL) |

# Table 3.2: Non-Spatial Data collection

Table 3.3: The Spatial data

| No. | Theme of spatial data            | Feature | Sources                         |
|-----|----------------------------------|---------|---------------------------------|
| 1   | Road Networks                    | Line    | GIS Vendor                      |
| 2   | Land use Map                     | Polygon |                                 |
|     | Residential Area                 |         |                                 |
|     | Commercial area                  |         | GIS Vendor                      |
|     | Open space                       |         |                                 |
|     | School zone                      |         |                                 |
| 3   | Market Location                  | Point   | SimeAerogreen Technology SdnBhd |
|     |                                  |         | (SIME FRESH)                    |
| 4   | Census data                      | Polygon | UM (Geography Department)       |
|     | Population                       |         |                                 |
| 5   | Digital map of Selangor district | Polygon | Department of cartographic and  |
|     |                                  |         | producing map in KL             |

# **3.3.3** Tool requirements

In this research, ArcGIS software is used to solving the problem of a complex network.

Using ArcGIS software with the extension of Network Analyst is the primary key to

exploring the data. ArcGIS developed by Environmental Systems Research Institute (ESRI), is a leading company on GIS solution to solve the problems. ArcGIS 9.3is used in this study.

ArcGIS software lets users explore and analyse the spatial data on their computers. Its basic mapping functionality and advanced GIS capabilities allow the user to create maps, showing the created maps, showing the data and integrate them and finally see the data in powerful new ways. ArcGIS enables to propose the research attractively. It can be found that communicating geographically is a powerful way to inform and motivate others.

ArcGIS is a complete system for designing and managing solutions through the application of geographic knowledge. It enables you to perform deep analysis, gain a greater understanding of your data, and make more informed, high-level decisions. Figure 3.5 illustrates the end users that can access maps, data and applications through different sources by using ArcGIS.



Figure 3.5: The Architecture of ArcGIS(Esri website)

Companies use GIS to achieve better information for good decision making. GIS shows the real world objects on map and user-friendly spatial tools to accomplish

complex task. GIS is used to display, manipulate and analyse spatial (map) data. Spatial data are data that contain a reference to a place (Nayati, 2008).

| Extension                     | Use  |  |
|-------------------------------|--|--|
| ArcGIS 3D Analyst             | 3D visualisation and analysis  |  |
| ArcGIS Geostatistical Analyst | Statistical tools and models for data exploration,<br>modelling, and probabilistic mapping |  |
| ArcGIS Network Analyst        | Routing, closest facility, and service area analyses                                       |  |
| ArcGIS Schematics             | Automatic schematic generation for ArcGIS  |  |
| ArcGIS Spatial Analyst        | Advanced raster GIS spatial analysis   |  |
| ArcGIS Survey Analyst         | Integrated survey management for ArcGIS  |  |
| ArcGIS Tracking Analyst       | Time-based data visualisation and analysis   |  |

 Table 3.4: ArcGIS Extension (ESRI, 2012)

Application of Network Analyst generates detailed directions across the route. All human beings have a natural tendency of select the shortest or fastest routes to go from one place to their destination. Transporters, as well as truck drivers both, prefer short and faster routes to cut operating cost and transportation time (Nayati, 2008).

# 3.3.4 Methodology

To facilitate managing spatial data features, GIS adopts layering technology (Heywood et al., 1998), which is commonly used in map making. Spatial features are grouped by feature type, and by the real world entities they represent. Each layer is handled by GIS separately, and has its set of spatial and attributes data (Heywood, Cornelius & Carver, 2002).

There are multi-layer operations, which allow combining features from different layers to form a new map and give the new information and features that were not present in the individual maps. The system development approach applying for this study and the group of procedures, in general, can be seen as following Figure 3.6. For this case, the first developing process for the spatial data contains the information about the base map (Digital map) such as:

- 1. Road networks layer (In the shape file format)
- 2. Land use layer (Shape file format)
- 3. Population (The unit is quantity, Person)
- 4. School Zone (Shape file as point)

All of these layers will overlay together and be classified to prepare the data for analysis.

Secondly, developing a process to prepare the non-spatial data consisting of:

- 1. The number and name of hypermarkets of the study location(Shape file, Points)
- 2. The speed limit for each route (kph)
- 3. Receiving Time Window (Time, Hour)
- 4. Drive time (kph)
- 5. Car Volume (Quantity of cars)
- 6. Time pattern of 7 to 10 am(kph)
- 7. Time pattern of 10 to 12 am (kph)

Figure 3.7 describes the combination of the spatial data and non-spatial data for preparing the information management system based on GIS tool. One of the benefits of using GIS is helping users think geographically as opposed to analysing data which is critical in distribution planning. The purpose of connecting the routes in this study is that the trucking company must determine the most efficient and effective routes of connecting stops for delivery fresh produce from starting points to all mentioned markets. A routing application has been developed to identify the most efficient routes for any set of pick-up locations.



Figure 3.6: Application developments with GIS

#### 3.4 Database Management System

DBMS is used to handle large volumes of data as well as to ensure the logical consistency and integrity of data, which also has become the major part of GIS system. Spatial data is mostly part of a complete work and information process. It needs to implement GIS functionality as part of a central database management system at least in conceptual level in which spatial data are maintained in one integrated environment (Zlatanova, Stoter, & Quak, 2004).

Making maps alone does not justify the high cost of building a GIS. Likewise, if the purpose is to generate tabular output, then a comprehensive database management system or a statistical package may be a more efficient solution. It is a spatial analysis that requires the logical connections between attribute data and map features, and the operational procedures built on the spatial relationships among map features. These capabilities make GIS a much more powerful and cost-effective tool than automated cartographic packages, statistical packages, or database management systems (Raju, 2003).

Database development is focused into two categories, a) spatial data b) non-spatial data. Spatial data analysis will propose by GIS technologies. From the part of spatialdatabase, three different layers can describe the model like road network, land use and Population layers. In non-spatial databases, the variables that play an important role are hypermarkets, the speed limit in restricted area drive time, car volume as well as time schedule. The Network Analyst module, analysis the road networks which store the information of travelling speed in different hours in the daytime. According to Winyoopradist and Siangsuebchart (1999), the travelling speed patterns for many roads in each direction are varied in the different hours in the morning.

The integration of these two databases it can make a good decision for distribution problems. It can develop various maps for producers of fresh vegetables to manage their goals better to improve delivery service. Two important aspects of DBMS functionality are addressed into two categories, firstly the spatial data and non-spatial data. The second aspect is DBMS functionality on spatial analysis. This aspect is also connected to network map and other resources related torouting information for the purpose of the better decision on solving distribution problems. Figure 3.7 illustrates how the data is prepared in GIS and development of an effective database for the project.

The figure is also devoted to the role of DBMS in a new generation of GIS architecture and focuses on the manner of spatial data that can be managed, stored and analysed in DBMS. For this study, DBMS is essential in the application in which large amounts of GIS data need to be maintained and managed. For the purpose of analysis, the GIS tool will interconnect with DBMS. The identified factors will be evaluated and analysed in the modelling process and the significant variables in pathfinding in term of time will be recognised and stored in the system for the further application.



Figure 3.7: Development of Database Management System

# 3.5 Conceptual framework

In this section, the research is focused on the conceptual design which is a systematic look at how users interact with the system, the functions they may apply and use and finally the database management system.

The planning and development of the model propose the set of routes for selecting optimal paths and achieving specific objectives of the study. The proposed model for this research is overcoming the distribution problem according to the time, the speed of the vehicle, distance, land use and population parameters which are used in the regression model. The first implementation is the function of selecting routes that can achieve the best time for transportation. Figure 3.8 describes the conceptual framework on how the processes of calculating the fast routes based on parameters are designed. According to this flowchart, there are two different sources of information about hypermarkets for decision makers to know the location of markets as well as the paths going through the final destination. All information regarding the road networks will be saved in the GIS database management system. The consideration of this thesis is to apply the regression analysis to find the most significant variables. All of these variables will develop the regression model in GIS. Finally, the decision makers will solve the distribution problem based on the parameter in the database system.

The data was enteredinto the database for all features from the digitised map of the network and other land use layers. The database contains the fields for which the data was collected. Gathering all the information collected from the routes and other related information for analysing the data will develop the database. The database in Figure 3.9 is proposed to analyse the network of routes and finally to solve the problem of distribution. Users can display objects according to the data in the database. Output options include cartographic quality maps as well as reports, lists, and final graphs.



Figure 3.8: The conceptual framework of capturing data for analysis

#### **3.6 Regression model**

A regression model relates a dependent variable to a number of independent variables in an equation, which can then be used for prediction or estimation (Rogerson, 2001). Regression model will estimate the fastest route for distribution of fresh vegetables based on significant parameters. The model is based on spatial and not spatial data collected from different sources as well as GIS departments. Results of such model may be exported into a database and import to GIS framework for visualisation and final decision making. The significant parameters will introduce the new equation in the regression model Tobe adopted for this research for the purpose of distribution between several destinations.

A simple linear regression is carried out to estimate the relationship between a dependent variable, *Y*, and a single explanatory variable, *x*, given a set of data that includes observations for both of these variables for a particular population. In the multiple linear regression models, there are p explanatory variables, and the following equation represents the relationship between the dependent variable and the explanatory variables:

$$Yi = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n$$

Where:

 $\beta_0$  is the constant term and

 $\beta_1$  to  $\beta_n$  are the coefficients relating the *p* explanatory variables to the variables of interest.

Multiple linear regression can be thought of as an extension of simple linear regression, where there are p explanatory variables, or simple linear regression can be thought of as a special case of multiple linear regression, where p=1. The term 'linear' is used because in multiple linear regression we assume that y is directly related to a linear combination of the explanatory variables.

As is the case with simple linear regression and correlation, this analysis does not allow us to make causal inferences, but it does allow us to investigate how a set of explanatory variables is associated with a dependent variable of interest. Regarding a hypothesis test, for the case of a simple linear regression the null hypothesis,  $H_0$  is that the coefficient relating the explanatory (x) variable to the dependent (y) variable is 0. In other words, there is no relationship between the explanatory variable and the dependent variable. The alternative hypothesis H1 is that the coefficient relating the x variable to the y variable is not equal to zero. In other words, there is some relationship between x and y.

In summary, we would write the null and alternative hypotheses as:

$$H0: \beta 1 = 0$$
$$H1: \beta 1 \neq 0$$

Figure 3.9lists the variables applied in this research. Those variables are adopted for this thesis after reviewing all the literature and research studied in the field of transportation and GIS for distribution planning. All of the variables are tested in the regression model to verify the variable that has the most effect on delivery time. The system will continue this process in loop and then all the variables will identify and use for the final mode. The parameters that have no influence will be eliminated from the system. The model will try to continue the processing for data evaluation, and if it realised that a variable is influential, the model would process the model further to find new results.

One of the most powerful uses of regression analysis is to predict the values of the dependent variable, sometime for a future period. By using the regression model the relationship between the variables and drive time will be specified. With a regression model, one can predict the value of each variable adopted in this research. The independent variable in this study is drive time. After data evaluation then the regression model will specify that which of the following variables affect the time for delivery.



Figure 3.9: Variables selected for regression analysis

# 3.7 Conclusion

This chapter outlines the research design and methodology used in conducting the research. A brief explanation of procedures for data analysis is included. This chapter also gives a brief explanation about the location of this study and also describes the hypermarkets for distribution process as the final destination. The flowchart describes the parameters for calculating the total distribution cost. This chapter also described the variables applied in this study and how to integrate them into mathematical equations.

The output of the study will produce the various maps showing the optimal routes for best distribution based on the study parameters. The main advantage of GIS differing from other database systems is its ability to show maps, visualising the planned results and makes them simple, easier and clearer to the decision makers especially for producers of fresh vegetables. Optimal path lets us analyse a grid file to find the best pathto a specific location which for this study optimal path is in terms of time and the closest start location as used in generating a cost surface. The computation is based on a surface cost file that will generate with cost surface. The selection in an overlay in polygons, lines and points enable the users to generate a map containing features and attributes of interest extracted from different layers for final decision in path finding. In other word, multi-layer operations are useful for manipulation of spatial data on multiple layer data such as land use layer.

Network analysis is designed especially for line features organised to connect networks, and the best applications are typically applied to transportation problems and optimal path finding. The network analysis tool can calculate the shortest travelling time between two locations by specifying the starting time.

This study focused on speed patterns for all the road network in two specific periods of time. For this case,routing is a major concern for transportation industry, especially in delivery services. The application of overlay analysis is that manipulates spatial data and organised them in different layers to create combined spatial features according to theological condition. The most common basic multi-layer operations are union, intersection, and identify operations. All three mentioned operations merge spatial features on separate data layers to create new features from the original coverage. The difference among these operations in GIS is the way that spatial features are selected for the purpose of processing and further analysis. According to data analysis DBMS plays an important role in thenew generation of GIS.

### **CHAPTER 4: DATA PREPARATION**

#### 4.1 Introduction

This chapter briefly details the data preparation and collection for the usage of the ArcGIS software. The next section will further extend these steps by explaining methods needed to input the spatial data into the database management system for analysis, and aiding the decision makers on solving the route problem for delivery of fresh vegetable. Section 4.3 shows the GIS models and their definitions, as well as its application in different case studies. The section also discusses the data's connectivity and its suitability for analysis. This is followed by a brief explanation regarding data preparation for statistical process vis-à-vis this thesis. This is followed by the explanation of the regression model conducted for this study, and finally the identification of the variables as key factors for modelling that aids in the decision-makingprocess for transportation planning. The final section discusses data preparation and the steps in processing the data analysis. Creating the new dataset for road network for the purpose of analysis and evaluation of data is explained step by step in this chapter. This chapter has extended the explanation about connectivity and Network Analyst and other related definition within ArcGIS. It clarifies the explanation about the existing models and how they approached the problem in transportation activities and delivery services.

# 4.2 Data preparation

In the initial stages, all of the raw data regarding the information of the road networks must be added to the database of the system in ArcGIS. Since the system itself focuses upon the most essential of road networks for distribution planning, the unwanted routes such as small roads will be duly removed from the system post-data evaluation.

There are different types of roads for solving the distribution problem of finding the routes with minimum drive time along the road network. The raw data was gathered from

a GIS vendor based on cartographic activities located in Kuala Lumpur. The combination of these road networks will be suitable for the preparation and analysis of a reliable model. Figure 4.1 addresses the types of roads.



Figure 4.1: Combination of all Road Layers

In order to combine all of the selected road networks for distribution planning, a decision maker would first have to click on the function labelled Toolbox, selecting data management tools, and finally settling upon the general tab. Under the general tab, there are several available options, and a user is to select a merging process to enable its analysis.

Next, the merged file is clicked upon, and the input dataset is selected and added to specific windows. The output dataset is renamed and saved in a specific folder. The
merging process takes a few minutes for all of the involved layers, with a small window popping up to show the progress of the merger in percentages terms. This window requires the details of all the necessarily data as input datasets and names the output dataset for the merged data file. Figure 4.2 displays the specification of data entry and output dataset as well as their functions in this stage.

| ➢ Merge   |   |
|---|---|
| Input Datasets  | Output Dataset<br>The output point, line, or<br>polygon feature class or<br>table that will contain all<br>merged input features. |
| C: (sis Data-Selangor &KL (principal highway Todo0_0.shp<br>C: (sis Data-Selangor &KL (principal highway _ 0_0.shp<br>C: (sis Data-Selangor &KL (street level.shp<br>Output Dataset | 0,  |
| Field Map (optional)  |   |
| LABEL (Text)  | *   |
| OK Cancel Environments << Hide Help   | Tool Help   |

**Figure 4.2: File dataset** 

The attribute table was developed for the total roads in KL and Klang Valley postmerger of all the types of selected roads, bringing the total to about 34617 roads for this study. This volume of road networks is the total number of all the roads emerged from the case study.

After data evaluation and study area preparation it was determined that the smaller parts of the street network, such as collector roads, were unsuitable for lorry delivery, and must subsequently be removed from the database management system. The collector roads are such small roads and not suitable for a lorry to pass through these roads. Figure 4.3 shows the finalised network dataset post-deletion of all the unnecessary roads, and the final result is the remaining roads suitable for the lorry in this case study to deliver fresh produce/vegetables. Figure 4.3 shows the remaining roads that will be processed in the database system to solve the distribution problem of route selection with respect to

time. The consideration of this study is to determine the routes with minimum drive time for the purpose of delivery.



Map of Finalized Road Networks



Figure 4.4 shows the attribute table of the total number of all of the selected roads for to solve the distribution problem. Figure 4.6 illustrates the number of roads for analysis in this study, which ended up being 2167 roads for further processing in the ArcGIS.

|   | 510  | Chang    | LAREL                         |   |  |
|---|------|----------|-------------------------------|---|--|
| H | FID  | snape    | LABEL                         |   |  |
|   | 2148 | Polyline | E18 silk                      |   |  |
|   | 2149 | Polyline | E18 silk                      |   |  |
|   | 2150 | Polyline | E18 silk                      |   |  |
|   | 2151 | Polyline | E18 silk                      |   |  |
|   | 2152 | Polyline | E18 silk                      |   |  |
|   | 2153 | Polyline | E18 silk                      |   |  |
|   | 2154 | Polyline | E7 cheras-kajang hwy (kajang) |   |  |
|   | 2155 | Polyline | E18 silk                      |   |  |
|   | 2156 | Polyline | E18 silk                      |   |  |
|   | 2157 | Polyline | E18 silk                      |   |  |
|   | 2158 | Polyline | E7 cheras-kajang hwy (cheras) |   |  |
|   | 2159 | Polyline | E7 cheras-kajang hwy (cheras) |   |  |
|   | 2160 | Polyline | E7 cheras-kajang hwy (kajang) |   |  |
|   | 2161 | Polyline | E18 silk                      |   |  |
|   | 2162 | Polyline | E18 silk                      |   |  |
|   | 2163 | Polyline | E18 silk                      |   |  |
|   | 2164 | Polyline | Ampang, j                     |   |  |
|   | 2165 | Polyline | Manis,2                       |   |  |
|   | 2166 | Polyline |                               |   |  |
|   | 2167 | Polyline |                               | - |  |
| ٠ | III  |          | •                             |   |  |
|   | Reco | rd: 14 4 | 1 ► ►I Show: All              | - |  |

Figure 4.4: The total number of road Networks

The next step involves setting up the attribute table by adding the variables identified in the literature review, which will be applied in the context of this study to create attribute tables for all the roads. The first is the length factor, which represents the distance for each of the selected road network. This data was gathered from a GIS vendor which is a private company located in Kuala Lumpur for map design and cartographic purposes. Preparing the result in this section requires calculations in the ArcGIS software with the help of geometry analysis. The results are shown in the window for road networks in this study. Figure 4.5 shows the table and columns that are representative of the fact that the length implies distance for each specific road. The length for the distance between two locations is in metres.

|   | Attributes of | of combine_ro | pad_rso                             | <u> </u>        | < |
|---|---------------|---------------|-------------------------------------|-----------------|---|
|   | FID           | Shape *       | LABEL                               | Lenght          |   |
|   | 0             | Polyline      | Exit to j kapar                     | 0.753625        |   |
| F | 1             | Polyline      | To shapadu hwy                      | 0.720727        |   |
|   | 2             | Polyline      | Ramp psn sultan ibrahim             | 0.243154        |   |
|   | 3             | Polyline      | Ramp psn sultan ibrahim to j langat | 0.381981        |   |
|   | 4             | Polyline      | Ramp psn sultan ibrahim             | 0.203056        |   |
|   | 5             | Polyline      | Ramp psn sultan ibrahim             | 0.277717        |   |
|   | 6             | Polyline      |                                     | 0.439268        |   |
|   | 7             | Polyline      | Shapadu - meru i/c                  | 0.866414        |   |
|   | 8             | Polyline      | Shapadu exit kapar                  | 1.079778        |   |
|   | 9             | Polyline      |                                     | 0.820921        |   |
|   | 10            | Polyline      |                                     | 0.394093        |   |
|   | 11            | Polyline      | Exit 502 pandamaran                 | 0.789177        |   |
|   | 12            | Polyline      | Exit 502 pandamaran                 | 0.773309        |   |
|   | 13            | Polyline      | Ramp psn sultan ibrahim             | 0.611434        |   |
|   | 14            | Polyline      | Ramp to federal high way            | 0.375867        |   |
|   | 15            | Polyline      | Ramp to tmn berkerley               | 0.414896        |   |
|   | 16            | Polyline      | Ramp psn sultan ibrahim             | 0.698781        |   |
|   | 17            | Polyline      | Toll plaza sg rasau                 | 0.012951        |   |
|   | 18            | Polyline      | Toll plaza sg rasau                 | 0.017878        |   |
|   | 19            | Polyline      |                                     | 0.37463         |   |
|   | 20            | Polyline      |                                     | 1.105236        |   |
|   | 21            | Polyline      | To meru                             | 0.390648        |   |
|   | 22            | Polyline      |                                     | 0.826715        |   |
|   | 23            | Polyline      | Ramp fed hwy to shapadu hwy         | 0.969899        |   |
|   | 24            | Polyline      | Ramp fed hwy to shapadu hwy         | 0.972858        | Ŧ |
| • |               |               | +                                   |                 |   |
|   | Record:       | 14 4          | 2 FI Show: All Selected Re          | cords (0 out of | • |

Figure 4.5: Attribute Table of the Distance

The second step is to calculate the average speed for all of the road networks in KL. First, all of the information regarding the average speed for the road networks in Kuala Lumpur was collected from the Department of Transportation and traffic management, under Dewan Bandaraya Kuala Lumpur (DBKL). The activities of DBKL are to monitor and handle the traffic and transportation problems. This gathered data is manually keyed into the attribute table in the ArcGIS software. Although there was an obligation to include the average speed of vehicles on all roads, this data was often not available for certain avenues and smaller road sat certain times.

Google maps display traffic conditions in real-time on major roads and highways. Google traffic can be viewed at the Google Mapswebsite or by using Google Maps application. Early versions of Google Maps provided information to users about how long it would take to travel a particular road based on the historic data. This information was not real-time and far from accurate (Al-Amin,2015).

Google maps also workby the concept of crowdsourcing, which refers to the process of information from a large group of people. Google Traffic works by analysing the determined locations transmitted to Google by a large number of mobile phone users. By calculating the speed of users along a length of road, Google can generate a live traffic map. Google processes the incoming raw data about mobile phone device locations and then excludes anomalies such as a postal vehicle which makes frequent stops. When a threshold of users in a particular area is noted, the overlay along roads and highways on the Google map changes colour (Al-Amin, 2015).

The first limitation of using such real-time data is that the information about drive time is available only for major roads and highways and not available for other roads. The second limitation is that the calculation of drive time is based on the speed of users along a length of road networks.

The data received from the Department of Transportation Management in Malaysia under (DBKL) were primarily transferred to the attribute table in ArcGIS with high precision. To compensate for the unavailable data mentioned earlier, we decided to arrange occasional field trips to locations where the data was missing. I have travelled by my car to the locations that the data was not available. As mentioned before, the average speed of the vehicle is available for the major roads such as highways. For the smaller roads where the speed of the vehicle was not available, the researcher recorded the data, and the result is the average speed of the vehicle for that particular road. Figure 4.8 shows the average speed of the vehicle for each road networks in KL that a truck might travel along the roads for the purpose of distribution of fresh vegetables. The unit for measuring the average speed of the vehicle, in this case, is Kilometre per hour. The abbreviation for Avg Speed in the following table stands for average speed in this case study.

95

|   | FID | Shape *  | LABEL                               | Lenght   | AvgSpeed |
|---|-----|----------|-------------------------------------|----------|----------|
| Т | 0   | Polyline | Exit to j kapar                     | 0.753625 | 80       |
| Þ | 1   | Polyline | To shapadu hwy                      | 0.720727 | 80       |
|   | 2   | Polyline | Ramp psn sultan ibrahim             | 0.243154 | 50       |
|   | 3   | Polyline | Ramp psn sultan ibrahim to j langat | 0.381981 | 80       |
|   | 4   | Polyline | Ramp psn sultan ibrahim             | 0.203056 | 50       |
|   | 5   | Polyline | Ramp psn sultan ibrahim             | 0.277717 | 80       |
|   | 6   | Polyline |                                     | 0.439268 | 80       |
|   | 7   | Polyline | Shapadu - meru i/c                  | 0.866414 | 80       |
|   | 8   | Polyline | Shapadu exit kapar                  | 1.079778 | 80       |
|   | 9   | Polyline |                                     | 0.820921 | 80       |
|   | 10  | Polyline |                                     | 0.394093 | 80       |
|   | 11  | Polyline | Exit 502 pandamaran                 | 0.789177 | 80       |
|   | 12  | Polyline | Exit 502 pandamaran                 | 0.773309 | 80       |
|   | 13  | Polyline | Ramp psn sultan ibrahim             | 0.611434 | 80       |
|   | 14  | Polyline | Ramp to federal highway             | 0.375867 | 80       |
|   | 15  | Polyline | Ramp to tmn berkerley               | 0.414896 | 80       |
|   | 16  | Polyline | Ramp psn sultan ibrahim             | 0.698781 | 50       |
|   | 17  | Polyline | Toll plaza sg rasau                 | 0.012951 | 80       |
|   | 18  | Polyline | Toll plaza sg rasau                 | 0.017878 | 80       |
| 1 |     |          |                                     |          |          |

Figure 4.6: Attribute Table of Average Speed for Road Networks

Figure 4.7 shows the average speed of the vehicle for all road networks in Selangor. The map illustrates the classification of the average speed of the vehicles for the entire road network. The speed for roads iscategorised into four classes. The maximum speed of the road networks is 80 Km/h according to the speed limit regulation in Malaysia for the Lorries. The minimum speed of the vehicle is 10 Km/h for the same roads in Kuala Lumpur. The advantage of the model is to consider the drive time as a dependent variable in regression analysis. The objective of this thesis is to find the key factors influencing on drive time. The calculation of drive time is based on the average speed of the vehicle in all road networks. To develop the database management system, the average speed for all the road networks was gathered to determine the efficient routes for delivery purpose. The statistical data for the average speed was available for one year, and the spatial database system was based on this information.



Figure 4.7: Map of Average Speed for Roads

When time is used as a travel expense, there are some factors affecting travel time between origin and destination such as school zone, traffic volume and one-way streets (Echols, 2003). The runtime of the vehicle is calculated by taking into account the length of the road and the speed of the vehicle on each road.

Winyoopradist and Siangsuebchart (1999) they proposed a GIS model to estimate shortest travelling time between two locations by specifying starting time. They consider travelling speed pattern for roads network which is varied in the morning and afternoon. It has been considered about traffic speed for each road and waiting time for each hour in lines and intersections. The weakness of this study is that it was quite difficult to apply the average speed of a vehicle for all the road network.

By the following theprocess, it describes the delivery routes and their directions in the 7 to 10am traffic time pattern, and the average speed of a lorry along the routes is collected

and saved into the attribute table. The travel times to all the destinations were determined by collecting data according to different time patterns. For data collection, we found that the Department of Transportation and Traffic Management under(DBKL) is the primary source of the data. Sime Darby start to deliver their first packs of fresh produce at 7 o'clock in the morning. They start to pack the fresh vegetables and send it to hypermarkets at 7 o'clock in the morning. The fresh produce should be received to hypermarkets before 12 for customers to meet in the marketplaces.

If the speed pattern for a vehicle along a route between 10-12 am is different, then the calculation for selecting the best route for the delivery of fresh vegetables will also be different. It has collected the average speed of a vehicle between 10-12 am and considered in the attribute table. Calculating the optimal routes for the delivery of fresh vegetables is based on time patterns in this model. The best decision making is then based on time of the day for the purpose of delivery and depends on hypermarkets request.

Figure 4.8shows the result obtained from data collection for an average speed of a vehicle in two time patterns. The attribute table considers two fields for the speed of a vehicle in the 7 to 10 am and 10 to 12 am. The total average speed in time between 7 to 10 is lower compared to 10 to 12 am. The caption for Vehicle\_SP stands for the car speed between 7 to 10 am and the caption for Vehicle\_1 stands for time 10 to 12 am.

| J | Vehicle_SP | Vehicle_1 |   |
|---|------------|-----------|---|
|   | 50         | 70        |   |
|   | 50         | 70        |   |
|   | 50         | 70        |   |
|   | 50         | 70        |   |
|   | 50         | 70        |   |
|   | 50         | 80        |   |
|   | 45         | 50        |   |
|   | 20         | 45        |   |
| Τ | 15         | 40        |   |
|   | 20         | 70        |   |
| 1 | 45         | 80        |   |
|   | 35         | 35        | - |
| ٠ |            | ► ►       |   |

Figure 4.8: Attribute Table of Average Speed for Two Time Patterns

Figure 4.9 describes the distribution of average speed of the vehicle along the road networks in time 7-10. The roads are classified into four classes. The maximum speed in this regard for the road is 70 Km/h. Routes could be generated during different hours during daytime in order to compare the total drive time between these predefined times intervals (Bhambulkar, 2011).



Figure 4.9: Map of Average Speed for Roads in Time Pattern 7-10

For the time pattern of 10-12, the other map will describe the speed of the vehicle for all the road networks. Figure 4.10 illustrates the speed of the vehicle for all the road network within KL. In this period, the maximum speed of the vehicle along the road is 80 Km/h.



Figure 4.10: Map of Average Speed for Road Networks in Time Pattern 10-12

According to the length of the roads then the information about the speed of vehicle in time frames are available. The result is derived from calculating the total length each route divided by the average speed of the vehicle in that particular road. The results obtained from this mathematical calculation are stored in an attribute table. The Length stands for the distance of each road networks, and the average speed is the speed of the truck on that particular road and the unit for 60 is in minutes.

# [Length / Average Speed based on time patterns] \* 60 (Minutes)

The results of the calculation are shown in the following attribute table in the ArcGIS. The fist column with a caption of Drive\_Ti\_1 stands for drive time which is based on average speed of a vehicle in the 7-10 time frame. The next column with a caption of Drive\_Ti\_2 stands for drive time based on average speed for vehicle in the 10 to 12 time frame.

|   | Attributes of Co | ombineRoadWit             | hPopWithLand |              |   |
|---|------------------|---------------------------|--------------|--------------|---|
|   | Vehicle_SP       | Vehicle_1                 | Drive_Ti_1   | Drive_Ti_2   |   |
| Þ | 50               | 70                        | 1.077697     | 0.769783     |   |
|   | 50               | 70                        | 0.667995     | 0.477139     |   |
|   | 50               | 70                        | 0.270956     | 0.19354      |   |
|   | 50               | 70                        | 0.617317     | 0.44094      |   |
|   | 50               | 70                        | 0.733611     | 0.524008     |   |
|   | 50               | 80                        | 1.243393     | 0.777121     |   |
|   | 45               | 50                        | 1.25613      | 1.130517     |   |
|   | 20               | 45                        | 2.499421     | 1.110854     |   |
|   | 15               | 40                        | 2.87339      | 1.077521     |   |
|   | 20               | 70                        | 3.995348     | 1.141528     |   |
|   | 45               | 80                        | 1.762691     | 0.991514     |   |
|   | 35               | 35                        | 2.552522     | 2.552522     |   |
|   | 30               | 45                        | 2.654144     | 1.769429     |   |
|   | 30               | 50                        | 0.066695     | 0.040017     |   |
|   | 45               | 80                        | 1.245224     | 0.700439     |   |
|   | 25               | 45                        | 1.215922     | 0.675512     |   |
|   | 35               | 80                        | 1.658244     | 0.725482     | - |
| R |                  |                           |              | •            |   |
|   | Record: II       | <ul> <li>▲ 1 ▶</li> </ul> | Show:        | All Selected | - |

Figure 4.11: Attribute of drive time for time frames

The average speed between 7 to 10 am is much slower compare 10 to 12 am. The total drive time between 7 to 10 am is greater than 10 to 12 am.

The next parameter is drive time, which needs a calculation based on the information derived from the attribute table shown in the proceeding figure. Right clicking on the drive time column will cause the appropriate window to appear, as illustrated in Figure 4.12

4.12.

For drive time variable the calculation is based on the following formula:

[Length / Average Speed] \* 60 (Minutes)

| Attributes of | of combine_ro | pad_rso                             |          |          |        | ≥ <b></b> } ¤ k 0 m ÷      |
|---------------|---------------|-------------------------------------|----------|----------|--------|----------------------------|
| FID           | Shape *       | LABEL                               | Lenght   | AvgSpeed | Drive_ | Time I .                   |
| 0             | Polyline      | Exit to j kapar                     | 0.753625 | 80       | 0.5    | Sort Ascending             |
| 1             | Polyline      | To shapadu hwy                      | 0.720727 | 80       | 0.5    | Sort Descending            |
| 2             | Polyline      | Ramp psn sultan ibrahim             | 0.243154 | 50       |        | A Advanced Section         |
| 3             | Polyline      | Ramp psn sultan ibrahim to j langat | 0.381981 | 80       | 0.2    | Z + Advanced Sorting       |
| 4             | Polyline      | Ramp psn sultan ibrahim             | 0.203056 | 50       |        | Summarize                  |
| 5             | Polyline      | Ramp psn sultan ibrahim             | 0.277717 | 80       | 0.2    |                            |
| 6             | Polyline      |                                     | 0.439268 | 80       | 0.3    | ∠ Statistics               |
| 7             | Polyline      | Shapadu - meru i/c                  | 0.866414 | 80       | 0.6    | Field Calculator           |
| 8             | Polyline      | Shapadu exit kapar                  | 1.079778 | 80       | 0.8    |                            |
| 9             | Polyline      |                                     | 0.820921 | 80       | 0.6    | <u>C</u> alculate Geometry |
| 10            | Polyline      |                                     | 0.394093 | 80       |        | Turn Field Off             |
| 11            | Polyline      | Exit 502 pandamaran                 | 0.789177 | 80       | 0.5    |                            |
| 12            | Polyline      | Exit 502 pandamaran                 | 0.773309 | 80       | 0.5    | Freeze/Unfreeze Column     |
| 13            | Polyline      | Ramp psn sultan ibrahim             | 0.611434 | 80       | 0.4    |                            |
| 14            | Polyline      | Ramp to federal highway             | 0.375867 | 80       | (      | X Delete Field             |
| 15            | Polyline      | Ramp to tmn berkerley               | 0.414896 | 80       | 0.3    |                            |
| 16            | Polyline      | Ramp psn sultan ibrahim             | 0.698781 | 50       |        | Propert <u>i</u> es        |

Figure 4.12: Field calculation for drive time

Manual input via keyboard seems to be the only way that the values in the table can be edited. In some cases, with different settings, a mathematical calculation might be needed to set a field value for a single or all of the records. Area, length, perimeter and other geometric properties of the fields in the attribute table can then be calculated.

A menu in the Field Calculator dialogue box allows the selection of how the fields will be listed. Clicking a small control to the right of the fields list accesses this menu, or it can also be accessed by right-clicking the fields list to provide the option of listing fields with their field aliases, instead of their underlying field names. When this option is chosen, the fields' list defers to the field alias properties when layers and tables are being worked with. Another option in this menu lets the user choose how the fields are sorted in the list. By default, the fields are still listed in their original order in the data source. Choosing either Sort Ascending or Sort Descending allows a user to locate the required field quickly.

The related formula that is needed in this section in ArcGIS needs to be determined, as stipulated in the appendix. The first step is to select the length from the field, followed by the suitable function for this process, culminating in the selection of this process in selecting the average speed and the final result will be multiplied with 60 to convert it into minutes.

The result of the calculation is shown in the following attribute table in the ArcGIS. The result for drive time is in minutes, which means the time that a vehicle will spend during a path going from one point to another destination is also based on minutes. Figure 4.13 illustrates the drive time column, as well as the information for all the roads in the degree of minutes.

|    | Attributes of | of combine_re | pad_rso                             | 1               | Providence     | <u> </u>   | ٢    |
|----|---------------|---------------|-------------------------------------|-----------------|----------------|------------|------|
| F  | FID           | Shape *       | LABEL                               | Lenght          | AvgSpeed       | Drive_Time |      |
| IC | 0             | Polyline      | Exit to j kapar                     | 0.753625        | 80             | 0.565219   |      |
| IC | 1             | Polyline      | To shapadu hwy                      | 0.720727        | 80             | 0.540545   |      |
| IC | 2             | Polyline      | Ramp psn sultan ibrahim             | 0.243154        | 50             | 0.291785   |      |
| iE | 3             | Polyline      | Ramp psn sultan ibrahim to j langat | 0.381981        | 80             | 0.286486   |      |
|    | 4             | Polyline      | Ramp psn sultan ibrahim             | 0.203056        | 50             | 0.243667   |      |
|    | 5             | Polyline      | Ramp psn sultan ibrahim             | 0.277717        | 80             | 0.208288   |      |
|    | 6             | Polyline      |                                     | 0.439268        | 80             | 0.329451   |      |
|    | 7             | Polyline      | Shapadu - meru i/c                  | 0.866414        | 80             | 0.649811   |      |
|    | 8             | Polyline      | Shapadu exit kapar                  | 1.079778        | 80             | 0.809834   |      |
|    | 9             | Polyline      |                                     | 0.820921        | 80             | 0.615691   |      |
|    | 10            | Polyline      |                                     | 0.394093        | 80             | 0.295569   |      |
|    | 11            | Polyline      | Exit 502 pandamaran                 | 0.789177        | 80             | 0.591883   |      |
|    | 12            | Polyline      | Exit 502 pandamaran                 | 0.773309        | 80             | 0.579982   |      |
|    | 13            | Polyline      | Ramp psn sultan ibrahim             | 0.611434        | 80             | 0.458575   |      |
|    | 14            | Polyline      | Ramp to federal highway             | 0.375867        | 80             | 0.2819     | Ŧ    |
| F  | (             |               |                                     |                 |                | +          |      |
|    | Record:       | 14 4          | 0 FI Show: All Selected Re          | cords (0 out of | 4804 Selected) | Option     | IS 1 |

Figure 4.13: Attribute table for Drive time

Another variable in this research is population. The following figure shows the attribute table that takes into consideration population as a factor. The focus of this study is to transportation of fresh vegetables and reducing the drive time is the main concern for this research. Therefore, selecting the routes passing through the area with less population is the concern of this research.

Land use and transportation planning system takes place in a highly dynamic system that involves many sub-factors. To review the sub-factors for land use, factors such as commercial places, residential places, and the river can be mentioned. Land use and transportation interaction, spatial analysis and visualisation function in GIS allows the examination of the different patterns caused by different forces related to land use and transportation, as well as their effects at different spatial and temporal scales within a specific urban area (Shaw & Xin, 2003). Land use is another parameter that will affect driving time for a vehicle along the road in this research. It has been discovered that the land use characteristics of transportation and delivery services affect the drive time for a vehicle travelling on a road. The land use is classified into many other features. The examples of land use are such as residential area, commercial area and agricultural land.

Figure 4.14 shows the land use parameter, along with sub-parameters such as commercial area, residential area, river and open spaces.

| Attributes of CombineRoadWithPopWithLanduseWithSchool |         |            |            |                 |            |                    |            |            |          |          |  |  |  |
|---|---------|------------|------------|-----------------|------------|--------------------|------------|------------|----------|----------|--|--|--|
| Г   | ONE_WAY | CAR_Volume | Vol_Vehicl | Vehicle_SP      | Vehicle_1  | Drive_Ti_1         | Drive_Ti_2 | Population | LANDUSE  | ^        |  |  |  |
| IC  |         | 23444      | 3400       | 60              | 80         | 0.751014           | 0.563261   | 0          |          |          |  |  |  |
| IC  | FT      | 12345      | 1234       | 60              | 80         | 1.5849             | 1.188675   | 0          |          |          |  |  |  |
|   | ]       | 12345      | 2345       | 60              | 80         | 0.522764           | 0.392073   | 0          |          |          |  |  |  |
|   |         | 34567      | 2345       | 60              | 70         | 0.550844           | 0.472152   | 0          |          |          |  |  |  |
|   |         | 23456      | 3456       | 60              | 80         | 0.762987           | 0.572241   | 0          |          |          |  |  |  |
| IC  |         | 23456      | 3456       | 50              | 70         | 1.360708           | 0.971934   | 0          |          |          |  |  |  |
|   |         | 0          | 0          | 50              | 70         | 0.726543           | 0.51896    | 1742       | Industri |          |  |  |  |
|   |         | 0          | 0          | 45              | 50         | 0.297381           | 0.267643   | 1742       |          |          |  |  |  |
|   |         | 0          | 0          | 50              | 70         | 0.494854           | 0.353467   | 1742       |          |          |  |  |  |
|   |         | 0          | 0          | 50              | 70         | 0.621623           | 0.444016   | 1742       | Kediaman |          |  |  |  |
|   |         | 0          | 0          | 50              | 70         | 0.423231           | 0.302308   | 1742       |          |          |  |  |  |
|   | FT      | 0          | 0          | 0               | 0          | 0                  | 0          | 0          |          |          |  |  |  |
| L   |         | 0          | 0          | 0               | 0          | 0                  | 0          | 0          |          |          |  |  |  |
|   |         | 0          | 0          | 40              | 45         | 0.659885           | 0.586565   | 2160       |          |          |  |  |  |
|   |         | 0          | 0          | 25              | 40         | 1.059302           | 0.662064   | 2160       |          |          |  |  |  |
| L   |         | 0          | 0          | 55              | 70         | 1.225198           | 0.962655   | 0          | Kediaman |          |  |  |  |
| L   |         | 0          | 0          | 45              | 65         | 2.900105           | 2.007765   | 0          | Kediaman |          |  |  |  |
| L   |         | 0          | 0          | 25              | 70         | 3.305074           | 1.180384   | 0          |          |          |  |  |  |
|   |         | 0          | 0          | 24              | 60         | 0.996368           | 0.398547   | 0          |          |          |  |  |  |
|   |         | 0          | 0          | 30              | 40         | 1.166867           | 0.875151   | 0          |          |          |  |  |  |
|   |         | 0          | 0          | 30              | 40         | 0.826014           | 0.619511   | 0          |          |          |  |  |  |
| L   |         | 0          | 0          | 30              | 40         | 0.450849           | 0.338137   | 0          |          |          |  |  |  |
| L   |         | 0          | 0          | 45              | 70         | 0.830735           | 0.534044   | 1742       |          |          |  |  |  |
|   |         | 0          | 0          | 45              | 75         | 1.0421             | 0.62526    | 1742       |          | <b>T</b> |  |  |  |
|   |         |            |            |                 |            |                    |            |            |          | +        |  |  |  |
|   | Record: | 14 4       | 1 🕨 🖬 Sh   | now: All Select | ted Record | s (0 out of 2168 s | Selected)  | Options 👻  |          |          |  |  |  |

Figure 4.14: Attribute Table for Population and Land Use

# 4.2.1 Adding school zones

Another parameter affecting the time for the delivery of fresh vegetables is the location of schools in the map during transportation procedure. This step involves the identification of all schools on the digital map in KL and Klang Valley, shown in Figure 4.15. Pinpointing the school location allows the creation of a 100-meter buffer zone from the main roads network, so the average speed for these types of roads within this distance are usually 35 Km/h. Also, the drive time will increase passing through this line.



Figure 4.15: Map of Schools' Locations

One of the most common examples of network attribute is cost, which functions as impedances over the network. Restriction attributes prohibiting traversal in both or one direction, for example in one-way roads. The one-way attribute table models the one-way traffic restriction that requires the adherence of drivers (ESRI, 2012).

The next stage defines the two-way variable, which covers the nature of the road; whether the road is either one or two-ways. This compels the provision of the attribute table for each road in order to see how the directions of each road are being considered.

Addressing the new variable is the identification of the car volume for each specific road. In the real world, car volume means the total capacity of each road in Kuala Lumpur. Figure 4.16 illustrates the number of cars in each road and the effect of this parameter on travelling time for the delivery of fresh vegetables involved in the network analysis

process. The caption for CAR\_Volume in attribute table stands for car volume for consideration in network analysis.

|     | Attributes of combine_road_rso |          |                                     |                   |               |            |         |            |   |  |  |  |
|-----|--------------------------------|----------|-------------------------------------|-------------------|---------------|------------|---------|------------|---|--|--|--|
| Ir  | FID                            | Shape *  | LABEL                               | Lenght            | AvgSpeed      | Drive_Time | ONE_WAY | CAR_Volume |   |  |  |  |
| IE  | 1                              | Polyline | To shapadu hwy                      | 0.720727          | 80            | 0.540545   | FT      | 2345       |   |  |  |  |
| JE. | 2                              | Polyline | Ramp psn sultan ibrahim             | 0.243154          | 50            | 0.291785   |         | 3234       |   |  |  |  |
| IE  | 3                              | Polyline | Ramp psn sultan ibrahim to j langat | 0.381981          | 80            | 0.286486   | FT      | 3444       | 1 |  |  |  |
|     | 4                              | Polyline | Ramp psn sultan ibrahim             | 0.203056          | 50            | 0.243667   |         | 434        | 1 |  |  |  |
| IE  | 5                              | Polyline | Ramp psn sultan ibrahim             | 0.277717          | 80            | 0.208288   | FT      | 3434       | 1 |  |  |  |
| IIC | 6                              | Polyline |                                     | 0.439268          | 80            | 0.329451   | FT      | 3434       | 1 |  |  |  |
|     | 7                              | Polyline | Shapadu - meru i/c                  | 0.866414          | 80            | 0.649811   | FT      | 65345      | 1 |  |  |  |
| IIC | 8                              | Polyline | Shapadu exit kapar                  | 1.079778          | 80            | 0.809834   | FT      | 34543      | 1 |  |  |  |
| IIC | 9                              | Polyline |                                     | 0.820921          | 80            | 0.615691   | FT      | 44345      | 1 |  |  |  |
|     | 10                             | Polyline |                                     | 0.394093          | 80            | 0.295569   |         | 45454      | - |  |  |  |
|     | Record:                        | • •      | 12 FI Show: All Selected Rec        | ords (0 out of 44 | 412 Selected) | Options    | -       |            |   |  |  |  |

Figure 4.16: Attribute Table for One-Way Roads and Car Volume

## 4.3 GIS regression model

Regression analysis allows you to model and examine the spatial relationship, and can help explain the factors behind observed. When the GIS regression used properly, this method is powerful and reliable statistics for estimating linear relationship.

Although many other models for improving delivery services in the field of fresh foods have been reviewed in Chapter2, those models are based on mathematical equations and are not widely approved among decision makers for a better and more accurate prediction of road traffic for selecting the right routes in delivery. The weakness of previous models conducted in GIS is that those models are validated by a small road network instead of the actual road network, and the cost of collecting those actual road network attributes is rather high.

This research has adopted the GIS tool to solve their work related problems such as retail businesses, which increases the efficiency in the delivery of goods and services. They would like to know the potential of a new store location, as well as realise the potential customers in their vicinity. These such questions are based on the number of spatially related factors, which can be solved with GIS technology (Bolstad, 2012).

This research concentrates on fresh vegetables and fast delivery services and route selection to avoid congested areas. The punctuation for this research is the primary key and reaching the market places in time and also to fulfil the demand requirements.

The main objective of this thesis is to select the fastest route for delivery purpose. This requires the prediction or estimation of the traffic situation in order to arrive at informed decisions regarding routes. In this regard, the regression model is suitable and reliable for this study, which requires predictions and estimates the time as a dependent variable to a number of independent variables in an equation determining the key factors affecting the delivery process with respect to time for fresh vegetables. The linear model usually uses GIS-based decision rules (Malczewski, 2000).

Regression models use map overlay operations in a GIS to combine all the independent variables needed for the analysis. According to the objective of this thesis, and the nature of the variables in the GIS field, the linear regression model is to be used, due to the fact thatdependent and independents variables are all numeric in nature. For this kind of study, different layers in the GIS are used, such as road networks, land use and population.

One of the reasons to adopt the regression model in this study is that it provides the decision makers on the delivery of foods with a powerful set of tool, which allows the prediction of the present or future events to be made via the knowledge accumulation of past or present events (Stockburger, 1996).



Figure 4.17: Distribution Cost Forecasting (SIMAFORE)

#### 4.3.1 Data observations

After conducting the suitable research methodology for this study, the next stage is to implement the appropriate tools to identify the relevant variables for the hypothesis test. By using appropriate statistical techniques without any conflict between the research method and various variables, the data will be collected, classified and analysed for the purpose of responding to the research questions (Khaki, 1995).

Data evaluation and justification is a process with multiple steps, involving data that has been collected via different sources and tasks that will duly summarise, classify and finally process the known relationships among the variables, providing the situation for hypothesis tests (Khaki, 1995).

Data analysis is a multi-level process, in which the information that has been gathered in different ways will be summarised, categorised and finally processed, in order to provide various analyses and gleam the relationships of the data with a view to examine the hypotheses. In this process, the data in both conceptual and experimental aspects are refined, and various statistical techniques play an important role in inferences and generalisation regarding statistical process (Khaki, 1995). Table 4.1 illustrates the information relating to the drive time for each road, the speed of a vehicle during two periods of time, the distance from each road, the population for each area, and the car volume that is the number of the cars in a specific area.

| X            | TIMESPEEDSPEED7-1010-12 |        | LENGTH  | Populatio<br>n | CARVOLUM<br>E |           |
|--------------|-------------------------|--------|---------|----------------|---------------|-----------|
| Mean         | 0.9988                  | 37.7   | 54.0821 | 1.140          | 2233.510      | 21860.51  |
| Median       | 0.6432                  | 40     | 60      | 0.677          | 1321.281      | 15678     |
| Maximum      | 4.9921                  | 70     | 80      | 4.837          | 15767.95      | 87996     |
| Minimum      | 0.0053                  | 5      | 11      | 0.004          | 500.932       | 1042      |
| Std. Dev.    | 0.934                   | 11.470 | 14.252  | 1.114          | 2327.574      | 17158.510 |
| Skewness     | 1.301                   | -0.360 | -0.366  | 1.141          | 2.548         | 1.120     |
| Kurtosis     | 4.481                   | 2.795  | 2.595   | 3.243          | 10.992        | 3.952     |
|              | 780.30                  |        |         |                |               |           |
| Jarque-Bera  | 0                       | 37.879 | 47.177  | 473.604        | 3934.478      | 147.695   |
| Probability  | 0                       | 0      | 0       | 0              | 0             | 0         |
| Observations | 2089                    | 1620   | 1620    | 2157           | 1051          | 598       |

 Table 4.1: Information Regarding Data Abundance

All the tables illustrating the quantity of all variables for this research as well as the description about other relevant statistical information are gathered in the appendix. In summary, the proper use of descriptive statistics can precisely describe the characteristics of the categorised information. Descriptive statistics are always used to determine the expression of characteristics of research data (Azar, 2005).

## 4.3.2 Table of abundance distribution

Abundance distribution is the organisation of data or observation into classes, along with their abundance in related level. To form a table of abundance distribution, the scope of the changes, the number and the size of classes must be calculated by the formula, and then followed by setting up the table with two columns X (column classes) and F (abundance classes). Creating an abundance distribution table is an economical and simple way to achieve the performance of a multitude and irregular types of data. However, in the classification, some data will be lost due to category errors, and the results are also reflective on the calculation of statistical indicators. However, these kinds of errors are minor and do not play any role in the statistical calculation.

The figures below show the tables of abundance distribution for all of the variables analysed in this research. The first figure 4.18, which is the distribution of average speed in the period between 07:00-10:00 AM, illustrating the information regarding the statistical value in this case, as shown to the right of the table.



Figure 4.18: Table of Abundance Distribution for Average Speed at7am to 10am

The second figure is the abundance distribution of the speed limit between 10:00-12:00, and illustrates the statistical description regarding this particular parameter.



Figure 4.19: Table of Abundance Distribution for Average Speed at10am to 12am

The next figure illustrates the distribution abundance of distance for each particular route and identifies the statistical information contributing to this thesis.



Figure 4.20: Table of Abundance Distribution for Distance

The figure below shows the distribution abundance of the population density in different zones for the study area.



**Figure 4.21: Table of Abundance Distribution for Population** 

Car volume is a parameter considered in this study, and the abundance distribution, as well as statistical information is shown in the figure below.



Figure 4.22: Table of Abundance Distribution for Car Volume

The two-way parameter is considered for the roads, which allows the trucks to commute in two directions on selected roads. The statistical information and distribution are shown in the figure below.



Figure 4.23: Table of Abundance Distribution for Two-Ways Directions

The obstacle in this study is school area, where speed limits restrict the drivers according to law. The speed limit for such roads is 35 Km/h. Such information gathered from transportation monitoring department. The distribution abundance for this parameter is shown in the figure below.



Figure 4.24: Table of Abundance Distribution for School Area

The residential area forms a sub-parameter of land use. In this case, the distribution abundance is illustrated in the appendix. The industrialarea is another sub-parameter from land use as well. The observation and samples regarding the information are tabulated in the appendix. Open space is also another sub-parameter for the land use variable.

# 4.3.3 Estimation of correlation

. In correlation analysis, without defining variables as dependent or independent, we are looking for the form of the relation of variables among themselves. Table 4.2 shows the results from the Pearson Correlation, along with their related probability values. A probability value of less than 5% means the acceptance of statistical correlation between the two variables above. Table 4.2 shows the correlation result between the variables in our model.

| Covar  | iance An  | alysis:  | Ordina   | ry       |           |         |        |        |        |      |
|--------|-----------|----------|----------|----------|-----------|---------|--------|--------|--------|------|
| Sampl  | e (Adjus  | ted): 2  | 2165     | •        |           |         |        |        |        |      |
| Includ | led Obser | rvations | s: 288 A | After Ad | ljustme   | nts     |        |        |        |      |
| Balan  | ced Sam   | ole (Lis | t wise ] | Missing  | Value     | Deletio | n)     |        |        |      |
| Durun  |           |          |          |          | rrelation |         | )      |        |        |      |
|        |           |          |          | Pr       | ohahilit  |         |        |        |        |      |
|        | TI        | 7-       | 10       | LE       | CA        | y<br>P  | SC     | R      | IN     | OP   |
|        | ME        | 10       | -12      | N        | R         | OP      | Н      | ES     | D      | EN   |
| TIME   | 1.00      |          |          |          |           |         |        |        |        |      |
|        |           |          |          |          |           |         |        |        |        |      |
|        |           | 1.       |          |          |           |         |        |        |        |      |
| 7-10   | 0.03      | 00       |          |          |           |         |        |        |        |      |
|        | 0.40      |          |          |          |           |         |        |        |        |      |
|        | 0.49      |          | 1        |          |           |         |        |        |        |      |
| 10-12  | 0.12      | 47       | 00       |          |           |         |        |        |        |      |
|        |           | 0.       |          |          |           |         |        |        |        |      |
|        | 0.04      | 00       |          |          |           |         |        |        |        |      |
| IFN    | 0.86      | 0.<br>24 | 0.<br>23 | 1.0      |           |         |        |        |        |      |
|        | 0.00      | 0.       | 0.       |          |           |         |        |        |        |      |
|        | 0.00      | 00       | 00       | -        |           |         |        |        |        |      |
|        |           | 0.       | 0.       | 0.4      | 1.0       |         |        |        |        |      |
| CAR    | 0.39      | 24       | 17       | 8        | 0         |         |        |        |        |      |
|        | 0.00      | 0.       | 0.       | 0.0      | -         |         |        |        |        |      |
|        |           | 0.       | -        | 0.5      | 0.1       | 1.0     |        |        |        |      |
| POP    | 0.60      | 05       | 0.05     | 6        | 9         | 0       |        |        |        |      |
|        | 0.00      | 0.       | 0.       | 0.0      | 0.0       |         |        |        |        |      |
| SCHO   | 0.00      | 43       | - 40     | 0.0      | -         | 0.1     | 1.0    |        |        |      |
| OL     | 0.19      | 0.06     | 0.24     | 5        | 0.01      | 6       | 0      |        |        |      |
|        |           | 0.       | 0.       | 0.3      | 0.8       | 0.0     |        |        |        |      |
|        | 0.00      | 28       | 00       | 8        | 5         | 1       | -      | 1.0    |        |      |
| RES    | 0.56      | 0.02     | 0.<br>04 | 0.5      | 0.2       | 0.5     | 0.0    | 1.0    |        |      |
| THE    | 0100      | 0.       | 0.       | 0.0      | 0.0       | 0.0     | 0.3    |        |        |      |
|        | 0.00      | 67       | 53       | 0        | 0         | 0       | 4      |        |        |      |
| IND    | 0.01      | 0.       |          | 0.0      | 0.1       | 0.0     | 0.0    |        | 1.     |      |
| IND    | 0.01      | 03       | 0.02     | 0.8      | 0.0       | 4       | 0.5    | 0.15   |        |      |
|        | 0.85      | 66       | 69       | 4        | 9         | 1       | 6      | 1      |        |      |
| OPEN   | - 0.56    | 0.<br>00 | - 0.03   | - 0.54   | - 0.28    | 0.34    | - 0.07 | - 0.59 | - 0.27 | 1.00 |
|        | 00        | 0.       | 0.       | 0.0      | 0.0       | 0.0     | 0.2    | 0.0    | 0.     | 1.00 |
|        | 0.00      | 95       | 67       | 0        | 0         | 0       | 3      | 0      | 00     |      |

|  | <b>Fable 4.2: Probabilit</b> | Value Result from | Pearson Correlation |
|--|------------------------------|-------------------|---------------------|
|--|------------------------------|-------------------|---------------------|

According to our results on estimation the correlation coefficients, the model does not show a high correlation between the independent variables, so it is possible to streamline these variables in the form of multiple linear regression model. According to the information in the table, LEN stands for length, means the distance between locations and pop stands for Population density, the number of people living within the area of each road network. SCH stands for school zone in the vicinity of each roads networks. RES means the classification of land use variables as a significant parameter in this research. IND means the Industrial area under land use classification. Open stands for Open space area.

## 4.4 Statistical calculation and prediction results

# 4.4.1 Durbin-Watson Test (D-W)

Durbin-Watson Test may be one of the most common tests for the inaccuracy of explanation of the model under study designed for identification of consecutive correlation in statistical information. Consecutive correlation is used to explain the condition in which the amounts of a variable are correlated to each other. Consecutive correlation or self-correlation both are used to indicate the breaking of independence assumption in disruption sentence. Conducting this test will improve the statistical credibility of the obtained estimations in any of the regression models since in this condition the obtained amounts of  $R^2$  can be considered as a correct criterion for explanatory nature of the independent variable (Khaki, 1995).

Table 4.3 shows the results of all variables after conducting the statistical test. The stars on the right side of each number mean that there is a statistical value in this regard. The (\*\*\*) and (\*\*) means that variables appear in a 1% and 5% error level afterexamining the test.

| Specification  | Intercept              | Intercept<br>and trend | None               |
|----------------|------------------------|------------------------|--------------------|
| Unit root test | Levin, Lin &<br>Chu t* | Levin, Lin &<br>Chu t  | Levin, Lin & Chu t |
| LTIME          | $-4.040^{***}$         | -4.292***              | -3.634***          |
| LSPEED7TILL10  | -8.762***              | -8.758***              | -3.860***          |
| LSPEED10TILL12 | -13.915***             | -14.032***             | -3.961***          |
| LLENGTH        | -5.949***              | -6.072***              | -3.989***          |
| TWO-WAY        | -8.544***              | -8.931***              | -2.185**           |
| SCHOOL         | -21.575***             | -21.575***             | -7.482***          |
| LCARVOLUME     | -5.847***              | -5.779***              | -4.355***          |
| RESIDENTIAL    | -8.290***              | -8.338***              | -5.921***          |
| INDUSTRY       | -43.711***             | -43.703***             | -29.539***         |
| OPENSPACE      | -8.175***              | -8.220***              | -3.262***          |

**Table 4.3: Statistical Test** 

# 4.5 Creating the network analyst from dataset

When creating a network dataset, choices that determine which edge and junction elements are created from source features will be made. Ensuring that edges and junctions are formed correctly is important to ensure the accuracy of the network analysis results.

The Network Dataset wizard guides a user through setting up a network dataset, which can be created from a shapefile or feature classes within a feature dataset. This wizard helps identify the feature classes that will be used as sources, and the role they play in the network, along with specifying the connectivity within the network, and identifying the network attributes (ESRI, 2012).Figure 4.25 shows the connectivity example in the road network.



Figure 4.25: Connectivity Example (Esri, ArcMap,2014)

The first step in this regard is to open Arc Catalog, which is one the applications within the ArcGIS software. Navigating to the Network Analyst dataset folder, which contains the shape file, as it is displayed by the following window. Figure 4.26presents the procedures.



Figure 4.26: Creating Network Analysis

The second step involves navigating to our folder in order to create a new data set for a related folder. A window will then appear detailing the name for our dataset. The name given for network dataset is CombineRoadWithPopWithLanduseWithSchoolND. It will proceed by clicking on next button. The example of this kind of figure is shown in the appendixes of this thesis.

To modify the connectivity policy of a source, click in the Connectivity Policy column for the source and choose the connectivity policy from the list. Edge source connectivity can be set to End Point or Any Vertex. For this research, it is required to select the Any Vertex option for further processing. The figure in appendixes at the end of this thesis illustrates the way on how to choose connectivity.

There are two options for selecting connectivity; one is any vertex, and the other one is the end point. For this study, the right choice is to connect to any vertex option. For this feature street class, all roads connect to each other at any vertex point.

In this stage of creating a new dataset, the dataset has elevation fields so make sure the elevation field data is chosen. Elevation settings in a network dataset further define connectivity. In this case, it has to select yes for model turns in creating the dataset. The global turn delay evaluator assigns a cost value for transitioning between two edge elements based on the deflection angle between the two edges and the hierarchy attribute value of each edge (ESRI, 2012).

| ( Ne | two    | ork Dataset            |                |                    | -                | 3 ×        |
|------|--------|------------------------|----------------|--------------------|------------------|------------|
| Spe  | cify t | the attributes for the | e network data | iset:              |                  | Add        |
| 1    | Θ      | Name                   | Usage          | Units              | Data Ty          |            |
|      | Θ      | Oneway                 | Restricti      | Unknown            | Boolean          | Remove     |
|      | 0      | lght<br>Drive_Time     | Cost<br>Cost   | Kilomet<br>Minutes | Double<br>Double | Remove All |
|      |        | AvgSpeed<br>Vehicle_SP | Cost<br>Cost   | Unknown<br>Unknown | Double<br>Double | Rename     |
|      |        | Vehicle_1<br>PENDUDUK  | Cost<br>Cost   | Unknown<br>Unknown | Double<br>Double | Duplicate  |
|      |        | CAR_Volume             | Cost           | Unknown            | Double           | Ranges     |
|      |        |                        |                |                    |                  | Parameters |
|      |        |                        |                |                    |                  | Evaluators |
|      |        |                        |                |                    |                  |            |
|      |        |                        |                | < Back             | Next >           | Cancel     |

Figure 4.27: Attributes for the Network Dataset

The next step is to click on the Network Dataset window and try to give a specific name for this new network dataset. This process will continue by clicking on the next button. Here, it has to modify the connectivity with elevation field data. To change the field used to provide elevation data, click under the Field column and click the correct field from the drop-down list. Figure 4.27 describes this process.

## 4.5.1 Directions

The next stage has to identify the directions for road network dataset. The information regarding this process is as following.

Directions are turn-by-turn instructions on how to navigate a route. They can be created for any route that is generated from a network analysis as long as the network dataset supports them. The minimum requirements for a network dataset to support directions are as follows:

- A length attributes with length units
- At least one edge source

• At least one text field on the edge source

The directions generated when a route is computed customisable at the dataset network level. This means that the street names used for reporting directions and other related information are stored with the network dataset schema. These settings can be modified to customise the directions. The units used to report directions, and the fields used to identify streets can be modified on the General tab, which is shown in the graphic below (ESRI, 2012). Figure 4.28 displays a summary of all the settings that has been made through the developing process of the network dataset for review. By pressing the Finish button, it will create the new dataset for research analysis. After this process, a progress bar opens showing that the Network Analyst is creating network dataset. Once the network is created, the system asks if you want to build it. The build process determines which network elements are connected and populates the attributes of the network dataset. You must build the network before you can perform any network analysis on it.

| New | v Network Dataset   | ? ×    |
|-----|---|--------|
|     | Summary:  |        |
|     | Name: CombineRoadWithPopWithLanduseWithSchool_ND<br>Type: Shapefile-Based Network Dataset               | ▲<br>E |
|     | Sources:<br>Edge Sources:<br>CombineRoadWithPopWithLanduseWithSchool                                    |        |
| 0   | Connectivity:<br>Group 1:<br>Edge Connectivity:<br>CombineRoadWithPopWithLanduseWithSchool (Any Vertex) |        |
|     | Tums:<br><global tums=""></global>  |        |
|     | Attributes:<br>Oneway:  | Ŧ      |
| -   |   |        |
|     | < Back Finish   | Cancel |

Figure 4.28: Summary of Network Dataset Establishment

#### 4.6 Conclusion

The objective of this chapter is to discuss the data preparation and identification of the variables for analysis in ArcGIS. From a statistical point of view, the regression model was utilised in order tofulfil the aim of this thesis. Finally, the dataset was created to determine the variables for further analysis in ArcGIS. Many companies related to delivery services of fresh foods have managed to substantially reduce costs when routes are managed using the GIS technology. The GIS software enhances the ability of a distributor board to deliver the fresh vegetables more efficiently to their respective destination. The integration of modelling with a statistical calculation to identify indicators is classic GIS applications, and well suited to this architecture.

The results of spatial data analysis in a GIS field often uncover the needs for more data evaluation. Hence, there are always several iterations through the collection, organisation, analysis, output and assessment steps before a final decision is made. The methodology proposed for this study can assist transportation manager of companies in optimally deliver their fresh vegetables to their customers in a proper way and time to reduce their total cost and to avoid the deterioration of their fresh produce. Now a day the GIS furnished with flow traffic information and optimisation concern offers a powerful platform for transportation organisation. The research presents a promising research area requiring further exploration.

What makes this created model useful is the fact that a transport manager can change the hypermarket locations based on demands from each market every day. The directions will duly change according to the numbers of markets entered into the model and time patterns.

The method of GIS-based regression analysis appears to offer an effective method for mapping the routes for delivery purpose and identify the most important key factors, which have a significant effect on optimal transportation. The travel time was identified for the road networks for delivery services and further process of analysis of the data. The travel times towards all of the destinations were determined by collecting data in accordance with different time patterns. The travel time on routes will be used to calculate travel time saving gained from vehicle trips. The pathfinding module of ArcGIS can be used to generate the shortest and fastest routes between points.

university

# **CHAPTER 5: RESULTS OF THE MODELS AND DISCUSSIONS**

#### 5.1 Introduction

This chapter describes the data evaluation and analysis in the ArcGIS software, which resulted in the output of several digital maps that enables better decision making. Section 5.2 expands upon the information regarding functions in a Network Analyst. This began with hypermarkets, and information regarding their location and receiving time-windows, respectively. The subsequent sections continue to describe the data analysis in different scenarios, based on selected variables. The selected routes are chosen by the decision maker in several phases, based on the distance, driving time, average speed, population density, land use, car volume and different daytime as impedances. Various maps describe the situation of fast routes, and take into account the restrictions, in order to narrow down the best choice. Several modelling of the optimal route was conducted for the delivery of fresh vegetables based on selected impedances studied in the literature review. The regression model has been applied in this study, and the variables were tested via this model. The most significant variables are taken into account in the model, and the one that has not contributed any value to the drive time is omitted from the analysis one by one. In the final process in the regression model, the proper formula is identified as the variables demonstrating the most significant place in the context of this study. The following section evaluates the statistical analysis and data assessment to estimate the percentage of each parameter that affects the drive time for delivery. The conceptual framework of the chapter is depicted in Figure 5.1.



**Figure 5.1: The Conceptual Framework of The Final Map** 

# 5.2 Functions in network analyst

The Network Analyst functions embedded in the ArcGIS software allows us to solve more common network problems. The functions are such as finding the closest facility, finding service areas and creating an OD (Origin-Destination) cost matrix. The consideration of this thesis is to solve the distribution problem of finding efficient routes by applying the Network Analyst function. The database management system was developed norder to facilities the both spatial and non-spatial in distribution planning.

## 5.2.1 Finding the best route

Whether finding a simple route between two locations, or one that visits several locations, people will always prefer the best route. However, the definition of best route is highly fluid and reliant on its current situation or settings. The best route can be the quickest, shortest, or most scenic route, depending on the selected impedance. If the impedance is time, then the best route is the quickest route. Hence, the best route can be defined as the route that has the lowest impedance, where the impedance is user-selected. Any valid network cost attribute can be used as the impedance during the determination of the best route (ESRI, 2012). In this research the impedance is time, and the objective is to determine the fast routes based on significant parameters obtained from the regression analysis.

# 5.3 Names of hypermarkets and their locations on the map

According to the research objectives and requirements for this thesis, the information regarding the locations of hypermarkets must be categorised in the map. In fact, there are target markets for delivery purposes around Kuala Lumpur. The lists of the hypermarkets and detailed information are summarised in Table 5.1.

 Table 5.1: The Information Chart for Hypermarkets as Destination for Delivery Purpose

| No. | DESTINATION                  | STATE                   | DELIVERY<br>(TIMES) | RECEIVING TIME<br>WINDOW |
|-----|------------------------------|-------------------------|---------------------|--------------------------|
| 1   | Jusco Bandar Utama<br>(AEON) | Selangor                | Daily               | 7:00AM- 12:00 PM         |
| 2   | Jusco Mid Valley (AEON)      | Willayah<br>Persekutuan | Daily               | 7:00 AM - 12:00 PM       |
| 3  | Jusco WangsaMaju         | Willayah    | Daily        | 7:00 AM - 12:00 PM |
|----|--------------------------|-------------|--------------|--------------------|
|    | (AEON)                   | Persekutuan |              |                    |
| 4  | Jusco Taman Maluri       | Selangor    | Daily        | 7:00 AM - 12:00 PM |
|    | (AEON)                   |             |              |                    |
| 5  | Jusco Kepong (AEON)      | Selangor    | Daily        | 7:00 AM - 12:00 PM |
| 6  | Giant Shah Alam          | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 7  | Giant Setiawangsa        | Willayah    | As per order | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 8  | Giant Bandar Sunway      | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 9  | Giant Amcorp Mall        | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 10 | Giant South City         | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 11 | Giant Maju Junction      | Willayah    | As per order | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 12 | Giant Taman Permata      | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 13 | Giant Bukit Antarabangsa | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 14 | Giant Ampang             | Willayah    | As per order | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 15 | Giant Batu Cave          | Selangor    | As per order | 7:00 AM – 4:00 PM  |
| 16 | AEON Subang Jaya         | Selangor    | As per order | 7:00 AM-12:00 PM   |
| 17 | AEON Mid Valley          | Willayah    | As per order | 7:00AM -12:00 PM   |
|    |                          | Persekutuan |              |                    |
| 18 | AEON Sri Petaling        | Willayah    | As per order | 7:00AM- 12:00 PM   |
|    |                          | Persekutuan |              |                    |
| 19 | AEON WangsaMaju          | Willayah    | As per order | 7:00AM- 12:00 PM   |
|    |                          | Persekutuan |              |                    |
| 20 | AEON Tropicana Mall      | Selangor    | As per order | 7:00AM- 12:00 PM   |
| 21 | AEON Cheras              | Selangor    | As per order | 7:00AM- 12:00 PM   |
| 22 | AEON Kepong              | Selangor    | As per order | 7:00AM- 12:00 PM   |
| 23 | Cold Storage Mutiara     | Willayah    | As per order | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 24 | Cold Storage Bangsar     | Willayah    | As per order | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 27 | Cold Storage KLCC        | Willayah    | Daily        | 7:00 AM – 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 28 | Cold Storage             | Selangor    | Daily        | 7:00 AM – 4:00 PM  |
|    | MutiaraDamansara         |             |              |                    |
| 29 | Central Hypermarket      | Willayah    | Daily        | 8:00 AM - 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 30 | Shop Taman Segar         | Willayah    | Daily        | 8:00 AM - 4:00 PM  |
|    |                          | Persekutuan |              |                    |
| 31 | ISETAN KLCC              | Willayah    | Daily        | 7:00 AM - 4:00 PM  |
|    |                          | Persekutuan |              |                    |

The receiving time window is also displayed for each market on the table. The exception is that for Carrefour markets, it is compulsory that the delivery services meet the customer's requirements before 12 noon. Selecting hypermarket as a destination for the delivery of fresh vegetables is based on the customer demand for that particular day. It has to be mention that all of the current Carrefour hypermarkets and supermarkets in

Malaysia will be fully re-branded into "ÆON Big". In this regard, the name is changed to AEON Big instead of Carrefour.

Figure 5.2 is the map of all of the hypermarkets in the KL district for this study, and from its initial point, it is going to be distributed throughout the city, and the purpose of this study is to determine the efficient route to these locations depends on customer demand. Decision making depends on the transportation manager whether to which hypermarkets will distribute the fresh produce and when the customer will receive the product based on calculating the travelling time. Every day the manager is responsible for running the GIS software based on consumer demand and location of hypermarket.



Figure 5.2: The Location of Hypermarkets (Cartographic and GIS Lab, UM)

The start point for delivery services is the entrance gate of Kuala Lumpur to the city centre. The start point here is the tollgate, aptly named Sungai Besi as it showing in Figure 5.3. This location is the start point for distribution the goods to KL city and from this

point, the road networks will be divided into many other addresses and places. As it has mentioned in the study, the main place for producing fresh vegetables is in Seremban, but there is only one main road until the tollgate in Sungai Besi.



Figure 5.3: Location of Starting Point (Sime Darby Bhd)

## 5.4 Street data processing

The street maps that are used in the ArcGIS Network Analyst require a special format. This special requirement demands extra attention, and can be time-consuming to create. Some streets can show passes in both directions, others only one. It also needs to take into consideration one-way roads. This particular issue is dependent upon two aspects, where first, how the sequences of the vertices are arranged in the digitising process and second, the value of the one-way field. The value of the one-wayfield is 'FT' and 'TF'. FT stands for From-To, while TF stands for To-From. If the value is FT, it means that the street is a one-way street, with the direction taken from the earlier digitised vertices to the latest digitised vertices.

### 5.4.1 Restrictions

Restrictions are identified for particular features, so during analysis, restricted objects cannot be traversed. For example, one-way streets can be modelled with a restriction attribute, so they can only traverse from one end to another, and not in the traverse direction. In all cases, a restriction attribute is defined using a Boolean data type (ESRI, 2010).

### 5.5 Scenarios

## 5.5.1 Phase one: Modelling of routes in terms of distance

The first phase of our study required the use of GIS tool to develop a spatial database that contains the detailed digital information about the routes in terms of distance. The purpose of this point is to find the best routes in terms of distance to selected supermarkets in Kuala Lumpur. The following map shows the best routes for the delivery of fresh vegetables to their respective final destinations. The data model, which is generated for this kind of model, is in the vector mode. The ArcGIS Network Analyst extension is capable of generating and analysing the network dataset. The street networks are stored in the form of a network data set that is created from the feature sources in the network. Network attributes are the properties of the network elements that control mobility, and they can be used to model restrictions and impedances.

Since the request from each hypermarket for fresh vegetables is varied day-by-day and depends on customer demand to be delivered, so this matter should be taken into consideration. Therefore, the decision manager of the company will calculate the best routes based on the demand from the hypermarkets every day. The following analysis finds best routes from the main point to the respected markets based on the customer demand when the company decides to deliver their fresh vegetables. Therefore, every day depends on the orders from the hypermarkets then the transportation manager makes a decision based on the total customer demand and total number of markets. So each day the decision maker has to calculate the best routes based on the total order by markets and solve the problem by ArcGIS tool.

The first selection of routes for the delivery of fresh vegetables is based on the distance as the impedance. Figure 5.4 shows the shortest routes in terms of length, from the main point for the delivery of fresh vegetables, to the selected hypermarkets in Kuala Lumpur. The final result which has obtained from the statistical analysis reveals that distance has significant value and influences on drive time for the delivery purpose. For this reason, we applied GIS model to calculate the shortest distance.



**Figure 5.4: Routes Selection Based On Distance** 

One-way streets also need to be taken into consideration. The one-way rule depends on how the sequence of the vertices are arranged in the digitising process, and secondly, the value of the One-Way field. To calculate the shortest routes for the delivery of fresh vegetables among the other roads, the software makes decisions based on the variables collected in the attribute table which shows that it does indeed affect the drive time.

| P Directions (Route)                               |                |       |
|--|----------------|-------|
| 223: At fork keep right on 28 mrr2 (e)             | 1.8 km 1 min   | Мар 🔺 |
| 224: At fork keep right on 28 mrr2 (e)             | 1.4 km 1 min   | Map   |
| 225: At fork keep left on Sg besi, j (besraya)     | 0.8 km < 1 min | Map   |
| 226: Continue on E9 besraya (serdang)              | 1.6 km 1 min   | Map   |
| 227: Finish at Graphic Pick 22, on the left        |                | Map   |
| Total time: 2 hr 8 min<br>Total distance: 107.5 km |                |       |
| Options Print Preview Save As                      | Print          | Close |

Figure 5.5: Total Time and Distance for Routes Based Distance and One-Way Road

As shown in Figure 5.5, the total time for the delivery of fresh vegetables, taking into account one-way streets, is 2 hour and 8 minutes. The total distance, in this case, is about 107.5 Km. Without taking into account the one-way road variable, the total distance and time will subsequently change. The result of those changes is demonstrated in the following figure 5.6. It means that the total time, as well as total distance, will both be reduced. So the total time for delivery in selected roads is 1 hour and 49 minutes while the total distance is 94.9 Km.

|   | 2 Directio           | ns (Route)   |         |                |            | x |
|---|----------------------|--|---------|----------------|------------|---|
|   | [-<br>] <u>Rou</u> t | te: Graphic Pick 1 - Graphic Pick 22                         | 94.9 km | 1 hr 49<br>min |            | Ê |
| 1 | <u>1</u> :           | Start at Graphic Pick 1                                      |         |                | Map        |   |
| l | <u>2</u> :           | Go north on E9 besraya (sg besi)<br>toward Sg besi, j (mrr2) | 2.4 km  | 2 min          | <u>Map</u> |   |
| l | <u>3</u> :           | Make sharp left on Nse to besraya ramp                       | 0.4 km  | $< 1 \min$     | Map        |   |
| ł | <u>4</u> :           | Make sharp left on Nse - kesas i/c                           | 0.2 km  | < 1 min        | <u>Map</u> | - |
|   | Options              | Print Preview Save As  | Print   |                | Close      |   |

Figure 5.6: Total Time and Distance with Consideration of One-Way Roads

#### 5.5.2 Phase two: Impedance modelling of optimal routes in terms of drive time

Considering the shortest time as the goal of route planning for the delivery purpose, the impedance functions will be integrated considering the length of the road and average speed of a transportation vehicle that affect the driving efficiency (Dong, 2012). The length and speed are the most important indicators of travel time in the urban transportation system. The shortest link length, the faster sections of speed, causes the shorter travel time (Dong, 2012).

The vehicle routing process will find the quickest routes. For the routing process, the street data must have a field that saves a value, which relates how much time is needed to pass through each street. In this case, the impedance is time, which frames the quickest route the best route. Hence, the best routes can be defined as the route that has the lowest impedance or lower cost, where the impedance is selected by the user.

Traffic condition is not fixed during daytime especially in central Kuala Lumpur, and it affects the speed of a vehicle. Hence, the travel time of the links changes continuously in a transportation network. In this case when a decision maker wants to select a path with minimum travel time based on the cost of travel then the software should find the optimal routes go through the roads with maximum travel speed. For this reason, the total drive time will constantly decrease for selected roads.

Here, the drive time takes into account the average speed for a vehicle divided by the length of the road. The speed of a vehicle is the average speed during the period of a 24-hour day. Drive time in this study is considered as the dependent variable to derive the most significant variables in transportation and delivery services of fresh vegetables. It plays as the dependent variable, and the analysis is to find the variables that have significant changes when talking about delivery of fresh produce. Figure 5.7 describes the best routes to distribute to requested hypermarkets around the city.



Figure 5.7: Best Routs Based on Drive Time

| Directions (Route)                                  |        |         |       | x        |
|---|--------|---------|-------|----------|
| 241: At fork keep right on 28 mrr2 (e)              | 1.8 km | 1 min   | Map   |          |
| 242: At fork keep right on 28 mrr2 (e)              | 1.4 km | 1 min   | Map   |          |
| 243: At fork keep left on Sg besi, j (besraya)      | 0.8 km | < 1 min | Map   |          |
| 244: Continue on E9 besraya (serdang)               | 1.6 km | 1 min   | Map   |          |
| 245: Finish at Graphic Pick 22, on the left         |        |         | Map   |          |
| Total time: 1 hr 53 min<br>Total distance: 120.3 km |        |         |       | (E)<br>• |
| Options Print Preview Save As                       | Print  |         | Close |          |

Figure 5-8: Total Time and Distance by Consideration of Average Speed for Drive Time

Without taking into account one-way roads in the ArcGIS analyst tool, the total distance travelled by the lorry and the total time will inevitably reduce. As the result shows that the total time is 1 hour and 37 minutes and the total distance becomes 106.6 Km.

| \$      | Directio   | ns (Route)   |             |                |            | x |
|---------|------------|--|-------------|----------------|------------|---|
| [-<br>] | Rout       | e: Graphic Pick 1 - Graphic Pick 22                          | 106.6<br>km | 1 hr 37<br>min |            | Ê |
|         | <u>1</u> : | Start at Graphic Pick 1                                      |             |                | Map        |   |
|         | <u>2</u> : | Go north on E9 besraya (sg besi)<br>toward Sg besi, j (mrr2) | 2.4 km      | 2 min          | <u>Map</u> |   |
|         | <u>3</u> : | Make sharp left on Nse to besraya<br>ramp                    | 0.4 km      | < 1 min        | <u>Map</u> |   |
|         | 4.         | Mala   | 0.01        |                |            | Ŧ |
|         | Options    | Print Preview Save As  | Print       |                | Close      |   |

Figure 5.9: Total Time and Distance Based on Average Speed Without Consideration of One-Way Road

## 5.5.3 Phase three: Impedance modelling of the primary time pattern—7 to 10

For the purpose of this project, only two different periods of time were considered for the impedance modelling of the delivery process. The Network Analyst tool allows for the analysis of the network of roads, which store information of travelling speed at different hours. The travelling speed patterns for many roads in each direction vary, such as slower in the morning, and average at noon. The purpose of this study is to calculate the shortest routes between two locations by taking into account different traffic conditions in the morning. It implemented the computation methodology within a realtime environment.

References to the previous research using regression model it has clarified that they have derived different parameters for transportation planning but the time pattern is not considered as part of their modelling. For example, Sofia (2013) analysed the problem of optimal routes associated with transportation. Variables taken into consideration were impedance for intersections, type of road and speed. Minimising the distance and drive time at the same time is the main goal of this research.

Figure 5.10 describes the delivery routes and their directions in the 7 to 10am traffic time pattern, and the average speed of a lorry along the routes is collected and saved into the attribute table.

For calculating the drive time based on average speed in the 7 to 10am a time frame of a vehicle then it has to consider the distance divided by average speed for each vehicle. The travel times towards all of the destinations were determined by collecting data in accordance with different time patterns. The routes selection is based on the average speed of the vehicle in this particular time frame. Due to the fact that the speed of cars is usually slow in the morning. The ArcGIS software will select routes that have least drive time on that particular route. This inevitably increases the total time obtained from a network analysis. The variables in the determination process are the location of the hypermarkets and delivery times during a single day.

In accordance with the nature of perishable produce, such as fresh vegetables, time is an important factor in this research. This makes the fastest routes the best possible choice in fulfilling the objective of this thesis. It is assumed from the statistical approach that times between 7 to 10 am significantly influences time, which prompts us to come up with a GIS model to determine the best routes for delivery purpose.



Figure 5.10: Fast Routes for Delivery in the 7 to 10Period

The total time and the total driving time in every possible, delivery direction at 0700-1000 are detailed in the following timetable. The final destinations or the hypermarkets are selected by the delivery manager based on their respective demands and the corresponding time. It was duly determined that the total time a vehicle would have to spend on a road is 3 hours and 2 minutes, at a total distance of 117.5 kilometres.

| Directions (Route)                                 |        |         |       | x        |
|--|--------|---------|-------|----------|
| 252: At fork keep right on 28 mrr2 (e)             | 1.8 km | 2 min   | Map   | *        |
| 253: At fork keep right on 28 mrr2 (e)             | 1.4 km | 2 min   | Map   |          |
| 254: At fork keep left on Sg besi, j (besraya)     | 0.8 km | < 1 min | Map   |          |
| 255: Continue on E9 besraya (serdang)              | 1.6 km | 2 min   | Map   |          |
| 256: Finish at Graphic Pick 22, on the left        |        |         | Map   |          |
| Total time: 3 hr 2 min<br>Total distance: 117.5 km |        |         |       | (E)<br>• |
| Options Print Preview Save As                      | Print  |         | Close |          |

Figure 5.11: Total Distance and Time for Delivery in the 7 to 10 Period

Without taking into account the one-way road, the total distance and time will inevitably change. The following information in the figure concerns both time and distance.

| Directions (Re                                      | oute)                |                 |        |         |       | x          |
|---|----------------------|-----------------|--------|---------|-------|------------|
| 239: Turr   | n right on Mrr2 to e | ewlink          | 0.2 km | < 1 min | Map   | *          |
| 240: Mak  | e sharp left on 28   | mrr2 (e)        | 3.3 km | 4 min   | Map   |            |
| 241: Con  | tinue on Sg besi, j  | (besraya)       | 0.8 km | < 1 min | Map   |            |
| 242: Con  | tinue on E9 besra    | ya (serdang)    | 1.6 km | 2 min   | Map   |            |
| 243: Finis  | sh at Graphic Pick 3 | 22, on the left |        |         | Map   |            |
| Total time: 2 hr 31 min<br>Total distance: 103.2 km |                      |                 |        |         |       | (III)<br>• |
| Options   | Print Preview        | Save As         | Print  |         | Close |            |

Figure 5.12: Total Distance and Time for Delivery Without Consideration of One-Way

# 5.5.4 Phase four: The impedance modelling of the secondary time pattern—10 to 12 am

If the speed of a vehicle along a route at1000-1200 differs, then the calculation for selecting the best routes for the delivery of fresh vegetables will differ as well. According to the statistical data, the average speed of a vehicle in this period will increase compared to the previous time patterns. In this time frame, it would seem to be a lot easier for the software to calculate the least travel time on the roads. This part of the analysis will result

in the decrease of drive time on the roads for the delivery of fresh vegetables. Figure 5.13 shows the best routes for delivery to selected hypermarkets located in Kuala Lumpur.



Figure 5.13: Best Routes for Delivery in the 10 to 12 Period

Figure 5.14 details the total distance and drive time that a vehicle must travel along a road for delivery fresh produces to all hypermarket in that particular period.

| 🔊 Directions (Route)                               |        |                  | x        |
|--|--------|------------------|----------|
| 258: At fork keep right on 28 mrr2 (e)             | 1.8 km | 2 min Map        | *        |
| 259: At fork keep right on 28 mrr2 (e)             | 1.4 km | 1 min Map        |          |
| 260: At fork keep left on Sg besi, j (besraya)     | 0.8 km | 1 min <u>Map</u> |          |
| 261: Continue on E9 besraya (serdang)              | 1.6 km | 2 min Map        |          |
| 262: Finish at Graphic Pick 22, on the left        |        | Map              |          |
| Total time: 2 hr 6 min<br>Total distance: 116.8 km |        |                  | (E)<br>* |
| Options Print Preview Save As                      | Print  | Close            |          |

Figure 5.14: Total Time and Distance for Delivery in Time 10 to 12

Without taking into account one-way routes, the total distance, as well as the total drive time will frequently change. Figure 5.15 shows the results obtained from ArcGIS Network Analyst.

| Directions (Route)                                  | l          | _ 0            | x        |
|---|------------|----------------|----------|
| 249: Turn right on Mrr2 to ewlink                   | 0.2 km < 1 | min <u>Map</u> | *        |
| 250: Make sharp left on 28 mrr2 (e)                 | 3.3 km 3   | min <u>Map</u> |          |
| 251: Continue on Sg besi, j (besraya)               | 0.8 km 1   | min <u>Map</u> |          |
| 252: Continue on E9 besraya (serdang)               | 1.6 km 2   | min <u>Map</u> |          |
| 253: Finish at Graphic Pick 22, on the left         |            | Map            |          |
| Total time: 1 hr 47 min<br>Total distance: 103.2 km |            |                | (E)<br>• |
| Options Print Preview Save As                       | Print      | Close          |          |

Figure 5.15: Total Time and Distance Without Consideration of One-Way Roads

It was determined that time between 10.00 - 12.00 significantly influence the total travel time. Therefore, it necessary for us to model the fast routes for delivery purposes based on the GIS aspect. The system is capable of adjusting the location of the hypermarkets as per the destinations at any possible time. Using the travel time information can help spread out the traffic congestion, because travellers might select another path, or opt for the use of public transportation.

## 5.5.5 Phase five: the population as an impedance for calculating driving time

Transport problems always occur with the growth of population Chen et al. (2008). The map outlines the average population for the different regions in KL. There is a direct relationship between the population and the total volume of cars. When the population is high, then the number of the cars that occupies the roads increases, and the average speed of the car will decrease, which makes increasing drive time in those routes. With the rapid urban development, the population will increase and so will the number of vehicles putting more pressure on traffic congestion.

Figure 5.16 shows the total density of population in Kuala Lumpur, divided into several polygons, with each polygon distinguished by the total number of population. Different colours will distinguish the density of population in the selected area by the following map. The polygons are representing the total numbers of people in each specific place. The colours describe the distribution of population individual person unit in KL. The lighter colour has a lower density of population. The dark colour represents the high density of population in that particular region. In the following figure, the dark brown colours indicate the higher populated density compare to lighter brown colour. The consideration of network analysis is to find the most efficient routes based on database management system which was developed by significant parameters as well as population density.



Figure 5.16: Population Density on the Map

The statistics show that the average speeds depend on the type of routes and the population density in their vicinity. The data regarding average speed is calculated from the speed of a medium vehicle during the day. The related data has been saved in the attribute table in ArcGIS as a dataset. This speed data is then used to divide the street length to determine the drive time value.

## 5.5.5.1 Weighting model for population density

For constructing a map that integrates roads and population; the following map shows the roads joined with a population polygon. This case merges the population for each region with selected roads. This step classifies the population layer into five distinct classes. The weighting method is applied for population density, and later calculated the best routes, based on the weight for each categorised number. Table 5.2 describes the classes of the population, the weighting for each class of population, taking into account the population density for all of the roads. The new field as a class is added to the attribute of the roads which is joining with population data.

In this study, population parameter is classified based on the density of people living in one region. The weighted average will be used for combination classified composite population. A weighted average method is average of between two quantity variable. These weightings determine the relative importance of each quantity on the average. Weightings are the equivalent of having many like items with the same value involved in the average.

| Population   | Description | Range     | Weighting |
|--------------|-------------|-----------|-----------|
| 0 - 994      | Very Low    | >550      | 1         |
| 995 - 2023   | Low         | 550-1650  | 2         |
| 2024 - 4060  | Moderate    | 1650-3950 | 3         |
| 4061 - 7181  | High        | 3950-9100 | 4         |
| 7182 - 15789 | Very high   | <9100     | 5         |

 Table 5.2: Classification of Population

In this research, the population was selected as a significant parameter in determining the fast routes. The density of population is various in different regions. According to the assessment of required data, in any region that the population is minimum then the drive time will decrease for those particular roads. For a description about population, it has categorised the range of population to very low, low, moderate, high and very high. It means that in a region with a minimum population then the description for those regions is very low and for areas with a maximum population then the description is very high.

Figure 5.17 describes the combination of road networks in KL and population density in each road area. The lighter colour describes the roads with the low density while the darker colour describes the roads with high population density.



**Figure 5.17: Classification of Population** 

It has to enter all the new information about the weighting number to the attribute table in ArcGIS. Then the new network dataset is created for the new attribute table, and the impedance is a class field for population density. The result for the best routes based on the class field according to attribute table is inFigure 5.18, which is run using ArcGIS software.



Figure 5-18: Result of Fast Routes for Delivery Based on Population Impedance

The total distance and total time that a vehicle must travel along the distribution roads based on population factor are shown in Figure 5.19. For this case, the total travel time is 2 hours and 31 minutes while the total distance is 131.6 Km.



Figure 5.19: Total Distance and Time for Delivery Based on Population Factor

Comparing the fast routes based on impedance as a population and the drive time shows that there are significant differences when the software calculates the best routes based on drive time as an impedance. The drive time along the populated area will increase significantly. Table 5.3 shows the differences in total distance and drive time for both population and drive time parameter. The differences for total time is about 24 minutes.

 Table 5.3: Comparison between Population and Drive Time

| Impedance  | Total time      | Total distance | Time differences |
|------------|-----------------|----------------|------------------|
| Population | 1 hr 51 Minutes | 90.9 Kilometre |                  |
| Drive time | 1 hr 27 Minutes | 89 Kilometre   | 24 Minutes       |

## 5.5.6 Phase six: Car volume as the impedance for selecting the best routes

The major urban traffic congestion is more concentrated on the central business district. Therefore, traffic congestion in many urban areas is spatially dispersed and in different location depends on the situation where time is different for each road network. Instead, congestion occurs is specific locations and propagates through the network over time. Congestion has a straight relation with car volume, and it causes the increasing the drive time along the roads. For this reason, it has calculated the total car volume for each road and then gathered all the information about the total cars in one attribute table. The

ArcGIS software calculates the best routes based on car volume, which is collected through road networks and saved in a specific attribute table.

Due to the rapid increase in the number of vehicles, traffic congestion has become an increasingly serious problem in large cities around the world. The travel time fluctuates during the day, due to being affected by different factors of traffic, such as vehicle volumes and routing path selection, which are the most important parameters for this study.

Figure 5.20 is the classification of average speed for all the roads in Kuala Lumpur. The places near to KLCC area are congested and because of the high density of cars along the roads so the average drive time for those roads is very slowly. It has classified all the roads to four classes of speed. The average maximum speed for a track or lorry for this study should not be more than 80 Km/h. This speed limit is the highest speed allowed by low for road vehicles in Malaysia. National speed limits are a set of speed limits applicable on Malaysian express ways, federal roads and municipal roads (Syafiq, 2007). Like any other country, failing to obey the speed limits on Malaysian roads and expressways is an offence as subject to Malaysian Road Safety Act 1987.



Figure 5.20: Map of Average Speed for All the Roads in KL

At this stage, the calculation of fastest routes in ArcGIS is based on the car volume for all the road networks in KL city. An effective approach to modelling the transportation based on car volume is Network Analyst tool. The Network Analyst model approach captures the relationship between users travel decisions and network performance. Adjustment occurs between user decision and network performance until a balanced pattern in achieved. Figure 5.21shows the fastest routes based on the total car volume on each route.



Figure 5.21: Best Routes Based on Car Volume

After data evaluation and assessment in statistical approach find that the car volume has significant value in drive time. The total distance and total time that a vehicle must travel along the distribution roads are shown in Figure 5.22.

|   | ¢ D | irectior     | ns (Route)  |        |         |            | x        |
|---|-----|--------------|---|--------|---------|------------|----------|
|   |     |              | ппк (е)   |        |         |            |          |
| 4 |     | <u>205</u> : | Continue on Exit 909 east-west link                                   | 0.3 km | < 1 min | <u>Map</u> |          |
| 1 |     | 206:         | Make sharp left on E9 besraya (serdang)                               | 4.8 km | 5 min   | Map        |          |
|   |     | <u>207</u> : | At fork keep right on E9 besraya<br>(serdang)                         | 0.7 km | < 1 min | <u>Map</u> |          |
|   |     | <u>208</u> : | Turn right at Sg besi, j (besraya) to stay<br>on E9 besraya (serdang) | 1.6 km | 2 min   | <u>Map</u> |          |
| I |     | <u>209</u> : | Finish at Graphic Pick 22, on the left                                |        |         | Map        |          |
|   |     |              | Total time: 3 hr 17 min<br>Total distance: 138.5 km                   |        |         |            | (E)<br>• |
|   |     | Options.     | Print Preview Save As   | Print  |         | Close      |          |

Figure 5.22: Total Time and Distance for Delivery Based on Car Volume

#### 5.5.7 Phase seven: The land use modelling to determine the optimal routes

Land use is another parameter that will affect driving time for a vehicle along the road. It has been discovered that the land use characteristics of transportation and delivery services affect the drive time for a vehicle travelling on a road.

The results obtained from the land use classification are a land use map over the area of study. This land use map was classified into three classes, which included a residential area, Industrial area and open spaces. After the evaluation and analysis of land use data in the statistical process, and running the regression model and identifying the output, the residential area seems to be the most significant factor affecting drive time on related roads. Therefore, the main consideration of this research focuses on the residential areas.

The different land use classes were used to differentiate the various terrains over the area of study, and these were used in the final judgement on the best alternative to determine the best route path. The usefulness of this result is that the land use map helps to understand what the different available land use over the area of study is being used for, as well as the available features that are present in the area.

Figure 5.23 describes the land use classification in Kuala Lumpur. The land use is classified into many other features. The examples of land use are such as residential area, commercial area and agricultural land. The following map shows the two different classes of land use. According to the variables appointed for this study, for the purpose of statistical analysis, the green describes the open space. Industrial areas are in yellow.





The following map describes the details of land use including the open space and industrial areas. Figure 5.24 zooms on the selected areas in KL. The yellow identifies the residential areas while the green stands for open space.



Figure 5.24: Land Use Map

The detailed information regarding a residential area shown in Figure 5.25. The key constraint is that the proposed path for delivery purposes must be sequestered from the residential area. In this case, the purpose is to find the most suitable road to pass through a non-residential area. This step applied the buffering function in selecting the best roads in order to avoid a congested area. The following map describes the residential area as the significant parameter for this study.



Figure 5.25: Location of Residential Areas

Figure 5.26 shows the map combining road networks with residential areas. This map shows the contribution of the residential area as a significant parameter in this study. It has been illustrated to overlay the two layers on one map. The layers include the road network and residential area as well.



Figure 5.26: Map of Road Combination and Residential Zones

# 5.5.7.1 Buffering function in ArcGIS

The buffer function is a feature of the GIS software that allows one to create a radial area of the desired distance around specific data points, or relate an interesting area to a more easily identified relationship among the data with regards to distance. The buffer function is used when someone wants to create a buffer to select features within a specific distance of a feature. The buffer tool creates a new coverage of buffer polygons around the input coverage features. Input features can be polygons, lines, points or nodes. The width of the buffer can be specified as a fixed distance from an attribute table, or from a distance table.

Initially, a base 100 metres from the main roads in KL and finally the roads are selected as the base within this area. The specified distance (100 m) is used for buffering in order to determine the specific roads around it. Figure 5.27 describes the results after the buffering function, and yellow are the zones falling within the residential area.



Figure 5.27: Buffering Zone Map

The point is to calculate the roads that do not pass through the residential area. Nearing the residential areas, the average speed will decrease, which inevitably increases the drive time.

After applying the buffering process, the remains residential areas are showing in Figure 5.28. It illustrates how the residential area zones that are buffering 100 metres from the main roads are selected. By this calculation, we can release the roads that are in the vicinity of the residential area.



Figure 5.28: Map of Residential Area Selection

# 5.5.7.2 Weighting model for land use classification

Reference to the research conducted by Sharifi et al. (2009) to evaluate different landfill sites index system was developed representing cumulative effects of relative importance weights and suitability values from field observation. At this step, it has to give weights to each feature in land use layer. For constructing a map that integrates roads with land use, Figure 5.29 shows the roads joined with a land use polygon. This case merges the land use for each region with selected road networks in Kuala Lumpur.



Figure 5.29: Combination of Roads with Land Use

This step classifies the land use layer into four distinct classes. The weighting method is appliedforland use layer, and later calculates the best routes, based on the weightage for each categorised number. Table 5.4 describes the classes of land use layer, the weighting to each class of land use, taking into account the land use density for all of the roads. The new field as a land uses added to the attribute of the roads which is joined with land use data.

| Land use    | Description | Classes |  |
|-------------|-------------|---------|--|
| Residential | High        | 4       |  |
| Industrial  | Moderate    | 3       |  |
| Open Space  | Low         | 2       |  |
| Others      | Very Low    | 1       |  |

Table 5.4: Land Use Classification and Description About Classes in AttributeTable

In this study, the land use parameter is classified based on land use features in one region. The weighted average will be used for combination classified composite land use. From the statistical result, it is assumed that residential area is the most significant parameter in land use. It means that in a residential area the drive time is much lower compare with other land use features. According to the assessment of required data, in any region that residential area categorises the land use then the drive time will decrease for those particular roads. These weightings determine the relative importance of each quantity on the average. Weightings are the equivalent of having many like items with the same value involved in the average.

It has to enter all the new information about the weighting number to the attribute table in ArcGIS. Then the new network dataset is created upon the new attribute table, and the impedance is based on the class field for land use layer. The result for the best routes obtained the class field according to the attribute table as following the map, which is running in ArcGIS software. When the software is running for new entry information for calculating the best routes, then the calculation is based on the best routes that show the minimum drive time for selected routes. The new field as a class is added to the attribute table. The following analysis is the result obtained from the attribute table in ArcGIS. The result for the best routes based on the weight class field according to the attribute table in Figure 5.30.



Figure 5.30: Result Obtained for Selecting Routes Based on Land Use Impedance

After data evaluation, land use layer significantly affects drive time. The description of total distance and the total time that a vehicle must travel along the distribution roads is shown in Figure 5.31.

| Directions (Route)                                 |                                |
|--|--------------------------------|
| 165: Continue northwest                            | 2 km 3 min Map 🔺               |
| <u>166</u> : Make sharp left at Residential        | 1.4 km 5 min Map               |
| <u>167</u> : Bear left at Residential              | < 0.1 < 1 min <u>Map</u><br>km |
| 168: Turn right at Residential                     | 0.5 km 1 min Map               |
| 169: Finish at Graphic Pick 23, on the right       | Map                            |
| Total time: 2 hr 17 min<br>Total distance: 65.9 km | (III)<br>•                     |
| Options Print Preview Save As                      | Print Close                    |

Figure 5.31: Result of Total Distance and Time for Routes in Delivery Services

Fuel consumption is one of the main concerns for most travellers, especially in the context of determining the fastest route with the minimum consumption of fuel. Many companies take this factor into account when providing vehicles, in order to adhere to the shortest travel time and distance. Traveller information and guidance reduces travel time, and may help delivery services avoid congestion and improve traffic efficiency. There is a strong need to make our route travel as efficient as possible in order to reduce the overall fuel consumption. Hence, there is a direct correlation between the travel time and fuel consumption.

# 5.5.7.3 Comparison result between distance and drive time

In the following Table 5.5, the comparison is among the results obtained from ArcGIS that determined the best routes based on distance and travelling time. The result is with consideration of determining the same routes for conducted analysis.

| Impedance  | Total time     | Total distance | Routes      |
|------------|----------------|----------------|-------------|
| Distance   | 1 h 39 Minutes | 80.1 Kilometre | Fast routes |
| Drive time | 1 h 21 Minutes | 86.4 Kilometre | Fast routes |

**Table 5.5: Comparison Result of Distance and Drive Time** 

According to the information in Table 5.5, the total difference for drive time for the related routes is 18 minutes, and the total difference in distance is about six kilometres.

## 5.5.7.4 Comparison results between time patters and drive time

The comparison is between two time frames in the morning for the aim of distribution. Time patterns are divided into the 7-10 time frame in the morning, and the other from the 10 to 12 am period. According to the schedule for delivery services for fresh vegetables to hypermarkets it has to consider that fresh produce must reach the markets before 12 noon. This time, regulation is based on the request from hypermarkets in KL. The company has to be aware of sending the vehicles during those periods of time. The critical reason to select those periods of time is that after 12 noon the fresh vegetables are not accepted by related hypermarkets. Comparison between the total times, taking into account the 0700-1000 time frame into the model and drive time for the delivery of fresh vegetables are shown in Table 5.6. The results obtained from statistical analyses showed that the drive time would increase when taking into account the time pattern of 07.00-10.00, compared to when the drive time is based on the average speed of a vehicle during the day.

| Impedance            | Total time            | Total distance | Routes      |
|----------------------|-----------------------|----------------|-------------|
| Time between 7 to 10 | 1 H and 59<br>Minutes | 85 Kilometre   | Fast routes |
| Drive time           | 1 H and 21<br>Minutes | 86.4 Km        | Fast routes |

Table 5.6: Comparison Results of Time Pattern of7 to 10 Driving Time

If the time pattern of 1000-1200 is regarded as impedance during modelling, the following table shows the effect of this variable on total drive time and distance. Table 5.7 illustrates the total time and distance between 1000-1200 as an independent variable and drive time as a dependent variable. Table 5.7 illustrates the total time at1000-1200 increasing when compared with a drive time of a vehicle with average speed.

Table 5.7: Comparison results of time pattern of 10 to 12 and driving time

| Impedance        | Total time             | Total distance | Routes      |
|------------------|------------------------|----------------|-------------|
| Time 10 to 12 am | 1 Hr and 28<br>Minutes | 86.4 Kilometre | Fast routes |
| Drive time       | 1 Hr and 21<br>Minutes | 86.4 Km        | Fast routes |

The results showed that the total drive time at 0700-1000 is greater than 1000-1200. The manager will decide on the most suitable time for delivery services based on time impedance and the total drive time for a vehicle. Table 5.8shows the comparison results for time pattern of time 7-10 and 10-12. The result showed that the total drive time for a vehicle at 0700-1000 is greater 1000-1200 for vehicles to travel. This isdue to the fact that the average speed of a truck is less during 0700-10.00 along the roads, compared to 10.00-1200. Consideration of fresh vegetables as a highly perishable product the time is

a critical factor to take into consideration of selecting the routes that haveleast amount of time. In this regard for the period of 10 to 12 the saving time is 31 minutes compared to 7 to 10. This amount of time is considerable for fresh produce such as fresh vegetables in this study.

| Time Pattern<br>Variables | 7 - 10          | 10 - 12         | Saving time comparison |
|---------------------------|-----------------|-----------------|------------------------|
| Distance                  | 85 Kilometres   | 86.5 Kilometres |                        |
| Drive time                | 1 hr 59 minutes | 1 hr 28 minutes | 31 Minutes             |

 Table 5.8: Compression Results of Time Pattern of 7 to 10 and 10 to 12

In Table 5.8, for the period of time in different speed patterns, it has been realised that integrating the obtained negative coefficient into the statistical assessment led us to conclude that the increased speed of automobile reduces the transmission of the products. Based on the amounts of coefficients, 1% increase of speed between 7 to 10 am will reduce the transportation time by up to 6%. Meanwhile, speed increases by 1% between 10 to 12 am, resulting 53% of the reduction in the transmission time. According to the results obtained from the ArcGIS, between 7-10 am, the total drive time is 1 hr and 59 minutes; but at 10-12 noon, the total drive time is 1 hour and 28 minutes for route selection.

When the software is running for new entry information for calculating the best routes, then the calculation is based on the best routes that show the shortest distance, and at the same time, the drive time will increase as well. In this situation, the ArcGIS will calculate the best routes, and the total distance will increase compared to the same roads.

Figure 5.32 describes the differences between the both time frames based on the directions. The blue is the directions for travelling in time 10 to 12 while the red colour is the directions of travelling along the roads in time 7 to 10. The Network Analyst has calculated the fast routes based on the average speed of the particular road. The map specifies the differences among the roads with different colours.



Figure 5.32: The Differences for Road Direction in Two Time Frames

# 5.6 Statistical data evaluation and analysis

To investigate the previously mentioned hypotheses, it is necessary that all variables within the framework of a multivariable regression be investigated. One of the reasons for adopting the regression model in this study is that it provides decision makers with a powerful set of tools, which allows prediction regarding present or future events. The outcome of this research is the accumulation of spatial and non-spatial information in database system. The consideration of this thesis is to develop a spatial database management system to solve the distribution problem of fresh vegetables. The intended regression model will be as follows.
$$LTIME = C(1) * LSPEED7TILL10 + C(2) * LSPEED10TILL12 + C(3)$$
  
\* CARVOLUME + C(4) \* LLENGTH + C(5) \* LPOP + C(6)  
\* TWO - WAY + C(7) \* SCHOOL + C(8) \* RESIDENTIAL  
+ C(9) \* INDUSTRIAL + C(10) \* OPENSPACE + C(11)

This regression model includes width from the origin (fixed coefficient) and coefficients of ten independent variables. C(i)s indicate the coefficients of independent variables. LTIME is the log of drive time as a dependent variable; and LLENGTH and LPOP indicate the log of distance for each road network and a log of the population in a specific area, respectively. LSPEED7TILL10 and LSPEED10TILL12 indicate the Log of two timeframes include time 7-10 and 10-12 in this study. CARVOLUME indicates the average number of cars in each road network. The proposed hypotheses have been designed on this basis that each coefficient in the regression model is significant.

 $H_0$  and H1 assumptions for any of the above hypotheses are as follows:

# {H0 : Coefficient Model is not significant} {H1: Coefficient Model is Significant}

Whenever in a regression model, a coefficient is not significant (H0 hypothesis is accepted); it means that variable has no effect on the dependent variable. If P-value is less than a = 0.05; the coefficient will be significant (H0 hypothesis is not confirmed) and if P-value is more than a = 0.05; the coefficient will not be significant (H0 hypothesis will be accepted). The results obtained from this regression analysis are shown in table 5.9:

| Dependent Variable: LTIME |                           |          |             |        |  |  |
|---------------------------|---------------------------|----------|-------------|--------|--|--|
| Method: Least Squares     |                           |          |             |        |  |  |
| Variable                  | Coefficient               |          | t-Statistic | Prob.  |  |  |
| LSPEED7TILL10             | -0.0658                   |          | -2.5856     | 0.0099 |  |  |
| LSPEED10TILL12            | -0.5404                   |          | -7.1962     | 0.0000 |  |  |
| CARVOLUME                 | 0.2664                    |          | 2.9735      | 0.0030 |  |  |
| LLENGTH                   | 0.6990                    |          | 8.1712      | 0.0000 |  |  |
| LPOP                      | 0.0202                    |          | 2.1225      | 0.0341 |  |  |
| TWO-WAY                   | 0.0597                    |          | 3.3905      | 0.0007 |  |  |
| SCHOOL                    | 0.1681                    |          | 5.9763      | 0.0000 |  |  |
| RESIDENTIAL               | 0.0286                    |          | 0.5089      | 0.6109 |  |  |
| INDUSTRIAL                | -0.0169                   |          | -0.3005     | 0.7639 |  |  |
| OPENSPACE                 | -0.0008                   |          | -0.0145     | 0.9884 |  |  |
| C                         | -0.5421                   |          | -0.4652     | 0.6419 |  |  |
| Regression Statistics     |                           |          |             |        |  |  |
| F-statistic               | 5635.386                  | Adjusted | R-squared   | 0.9866 |  |  |
| Prob(F-statistic)         | 0.0000 Durbin-Watson stat |          | 1.3571      |        |  |  |

**Table 5.9: First Round of Regression Results** 

#### 5.6.2 Substituted coefficients

By inputting the obtained coefficients in the above regression, we will have the final result from data analysis

LTIME = -0.0658 \* LSPEED7TILL10 - 0.5404 \* LSPEED10TILL12 + 0.2664 \* CARVOLUME + 0.6990 \* LLENGTH + 0.0202 \* LPOP + 0.0597 \* TWO - WAY + 0.1681 \* SCHOOL + 0.0286 \* RESIDENTIAL - 0.0169 \* INDUSTRIAL - 0.0008 \* OPENSPACE - 0.5421

However, by considering the amounts of statistic t and its probability amount, it is observed that the variables associated withresidential area, industrial and open space are not significant and effective at this statistical level. On this basis, it is necessary that the above regression is met once again by subtracting one variable that its probability is higher compared to residential and industrial from the total independent variables. For this purpose, it is necessary that the insignificant independent variables are subtracted one by one. In this way, regression is met once again without considering the virtual variable of open space, and the following results are obtained:

| Dependent Variable: LTIME |             |                           |   |       |  |  |  |
|---------------------------|-------------|---------------------------|---|-------|--|--|--|
| Method: Least Squares     |             |                           |   |       |  |  |  |
| Variable                  | Coefficient | t-Statistic               | ] | Prob. |  |  |  |
| LSPEED7TILL10             | -0.0658     | -2.5874                   | 0 | .0099 |  |  |  |
| LSPEED10TILL12            | -0.5404     | -7.2022                   | 0 | .0000 |  |  |  |
| ICARVOLUME                | 0.2664      | 2.9755                    | 0 | .0030 |  |  |  |
| LLENGTH                   | 0.6990      | 8.1767                    | 0 | .0000 |  |  |  |
| LPOP                      | 0.0202      | 2.1256                    | 0 | .0339 |  |  |  |
| TWO-WAY                   | 0.0597      | 3.4019                    | 0 | .0007 |  |  |  |
| SCHOOL                    | 0.1681      | 5.9841                    | 0 | .0000 |  |  |  |
| RESIDENTIAL               | 0.0294      | 2.2717                    | 0 | .0234 |  |  |  |
| INDUSTRIAL                | -0.0161     | -0.8020                   | 0 | .4228 |  |  |  |
| С                         | -0.5429     | -0.4668                   | 0 | .6407 |  |  |  |
| Regression Statistics     |             |                           |   |       |  |  |  |
| F-statistic               | 6269.88     | Adjusted R-squared 0.9867 |   |       |  |  |  |
| Prob(F-statistic)         | 0.0000      | Durbin-Watson stat 1.357  |   |       |  |  |  |

 Table 5.10: 10:Second Round of Statistical Regression Results

The coefficients input in the regression model and then it has obtained following result:

*LTIME* = -0.0658 \* *LSPEED7TILL10* 0.5404 \* *LSPEED10TILL12* + 0.2664

\* *CARVOLUME* + 0.6990 \* *LLENGTH* + 0.0202 \* *LPOP* + 0.0597

\* TWO – WAY + 0.1681 \* SCHOOL + 0.0294 \* RESIDENTIAL

- 0.0161 \* *INDUSTRIAL* - 0.5429

However, the interesting point is that this time, the virtual variable of the residential zone has appeared in a 5% error level. On this basis, this time, the virtual variable of industrial zones will be removed from the model. Estimation of regression for the third round was obtained:

| Dependent Variable: LTIME |         |      |               |                |  |  |  |
|---------------------------|---------|------|---------------|----------------|--|--|--|
| Method: Least Squares     |         |      |               |                |  |  |  |
| Variable                  | Coeffic | ient | t-Statistic   | Prob.          |  |  |  |
| LSPEED7TILL10             | -0.0664 |      | -2.6014       | 0.0095         |  |  |  |
| LSPEED10TILL12            | -0.5387 |      | -7.1919       | 0.0000         |  |  |  |
| ICARVOLUME                | 0.2663  |      | 2.9740        | 0.0030         |  |  |  |
| LLENGTH                   | 0.6984  |      | 8.1765        | 0.0000         |  |  |  |
| LPOP                      | 0.0203  |      | 2.1327        | 0.0333         |  |  |  |
| TWO-WAY                   | 0.0605  |      | 3.4513        | 0.0006         |  |  |  |
| SCHOOL                    | 0.1681  |      | 5.9863        | 0.0000         |  |  |  |
| RESIDENTIAL               | 0.0317  |      | 2.5219        | 0.0119         |  |  |  |
| С                         | -0.5497 |      | -0.4730       | 0.6363         |  |  |  |
| Regression Statistics     |         |      |               |                |  |  |  |
| F-statistic               | 7058.5  |      | Adjusted R-sq | uared 0.986718 |  |  |  |
| Prob(F-statistic)         | 0.0000  | X    | Durbin-Watso  | on stat 1.3539 |  |  |  |

 Table 5.11: Final Results of Regression Model

Table 5.11 shows the significant variables for the purpose of delivery services in terms of time. The mentioned variables are obtained from statistical evaluation and data preparation by using regression analysis. These variables are useful in making the best decision for choosing the fast routes for the delivery of fresh vegetables.

## 5.6.3 Final substituted coefficients

The following model in the regressing analysis is the final equation contributed for this research. The final key factors, which are the most significant parameters affecting on drive time are identified as the following formula.

LTIME = -0.0664 \* LSPEED7TILL10 - 0.5387 \* LSPEED10TILL12 + 0.2663 \* CARVOLUME + 0.6984 \* LLENGTH + 0.0203 \* LPOP + 0.0605 \* TWO - WAY + 0.1681 \* SCHOOL + 0.0317 \* RESIDENTIAL - 0.5497

#### 5.6.4 Coefficients evaluation obtained from final regression model

The most significant variable among all variables in the final regression analysis is distance. The coefficient of distance, which has obtained from the formula, is 0.6984. It means that distance has the most impact among all variables on drive time. Further explanation about this matter is that shortest distance is not the solution for selecting the best routes in all cases. It means that the shortest distance is not usually the best routes in terms of drive time. Because of traffic condition sometimes, the shortest distance is not the optimal solution for selecting the roads. Therefore, the highways probably are faster compare to those roads that the distance is shorter. The shorter roads always pass through the city centre. The city centre always face with the traffic density, especially in this study area. The coefficient obtained from the formula is higher compares to another coefficient in the regression model. The second higher coefficient in the formula is the speed of the vehicle at the time 10-12. The coefficient obtained for this variable is 0.5387. The third is car volume in terms of ahigher coefficient obtained in the regression analysis. The coefficient obtained in this regard is 0.2663. The fourth is the school zones, which the coefficient obtained for this variable is 0.1681. The fifth is the speed of the vehicle in time 7-10 and the coefficient obtained for this variable is 0.0664. The sixth is at wo-way restriction. The coefficient obtained for this kind of variable is 0.0605. The seventh is residential areas. The coefficient obtained for the residential area is 0.0317. The eighth and the last parameter is population density and the coefficient obtained for this variable is 0.0203. The population has the lowest coefficient obtained among all the variables and has the lowest impact on drive time.

#### 5.6.5 Research philosophy

The advantage of current research is that the new parameter as the average speed of a vehicle within two periods of time in the morning is a new method applied to this research. Also, with consideration of the speed limit for the road networks, it is calculated the average speed of a vehicle within two periods. In many types of research, they have considered only the speed limit for the road networks in their estimation, which is the regulation and rules for the road traffic, and highways and drivers have to follow these kinds of rules. The value of this model is to integrate the spatial and non-spatial data such as social and geographical data to analyse the sufficient delivery routes for fresh vegetables. Table 5.12 shows the final significant variables obtained from the regression analysis.

| Variables                   | Sub-Variables |
|-----------------------------|---------------|
| Speed within the 7 to 10 am |               |
| Speed within 10 to 12 am    |               |
| Distance                    |               |
| Two-way routes              |               |
| The volume of traffic       |               |
| Land use                    | Residential   |
| School Zone                 |               |
| Population                  |               |

Table 5.12: Finalised Significant Variables in Regression Model

## 5.7 The final map result of the fast routes

Figure 5.33 depicts the fast routes through considering the combination of all variables inroad networks. The average drive time is calculated by dividing the speed of the vehicle to the travelled length. The average speed of the vehicle is affected by several parameters that have been discussed in the methodology section. It is concluded that volume of automobiles, a distance of routes, speed within time 7-10, speed within time 10-12 population in the routes, two-way route, the existence of schools within 100 meters and existence of a residential unit are the eight explanatory variables of the time variation observed during the transportation of fresh produces. For verification of the final routes

obtained from the GIS model in this research, it has been tested that seven out of ten routes have the correct time compare with actual road networks.



Figure 5.33: The final Map of Fast Routes Based on All Variables

# 5.8 Statistical results

According to the coefficients obtained for the independent variables, two variables of the velocity of the vehicles are observed in a significant error level of 1%, and the sign obtained from them conforms to this anticipation. In other words, considering the obtained negative coefficient, it is concluded that the increased speed of automobile reduces transmission of the products. However, since the absolute value of the obtained coefficient for the speed during 10 to 12 am is higher than the speed during 7 to 10 am, it is concluded that change of speed during 10 to 12 am has a higher effect on the reduction of necessary time for transportation of vegetables.

Based on the amounts of coefficients, 1% increase of speed within the 7 to 10 am will reduce transportation time up to 6%. This is while 1% increase of speed during 10 to 12 am has caused 53% reduction in the transmission time. In this way, the period 10 to 12 is a considered as a more suitable time for transportation of vegetables. Another explanatory variable, i.e., the volume of machines has also shown a positive and significant effect with an error level of 1%. Based on the coefficient obtained for this variable, 1% increase in the volume of machines in the investigated routes has caused 26% increase in the period required for transportation of products. The travelled distance variable also has an increasing effect according to the research prediction. Based on the coefficient of this variable, 1% increase in the travelled distance will cause 70% increase in the necessary time for transportation of products. Moreover, population variable has shown 2% increase in the necessary transmission time. However, in this regard, despite the anticipation of this study, the virtual variable of two-way routes has an increasing effect as compared to the one-way routes.

Although according to our hypothesis, a two-way route avoids mandatory turnings of automobiles and increased route and should have a negative effect on the necessary time, the positively obtained coefficient means that existence of two-way routes will cause 6% increase in the necessary time for transmission as compared to the one-way route. Another explanatory variable was the existence of schools in a distance of 100 meters from the transmission routes. The positive and significant coefficient obtained for this variable indicate that the existence of schools near transmission routes caused 16% increase in the transmission time. The virtual variable of the residential zone is also effective and significant, in the manner that if the transmission route of products is a residential unit, it will increase the transmission time up to 3%.

In this way, speed within the 7 to 10 am, speed within 10 to 12 am, volume of automobiles, distance of routes, population in the routes, two-way route, existence of schools within 100 metres and existence of a residential unit are the eight explanatory variables of the time variation observed during the transportation of fresh produces.

## 5.9 Conclusion

Real-time traffic information for all of the major roads, especially congested roads in urban areas, is some of the most important information that is necessary to produce a dynamic route planning system. It brings our attention to this point that GIS has been used with traffic and transportation models for the purpose of various transportation system analyses. GIS can both model transportation networks and integrate the association of network characteristics directly into a database, and it seems that the next generation of GIS application for transportation will be advanced transportation network analysis tools, taking advantages of GIS-based datasets.

By using the weighted average method, one can estimate the drive time along the routes with population density. A combination of methods can be used for combing the population density with the roads in Kuala Lumpur to estimate the total drive time for delivery services. The weighted average of population layer obtained from classified parameter and weighting were evaluated in a polygon attribute table and classified into five classes. It has been discovered that the land use characteristics of transportation and delivery services affect the drive time. Time is an important factor to consider when determining the best routes for delivery purposes. The weighting average method is applied for land use layers, and by the next step, the calculation for selecting the fast routes is based on the weight categories. Congestion is a temporal phenomenon, affecting some periods more than others. Measures of drive time are based on average speed, time patterns in the different day times, residential area, population, the total car volume and

distance as well. This paper presented the implementation of statistical analysis using a regression model to identifies the key factors affecting on driving time.

#### **CHAPTER 6: CONCLUSION AND RECOMMENDATIONS**

#### 6.1 Introduction

This chapter intends to review the methodology and research process and its findings, provide a suggestion for future improvements, and conclude the work. Subsequent sections expand upon the purpose of the research and provide a better understanding of the work that has been concluded with thepurpose of realising the research objectives. The significance of the study and also the contribution of GIS modelling in this researchare discussed as well. The results of this research and the outcome of GIS model will be detailed as well. The suggestions for further analysis are also explained.

## 6.2 Research findings

The results obtained from the regression model shows distance for the delivery routes have a positive effect on the delivery time of vegetables. Based on the coefficient of this variable, a 1% increase in the travelled distance will precipitate a 70% increase in the necessary travel time for the transportation of products. Taking into account the shortest time as the goal of route planning for delivery purposes, the impedance functions will be integrated by considering the length of the road and the average speed of a transportation vehicle's effect on driving time.

For the purpose of this project, only two different periods of time were considered for the impedance modelling of the delivery process. The purpose of this study is to calculate the shortest routes between two locations by taking into account the different traffic conditions in the morning. According to the coefficients that were obtained for the time pattern, it is observed that two variables of the velocity of the vehicles are observed in a significant error level of 1%. However, since the absolute value of the obtained coefficient for the speed during 10 to 12 am is higher than the speed during 7 to 10 am, it is concluded that the change of speed during 10 to 12 am has a higher effect on the reduction of necessary time for the transportation of vegetables. It is assumed from the statistical approach that time at 07.00-10.00 significantly influences travel time, which prompts us to come up with a GIS model to determine the best routes for delivery purpose.

There is a direct relationship between the population and the total volume of cars. When the population is high, the number of the cars that occupies the roads increases, and the average speed of the car correspondingly decreases, which cuts down on the drive time of these routes. It is determined that the volume of traffic on the route has an increasing effect on travelling time. Moreover, the population variable showed a 2% increase in the transmission of necessary time.

The volume of the car is another parameter that affects drive time. Congestion is directly proportional to car volume, and it increases the drive time along the roads. For this reason, the total car volume for each road was calculated and recorded into a single attribute table. The travel time fluctuates during the day, due to it being affected by factors such as traffic, vehicle volumes and routing path selection, which correspond to the most important parameters in this study. After data evaluation and assessment in the statistical approach, it was discovered that the car volume has significant value in drive time. Based on the coefficient obtained for this variable, a 1% increase in the volume of machines in the investigated routes precipitated a 26% increase in the period required for the transportation of the products.

It has been discovered that the land use characteristics of transportation and delivery services affect the drive time for a vehicle travelling on a road. The land use map considered in this research was classified into three, which includes the residential area, industrial area and open spaces. After the evaluation and analysis of the land use data in the statistical process, running the regression model and identifying the output, the residential area seems to be the most significant factor affecting drive time on designated roads. The key constraint is that the proposed path for the purpose of delivery must be sequestered from residential areas. The purpose of this particular section is to find the most suitable road to pass through a non-residential area. The results show that when a driver gets close to a residential area, their average speed will decrease, which inevitably increases drive time. The data evaluation in the statistical approach was conducted to determine which residential area in the Land use layer has the most significant value in drive time.

In a Network Analyst tool, the special requirement demands extra attention, and can be time consuming to create. Some streets can show passes in both directions, while others only one. It also needs to take into consideration the one-way roads. Restrictions are identified for particular features, so during analysis, restricted objects cannot be traversed.

#### 6.3 Outcome of the research

In summary, the research has listed three specific objectives of the study:

- **1.** To identify the key factors that influence the selection of the fastest routes for distribution.
- **2.** To develop a Spatial Database Management System for analysing the road networks dataset.
- **3.** To develop and propose a GIS model for distribution problem and determine the fastest delivery routes.

## 6.3.1 Key influential factors on route selection

The required data was analysed and tested via a regression model. After evaluating the data, the results obtained and the most significant variables were identified as the following.

In this way, the speed within the 7 to 10 am, the speed within 10 to 12 am, the volume of automobiles, the distance of routes, the population in the routes, two-way route, the

existence of schools within 100 metres, and the existence of a residential unit are the eight explanatory variables of the time variation that were observed during the transportation of fresh produces.

## 6.3.2 Spatial Database Management System

Database development was focused into two categories, a) spatial data b) non-spatial data. The GIS technologies proposed spatial data analysis.

After identifying the key factors for determining the optimal route in terms of time, the database management system will then be implemented. For this study, DBMS is essential in the application where large amounts of GIS data need to be maintained and managed. For the purpose of analysis, the GIS tool will be interconnected with the DBMS. The identified factors have been evaluated and analysed during the modelling process, and the significant variables in pathfinding in terms of time have been recognised and stored in the system for further applications.

A sample of the implemented spatial database management system is displayed in the Appendix F and the complete version is saved in CD and coming with this thesis.

# 6.3.3 The contribution of the GIS model in this research

The application of GIS has been used with traffic and transportation models for the purpose of various transportation system analyses. GIS can both model transportation networks, and integrate the association of network characteristics directly into a database. The measures of drive time are based on average speed, time patterns in differentday times, residential area, and population, the total car volume, and distance as well.

The main advantage of the developed model as compared to the existing models on route efficiently based on travel time relates to its capability in capturing the diversity of data from different fields of the area for solving distribution problem. This diversity includes the data from social and spatial aspect as well as traffic and transportation field. One of the contributions of this model compare to others is to consider about the speed of the vehicle in the vicinity of a school zone. The consideration is given to the limitation of the speed for the roads near to school zones.

Current research concentrates on the GIS tool to solve the distribution problem via the utilisation of the Network Analyst tool. This tool helps the decision makers determine the best routes among all of the existing road networks for transportation and delivery services.

Subsequent to statistical analysis with a regression model, prominent variables most significant in altering drive time were identified. The linear regression model is the most proper model for this research, and the variables that have been mentioned in literature review are tested via this model. The most significant variables are taken into account, and the one that has not contributed any value to the drive time is omitted from the analysis. The regression model will then finally contribute to the final equation for this research.

## 6.4 The significance of this research

1) The novelty about this research is to integrate the social data such as population and spatial data such as land use in developing the spatial database management system.

2) The consideration of this thesis is to add the value of traffic and transportation data such as car volume and average speed of the vehicle for all road networks. The contribution was to add these data into are gression analysis to find the significance of drive time.

3) The contribution is to connect the information about the spatial and non-spatial data with digital road networks data to solve the distribution problem to select the efficient and effective route for delivery purpose. The final output of this thesis is producing different maps integrated with road networks for decision makers to deliver their fresh vegetables.

#### 6.5 Limitations of the study

The time pattern for calculating the best routes for delivery services are included in two periods of time in the morning. The data, which has already been collected in the attributes table for average speed, resulted from two times travelling along the road, and it is more accurate if it considers other time patterns during the day and the weekends.

Due to the limitations of the data sources, this research only considers hypermarkets that are involved in the delivery services for the mentioned company. By adding more hypermarkets to the list of customers for delivery purposes, we can then calculate the most efficient routes, and get more accurate results, depending on the cost impedances identified for the attribute table.

The study is subject to the usual limitations that are associated with survey research, such as low response rate and the lack of data sources for calculating drive time. The average speed was not available for all of the roads in Kuala Lumpur. Due to the fact that calculating best routes for transportation require the speed for all the roads and the delivery services to be based on roads with minimum drive time, for the data were gathered in order to determine the average speed during the day.

A central dispatcher such as traffic monitoring under (DBKL) would observe the realtime traffic information for updating routing instructions every month. Of course, to fully reflect recent traffic flows, we would need to update the database of historical traffic information periodically.

A method for calculating the drive time that takes into consideration the car volume has been applied and a map of the average speed for all of the roads in KL is clarified. For car volume, the data was unavailable for all of the roads in Kuala Lumpur. The data that was collected through the road traffic and monitoring takes into account only the roads located in central KL. Other roads will require more calculations using ArcGIS.

#### 6.6 Suggestions for future research

There have been a lot of important tasks accomplished in this project, and these results provide a solid groundwork for future exploration in the area of optimal delivery routing. The outcome of this project has helped build foundation and rationale for future investigation and discussion with regards to the optimal distribution of fresh vegetables to hypermarkets in Kuala Lumpur.

There are more aspects of this research that can continue to be built upon going forward. It is recommended that in the future, the development of modelling cost scenarios for selecting optimal routes consider factors such as roads that are under construction and prohibited turns. These constraints provide a better solution in data evaluation and analysis to find the path for distributing fresh vegetables.

In addition to these concrete advancements, there are various new challenges for next generation route planners that arise from the considerably increasing availability of dynamic road data on the current and the upcoming traffic situation and the client's demand for route planning tailored to his individual needs. One challenge is to deal with a massive number of updates to the cost function. These updates reflect the current traffic situation, in particular, unexpected events like traffic jams and their effects on the surrounding area. Another challenge is to incorporate predictions for upcoming traffic conditions. Such predictions are based on statistical/historical data and are expressed by time-dependent cost functions.

Real-time traffic information for all of the major roads, the especially congested road in urban areas, is some of the most important information that is necessary to produce a dynamic route planning system. It is suggested in the future to apply the real road networks data for analysis approach. It can perform network analysis and access live traffic feeds through ArcGIS online services. It can also include historical traffic data that makes it possible to view travel speeds for different times of day and solve time dependent network problem.

It will be suggested on the application of real-time traffic information on speed incorporated with historical traffic data that can significantly reduce the expected drive time and vehicle usage during the time of potential heavy congestion while satisfying or improving delivery service for just in time delivery.

## 6.7 Conclusion

This research demonstrates how the GIS application is conducted obtain the best result in the calculation for obtaining fast routes in the delivery services. The regression model can provide information needed to estimate the values of a dependent variable when the values of the independent variables are known, and the relationship exists between the two variables. The regression model is used to describe the strength or weakness of this relationship between two variables.

After exploring the existing work and the data that was collected through different sources and organisations, it is noted that there is a lack of related data to process the analysis for finding fast routes. To address this concern, surveys have been carried out. This research had to determine the methodology that meets the timeframe and deliver acceptable results. By using the weighted average method, one can estimate the drive time along the routes with a population density.

It has been realised that sometimes, the shortest distance could not provide the best solutions, due to extended time consumptions that are caused by traffic conditions during the way. To calculate the best routes regarding land use, it has to be classified into different features. Hence, from the results that were obtained, the residential area is the most significant factor in selecting the fast routes. Then the calculation finds the routes that pass through the non-residential area. It has applied by giving the weights to each class of land use. These weightings value determine the relative importance of each average quantity.

This research demonstrated how a decision maker should choose the best orders to deliver fresh vegetables based on the criteria and constraints set upon in this research. Network analysis helps users to identify the shortest distance, and in the same time, the fastest routes between markets, and also provide information about the directions for the roads based on data collected in the attribute table. Network analysis can also calculate the drive time between the specific market places and show routes, which has lowest drive time among the other existing roads.

In this research, the purpose of using GIS technology was to accurately locate tasks and personnel and calculate the optimal minimal cost paths between the tasks and personnel. The derivation of the paths along a topologically consistent spatial network (i.e. roads) can take into account impedance gradients such as travel speed and distance, as well as other user specified constraints. GIS-based routing applications can solve the complex routing and scheduling problems, such as travel time, distance, appointment time, as well as other user defined constraints.

#### REFERENCES

- Akay, A. E., Wing, M. G., Sivrikaya, F., & Sakar, D. (2012). A GIS-based decision support system for determining the shortest and safest route to forest fires: a case study in Mediterranean Region of Turkey. *Environ Monit Assess*, 184(3), 1391-1407.
- Akbari, A., Abu Samah, A., & Othman, F. (2012). Integration of SRTM and TRMM date into the GIS-based hydrological model for the purpose of flood modelling. *Hydrol. Earth Syst. Sci. Discuss.*, 9(4), 4747-4775.
- Al Amin, M. D., & Uddin, M. (2015). Real time traffic monitoring system using crowd sourced GPS data (Doctoral dissertation, BRAC University).
- Alazab, A., Venkatraman, S., Abawajy, J., & Alazab, M. (2011). An optimal transportation routing approach using GIS-based dynamic traffic flows. In 3rd International Conference on Information and Financial Engineering IPEDR (Vol. 12).
- Alivand, M., Alesheikh, A., & Malek, M. (2008). New method for finding optimal path in dynamic networks. *World Applied Sciences Journal*, *3*(1), 25-33.
- Andersson, G., Flisberg, P., Lidén, B., & Rönnqvist, M. (2008). RuttOpt a decision support system for routing of logging trucks. *Canadian Journal of Forest Research*, 38(7), 1784-1796.
- Arampatzis, G., Kiranoudis, C. T., Scaloubacas, P., & Assimacopoulos, D. (2004). A GIS-based decision support system for planning urban transportation policies. *European Journal of Operational Research*, 152(2), 465-475.
- Ariffin, R. N. R., & Zahari, R. K. (2013). The Challenges of Implementing Urban Transport Policy in the Klang Valley, Malaysia. *Procedia Environmental Sciences*, 17(0), 469-477.
- Aronoff, S. (1989). Geographic information systems: A management perspective. *Geocarto International*, 4(4), 58-58.
- Aronoff, S. (1991). Geographic information systems: a management perspective. WDL *Publications*. Ottawa.
- Arvanitoyannis, I. S., & Vaitsi, O. B. (2007). A Review on Tomato Authenticity: Quality Control Methods in Conjunction with Multivariate Analysis (Chemometrics). *Critical Reviews in Food Science and Nutrition*, 47(7), 675-699.
- Avelar, S. (2008). Visualizing public transport networks: an experiment in Zurich. *Journal of Maps*, 4(1), 134-150.

- Ayalew, L., Yamagishi, H., & Ugawa, N. (2004). Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides, 1(1), 73-81.
- Azar, A. (2005). *Application of Stattistics in Management* (12 Edition ed. Vol. Firts Volume). Tehran: Samt Publisher.
- Badgie, D., Mhd Armi Abu samah., Latifah, A. M., Azizi, B. M. (2012). Assessment of Municipal Solid Waste Composition in Malaysia: Management, Practice, and Challenges. *Pol. J. Environ. Stud*, 21(3), 539-547.
- Battarra, M., Monaci, M., & Vigo, D. (2009). An adaptive guidance approach for the heuristic solution of a minimum multiple trip vehicle routing problem. *Computers & Operations Research*, 36(11), 3041-3050.
- Beck, N., & Katz, J. N. (1996). Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models. *Political Analysis*, 6(1), 1-36.
- Belenguer, J. M., Benavent, E., & Martínez, M. C. (2005). RutaRep: a computer package to design dispatching routes in the meat industry. *Journal of Food Engineering*, 70(3), 435-445.
- Berger, R. T. (2007). Location-Routing Problems with Distance Constraints. *Journal of Transportation Science*. 41, 29-43. Bethlehem, Pennsylvania
- Bhambulkar, A. (2011). Municipal solid waste collection routes optimized with arc gis network analyst. *International Journal Of Advanced Engineering Sciences And Technologies*, 11(1), 6.
- Bohari, A. M., Rainis, R., & Marimuthu, M. (2012). Fresh Food Product as a Strategy of Sustaining the Hypermarket Lifetime Value: An Evident from Survey and Un-Structured Interview's Session. OIDA International Journal of Sustainable Development, 3(7), 39-46.
- Bolstad, P. (2012). *GIS Fundamentals: A First Text on Geographic Information Systems*. (4th Edition ed.). White Bear Lake, Minnesota: Eider Press.
- Bräysy, O., Dullaert, W., & Nakari, P. (2009). The potential of optimization in communal routing problems: case studies from Finland. *Journal of Transport Geography*, *17*(6), 484-490.
- Brimicombe, A. (2009). *GIS, Environmental Modeling and Engineering, Second Edition*: Taylor & Francis.
- Brosi, P., & Bast, H. (2014). Real-Time Movement Visualization of Public Transit Data (Doctoral dissertation, MS thesis, Faculty of Engineering Department of Computer Science, Univ. of Freiburg, Freiburg, Germany).

- Burrough, P.A. & Frank, A.U. (1995). Concepts and paradigms in spatial information: Are current geographical information systems truly generic?, *International Journal of GeographicalInformation Systems*. 1995, 9:101-116.
- Burroughs, P.A. (1986). Principles of Geographic Information Systems for Land Resources Assessment. Oxford University press.
- Butler, M., Herlihy, P., & Keenan, P. B. (2005). Integrating information technology and operational research in the management of milk collection. *Journal of Food Engineering*, *70*(3), 341-349.
- Cairns, S. (1998). Promises and problems: using GIS to analyse shopping travel. *Journal* of *Transport Geography*, 6(4), 273-284.
- Cascetta, E., Russo, F., Viola, F. A., & Vitetta, A. (2002). A model of route perception in urban road networks. *Transportation Research Part B: Methodological*, *36*(7), 577-592.
- Çela, L., Shiode, S., & Lipovac, K. (2013). Integrating GIS and spatial analytical techniques in an analysis of road traffic accidents in Serbia. *International Journal for Traffic and Transport Engineering*, 3(1), 1-15.
- Chamhuri, N., & Batt, J. (2013). Exploring the Factors Influencing Consumers' Choice of Retail Store When Purchasing Fresh Meat in Malaysia. 3 International Food and Agribusiness Management Association (IFAMA), (pp. Volume 16, Special Issue 3). Atlanta.
- Chang, K. T. (2008). Introduction to Geographic Information Systems: McGraw-Hill.
- Chen, H. K., Hsueh, C. F., & Chang, M. S. (2009). Production scheduling and vehicle routing with time windows for perishable food products. *Computers & Operations Research*, *36*(7), 2311-2319.
- Chen, T. Y., Chang, H. L., & Tzeng, G. H. (2001). Using a weight-assessing model to identify route choice criteria and information effects. *Transportation Research Part A: Policy and Practice*, 35(3), 197-224.
- Chen, Y. W., Wang, C. H., & Lin, S. J. (2008). A multi-objective geographic information system for route selection of nuclear waste transport. *Omega*, *36*(3), 363-372.
- Chrisman, N. R. (2002). *Exploring geographic information systems*. The University of Michigan: Wiley.
- Clarke A M and F G Thomas (1991). The Geography of the 1991 Census, *Population Trends* 60; 9-15
- Coppock, J. T., & Rhind, D. W. (1991). The history of GIS. *Geographical information* systems: Principles and applications, 1(1), 21-43.

- Cowen, D. J. (1988). GIS versus CAD versus DBMS: What are the differences? *Photogramm. Eng. Remote Sens.*, 54(11), 1551-1555.
- Dannenberg, P., Kunze, M., & Nduru, G. M. (2011). Isochronal Map of Fresh Fruits and Vegetable Transportation from the Mt. Kenya Region to Nairobi. *Journal of Maps*, 7(1), 273-279.
- Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management* science, 6(1), 80-91.

DeMers, M. N. M. N. (2002). GIS modeling in raster (No. 910.285 D4).

Dessing, D., de Vries, S. I., Hegeman, G., Verhagen, E., Van Mechelen, W., & Pierik, F. H. (2016). Children's route choice during active transportation to school: difference between shortest and actual route. *International journal of behavioral nutrition and physical activity*, 13(1), 1.

Dicken, P. (1998). Global shift: transforming the world economy.

- Dissanayake, D., Kurauchi, S., Morikawa, T., & Ohashi, S. (2012). Inter-regional and inter-temporal analysis of travel behaviour for Asian metropolitan cities: Case studies of Bangkok, Kuala Lumpur, Manila, and Nagoya. *Transport Policy*, *19*(1), 36-46.
- Dong, J., & Shen, G. (2012). A Weighted-based Road Impedance Function Model. Paper presented at the 2nd International Conference on Computer and Information Application (ICCIA 2012), Paris, France.
- Duran-Fernandez, R., & Santos, G. (2014). A GIS model of the National Road Network in Mexico. *Research in Transportation Economics*, 46, 36-54.
- Echols, N. (2003). *Finding the Fastest Route*. Paper presented at the ESRI User Conference, URL:http://gis.esri.com/library/userconf/proc03/p0320. pdf, Redlands, California (last date accessed 29 January 2007).
- ESRI, A. (2012). 10.1. Environmental Systems Research Institute, Redlands, CA.
- Faulin, J. (2003). Applying MIXALG procedure in a routing problem to optimize food product delivery. *Omega*, 31(5), 387-395.
- Fazal, S. (2008). GIS Basics: New Age International (P) Limited, Publishers.
- George, M. (1991). *Transportation GIS applications*. Paper presented at the Proceedings of teh 1991 Geographic Information Systems (GIS) for Transportation Symposium.

- Ghose, M. K., Dikshit, A. K., & Sharma, S. K. (2006). A GIS based transportation model for solid waste disposal – A case study on Asansol municipality. *Waste Management*, 26(11), 1287-1293.
- Gkolias, M., & Boile, M. (2010). Truck Demand Modeling with ArcInfo 9.0. A Proposal Report by Center for Advanced Infrastructure and Transportation (CAIT), Rutgers University
- Gohari, A. (2010). Route Planning System based on Geographic Information System. (Master of Sceince (M.Sc)), University Technology Malaysia (UTM), University Technology Malaysia
- Goodchild, M. (2000). GIS and Transportation: Status and Challenges. *GeoInformatica*, 4(2), 127-139.
- Goodchild, M. F. (1992). Geographic Information Science, International Journal of Geographic Information Systems, 6:31-45.
- Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3), 239-260.
- Hashim, R., Mohamad, S., Haron, S., Hassan, F., Hassan, N. A., & Kasa, A. (2013, April). Student satisfaction with the campus bus services at UiTM Shah Alam, Malaysia. In Business Engineering and Industrial Applications Colloquium (BEIAC), 2013 IEEE (pp. 723-727). IEEE.
- Hassan, H., & Rahman, M. S. (2012, May). Transformation of hypermarket retailing industry in Malaysia. In *Innovation Management and Technology Research* (*ICIMTR*), 2012 International Conference on (pp. 513-516). IEEE.
- Hassan, H., Bakar Sade, A., & Sabbir Rahman, M. (2013). Malaysian hypermarket retailing development and expansion. *International Journal of Retail & Distribution Management*, *41*(8), 584-595.
- Hato, E., Taniguchi, M., Sugie, Y., Kuwahara, M., & Morita, H. (1999). Incorporating an information acquisition process into a route choice model with multiple information sources. *Transportation Research Part C: Emerging Technologies*, 7(2–3), 109-129.
- Hawas, Y. E. (2013). Simulation-based regression models to estimate bus routes and network travel times. *Journal of Public Transportation*, 16(4).
- Healey, R. (1991). Database management systems. *Geographic Information Systems: Principles and Applications, 1,* 251-267.
- Hensher, D. A., & Button, K. J. (2000). Handbook of Transport Modeling (pp. 1-9). Oxford, UK: Elsevier Science Ltd.

- Hensher, D. A., Button, K. J., Haynes, K. E., & Stopher, P. R. (2004). Handbook of transport geography and spatial systems. The School of Transport Studies, University of Sydney: Elsevier Science Ltd.
- Henstra D, C Ruijgrokand & L Tavasszy (2007). Globalized trade, logistics and intermodality: European perspectives. In Leinbach TR & C Capineri (ed). Globalized Freight Transport, 135 163. Edward Elgar Publishing, Cheltenham.
- Heywood, I., Cornelius, S., & Carver, S. (2002). An Introduction to Geographical Information Systems (pp. 336).
- Horner, M., & Downs, J. (2007). Testing a Flexible Geographic Information System-Based Network Flow Model for Routing Hurricane Disaster Relief Goods. *Transportation Research Record: Journal of the Transportation Research Board*, 2022(-1), 47-54.
- Hsu, C., Hung, S., & Li, H. (2007). Vehicle routing problem with time-windows for perishable food delivery. *Journal of Food Engineering*, *Volume 80, Issue 2*, 465–475.
- Ichoua, S., Gendreau, M., & Potvin, J.-Y. (2003). Vehicle dispatching with timedependent travel times. *European Journal of Operational Research*, 144(2), 379-396.
- Indati, M. S., Ghate, A. T., & Leong, Y. P. (2013). Towards greener environment: Energy efficient pathways for the transportation sector in Malaysia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 16, No. 1, p. 012122). IOP Publishing.
- Ioannou, G., Kritikos, M. N., & Prastacos, G. P. (2002). Map-Route: A GIS-Based Decision Support System for Intra-City Vehicle Routing with Time Windows. *The Journal of the Operational Research Society*, *53*(8), 842-854.
- Ismail, M. A., & Said, M. N. (2014). Integration of geospatial multi-mode transportation Systems in Kuala Lumpur. In *IOP Conference Series: Earth and Environmental Science* (Vol. 20, No. 1, p. 012027). IOP Publishing.
- Iwan, G., (2003). Geographic Information Systems Markets and Opportunities. Handout. Master of Science in Information Technology for Natural Resources Management. Bogor Agricultural University, Bogor, Indonesia.
- Jankowski, P. (1995). Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Systems*, 9(3), 251-273.
- Jenelius, E. (2010). *Large-scale road network vulnerability analysis* (Doctorate), KTH royal institute of technology, Sweden.

- Kamath, P., & Godin, C. (2001). French Carrefour in South-East Asia. British Food Journal, 103(7), 479-494.
- Kamil, F. M. (2005). Agricultural policy and sustainable development in Malaysia: Kolej Universiti Sains dan Teknologi Malaysia.
- Khaki, Gh. (1995). Research Method with Thesis Writing Approach, Tehran, Ministry of Culture and High Education, The Center of National Scientific Researches.
- Khalid, U. A., Bachok, S., Osman, M. M., & Ibrahim, M. (2014). User Perceptions of Rail Public Transport Services in Kuala Lumpur, Malaysia: KTM Komuter. *Procedia-Social and Behavioral Sciences*, 153, 566-573.
- Khan, N., Salman, M. S., Khan, M. M. (2014). Agricultural Marketing, Economic Benefits and Social Groups: Exploring Vegetable Cultivation in Rural India. *Euro-Asian Journal of Economics and Finance*, 2(4), 335-346.
- Kim, H. Y., Wunneburger, D., Neuman, M., & An, S. Y. (2014). Optimizing high-speed rail routes using a Spatial Decision Support System (SDSS): the Texas Urban Triangle (TUT) case. *Journal of Transport Geography*, 34, 194-201.
- Kuby, Michael, et al. "An efficient online mapping tool for finding the shortest feasible path for alternative-fuel vehicles." *International Journal of Hydrogen Energy* 39.32 (2014): 18433-18439.
- Lam Shan, V. (1997). Network and Urban Form Analyses: An Approach to Routing Bus Transit in Geographic Information Systems. (Master of Philosophy), University of Hong Kong, Hong Kong.
- Lam, T.C., Small, K.A., (2001). The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E* (37), 231–251.
- Leonard, J., & Oliveira, M. (2000). *Towards an areawide service measure*. Paper presented at the Proceedings of the fourth international symposium on highway capacity.
- Li, H., Guensler, R., & Ogle, J. (2005). Analysis of Morning Commute Route Choice Patterns Using Global Positioning System-Based Vehicle Activity Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1926(1), 162-170.
- Li, Q., & Zeng, Z. (2008, October). A Voronoi-based hierarchical graph model of road network for route planning. In 2008 11th International IEEE Conference on Intelligent Transportation Systems (pp. 599-604). IEEE.
- Liu, Y. (2008). Assessing the Reliability of NFIRS Database in Exploring Emergency Response Route-Selection: Using the Fire Department of the City of Lebanon, Ohio as an Example. University of Cincinnati. Retrieved from http://rave.ohiolink.edu/etdc/view?acc\_num=ucin1250045718

- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic Information Systems and Science*: Wiley.
- Lopes, S. B., Brondino, N. C. M., & Rodrigues da Silva, A. N. (2014). GIS-based analytical tools for transport planning: Spatial regression models for transportation demand forecast. *ISPRS International Journal of Geo-Information*, 3(2), 565-583.
- Malczewski, J. (2000). On the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. *Transactions in GIS*, 4(1), 5-22.
- Manheim, M. L. (1980). Understanding "supply" in transportation systems. *Transportation Research Part A: General*, 14(2), 119-135.
- Martin, D. (1991). Geographic information systems and their socioeconomic applications, London, Routledge.
- Martin, D. (1996). *Geographical Information Systems: Socio-economic Applications* (2<sup>nd</sup> ed.). Routledge, London.
- Martin, J. (1991). Rapid application development: Macmillan Publishing Co., Inc.
- Memon, I. A. (2005). Application of geographic information system in transportation for road network analysis. Universiti Teknologi Malaysia, Faculty of Civil Engineering.
- Menard, S. (1995). Applied logistic Regression analysis. Series on Quantitative Applications in the Social Science. *SAGE*, *106*, 98.
- Moon, T. and Hagishima, H. (2001). Integrated Simulation System of GIS and ANN for Land Price Appraisal. Paper presented at the 7th International Conference on CUPUM, University of Hawaii, Honolulu, 18-21 July.
- Mukund, Rao., et al. (1992). Network module for ISROGIS: *design consideration and specifications*. Space Applications Center.
- Musyoka, S., Mutyauvyu, S., Kiema, J., Karanja, F., & Siriba, D. (2007). Market segmentation using geographic information systems (GIS): A case study of the soft drink industry in Kenya. *Marketing Intelligence & Planning*, 25(6), 632-642.
- Nasir, M (2005). Changing Retail Food Sector in Malaysia. *Thesis*, University Putra Malaysia, Malaysia.
- National Oceanic and Atmospheric Administration [NOAA]. (1996). Estuarine eutrophication survey. Volume 1: South Atlantic Region. Office of Ocean Resources Conservation Assessment, *National Ocean Service, Silver Spring*, MD.

- Nayati, M. A. K. (2008). School bus routing and scheduling Using GIS. Thesis Advanced level (degree of Magister), University of Gävle, Department of Technology and Built Environment.
- NCGIA, (1990). National Center for Geographic Information and Analysis, *University of California*, Santa Barbara, USA. 3 volumes.
- Nielsen, F. (1994). GIS database design and organization: Training course on GIS for resources management and development planning.
- Noonan, J., & Shearer, O. (1998). Intelligent Transportation systems Field operational test cross-cutting study: Advance traveler information systems.
- O'Brien, L. (1992). Introducing Quantitative Geography: Measurement, Methods, and Generalised Linear Models: Routledge.
- Osvald, A., & Stirn, L. Z. (2008). A vehicle routing algorithm for the distribution of fresh vegetables and similar perishable food. *Journal of Food Engineering*, 85(2), 285-295.
- Panwhar, S. T., Pitt, R., & Anderson, M. D. (2000). Development of a GIS-based hazardous materials transportation management system. UTCA report, 99244.
- Papinski, D., & Scott, D. M. (2011). A GIS-based toolkit for route choice analysis. Journal of Transport Geography, 19(3), 434-442.
- Parsafard, M., Esmaeel, A., Masoud, K., Mohammadreza, N., & Li, X. (2015). Practical Approach for Finding Optimum Routes for Fuel Delivery Trucks in Large Cities. *Transportation Research Record: Journal of the Transportation Research Board*, (2478), 66-74.
- Pimenta, L. (2010). *The Logistics of milk collection: an exploratory case study between New Zealand and Brazil.* (M.Sc), Massey University New Zealand.
- Prindezis, N., Kiranoudis, C., & Marinos-Kouris, D. (2003). A business-to-business fleet management service provider for central food market enterprises. *Journal of Food Engineering*, 60(2), 203-210.
- Raju, P. (2003). Spatial data analysis. Satellite Remote Sensing and GIS Applications in Agricultural Meteorology, 151-174.
- Rezai, G., Mohamed, Z., & Shamsudin, M. N. (2011). Malaysian consumer's perceptive towards purchasing organically produce vegetable. Paper presented at the 2nd International Conference on Business and Economic Research, Holiday Villa Beach Resort and Spa, Langkawi, Kedah, Malaysia. Retrieved from http://www. internationalconference. com. my/proceeding/2ndicber2011\_proceeding/310-2nd% 20ICBER.

- Rice, M. N. (2005). A new hybrid computational intelligence algorithm for optimized vehicle routing applications in geographic information systems. University of Georgia.
- Rich, S., Davis, K. H. (2010). Geographic Information Systems (GIS) for Facility Management. *IFMA Foundation*
- Rittgers, C. P. (2013). Global Agricultural Information Network Kuala Lumpur, Malaysia

Rogerson, P. (2001). Statistical Methods for Geography: SAGE Publications. USA.

- Rout, U. K., Akimoto, K., Sano, F., Oda, J., Homma, T., & Tomoda, T. (2008). Impact assessment of the increase in fossil fuel prices on the global energy system, with and without CO2 concentration stabilization. *Energy Policy*, 36(9), 3477-3484.
- Sadeghi-Niaraki, A., Varshosaz, M., Kim, K., & Jung, J. J. (2011). Real world representation of a road network for route planning in GIS. *Expert Systems with Applications*, 38(10), 11999-12008.
- Salonen, M., & Toivonen, T. (2013). Modelling travel time in urban networks: comparable measures for private car and public transport. *Journal of Transport Geography*, *31*, 143-153.
- Samet, H. (1990). *The design and analysis of spatial data structures* (Vol. 85): Addison-Wesley Reading, MA.
- Sanjeevi, V., & Shahabudeen, P. (2016). Optimal routing for efficient municipal solid waste transportation by using ArcGIS application in Chennai, India. Waste Management & Research, 34(1), 11-21.
- Schwarcz, S. (2003). Public Transportation in Kuala Lumpur, Malaysia. MTRG, Malaysia.

Schwarcz, S. L. (2003). Securitization Post-Enron. Cardozo L. Rev., 25, 1539.

- Selten, R., Chmura, T., Pitz, T., Kube, S., & Schreckenberg, M. (2007). Commuters route choice behaviour. *Games and Economic Behavior*, 58(2), 394-406.
- Shahnaei, S. (2012). The Impact of Individual Differences on Green Purchasing of Malaysian Consumers. International Journal of Business and Social Science, 3(16), 132-140.
- Shamsudin, I. (2008). Transportation Optimization Model Of Palm Oil Products For Northern Peninsular Malaysia. Universiti Sains Malaysia.
- Shamsudin, M., & Selamat, J. (2005). *Changing retail food sector in Malaysia*. Paper presented at the PECC Pacific food System Outlook Annual Meeting, held.

- Sharifi, M., Hadidi, M., Vessali, E., Mosstafakhani, P., Taheri, K., Shahoie, S., & Khodamoradpour, M. (2009). Integrating multi-criteria decision analysis for a GIS-based hazardous waste landfill sitting in Kurdistan Province, western Iran. *Waste Management*, 29(10), 2740-2758.
- Shaw, S.-L., & Xin, X. (2003). Integrated land use and transportation interaction: a temporal GIS exploratory data analysis approach. *Journal of Transport Geography*, 11(2), 103-115.
- Sheng, T. Y., Shamsudin, M. N., Mohamed, Z. A., Abdullah, A. M., & Radam, A. (2008). Food Consumption Behavior of the Malays in Malaysia. *International Journal* of Economics, Management and Accounting, 16(2).
- Shokri, F., Chu, M.-Y., Mokhtarian, H. R., Rahmat, R., & Ismail, A. (2009). Best route based on fuel-economy. *European Journal of Scientific Research*, 32(2), 177-186.
- Siangsuebchart, S. Winyoopradist, S. (2000). Integration of two scales road network analysis (Country and city). ESRI. User Conference Proceedings.
- Silva, F., Minette, L. J., Souza, A. P. D., Moraes, Â. C. D., & Schettino, S. (2016). CLASSIFICATION OF FOREST ROADS AND DETERMINATION OF ROUTE USING GEOGRAPHIC INFORMATION SYSTEM. *Revista* Árvore,40(2), 329-335.
- Sloof, M., Tijskens, L. M. M., & Wilkinson, E. C. (1996). Concepts for modelling the quality of perishable products. *Trends in Food Science & Technology*, 7(5), 165-171.
- Sofia, A. S. S. D., Nithyaa.R, & Arulraj.G, P. (2013). Minimizing the Traffic Congestion Using GIS. International Journal of Research in Engineering & Advanced Technology, 1(1), 1-6.
- Song, S. H., Lee, K. S., & Kim, G. S. (2002). A practical approach to solving a newspaper logistics problem using a digital map. *Computers & industrial engineering*, 43(1), 315-330.
- Stockburger, D. (1996). *Introduction to Statistics: Concepts, Models and Applications*: On-line at URL: http://www.psychstat.smsu.edu/sbk00. Htm
- Suárez-Vega, R., Santos-Peñate, D. R., Dorta-González, P., & Rodríguez-Díaz, M. (2011). A multi-criteria GIS based procedure to solve a network competitive location problem. *Applied Geography*, 31(1), 282-291.
- Syafiq, M., A., S. (2007). An Expert System for Recommending Speed Limits. (BSC. (HONS.) INTELLIGENT SYSYTEM), University Technology of MARA Malaysia

- Tarantilis, C., & Kiranoudis, C. (2001). A meta-heuristic algorithm for the efficient distribution of perishable foods. *Journal of Food Engineering*, 50(1), 1-9.
- Tarantilis, C., & Kiranoudis, C. (2002). Distribution of fresh meat. *Journal of Food Engineering*, 51(1), 85-91.
- Tarantilis, C., Diakoulaki, D., & Kiranoudis, C. (2004). Combination of geographical information system and efficient routing algorithms for real life distribution operations. *European Journal of Operational Research*, 152(2), 437-453.
- Tatomir, B., Rothkrantz, L., & Suson, A. (2009). Travel time prediction for dynamic routing using ant based control. *Paper presented at the Simulation Conference (WSC)*.
- Tey, L. (2003). Trip Generation Model for University Technology Malaysia (UTM) Skudai. University Technology Malaysia (UTM), Malaysia.
- Thill, J.-C. (2000b). Geographic information systems for transportation in perspective. *Transportation Research Part C: Emerging Technologies*, 8(1–6), 3-12.
- Toth, P., & Vigo, D. (2002). Models, relaxations and exact approaches for the capacitated vehicle routing problem. *Discrete Applied Mathematics*, 123(1), 487-512.
- Tram, H. (1999). Using a distrubution management system to improve asset management. *Network Operations Management*.
- Tram, H., Engelken, L., & Gay, A. (1999). Implementation of an Information Technology Architecture for Energy Delivery Utility Mergers. In *DistribuTech Conference*, *San Diego*.
- Trochim, W. M. K. (2006). Research Methods Knowledge Base. Cornell University, USA. fromhttp://www.socialresearchmethods.net/kb/index.php
- Turban, E. (1993). Decision Support and Expert Systems: Management Support Systems: Macmillan.
- Vasserman, S., Feldman, M., & Hassidim, A. (2015, June). Implementing the wisdom of waze. In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015).
- Wang, Z., Che, O., Chen, L., Lim. A. (2006). An Efficient Shortest Path Computation System for Real Road Networks. Springer-Verlag Berlin Heidelberg, 40(31), 711-720.
- Werner, C. (1985). Spatial transportation modeling. Sage Publications, Incorporated, Beverly Hills, USA.

- Winyoopradist, S., & Siangsuebchart, S. (1999). *Network Analysis for variable traffic speed*. Paper presented at the ESRI User Conference.
- Wong, A., Taylor, G. D., Taha, H. A., & Wootten, J. (1997). GIS-BASED FREIGHT DENSITY AND CAPACITY MODELING (pp. 33): University of Arkansas.
- Xiaolin, X. (2006). *Optimal decisions in a time-sensitive supply chain with perishable products* (Vol. 68).
- Xue, W., & Cao, K. (2016). Optimal routing for waste collection: a case study in Singapore. International Journal of Geographical Information Science, 30(3), 554-572.
- Zailani, S., Iranmanesh, M., Nikbin, D., & Jumadi, H. B. (2014). Determinants and environmental outcome of green technology innovation adoption in the transportation industry in Malaysia. Asian Journal of Technology Innovation, 22(2), 286-301.
- Zlatanova, S., Stoter, J., & Quak, W. (2004). *Management of multiple representations in spatial DBMSs.* Paper presented at the Proceedings of the 7th AGILE Conference on Geographic Information Science, Heraklion, Greece.
- Zhu, S., & Levinson, D. (2015). Do people use the shortest path? An empirical test of Wardrop's first principle. *PloS one*, *10*(8), e0134322.