UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS

School of Civil Engineering and the Environment

Journey Time Estimation and Incident Detection Using GPS Equipped Probe Vehicle

by

Yanying Li

Thesis for the degree of Doctor of Philosophy

March 2004

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF CIVIL ENGINEERING AND THE ENVIRONMENT

Doctor of Philosophy

JOURNEY TIME ESTIMATION AND INCIDENT DETECTION
USING GPS EQUIPPED PROBE VEHICLE

by Yanying Li

This thesis presents a study of the use of GPS equipped probe vehicle to collect traffic data on a motorway network. The performance of the GPS information in journey time estimation has been studied by comparing the results against video camera data and the various factors affecting estimation accuracy have been discussed. By discontinuing the use of Selective Availability, one of the main error sources of GPS, current positioning accuracy without Differential GPS is sufficient for journey time estimation.

Two types of GPS equipped probe vehicles, active and passive, have been studied. A passive probe vehicle was considered to provide only link journey time and a minimum number of probe vehicles is required for reliable estimation. This research has studied the distribution of individual journey times and calculated the sample size of probe vehicles required in different traffic conditions. The sample size has shown to be generally stable for the same link, but may decrease in heavier traffic. The use of real-time estimates of journey time by probe vehicles has been studied for incident detection and journey time prediction. Link journey times at current time intervals and the differences in journey times between two adjacent time intervals have been shown to be bivariate-normally distributed in incident-free traffic. Outliers of the distribution were considered to be observed in incident traffic. A bivariate model has been developed for incident detection and a satisfactory detection and false alarm rates have been achieved. Journey times were predicted based on current observations and historic data for incident and incident-free conditions.

An active probe vehicle was found to be able to determine vehicle position and speed at 1 Hz frequency over an entire journey. By analysing the speed profile of probe vehicles, journey times can be estimated from fewer probe vehicles than normally required. In this research, a fuzzy model was developed to analysis speed profiles, and journey time could be estimated using a single probe vehicle. Satisfactory estimates were obtained in both non-incident and incident conditions. Combinations of average speed and deceleration rates were used for incident detection.

Contents

Abstract i
Contents ii
List of Figures vi
List of Tables viii
Declaration of authorship ix
Acknowledgements x
Definitions of Symbols xi
Abbreviations xii

1. Introduction **1**
1.1 Background 1
1.2 Objectives 3
1.3 Approach 3

2. GPS equipped probe vehicle **5**
2.1 GPS technology 5
2.1.1 Overview of GPS 5
2.1.2 Applications in transportation 6
2.2 Probe vehicle 7
2.2.1 Active probe vehicle 8
2.2.2 Passive probe vehicle 8
2.3 Applications of probe vehicle 9
2.3.1 Data collection techniques 9
2.3.2 Previous experiences 10
2.3.2.1 Beacon-based automatic vehicle location 10
2.3.2.2 Automatic vehicle identification 11
2.3.2.3 Cellular phone positioning 12
2.3.2.4 GPS equipped probe vehicle 14
2.4 Summary 16

3. Data collection **18**
3.1 Introduction 18
3.2 Survey Site 18
3.3 Video camera data 19
3.3.1 ANPR Technology 19
3.3.2 ANPR Data Component 22

3.3.3 ANPR Data Applications 23
3.4 Loop data 25
3.4.1 Loop data component and reliability 25
3.4.2 Missing data replacement 26
3.5 Probe vehicle data 26
3.5.1 Main surveys 26
3.5.2 GPS data component 28
3.5.3 GPS data conversion 28
3.6 Cleaning of ANPR Data 29
3.6.1 Assessment of ANPR data quality 31
3.6.2 Outlier detection 33
3.6.2.1 Previous researches 33
3.6.2.2 Outlier detection method 34
3.7 Summary 37

4. Journey time estimation 38
4.1 Journey time calculation 38
4.2 Accuracy of journey time estimation by GPS 40
4.3 Effect of GPS update frequency 41
4.3.1 Accuracy at various GPS update frequency 42
4.3.2 Accuracy at various GPS update frequency with correlation 44
4.4 Journey time report 47
4.5 Summary 47

5. Sample size of probe vehicles 49
5.1 Introduction 49
5.2 Literature review 50
5.2.1 Empirical data 50
5.2.2 Based on simulation data 52
5.3 Determining probe vehicle sample size 54
5.4 Vehicle journey time distribution 56
5.4.1 Normal distribution 56
5.4.1.1 Normal distribution test: an example 56
5.4.1.2 Percentage of normal distribution on each link 58
5.4.2 Non-normal distribution 58
5.5 Determining sample size of probe vehicles 61
5.5.1 Sample size and average speed 61
5.5.2 Sample size and traffic flow 64
5.5.3 Probe vehicle sample size for different links 65
5.5.4 Sample size for non-normality period 68

5.6 Summary 69

6. Incident detection 70

6.1 Introduction 70
6.2 Literature review 72
6.2.1 Methods for evaluating algorithm 72
6.2.2 Automatic incident detection algorithms 73
6.2.2.1 AID algorithms based on loops 73
6.2.2.2 AID algorithms based on probe vehicles 78
6.2.3 Summary of AID algorithms 80
6.3 Methodology 80
6.3.1 Incident data 80
6.3.2 Model input 81
6.3.3 Proposed algorithm 82
6.4 Model Development 83
6.4.1 Variable Definition 83
6.4.2 Bivariate Analysis 84
6.4.3 Parameter estimation 86
6.4.3.1 Link journey time 86
6.4.3.2 Difference of link journey time between adjacent time intervals 87
6.4.3.3 Correlation Coefficient 88
6.5 Model evaluation 90
6.5.1 Model performance 90
6.5.2 Performance analysis 91
6.6 Summary 92

7. Journey time prediction 93

7.1 Introduction 93
7.2 Literature review 94
7.3 Journey time prediction in non-incident condition 95
7.3.1 One-step-ahead prediction 95
7.3.2 Multi-step-ahead prediction 98
7.4 Journey time prediction in incident condition 100
7.4.1 Breakdown stage 104
7.4.2 Recovery stage 107
7.4.3 Summary of prediction procedure 108
7.4.4 An example of prediction 109
7.5 Summary 110

8. Active probe vehicle 112

8.1 Introduction 112

8.2 Speed profile development 113

8.3 Journey time estimation using single GPS equipped probe vehicle 113

8.3.1 Driving pattern 113

8.3.2 Definition of variables 115

8.3.3 Parameter estimation 117

8.3.4 Fuzzy model development 120

8.3.5 Model validation 122

8.4 Incident detection using single GPS equipped probe vehicle 124

8.4.1 Methodology 125

8.4.2 Result analysis 127

8.5 Summary 128

9. Summary and conclusions 130

9.1 Main findings of the research 131

9.2 Potential applications 133

9.3 Further work 134

Appendix A: Converting ellipsoidal latitude and longitude to grid eastings and northings 136

Appendix B Sampling of normal distribution 138

B1 Notation of $z_{\alpha/2}$ 138

B2 Determination of sample size 138

Appendix C Bivariate data analysis 141

C1 Bivariate normal distribution 141

C2 Assessing bivariate normality 143

Appendix D Speed change in incident period 145

References 148

List of Figures

Figure 1.1 Structure of the thesis 4
Figure 2.1 The floating vehicle data system 16
Figure 3.1 Survey site map with locations of video cameras 20
Figure 3.2 The paths of the journeys as recorded by the GPS device 30
Figure 3.3 Examples of raw journey time data on link 3: Thursday, June 12 2001 30
Figure 3.4 The path of the journey as recorded by the GPS device 32
Figure 3.5 Identification of outliers on link 3, August 27, 2002 35
Figure 3.6 Cleaned journey time data on link 3: Thursday, June 12 2001 36
Figure 4.1 Format of raw VRN data 40
Figure 4.2 Comparison of journey times estimated from GPS and ANPR 41
Figure 4.3 An example of sampling for 5-s interval 42
Figure 4.4 Errors in journey time estimation at 5 second interval 43
Figure 4.5 Estimating the closest time stamp to a checkpoint 44
Figure 4.6 Speed profiles at various update intervals 45
Figure 4.7 Distribution of journey time difference at various update intervals 46
Figure 5.1 Standard error versus. number of probes 51
Figure 5.2 Probe vehicle percentage versus demand level 53
Figure 5.3 Normal probability plot 57
Figure 5.4 Observed speed-flow relationship in link 7 59
Figure 5.5 Link journey time with flow data in link 7 59
Figure 5.6 Non-normal distribution intervals, July 9 2001 60
Figure 5.7 The journey time distribution for each minute 61
Figure 5.8 Mean journey time and standard deviation for link 3 62
Figure 5.9 Relationship between the mean link speed and required minimum sample size for link 3 ... 63
Figure 5.10 Traffic flow and probe vehicle sample size 65
Figure 5.11 Probe vehicle sample size for different links 66
Figure 5.12 Comparison of required sample size observed and estimated from the traffic flow 67
Figure 6.1 Phases of a traffic incident 70
Figure 6.2 99% and 99.9% coverage contours 85
Figure 6.3 Mean of link journey time at each time interval 87
Figure 6.4 Standard deviation of link journey time at each time interval 87
Figure 6.5 Mean and standard deviation of difference of journey time at two adjacent time intervals .. 88
Figure 6.6 Bivariate normal density of JT_i and DT_i 90
Figure 7.1 Journey time on different days 95
Figure 7.2 Scatterplot of ZJT_i vs. ZJT_{i+1} 96
Figure 7.3 Comparison of observed and predicted journey times on link 2, Sept 30 2002 97

Figure 7.4 Multi-step-ahead journey time predictions on link 2, Sept 30 2002 101
Figure 7.5 Prediction accuracy for various time intervals 102
Figure 7.6 Journey time changes during an incident period on link 2, October 11 2002 102
Figure 7.7 Average speed during an incident period on link 2, October 11 2002 103
Figure 7.8 Speed and fitted curves in breakdown stage 105
Figure 7.9 Length of recovery stage and mean speed 107
Figure 7.10 Comparison of observation and predicted speed on link 1, Oct 11 2002 110
Figure 8.1 Time-speed and space-speed profiles 114
Figure 8.2 Different driving patterns 115
Figure 8.3 Membership functions of average speed 117
Figure 8.4 The membership functions of average speed for link 2 and the interval of 7:50-7:55 119
Figure 8.5 The membership functions of maximum continuous acceleration 119
Figure 8.6 The membership functions of the output variable 120
Figure 8.7 Graphic presentation of the fuzzy model 120
Figure 8.8 Comparison of estimated journey times and ANPR data for link 1 122
Figure 8.9 Comparison of estimated journey times and ANPR data for link 2 123
Figure 8.10 Speed profiles in incident and non-incident condition 124
Figure 8.11 Incident detection on link 1 127

List of Tables

Table 1.1 Summary of traffic detectors 2
Table 2.1 Comparison of probe vehicle data collection techniques 10
Table 3.1 Link length and average speed 19
Table 3.2 Video camera data weeks 22
Table 3.3 An example of individual vehicle journey time data 23
Table 3.4 Outliers of ANPR journey time 31
Table 4.1 Statistics of errors in journey time estimation at various update intervals 46
Table 5.1 Values of $z_{\alpha/2}$ 55
Table 5.2 One-sample Kolmogorov-Smirnov Test 57
Table 5.3 Normality percentage for each link 58
Table 5.4 Sample size for each link 66
Table 5.5 Minimum number of probe vehicles for 1-minute interval 68
Table 6.1 Incident magnitudes 71
Table 6.2 Typical capacity reductions of during incident conditions 71
Table 6.3 Definition of parameters used in California algorithms #8 75
Table 6.4 Critical journey time ratios for declaring incidents 78
Table 6.5 Critical SND values and corresponding percentage values 79
Table 6.6 Required number of probe vehicles for each link 81
Table 6.7 Sample of JT_{23} 86
Table 6.8 Model performance 91
Table 6.9 Model performance summary 91
Table 7.1 One-step-ahead prediction error 98
Table 8.1 Fuzzy role matrix 121
Table 8.2 An example of fuzzy output 121
Table 8.3 Comparison of link 1 and link 2 estimation results 123
Table 8.4 Incident detection result on link 1 and link 2 127

Acknowledgements

I am deeply grateful to Professor Mike McDonald who supervised the work, for his guidance, help and support over the years; also, for providing an excellent research environment which enabled me to enjoy completing this work.

The work presented in this thesis has benefited greatly from the input and assistance of many individuals. Dr. Mark Brackstone was helpful in the use of the instrumented vehicle. Eamonn Kennard, Dr. Jinan Piao, Doug Robinson, and Dr. Jianping Wu volunteered in driving the instrumented vehicles to collect journey time data in very early morning. Thanks to Tim Felstead for help in data analysis.

I am grateful to my colleagues, John Armstrong, Jonny Crockett and Grant MacKinnon, who gave me support and encouragement throughout the study period.

Stephen Clark in Leeds City Council was helpful in processing the video camera data. Gary Gates in ITIS Holdings plc. provided Figure 2.1.

I would like to acknowledge the financial support of ORS Award and departmental scholarship of School of Civil Engineering and the Environment. I am also grateful to Tsinghua University who provided my masters study and research experiences in 1997-2000. I am extremely fortunate to have been exposed to so many great teachers.

Last and most importantly I am indebted to my parents who shared all my study experiences and supported me in all aspects.

Definitions of Symbols

DT_i = difference of journey time between the ith and the $(i-1)$th time interval.

e = permitted error of estimation

e_f = error of journey time estimation using fuzzy model

e_p = error of journey time prediction

f = ratio of probe vehicle journey time to average journey time

JT_i = journey time at the ith time interval

l = length of GPS update interval

L = length of a link

n = sample size of probe vehicles

n_f = sample size of probe vehicles represented by percentage of traffic flow

N_r = number of time interval of recovery stage during an incident period

q = traffic flow

r = correlation coefficient of determination

v_i = space mean speed at the ith time interval for a link

$v_{\min}$ = minimum speed in an incident period

x = easting coordinate in the National Grid system

y = northing coordinate in the National Grid system

ϕ = latitude

λ = longitude

μ = population mean

σ = population standard deviation

ρ_i = correlation coefficient of JT_i and DT_i

Abbreviations

AID	Automatic Incident detection
ANPR	Automatic Number Plate Recognition
AVI	Automatic Vehicle Identification
AVL	Automatic Vehicle Location
CA	Continuous Acceleration
CD	Continuous Deceleration
DGPS	Differential Global Positioning System
DR	Detection Rate
FAR	False Alarm Rate
HA	The Highways Agency
FHWA	The US Federal Highway Administration
GPS	Global Positioning System
JTMC	Journey time measurement system
MCA	Maximum value of Continuous Acceleration
MDA	Maximum value of Continuous Deceleration
MF	Membership Function of Fuzzy Variable
OS	Ordnance Survey
SA	Selective Availability
VRN	Vehicle Registration Number

Chapter One

Introduction

1.1 Background

Motorways were originally conceived and designed to provide continuous, free-flow, high-speed movement of traffic on limited-access facilities. Initially, little consideration was given to providing for the needs of traffic management and control systems to maintain a high level of mobility on these facilities. However, as traffic continued to grow, motorways became more congested. Today, it is recognised that the previous approach of constructing more motorway lane-miles to relieve congestion is often politically and socially unacceptable and economically infeasible (Carvell et al., 1997). Motorway management systems are a primary means of making the best possible use of the existing motorway network. Motorway management systems make use of control strategies, and operational activities such as information dissemination and incident management to reduce the occurrence of congestion and lessen its duration and extent.

In a motorway management system, a sound and timely database is key to performing management functions, such as measuring traffic conditions and making control decisions. Many technologies are available for collecting traffic data. Although inductive loop detectors are currently used most frequently, other technologies are beginning to replace loop detectors in many applications (Nelson, 2002). Table 1.1 provides summaries of the characteristics for some embedded and non-intrusive detectors. The cost estimation is mainly based on the Freeway Management Handbook (Carvell et al., 1997). However, because of the decreasing cost of wireless communication in the last few years, the cost estimation for Automatic Vehicle Location (AVL) is according to the PRELUDE project (Kroes et al., 1999).

Disadvantages of each detector based on current techniques and market penetration have been summarised by Suennen (2000) (Table 1.1).

Table 1.1 Summary of traffic detectors

Detector Type	Detector	Estimated Annual Operation and Maintenance cost	Main Disadvantages
Embedded	Inductive Loop	£150-£200 per station	Installation and maintenance require lane closure
Non-intrusive	Microwave Radar	£150-£200 per station	May lock on to the stronger signal (e.g. large truck)
	Infrared	£150-£200 per station	Operation affected by precipitation (e.g. rain, fog)
	Ultrasonic	£150-£200 per station	Performance affected by environmental conditions (e.g. temperature, humidity)
	Acoustic	£150-£200 per station	High level of special maintenance capacity is required
	Video Imaging	£350 per station	Performance affected by weather and light
Probe Vehicles	Automatic Vehicle Identification	£350 per station	Generally rely on automatic toll collection systems
	Global Positioning System	£150-£200 per vehicle	Insufficient sample size in the traffic stream

It can be found from Table 1.1 that Global Positioning System (GPS) equipped probe vehicles have the advantages of being cost effective and are not limited to specific data collection sites over other traffic detectors. The primary disadvantage of GPS equipped probe vehicles is the limited number of vehicles equipped with the device. A statistic in 2000 has shown GPS device prices have fallen by about 15 to 20 per cent a year (Barnes 2000). The decreasing cost has brought a continued growth of the global GPS market. As GPS locations become more commonplace in vehicles, the ability of GPS tracking to collect vehicle roadway data will continue to increase.

1.2 Objectives

This research aims at journey time estimation and incident detection using GPS-equipped probe vehicles. It has had the following objectives:

- Link journey time estimation: to calculate motorway link journey time from vehicle location data at various GPS sampling frequencies; to evaluate estimation accuracy and analyse factors which have a potential impact on accuracy.
- Sample size determination: to study the journey time distribution of individual vehicles under different traffic conditions; to determine the required minimum number of probe vehicles using statistical sampling theory.
- Incident detection: to develop a model using measures of probe vehicles for motorway incident detection; to evaluate the model using field incident data.
- Journey time prediction: to develop a model for predicting motorway link travel times under non-incident traffic condition at different prediction intervals; to develop a model for predicting duration of incident.

1.3 Approach

This research starts with a review of current applications of probe vehicles and the existing technologies of location and communication used with probe vehicles. In application, there are two types of vehicle techniques used in collecting traffic data. Active probe vehicles, named as test vehicle in the travel time data collection handbook (Turner et al., 1998), are specially equipped for traffic data collection purpose, whilst passive probe vehicles have the location and communication equipment installed for other purposes, e.g. navigation.

Although both active and passive probe vehicles have been studied in this research, the primary focus has been on passive probe vehicles. In passive probe vehicle research, the sample size of probe vehicles is a key issue. By analysing the journey time distribution of individual vehicles, the sample size in different traffic conditions has been obtained using statistical sampling principles. Since an incident can cause significant delays in journey times, an incident detection model is developed based on journey time estimates using required number of probe vehicles. Because there are not

enough probe vehicles in the traffic stream, journey time prediction is required over a short time interval when there is insufficient sample of probe vehicles. In research on active probe vehicle, GPS data is used to build detailed speed profiles. Therefore, more data in addition to journey time can be obtained. By analysing the speed profiles, journey time might be estimated by relatively fewer GPS equipped probe vehicles. Since an incident may influence some features of speed profiles, analyses of speed profiles are also applied to incident detection. The structure of the thesis is illustrated in Figure 1.1.

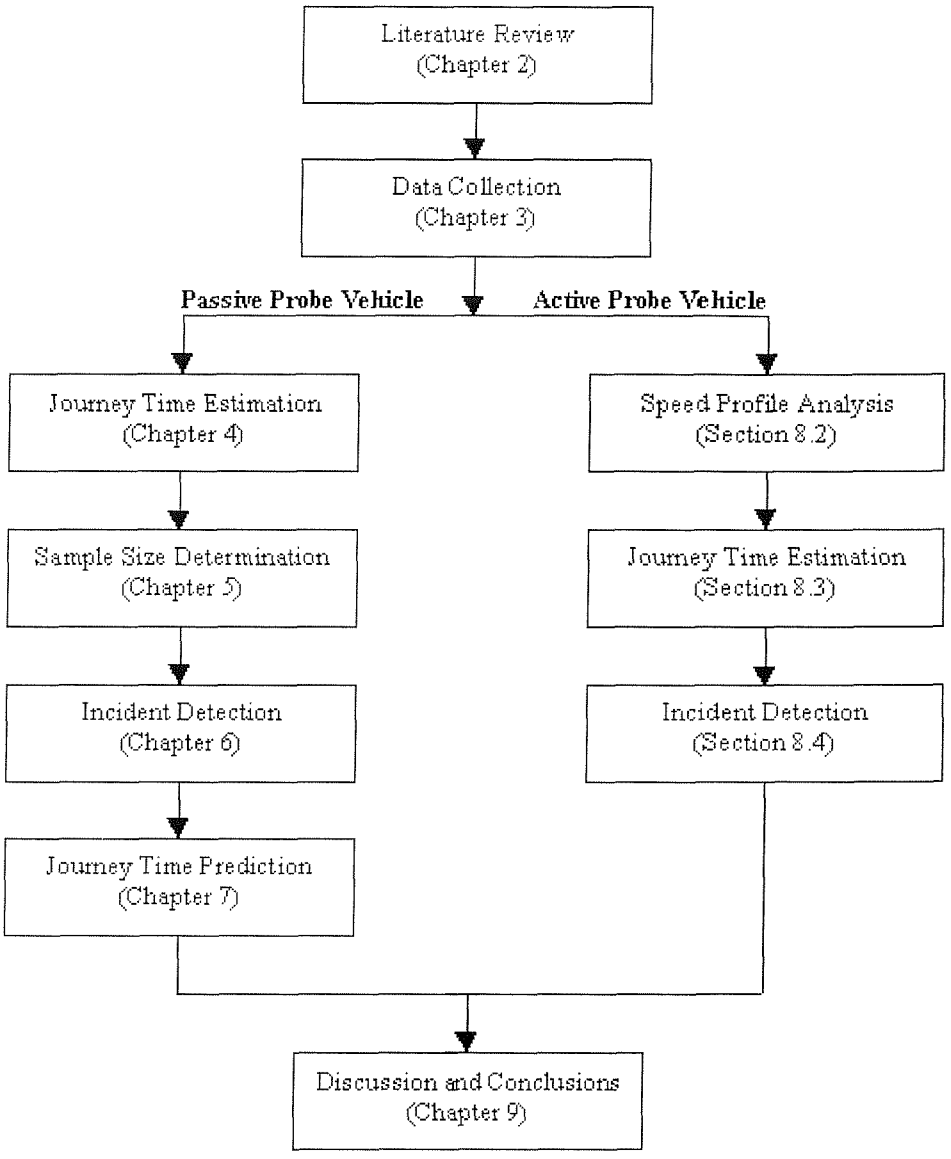


Figure 1.1 Structure of the thesis

Chapter Two

GPS Equipped Probe Vehicles

2.1 GPS technology

2.1.1 Overview of GPS

GPS is a US owned space-based system of satellites providing 24 hour, all weather 3D position, velocity and time all over the world. The full operational capacity of GPS was achieved in 1995 with 24 satellites uniformly distributed in six orbital planes, at an altitude of approximately 20,200km. This normal operational configuration ensures that at least four satellites are visible at any time and from any point on the earth's surface. Note that with the current constellation (i.e. 27 satellites and rising) at least seven satellites are visible (Ochieng and Sauer, 2002). Although civilians can access GPS signals free, the performance of GPS had in the past been limited by the artificial degradation of the signal through the process of Selective Availability (SA). The U.S. government switched off SA on the first May 2000 to encourage the acceptance and applications of GPS. This improves the positioning accuracy of GPS for civilian users from within 100 meters to within 20 meters for 95% of the time (US DoC, 2000). In many cases, real-world users find the accuracy to be even better.

GPS satellites transmit specially coded signals that can be decoded by a GPS receiver to determine time, position and velocity of the receiver. This is one-way broadcast system, so receivers do not transmit any signals back to the satellites. Since it is a passive system, GPS can support an unlimited number of users. Positioning measurement of GPS is based on the principle of time of arrival ranging. The time interval taken for a signal transmitted from a satellite at a known location to reach a GPS receiver is multiplied by the speed of the signals to obtain distance between the satellite and the receiver. There are three unknowns (X, Y, Z) to determine position of an object. Since the receiver clock is not so precise as satellite clock, the receiver clock

bias is considered as the fourth unknown. In a GPS receiver, signals from a minimum of four satellites are required to solve the four unknowns (Hofmann-Wellenhof et al., 1993).

In addition to positioning data, GPS also provides speed measurement. There are two methods to estimate speed by GPS. The first is to derive speed from differences in position. The second method is to use the Doppler effect. The Doppler shift in the frequency measures the relative velocity between the receiver and the satellite along the line between them. Velocity measurement using the Doppler effect is almost instantaneous and essentially independent of positioning data. Most GPS receivers products provide output of the Doppler speed and the measurement accuracy is 0.1m/s with 95% confidence, after the termination of SA (Garmin, 2003).

Although GPS has many advantages, it suffers from several weaknesses. Civilian GPS receivers have potential position errors primarily due to some of the following sources:

- Ionosphere and troposphere delays — The satellite signal slows as it passes through the atmosphere;
- Signal multi-path — Occurs when the GPS signal is reflected off objects such as tall building or large rock surface before it reaches the receiver. This increases the journey time of the signal, thereby causing errors;
- Receiver noise;
- Orbital errors.

2.1.2 Applications in transportation

Since positioning plays an essential role in transportation, GPS can be an effective tool in the transportation industry. On-board navigation may be the most visible use of GPS technology in transportation, and GPS based Automatic Vehicle Location (AVL) is also applied to a variety of areas, e.g. commercial fleet monitoring, public transport management and emergency response, etc. AVL combines a GPS receiver with an outbound communication link to provide real-time position data for persons and vehicles. The communication link can be accomplished through cellular phone networks, terrestrial radio or through a separate satellite communication network. Because AVL provides real-time location and status of vehicles, dispatchers can make

informed decisions. In addition, the real-time map display of AVL can allow dispatchers to help guide drivers through unfamiliar areas to reach their destinations faster, vital in an emergency response. With support for schedule adherence and delivery time prediction, GPS-based AVL can improve public transportation management and commercial fleet operations (Shrestha, 2003). A new approach for electronic road pricing (ERP) using GPS has been proposed. Different charges could be set depending on which roads the driver uses and the time of day. It could cost more to use congested routes and during peak conditions (Catling, 2000; Srinivasan et al. 2002).

2.2 Probe vehicle

The floating-car has been established as a method of collecting journey time data on roads since the late 1920s. Traditionally, this technique has involved the use of a vehicle within which an observer (passenger) records cumulative journey time at predefined checkpoints along a travel route. This information is then converted to journey time, speed, and delay for each segment along the survey route. The driver attempts to travel at the speed of the traffic stream and maintain the number of overtaking vehicles the same as those overtaken. In Drew (1968), the floating-car method was applied to evaluate level of service (LOS) by measuring acceleration noise. The rapid development of AVL and automatic vehicle identification (AVI) technologies has enabled continuous automatic traffic data collection. Faghri et al. (1999) compared the accuracy of journey time data collected by a GPS receiver with data collected manually. The conclusion proved that GPS was a more efficient and more accurate means of collecting data than manual records. When a vehicle is instrumented specially for traffic data collection, it is referred to as an "active" probe vehicle. Conversely, "passive" probe vehicles are vehicles that are already in the traffic stream for purposes other than data collection.

To provide real-time traffic data, both active and passive vehicles should maintain frequent communications with a central computer which tracks the vehicles along travelled route (Kroes et al., 1999). In this research, a probe vehicle is defined as "a vehicle, which is equipped with positioning and wireless communication systems

providing real-time traffic data by running on road network." Definitions of active and passive probe vehicles are summarised in the following sections with characteristics.

2.2.1 Active probe vehicle

An active probe vehicle is defined as a specially equipped vehicle for traffic data collection that actively gathers information through the monitoring of vehicle attributes. Currently, GPS is the common positioning technique used in active probe vehicles, as it provides accurate and detailed data over wide areas with relatively low installation and operation costs. In most travel time studies involving active probe vehicles, only one vehicle is used to characterise traffic flow along a pre-defined route. Therefore, errors can occur from human or equipment failures and adequate quality control is needed. In addition, detailed data collection (e.g. every second) can cause data storage difficulties (Clark and McKimm, 2003).

2.2.2 Passive probe vehicle

A passive probe vehicle is one where the location and communication equipment has been installed for another purpose, e.g. navigation. A passive probe vehicle provides journey time only between two points rather than detailed descriptions of vehicle movements over a journey. Passive probe vehicles allow continuous data collection with minimal human interaction. If the infrastructure is permanently installed, data are collected as long as the probe vehicles continue to travel through the system.

Probe vehicles can be buses, taxis, commercial vehicles or private cars. However, journey time estimates from some vehicles may be biased, for example, heavy transit vehicles may take much longer to travel over a road segment than an average value, if the traffic stream consists mainly of passenger cars. Buses may have priority by using bus lanes. The use of passive probe vehicles is increasing rapidly, for example, in Singapore, more than 10,000 taxis from major taxi companies are now used as probe vehicles for link journey time and speed estimation. The experience from the system is that unloaded taxis usually tend to cruise at a lower speed when looking for passengers (Xie et al., 2001). Hall et al (1999) evaluated the OCTA (Orange County Transit Authority) Transit Probe Project in California. The OCTA Transit Probe Project used

tracking data from GPS equipped buses for multiple purposes including bus schedule adherence and information on roadway traffic. One of the research issues was whether the tracking data could be effectively integrated into existing traffic management systems. The study found significant problems with the use of buses as probes primarily due to differences in the travel behaviour of buses and cars. Turner et al. (1998) indicated that a major problem of using private cars as probe vehicles relates to privacy issues, as in probe vehicle systems, motorists are monitored throughout their entire journey.

2.3 Applications of probe vehicle

2.3.1 Data collection techniques

Probe vehicles can be instrumented with different types of electronic equipment, but a common feature is the frequent reporting of vehicle location. Turner et al. (1998) provided a comparison of five probe vehicle data collection techniques:

- Beacon-based Automatic Vehicle Location (AVL)
- Automatic Vehicle Identification (AVI)
- Ground-based radio navigation
- GPS
- Cellular phone positioning

Beacon-based AVL systems use several beacons at known locations along the route. Beacon-based AVL systems are operated with bus systems in London and Southampton (Hounsell et al., 2000). In AVI systems, probe vehicles are equipped with electronic tags, which communicate with roadside transceivers to uniquely identify vehicles and collect journey times between transceivers. Both beacon-based AVL and AVI transmit information from a probe vehicle at fixed points to roadside devices, whilst the other techniques can locate a probe vehicle anywhere along a route and transmit its position and characteristics at a regular frequency. Ground-based radio navigation systems, such as LORAN C, Datatrak, and Omega, had been widely used as main navigation means in the past. However, due to vulnerabilities in cost and accuracy, they have been replaced by GPS completely, although they are still referred

in recent literature as available positioning techniques (Shrestha, 2003). Compared with other techniques, GPS has many strong points, e.g. high accuracy, good reliability and low cost of device. However, GPS provides only a positioning function and requires an additional communication system for transmitting real-time information. Many communication systems can be used in GPS equipped probe vehicles and the cellular phone is considered to be complementary to GPS. Cellular phones can also be used independently for both positioning and reporting. On GSM network, currently sub-50 metre accuracy has been delivered. For the third generation network (3G), an accuracy of sub-20 meters is expected, comparable with GPS (Bartlett and Morris, 2002). In Section 2.3.2, several previous research and projects involved probe vehicles using beacon-based AVL, AVI and cellular phone are briefly described and applications of GPS equipped probe vehicles are discussed later in the section.

Table 2.1 Comparison of probe vehicle data collection techniques

Technique	Costs				Data Accuracy	Constraints
	Capital	Installation	Data Collection	Data Reduction		
Beacon-Based AVL	High	High	Low	High	Low	Beacon density and placement, no. of probes
AVI	High	High	Low	Low	High	No. of probes and tag placement
Ground-Based Radio Navigation	Low	Low	Low	Low	Moderate	No. of probes and size of service area
Cellular phone	High	High	Low	Moderate	Low	No. of mobile users
GPS	Low	Low	Low	Moderate	High	No. of probes

Source: Turner et al., 1998

2.3.2 Previous experiences

2.3.2.1 Beacon-based Automatic Vehicle Location

The Road Traffic Adviser (RTA) project was based on applications of Dedicated Short Range Communications (DSRC) beacons, on a demonstration site running the length of the M25 and M4 motorways between the M23/M25 junction to the south of London, to Wales in the west (McDonald et. al., 2002). Operating at 5.8G Hz, the DSRC beacons allowed two-way data transfer between a moving vehicle and a stationary

beacon. This technology provides location specific information relevant to the road and direction being driven by the user and collects information about recent past performance on that road for both road management and driver information purposes. Over 70 DSRC beacons are deployed along the test site and linked to two network control centres, one in Wales and one in England. Within the RTA project, vehicles equipped with on-board units to communicate with beacons played the role of probe vehicles to measure journey time and speed, and in turn benefited from information collected by other probe vehicles. Within the RTA project, information has been passed regarding speed profiles of vehicles between beacons, hence overcoming spatial 'granularity' problems. However, there would be a delay in the provision of some types of information in some cases. For example, data indicating a queue several kilometres before a beacon may be out of date when it is transmitted, due to the probe vehicle itself being delayed in traffic (Koelbl et. al., 2002).

As illustrated in Table 2.1, the constraints for systems using beacon-based AVL are beacon placement and density, as well as sample size of probe vehicles. Through empirical experiment with the RTA project, Brackstone et. al (2001) found that at least 0.25% of the vehicle population using that road would need to be equipped to ensure that representative speed measurements are available. Implementing a system with beacon spacing below typical junction to junction distances may have minimal impact on our ability to formulate an accurate picture of the average speeds on such links.

2.3.2.2 Automatic Vehicle Identification

Many toll agencies in the U. S. and Europe (Nelson, 2003) are using automatic vehicle identification (AVI) technology for electronic toll collection (ETC). To validate ETC, each vehicle's presence not only has to be detected (identified), but also must be recognised as unique. Because an AVI-equipped probe vehicle is uniquely identified, its journey time between two roadside readers can be calculated. Some existing ETC agencies have begun to use AVI technology for journey time data collection in addition to toll processing. Some ETC systems experience between 25 percent to 100 percent of all tollway vehicles equipped with ETC on-board units. This provides a large sample to collect representative journey time data.

Houston was the first city to apply AVI technology for monitoring traffic condition (Houston TranStar, 2003). The Texas Department of Transportation has helped to develop the TranStar system in Houston which operates an AVI system in order to monitor traffic conditions, detect incidents, distribute travel information, and archive journey time data. Roadside reader units are being placed at 1.8 to 8.0 kilometre intervals along all major freeways in the Houston area, including over 483 kilometres of highway. Several toll roads in the area have automated toll booths, encouraging the acquisition of thousands of AVI transponders by motorists in the area. Over 200,000 ETC equipped vehicles have been distributed in the area (Turner et al., 1998).

In New York and New Jersey, the TRANSCOM agency operates the TRANSMIT system to monitor traffic conditions with AVI technology. Fifteen roadside readers have been deployed at 0.8 to 3.7 kilometres spacing on 29km of highway in the area (Chien and Kuchipudi, 2002).

2.3.2.3 Cellular phone

The mobile phone is the most popular public communication means today. Cellular phones have been used to report journey time manually (Balke et al, 1996), in which volunteer drivers call a central facility when they pass checkpoints along the route. Location based services in mobile networks is widely considered to be a growing opportunity. In the USA the Federal Communication Commission (FCC) has mandated the introduction of technology that will enable a caller's position to be pinpointed to better than 100m 67% of the time and better than 300m 95% of the time when an emergency call is made (Feng and Law, 2002). Europe has started looking at requirements for an equivalent system to pinpoint the location of emergency calls from mobile phones. New commercial applications are also emerging in which location provides an important ingredient to make the service more attractive to users.

Cellular phone positioning is a kind of radio positioning that uses the propagation characters of radio waves. Recently, several cellular phone location technologies have been developed. The commonly used cellular phone positioning methods are cell identification, signal level, angle of arrival, time of arrival, distance measurement, and phase measurement. Some of these methods have been implemented in trial systems

and some commercial products have already been introduced. However, no method that is superior to all other has been found yet (VTT, 2003).

Cambridge Positioning Systems Ltd (CPS) in the UK now offers sub-100m performance of mobile phone positioning on GSM networks with industry plan for sub-50m next year. On a mobile network, the positioning accuracy depends on the number of Base Stations (BS) are used in a location calculation. The more BS that can be measured the better the results. Therefore, better accuracy is achieved in urban areas with more BS than in rural areas. CPS tested that 90% of location estimates are within 100m from the true location in the city centre and within 200m in a suburban area (Bartlett and Morris, 2002). Ygnace, Drane and Yim (2000) estimated that with 80 millions cellular phones subscribers in the U. S., the percentage of cars travelling on major roads and motorway corridors with a phone switched in the "on" was high enough to give a good sample of the travelling population. However with current technologies and infrastructure, simulation results (Yim and Cayford, 2002) found that in probe vehicle application, current accuracy of cellular phone positioning could provide journey time information for only 68% of freeway segments.

Current mobile phone location technologies have been anticipated to achieve better accuracy on third-generation networks (3G). 3G networks are already a reality in many parts of the world. Japan launched the world's first commercial 3G networks in 2001, and similar networks are now operating commercially in Austria, Italy, Sweden and the UK with more launches anticipated during 2003-2004 (Dunne, 2003). CPS announced that CPS has developed high-accuracy location technology for 3G networks, which is based on principle of observed time of arrival and promises an accuracy of 10-20m, comparable with GPS. However, the accuracy will be only achieved in urban areas with sufficient Base Stations.

Despite the low accuracy of cellular positioning, cellular phone equipped probe vehicle systems have advantages, such as low establishment costs, two-way communication link and great number of users. Yim and Cayford (2002) suggested that widely deploying GPS in cellular phones may make probe vehicle methods more attractive and realistic. As discussed above, mobile phone solution can provide good accuracy in city centres bur worse accuracy in rural areas, while GPS has better performance in rural

areas than in cities. Hybrid solution using the complementary nature of both approaches to overcome situational weakness experienced by either mobile phone network or GPS working alone has been developed. Benefits of the hybrid solution (Feng and Law, 2002) include maximum availability, increased sensitivities and reduced handset cost and complexity. The hybrid solution is considered to be widely installed in future 3G handsets. Therefore, vehicles carrying 3G mobile phones will be potential probe vehicles to provide traffic information (QCT, 2003).

2.3.2.4 GPS equipped probe vehicle

Many researches have used active GPS equipped probe vehicles for journey time studies (Quiroge and Bullock, 1998) and measuring traffic system performance (D'Este et al, 1999). In such research, only one probe vehicle was involved, collecting positioning and speed data at short intervals, e.g. 1-s. The active probe vehicle method is based on the assumption that a probe vehicle represents a good average of the traffic stream. Average speed, running time, average speed, variation of speed have been calculated to measure traffic system performance and congestion. By repeatedly running the probe vehicle along the same route, journey time statistics for different time of day and different day were obtained (Clark and McKimm, 2003).

The earliest large scale application of probe vehicles was the ADVANCE (Advanced Driver and Vehicle Advisory Navigation ConcEpt) project in the northwest suburb of Chicago, Illinois (Sen et al., 1997). The main aim of ADVANCE system was to provide dynamic route guidance to vehicles in study areas. Vehicles involved were equipped with GPS-based navigation system and radio frequency modem for transmitting and receiving message. With two-way communication equipment, in-vehicle navigation systems could be used by the traffic information centre to locate each equipped vehicle across the network. The equipped vehicles were therefore used as "probes" to collect journey time information which was fused with inductive loop data to provide real-time estimates. The fused data has been proven to be more efficient in journey time estimation and prediction, as well as incident detection. Although probe vehicles have some advantages over other technologies, probe vehicles provide only speed and journey time, which may be not sufficient for traffic information systems. Data fusion that combines information from multiple technologies, including fixed sensors and probe vehicles may be a good solution to meet all requirements of various

users. Probe vehicle data can be fused with existing detectors, such as MIDAS loops spaced at 500m on the M1 between junction 9 and 19 (McDonald et al., 2000). The data fusion will provide more sound traffic information more than any source working alone.

The PRELUDE project (Kroes et al., 1999) in Rotterdam in 1997-1999 aimed to pilot the use of floating car data, collected using GPS device to provide historical and real-time information on the Dutch road network. Differential GPS devices on each probe vehicle determined location and speed every 10 seconds and the resulting information was stored together with time-stamp in an on-board computer. GSM telecommunication technology was used to transmit accumulated series of recorded time and location data to a central computer every five minutes. This information was then used to update a demonstration system on a central computer which displayed journey time in almost real-time.

In the UK, ITIS Holdings Plc has implemented the largest Floating Vehicle Data (FVD) system in the world (Cowan and Gates, 2002) to provide journey time statistics and real time traffic management information. The FVD system has been collecting and storing traffic data since February 2000 with initially only a limited number of probe vehicles on the network. Now the system has in excess of 30,000 probe vehicles of various characteristics contributing to the gathering of data on live and historical traffic conditions. Probe vehicles equipped with GPS and GSM technology regularly send data on their current position and speed. The Floating Vehicles include commercial trucks, National Express coaches and passenger cars. The information is collected and centrally analysed, then transmitted to subscribers of the services in the form of up-to-date traffic information.

The ITIS FVD system is the UK's first commercial application of the probe vehicle concept. Commercial provision of the data gathered by FVD commenced during 2002 and is now providing an ever-growing source of revenue with new customers and uses of FVD being identified (Simmons et al., 2002). FVD captures data for motorways, urban motorways, A roads and some B roads of the UK road network. These data are aggregated into "road timetable" and "congestion schedule" by road and day/time category. The FVD system identifies recurring congestion and uses these patterns to

predict the future, thus enhancing route planning and navigation. ITIS FVD system delivers dynamic traffic content and integrates this with navigation systems. ITIS have installed Data Collection Units across the entire National Express coaches as part of the strategy for acquiring FVD. National Express is also able to use this to monitor the reliability of its services through the use of historical data and to provide better customer information services. Control operators of National Express can locate the exact position of coaches and the traffic conditions they are driving in.

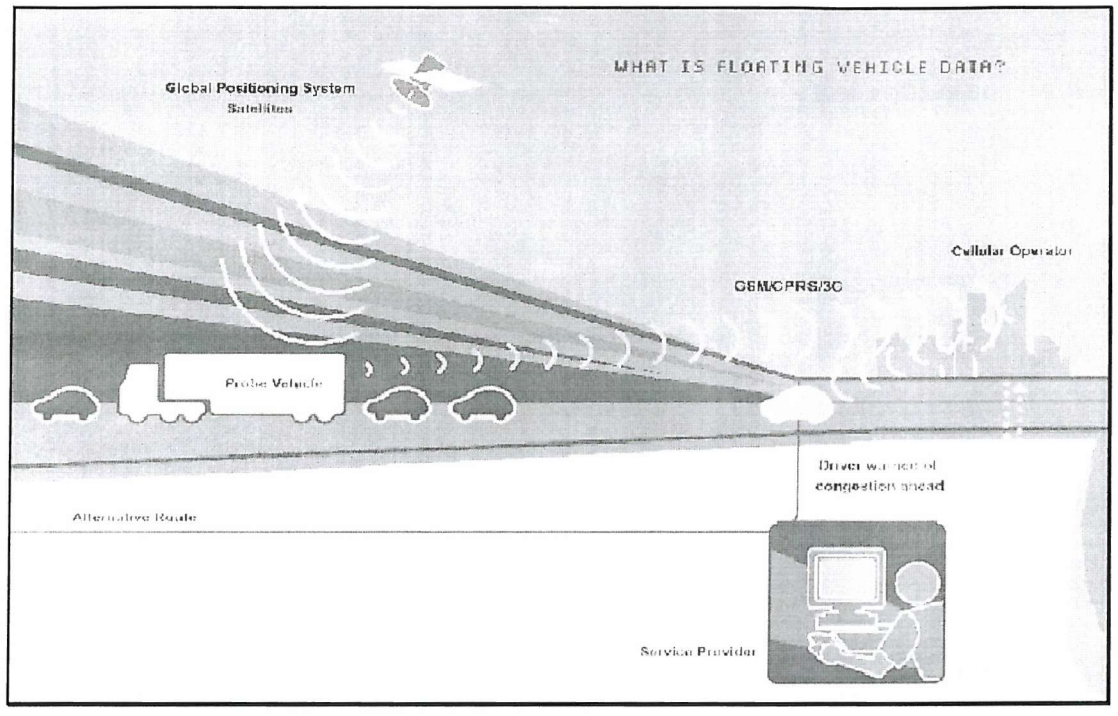


Figure 2.1 The Floating Vehicle Data System
(Source: Cowan and Gates 2002)

2.4 Summary

Probe vehicle has long been considered as an extremely cost effective means of monitoring journey times when compared with the alternative of installing fixed detectors. The concept is based on reporting location and speed information from vehicles travelling the road network to a central information operator. Two types of vehicle can be used in a 'probe' capacity to collect traffic data. There are: specially equipped vehicle that actively gather information through the monitoring of vehicle movement (active probe vehicle), and public transport vehicle, goods fleets or general cars that have a passive role (passive probe vehicle).

After discontinuing the use of Selective Availability (SA), GPS provides accurate position and speed measurement. Currently, active probe vehicle uses GPS to collect both position and speed data while passive probe vehicle can use many technologies, e.g. automatic vehicle identification (AVI), beacon-based automatic vehicle location (AVL), cellular phone positioning and GPS technology. Although each technology has its own advantages and disadvantages, GPS is superior to the others with low installation and operation costs, good accuracy and unlimited coverage area. In addition, GPS provides detailed data continuously along the entire route. Since GPS is becoming increasing available as a consumer product for in vehicle navigation and monitoring, large samples of passive probe vehicles can be obtained. Large scale applications of GPS equipped probe vehicles in the ADVANCE project have shown encouraging results for journey time estimation and incident detection. A number of commercial services based on GPS/GSM equipped probe vehicles are already operating in the UK. The ITIS Holding Plc uses in excess of 30,000 commercial and other vehicles as probes for the gathering of live and historical traffic data for UK's major roads and motorway corridors.

This PhD research focuses on using GPS-equipped probe vehicles for journey time estimation and incident detection. Both active and passive probe vehicles are studied with main attention on passive probe vehicle. Among following chapters, Chapter 4, 5 and 6 are based on passive probe vehicle and Chapter 7 studies active probe vehicle. Journey time prediction discussed in Chapter 8 is based on the two types of probe vehicles. Since the main part of this research focuses on passive probe vehicles, in the following discussions, probe vehicles denote only passive probe vehicles while active probe vehicles will be specially referred.

Chapter Three

Data Collection

3.1 Introduction

Data from motorways have been used in this research: the M27 between Southampton and Portsmouth, and the M3 between Southampton and Winchester. Details of the survey sites are given in the following Section 3.2. Three different types of data were used: loop data, video camera data, and probe vehicle data. Loops and video cameras were installed by the Highways Agency (HA) for daily traffic control and management. Loops provided traffic volume data such as vehicle count per minute. Video cameras enabled journey times of individual vehicles to be obtained using Automatic Number Plate Recognition (ANPR) techniques. GPS equipped vehicles were driven through the survey sites as probe vehicles to measure journey times. The three data sources have been used to study journey time calculation methods, assess measurement accuracy, and determine sample size requirements, etc. In addition, incident reports were provided by the HA and the Hampshire ROMANSE (Road MANagement System for Europe) Office. Details of the incident data are described in Chapter 6 (Incident Detection).

3.2 Survey site

In 1999, a 50-camera real-time journey time measurement system was supplied by Initial Systems Ltd for use by the HA. This system is being used as part of the Ramp Metering Trial conducted on the M3 (Junction 11 to 14) and M27 (Junction 2 to 11) motorways in the UK (Adaway, 2001). Ramp metering controls the rate at which traffic joins a motorway from the slip roads according to the traffic flow on both slip road and the main carriageway. Data gathered by the loops and cameras have been

used to assess the operational performance of the ramp metering system (Gould et al., 2002). The loops provided traffic volume data for 24 hours per day, while the cameras provided individual journey times for the morning peak hours of 6:00-9:30 and the evening peak hours of 15:00-19:30, on working days throughout a year. Seven links are covered by ANPR cameras as shown in Figure 3.1. The seven links have different lengths and geometric characteristics as well as different levels of traffic flow, enabling probe vehicle applications to be compared for different links. The survey site map with locations of video cameras is shown in Figure 3.1, and the length and average speed of each link are shown in Table 3.1.

Table 3.1 Link length and average speed

Link Number	Length (km)	AM Peak Hour Average Speed (km/h)	PM Peak Hour Average Speed (km/h)
1	1.44	96.5	88.9
2	2.21	75.3	89.5
3	4.51	67.2	79.3
4	3.40	94.1	95.6
5	3.56	73.5	94.9
6	3.31	97.5	99.9
7	3.72	93.7	100.7

3.3 Video camera data

3.3.1 ANPR technology

Automatic Number Plate Recognition or ANPR, also known as Automatic License Plate Recognition or (ALPR), has been available for a number of years. Recent advances in image processing techniques combined with the advent of low cost high performance computing devices have led to the development of several journey time measurement systems (JTMS). Typical systems are in UK motorway networks and ring road around Helsinki in Finland (Ellis, 2002; Frith and Pearce, 2002; Eloranta et al., 2000).

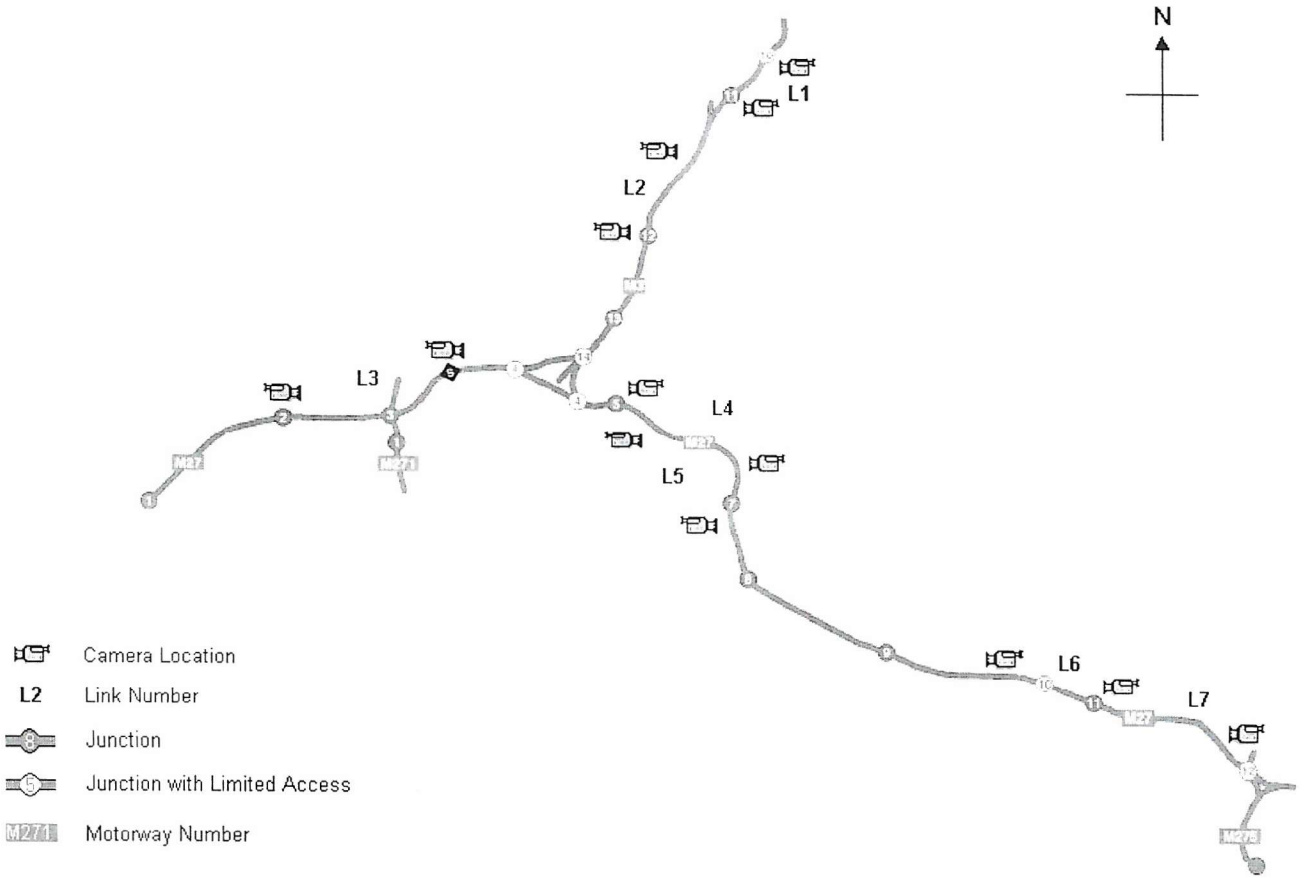


Figure 3.1 Survey site map with locations of video cameras

ANPR is an image processing task, which incorporates the following processes (Bunn and Barrett, 1997):

- detection of a vehicle in the scene
- detection of the number plate on the vehicle
- recognition of the individual characters within the plate area
- checking rules relating to the number plate format
- reporting result with confidence scores and time.

In a JTMS, video cameras are set at origin and destination locations, and cameras are located above each lane of traffic in one direction only. Full vehicle registration numbers are recorded and recognised by ANPR software. If a vehicle is recorded by an 'origin' camera, it is checked against corresponding records by 'destination' cameras. If a match is found then the vehicle journey time is recorded in the database.

ANPR cameras cannot capture all passing vehicles for many reasons, such as dirty or damaged license plates. Severe weather may also affect ANPR performance with snow, water spray, low sun angles and so forth. Instances of poor capture occur shortly after the onset of rain. Fast travelling vehicles throw out a curtain of spray, this spray then obscures the number plate of the following vehicle. Not until the road surface has dried out sufficiently does capture rate recover. Because of dry weather in summer, data quality in summer is generally better than in winter. An average recognition rate of 85% to 95% for unobstructed license plates in real conditions is reported in literature, and it may reach 97% in ideal circumstances (Wiggins, 1999; Bibaritsch and Egeler, 2002). In this research, cameras at a particular site usually recognised about 60% of license plates by comparing with traffic counts provided by inductive loop detectors. To collect point-to-point journey time of a vehicle, the vehicle should be captured by both 'origin' and 'destination' cameras. Assume recognition rate of each camera independent, capture rate of vehicle journey times was about 36% (60%×60%). However, a certain group of vehicles with dirty or non-standard license plates was missed by all cameras. The proportion of such vehicles was about 5%, that slightly increased the capture rate of vehicle journey times.

ANPR data for the eighteen weeks in 2001 and 2002 listed in Table 3.2, were chosen for use in this study. Average capture rate of vehicle journey times was about 40%, i.e. journey times of about 40% of all vehicles on a link were recorded. The minimum capture rate of vehicle journey times used in this research was 23%.

Journey times logged by ANPR cameras may include outliers, and the outlier data have been checked and removed where appropriate. Outlier detection and cleaning is dealt with in Section 3.5.

Table 3.2 Video camera data weeks

Year 2001	Year 2002
21-27May	21-27Jan
04-10June	28Jan-03Feb
11-17June	05-11Aug
18-24June	19-25Aug
25June-01July	26Aug-01Sept
02-08July	02-08Sept
09-15July	16-22Sept
16-22July	30Sept-06Oct
23-29July	07-13Oct

3.3.2 ANPR data component

An example of a subset of the ANPR data is shown in Table 3.3 (the license plates have been modified to preserve anonymity), and the format can be summarised as follows:

- License plate
- Location of start cameras
- Location of end cameras
- Journey time
- Date
- Time at start cameras

Table 3.3 An example of individual vehicle journey time data

VRN	Location 1	Location 2	Journey Time (hh:mm:ss)	Date	Time
M212LJT	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:34	19/06/01	08:07:05
R838GGH	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:17	19/06/01	08:07:07
L583LVS	M27 J3 8/3A	M27 J3/J4 14/0A	00:14:14	19/06/01	08:07:07
D254WDP	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:29	19/06/01	08:07:10
R932UOR	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:37	19/06/01	08:07:10
P47AJT	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:22	19/06/01	08:07:12
R976NRW	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:36	19/06/01	08:07:13
F744LRR	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:28	19/06/01	08:07:14
K324EWK	M27 J3 8/3A	M27 J3/J4 14/0A	00:02:21	19/06/01	08:07:16

Average journey times for a time interval on a link could be calculated from individual journey times. For example, average journey time on a link for period of 9:00-9:05 is obtained from the mean of journey times of all captured vehicles that entered the link from 9:00 to 9:05. In every morning peak hour of 6:00-9:30, using five-minute interval, there are a total of 42 observations of average journey time, denoted by JT_i, $i = 1,2,...,42$, where JT_1 denotes the average journey time of 6:00-6:05 on a given link, and JT_{42} is denotes the average journey time of 9:25-9:30. Average journey time is therefore considered to be a discrete time-series. For purposes of journey time prediction and incident detection, time series models are developed for analysis and forecasting.

3.3.3 ANPR data applications

ANPR data play a very important role in this research. The data are used for many purposes, including:

1. Providing "real" journey times of probe vehicles.

In this research, GPS-equipped probe vehicles were used to collect journey time data by travelling on the survey site. When the probe vehicles were running on the site, registration numbers of the probe vehicles were sometimes recorded in the ANPR database. The journey time observations by ANPR were compared with journey times calculated from GPS data to assess accuracy of journey time estimation.

2. Identifying distribution of individual vehicles' journey times.

One important issue for probe vehicle applications is sample size, e.g. minimum number of probe vehicles required for any particular estimate of accuracy. To determine the minimum number of probe vehicle, it is necessary to study the distribution of individual journey times. All vehicles on each link during a period can be considered as the population being surveyed, the vehicles captured by video cameras therefore can be considered as samples. As introduced in Section 3.3.1, the minimum capture rate of vehicle journey times used in this study was 23%, i.e. journey times of at least 23% of all vehicles on a link were recorded by ANPR cameras. Therefore, a sample size of more than 23% was achieved and the sampling was used to identify the distributions of vehicle journey times as well as other statistical characteristics.

3. Providing average journey time and developing historical database.

Statistically, the larger the sample size, the more likely the sample mean will be an accurate representation of the mean of the whole data population. Since the capture rate of ANPR was far beyond potential sample size of probe vehicles, the average journey times provided by ANPR were considered as "real" average journey times'. The "real" average journey times were used to analyse impacts of sample size of probe vehicles and develop historical journey time database. Three weeks' data were selected to develop an historical journey time database: 21-25 May 2001, 2-9 Sept. and 16 -22 Sept 2002. This was used for the development of incident detection model and the prediction of journey time.

4. Providing simulated probe vehicles.

While several vehicles were instrumented and used for data collection, the probe vehicle database was relatively small. Therefore, a number of individual vehicles recorded by the cameras were randomly selected to be simulated probe vehicles. Mean journey times of these selected "probe vehicles" were compared with the "real" average journey times to study sample size of probe vehicles required and used as input to the journey time prediction and incident detection models.

3.4 Loop data

3.4.1 Loop data component and reliability

In this research, loops provided traffic volume data as vehicle count on each lane in one minute. The traffic volume data were used to study the relationship of required sample size of probe vehicles and traffic volume. Loop detector is mature technology and the count accuracy is always within 1 to 2 per cent of relative error compared with manual counts (Middleton and Parker, 2000). Loops recorded data for 24 hours of day and data included:

- Date
- Site (loop number)
- Time (1-minute commencing)
- Vehicle count for vehicle length 1
- Vehicle count for vehicle length 2
- Vehicle count for vehicle length 3
- Vehicle count for vehicle length 4
- Total vehicle count
- Average speed for each lane in km/h
- Vehicle count for each lane
- Average occupancy in percent for each lane
- Average headway (tenths of a second) for lane 1

There were records of zero for a number of minutes in the loop data files. Loop data of zero could be due to no traffic or loop faults. For missing data caused by loop faults, replacement is required. It is therefore necessary to distinguish between no traffic and missing data. In this research, since only loops close to a group of ANPR cameras were used to provide vehicle counts, ANPR data were used to distinguish between no traffic and no data. For example, loop data was zero on the entrance of link 7, at 7:48, 9th July 2001. At the same time, ANPR cameras recorded 46 vehicles entering the link in the one minute interval of 7:48-7:49. Therefore, missing data was supposed rather than no traffic for the whole minute. In the loop data file, the failure rate was very low, less than 1 %. Missing data can be replaced by estimation from adjacent minutes (Abou-Rahme et al., 2002).

However, zero record could also occur when traffic comes to a halt. Replacement of such gaps leads to overestimation of traffic volume (McDonald et al., 2000). In this case, very low speeds should be observed in adjacent minutes, e.g. lower than 10 km/h. If the average speed at the previous minute is lower than 10 km/h, zero loop data are considered as traffic stop and no replacement is needed. However, this situation was not found in the loop data file. Of data collected, very small but non-zero counts occurred in severe congestion, e.g. 6 vehicles/minute.

3.4.2 Missing data replacement

Missing data can be replaced by estimation from adjacent minutes. Up to three consecutive missing minutes were found in the loop data files. The missing data were replaced by the following steps:

- For each loop, look through the data, minute by minute and identify where there are gaps.
- For each gap, recognise whether the gap is missing data or no traffic.
- For missing data, see where there is available data: either the previous minute or the next minute, or both sides. If data are available at one side only, use that data to replace the minute gap. If data are available at both sides, use the average of the data either side.
- If no data is available either side, first replace vehicle counts for the previous cases, then perform another iteration to see if, once data have been replaced, there are some data near to the missing minute and fills it.

3.5 Probe vehicle data

3.5.1 Main surveys

Three surveys involving GPS equipped vehicles were conducted on selected links of the survey site. However, of the three survey databases, only one was specially collected for this probe vehicle research. The original purposes of the other two surveys were to study impacts of ramp metering, and the GPS application for automatic car following.

Survey 1: Ramp metering survey

To study merging behaviour and impact of ramp metering, a survey was carried out in a period of two months, from May 21 to July 17, 2001 on the M27 Junction 11. A combination of instrumented vehicle and camera technology was used to observe the whole process of merging. For complete details of the survey, the reader should refer to Zheng (2002).

During the survey period, the TRG instrumented vehicle (IV) was driven by different drivers from Junction 10 to Junction 12 on the M27, on weekday morning peak hours from 7 am to 9 am. Equipped with a GPS receiver, the IV was used as a probe vehicle to collect journey time data. In total, 105 valid observations of journey time on link 7 were obtained. Since the instrumented vehicle was driven by different drivers every morning, the impact of driver behaviour on journey times can be studied (Brackstone et al., 2002).

Although the survey provided sufficient journey time data measured by GPS, the journey times collected do not represent the real traffic conditions. Since the purpose of the survey was to study merging behaviour, the IV drivers were required to drive the vehicle on the motorway lane 1 for as long as possible. Therefore, most of the journey time observations were longer than the average.

Survey 2: Car Following Survey

Automated car following is the core technology for various intelligent transport systems, e.g. autonomous cruise control and convoy driving. The key parameters required for automated car following are the separation and relative speeds of successive vehicles. Most current automated car following systems rely on front facing radar systems. TRG has studied the potential of using GPS in car following system. During 15:30-17:30 on the 26th and the 27th of July 2001, six cars were driven on the M27. The six cars departed in different intervals from 1 to 120 seconds each cycle and ran 3 cycles from Junction 8 to Junction 12 of the M27 each day (McDonald et. al,. 2002). The data are used to study difference of journey times collected by different vehicles under the same traffic conditions.

Survey 3: Probe Vehicle Survey

Data of the two surveys discussed above were collected on the M27. A new survey was designed and conducted on the M3 to study probe vehicle application on different motorways. During morning peak hours of 7:00-9:00 from the 1st to the 11th October 2002, on each weekday morning, a GPS equipped probe vehicle was driven by different drivers from Junction 10 to Junction 14 of the M3. 84 observations of journey time were obtained: 42 journey times from J12 to J11 northbound, and 42 journey times from J11 to J12 southbound. The data are used to assess accuracy of journey time estimation, study daily change of journey time and identify recurrent traffic congestion.

3.5.2 GPS Data component

GPS receivers used in the above surveys were Garmin 35 receivers in stand-alone mode. The Garmin GPS 35 is a low-cost GPS receiver (current price about £100) and made for navigation purpose. A Garmin 35 receiver can track up to 12 satellites and the update rate is 1Hz (Garmin 2003). The main output of a Garmin 35 GPS receiver is:

- UTC (Universal Time Coordinated): data and time of day;
- Position data: latitude, longitude and height;
- Speed data: speed over ground, three-dimensional velocities: east, north and up;
- Estimated error information: estimated horizontal and vertical position errors;
- Satellites in view: total number of satellites in view and total number of satellites to be used for positioning.

3.5.3 GPS data conversion

GPS position data are based on an ellipsoidal reference system, WGS-84 (World Geodetic System 1984), and expressed by latitude and longitude. For journey time estimation, a vehicle should be located on a plane map using a simple 2-D Cartesian coordinate system in which the two axes are known as eastings and northings. The coordinates of a point on the plane map can be converted from its ellipsoidal latitude and longitude by a standard formula known as a *map projection*, which is a function to convert ellipsoidal coordinates to plane coordinates. The projection can be expressed as:

$$[y, x] = f_{projection}(\phi, \lambda) \qquad (3.1)$$

where x and y denote the easting and northing on the plane map, ϕ and λ denote latitude and longitude. Ordnance Survey (OS) maps use a type of projection known as the Transverse Mercator (TM). The same type of projection is used in a worldwide mapping standard known as Universal Transverse Mercator (UTM). The TM projection can be thought of as a sheet of paper carrying the mapping grid (of eastings and northings), which is curved so as to touch the ellipsoid along a certain line. The line of contact is chosen to be north-south central meridian. Points on the ellipsoid are projected onto the curved sheet, giving easting and northing coordinates for each point. In different plane coordinate systems, different parameters are used in the TM projection.

In this research, a map with the National Grid coordinate system was selected. Thus, the National Grid TM was used to convert latitude and longitude data surveyed by GPS (Ordnance Survey 2000). The TM projection for Ordnance Survey maps has a central meridian at longitude 2°West and latitude 49°North. The two lines of true scale are about 180 km to the east and west of the central meridian. The stated scale of an Ordnance Survey map is only true on these lines of true scale, but the scale error elsewhere is quite small. For instance, the true scale of Ordnance Survey 1:50,000 scale map sheets is actually between 1:49 980 and 1:50 025 depending on easting. The equations and parameters of the National Grid TM projection are given in Appendix A.

Using equation 3.1, latitude and longitude data are transferred to Cartesian coordinate (x,y), which is compatible with OS National Grid maps. After coordinate conversions, journeys of GPS equipped vehicles can be displayed on a OS National Grid map. Figure 3.2 shows the paths of eight journeys of Survey 3 recorded by the GPS device. The route is clearly visible. However, some data are displayed on the wrong side of the roads. For journey time estimation, because lane differentiation is not required, the current GPS accuracy without DGPS correlation is sufficient.

3.6 Cleaning of ANPR data

The availability of efficient and accurate ANPR technology has provided a valuable source of data for traffic engineers (Eloranta et al., 2000; Ellis, 2002). ANPR data have been assumed to present the real journey times and used as a reliable reference source.

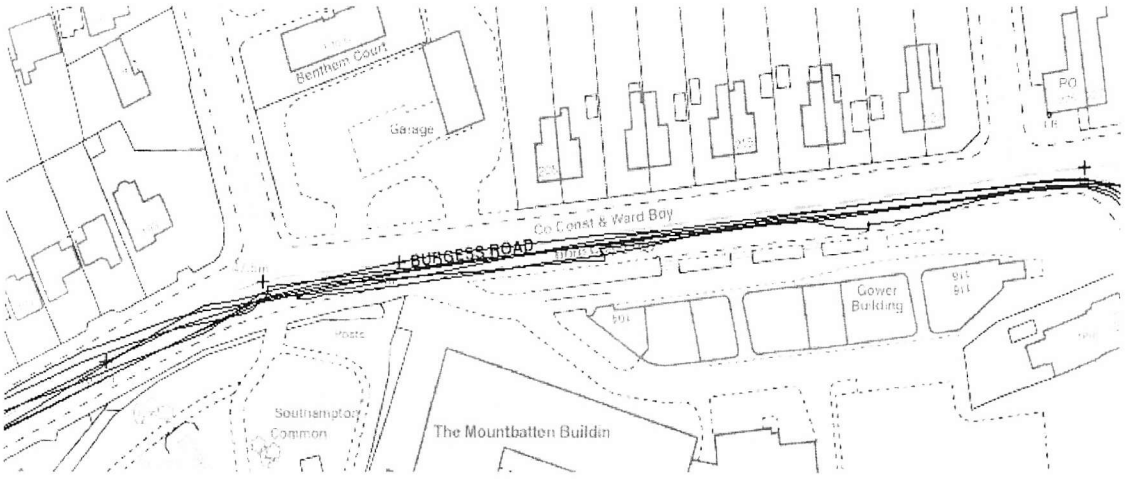


Figure 3.2 The paths of the journeys as recorded by the GPS device

For example, ANPR data have been used to validate the MIDAS journey time estimation algorithm (Abou-Rahme et al, 2002). However, journey times logged by ANPR are not 100% accurate. For example, a journey of the TRG instrumented vehicle (registration number L583 LVS) from the M27 J11 to J12 on June 19 was logged by the video cameras. Journey time of the vehicle measured by the video cameras was 14 mins 14 seconds (854 seconds), but the real journey time was only 3 mins 30 seconds (210 seconds). Among all journey time records, some extremely long journey times have been found. An example plot of the journey times in link 3 for a Thursday in June 2001 is shown in Figure 3.3.

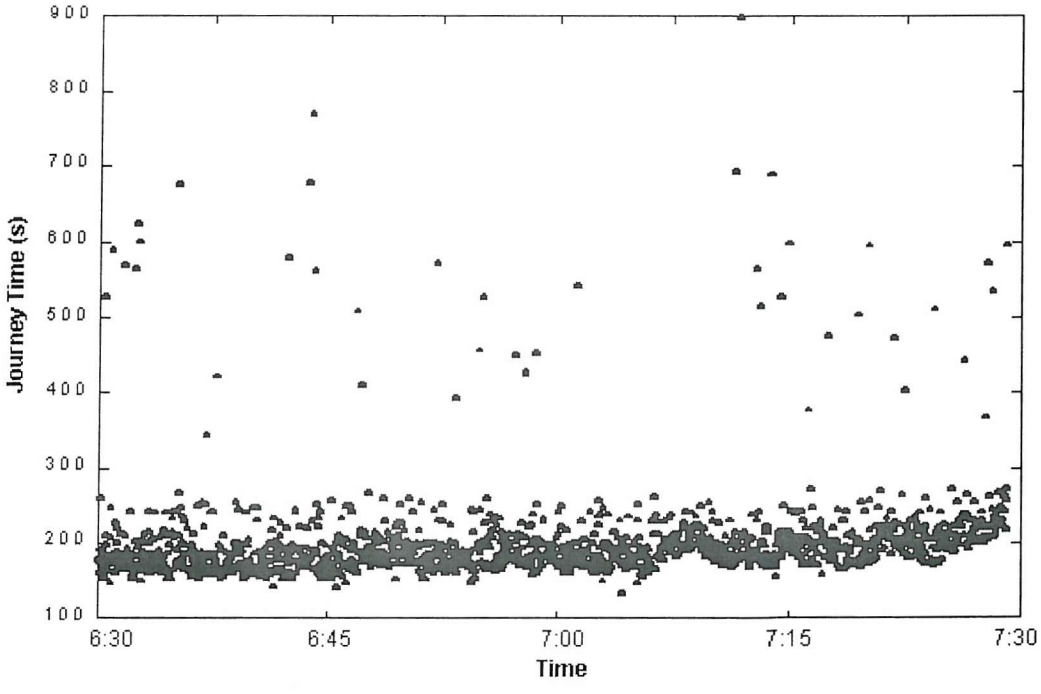


Figure 3.3 Example of raw journey time data on link 3: Thursday, June 12, 2001

The bulk of the data are within a band of 150 to 260 s (108 km/h to 62 km/h), but there are some journey times substantially longer. Short journey times are also considered. The shortest journey time in Figure 3.3 is 133 seconds, which corresponds to mean space speed of 122 km/h (75.8mph). In the probe vehicle surveys, the highest mean space speed achieved by the IV was 126 km/h (78.7 mph) in the morning peak hour. Therefore, the journey time of 133 seconds is considered to be feasible. In this circumstance, outliers were considered to occur only in the upper extreme (long journey time), rather than in both extremes. The outliers can occur for a number of reasons, including transcription errors, mismatching and short diversions (Bunn and Barrett, 1997). In the following section, quality of ANPR journey time data will be assessed and reasons of outlier occurrence will be discussed. An outlier detection method is developed in Section 3.6.2 towards the reasons of outlier occurrence.

3.6.1 Assessment of ANPR data quality

Since individual observations of journey time will be used to identify journey time distribution, it is necessary to assess quality of ANPR data before any analysis is undertaken. GPS logged journey times are therefore used as a reliable reference to assess the quality of ANPR data. Furthermore, ANPR data are also used to assess accuracy of journey time estimated by GPS. During the ramp metering survey, the ANPR cameras recorded 84 of the 189 journeys made by the TRG instrumented vehicle, i.e. 44%. There were 5 extremely long journey times in all the 84 records, as shown in Table 3.4. For the other 79 records, GPS data and ANPR data are very similar, and the maximum difference is only 5 seconds.

Table 3.4 Outliers of ANPR journey time

Date	GPS data		ANPR data	
	Time at start point	Journey Time (s)	Time at start point	Journey Time (s)
24/05/01	08:08:35	170	08:08:33	803
19/06/01	08:07:08	207	08:07:07	854
19/06/01	08:33:16	160	08:33:15	897
29/06/01	08:31:48	179	08:31:47	822
12/07/01	07:40:15	169	07:40:13	870

Short diversion has been accounted to be the main reason to produce outliers of large journey times by Clark et al. (2002). By studying all the above records, the outliers with large journey times were observed during the ramp metering survey. In the ramp survey (Wu et al., 2002), two test courses were performed to collect required data with the TRG instrumented vehicle (IV): mainline route survey and merging route survey. The IV first jointed the eastbound M27 at J10 and drove along the M27 down to J12 and returned to J11 along the westbound M27 (mainline route). Then, the IV joined the eastbound M27 from slip road at J11 and drove along the eastbound M27 down to J12 (merging route). For example, on June 19 2001, the IV passed J11 on the mainline route (i.e. from J10 to J12) at 8:07:09, which was recorded by ANPR cameras on J11. However, ANPR cameras on J12 did not capture the time when the IV passed the cameras. The IV returned to Junction 11 along westbound and rejoined the eastbound from slip road at 8:17:45, which could not be recorded by ANPR cameras since the cameras were only able to capture vehicles on main carriageways. The ANPR cameras on J12 captured the IV when it passed the cameras at 8:21:24. Therefore, a match was found by ANPR and a journey time was obtained 854 s, similar to 855 s, i.e. the difference between 8:21:24 and 8:07:09. All the extremely long journey times shown in Table 3.4 were obtained because of the same reason. Thus, this research considered that outliers in database were caused by short diversion and other error sources were ignored.

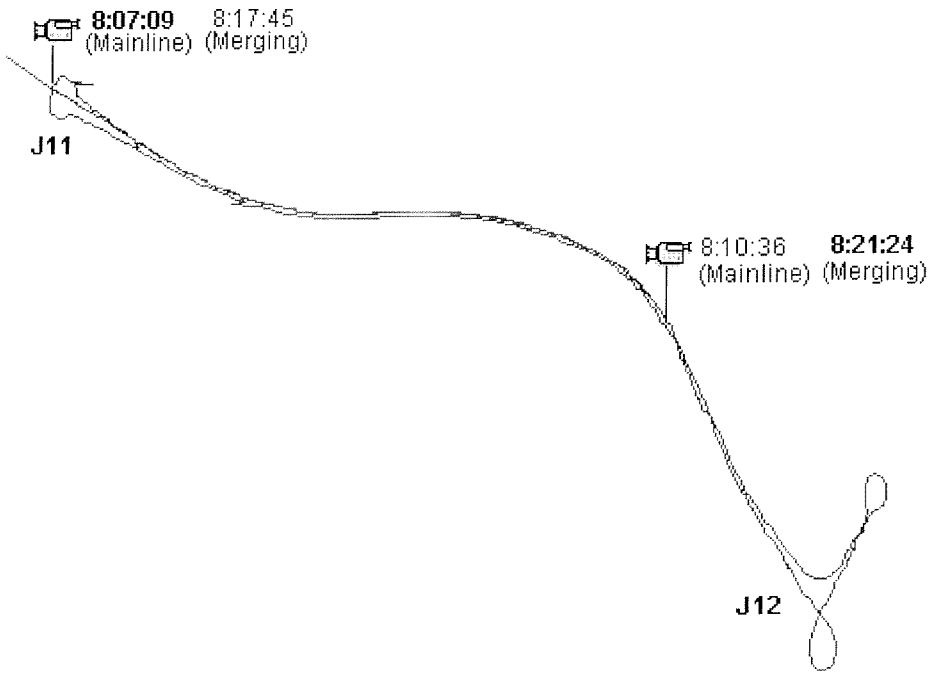


Figure 3.4 The path of the journey as recorded by the GPS device

In the ANPR database, very few outliers occurred on link 1, link 2, link 6 and link 7, while about 1-2 per cent of extremely long journey times were observed on link 3, link 4 and link 5. On average, about 1.38% of all raw journey times on the three links may be considered as outliers. On the three links, there are some exits and entrances between the origin and destination. It could be supposed that a lot of people use the M27 for short and utility trips, and people may leave at one exit for a while and then rejoin the motorway.

Two conclusions can be drawn from the above results:

- ANPR provided reliable journey time data.
- Outliers with large journey time occurred mainly because of short diversions.

3.6.2 Outlier detection

By considering vehicle journey times logged by the ANPR cameras in a five minute interval as a sample, very long journey time are considered as outliers of the sample. In the ramp metering survey, drivers were required to drive on motorway lane 1 as long as possible, the journey times obtained may be the longest on the link for the period. The data were used to develop and validate outlier detection methods.

3.6.2.1 Previous researches

Three statistical methods have been developed for cleaning of matched license plate data by Clark et al. (2002):

- Percentile test: to define as an outlier all observations that fall outside a range set by the 10^{th} and the 90^{th} percentiles.
- Mean absolute deviation test: to identify as outlier any observation beyond the limits:

$$M_e \pm 3MAD \qquad (3.2)$$

where, M_e = median of individual journey times in the 5-minute interval

$$MAD = \frac{\sum_{i=1}^{n} |JT_i - M_e|}{n} \qquad (3.3)$$

JT_i = individual journey time

n = number of observations in this interval.

- The quartile deviation test: to identify as outlier any observation outside the limits:

$$M_e \pm F_1(n) t_{0.975,n^*} QD \qquad (3.4)$$

where, QD = (inter-quartile range)/1.34898, and $t_{0.975,n^*}$ is the appropriate t-statistic with n^* degrees of freedom at the 95% level. n^* is corrected sample size, computed by $n^* \approx \frac{2}{5}n$. The value of $F_1(n)$ depends on whether n is odd or even and is given by the following equations:

$$F_1(n) = \sqrt{\frac{1}{\frac{2}{\pi} + \frac{1}{n}(\frac{6}{\pi} - 1)}} \qquad n \text{ even} \qquad (3.5a)$$

$$F_1(n) = \sqrt{\frac{1}{\frac{2}{\pi} + \frac{1}{n}(\frac{4}{\pi} - 1)}} \qquad n \text{ odd} \qquad (3.5b)$$

For complete details of the three tests, the reader should refer to Clark, Grant-Muller and Chen (2002). Clearly, the first test will be extreme, because 20% of all the observations will be classed as outliers by its application. It is unsuitable for this research since the outlier rate of ANPR data is only about 1.38%. Applying the second and third test to journey time data in link 7, 7:15-7:20, July 12 2001, the limits are calculated:

the second test: [89s, 168s]
the third test: [79s, 179s]

However, journey time of the instrumented vehicle on the link in the time period was 188 s, which will be identified as an outlier by both tests. Thus, those tests are unsuitable for this research, and a new method for outlier detection is required.

3.6.2.2 Outlier detection method

A traditional statistical method of outlier detection using inter-quartile range was first tested. The inter-quartile range, Q, measures the range between the first and third quartiles. The first quartile, Q_1, means that 25% of the observations are below Q_1, and the third quartile, Q_3, means that 25% of the observations are above Q_3. The inter-quartile range is therefore $Q = Q_3 - Q_1$ which is the range of the central half of the data. First, a distance of 1.5 inter-quartile range, 1.5Q, is chosen for detection of outliers, i.e.

an observation is detected as outlier if it is outside $[Q_1 - 1.5Q,\ \ Q_3 + 1.5Q]$. In this research, since outliers were considered to occur only in the upper extreme, an individual journey time JT_i, which was longer than $Q_3 + 1.5Q$, was identified as an outlier. For journey time data on link 7, 7:15-7:20, July 12 2001, the threshold of the test was 183s and the journey time of 188 s obtained by the TRG IV, was still outside the limit. Also, since in busy traffic most of individual journey times may be in a small range and the value of Q is very small, some long but valid journey times may be identified as outliers.

An alternative way is using range between 15th percentile (Q_L) and 85th percentile(Q_H), i.e. $Q_a = Q_H - Q_L$. An individual journey time is identified as invalid, if it is longer than $Q_H + 1.5Q_a$. An example is shown in Figure 3.5 with journey times on link 3, in period of 7:25-7:30, August 27 2002. For this case,

$$Q_L = 188\,\text{s}$$
$$Q_H = 218\,\text{s}$$
$$Q_a = Q_H - Q_L = 30\,\text{s}$$
$$Q_H + 1.5Q_a = 263\,\text{s}$$

A journey time longer than 263 seconds is identified as an outlier.

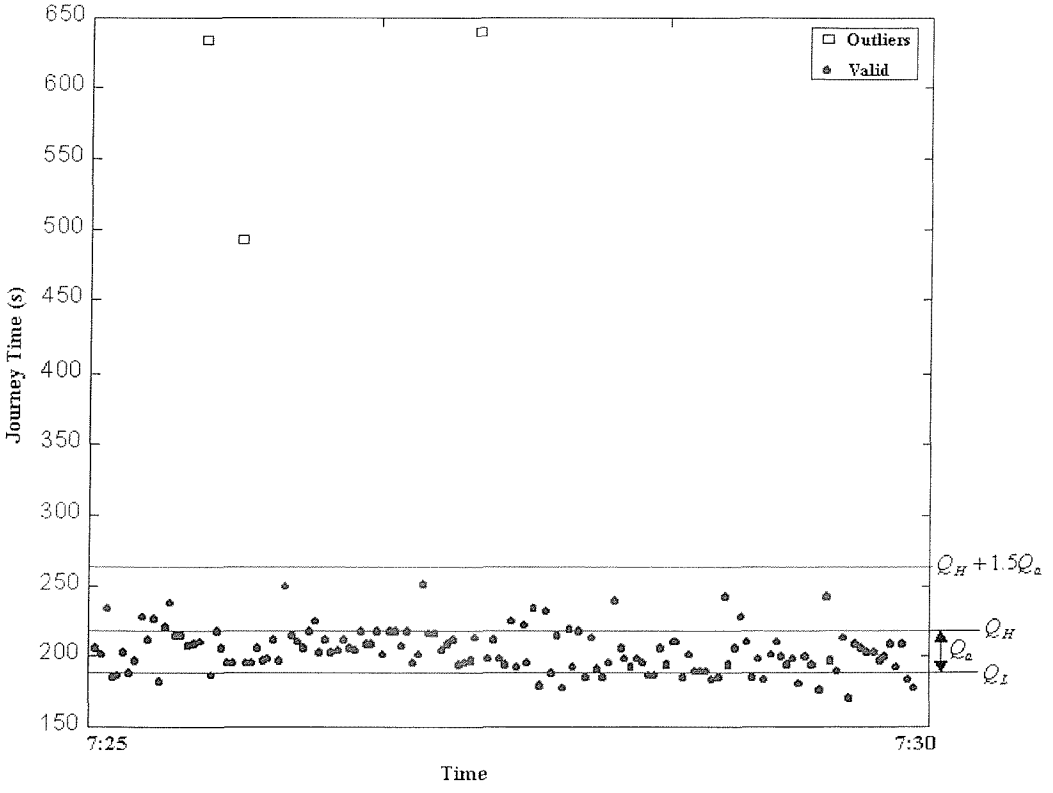


Figure 3.5 Identification of outliers on link 3, August 27, 2002.

An example of this is shown in the 5-min period containing the following 10 observations: (196, 211, 195, 250, 187, 494, 187, 210, 225, 194). The threshold of 263 is sufficient to reject the journey time of 494 but not 250.

This method is less sensitive than other methods. For example, for journey time data on link 7, in period of 7:15-7:20, July 12 2001, the threshold $Q_H + 1.5Q_a$ = 207s. The journey time of 188 s, produced by the TRG IV, is inside the limit and the higher threshold is considered to be appropriate. Since in the ramp metering survey, drivers were required to drive on motorway lane 1 as long as possible, the journey times obtained may present the longest on the link for the period, which are generally produced by heavy vehicles. This method can retain long but valid journey times and detect outliers. By applying this method, outliers in Figure 3.3 have been detected and removed and cleaned journey times are shown in Figure 3.6. The application of the outlier detection method clearly identified the outliers in the data while avoiding detecting genuine observations as outliers.

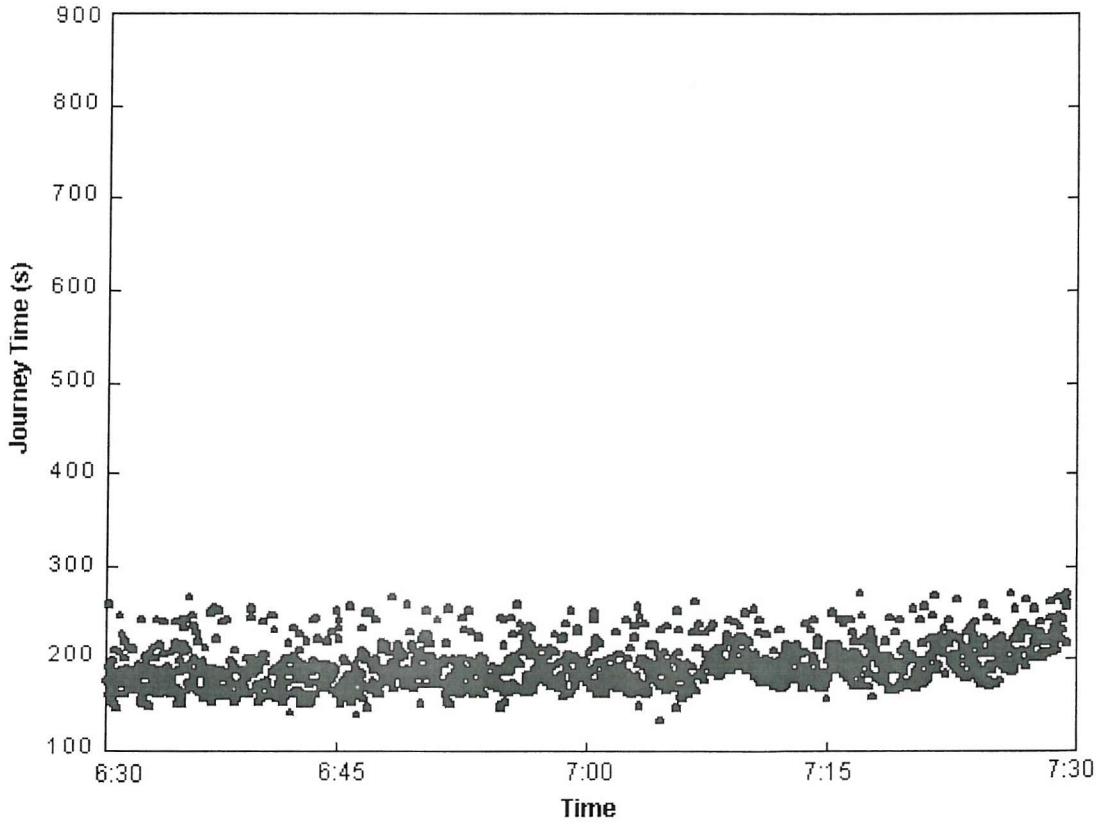


Figure 3.6 Cleaned Journey Time data on link 3: Thursday, June 12 2001

3.7 Summary

This chapter has described the data collection process involved in this research in detail. The data collection was carried out for seven links on two motorways around Southampton: the M3 and the M27. Data were collected by video cameras and GPS equipped probe vehicles.

Individual journey times were logged by video cameras using Automatic Number Plate Recognition (ANPR). Data from eighteen weeks in 2001 and 2002 were used in this research. By comparison with journey times measured by GPS, ANPR data have been shown generally reliable and accurate but with outliers, which occurred mainly for short diversions. A method of outlier detection using range between 15^{th} and 85^{th} percentile has been developed and applied to clean ANPR data.

GPS equipped vehicles were driven on some links of the survey site as probe vehicles to collect journey time data. GPS data are used for studying performance of GPS and efficiency of probe vehicle for journey time estimation. The GPS data have been converted into plane coordinates that are usable for journey time calculation. The converted data were then ready to match the map and calculate journey times. The journey time calculation based on converted GPS data is introduced in next chapter.

Chapter Four

Journey Time Estimation

4.1 Journey time calculation

For a GPS equipped probe vehicle, a link journey time is the time taken for the probe vehicle to pass locations of successive checkpoints of the link which have been accurately surveyed by GPS and stored in a database. The first step of computing journey time is to search location data of a probe vehicle to find the point whose coordinates are the closest to each checkpoint. Then, journey time is simply time difference between two successive checkpoints.

Individual link journey time of JT_{ij} and *speed* v_{ij} are estimated as follows:

$$JT_{ij} = T_{ij}^{out} - T_{ij}^{in} \qquad (4.1)$$

$$v_{ij} = \frac{L}{JT_{ij}} \qquad (4.2)$$

where

JT_{ij} = Journey time of probe vehicle j for the ith time interval

T_{ij}^{in} = Time probe vehicle j enters the link

T_{ij}^{out} = Time probe vehicle j exits the link

v_{ij} = Link speed of probe vehicle j for the ith time interval

L = Length of link

In this research, since journey times calculated from GPS data will be compared with Automatic Number Plate Recognition (ANPR) data, a link is considered as road segment between two video cameras rather than two adjacent junctions, and position of each camera has been surveyed by GPS. The capture zone of an ANPR camera is

located between 5.5-7.3 meters (18 and 24 feet) from the camera. Thus, the passing time of a vehicle recorded by the camera (T_{camera}) is the time when the vehicle passed the capture zone rather than the camera position. However, the time stamp calculated from GPS data (T_{GPS}) is the time when the vehicle passed the camera position. The relationship between the two time stamps can be estimated by:

$$T_{GPS} = T_{camera} + \frac{D_{capture}}{v} \qquad (4.3)$$

where $D_{capture}$ is the capture distance from the camera and v is the vehicle speed to traverse the capture distance. Assume $v = 22.2m/s$ (80km/h), T_{GPS} is about 0.3 s behind T_{camera}. The lower the vehicle speed, the larger the difference between the two time stamps. The lowest speed when vehicle passed a camera observed in this research was *2.45m/s,* in that case T_{GPS} was 3 seconds behind T_{camera}. In general, since vehicle speed changes in a small range over such short period, time stamp calculated by GPS data can be corrected by Equation 4.4 for comparison of journey time surveyed by GPS and ANPR cameras:

$$\begin{aligned} \widetilde{T}_{GPS} &= T_{GPS} + \frac{D_{capture}}{\bar{v}_T} \\ \bar{v}_T &= \frac{1}{T+1} \sum_{t=T_{GPS}-T}^{T_{GPS}} v_t \\ T &= \frac{D_{capture}}{v_{T_{GPS}}} \end{aligned} \qquad (4.4)$$

Most speed observations on motorway collected in this research are faster than 15 m/s (54 km/h), and the difference between GPS surveyed time and ANPR recorded time is less than 1 second. Thus, the correction equation, i.e. Equation 4.4, need only be applied only in severe congestion.

The probe vehicles starting their journey on a link within a given interval are used to estimate average link journey time and average link speed. For example, to estimate 5-minute average link journey time and speed, if for a period of 9:00-9:05, all the vehicles which enter that link between 9:00 and 9:05 are included. The average link journey time and speed are estimated as:

$$JT_i = \frac{\sum_{j=1}^{n} JT_{ij}}{n} \qquad (4.5)$$

$$v_i = \frac{L}{JT_i} \qquad (4.6)$$

where n is the number of probe vehicles for the *ith* time interval.

4.2 Accuracy of journey time estimation by GPS

For an update rate of 1 Hz, i.e. a GPS receiver surveys its position and speed at 1-second interval, times when vehicle passes each checkpoint can be directly obtained from GPS position data, and journey time can be calculated by Equation 4.1. To assess accuracy of journey time estimation by GPS, journey time obtained from GPS data are compared with ANPR data. Data collected in the car-following survey are used for the comparison. In the car-following survey, as described in Section 3.4.1, six cars were driven from the M27 J8 to J12 and a group of cameras were temporarily set on a bridge downstream at J8 in addition to the HA cameras. Thus, during a journey, each car passed four locations of ANPR cameras: J8, J10, J11 and J12. To obtain accurate individual journey times from ANPR, raw vehicle registration number (VRN) data logged by ANPR rather than matched journey time are used. An example of the raw VRN data is:

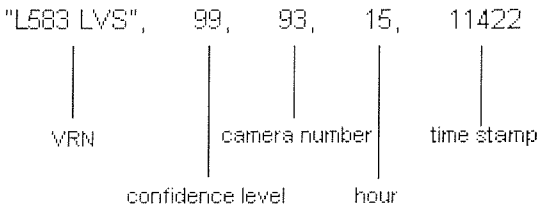


Figure 4.1 Format of raw VRN data

The 'time stamp' was recorded at frequency 5 Hz, thus, there are 18,000 (60×60×5) time stamps for one hour. And '11422' equals 38 minutes 4.4 seconds as follows:

$$\frac{11422}{60 \times 5} = 38.07 \longrightarrow \text{38 minutes}$$

$$\frac{11422 - 38 \times 60 \times 5}{5} = 4.4 \longrightarrow \text{4.4 seconds}$$

The record shown in Figure 4.1 represents camera 93 (Lane 3 at J10) having captured a vehicle with registration number L583LVS at 15:38:04 with a confidence level of 99%.

In the car-following survey, there are 79 journey times derived from raw VRN data. The 79 journey times were compared with corresponding results from GPS, as shown in Figure 4.2.

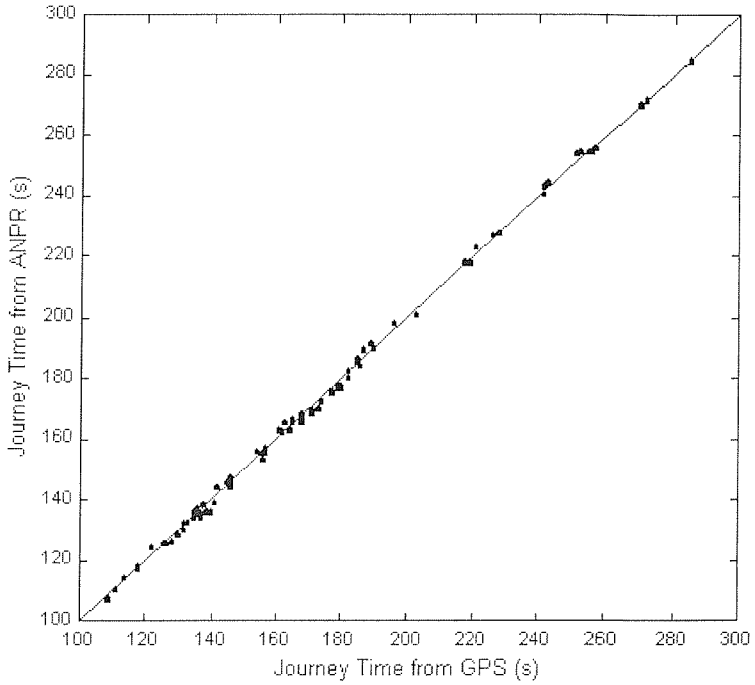


Figure 4.2 Comparison of journey times estimated from GPS and ANPR

It may be seen from Figure 4.2, that GPS can be used to estimate journey time of a probe vehicle accurately. The correlation between the two measures is 0.9908 and the mean absolute difference is near 2 seconds with a standard deviation of only 1 second. The results have shown that probe vehicles equipped with stand-alone GPS receivers using an update interval of 1 second can estimate motorway journey time accurately.

4.3 Effect of GPS update frequency

Although the update interval of 1 second (update rate of 1 Hz) is generally used by current GPS products, probe vehicles may use longer update intervals due to storage and computing capacity limits. For example, in the PRELUDE project in Rotterdam (Kroes et al, 1999), GPS devices on each probe vehicle determined location and speed every 10 seconds and the resulting information was stored together with a time-stamp

in an on-board computer. Clark and McKimm (2003) used GPS devices for journey time surveys within West Yorkshire area. In the study, the GPS device used had a limited capacity to store such points in their memory and a time resolution of 15 seconds was calculated as the finest resolution possible in order to accommodate entire shifts. In the ITIS FVD system, each in-vehicle data collection unit can record and store vehicle positions and speed at a configurable interval, but at a minimum of one minute. It is therefore necessary to study the GPS performance in estimating journey time using longer update intervals.

Currently, in-car navigation devices update GPS data generally every 1, 5 or 10 seconds. For bus management systems, in Cardiff, bus locations are updated and reported by GPS every 20 seconds, and in Maidstone, GPS devices determine bus locations at 30-second interval (Shrestha, 2003). Fleet monitoring products mostly report vehicle positions every one minute (Simmons et al., 2002). Therefore, time intervals of 5 s, 10 s, 20 s, 30 s and 60 s are studied. Data used for accuracy assessment were collected on the M3 (Survey 3) and 84 journeys were studied. Original data collected at 1-s were sampled according to a required interval. An example of sampling for 5-s interval is shown as below:

Time	X (m)	Y (m)	Speed (m/s)
7:00:00	41934.4	5644915.7	23.81
7:00:01	41926.9	5644894.7	22.30
7:00:02	41920.2	5644874.7	21.17
7:00:03	41913.9	5644855.7	20.08
7:00:04	41908.0	5644837.9	18.89
7:00:05	41902.5	5644821.1	17.61
7:00:06	41897.6	644805.9	16.14
7:00:07	41893.4	5644791.8	14.67
7:00:08	41889.7	5644779.0	13.50
7:00:09	41886.3	5644766.9	12.50

→

Time	X (m)	Y (m)	Speed (m/s)
7:00:00	41934.4	5644915.7	23.81
7:00:05	41902.5	5644821.1	17.61

Figure 4.3 An example of sampling for 5-s interval

4.3.1 Accuracy at various GPS update frequency

For an update interval more than one second, a vehicle can pass a checkpoint in anytime of an interval with equal probability. Therefore, the error in journey time estimation can be any value of the range of the interval. To assess the estimation accuracy, original GPS data collected at 1-s are sampled according to the required intervals, e.g. 5-s, 10-s, 20-s, 30-s and 60-s. The data used for accuracy assessment were collected on the M3 and 84 journeys were studied.

It has been shown that good accuracy of journey time estimation can be achieved using GPS update interval of 1-s. Journey times estimated at 1-s interval are considered to be "real" journey times and journey times estimated using longer intervals are compared with the "real" values. The errors in journey time estimation using 5-second interval are shown in Figure 4.4. Range of the errors is [0, 1, 2, 3, 4] and probabilities to achieve each value in the range are equally likely.

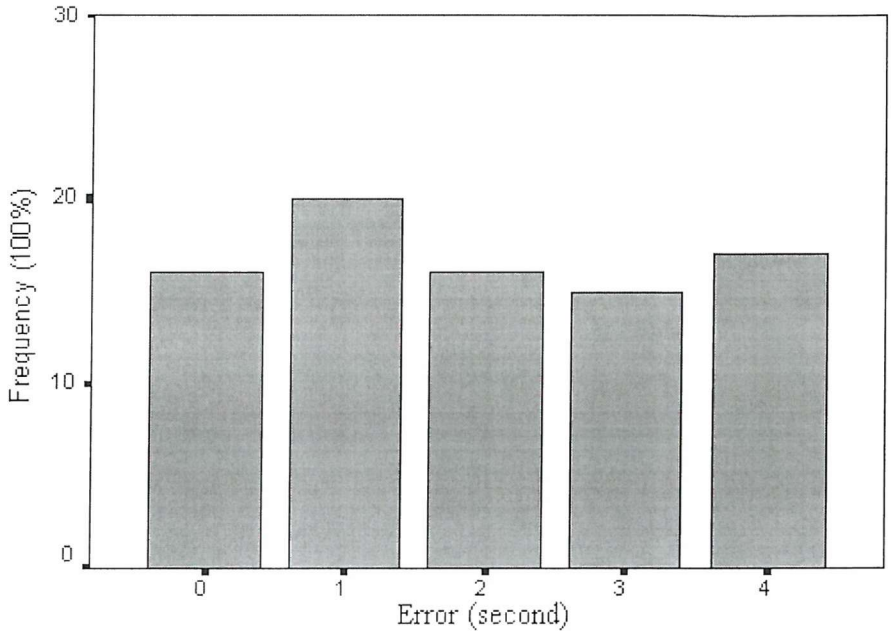


Figure 4.4 Errors in journey time estimation at 5-second interval

The errors are therefore estimated to have a discrete uniform distribution and the probability function is:

$$P(error = p) = \begin{cases} \frac{1}{l} & if \ \ p = 0,1,\ldots,l-1 \\ 0 & otherwise \end{cases} \qquad (4.7)$$

where l is the length of the update interval. For example, if positions of a probe vehicle are determined every 60 seconds, the errors of journey time estimation are uniformly distributed on [0,59]. The distribution indicates that probability to achieve error of zero is 1/60, and probability to obtain an error which is less than 30 seconds is 50%.

4.3.2 Accuracy at various GPS update frequency with correlation

For an interval longer than one second, a higher accuracy can be achieved using a simple computing method (Khan and Thanasupsin, 2000). The time stamp to pass a checkpoint can be estimated based on linear interpolation between two consecutive reports of location. The last reporting time before the checkpoint, t_1, and the first reporting time after the checkpoint, t_2, are required, and corresponding distances D_1 and D_2 are calculated (Figure 4.5).

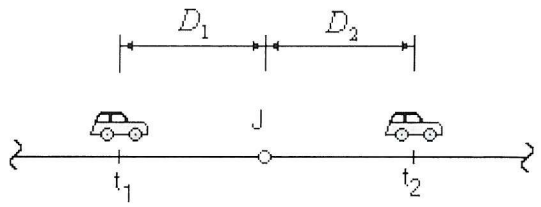


Figure 4.5 Estimating the closest time stamp to a checkpoint

The closest time stamp when the vehicle passed the checkpoint J is calculated by Equation 4.8:

$$t_J = t_1 + (t_2 - t_1)\frac{D_2}{D_1} \qquad (4.8)$$

This method is based on the assumption that vehicle speed is stable over a short interval. Good accuracy can be achieved if a probe vehicle maintains stable speed over a sampling period. Since the data used in this research were collected on motorways and vehicle speed on motorway is generally stable over a short period, good accuracy of journey time estimation can be obtained using a relatively shorter update interval, such as 5-s, 10-s, 20-s and 30-s. However, 1-minute interval may cause large errors for journey time measurement since speed may change greatly during a 1-minute interval. Speed profiles at various update intervals are shown in Figure 4.6. It can be shown that 5-second and 10-second sampling intervals can retain the original information of speed change. The longer the interval adopted, the more information is lost. Thus, the accuracy of journey time estimation using this method depends on the stability of vehicle speed.

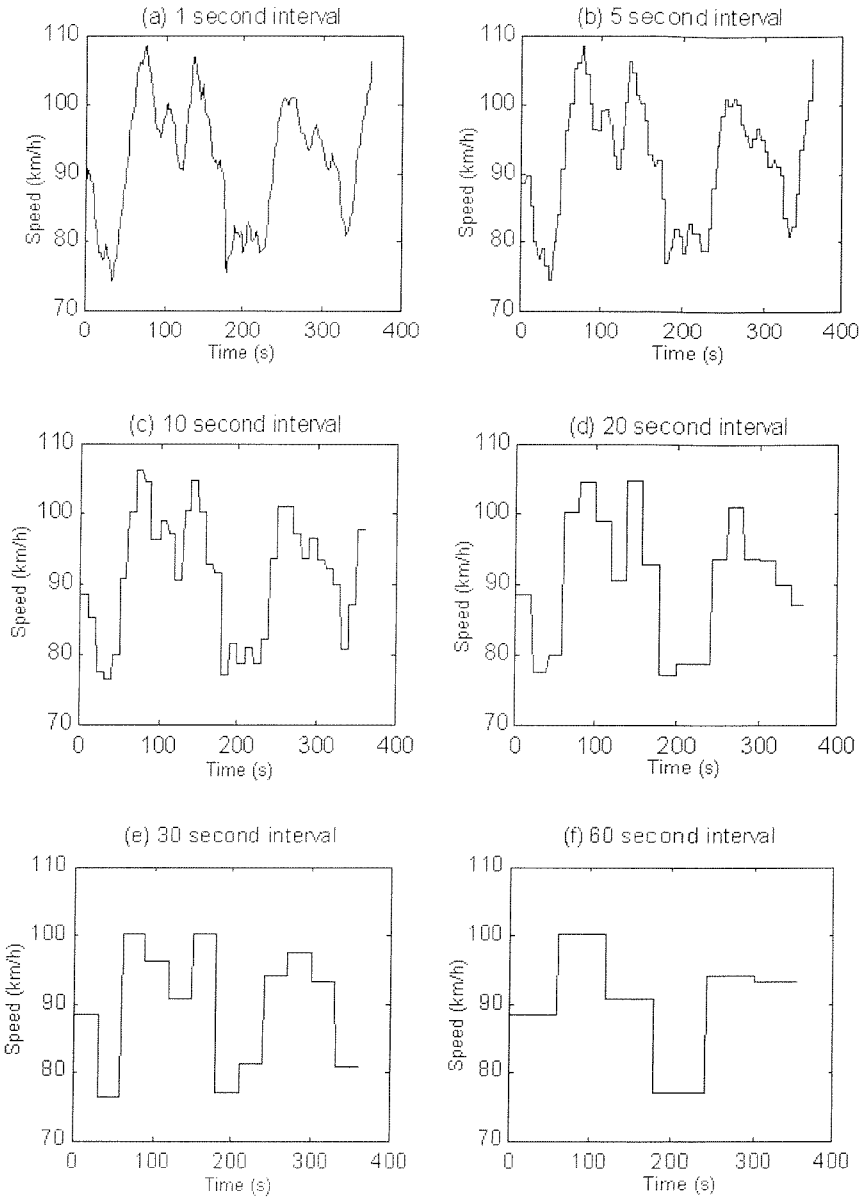


Figure 4.6 Speed profiles at various update intervals

Journey times of the 83 journeys on the M3, carried out in Survey 3, are calculated at intervals of 5 s, 10 s, 20 s, 30 s, 60 s and compared with the 'real' value, i.e. estimated journey times at 1-second interval. Journey time differences at various update intervals are shown in Figure 4.7. The accuracy results of journey time estimation with different update intervals, including mean error, standard deviation of error, observed maximum of error and the percentage of zero error are summarised in Table 4.1. The results shown in Table 4.1 indicate that relatively short update intervals, such as 5 seconds, 10 seconds, 20 seconds and 30 seconds, do not have a great effect on the accuracy of journey time estimation, while large errors are observed with 1-minute intervals.

It is not surprising that for 60-second intervals, the worst estimates (with an error of 18 seconds) were obtained during traffic flow breakdown. That is because the vehicle speed changed greatly over one minute. However, some very good estimates (errors of zero) were also observed in flow breakdown since the vehicle travelled in severe congestion with very slow but stable speed.

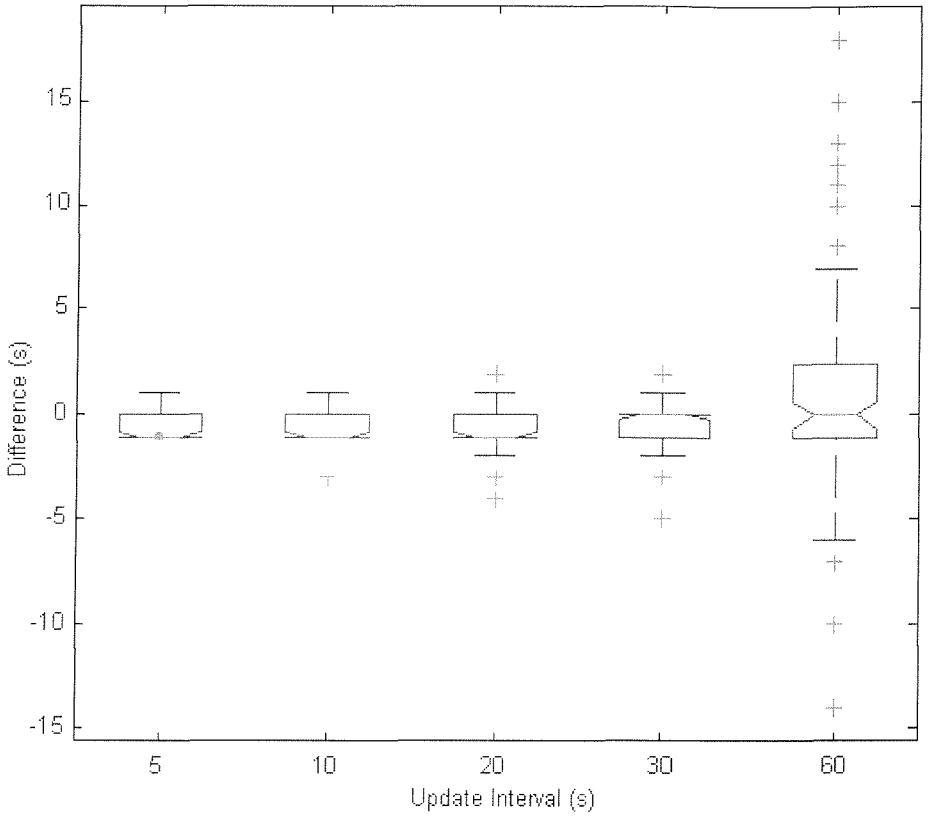


Figure 4.7 Distribution of journey time difference at various update intervals

Table 4.1 Statistics of errors in journey time estimation at various update intervals

Interval (s)	Mean (s)	Std. Deviation (s)	Maximum (s)	Percentage of Zero Error*
5	0.57	0.498	1	42.9%
10	0.60	0.563	3	42.7%
20	0.80	0.753	4	35.7%
30	0.85	0.853	5	35.7%
60	3.69	4.502	18	15.5%

*Zero error means that estimated journey time equals to 'true' journey time.

4.4 Journey time report

To provide real-time travel time information, probe vehicles need frequent communication with a central computer. Generally, communication frequency is much lower than GPS update frequency. Although collecting GPS data at such short intervals has a very minor effect in terms of cost, communication and data management may be significant. In the PRELUDE project in Rotterdam (Kroes et al., 1999), probe vehicles determined and stored location and speed data every 10 seconds. The accumulated information was transferred to a central computer by GSM every five minutes. In such systems, vehicles are equipped with only GPS and communication systems, i.e. without digital map and GIS (Geographic Information Systems) software, and travel times are calculated by central computer from the vehicle location data. For probe vehicles with navigation systems consisting of on-board computer and GIS software, a possible solution would be to report link travel time along a journey. Navigation computers in ADVANCE equipped vehicles carried on tasks of navigation and estimation of link traversal times, which were transmitted by radio frequency to Traffic Information Centre every 5 minutes (Sen et al., 1997).

4.5 Summary

This chapter has described the process of journey time estimation from position data surveyed by GPS. Journey times calculated by GPS data have been compared with video camera data, which provide a positioning measurement error within 0.8m. The comparison has shown that the current performance of GPS enables stand-alone GPS receivers to be used for journey time measurement. The results will encourage the acceptance and application of GPS equipped probe vehicles.

Some potential GPS equipped probe vehicles use different intervals to determine vehicle position. For example, in some bus management systems, an on-board GPS device determines bus position each 20 or 30 seconds. It is therefore necessary to study the impact of different update intervals on accuracy of journey time estimation. For intervals larger than 1 second, without correlation, the error is directly related to the length of the interval. With correlation, which is based on linear interpolation using the two closest time stamps before and after a checkpoint to calculate the 'true' passing time, relatively shorter intervals such as 5 s, 10 s, 20 s and 30 s can also produce

accurate results. One-minute interval results in unstable accuracy because of the variations in speed within such intervals. Traffic flow has great influence on the estimate accuracy, e.g. poor estimates may be obtained during traffic flow breakdown. The results will encourage more applications of probe vehicles. Vehicles equipped with GPS devices using low updating frequency can be used to provide good estimates of journey times.

Chapter Five

Sample size of probe vehicles

5.1 Introduction

Although the results described in Chapter 4 have shown the ability of GPS to measure link journey time of an individual vehicle accurately, the journey time of any one vehicle does not represent average journey time. In general, the larger the sample size, the greater the accuracy of representation of the mean of the whole data population. Previous studies and applications have indicated that a relatively small number of probe vehicles travelling in the traffic stream can provide potentially valuable information about current journey times (Turner et. al., 1998; Srinivasan and Jovanis, 1996). However, too few probe vehicles may provide erroneous or misleading data. For example, in the ADVANCE system (Sen et al., 1996), probe vehicles were used to estimate real-time journey times to navigation assistants. However, in general, the navigation assistants computed the desirable routes using default journey times based on historical data rather than the real-time estimates. The navigation assistants used real-time observations only when the real-time estimates differed significantly from the default estimates. In such situations, incidents were suspected. In incident free traffic, the historical data were considered to be more reliable since the real-time journey times estimated by few observations may have large variance.

Although the proportion of GPS-equipped vehicles is expected to increase with the gradual implementation of in-vehicle navigation systems, the capacity and cost of wireless communication links between in-vehicle devices and the traffic management centre will be likely to still limit the available sample size of probe vehicles. Therefore, it is necessary to determine the minimum number of vehicles that should be equipped as "probes" to estimate journey time or speed within a satisfactory statistical accuracy.

Probe vehicle sample size has two separate meanings: (1) overall probe vehicle percentage of the total vehicle population to estimate journey times for a desired reliability and proportion of link coverage over an entire network; and (2) number of probe vehicles sampled in a link for a required time interval and a desired statistical accuracy. This research focuses on the second, i.e. determining the minimum number of probe vehicle in a link to estimate reliable link journey time.

Previous researches have provided various estimates of the number of probe vehicles required on roads of different character. However, such researches to determine the number of probe vehicles have used simulation data (Chen and Chien, 2000; Cheu et al., 2002) or a small proportion of probe vehicles (Turner and Holdener, 1995). These approaches lack a detailed description of individual journey times. Since video cameras can log a large proportion of individual vehicles using ANPR (automatic number plate recognition) technology, ANPR data have been used in this research to provide a comprehensive description of individual journey times. Statistical characteristics of individual journey times can be estimated from ANPR data, which are used to study two main questions related to probe vehicle sample size: (1) How few probe vehicles in a link are needed to maintain the desirable statistical accuracy of link journey time estimation? (2) Which factors affect the sample size?

5.2 Literature Review

A number of studies into probe vehicle sample size have been reported in the literature. Statistical analyses have been used to examine journey time data and determine sample size for different sampling intervals and desirable accuracy. Two main data sources have been employed: empirical data and simulation data.

5.2.1 Empirical data

In the ADVANCE project, probe vehicle sample size was calculated using a matrix of trips between every origin and destination (O-D matrix). Sample sizes were calculated based on having at least one probe vehicle traverse a certain percentage of roadway links during different time intervals (5,10,15 and 20 minutes). For example, with 5,000

probe vehicles, 60% of all links in a 200-square mile test area could be traversed by one vehicle per 15 minute.

In the ADVANCE project, the relationship between sample size and accuracy of link journey time estimates has been also studied. The effect of sample size on standard error of means of probe reports is shown in Figure 5.1. The standard errors of the mean of probe reports do not go to zero and there is a minimum value below which standard errors never fall, no matter how large the number of probes becomes. While this minimum value is link-specific, the general shape of the relationship between standard error and sample size is similar for all moderately or heavily congested signalized links.

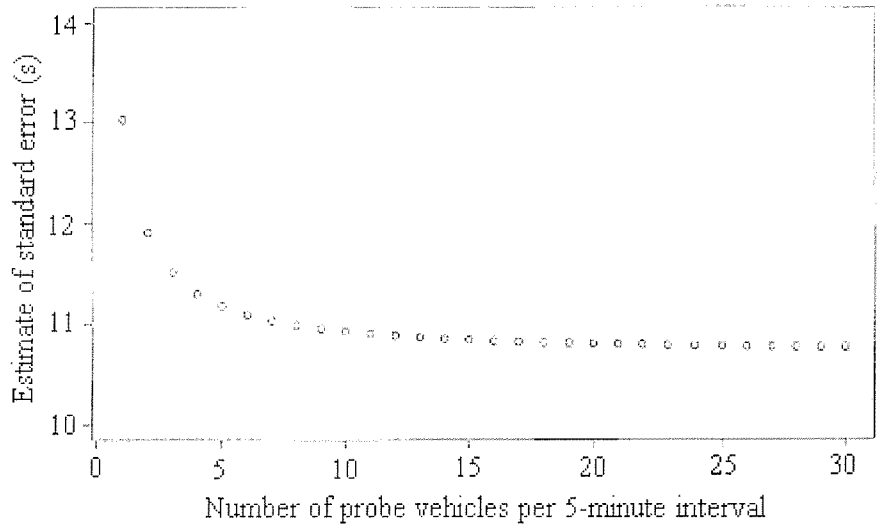


Figure 5.1 Standard error vs. number of probes (Source: Sen, Soet and Berka 1997)

Turner and Holdener (1995) provided recommendations about probe vehicle sample sizes using empirical journey time data from Houston AVI (Automatic Vehicle Identification) traffic monitoring systems. In the Houston area, AVI reader units were placed at 1.8- to 8.0 kilometre intervals along all the area's major freeways, eventually including over 483 kilometres of highway and 161 kilometres of HOV (high-occupancy vehicle) lanes. Based upon the AVI checkpoints, there were a total of 60 freeway segments for which journey time data were analysed. It has been estimated that approximately 40,000 AVI tags have been distributed in the Houston area and the average number of probe vehicles providing a journey time on each segment ranged from 1 to 7 every 5 minutes, or 2 to 20 every 15 minutes. The coefficient of variation (c.v.) of journey times were calculated and sampling theory was applied to determine

the sample size based on the c.v. values. Results showed that for 5-minute periods, a 95% confidence level and a 10% relative error, the sample sizes needed ranged from 1 probe vehicle every 5 minutes for free-flow conditions (HOV lane segment) to 6 probe vehicles every 5 minutes for severely congested conditions. Sample sizes were slightly lower using a 90% confidence level and 10% relatively error. For 15-minute periods, a 95% confidence level and a 10% relative error, the sample sizes needed ranged from one probe vehicle to 8 probe vehicles every 15 minutes. For a 90% confidence level and 10% relatively error, sample sizes from 1 to 6 were required.

Since the required sample sizes calculated for the Houston AVI system were directly related to the journey time variation, the relationship between average speed and journey time variation was studied and a regression equation was then obtained that predicted journey time variation using average speed:

$$85^{th} \text{ percentile c.v.} = 33.9\text{-}0.27\times\text{Average Speed (km/h)} \qquad (5.1)$$
$$r^2 = 0.60$$

where r is the correlation coefficient of observed and estimated values. The 85^{th} percentile c.v. were then used to estimated the require sample sizes for the desired statistical accuracy.

5.2.2 Based on simulation data

Simulation models have been used to determine probe vehicle sample sizes with respect to (1) overall probe vehicle percentage; and (2) the number of probe vehicles sampled in a link.

Srinivasan and Jovanis (1996) developed a heuristic algorithm to estimate the total number of probe vehicles required for the Sacramento network, in California. A heuristic algorithm was implemented using a simulation procedure. Based on three simplifying assumptions: the normality of journey time distribution, constant journey time variations over all links and constant journey time variations over all measurement periods. The simulation procedure was run to determine the total number of probe vehicles towards two scenarios: 4 probe vehicles on each link during 10 minutes, and 6 probe vehicles on each link during 15 minutes. However, in the absence of actual

journey time data, it was not clear how the assumptions would affect the simulation results.

More recently researchers used microscopic simulation to estimate the number of probe vehicles. Chen and Chien (2000) used a microscopic model, CORSIM, to simulate traffic flow on a freeway segment of I-80 in New Jersey and determine the minimum number of probe vehicles on a link for a 5 minute interval. Journey times of all vehicles on each link of the freeway segment were recorded from the CORSIM output, and distributions of link journey times were then obtained from these statistics. The simulation results reported that vehicle journey times on some links were not normally distributed. Factors affecting the vehicle journey time distribution were considered to be geometric condition and traffic volume on the link. It was observed that longer links tended to absorb the impact of various traffic flows. To study the impact of traffic volume, five demand levels were chosen to generate various traffic volumes from free-flow to near-capacity over the freeway links. Results showed that the type of link journey time distribution could vary with the traffic volume it carries and non-normal distributions were likely to be found under the highest level of demand (level 5).

The results also studied the impact of traffic volume on the percentage of probe vehicles required to provide accurate estimates on link journey times. With 5-minute time interval, 5 percent maximum relative error, the minimum percentages of probe vehicles under five traffic volume levels are shown in Figure 5.2.

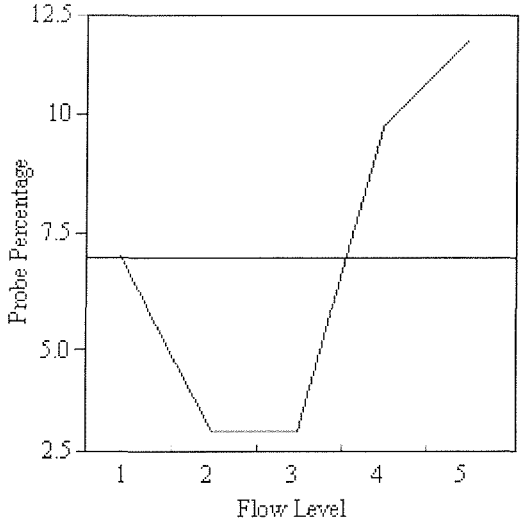


Figure 5.2 Probe vehicle percentage versus demand level

(Source: Chen and Chien 2000)

It was shown that for light or heavy traffic flow, the minimum required probe vehicle percentage was higher than at the medium demand levels.

In Singapore, a large-scale, nationwide travel speed information acquisition and dissemination system has already been in operation using a large fleet of taxis equipped with DGPS devices. Cheu et al. (2002) used a simulation approach to study the minimum number of probe vehicles in a link and overall probe vehicle percentages for the road network of the Clementi town area in Singapore. The simulation used GPS equipped probe vehicles to report journey time/link speeds. A GPS receiver was assumed to continuously survey its position at 1 sec or 2 sec intervals. The instantaneous position was then compared with an on-board digital map database and the vehicle position and time stamp passing each checkpoint along the route were stored and transmitted to the traffic management centre every 700 seconds. The data collection period was therefore 700 seconds. Results indicated that for an absolute error in estimated average link speed to be less than 5km/h at least 95% of the time, the network needed 4% to 5% of total traffic as probe vehicles or at least ten probe vehicles on a link within 700 seconds.

5.3 Determining probe vehicle sample size

It is important to able to rely on as few probe vehicles as possible for satisfactory statistical accuracy because probe vehicles are costly in both equipment and real-time communication. The objective of this chapter is to develop a methodology to determine the minimum number of probe vehicles in a motorway link that would provide link journey time for a prescribed time period and desired accuracy.

If 5 minute interval is selected, journey times of all vehicles on a link during a 5 minute interval are the population being surveyed. Probe vehicles are considered as samples. The object of the sampling process is to estimate the population mean, e.g. average link journey time. The statistical sampling methodology can be used to determine the minimum required number of probe vehicles that would provide reliable link journey time estimates. For a population, parameters of the population are denoted by:

Population mean = μ

Population standard deviation = σ

Statistical inferences about the population mean are based on the sample mean

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

For a *permitted error* e, expressed as a percentage of the population mean, and a confidence level $1-\alpha$, the statistical sampling methodology provides an equation to determine the required sample size:

$$n = \left[\frac{z_{\alpha/2}\sigma}{e\mu} \right]^2 \qquad (5.2)$$

where $z_{\alpha/2}$ denotes the upper $\alpha/2$ point of the standard normal distribution. A detailed illustration of the notation $z_{\alpha/2}$ is shown in Appendix B1. Commonly used values of $z_{\alpha/2}$ are shown in Table 5.1 for easy reference. The Equation 5.2 determines the required sample size to be $100(1-\alpha)\%$ to ensure that the error of estimation $\frac{|\overline{X} - \mu|}{\mu}$ does not exceed e. See Appendix B2 for a rationale of the equation.

Table 5.1 Values of $z_{\alpha/2}$

$1-\alpha$	.80	.85	.90	.95	.99
$z_{\alpha/2}$	1.28	1.44	1.654	1.96	2.58

According to Equation 5.2, selection of permitted error and confidence level will directly affect the minimum number of probe vehicles required. This research calculated the required sample size of probe vehicles based on a permitted error of ±10%, which was widely used in previous probe vehicle studies (Turner and Holdener, 1995; Srinivasan and Jovanis, 1996; Chen and Chien, 2000; Cheu et al., 2002). The confidence level is the probability associated with the permitted error. 90% and 95% confidence levels are commonly used in statistics. However, the same methodology can be used for any specified permitted error and confidence level.

Equation 5.2 can be applied only when the distribution of $\overline{X}$ is normal. It has been proven that distribution of $\overline{X}$ is normal when sampling from a normal population.

When sampling from a non-normal population, the central limit theorem (Johnson and Bhattacharyya, 2001) states that when the sample size n is large, the distribution $\bar{X}$ is approximately normal, regardless of the shape of the population distribution. In practice, the normal approximation is usually adequate when n is greater than 30. Since in this research, the number of probe vehicles in a link for a 5-minute interval generally may not exceed 30, it is necessary to study vehicle journey time distribution and decide whether Equation 5.2 can be applied.

5.4 Vehicle journey time distribution

By considering all individual journey times on a link over a 5 minute interval as the population, individual journey times logged by ANPR cameras in the 5 minute interval are a sample of the population to estimate the population distribution and parameters. As introduced in Section 3.3.1, ANPR cameras can log individual journey times of a substantial proportion of all vehicles and the sampling has been assumed to represent the sampled population.

The distribution of journey times of individual vehicles on a link in a time interval has been assumed to be normal by the majority of previous researches (Turner and Holdener, 1995; Srinivasan and Jovanis, 1996; Cheu et al., 2002). However, Chen and Chien (2000), analysing results generated by microscopic simulations, observed non-normal distribution of journey times on a link during certain intervals. In this research, the distribution of journey times on links with different characteristics in different traffic conditions have been studied. The objective of this section is to reveal relationships between journey time distribution and road characteristics as well as traffic conditions.

5.4.1 Normal distribution

5.4.1.1 Normal distribution test: an example

In this research, individual journey times logged by video cameras were used to infer distribution type and parameters. An example of the approach is described using link 7 at 9:20-9:25 on the 11th June 2001. In the 5 minute period, ANPR cameras logged 196

individual journey times, in a total vehicle count of 390. Thus, with a sample size of 50.26%, the population distribution could be estimated from the sampling distribution.

A graphical normality test is shown in Figure 5.3, in which, if the samples came from a normal distribution, the plot would appear linear. Otherwise the plot would appear curved. In Figure 5.3, the individual journey times tend to follow the linearity of the normal quantile plot, indicating that the journey time distribution can be considered as normal.

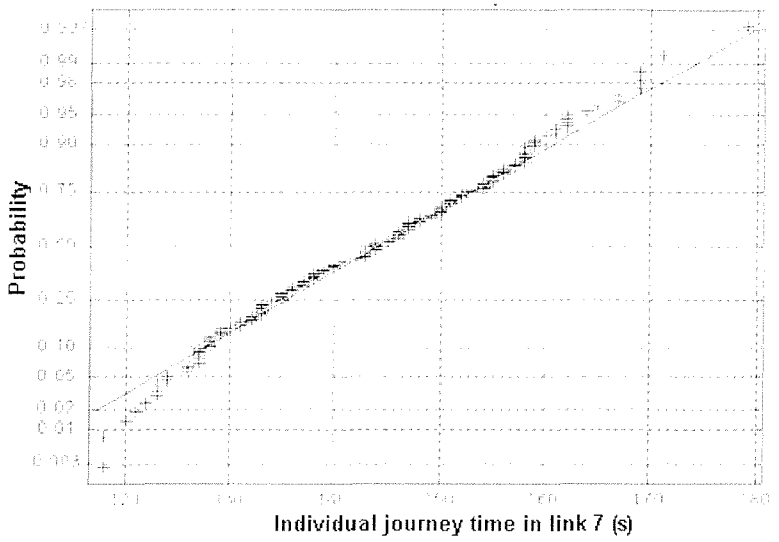


Figure 5.3 Normal Probability Plot

The Kolmogorov-Smirnov (K-S) Test was used to obtain a quantitative assessment of normality. The output of the SPSS K-S test is given in Table 5.2. In a K-S test, large significance values (>0.05) indicate that the observed distribution corresponds to the expected normal distribution.

Table 5.2 One-Sample Kolmogorov-Smirnov Test

		V1
N		196
Normal Parameters[a,b]	Mean	145.2245
	Std. Deviation	13.91883
Most Extreme Differences	Absolute	.050
	Positive	.050
	Negative	-.033
Kolmogorov-Smirnov Z		.702
Asymp. Sig. (2-tailed)		.707

a. Test distribution is Normal.

b. Calculated from data.

5.4.1.2 Percentage of normal distribution on each link

In this study, data from each of the seven links has been examined for 3 hours, i.e. 36 periods, selected from different times of morning and evening peak period in different seasons. The percentage of time in which journey times is not statistically different to normal distributions on each link are listed in Table 5.3.

Table 5.3 Normality percentage for each link

Link Number	Percentage of Normality
1	80%
2	75%
3	74%
4	88%
5	85%
6	81%
7	83%

The results in Table 5.3 indicate that in most cases, journey times of all vehicles in a link over a 5-minute interval have a normal distribution. A correlation of 0.095 between the percentage of normality and link length indicates that there is no significant relationship between the two variables. Relatively low percentages were observed for link 2 and link 3 which carry heaviest traffic of the seven links. The impact of traffic volume on individual journey time distribution can be studied, since on each of the seven links studied, groups of inductive loops provided traffic flow data such as vehicle count per minute.

5.4.2 Non-normal distribution

As introduced in Section 5.5.5, Chen and Chien (2000) observed non-normality in link journey time distributions using microscopic simulation data, and non-normal distributions were found mostly under the highest level of demand. The results may have been affected by the mechanisms of the simulation process. However, results from this study also showed that journey time distribution remained normal in heavy traffic, and non-normality was observed to be most likely in some special stages of congestion.

In non-congested traffic, average link speeds were found to decline slightly with increasing flow. When the traffic flow exceeded the estimated capacity of the location, average link speed began to decline more quickly with increasing flows, and

breakdowns in flow occurred. An example of the relationship between traffic flow and average speed is shown in Figure 5.4, in which traffic on link 7 during the morning peak hours on the 9th July 2001 is displayed.

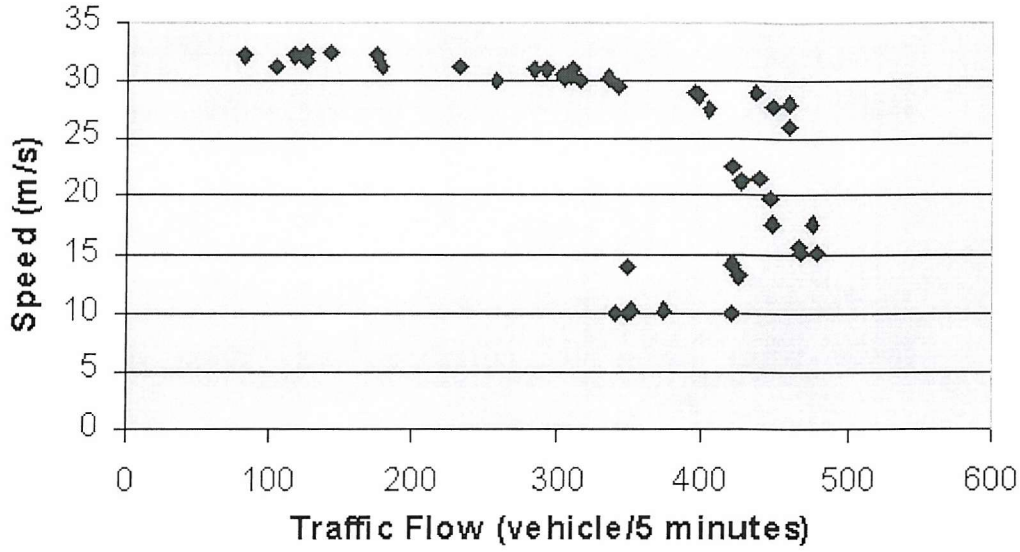


Figure 5.4 Observed speed-flow relationship on link 7

Since the relationship between flow and speed is not linear, increase of traffic flow may not affect link speed (or link journey time) instantaneously. Average link journey times and flow data on link 7 during the morning peak on the 9th July 2001 are shown in Figure 5.5.

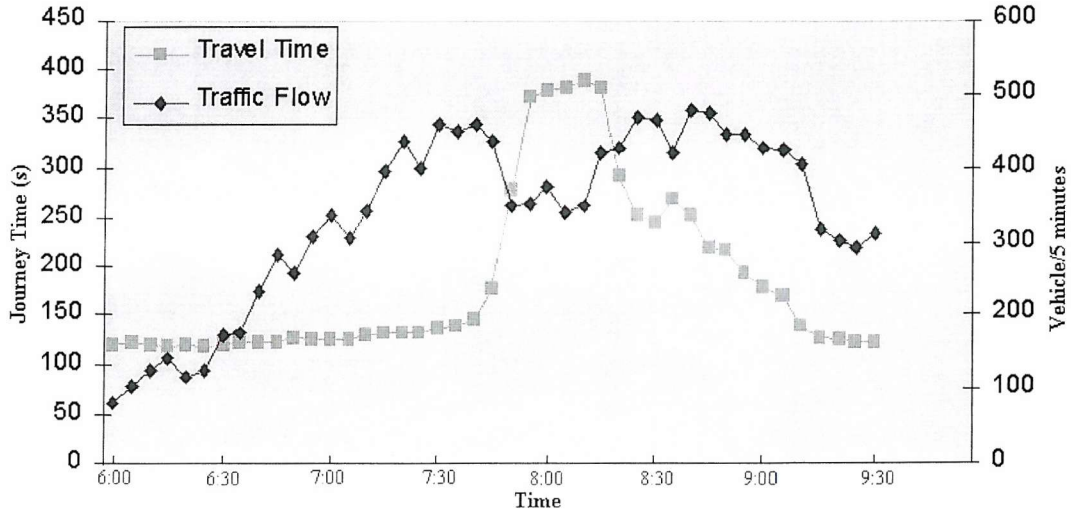


Figure 5.5 Link journey time with flow data on link 7, 09/07/2001

It may be seen from the figure that in the early part of the peak, link journey time only increased by a small amount with increasing flow. Later when traffic flow data exceeds the expected capacity flow, link journey times increased rapidly and traffic flow declined. Journey time then remained at a high level for some time, before gradually decreasing. Individual journey times in the link for each period from 6:00 to 9:30 have been studied and the non-normality was only observed in five periods: 7:40-7:45, 7:45-7:50, 7:50-7:55, 8:20-8:25, 8:25-8:30, as shown in Figure 5.6. In the five periods, journey time increased or decreased significantly from the previous perod (i.e more than about 10%). The distribution of journey times for a 5 minute interval between 8:00 and 8:20 was normal, when journey time was the highest, since journey time was relatively stable over the period. Compared with the flow data shown in Figure 5.5, non-normality begins to occur when traffic flow achieves the maximum value, i.e. the occurrence of non-normality may indicate a breakdown in traffic flow.

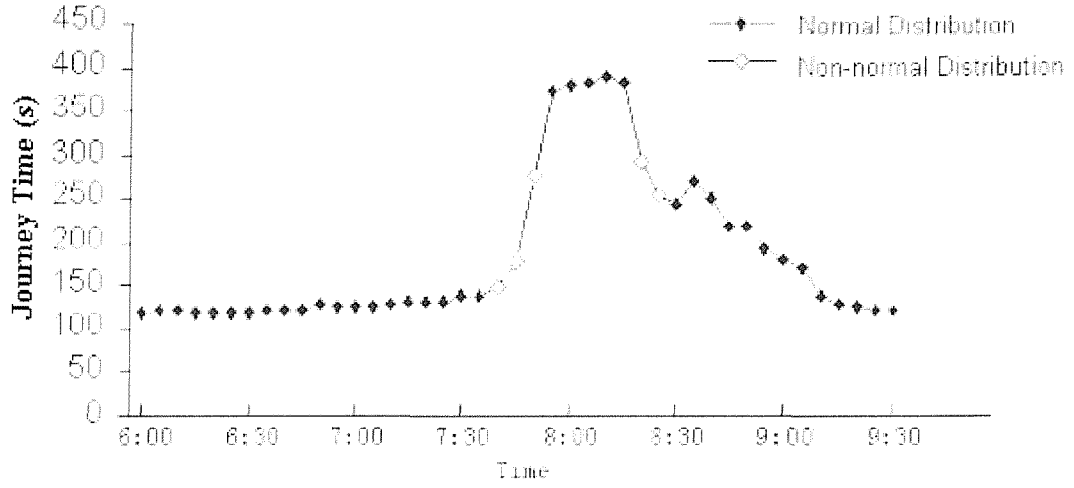


Figure 5.6 Non-normal distribution intervals, 09/07/2001

To explain the observed non-normality, all individual journey times logged during these 5-minute intervals with non-normality were studied. When the update interval was changed to 1 minute instead of 5 minutes, a normal distribution was obtained for each of the 1-minute intervals. However, the means and standard deviations of each interval would be very different. Since the journey time in a 5-minute interval can change greatly, a 5-minute interval is not suitable for journey time updating. The journey time distribution of each minute from 7:50 to 7:55 is shown in Figure 5.7. Obviously, when journey time changes very rapidly, the journey time for even a minute interval can have a non-normal distribution. In such cases, the mean and standard

deviation of journey time are time-series rather than static variables, and the journey time should be estimated by time series analysis methods instead of static sampling theory.

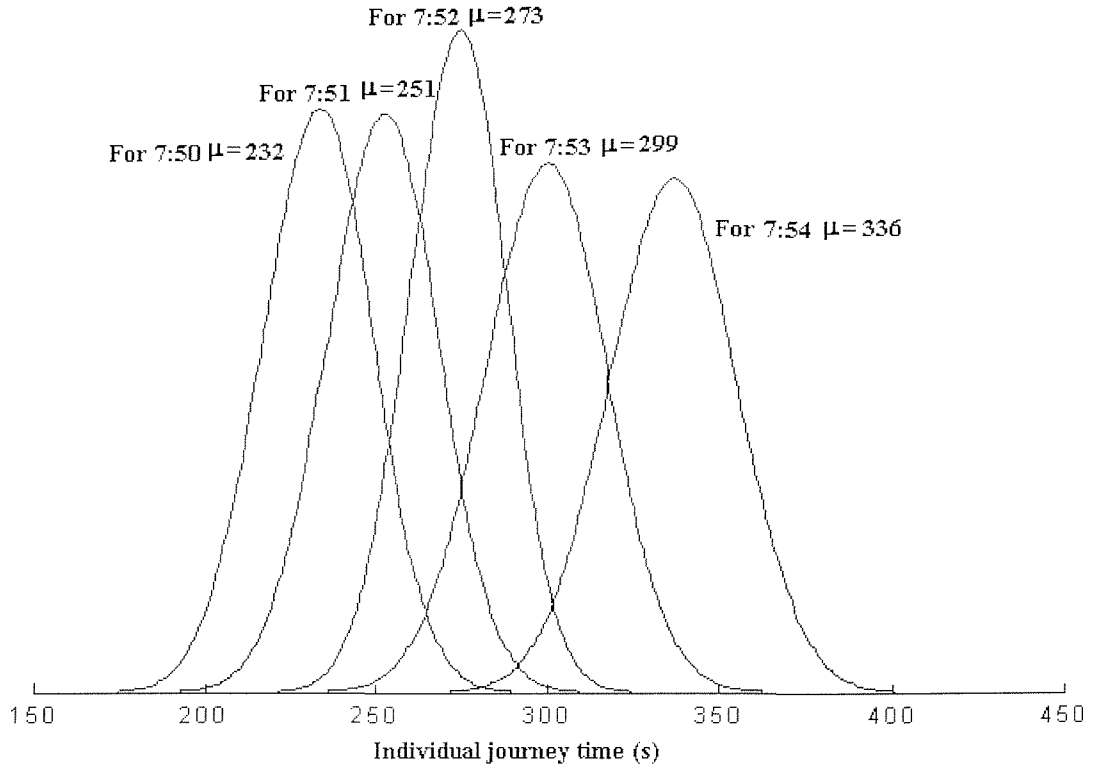


Figure 5.7 The journey time distribution for each minute

5.5 Determining sample size of probe vehicles

5.5.1 Sample size and average speed

After identifying the normal distribution of vehicle journey times, probe vehicle sample size can be determined by Equation 5.2. As discussed in Section 5.3.2, the permitted error used in this research has been defined as 10%. For a confidence level of 95%, according to Table 5.2, $z_{0.05/2} = 1.96$, Equation 5.2 can be written as:

$$n = [19.6(\sigma / \mu)]^2 \qquad (5.3)$$

For a confidence level of 90%, $z_{0.1/2} = 1.654$, probe vehicle sample size is determined by:

$$n = [16.45(\sigma / \mu)]^2 \qquad (5.4)$$

Equations 5.3 and 5.4 indicate that probe vehicle sample size is only related to the coefficient of variation (c.v.), a relative measure of dispersion obtained by dividing the standard deviation by the mean.

Individual journey times on link 3 were calculated for 20 hours, i.e. 240 time intervals. Mean journey time and standard deviation for each time interval are shown in Figure 5.8.

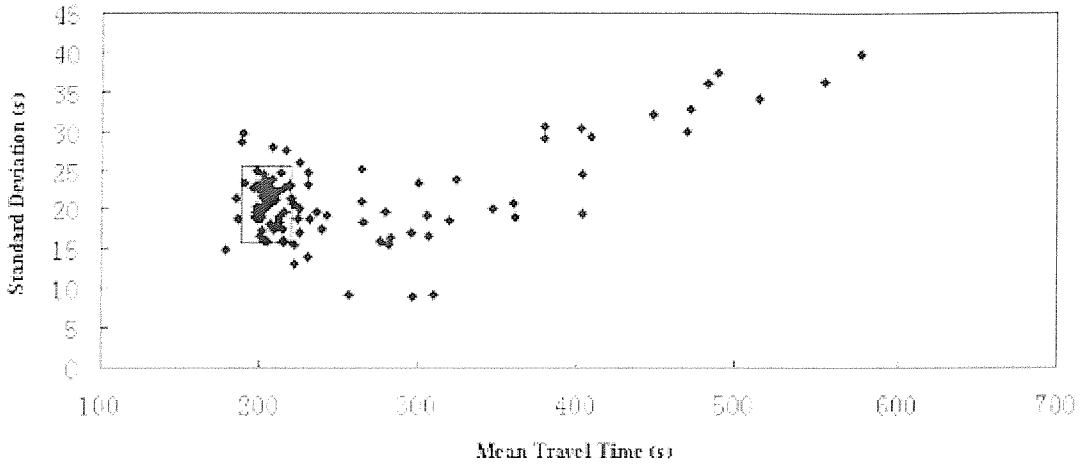


Figure 5.8 Mean journey time and standard deviation for link 3

These data indicate that:

(i) For the same link and under non-congestion conditions, mean journey times and standard deviations change only within a small range. 75% of all cases have mean journey times in a range of 190-220 seconds, and standard deviations in a range of 16-25 seconds, i.e. inside the rectangle shown in Figure 5.8. That indicates that most c.v. values are in a small range, furthermore, sample size requirements are similar for most cases according to Equations 5.3 and 5.4.

(ii) Larger values of standard deviation occur with lower mean journey times, and more probe vehicles are required in lighter traffic. This is in agreement with simulation results obtained by Chen and Chien (2000). It is considered that, in the situation of light traffic, large variations in link journey time can be attributed mainly to the variation in driving behaviour, e.g. free speed selections.

(iii) When mean journey time is increasing due to increasing levels of traffic volume, standard deviation will be also increasing. However, the range of increase in journey time is greater than that of standard deviation.

By applying Equation 5.3, the minimum numbers of probe vehicles for a confidence level of 95% can be determined and falls within the range of probe vehicle sample size 1 to 10 for link 3 for various traffic conditions. The relationship between mean link speed and the minimum number of probe vehicles required is shown in Figure 5.9, which matches the above analyses about Figure 5.8.

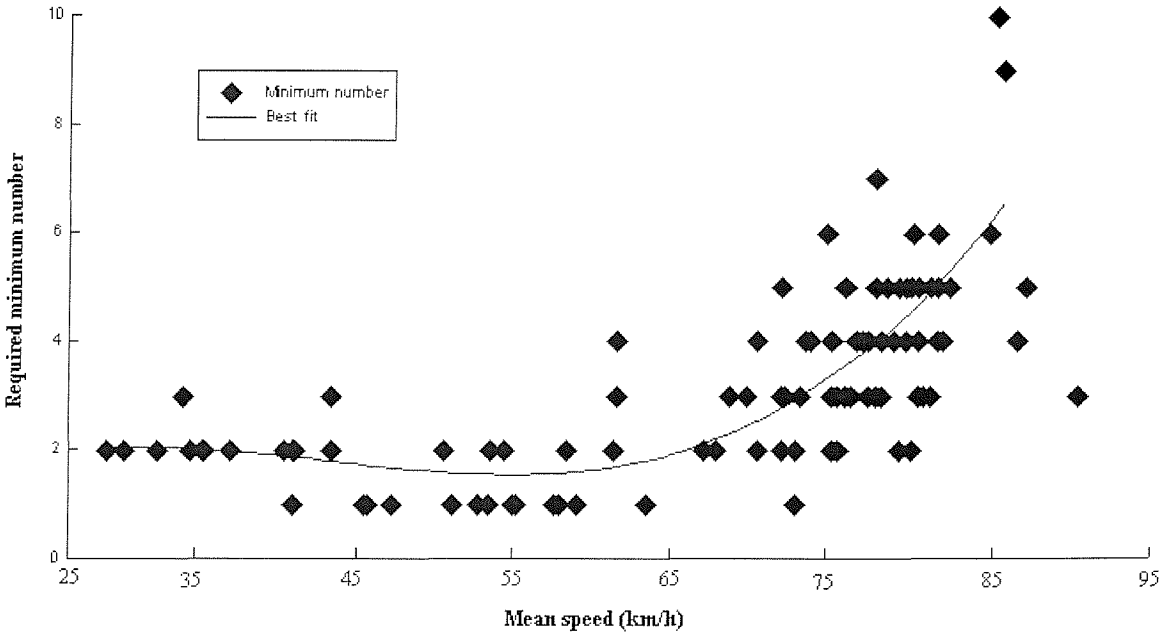


Figure 5.9 Relationship between the mean link speed and required minimum sample size of probe vehicles for link 3

For those time intervals with non-normal journey time distribution (as shown in Figure 5.6), i.e. flow achieved the highest level, as shown in Figure 5.5, the required sample size should be calculated for each minute of a five-minute interval. For the time interval shown in Figure 5.7, two probe vehicles are required for each minute and in total, 10 vehicles are required for the five-minute intervals, which is generally higher than under other circumstances.

Although the range of probe vehicle sample size is link-specific, the general shape of the relationship between mean speed and the required probe vehicle sample size was

found to be similar for all the seven links studied. The following comments may be made to describe the relationship of sample size and traffic condition:

(i) The required minimum sample size remains stable in most cases.

(ii) A larger number of probe vehicles is required when mean travel speed is high, i.e. more probe vehicles are needed for light traffic. This is because of the variability in desired speeds.

(iii) The largest sample size are required in the period between stable and unstable flow, when journey times are increasing or decreasing rapidly and are non-normally distributed.

(iv) In heavily congested traffic, i.e. cases with very low mean speeds, relatively fewer probe vehicles are required. In this case, however, traffic flow is relatively low, i.e. at medium level, as shown in Figure 5.5. This is in agreement with simulation results by Chen and Chien (2000) that the smallest sample size is required at the medium flow level (see Figure 5.2).

5.5.2 Sample size and traffic flow

As shown in Figure 5.4, the same average speed can be observed with different values of traffic flow. Therefore, the relationship between average speed and probe vehicle sample size may differ from the relationship between traffic flow and probe vehicle sample size. Furthermore, since a traditional parameter to describe sample size of probe vehicles is as a percentage of total traffic carried on a link during a period, it is necessary to study the relationship between traffic flow and probe vehicle sample size. Data from link 7 during morning peak hours on the 9th July 2001 was taken as an example to study the relationship between traffic flow and probe vehicle sample size. Data from the non-normal periods in the morning peak (as shown in Figure 5.6) have been removed. The resulting sample sizes of probe vehicles, represented by percentage of traffic flow with traffic flow data, are shown in Figure 5.10. The relationship between traffic flow and probe vehicle sample size can be described by a power regression equation that determines probe vehicle sample size using traffic flow data:

$$n_f = 64210q^{-1.8514} \qquad (5.5)$$
$$r^2 = 0.86$$

where q denotes traffic flow (vehicle/5 minute), n_f denotes probe vehicle sample size, represented as a percentage of traffic flow, and r^2 denotes the square of the correlation coefficient of observed and estimated values by the equation.

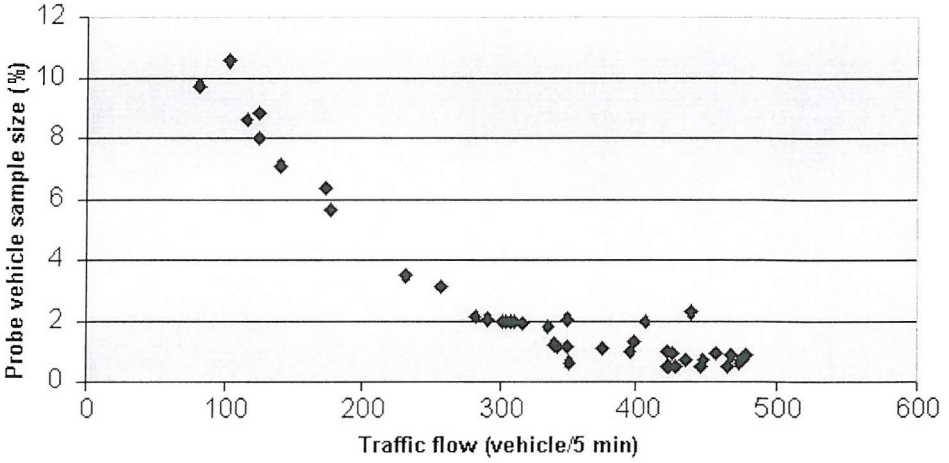


Figure 5.10 Traffic flow and probe vehicle sample size

5.5.3 Probe vehicle sample size for different links

Previous research (Turner and Holdener 1995, Chen and Chien 2000) has shown that sample size of probe vehicles is link-specific, and factors affecting sample size can include geometric conditions as well as traffic flows on a link. Since the seven links studied in this research have different lengths and carry different traffic flows, probe vehicle sample sizes for each of the seven links were calculated to study these factors. Sample size of probe vehicles on each link have been calculated for 84 intervals, i.e. 42 intervals in the morning peak on May 23, 2001 (Wednesday), and 42 intervals in the evening peak on 30 Oct. 2002 (Monday). In the two time periods, there was no non-recurrent congestion on the seven links, i.e. traffic flow was as usually expected. Box plots of resulting sample sizes for 10% permitted error and 95% confidence level are shown in Figure 5.11.

Larger inter-quartile ranges of sample sizes were found on link 1 and link 2 compared to others links (Figure 5.11). Since these two links are shorter, length is considered to be an important factor, in that, with the same level of traffic, a short link requires more probe vehicles. As discussed above in 5.5.2, an increase in traffic flow leads to a reduced probe vehicle sample size. Therefore, busier links require fewer probe vehicles. Link 3 which is the longest and busiest link requires the smallest sample size.

Sample sizes of probe vehicles for each link have been calculated with confidence levels of 90% and 95%. Mean sample sizes and corresponding 85% percentile values are shown in Table 5.4.

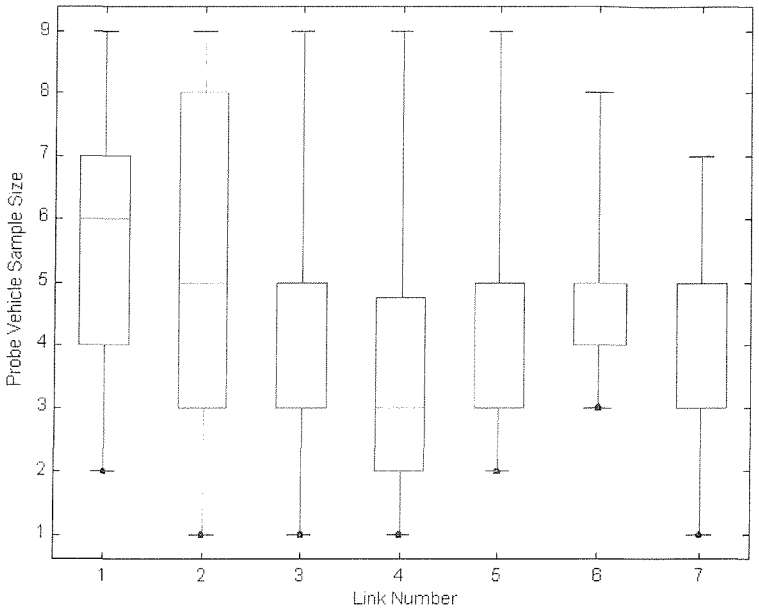


Figure 5.11 Probe vehicle sample size for different links

Table 5.4 Sample size for each link

Link Number	90% confidence level		95% confidence level	
	Mean Sample Size	85% Percentile Sample Size	Mean Sample Size	85% Percentile Sample Size
1	5	6	6	8
2	4	6	6	8
3	4	4	4	5
4	3	4	4	6
5	4	5	5	6
6	4	4	5	6
7	3	4	4	6

Shorter links were more sensitive to changes in traffic flow and Equation 5.5 is unsuitable for links shorter than link 3. However, the general shape of the relationship between traffic flow and probe vehicle sample size was found to be similar for all the seven links. However, shorter links required slightly more probe vehicles for the same traffic flow. The power equation can still be used to estimate required percentage of probe vehicles but needs an adaptive factor of link length:

$$n_f = 64210 \times (1 + \frac{L_3}{L} \times 0.05) q^{-1.8514} \qquad (5.6)$$

where q denotes traffic flow (vehicle/5 min), n_f denotes probe vehicle sample size, represented by percentage of traffic flow, L_3 denotes the length of link 3 and L denotes the length of the calculated link. The Travel Time Data Collection Handbook (Turner et. al., 1998) recommends the segment length for travel time data for freeways/expressways should be in a general range of 1.6 – 4.8 km (1 to 3 mile). Equation 5.6 is suitable for a link whose length is in this range. For instance, the sample size on link 2 for a permitted error of 10% and confidence level of 95% can be estimated based on traffic flow by:

$$n_f = 64210 \times (1 + \frac{4.51}{2.11} \times 0.05) q^{-1.8514} = 71072 q^{-1.8514} \qquad (5.7)$$

A comparison of required sample sizes calculated from ANPR data and estimated using Equation 5.7 is shown in Figure 5.12.

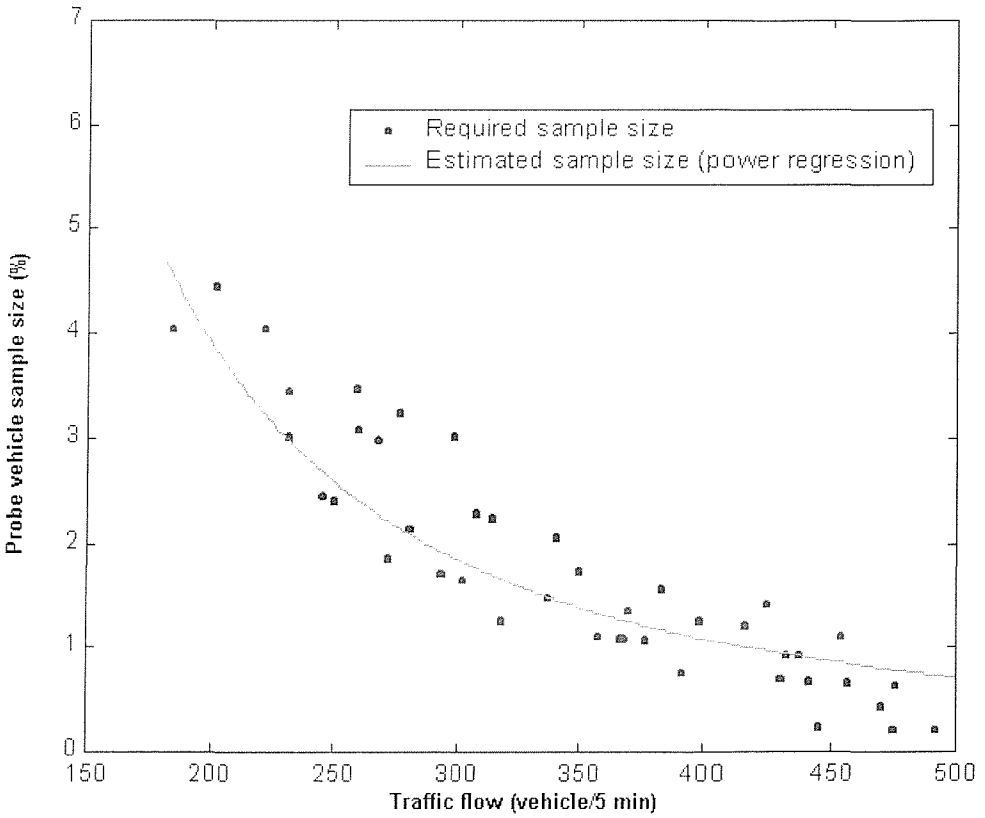


Figure 5.12 Required sample size calculated from ANPR data and estimated using power regressions of traffic flow

It can be found that for the same link, different number of probe vehicles may be required under the same traffic flow. However the sample size of probe vehicles estimated by Equation 5.6 can generally represent average sample size required for various traffic flow. Therefore, the Equation 5.6 is considered to be useful in early

planning stages of probe vehicle deployment. For developing and implementing a data collection plan of a route, the route should be sub-divided into links with the length in the range 1.6 – 4.8 km. A traffic flow survey for different time interval is needed and the flow data can be applied to the Equation 5.6 for estimating required sample size. An alternative way is using video cameras to record journey times of individual vehicles on a link and calculate the parameters of journey time distribution, applying Equation 5.2 to estimate the sample size of probe vehicles based on desirable permitted error and confidence level.

5.5.4 Sample size for non-normality period

As discussed above in Section 5.4.2, the journey time distribution over a 5-minute period will not be normal when the mean of journey time varies greatly during the 5 minutes. In that case, the analysis period should be shortened, for example, using 1-minute period instead. For the period described in Figure 5.7, the sample size of probe vehicles required for each minute is shown in Table 5.5.

Table 5.5 Minimum number of probe vehicles for 1-minute interval

Time	Minimum required sample size
7:50	2
7:51	2
7:52	1
7:53	1
7:54	1

For the 5-min period, 7 probe vehicles would be required to estimate mean link journey time. However, the sampling is dependent, e.g. 7 vehicles should be sampled from different sub-intervals according to Table 5.4. In this case, shorter intervals to update journey time information is recommended, and journey times at intervals with inadequate probe vehicles should be estimated by data from adjacent intervals.

In practice, non-normality cannot be identified with a small number of probe vehicles. Thus, when resulting mean journey time estimated from probe vehicles at an interval differs greatly from estimates at previous intervals, a shorter update period should be adopted.

5.6 Summary

This chapter has used individual journey times logged by video cameras to study journey time distribution. It has been generally shown to be true that individual journey times are normally distributed, although non-normality was observed in some cases. Individual journey times were most likely to be non-normally distributed when average journey time changed rapidly and substantially, usually when congestion occurred. In such situations, it would be inappropriate to use a 5-minute time interval to express journey time. Thus, shorter intervals and a different sampling strategy should be adopted.

This chapter studied the relationship between average speed and required probe vehicle sample size. In most cases, changes in mean and standard deviation values of individual journey times on the same link over a 5-min interval were found to be in a narrow range, and the required sample size for the same link remained stable. However, in congested traffic, i.e. very low mean speeds, fewer probe vehicles were needed, whilst in light traffic, an extremely large sample size of probe vehicles may be required. By using the percentage of traffic flow as parameter to describe the required sample size of probe vehicles, the relationship between traffic flow and the required probe vehicle sample size can be represented by a power regression equation. The equation is useful in early planning stages of probe vehicle deployment.

It is not surprising that different links require different numbers of probe vehicles for a desirable statistical accuracy. Factors affecting probe vehicle sample size include link length and traffic condition on the link. The results indicated that shorter links require a larger number of probe vehicles. With a similar length, a link that carries heavier traffic requires a smaller sample size of probe vehicles. The sample size will also be determined by the use to which it is put. If traffic flows are low and there is no congestion, an accurate journey time may not be necessary with for management proposes or as information to drivers, who can select their own desired speeds.

Chapter Six

Incident Detection

6.1 Introduction

Incidents are responsible for a significant proportion of delays and costs to the motoring public. In addition to the duration of an incident, an incident history can be considered as four critical phases (shown in Figure 6.1). It is estimated that peak-period incidents are responsible for more delay than recurrent peak-period congestion.

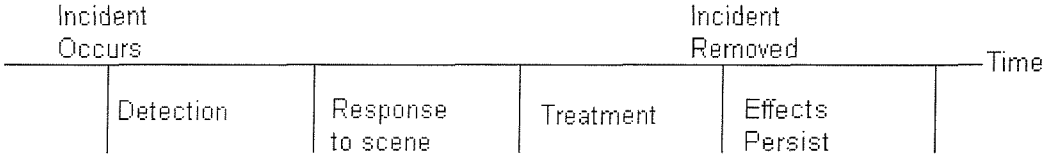


Figure 6.1 Phases of a traffic incident

Source: Highway Capacity Manual (1994)

Although incidents have been discussed widely, their complexity has meant that there is no common agreement on the definition of an incident. In this research, an incident has been defined as "any non-recurrent event which causes reduction of roadway capacity or abnormal increase in demand" (Weil et al., 1998). According to the magnitude of the impact, incidents can be divided into two categories: major and minor (see Table 6.1). The majority of incidents are minor, such as flat tyres, overheating and out of fuel. Minor incidents will, in general, only result in a vehicle being parked on the hard shoulder. Although a minor incident lasts less than half hour, 65% of the total delay caused by incidents are attributable to this category. Major incidents contribute 35% of the overall incident caused delays, constituting severe capacity reduction according to how many lanes are blocked or whether there are accompanying injuries associated with the accident. One study of FHWA (1998) showed that an incident removed to the shoulder on a three-lane facility still reduced capacity by one-third; a

single-lane blockage reduced capacity by 50 percent; a two-lane blockage reduced capacity by 79 percent. Table 6.2 summarises the effect of stalls and accidents on capacity for motorway sections with three travel lanes per direction.

Table 6.1 Incident magnitudes

Characteristic	Minor	Major
Duration	< ½ hour	≥ ½ hour
Blockage	Hard shoulder area only	One or more travelled lanes
Contribution to overall Incident-caused delay	65%	35%

Table 6.2 Typical capacity reductions during incident conditions

Type of incident	Number of lanes blocked	Capacity reduction
Accident on shoulder	0	26%
Vehicle stall	1	48%
Non-injury accident	1	50%
Severe accident	2	79%

(Source: FHWA 1998)

Incident detection is the process that brings an incident to the attention of agencies responsible for maintaining traffic flow and safe operations. Impacts of incidents can be reduced through a variety of actions including broadcasting traffic information, ramp restrictions or closure, and alternative route suggestions. Detection of incidents also helps agencies respond more quickly to remove the problem and to warn the oncoming traffic, thereby reducing the danger of secondary incidents. In the MIDAS (Motorway Incident Detection and Automatic Signalling) system on the M1, after detecting the presence of moving or stationary traffic, signals of 50mph advisory speed limit is set upstream of traffic queues. Analyses have found that a net reduction of 18% in personal injury accidents and a raw reduction of 28% in accidents could be attributed to the scheme (McDonald et al., 2000).

There are a number of automatic incident detection (AID) algorithms reported in the literature. The common source of data for AID algorithms is inductive loops, which

provide traffic measures such as occupancy and volume on the roadway. For example, the MIDAS system uses the "High occupancy algorithm" (HIOCC) to process signals from inductive loops spaced at 500m intervals. With the introduction of automatic vehicle location (AVL) and automatic vehicle identification (AIV), probe vehicle based AID algorithms have been developed and most of them are based on large sample sizes of probe vehicles (Mahmassani et al., 1999).

In this study, the sample size of probe vehicles is assumed to achieve only the required minimum number estimated in Chapter 5, and measure of probe vehicles is journey time for a given roadway segment and a time interval. It is expected that an incident can cause significantly higher journey time than normally experienced at that specific time of day. The aim in this chapter is to use the average journey time data measured by probe vehicles for incident detection. Since probe vehicles provide only journey time, an incident cannot be detected unless the incident has caused delay in journey time.

6.2 Literature Review

6.2.1 Methods for evaluating algorithms

Standard evaluation of incident detection has been widely accepted. Three quantitative measures are commonly used to evaluate freeway incident detection algorithms:

- Detection Rate (DR): defined as the ratio of the number of incident cases correctly detected by the algorithm to the total number of incident cases known to have occurred.
- False Alarm Rate (FAR): defined as the ratio of the number of false alarm cases to the total number of applications or decisions made by the algorithm;
- Detection time: defined as the time it takes the algorithm to signal the incident after its occurrence.

Traditionally, an incident detection algorithm is evaluated by detection rate versus false alarm rate curves—where the higher the curve, the better the algorithm. This evaluation has been and continues to be a useful tool for practitioners and traffic operators. DR-FAR curves treat all incidents with equal importance. That is, failing to detect a low-

impact breakdown on the shoulder contributes equally towards the non-DR as missing a major accident that causes hours of congestion. From a practical point of view, this is a fundamental flaw. Petty et. al. (2002) has proposed a new evaluation method of a cost-benefit analysis where cost mimics the real costs of implementing algorithm and benefits from reducing congestion.

6.2.2 Automatic incident detection algorithms

The following data sources are commonly used to detect incidents:

- Roadway detectors
- Video and closed circuit television
- Probe vehicle: Automatic Vehicle Identification (AVI) and Automatic Vehicle Location(AVL)
- Cellular telephone (Emergency Phone) and motorist aid call boxes
- Service patrols and law enforcement

The most common data source for AID algorithms is loops, which provide traffic measures such as occupancy and volume on the roadway. Other roadside sensors, such as ultrasonic transmitter/receivers, microwave transmitter/receivers are also used in incident detection. A recent development has been the application of image processing techniques that use video cameras to monitor a section of the road and detect incidents by pattern recognition. Since probe vehicles provide direct journey times of individual vehicles, existing AID algorithms using probe vehicle data is based on the premise that incidents cause the journey times to increase significantly over the journey time normally experienced at that specific time of day.

6.2.2.1 AID algorithms based on loops

Numerous AID algorithms have been developed and reported in the literature. Most of these algorithms are based on inductive loops and use the lane occupancy and volume values averaged over a time interval. Based on video image processing, video detectors can also be used for the same purpose. Michalopoulos et. al. (1993) described a machine vision-based algorithm in which a single camera could provide similar traffic measures as an inductive loop in multi-locations within the camera's field of view, thus

replacing many loops. With the development of image processing technology, video detectors can now provide more traffic information, such as vehicle journey time and stopped vehicle detection (Sachse, 2002).

Stephanedes et al. (1992) and Sethi et al. (1995) divided existing AID algorithms into two broad categories: comparative (or pattern recognition) algorithms and time-series algorithms. Comparative algorithms rely on the principle that an incident is likely to increase significantly occupancy upstream while reducing the occupancy downstream. Current values of occupancy are compared against preselected thresholds to detect an incident. Time series algorithms forecast short-term traffic flow based on observed traffic flow in previous time intervals. If the predicted flow differs significantly from the observed flow (i.e. the deviation is greater than a predefined threshold), an incident is declared.

The California algorithms are widely known comparative algorithms. As many as 10 variations of these algorithms have been developed since 1970s. All of these algorithms use the lane occupancy values at one or two adjacent stations as input and compare them with preselected thresholds to characterize the state of the traffic flow. In the original California algorithm, California algorithm #1, traffic flow is characterized as either incident or incident-free states based on three simple comparisons to preselected thresholds. An incident is detected when upstream occupancy is significantly higher than downstream occupancy (Test 1), and upstream occupancy has increased during the past 2 min (Test 2) as well as downstream occupancy has adequately decreased during the past 2 min (Test 3). Test 3 distinguishes an incident from a bottleneck situation by indicating that a reduction in downstream occupancy has occurred over a short period of time as a result of the incident (Mahmassani et al., 1999).

Later California algorithms extended this simple logic by increasing the number of logic decisions made and the number of traffic flow states reported by the algorithm. California algorithm #8 uses both temporal and spatial occupancy values as input and classifies traffic into five states: incident-free, compression wave, tentative incident, incident confirmed, and incident in progress. Since the algorithm suppresses the signalling of an incident for a specified number of time periods after a compression

wave is detected, it reports an incident only after the incident condition has persisted for a while. The algorithm needs six parameters, presented in Table 6.3, for calibration. Five of them (P_1 to P_5) are thresholds for occupancy-based values, while parameter P_6 specifies the number of time periods the algorithm will wait for a compression wave condition to persist before signalling an incident.

Table 6.3 Definition of Parameters Used in California Algorithm #8

Parameter	Definition
P_1	Threshold of occupancy difference between consecutive stations
P_2	Threshold of percent occupancy change at downstream station
P_3	Threshold of percent occupancy difference between consecutive stations
P_4	Threshold 1 of occupancy at downstream station
P_5	Threshold 2 of occupancy at downstream station
P_6	Number of compression wave suppression periods

(Souce: Karim and Adeli 2002)

Time-series algorithms employ a time-series model to provide short-term forecasts. Significant deviation between observed and forecast values are attributed to incidents. The first three algorithms in this class, the standard normal deviation algorithms, calculate the mean and standard deviation of occupancy for the last 3 to 5 minutes and detect an incident when the present value differs significantly from the mean in units of standard deviation.

Ahmed and Cook (1982) developed the autogressive integrated moving average (ARIMA) algorithm, in which an ARIMA model provides short-term forecasts of occupancy and associated 95 percent confidence limits. An incident is detected when the observed occupancy values lies outside the confidence limits.

Unlike the other algorithms that use mainly occupancy data, the McMaster algorithm is based on a two-dimensional analysis of the traffic data. It proposes separating the flow-occupancy diagram into four areas corresponding to different states of traffic

conditions. Incidents are detected after observing specific changes of the traffic state in a short time period (Mahmassani et al., 1999).

Besides the previous approaches, which use aggregate traffic data average over 30 to 60 sec, Collins et al. developed the HIOCC algorithm at the UK Transportation Research Laboratory (TRL) on the basis of one-second instantaneous occupancy data. The algorithm seeks several consecutive seconds of high detector occupancy in order to identify the presence of stationary or slow-moving vehicles over individual detectors (Abou-Rahme et. al., 2000). A weakness of this method is the lack of an effort to distinguish incidents from other congestion-produced traffic phenomena (Michalopoulos, et. al., 1993).

AID algorithms also take advantage of insights gained from research in traffic flow modelling. Willsky et al. (1980) proposed using macroscopic traffic modelling to describe the evolution of spatial average traffic variables (speed, flow and density), thus capturing the dynamic aspect of the traffic phenomena to alleviate the false alarm problem.

Stephanedes et al. (1992) concluded certain limitations of the above algorithms. The limitations result mainly from two sources: (a) the use of row data with only limited filtering and (b) the lack of effort, or effectiveness of effort, in distinguishing incidents from incident-like traffic situation.

Over the last decade, the development of incident detection algorithms has gained an advantage over conventional techniques by emergent mathematical tools, such as neural network, fuzzy logic and wavelet transform (Cheu and Ritchie, 1995; Sachse, 2002; Adeli and Samant, 2000). Although still using flow speed and occupancy data from loop detectors or video cameras, a number of new algorithms have been developed based on new techniques.

Neural networks use parallel and distributed information processing structures that mimic the simplified operation of a human brain. Consequently, neural networks are capable of performing a non-linear mapping between inputs and outputs. For example, associating patterns in traffic data with various traffic conditions. Cheu and Ritchie

(1995) employed three types of neural network models for freeway incident detection. The models were developed and tested with simulation and field data from a study site. Test results with simulation and field data have shown that the neural network models had lower false alarm rates and lower detection rates than California #8. Newly developed algorithms of neural network have been applied rapidly to AID algorithms. Dia and Rose (1997) investigated a multi-layer feedforward neural network AID model using speed, flow and occupancy data. Jin et al. (2002) developed a constructive probabilistic neural network AID algorithm based on a mixture of Gaussian models.

Fuzzy logic is a branch of mathematics that allows the introduction of a quantifiable degree of uncertainty into the modelled process in order to reflect 'natural' or subjective perception of real variables. The way fuzzy logic works is through the use of fuzzy sets, which are different from traditional sets, which can be described as 'yes/no' or 'black/white'. Traditional sets impose rigid membership requirements as an object is in the set or not. In contrast fuzzy sets have more flexible membership requirements that allow for partial membership in a set. A fuzzy logic AID model has been applied to a motorway control system on the N2/N3 motorway near Basel, Switzerland (Sachse 2002). Good progress has been made with the fuzzy approach with decreasing detection time and rate of false alarms. The fuzzy membership functions and key parameters can be determined with neural networks and expert systems. Lin and Chang (1998) proposed a fuzzy-expert system, which functioned to detect not only the occurrence of incidents, but also their located lanes and the resulting type of severity. With such information, the traffic control centre can better advise drivers to make necessary lane changes and take timely actions to minimise the impacts of incident on traffic conditions.

Recently, researchers have investigated discrete wavelet transformation-based incident detection algorithms. Karim and Adeli (2002) created a fuzzy-wavelet Radial Basis Function Neural Network (RBFNN) freeway incident detection algorithm. The algorithm was a single-station pattern-based freeway incident detection model. The characterising pattern used was a time-series of upstream lane occupancy and speed. Wavelet-based de-noising, fuzzy clustering, and neural network classification were used to reliably identify incident and non-incident conditions from the time-series pattern. Using the algorithm, a zero false alarm rate was reported.

6.2.2.2 AID algorithms based on probe vehicles

Sethi et. al. (1995) developed a probe vehicle algorithm based on two cases: using large numbers of probe vehicles and using fewer probe vehicles. The algorithm compared current journey time reports to historical average journey times for the corresponding link, day type, and time period to infer the presence of incidents. The primary effect of reducing the number of probe vehicles was to increase the critical journey time ratio (the ratio of current to historical journey time above which an incident is declared). This increase in the critical value is shown in Table 6.4.

Table 6.4 Critical journey time ratios for declaring incidents

Number of probe reports	Journey time ratio
1	3.50
2	3.45
3, 4	2.80
5, 6, 7	2.60
8, ..., 15	2.40
15, 16, ...	1.45

(Source: Sethi et al, 1995)

Balke et al. (1996) reported an algorithm which used probe- measured average journey time for every 15-min interval. The algorithm was developed using the statistical principle of *standard normal deviate* (SND). The standard normal deviate for the normal distribution is referred to as the Z value. The SND principle, when used with journey times, compares a current probe-provided journey time with the expected journey time derived from historical data using the following formula:

$$SND = \frac{JT - \overline{JT}}{s} \qquad (6.1)$$

where JT = journey time measured by the probe vehicle at a given time;

$\overline{JT}$ = historical journey time on link for a given time interval of day; and

s = standard deviation about the historical journey time for the given time interval of day.

Essentially, the SND establishes confidence intervals of historical journey time. Reorganising Equation 6.1, the SND algorithm takes the form:

$$JT = \overline{JT} + (SND)(s) \qquad (6.2)$$

Using this formulation, an incident alarm is declared if reported journey time x exceeds the confidence interval of $\overline{JT} + (SND)(s)$. The algorithm provided a mechanism for detecting when a probe-measured journey time was outside a range that could be considered typical for the link at a specific time of day and day of the week. If journey times are normally distributed, SND value of 2 means that 97.72% of the distribution can be expected to fall within the interval defined by the SND value. Five selected critical SND values and corresponding percentage values are list in Table 6.5:

Table 6.5 Critical SND values and corresponding percentage values

SND	Percentage Value
2.0	97.2%
2.5	99.38%
3.0	99.87%
3.5	99.98%
4.0	99.9968

Hellinga and Knapp (2000) described three time-series algorithms using AVI data: the confidence limit algorithm, the speed and confidence limit algorithm and the dual confidence limit algorithm. The three algorithms use 20-s time interval for journey time reports that is equal to fifteen probe vehicles for a 5-minute interval. The premise for all three models is that the journey time experienced by vehicles over a section of roadway increases more rapidly as a result of a change in capacity than as a result of a change in demand. Therefore, each of these algorithms attempts to characterise the mean and variance of the journey times associated with the traffic conditions before and after an incident. When an incident occurs, the statistical characteristics of journey times change. Thus, the journey times before and after an incident can be thought of as belonging to different populations. The algorithms attempt to determine if reported journey times are outside of the confidence limits associated with the current population, and if so, it is assumed that an incident has occurred.

6.2.3 Summary for AID algorithms

Traditionally, AID algorithms have been divided into two broad categories: comparative (or pattern recognition) algorithms and time-series algorithms. Comparative algorithms compare traffic measurements upstream with those downstream to identify an incident while time series algorithms compare current measurements with observed data in previous time periods. However, there are now pattern-based time-series AID algorithms which use principles of both comparative and time-series algorithms (Adeli and Karim, 2000). In the last 10 years, a number of modern mathematical tools, such as neural network, fuzzy logic and wevelet transformation, have been applied in AID algorithms. The algorithms using one or more modern mathematical techniques rely on highly complex computation models rather than traditional mathematical functions. This characteristic discriminates these algorithms from traditional algorithms. Therefore, the author considers AID algorithms as falling into two new categories: classic algorithms and modern algorithms. Classic algorithms are wholly model-based approaches using statistical principles for characterising incident and non-incident traffic conditions. In general, modern algorithms are computationally intensive algorithms, which may require high cost for software.

6.3 Methodology

6.3.1 Incident Data

Logs of incident descriptions provided by the Highway Agency and Hampshire ROMANSE Office were used for model evaluation. Information entered for an incident reported included the following:

- Date and time the incident was reported
- Location of the incident
- Type of incident (whether it was an accident, a stalled vehicle, etc.)
- Brief description of the incident and the resulting traffic conditions.

The incident logs contained firsthand accounts of accidents, stalled vehicles, and other capacity-reducing events. However, since the logs were not always involved in every incident on the motorway area studied, it was difficult to obtain real values of the false

alarm rate. Furthermore, the exact times of occurrence of incidents were unknown, therefore detection time could not be determined in this research. Since incidents reported in the log occurred only on link 1, link 2, link 3 and link 5, only these four links were studied for AID algorithms. In addition, since major incidents on the four links occurred only during morning peak periods, morning peak hours of 6:00-9:30 were used for model development and validation.

6.3.2 Model Input

In this research, an AID model has been developed using average journey time measured by probe vehicles. Since sufficient probe vehicles were not deployed on the survey site, individual vehicles recorded by Automatic Number Plate Recognition (ANPR) cameras were randomly selected to simulate 'probe vehicles'. The number of probe vehicles for each link was assumed to achieve only the required sample size estimated in Chapter 5. As discussed in Chapter 5, in a 5-min interval, even for the same link, different sample size may be required in different traffic. However, a uniform sample size for each link was used to simplify computation of model input. The sample size used is the mean sample size estimated for 95% confidence level and 10% permitted error, as shown in Table 6.6.

Table 6.6 Required number of probe vehicles for each link

Link number	Required number of probe vehicles
1	6
2	6
3	4
4	5

For a 5-min period, the required number of individual vehicles was randomly selected from the ANPR database and average journey times of the selected vehicles were used as the model input. It was assumed that there were the required minimum number of probe vehicles to estimate journey times in each 5-min interval in the morning peak hours of 6:00 - 9:30. Therefore, in every morning, there were totally 42 estimates of average link journey time, denoted by JT_i, $i=1,2,...,42$. For example, JT_1 denoted the link journey time of 6:00-6:05 on the given segment, and JT_{42} denoted the link journey time of 9:25-9:30.

Duration of an incident can be affected by many uncertain factors, depending on the incident characteristics, severity etc. The incident data collected in this study has shown that the average duration of an incident was about one hour. Therefore, once an incident was declared, it was assumed to last one hour and the detection task was resumed after the assumed incident duration had elapsed. Of all incident data collected in this study, the longest duration of incident was 2.5 hours. Within 1.5 hours after resuming the detection task, if further incident was declared, journey time data from ANPR was used to identify whether it is the same incident. If it was the same incident, i.e. the effect of the incident lasted longer than one hour, the test was considered invalid. If the effect of the incident had been eliminated before and there was no new report of incident, the declaration was considered a 'false alarm'.

6.3.3 Proposed algorithm

Not all motorway incidents impede traffic flow enough to affect journey time on a link. An incident was assumed to affect link journey time if it blocked part or all of the link or resulted in 'rubbernecking' by drivers. Rubbernecking refers to the action of drivers, who pass an incident slowly and observe the incident scene. Since probe vehicles provide only journey time (or mean space speed) on a motorway link and in a time period, an incident can not be detected unless the incident has caused delay in journey time.

The research described in this chapter relates to incident detection using average link journey times measured by probe vehicles. An incident detection model has been developed based on the premise that an incident causes link journey times to increase significantly over the link journey time normally experienced at that specific time of day. Previous algorithms have detected an incident using few probe vehicles, when a real-time observation of journey time exceeded a threshold expected according to historical data. However, since longer link journey time can be obtained also under non-incident congested conditions, using link journey time data for incident detection may be unable to distinguish incidents from other congestion-produced traffic phenomena.

When an incident occurs, the capacity at that location decreases. Journey time over the road segment increases more rapidly as a result of a change in capacity than as a result of a change in demand, i.e. the reduction in capacity that results from the occurrence of an incident causes larger magnitude of increase in link journey time over a time interval than increase of traffic flow does. Therefore, the magnitude of increase in link journey time can be used to characterise incident and non-incident condition. An incident can be characterised by testing: (1) How much does link journey time increase? (2) How rapidly does link journey time increase?

Two variables were used to characterise an incident: link journey time and *difference in journey time* between two adjacent time intervals. Bivariate analyses were used to study statistical characteristics of two variables as well as their relationship in incident and non-incident traffic. A BivariatE Analysis Model (BEAM) was developed using two variables. The model identified an incident by comparing current link journey time measured by probe vehicles not only with historical link journey time for the specific time interval of day but also with link journey time obtained in the previous time interval.

6.4 Model development

6.4.1 Variable definition

Two variables are used to describe the characteristics of journey time in incident condition: link journey time JT_i, and the difference of link journey time between two adjacent time intervals, which is defined as:

$$DT_i = JT_i - JT_{i-1} \qquad (6.3)$$

As introduced in Section 6.3.2, with 5-min interval, in every morning there is a total of 42 observations of link journey journey time, denoted by JT_i, $i=1,2,...,42$. Since DT_i expresses the change of link journey time from the previous time interval, there are 41 observations, denoted by DT_i, $i=2,3,...,42$.

With the definitions of the two variables, a time series of link journey time can be transformed from one-dimension data into two-dimension. JT_i and DT_i observed on a

link at a time interval are considered as a bivariate sample. Thus, bivariate analysis method of statistics can be used to characterise journey time in incident and non-incident conditions.

6.4.2 Bivariate Analysis

Data collected in non-incident condition will be shown that the joint distribution of JT_i and DT_i is bivariate normal and detailed analyses are referred to Section 6.4.3.3. The joint distributions of JT_i and DT_i describe not only spreads of JT_i and DT_i but also the relationship between JT_i and DT_i. A detailed description of bivariate normal distributions can be found in Appendix C1. JT_i and DT_i at a time interval in non-incident traffic condition can be considered as one sample point of the bivariate normal distribution, while data observed in incident traffic can be considered as outliers of the distribution. The task of incident detection is therefore one of checking for outliers in bivariate normal distribution.

The bivariate normal density is given by:

$$f(JT_i, DT_i) = c_1^{-1} \exp\left\{-\frac{1}{2}\begin{bmatrix} JT_i - \mu_{JT_i} \\ DT_i - \mu_{DT_i} \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} JT_i - \mu_{JT_i} \\ DT_i - \mu_{DT_i} \end{bmatrix}\right\} \qquad (6.4)$$

where $c_1 = 2\pi\sigma_{JT_i}\sigma_{DT_i}(1-\rho_i^2)^{1/2}$ and the covariance matrix $\Sigma = \begin{bmatrix} \sigma_{JT_i}^2 & \sigma_{JT_iDT_i} \\ \sigma_{JT_iDT_i} & \sigma_{DT_i}^2 \end{bmatrix}$.

A measure of the strength and direction of association between variables is provided by the *correlation coefficient*:

$$\rho_i = \sigma_{JT_iDT_i} / \sigma_{JT_i}\sigma_{DT_i} \qquad (6.5)$$

Therefore, the bivariate normal density contains five parameters $\mu_{JT_i}, \mu_{DT_i}, \sigma_{JT_i}, \sigma_{DT_i}$ and ρ_i. For bivariate normal distribution, elliptical contours can be used to study the distribution of the density over the JT_i - DT_i plane. The equation

$$\begin{bmatrix} JT_i - \mu_{JT_i} \\ DT_i - \mu_{DT_i} \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} JT_i - \mu_{JT_i} \\ DT_i - \mu_{DT_i} \end{bmatrix} = \kappa \qquad (6.6)$$

describes an ellipse in the JT_i-DT_i plane with centre at (μ_{JT_i}, μ_{DT_i}). As κ increases, the area of the ellipse increases. For the bivariate normal density, the constant k on the right-hand side of Equation (6.6) is equal to $\chi^2_{\alpha,2}$. The elliptical contour contains average 100(1-α)% of the sample points. If a sample plot of (JT_i, DT_i) lies in the elliptical contours, the Equation (6.7) should be fulfilled and the sample is considered as non-incident data. Otherwise, when the sample plot lies outside the elliptical contours, an incident is declared:

$$\begin{bmatrix} T_i - \mu_{T_i} \\ DT_i - \mu_{DT_i} \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} T_i - \mu_{T_i} \\ DT_i - \mu_{DT_i} \end{bmatrix} \leq \kappa \qquad (6.7)$$

In this research, two elliptical contours containing 99% and 99.9% of the distributions respectively were constructed. Each observation of (JT_i, DT_i) was compared with corresponding elliptical contours to check for outliers (incidents). Figure 6.2 shows observations of 10 days on link 2 and the time interval of 7:40-7:45. Some plots are displayed inside both contours and these indicate non-incident condition. The plot outside the external elliptical boundary would result in an incident being declared. The plot displayed outside the 99% coverage contour but inside the 99.9% coverage contour results in an incident being declared with 99% contour but not with 99.9% contour. The sensitivity of different contours has been studied in the following evaluation.

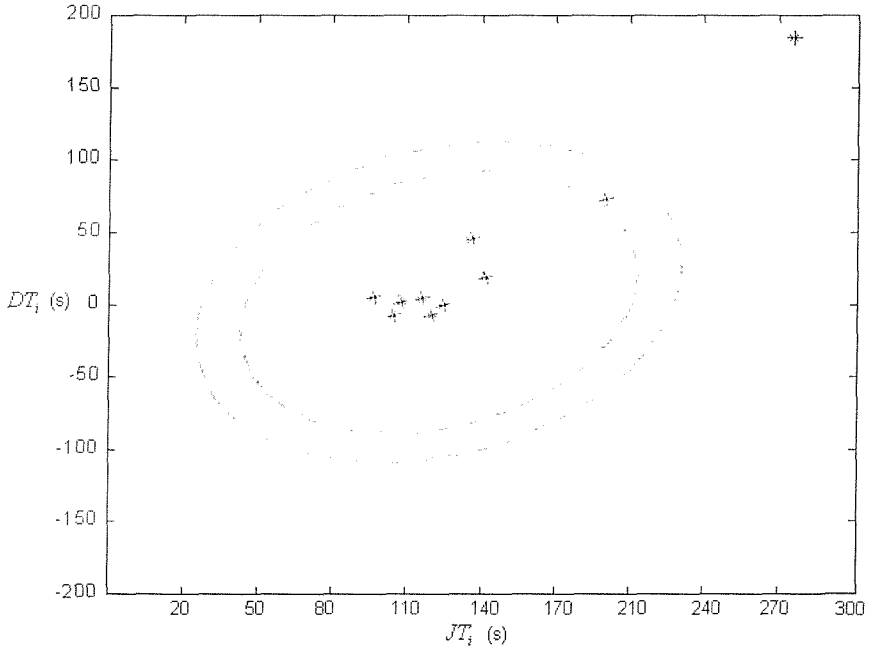


Figure 6.2 99% and 99.9% coverage contours

6.4.3 Parameter Estimation

6.4.3.1 Link Journey Time

Observations of average journey time on the same link for the same interval were considered as the population studied, and average journey times at each of the 15 days, provided by ANPR, were considered as a sample of the population. For example, to estimate the journey time distribution on link 2 at the 23rd time interval (7:45-7:50), observations of JT_{23} at each of the 15 days were used, as shown in Table 6.7.

Table 6.7 Sample of JT_{23}

Day	21/5	22/5	23/5	24/5	25/5	2/9	3/9	4/9	5/9	6/9	16/9	17/9	18/9	20/9	21/9
JT_{23} (s)	91	91	132	88	90	141	94	97	196	166	106	145	101	124	84

The Kolmogorov-Smirnov Test (K-S test) was used to assess and confirm that journey times were distributed normally. The mean and standard deviation of JT_{23} were estimated from the sample:

$\mu_{JT_{23}} = 116s$

$\sigma_{JT_{23}} = 33.4s$

For each of the 42 time intervals in the morning peak hours, JT_i, $i=1,2,...,42$, have been shown to be normally distributed. The normal distributions have time-varying means and standard deviations. Due to the limited sample size, the directly estimated means may include noise and a smoothing process was needed. Figure 6.3 shows the smoothed mean of JT_i over the morning peak period. Smoothed standard deviations of JT_i at each time interval are shown in Figure 6.4. The results indicated that larger values of standard deviation are most likely to be obtained at busier time, i.e. 7:30-8:30. That is, average journey time for the same time interval can be very different from day to day. Average journey time and standard deviation discussed here are different from those discussed in Section 5.5.1. In Section 5.5.1, average journey time is the mean journey time of all vehicles on a link in a time interval, and standard deviation is calculated from journey times of all vehicles on the link in the time interval. In this section, however, mean journey time and standard deviation describe

how different average journey times of all vehicles for the same link in the same time interval can be from day to day.

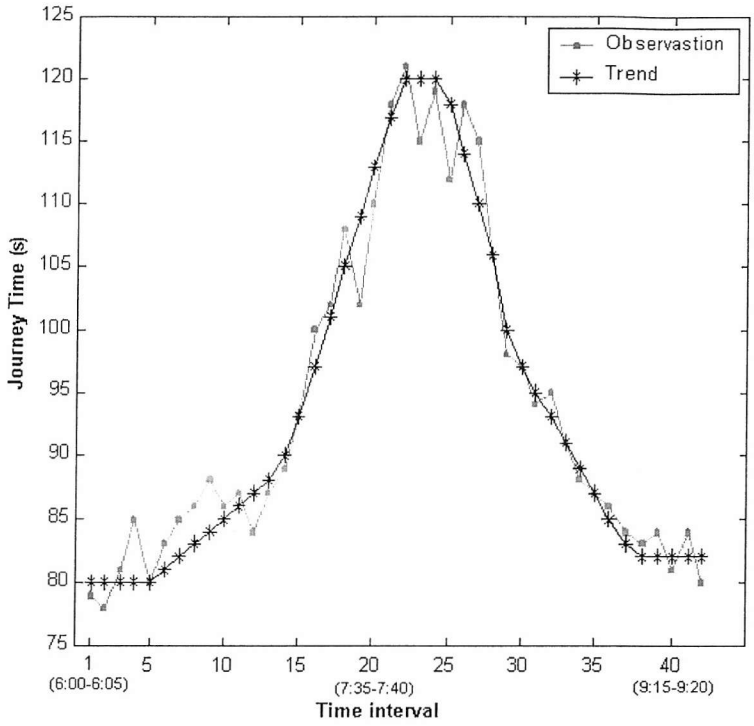


Figure 6.3 Mean of link journey time at each time interval

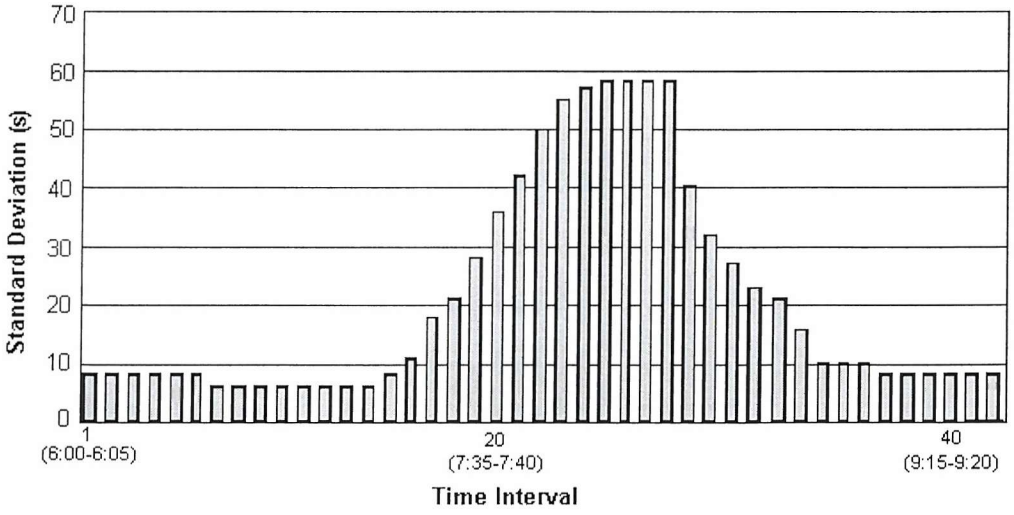


Figure 6.4 Standard deviation of link journey time at each time interval

6.4.3.2 Difference of link journey time between adjacent time intervals

The distribution of DT_i at each time interval was studied using data from the same 15 days. Normal distributions have been shown for each time interval. Mean and standard

deviation of DT_i are shown in Figure 6.5. For time intervals before 6:30 and after 9:00, DT_i is zero due to the stability of journey time in this period. Increasing of journey time from 7:00 to 8:00 gave DT_i positive values, while negative values of DT_i in the period of 8:00-9:00 were obtained due to decreasing of journey time. Similar to JT_i, larger values of standard deviation were obtained in busier time.

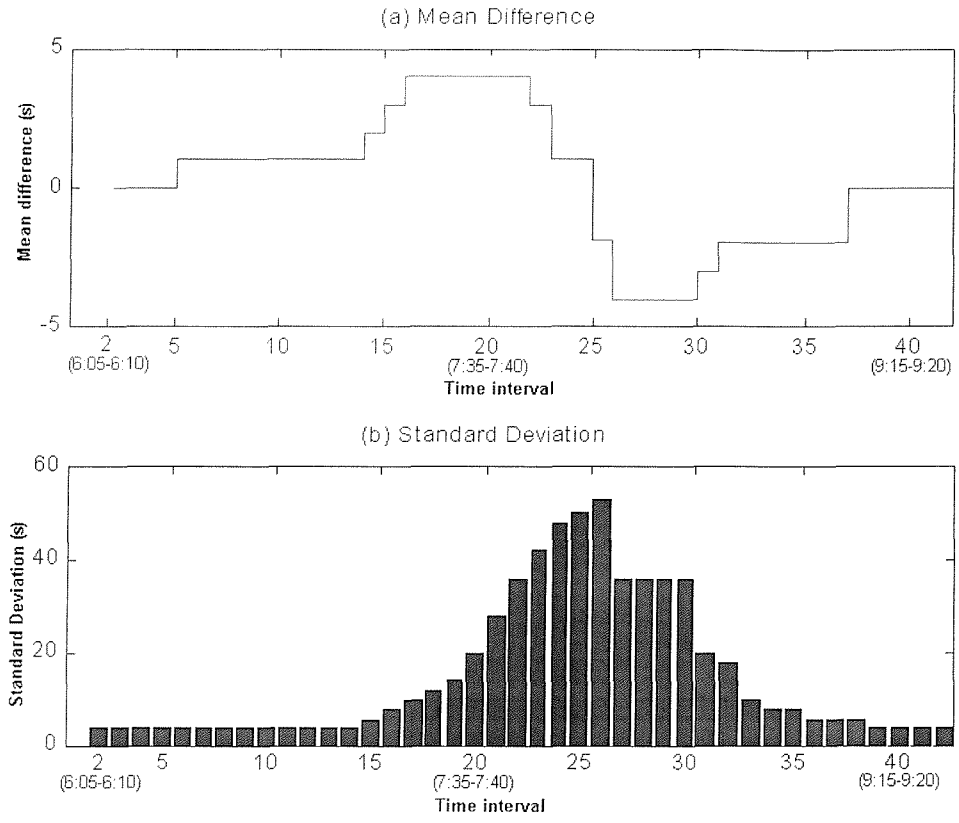


Figure 6.5 Mean and standard deviation of difference of journey time at two adjacent time intervals

6.4.3.3 Correlation Coefficient

Both JT_i and DT_i have been shown to be univariate normal. However, it is not necessarily true that JT_i and DT_i will be bivariate if both JT_i and DT_i are univariate normal. The correlation coefficient ρ_i measures the strength and direction of association between JT_i and DT_i. For the time period when the journey time is stable, i.e. before 6:30 and after 9:00, $\rho_i = 0$, i.e. JT_i and DT_i are independent. In this case,

the joint density can be written as the product of the densities of JT_i and DT_i, and hence the joint density is bivariate normal.

During the time period when the journey time increases or decreases substantially and quickly, large positive values of ρ_i were obtained, indicating a strong positive relationship between JT_i and DT_i. Relatively weak relationship between JT_i and DT_i was obtained in the time intervals when JT_i was increasing or decreasing slightly. The correlation coefficient at each time interval were estimated as the following:

$$\rho_i = \begin{cases} 0 & i \in [2,6] \cup [37,40] \\ 0.2 & i \in [7,12] \cup [32,36] \\ 0.5 & i \in [19,31] \\ 0.8 & i \in [13,18] \end{cases} \tag{6.8}$$

For the time intervals with non-zero correlation, the distributions were determined be bivariate normal by using Mahalanobis Distance. The assessment of bivariate normality is introduced in detail in Appendix C2. After estimation of the five parameters of the bivariate normal distribution, $\mu_{JT_i}, \mu_{DT_i}, \sigma_{JT_i}, \sigma_{DT_i}$ and ρ_i, the bivariate normal density of the distribution could be obtained. The surface of bivariate normal density at the 23rd time interval, i.e. 7:50-7:55, is shown in Figure 6.6. The three dimensional picture shows that the density at (JT_{23}, DT_{23}) is given by the height $f(JT_{23}, DT_{23})$. According to the Equation 6.4, $f(JT_{23}, DT_{23})$ decreases exponentially with the square of the distance between (JT_{23}, DT_{23}) and the centre (μ_{JT_i}, μ_{DT_i}), then the surface looks like a bell. The marginal densities for JT_{23} and DT_{23} are the familiar normal densities:

$$\begin{aligned} f_{JT_{23}}(JT_{23}) &= \frac{1}{\sqrt{2\pi\sigma^2_{JT_{23}}}} \exp[-\frac{1}{2}(\frac{JT_{23} - \mu_{JT_{23}}}{\sigma_{JT_{23}}})^2] \\ f_{DT_{23}}(DT_{23}) &= \frac{1}{\sqrt{2\pi\sigma^2_{DT_{23}}}} \exp[-\frac{1}{2}(\frac{DT_{23} - \mu_{DT_{23}}}{\sigma_{DJT_{23}}})^2] \end{aligned} \tag{6.9}$$

Contours of the surface on the JT_i- DT_i plane are ellipses, as shown in Figure 6.2.

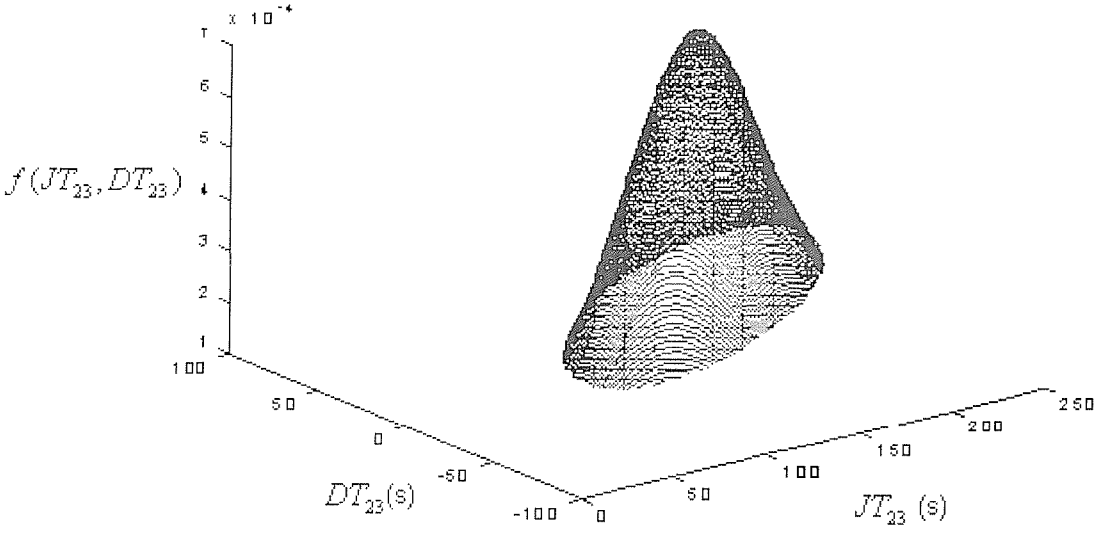


Figure 6.6 Bivariate normal density of JT_i and DT_i

6.5 Model evaluation

In this research, the logs of incident descriptions provided by the Highway Agency and Hampshire ROMANSE Office were used to evaluate model performance. However, since the log data were not always available for every incident on the survey site, the FAR obtained may be higher than the actual value. Furthermore, due to an unknown exact occurrence times of incidents in the logs, detection time could not be evaluated.

6.5.1 Model performance

For each of the four links in the study motorway area, 9 weeks data were used for model evaluation. The evaluation was carried out using 99% and 99.9% coverage contours respectively. The results of the test are summarised in Table 6.8. The best detection rate and false alarm rate were obtained on link 1. This was because journey time was strongly stable on link 1 and an incident could produce more obvious impact. The lowest detection rate was on link 5. However, it may not be associated with carried traffic or geometric conditions of the link. It may be that more minor incidents occurred on link 5. Table 6.9 summarises the results on the four links using different contours.

Table 6.8 Model performance

Contour	Link 1		Link 2		Link 3		Link 5	
	99%	99.9%	99%	99.9%	99%	99.9%	99%	99.9%
Number of Incidents	4		4		5		6	
Number of Incidents Detected	4	4	4	3	5	4	4	4
Number of Tests	1678		1698		1668		1612	
Number of False Alarms	6	4	12	10	17	14	12	10
Detection Rate (DR)	100%	100%	100%	75%	100%	80%	66.7%	66.7%
False Alarm Rate (FAR)	0.36%	0.24%	0.71%	0.59%	1.02%	0.84%	0.74%	0.62%

Table 6.9 Model Performance Summary

	99% Contour	99.9% Contour
Detection Rate (100%)	90.5%	85.7%
False Alarm Rate (100%)	0.71%	0.57%

6.5.2 Performance Analysis

The four links studied carry heavy traffic in morning peak hours. Therefore, an incident occurring on these links can produce significant impact on journey time of any individual vehicle. Consequently, a high detection rate was achieved by probe vehicle algorithm. However, a higher false alarm rate was obtained on the same links. In general, algorithms with probe vehicle journey times have experienced a considerably higher false alarm rate than the other algorithms using loop detector data (Balke et al., 1996). The false alarms are likely to occur when one or two of the probe vehicles have unusually large journey times due to driver' behaviour or vehicle condition.

The length of a link can have an impact on model performance. With relatively short lengths, links 1 and 2 have higher detection rates than the others. That is because an incident may produce more significant increases in journey time on a shorter link. Previous research (Hellinga and Knapp, 2000) has shown that an incident is likely to create congestion upstream and reduce the flow downstream. The decrease in flow downstream of the incident is in turn likely to allow an increase in speed. Therefore, the speeds of vehicles exiting the incident area will increase. In that case, loss of journey time due to a minor incident may be made up over a longer journey. The same

AID algorithm may be more sensitive to an incident for shorter links than for longer links.

6.6 Summary

In this chapter, existing AID algorithms have been reviewed. Based on experiences of previous researchers, the author chose statistical principles to develop an AID model using average journey time measured by probe vehicles. The premise of the model is that link journey time increases more rapidly as a result of a change in capacity (i.e. such as the reduction in capacity that results from the occurrence of an incident) than as a result of a change in demand. In this case, incident conditions were suspected if probe-measured journey time at a time interval was not only significantly higher than normally expected, but also higher than the journey time measured at the previous interval. Two variables, average journey time and journey time difference between two adjacent time intervals, have been used to characterise incident condition. Statistical principles of the bivariate analysis have been used to describe the relationship of the two variables. The joint distribution of the two variables obtained in non-incident traffic is considered to be bivariate normal, while observations of the two variables in incident traffic are considered as outliers of the bivariate normal distribution. The task of the BivariatE Analysis Model (BEAM) is to check observations for outliers. Four links have tested the feasibility of the model with 99% and 99.9% coverage contours respectively. High detection rates and reasonable false alarm rates have been obtained.

Chapter Seven

Journey Time Prediction

7.1 Introduction

It has been shown in Chapter 4 and Chapter 5 that GPS equipped probe vehicles can measure journey time accurately, and only a small sample size is required to estimate mean journey time reliably on a motorway link over a period. However, where there are insufficient probe vehicles to meet this minimum criterion, an alternative approach is to predict journey time. Moreover, although probe vehicles can provide accurate real-time or near real-time journey time, many ITS applications (e.g. in-vehicle navigation systems) need sound estimates of future conditions. Short term prediction of journey times can be particularly important for motorway management/operations and dynamic route planning. Real-time and predictive traveller information has become possible by the convergence of information technology and wireless communications, for example delivering information through the SMS (Short Message Service) function of mobile phones. Individual travellers will have an opportunity to review their travel options every time just before they travel. A report (Karl and Trayford, 2000) found that 33 per cent of peak hour travellers have the opportunity to start their journeys at flexible times and predictive journey time information benefits this group of motorists as they can decide what time and route to take.

This chapter is concerned with predicting motorway link journey times from real-time journey times estimated by probe vehicles and historic travel times. Different models have been developed to forecast journey times for incident and incident-free traffic conditions. For incident-free conditions, the predictive periods will be from 5 to 30 minutes. As discussed in Chapter 6, since the duration of delay caused by an incident may be more than 30 minutes, a longer term prediction is required for incident conditions. In this chapter, journey times for each time interval of one hour after an

incident have been forecast based on impact of the incident, i.e. if an incident occurred at 7:30, journey times at each time interval from 7:30-8:30 should be forecast.

7.2 Literature review

There has been much research into journey time prediction. Examples include time series models (D'Angelo et al., 1999), Kalman Filter Model (Chen and Chien, 2001), simulation models (Chien et al., 2003), and neural network models (Lint et al., 2003; Krikke, 2002). Prediction models can be based on inductive loops (Huisken and Berkurn, 2002;), probe vehicle data (Sun et. al, 2003), and fused data of both sources (Nanthawichit et al., 2003).

Sun et al. (2003) proposed a local linear regression model for short term speed forecasting based on AVI data. The model has been developed based on the assumption that speed data on different days may show similar trends, although the shifts are different on the time axis. Therefore, patterns could be matched after sliding the data in the time dimension.

In Chien and Kuchipudi (2002), a Kalman filter algorithm has used in predicting travel time. The system model was written as:

$$\begin{cases} x(t) = \phi(t-1)x(t-1) + w(t-1) \\ z(t) = x(t) + v(t) \end{cases} \qquad (7.1)$$

where, $x(t)$ denotes the travel time at time interval t that is to be predicted, $w(t)$ denotes a white noise process, $z(t)$ denotes the observation of travel time on time interval t and $v(t)$ denotes the measurement error. The transition parameter $\phi(t)$ was determined using historic data, i.e. travel time from the same time interval of a previous data with a similar traffic condition was used to obtain $\phi(t)$. This assumes that the pattern of journey time variation over time remains similar between two days.

However, data collected on links on the M3 and the M27 have shown very different patterns of journey times from day to day. Figure 7.1 shows journey times on link 2 in the morning peak time on three days: 30 Sept. (Monday), 1 Oct. (Tuesday), and 2 Oct.

(Wednesday), 2002. It can be found that journey time could achieve its highest value at any time between 7:30 and 8:30, in which journey time has larger variance (as shown in Figure 6.3). Therefore, the assumption that the pattern of journey time variation over time remains similar between days may not be supported and the transition parameter $\phi(t)$ obtained from a previous day may not reflect journey time change in another day.

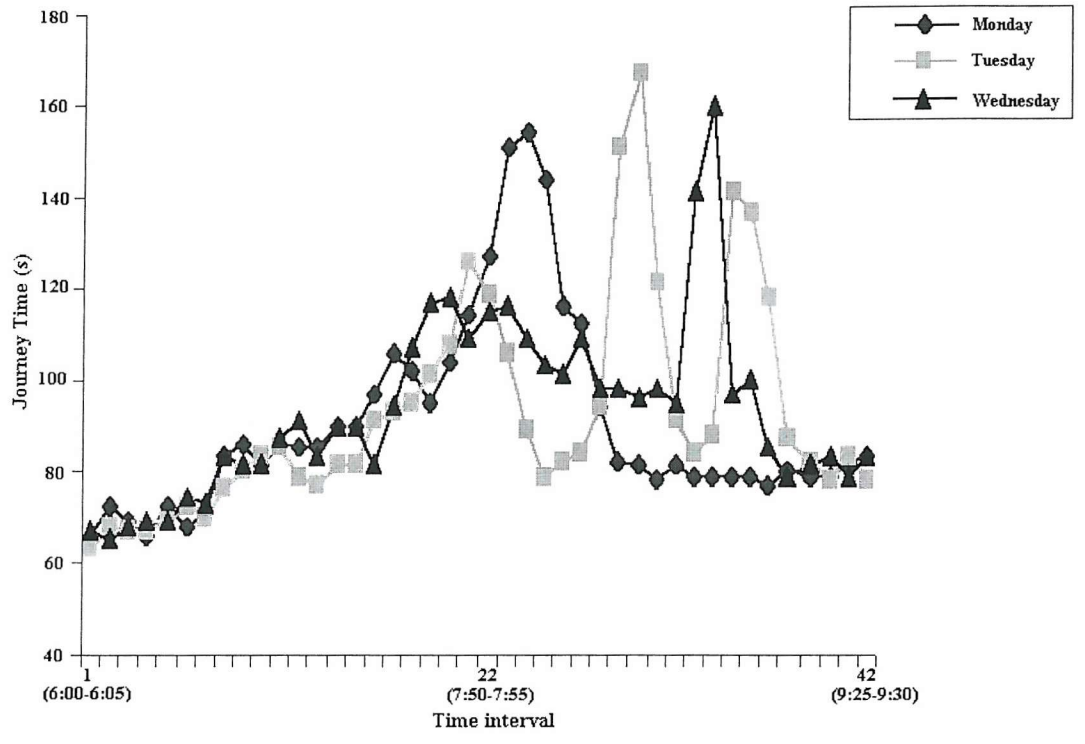


Figure 7.1 Journey time on different days

7.3 Journey time prediction in non-incident condition

7.3.1 One-step-ahead prediction

One-step-ahead prediction in this study is used to forecast average journey time of the next 5-minute interval. As introduced in Section 3.3.2, in the morning peak period of 6:30-9:30, using five minute interval, there are a total of 42 observations of average journey time, denoted by JT_i, $i=1,2, \ldots, 42$. Average journey time is considered to be a discrete time-series. For prediction, average journey time is considered as a first-order autoregressive (Markov) process:

$$JT_{i+1} = a_i JT_i + b_i \qquad (7.2)$$

where b_i is an independent white noise process with mean zero and variance Q_i . The journey time prediction consists of estimating the time-variant mean of a stochastic system.

Analyses in Section 6.4 have shown for the same time interval, average journey times obtained from different days are normally distributed, although mean and variation for each time interval are different (Figure 6.3 and 6.5). Therefore, the relationship between JT_i and JT_{i+1} depends on the parameter *i*. Journey times in each time interval can be transformed to standard normal distributions by the formula:

$$ZJT_i = \frac{JT_i - \mu_i}{\sigma_i} \qquad (7.3)$$

It can be found that the relationship between ZJT_i and ZJT_{i+1} is approximately linear, shown in Figure 7.2.

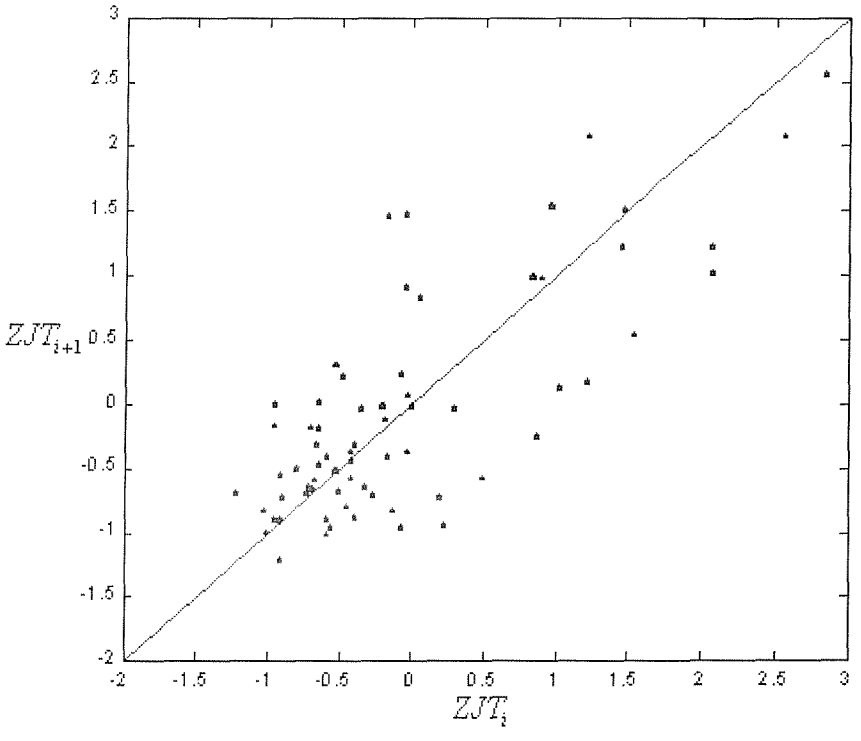


Figure 7.2 Scatterplot of ZJT_i vs. ZJT_{i+1}, *i=18,19, ..., 23*

The linear relationship can be explained by noting that the deviation of journey time from its mean in the current time interval is the same as in the previous time interval. The relationship can be described by the following equation:

$$\frac{JT_{i+1} - \mu_{i+1}}{\sigma_{i+1}} = \frac{JT_i - \mu_1}{\sigma_i} \qquad (7.4)$$

Journey time in the next time interval can be forecast by:

$$JT_{i+1} = \mu_{i+1} + \frac{\sigma_{i+1}}{\sigma_i}(JT_i - \mu_i) \qquad (7.5)$$

An example of the observed and predicted journey times is shown in Figure 7.3. It may be seen that the predicted journey times closely follow the propagation of the observed journey times over the entire morning peak period. For non-congested time slices before 7:30 and after 9:00, the predictions match. Throughout the congested period, journey time predictions take time to adjust to significant increases/decreases.

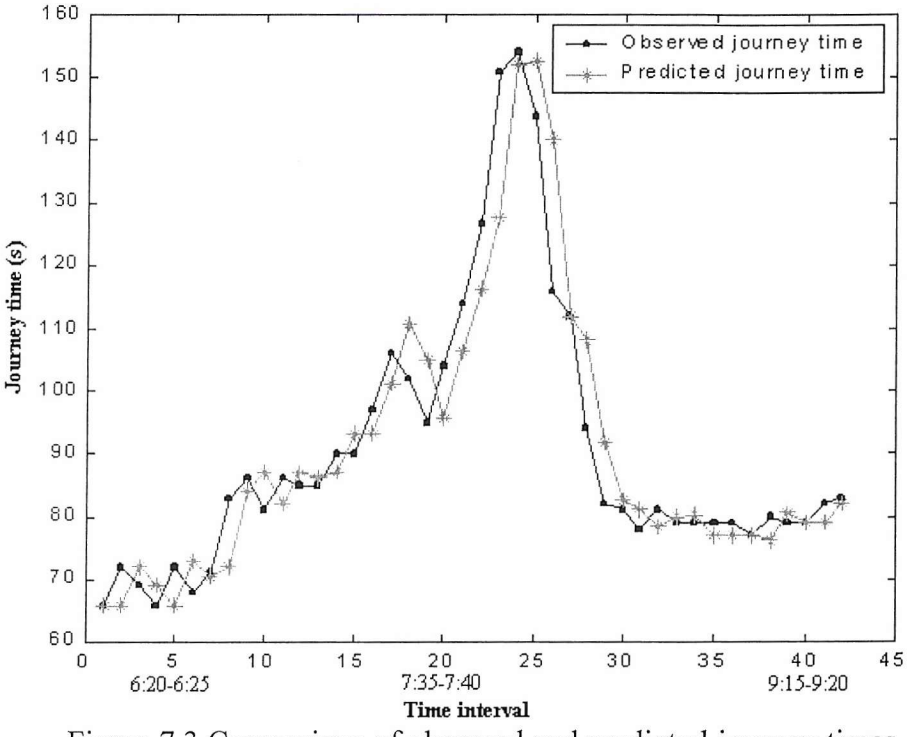


Figure 7.3 Comparison of observed and predicted journey times on link 2, Sept. 30 2002

Prediction accuracy is assessed by *prediction error*, which is estimated by:

$$e_p = \frac{\left\lfloor JT_{predicted} - JT_{observed} \right\rfloor}{JT_{observed}} \qquad (7.6)$$

Equation 7.4 was applied to morning peak period data from 30 Sept. to 4 Oct. 2002 and the calculated prediction errors for the five weekdays are listed in Table 7.1. Over the five days, 90% of all predictions errors of journey times fell within 10%, and the average prediction error was 6.64%. Larger prediction errors occur at the temporal boundaries of congestion due to the property of delay in the linear prediction model.

Table 7.1 One-step-ahead prediction error

Day	Average Prediction Error	Maximum Prediction Error
30/09/02	5.28%	20.69%
01/10/02	9.56%	39.26%
02/10/02	7.86%	60.89%
03/10/02	5.46%	40.67%
04/10/02	5.07%	17.46%

7.3.2 Multi-step-ahead prediction

In practice, the time scales of interest are longer than those applied in one-step-ahead prediction, typically 10 to 30 minutes (Karl and Trayford, 2000). In this study, three prediction cases are considered, 10-minute, 20-minute and 30-minute predictions. Since journey time may change substantially in 10 minutes, use of linear regression equation (Equation 7.3) may result in larger prediction errors, thus a more comprehensive model is required for multi-step-ahead prediction. In the model, the current observation of journey time is still of prime interest but the weight of historic data increases.

For multi-step-ahead prediction, journey time is considered to be generated by a random walk plus drift:

$$JT_i = JT_{i-k} + \beta_i + \eta_i \qquad (7.7)$$

where β_i denotes the difference of mean journey time between the two time intervals:

$$\beta_i = \mu_i - \mu_{i-k} \qquad (7.8)$$

Although parameter β_t is known and constant, β_t is still treated as part of the state variables, by defining the state variable as $\alpha_t = [JT_t \quad \beta_t]^T$. Equation 7.7 can be written in the state space form as:

$$y_t = [1 \quad 0]\alpha_t + \varepsilon_t \tag{7.9}$$

$$\alpha_t = \begin{bmatrix} JT_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} JT_{t-k} \\ \beta_{t-k} \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \end{bmatrix} \tag{7.10}$$

Equation 7.9 is the observation equation and Equation 7.10 is the transition equation. The Kalman filter can be applied to solve the equations in two phases. The first step is to predict the next value of the state and the variance of the state:

$$\hat{\alpha}_{t|t-k} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\hat{\alpha}_{t-k|t-k} \tag{7.11}$$

$$P_{t|t-k} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}P_{t-k|t-k}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T + Q_t \tag{7.12}$$

where Q_t is the covariance matrix of $[\eta_t \quad 0]^T$. As the equations are first order stochastic difference equations, starting values are required. Thus the initial state, α_0, and its covariance matrix, $P_{0|0}$, are assumed to be known.

The second step in the Kalman filter is to update the value of the state by reviewing the current predictions when new information is available:

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-k} + K_t(y_t - [1 \quad 0]\hat{\alpha}_{t|t-k}) \tag{7.13}$$

$$P_{t|t} = (I - K_t[1 \quad 0])P_{t|t-k} \tag{7.14}$$

$$K_t = P_{t|t-k}[1 \quad 0]^T([1 \quad 0]P_{t|t-k}[1 \quad 0]^T + \sigma_t^2)^{-1} \tag{7.15}$$

where σ_t^2 is the variance of ε_t.

Rearranging the terms on the right hand side of Equation 7.13 yields:

$$\hat{\alpha}_{t|t} = K_t y_t + (I - K_t[1 \quad 0])\hat{\alpha}_{t|t-k} \tag{7.16}$$

The updated state in Equation 7.16 is expressed as a weighted average of y_t and $\alpha_{t|t-k}$. In this study, the new observation, y_t, is the historic journey time for the same time

interval. Thus, the Kalman filter model provides an estimate of journey time in the ith time interval from journey time in the $(i-k)th$ time interval and historic data.

Prediction results for the same day predicted in Figure 7.3 are shown in Figure 7.4. As journey times increase, the error of prediction becomes greater, and the level of delay increases with the length of the prediction period. Compared with one-step-ahead prediction, multi-step-ahead predictions are unable to forecast significant change in journey time since the Kalman model adheres to historic data.

The Kalman filter model has been applied to morning peak period data from 30 Sept. to 4 Oct. 2002. Average prediction errors for various prediction periods are summarised in Figure 7.5. It can be seen that the prediction error increases with the length of the prediction time period. With the prediction time of 20 minutes, the error approached 10%. For the prediction time of 30 minutes, average prediction error of 11% is obtained.

7.4 Journey time prediction in incident conditions

Journey time prediction in incident-free conditions is based on current observations and historic data, and the prediction period is from 5 minutes to 30 minutes. However, in incident conditions, journey time will have very different characteristics and those prediction models developed in Section 7.2 are unable to predict journey times in incident conditions. Moreover, after an incident is identified, traffic operators and travellers may want to know the impacts of the incident, such as duration of the incident. Therefore, a prediction model should aim to forecast the duration as well as journey times during the period, and a different prediction strategy is required.

Analyses in Chapter 6 have shown that an incident may cause a considerable increase in journey time and, by studying the increasing magnitude, an incident can be identified. A typical example of changes in journey times is shown in Figure 7.6. This displays an incident on link 2 in the morning peak period on the 11th Oct. 2002.

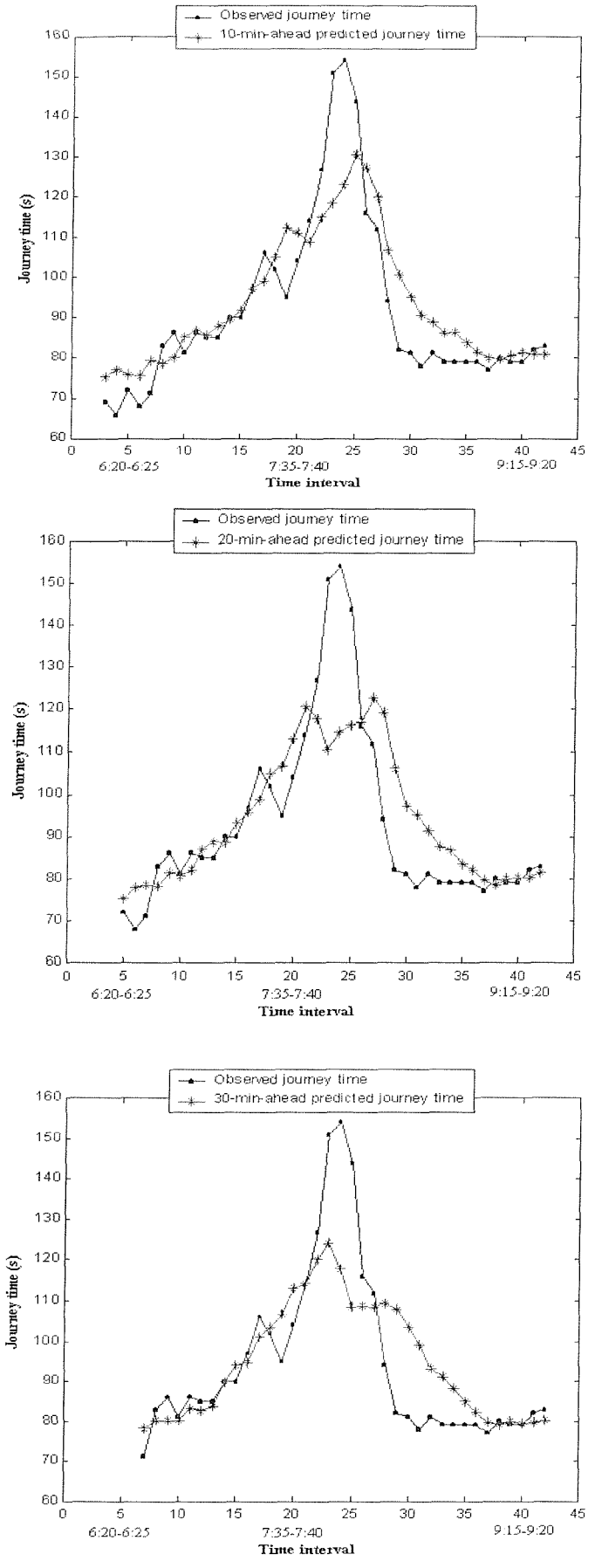


Figure 7.4 Multi-step-ahead journey time predictions on link 2, Sept. 30 2002

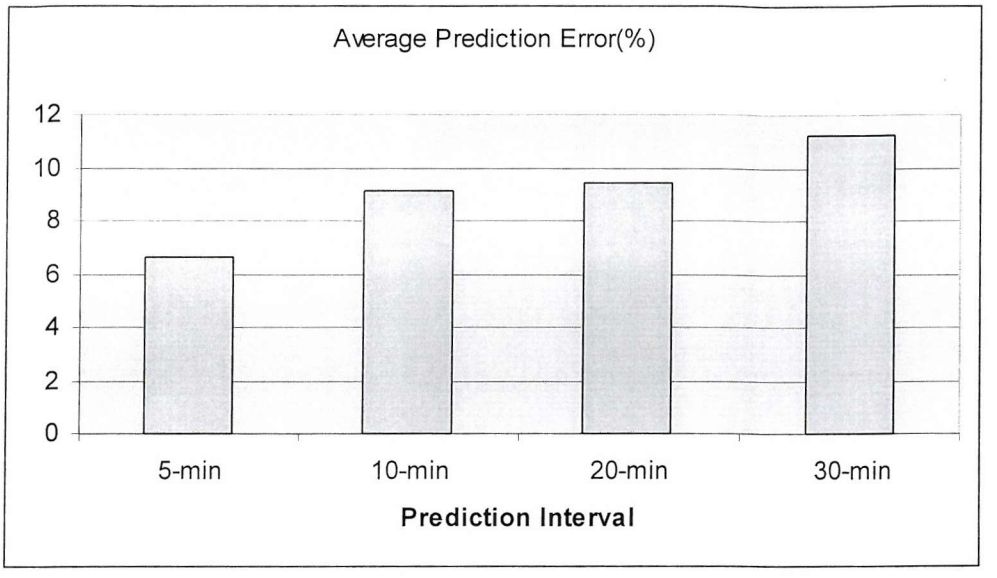


Figure 7.5 Prediction accuracy for various time intervals

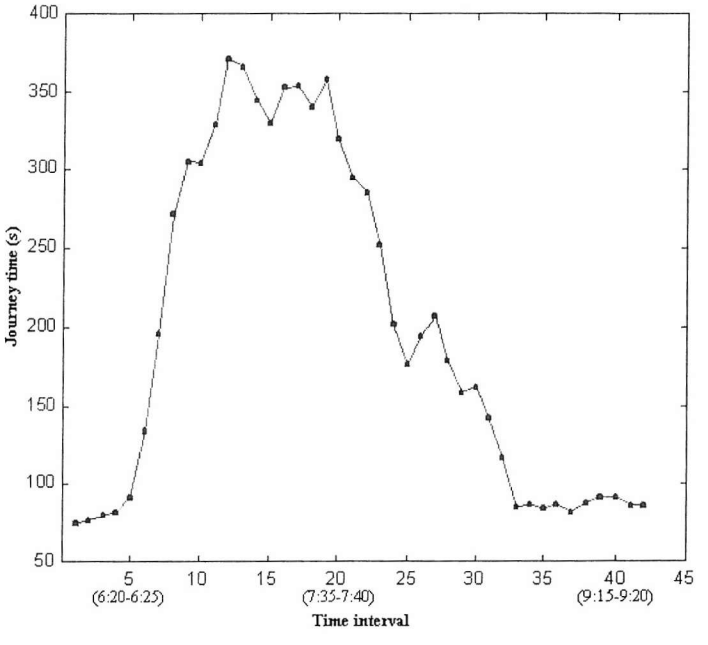


Figure 7.6 Journey time changes during an incident period on link 2, October 11 2002

It can be seen that in the early stage of the incident, journey times increase rapidly. It is then maintained at a high level before gradually decreasing. The corresponding average speed during the incident is shown in Figure 7.7. More incident data are presented in Appendix D. Although different incidents produce different impacts, changes in average speed during an incident period have similar characteristics. By studying the

characteristics of speed changes, it is possible to predict the duration of an incident and journey times in each time interval during the incident period.

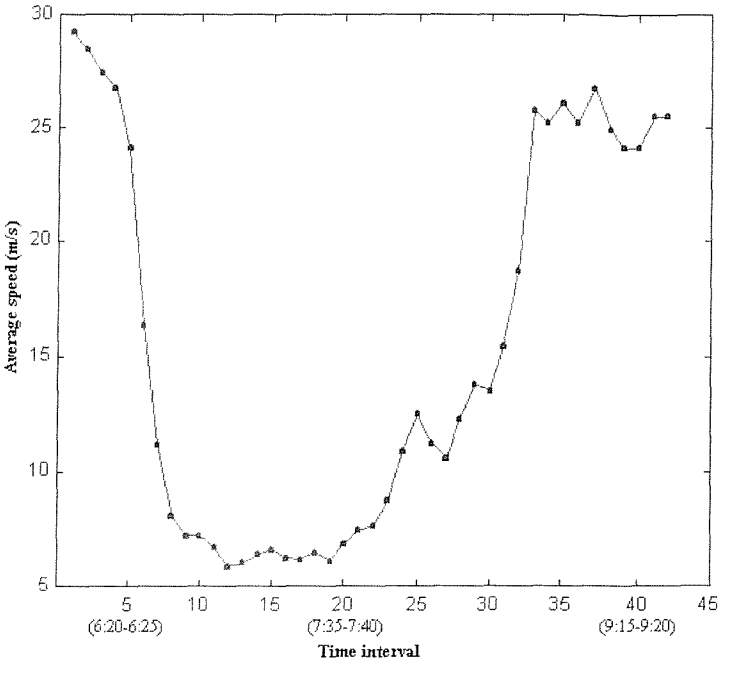


Figure 7.7 Average speed during an incident period on link 2, October 11 2002

In general, the same incident may cause longer delay in busy traffic than in light traffic. However, impacts of incidents may also be affected by many uncertain factors, such as driver behaviour and traffic operators' response. Therefore, it is difficult to precisely predict the duration of an incident at the beginning. Moreover, due to very limited incident database, the statistical characteristics of an incident process obtained may be biased.

Despite such disadvantages of journey time prediction in incident conditions, a trial of prediction has been carried out in this section. The prediction model aims to predict journey time by forecasting average speed. Speed change in an incident period has been divided into two stages: breakdown stage and recovery stage. Different mathematical functions have been used to reflect the different characteristics of the two stages of an incident.

7.4.1 Breakdown stage

A rapid decrease in the average speed can be found at beginning of a breakdown stage, with a minimum speed approached after only a few time intervals. This is considered to be the breakdown stage. For the incident process shown in Figure 7.7, the BEAM model identified the incident in the 6th time interval, and the minimum speed was achieved in the 12th time interval. Therefore, the breakdown stage includes seven time intervals, i.e. from the 6th to the 12th time interval. Incident data shown in Appendix D indicate that length of the breakdown stage for different incidents is generally from 25 minutes (5 time intervals) to 35 minutes (7 time intervals). A fixed length of 30 minutes is used for the prediction. The minimum speed in an incident period depends on many factors, such as traffic conditions and incident type. Data collected during this research (Appendix D) has shown that minimum speeds for different incidents are distributed from 2.65 m/s to 10.7 m/s with most of them are around 6 m/s. Since no dominating factors have been found, a fixed minimum speed of 6 m/s has been used for the prediction.

For the incident process shown in Figure 7.7, the breakdown stage is considered to be from the 6th time interval to the 11th time interval. Speeds in the six time intervals are shown in Figure 7.8. Five types of functions have been chose to fix the data as following:

1. Linear Function

$$y = -1.6889x + 14.989 \qquad (7.17)$$

2. Logarithmic Function

$$y = -5.2827\ln(x) + 14.87 \qquad (7.18)$$

3. Polynomial Function

$$y = -0.6173x^2 - 6.01x + 20.751 \qquad (7.19)$$

4. Exponential Function

$$y = 15.55\exp(-0.169x) \qquad (7.20)$$

5. Power Function

$$y = 15.135x^{-0.5169} \qquad (7.21)$$

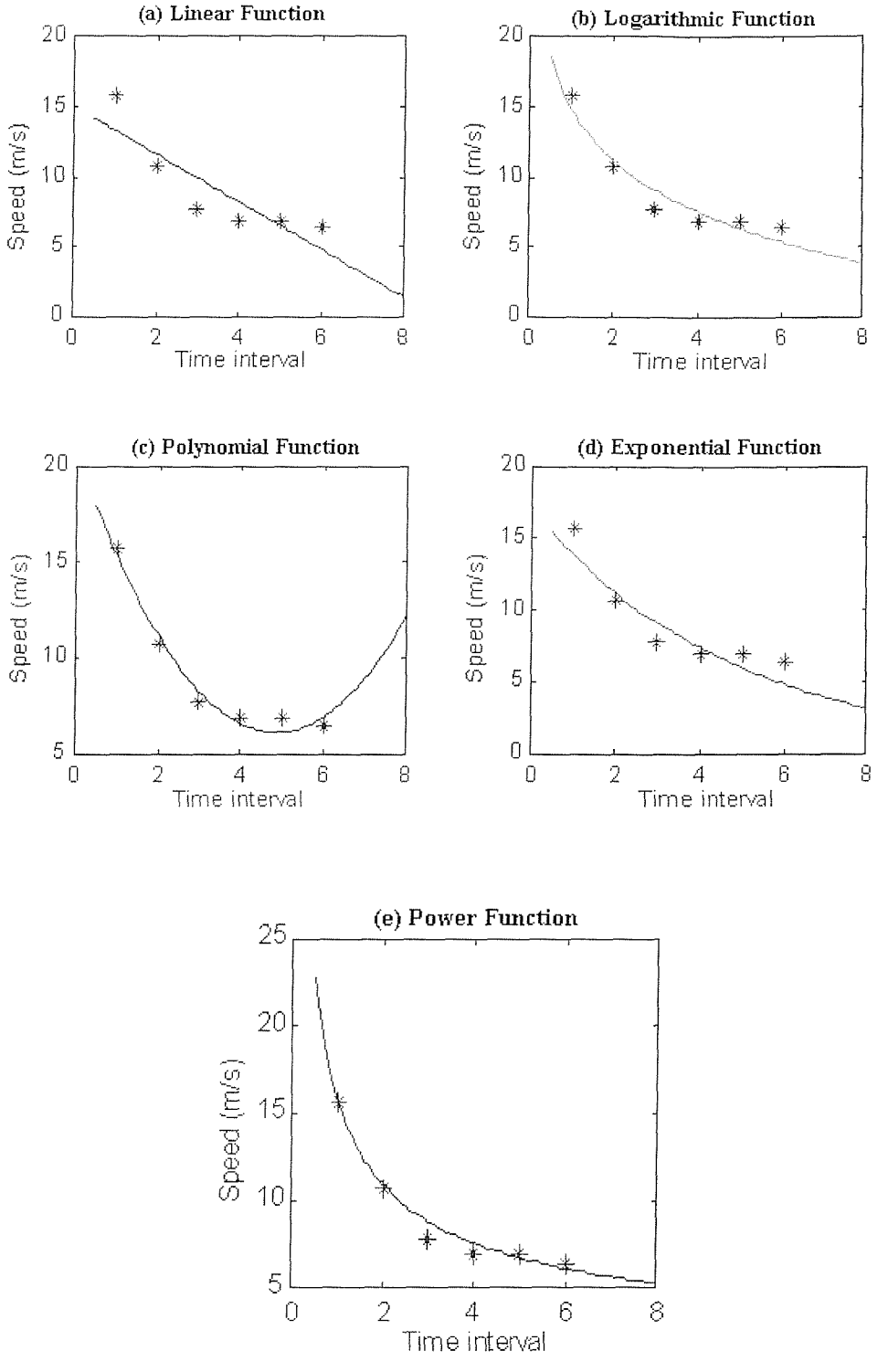


Figure 7.8 Speed and fitted curves in breakdown stage

Figure 7.8 shows fitted curves of the five functions and the best fitting function is the power function. For the power function, two parameters should be determined:

$$y = \alpha x^{\lambda} \qquad (7.22)$$

Speed at the first interval of the breakdown stage is known. Speed at the last interval of the breakdown stage is supposed to be the minimum speed of the incident process, i.e. 6m/s:

$$y(1) = \alpha = 15.75$$
$$y(6) = \alpha(6)^{\lambda} = 6 \qquad (7.23)$$

Solving the Equation 7.23, the two parameters can be obtained:

$$\alpha = 15.75$$
$$\lambda = -0.538$$

Therefore, the power function used for estimating speeds in the breakdown stage is:

$$y = 15.75x^{-0.538} \qquad (7.24)$$

It can be found that the parameters of Equation 7.24 are similar to parameters of Equation 7.21 which were estimated by known data in breakdown stage. For the incident shown in Figure 7.8, the breakdown stage is from the 6th to the 11th time interval and estimated speed at each time interval for the period is calculated by:

$$v_i = 15.75(i-5)^{-0.538} \qquad i = 6, 7, \ldots, 11 \qquad (7.25)$$

Incident data shown in Appendix D support the concept that average speed changes in the breakdown stage are generally similar to Figure 7.8. Thus, the power function can be used to predict the rapid decrease in speed in the breakdown stage. Two parameters, α and λ, are required. The parameter α is estimated from speeds at the early stage of an incident. For example, an incident causes a significant increase in journey time in the k*th* interval and the BEAM model will identify the incident in the (k+1)*th* interval. The parameter α is estimated as speed at the (k+1)*th* time intervals:

$$\alpha = v_{k+1} \qquad (7.26)$$

Since a fixed length of 30 minutes is used for the prediction, the 6th time interval of the breakdown stage is the minimum speed of the incident process:

$$v_{\min} = v_{k+1} 6^{\lambda} \qquad (7.27)$$

λ is calculated by solving Equation 7.27:

$$\lambda = \frac{\ln(v_{\min}) - \ln(v_{k+1})}{\ln(6)} \qquad (7.28)$$

7.4.2 Recovery Stage

The recovery stage of an incident is the period in which speed increases from the minimum speed to the normal link speed. The duration of the recovery stage also depends on many factors, such as the minimum speed of the incident and traffic flow on the link during the incident period. Data collected in incident conditions have shown that the length of the recovery stage substantially relates to the traffic flow during the time period when the incident occurred. In general, the busier the traffic, the longer period required for recovery from an incident. If it is assumed that the recovery stage lasts 30 minutes, i.e. six time intervals, the mean speed of the six time intervals can be calculated using historic journey time data. An approximately linear relationship between the mean speed and length of the recovery stage has been found, as shown in Figure 7.9.

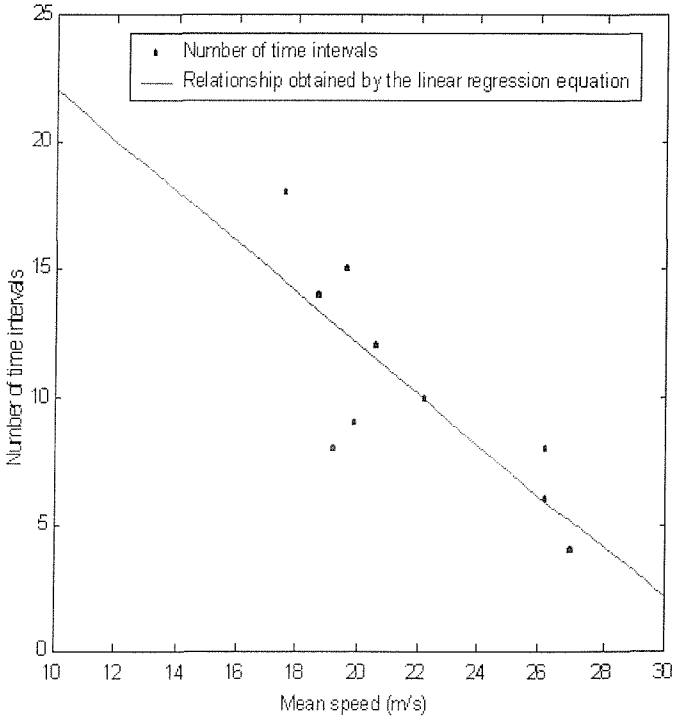


Figure 7.9 Length of recovery stage and mean speed

The length of the recovery stage, presented in the number of time intervals, is estimated by a linear regression as:

$$N_r = -0.996S_6 + 32.05 \qquad (7.29)$$
$$R^2 = 0.65$$

where S_6 is mean travel speed on the link for the six time interval, estimated from historic data. Thus, length of the recovery stage can be estimated using historic journey time data. Increase in speed in the recovery stage can be represented by a hyperbolic cosine function:

$$v = v_{\min} \cosh(\beta x) = v_{\min} \frac{e^{\beta x} + e^{-\beta x}}{2} \qquad (7.30)$$

The function describes the speed change from the minimum speed to the speed before the incident occurred:

$$\begin{aligned} v(0) &= v_{\min} \\ v(N_r) &= v_{N_r} \end{aligned} \qquad (7.31)$$

where $v_{\min}$ denotes the minimum speed in the incident period, i.e. 6 m/s, and v_{N_r} is speed in the corresponding time interval, estimated from historic data. Parameters of the Equation 7.29 are therefore estimated from the following equations:

$$\begin{aligned} v_{\min} &= 6 \\ \beta &= \frac{1}{N_r} \ln(\frac{v_{N_r}}{3}) \end{aligned} \qquad (7.32)$$

In summary, the duration of the recovery stage, presented as the number of time interval, is estimated by:

$$N_r = -0.996 S_6 + 32.05$$

where S_6 is mean travel speed on the link for the six time intervals, estimated from historic data. Average speed in each time interval of the recovery stage is predicted by:

$$v_i = 6\cosh[\beta(i-k-6)] \qquad i = k+7,\ k+8,\ \ldots,\ k+6+N_r$$

where $\beta = \frac{1}{N_r} \ln(\frac{v_{N_r}}{3})$.

7.4.3 Summary of Prediction Procedure

When an incident causes a significant increase in journey time in the k*th* interval, and the BEAM model will identify an incident in the (k+1)*th* interval, the following prediction procedure is applied:

Step 1. The breakdown stage is from (k+1)*th* interval to (k+6)*th* interval;

Step 2. Assume the recovery stage is from the (k+7)*th* k+7th to the (k+12)*th* time interval, and calculate historic mean speed of the six time intervals S_6;

Step 3. Recalculate length of the recovery stage by:

$N_r = -0.996S_6 + 32.05$

and the recovery stage is estimated from the (k+7)*th* time interval to the $(k+6+N_r)th$ time interval;

Step 4. Estimate the parameters $\alpha = v_{k+1}$ and $\lambda = \frac{\ln(v_{\min}) - \ln(v_{k+1})}{\ln(6)}$;

Step 5. Use $v_i = v_{k+1}(i-k)^{\lambda}$ to calculate speed in each time interval of the breakdown stage;

Step 6. Estimate historic speed in the $(k+6+N_r)th$ time interval v_{N_r}

Step 7. Calculate $\beta = \frac{1}{N_r}\ln(\frac{v_{N_r}}{3})$;

Step 8. Use $v_i = 6\cosh[\beta(i-k-6)]$ to calculate speed in each time interval of the recovery stage.

7.4.4 An example of prediction

In this section, an incident which occurred on the 11th October 2002 on link 1 has been taken as example to illustrate the prediction procedure. The BEAM model declared the incident in the 16th time interval, i.e. 7:15-7:20. The breakdown stage is estimated from the 16th to the 21st time interval. By obtaining S_6= 26.22 m/s from historic data, length of the recovery stage is estimated as:

$$N_r = -0.996 \times 26.22 + 32.05 = 5.9$$

The recovery stage involves six time intervals, i.e. from the 22nd to the 27th time interval. The historic speed in the 27th time interval is v_{27} =26.9 m/s.

Parameters α and λ are estimated as:

$\alpha = 14.23$

$\lambda = -0.4823$

Therefore, the speed in each time interval in the breakdown stage is estimated by:

$$v_i = 14.23(i-15)^{-0.4823} \qquad i = 16, 17, \ldots, 21$$

The speed in each time interval in the recovering stage is estimated by:

$$v_i = 6 \times \cosh[0.3656(i-21)] \qquad i = 22, 23, \ldots, 27$$

The incident period is considered from the 16th to 27th time intervals, i.e. 60 minutes. Comparison of observed and predicted speeds in the period is shown in Figure 7.10. Because of lighter traffic on link 1 than on other links in the morning peak time, the incident period is likely to last a shorter time on link 1.

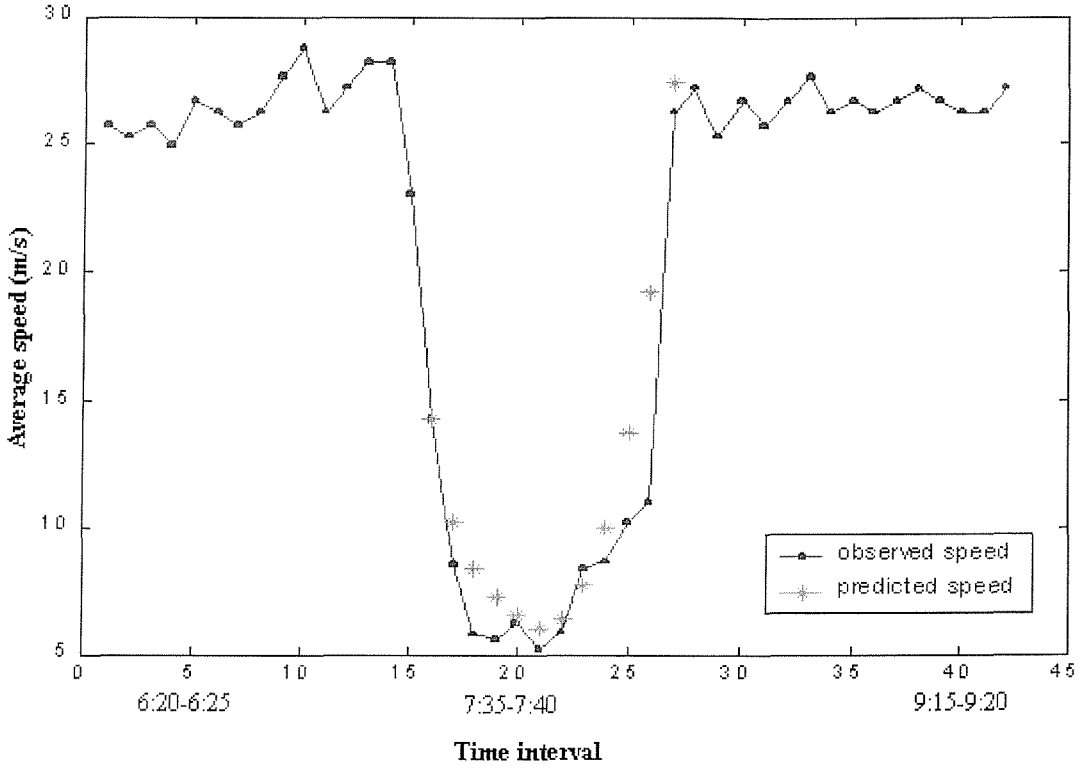


Figure 7.10 Comparison of observed and predicted speed on link 1, Oct. 11 2002

7.5 Summary

This chapter has developed a model to predict motorway link journey times based on journey time data measured by probe vehicles. Consideration was given to the fact that probe vehicle data provides only direct journey time with good accuracy. Average link journey times expressed by 5-minute intervals were considered as stationary stochastic processes. In incident-free traffic flow, a linear model has been developed for one-step-

ahead prediction (5 minutes ahead), and the Kalman filter algorithm has been used for multi-step-ahead prediction. The Kalman filter model developed in this research has been shown to be effective in predicting journey time and acceptable errors have been obtained using this model. The average relative prediction error was found to increase with the length of the prediction interval. With a 20-minute interval, the prediction error remains under 10%. However, significant increases in the prediction error were observed at 30-minute prediction period.

As the above predictions are based on historic journey time data, the predictions are unsuitable for incident condition. Therefore, a different prediction strategy was required for use in incident situation. After an incident is identified, prediction work aims to forecast journey times over an entire process of an incident rather than in a certain time interval. According to characteristics of speed in incident traffic, a process of an incident has been divided into two stages: breakdown stage and recovery stage. Length of the breakdown stage has been considered as fixed, i.e. six time intervals, while length of the recovery stage has been estimated by historic journey times. A power equation has been used to estimate speed in each time interval in the breakdown stage, and a hyperbolic cosine equation has been used for the recovery stage. Due to complexity of change in journey time with various incidents, the model may not provide accurate prediction results universally for all cases. However, the model can give an entire predictive description of an incident at the beginning of the incident. Although the predictive journey times may not match the real journey times precisely, the model will be helpful for traffic operators and travellers, especially before further information becomes available.

Chapter Eight

Active Probe Vehicle

8.1 Introduction

Two types of vehicle can be used in a 'probe' capacity: specially equipped probe vehicles for traffic data collection (active probe vehicles), and general vehicles installed with the location and communication equipment for another purpose other than data collection (passive probe vehicles). Historically, the manual method has been the most commonly used travel time data collection technique. This method requires a driver and a passenger to be in the vehicle. The driver operates the vehicle while the passenger records time information at pre-defined checkpoints. Now GPS has become the most common technology used for active probe vehicle. As introduced in Section 2.2.1, in most travel time studies involving active probe vehicles, only one vehicle is used to measure journey time along a specific direction of travel. Since the driver of the probe vehicle is a member of the data collection team, driving styles and behaviours can be controlled to match desired driving behaviour. Most travel time studies ask drivers to travel at the speed of the traffic stream and maintain the number of overtaking vehicles the same as those overtaken. And journey time of the active probe vehicle will be directly taken as the average journey time. In practice, however, drivers may be unable to achieve the requirement because of the inherent difficulties of keeping track of passed and passing vehicles.

Since GPS device can measure vehicle location and speed at 1 Hz, GPS data can be used to build detailed speed profiles of the probe vehicle over an entire journey. By combining the high time resolution data, it is possible to assemble a detailed picture of the movements of the probe vehicle. This study seeks insight into various aspects of traffic flow by measuring and analysing the movement of the probe vehicle. In this chapter, journey time will be calculated based on the analysis of the speed profiles rather than directly taking journey time of the probe vehicle. In this case, drivers of

probe vehicles do not have to attempt to drive at the speed of the traffic stream, i.e. drivers can drive as they would normally.

8.2 Speed profile development

Speed measurement is a basic output of GPS, although it has been ignored by most users. GPS speed measurement is almost instantaneous and essentially independent of position fixes (D'Este et al., 1999). After the termination of the Selective Availability (SA), the speed measurement can achieve to 0.1m/s with 95% confidence. By combining GPS observations of location, time and speed, GPS can offer a cost-effective way of collecting a stream of traffic information that can be processed to derive journey time estimation and incident detection. Two speed profiles can be derived from GPS data: time-speed profile, and space-speed profile. Examples of the two speed profiles is shown in Figure 8.1, in which a journey on link 2 in the morning peak hour of the 1st. October 2002 is described. Speed profiles are likely to differ by vehicle type, road type and driver style as well as traffic flow conditions. Any study based on speed profiles should extract traffic flow characteristics and filter influence of other factors.

8.3 Journey time estimation using single GPS equipped probe vehicle

8.3.1 Driving pattern

Speed profiles are determined not only by traffic flow but also driver behaviour. Speed profiles of four vehicles on link 2 at 15:50-15:55, 26th July 2001 are shown in Figure 8.2. The figure illustrates that in the same general traffic condition, speed profiles can be very different. In this study, three driving patterns have been classified: fast, medium and slow. Assuming that average journey time on a link in a time interval is known, fast driving is classified when the journey time of a vehicle is shorter than the average journey time in the same time interval and on the same link.

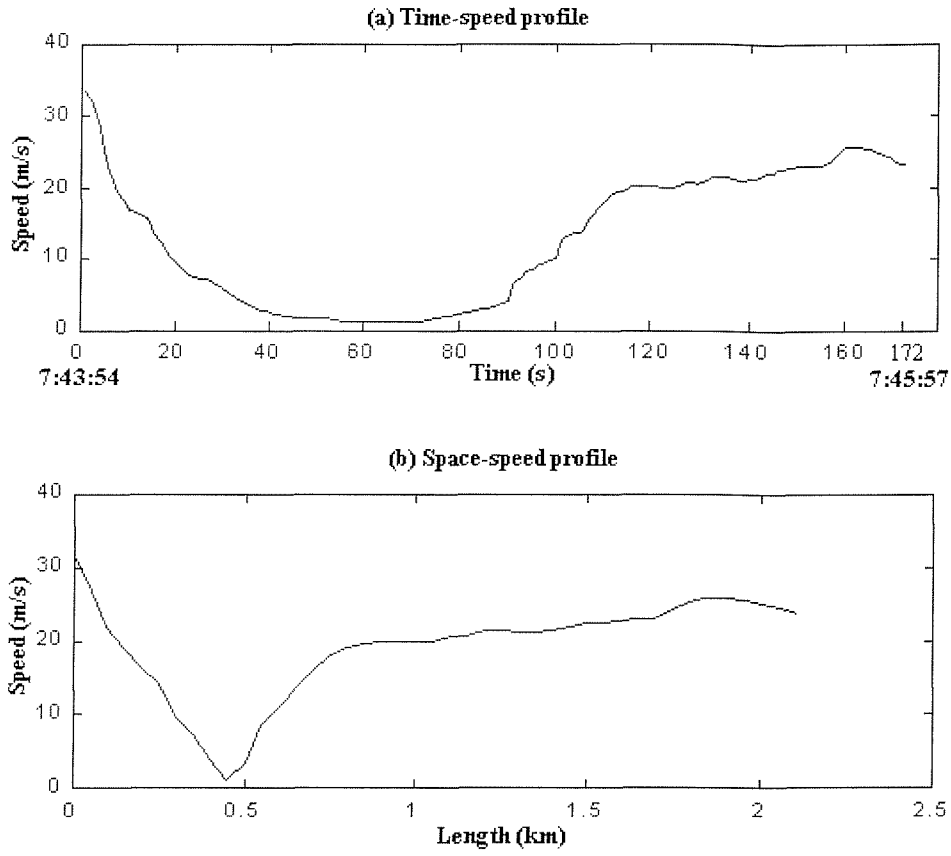


Figure 8.1 Time-speed and space-speed profiles

Conversely, slow driving is classified if journey time of the vehicle is longer than the average journey time. Medium driving means that journey time of the vehicle is similar to the average journey time, i.e. the difference between the two journey times is less than 5 seconds. The driving pattern in this study is associated only with the difference between vehicle journey time and the average journey time. Therefore, the same driver can be classified as belonging to different groups of driving pattern for different journeys. As shown in Figure 8.2, different driving patterns have different characteristics in acceleration and deceleration. Therefore, driving patterns can be identified from features of speed profiles.

Data collected in the ramp metering survey have been used to characterise speed profiles for different driving patterns. As introduced in Section 3.4.1, in the ramp metering survey, the TRG Instrumented Vehicle (IV) carried out 105 valid journeys on link 7. During the survey, the IV was driven by a variety of drivers. However, since in the ramp metering survey, drivers were required to drive the vehicle on the motorway

lane 1 for as long as reasonably possible, most journey time observations were longer than the average. Of all 105 journeys, 81 were classified as slow, 21 medium and only 3 fast driving patterns. Characteristics of fast driving cannot be obtained due to the small sample.

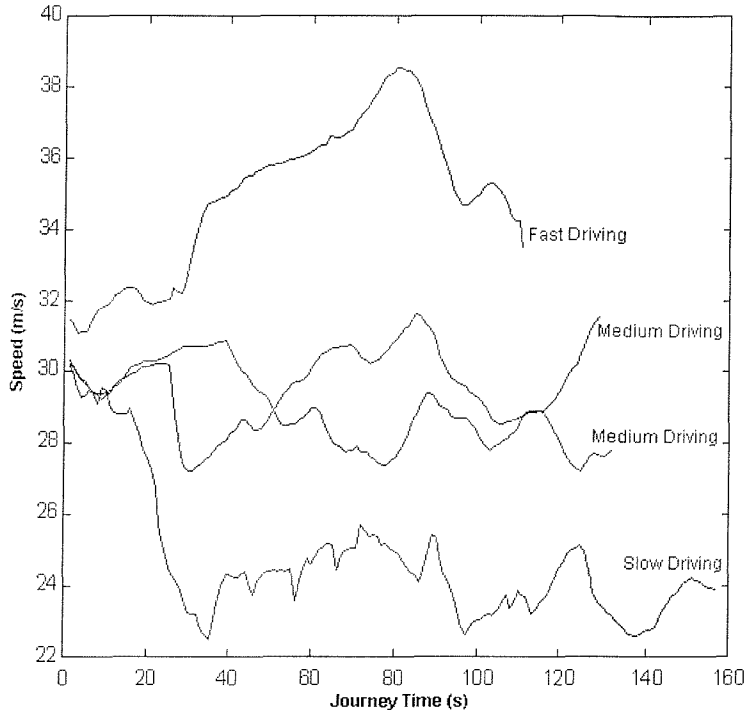


Figure 8.2 Different driving patterns

As described in Section 3.4, the car following survey provided some speed profiles which can be used to study features of driving patterns. 69 journeys on link 7 from the survey have been used, including 30 fast, 18 medium and 15 slow driving patterns. The data collected from the two surveys have been used to develop a model to estimate the average journey time from a speed profile obtained by a GPS equipped probe vehicle. By identifying driving pattern of a journey, the model will also estimate the difference between journey time of the probe vehicle and average journey time, therefore obtain the average journey time.

8.3.2 Definition of variables

Speed and acceleration have been used to describe the driving style and performance of individual drivers (Drew, 1968; Lu, 1992; Sermons and Koppelman, 1996; Ericsson, 2001). In this research, two variables associated with speed and acceleration

respectively have been used to characterise speed profiles of different driving patterns. Average speed is a common parameter to describe traffic conditions and driving style. For a GPS equipped probe vehicle, the average speed of a journey is calculated from a speed profile as follows:

$$\bar{v} = \frac{1}{T_{out} - T_{in}} \int_{T_{in}}^{T_{out}} v(t)dt \qquad (8.1)$$

where T_{in} is start time of the journey, T_{out} is the end time of the journey, $v(t)$ is the speed of probe vehicle at time t.

Traditionally, both acceleration and deceleration were considered to have the same weighting in characterising vehicle movement (Drew, 1968). However, research on 'Acceleration Signature' (Robertson et al, 1992), which used acceleration to describe the style of a driver, has shown that deceleration rate of most drivers is very similar, even though they have various acceleration rate. That means that deceleration rate may not represent the characteristics of driving styles as efficiently as acceleration rate does. Therefore, in this research, only acceleration rate is used to classify driving patterns. A new variable, *Continuous Acceleration* (CA), has been defined to described characteristics of acceleration:

$$CA = \frac{v_{t_2} - v_{t_1}}{t_2 - t_1} \qquad (8.2)$$
$$t_2 - t_1 \geq 5s$$

Equation 8.2 calculates average accelerations which last at least 5 seconds. In an 'Acceleration Signature', it was required to place more priority on accelerations at high speed than at low speed, since high acceleration is easier to achieve at low speed. To combine acceleration with speed, a speed factor was introduced, and Equation 8.2 can be written as:

$$CA = \left(\frac{v_{t_2} - v_{t_1}}{t_2 - t_1}\right)\left(\frac{v_{t_1}}{S}\right) \qquad (8.3)$$
$$t_2 - t_1 \geq 5s$$

where S is mean travel speed on the link in the time interval, estimated from historic data. For a journey, more than one CA may be obtained and the *Maximum Continuous*

Acceleration (MCA) has been used to identify driving pattern. If in a journey, there is no acceleration which lasts more than 5 seconds, the value of MCA is zero.

A fast driver may have a high average speed. However, high average speed may also be achieved by a medium driver if the traffic is light. Moreover, in relatively busy traffic, a fast driver may produce a low average speed but still with high value of MCA. Fuzzy logic has been used to recognise driving pattern using average speed and MCA and a fuzzy model has been developed to estimate journey time from the speed profile of a GPS equipped probe vehicle. Inputs of the model are average speed and MCA, derived from the speed profile. Output of the model is the ratio of journey time of the probe vehicle to average journey time of all vehicles on the link in the time interval:

$$f = \frac{JT_{PV}}{\overline{JT}} \tag{8.4}$$

If $f > 1$, journey time of probe vehicle is longer than the average journey time and driving pattern is slow. If $f < 1$, journey time of probe vehicle is shorter than the average journey time and driving pattern is fast. $f = 1$ indicates a medium driving pattern. Of the data collected, the maximum value of f was 1.35, and the minimum value is 0.75. Consequently the range of the output is determined to be from 0.75 to 1.35.

8.3.3 Parameter estimation

Each input variable has five membership functions (MF): "very slow", "slow", "medium", "fast", "very fast". Triangular MF is used for "slow", "medium" and "fast", while trapezoidal MF is used for "very slow" and "very fast". Membership functions of average speed are shown in Figure 8.3.

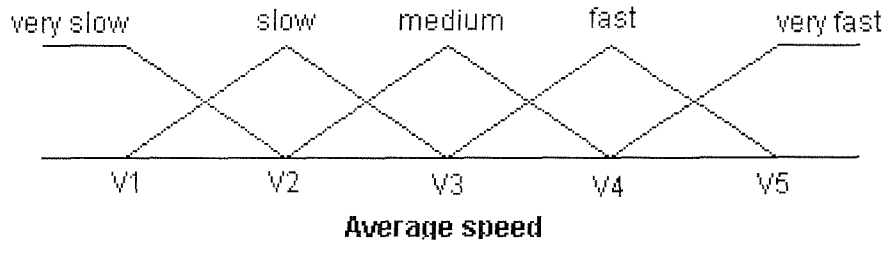


Figure 8.3 Membership functions of average speed

For membership functions of average speed, there are five parameters, which are estimated by historic journey time data. As introduced in Section 6.4.3, historic journey time was obtained from 15 incident free days. V_1, V_3 and V_5 are given by:

$$\begin{aligned} V_1 &= \frac{L}{JT_{\max}} \\ V_3 &= \frac{L}{\overline{JT}} \\ V_5 &= \frac{L}{JT_{\min}} \end{aligned} \qquad (8.5)$$

where, $JT_{\max}$ is the maximum journey time of the 15 days, $JT_{\min}$ is the minimum journey time of the 15 days, and $\overline{JT}$ is the mean journey time. V_1 and V_2 are given by:

$$\begin{aligned} V_2 &= \frac{V_1 + V_3}{2} \\ V_4 &= \frac{V_3 + V_5}{2} \end{aligned} \qquad (8.6)$$

It can be found from Equation 8.5 and 8.6 that the membership function values for average speed are link-specific. For the same link, the membership function values also vary for different intervals. For instance, the five parameters for link 2 and time interval of 7:50-7:55 are determined by the journey time data presented in Table 6.5. The maximum journey time in Table6.5 is 196 seconds, the minimum journey time is 84 seconds and the mean journey time is 115 seconds. The five parameters are calculated as:

$$\begin{aligned} V_1 &= \frac{2208}{196} = 11.27 m/s \\ V_3 &= \frac{2208}{115} = 19.2 m/s \\ V_5 &= \frac{2208}{84} = 26.29 m/s \\ V_2 &= \frac{V_1 + V_3}{2} = 15.23 m/s \\ V_4 &= \frac{V_3 + V_5}{2} = 22.74 m/s \end{aligned}$$

The membership functions of average speed for link 2 and time interval of 7:50-7:55 are presented in Figure 8.4.

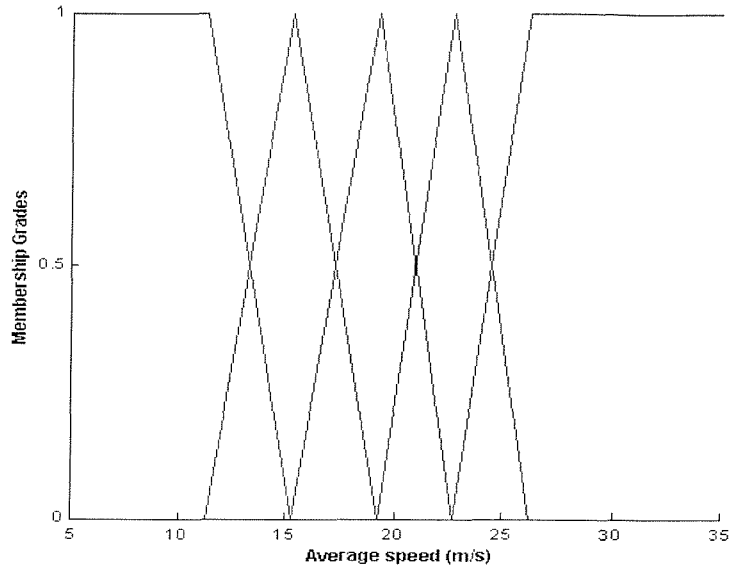


Figure 8.4 The membership functions of average speed
for link 2 and time interval of 7:50-7:55

The data collected from the ramp metering survey have been used to estimate the membership function values of MCA. The same membership functions of MCA have been used for different time intervals. The membership functions of the MCA with the corresponding parameters are presented in Figure 8.5.

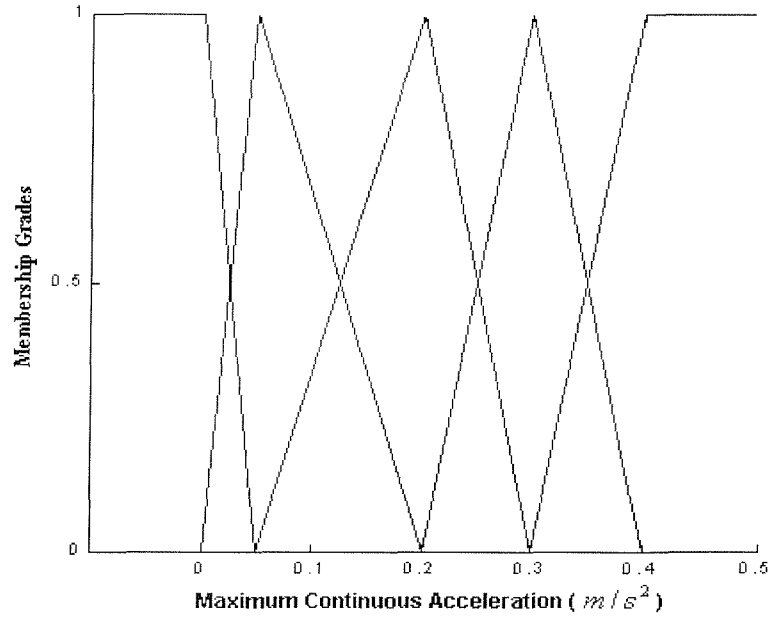


Figure 8.5 The membership functions of Maximum Continuous Acceleration

Differing from the input variables, the output variable *f*, has only three triangular membership functions: fast, medium and slow, as shown in Figure 8.6.

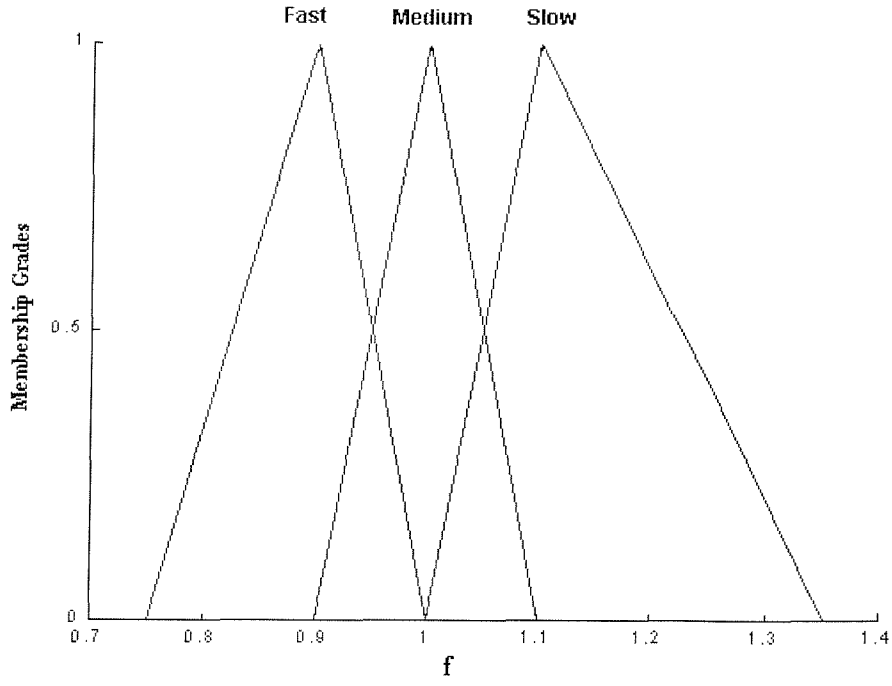


Figure 8.6 The membership functions of the output variable

8.3.4 Fuzzy model development

Two variables, average speed and maximum continuous acceleration (MCA), have been extracted from the speed profiles of probe vehicles. Fuzzy logic was used to describe the relationship between the two variables and the driving patterns. A model using fuzzy logic has been developed to identify a driving pattern of a journey using the two variables. A graphical presentation of the potential fuzzy model is shown is Figure 8.7.

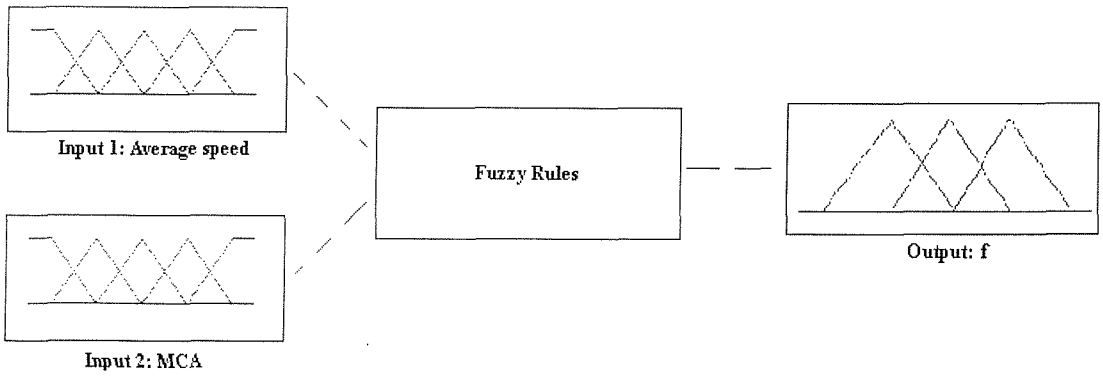


Figure 8.7 Graphic presentation of the fuzzy model

The fuzzy rule presents the relationship between the input variables, i.e. average speed and MCA, and the output variable, i.e. driving pattern. The data collected on link 7 have indicated that fast driving generally results in a fast average speed and large value of MCA. However, fast driving may have a fast average speed but a very low MCA. This situation occurs when speed in the journey is maintained at a high level and it is difficult to be increased in the traffic condition. Fast and very fast average speed may also present medium driving pattern, if with a medium MCA. This is generally because of light traffic. Slow average speed with low MCA indicates slow driving. However, very slow average speed is generally observed in congested traffic. In congested traffic, since vehicle movement may be dominated by traffic flow, i.e. it is difficult to drive too fast or too slow in congestion, driving patterns in congestion are generally medium, even with a high MCA. The fuzzy rules of the model are summarised in Table 8.1

Table 8.1 Fuzzy rule matrix

		Average speed				
		very slow	slow	medium	fast	very fast
MCA	very slow	slow	slow	slow	medium	fast
	slow	medium	slow	medium	medium	medium
	medium	medium	medium	medium	medium	fast
	fast	medium	medium	medium	fast	fast
	very fast	medium	fast	fast	fast	fast

An example of output of the fuzzy model which parameters indicated in Figure 8.4 and Figure 8.5 is shown in Table 8.2.

Table 8.2 An example of fuzzy output

Average Speed (m/s)	MCA (m/s^2)	f	Driving pattern
30.6	0.16	0.905	Fast
29.1	0.39	0.883	Fast
26.5	0.28	0.905	Fast
26.5	0.16	1	Medium
20.8	0.39	0.925	Fast
20.8	0.22	1	Medium
21.5	0.05	1.03	Slow
21.5	0.22	1	Medium
17.5	0.15	1.1	Slow

8.3.5 Model validation

Data collected on link 1 and link 2 are used to validate the fuzzy model. The data were collected in the morning peak hours of 7:00-9:00, from the 1st to the 11th October 2002. Each weekday morning, a GPS equipped probe vehicle was driven by a different driver between Junction 10 and Junction 14 on the M3. 83 journeys, 42 on link 1 and 41 on link 2, were carried out. For each journey, average speed and MCA are calculated and used as the input to the fuzzy model. Average journey time was then estimated by:

$$\overline{JT}_f = \frac{JT_{PV}}{f} \qquad (8.7)$$

where JT_{PV} is journey time of the probe vehicle, f is output of the fuzzy model, and $\overline{JT}_f$ is estimated average journey time.

To estimate the accuracy of the fuzzy model, average journey times measured by Automatic Number Plate Recognition (ANPR) were considered as 'real journey time' and compared with results estimated by the fuzzy model. Estimation results for all 42 journeys on link 1 are shown in Figure 8.8. Journey times on link 1 are generally stable. The large journey times in journey 41 and 42 were obtained in incident conditions. Results in Figure 8.8 show that the fuzzy model can produce good results in general traffic as well as in incident conditions.

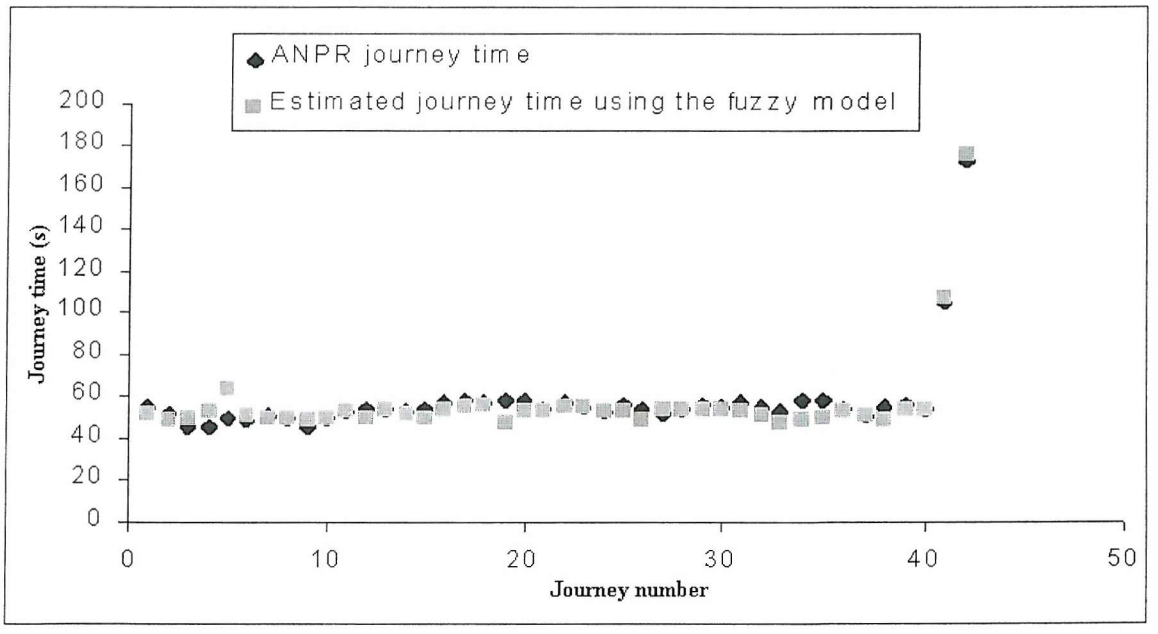


Figure 8.8 Comparison of estimated journey times and ANPR data for link 1

Journey times on link 2 vary more than on link 1 and more congestion occurs on link 2. Results shown in Figure 8.9 indicate that the fuzzy model can perform better in busier traffic than in lighter traffic. Since traffic on link 2 is generally busier than on link 1, more accurate estimation results was obtained for link 2.

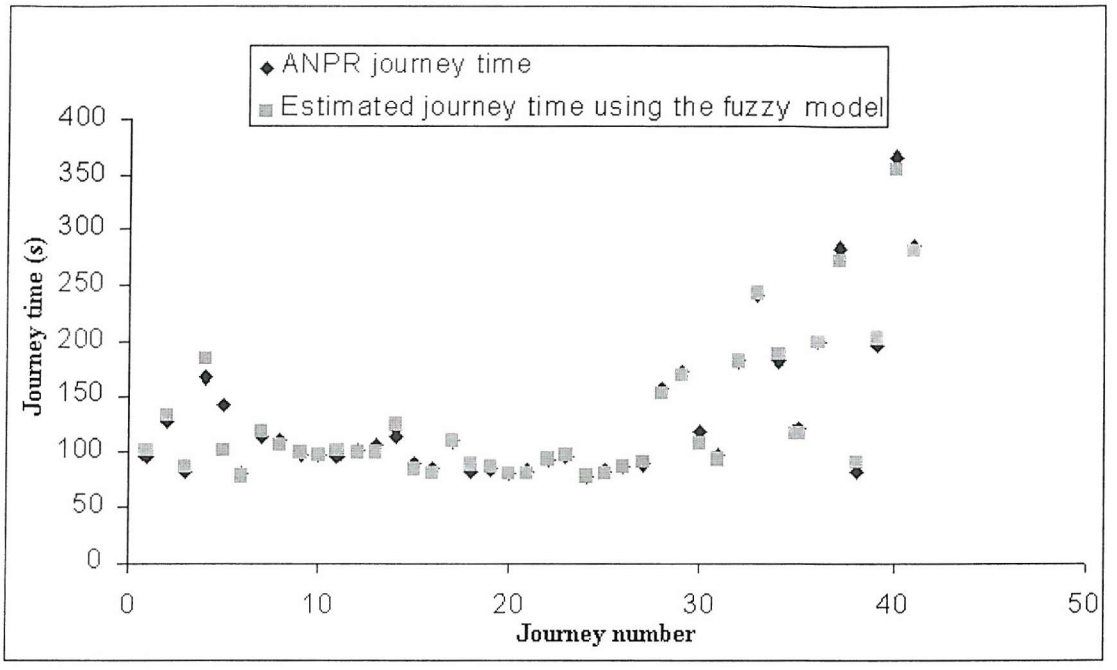


Figure 8.9 Comparison of estimated journey times and ANPR data for link 2

The model performance on link 1 and link 2 are summarised in Table 8.3, in which *estimation error* is calculated by:

$$e_f = \frac{\left|\overline{JT}_f - \overline{JT}_{ANPR}\right|}{\overline{JT}_{ANPR}} \tag{8.8}$$

where $\overline{JT}_{ANPR}$ denotes ANPR average journey time.

Table 8.3 Comparison of link 1 and link 2 estimation results

	Number of journeys			Mean Estimation error
	Total	Estimation error ≤ 10%	Estimation error = 0	
Link 1	42	36	9	5.56%
Link 2	41	39	4	4.10%

8.4 Incident detection using single GPS equipped probe vehicle

In Section 8.3, it has been shown that using single GPS equipped probe vehicle, the average link journey time of all vehicles on a link and in 5-min intervals can be estimated accurately. The resulting journey times can be used for incident detection as input of the BEAM model developed in Chapter 6. The BEAM model requires journey times for both previous and current time intervals. In most studies involving active probe vehicles, only one probe vehicle may be used and journey time cannot be measured for each time interval. Since active probe vehicles can collect much more data than link journey time, a new method for incident detection using a single probe vehicle is required.

Speed profiles of the probe vehicle can be also used to detect incidents by analysing characteristics associated with each traffic flow condition. As introduced in Section 8.3, characteristics of speed profiles can be affected not only by traffic flow but also by driving style and vehicle type. However, in incident conditions, traffic flow will dominate speed profile over other factors. Figure 8.10 shows speed profiles of three journeys in link 2 on the 8th October 2002, in which an incident was reported at 7:40 on the link. Therefore, speed profiles from single probe vehicles will be able to detect the presence of slow-moving or stationary traffic.

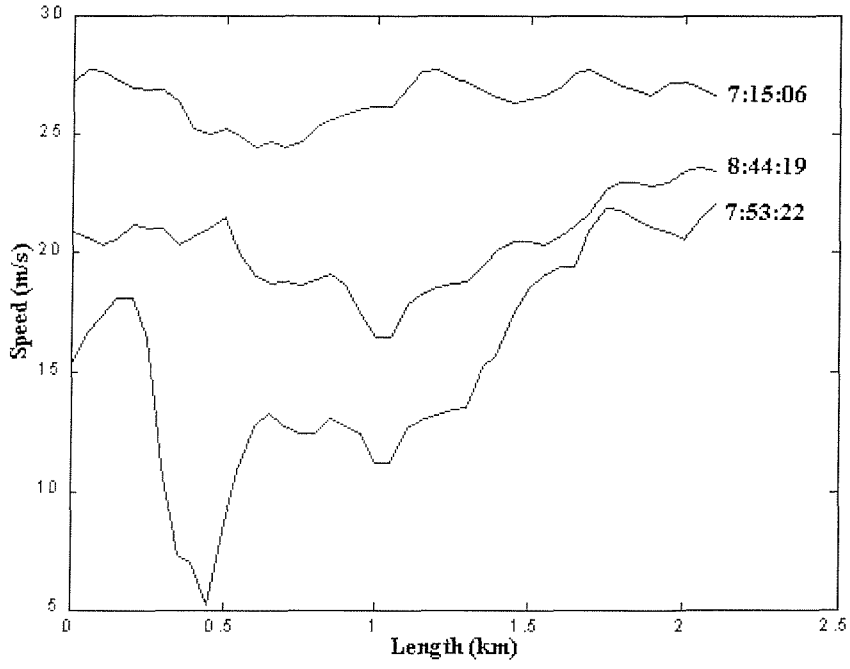


Figure 8. 10 Speed profiles in incident and non-incident condition

8.4.1 Methodology

In Section 6.2.2.2, some AID algorithms based on probe vehicles have been reviewed. Since incidents can cause larger journey time than usually experienced, previous researchers have compared current and historical journey time to identify incidents. In Bhandari et al. (1995), for one probe vehicle report, when the ratio of current journey time to historic data rises above 3.5, an incident is declared (see Table 6.4). Although an incident causes rapid increases in journey time, such high level will not be achieved immediately. Therefore, if a probe vehicle is travelling on a link where an incident has just occurred, the probe vehicle journey time is unlikely to be declared. For example, on the 8th Oct. 2002, an incident occurred at 7:40 on link 2. The TRG IV travelled on the link at 7:50 with a journey time of 156 seconds, which is only 1.35 times the historic journey time.

Sermons and Koppelman (1996) investigated the effectiveness of individual vehicle movement measures in the detection of incidents on urban arterial road segments. Several measures of varying complexity, including average speed, running time and speed, and coefficient of variation of speed, were calculated from vehicle positioning data recorded at 1-s time intervals. Analysis showed that incident journeys result in higher running times, total times and coefficients of variation of speed, but lower average speeds and running speeds. A number of discriminant models using the measures were developed, and these require sufficient incident data and incident-free data to classify the two traffic conditions. In this research, since only a few journeys were conducted in incident conditions, using existing data will not enable a discriminant model to be developed.

The BEAM model introduced in Section 6.4 was developed based on the premise that an incident can cause significant increase in journey time. The BEAM model uses two input variables: link journey time and difference of journey time between two adjacent time intervals. Since link journey time can be measured by an active probe vehicle, another alternative variable should be defined to reflect the change in journey times between two adjacent time intervals. At the beginning of an incident, the incident will not have affected upstream traffic. Thus, a vehicle may travel with normal speed from upstream, and then decelerate on approach to the incident location. Therefore, a large

deceleration rate can be observed. Drew (1968) showed that in congestion, the traffic condition dominated acceleration and deceleration rates over driver behaviours and it could be considered that a large deceleration rate is obtained mainly as a result of traffic conditions. Thus, deceleration rate can be used as an input variable of the BEAM model to replace the difference of journey time between two adjacent time intervals.

Similar to the variable used in the estimation fuzzy model, Maximum Continuous Acceleration (MCA), a new variable, *Continuous Deceleration (CD)*, has been defined to describe the characteristics of deceleration in incident condition:

$$CD = \left(\frac{v_{t_1} - v_{t_2}}{t_2 - t_1}\right)\left(\frac{S_d}{v_{t_2}}\right) \qquad (8.9)$$
$$t_2 - t_1 \geq 5s$$

where S_d is a constant with a value of $26m/s^2$. Equation 8.9 is used to calculate average deceleration which lasts at least 5 seconds. Differing from Equation 8.2, speed at the end of the deceleration process rather than speed at the start point is used in Equation 8.9. The lower the speed at the end point, the larger the value of the continuous deceleration. For a journey, the *Maximum Continuous Deceleration (MCD)* will be used as input of the BEAM model. Speed at the end of a deceleration process is not associated with average link speed, and a constant value is used to maintain the same form as the definition of the MCA. Thus, the MCD is not link-specific as the MCA. If there is no deceleration process lasting more than 5 seconds in a journey, the value of MCD is zero.

Values of MCD for six incident-free journeys on link 1 and link 2 respectively have been calculated to estimate the distribution of MCD. The values of MCD from different journeys in incident-free condition are normally distributed and the parameters of the distribution are:

$$\mu_{MCD} = 0.39m/s^2$$
$$\sigma_{MCD} = 0.52m/s^2$$

The data have been shown to be independent to journey time, i.e. $\rho = 0$. Due to the independence between the two variables, the bivariate normal distribution is simply the

product of two univariate normal distributions. Therefore, an incident will be declared if either journey time or MCD exceeds the confidence level of the normal distribution.

8.4.2 Result analysis

The BEAM model using MCD is applied to all 42 journeys on link 2 and 41 journeys on link 2. Since journey time on link 1 maintains stable over the morning peak hours and the journey time distribution for different time intervals in the morning peak hours are same. Detection results using 99.9% coverage contour for the 42 journeys on link 1 are shown in Figure 8.11.

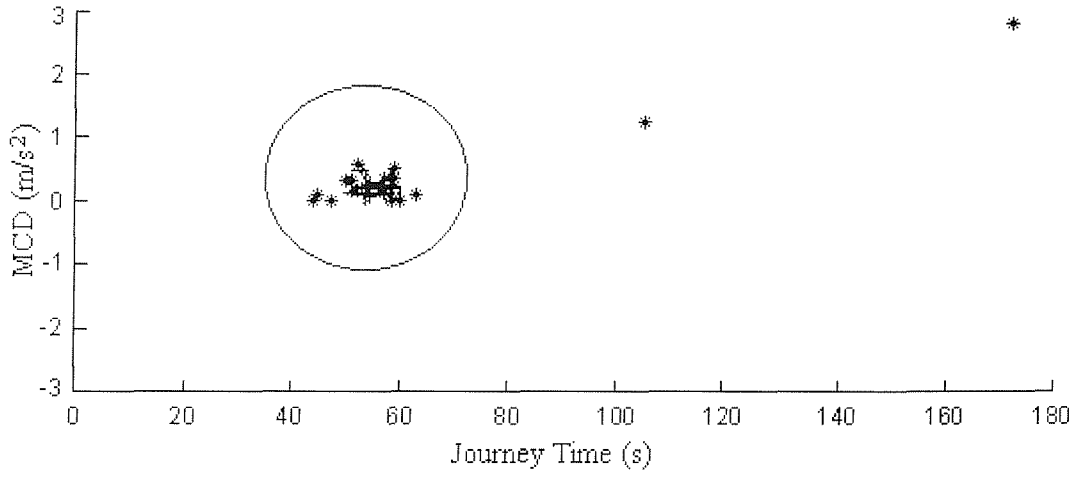


Figure 8.11 Incident detection on link 1

There are two incidents declared. The detection rate is 100% and the false alarm is zero. As introduced in Section 6. 5, good performance of the BEAM model has also been obtained due to stable traffic on link 1. Applying the model to the journeys on link 2, large false alarm rate is obtained while the detection rate remains good. Detection results on link 1 and link 2 are summarised in Table 8.4.

Table 8.4 Incident detection result on link 1 and link 2

	Link 1	Link 2
Detection Rate	100%	100%
False Alarm Rate	0%	14.63%

For link 2 journey conducted at 7:50 on the 8th Oct. 2002, although the measured journey time of 156 seconds is in the confidence level of the historical journey time, the MCD of 4.71 indicates occurrence of incident. By using MCD, the model enables detecting an incident in a very early stage. If a probe vehicle is travelling in middle stage of an incident, a small value of MCD may be obtained but long journey time can declare the incident. By combining journey time and MCD, the model performs well in detection rate. However, the model is unable to classify severe congestion caused by busy traffic and high false alarm rate can be obtained when applying to links carrying busy traffic.

As discussed in Section 6.4.2, 99% coverage contour can produce better incident rate but worse false alarm than 99.9% contour. Since using 99.9% contour has achieve detection rate of 100%, 99% coverage contour is not used for active probe vehicle.

8.5 Summary

The study described in this chapter uses active probe vehicle to characterise traffic flow. Since GPS can provide instantaneous and accurate speed measurement, GPS equipped probe vehicle can provide an efficient tool for journey time estimation and incident detection based on detailed analyses of speed profiles.

In contrast to research on passive probe vehicles, in which journey times estimation directly records the journey times of probe vehicle and calculates the mean of journey times from a number of probe vehicles, this chapter describes a new approach to estimate journey time using a single probe vehicle. According to the features extracted from the speed profiles, the driving pattern of a probe vehicle is classified by using fuzzy sets. Differing from the traditional concept in research of driving behaviour, the driving pattern in this study is only associated with difference between journey time of the probe vehicle and mean journey time. A new variable, the Maximum Continuous Acceleration (MCA) is introduced to reflect acceleration characteristics of the driver by combining continuous acceleration and the speed at the acceleration start point. MCA and average speed of probe vehicle are taken as the input variables of fuzzy sets. The fuzzy memberships are determined by historical traffic data. Journey time is calculated

by corresponding equation for different driving patterns. Comparison of estimated journey time and actual average journey time illustrates the value of the approach.

The BEAM model developed in Chapter 6 is still used in active probe vehicle data. One input variable of the model, difference of journey time between two adjacent time intervals is replaced by the Maximum Continuous Deceleration (MCD), a variable derived from speed profile of the probe vehicle. Values of MCD obtained from different journeys have been shown to be normally distributed in incident-free condition and the distribution is uniform for different time intervals and different links. Combination of MCD and journey time of probe vehicle can achieve good detection rate. A weakness of the model is the lack of an effort to distinguish incidents from other congestion producing traffic phenomena. Thus, poor performance in false alarm rate has been obtained when the model was applied to links carrying busy traffic.

Chapter Nine

Summary and Conclusions

This research has studied GPS-equipped probe vehicles in collecting traffic data. Since this research started after discontinuing the use of Selective Availability (SA), the main error source of GPS, stand-alone GPS receivers were used in data collection. Performance of GPS in journey time estimation has been studied. Journey times estimated by GPS were compared with results from Automatic Number Plate Recognition (ANPR) cameras for assessing estimation accuracy. Data were collected on the M3 Junction 11 to 14, and the M27 Junction 2 to 11, where ANPR cameras have been operational. GPS equipped vehicles were driven by different drivers on the survey site.

This research has studied two types of probe vehicles: active and passive. In this research, a passive probe vehicle was considered to provide only point-to-point journey time, while an active probe vehicle could provide continuous GPS data over an entire journey, including time, position and speed at 1 Hz frequency.

A key issue concerning passive probe vehicles is sample size, which can be affected by many factors, e.g. traffic flow and link geometric condition. Statistical sampling theory has been used to study journey time distribution of individual vehicles and estimate the required sample size in different traffic conditions. Incident detection using journey time measured by probe vehicles was based on the premise that an incident causes more significant increase in journey time than busy traffic. A BivariatE Analysis Model (BEAM) has been developed using journey time at current time interval and difference of journey times between two adjacent time intervals. Although probe vehicles provide real-time and near real-time journey time, many ITS applications (e.g. vehicle navigation systems) require estimate of future journey times. In this research, journey time predictions have been studied for incident and incident-free traffic

conditions. For non-incident traffic, predictions are based on historic data and current observation. After declaration of an incident, journey times were predicted according to the magnitude of increase in journey time at beginning of the incident.

Active probe vehicle can provide speed profile through a journey rather than journey time only. By analysing speed profile, journey time can be estimated from fewer probe vehicles than normally required. In this research, a fuzzy model has been developed to analyse the speed profile, and journey time could be estimated using a single probe vehicle. Good estimates were obtained in both non-incident and incident conditions. For active probe vehicle, BEAM model was used in incident detection. One variable representing deceleration calculated from speed profiles was used as input of the model to replace the difference of journey time between two adjacent time intervals.

9.1 Main findings of the research

- Data collected in this research have shown that in motorway area, GPS performed well in positioning accuracy and availability of service. Journey time can be estimated accurately using GPS positioning data at updating frequency of 1 Hz. Good estimates can also be obtained using long updating intervals up to 30 seconds. For updating interval of one minute, estimation accuracy is associated with traffic flow condition. Poor estimates may be obtained in traffic flow breakdown due to non-stability of speed in such a long interval.
- The journey time distribution of individual vehicles in a 5 minute interval on the same motorway link has been studied. The common assumption that link journey time has a normal distribution has been shown, although non-normality has been observed in some cases. Non-normality of journey time distribution is most likely found when the link journey times change rapidly and greatly. The minimum sample sizes of probe vehicles were calculated for the seven links studied in this research. In general, 4-6 probe vehicles on the motorway links can estimate journey times for 10% permitted error and 95% confidence level. Results have shown shorter links may require more probe vehicles. For the same link, the sample size decreases with increasing traffic flow.

- Joint distribution of link journey time at current time interval and difference of journey times between two adjacent time intervals have been shown to be bivariate normal in incident-free traffic. Occurrence of an outlier in the distribution can be considered to declare an incident. The method can achieve detection rate of 90.5% and false alarm rate of 0.71%. The method performed better in shorter links since the increase of journey time caused by an incident was more significant in shorter links than in longer links.
- Journey time at the next 5 minute interval can be predicted by a linear model using journey time observation at the current interval. Kalman filter algorithms could be used to predict journey time for next thirty minutes. Prediction errors have been found to increase with the length of prediction interval. Up to 20 minute intervals, the average prediction errors remains under 10%.
- Journey time in incident traffic has been shown to have different characteristics and a different prediction strategy should be used after an incident declared. Change of journey time during an incident period can be considered as breakdown stage and recovery stage. Decreasing journey time in breakdown stage may be described by a hyperbolic cotangent function and increasing journey time in recovery stage may be described by an exponential function. However, due to complexity of change in journey time with various incidents, the method is not universal.
- A GPS equipped probe vehicle can provide continuous speed measures of the vehicle at 1 Hz frequency. Fuzzy logic can be used to analyse the speed profiles and identify the driving style. By removing the influence of driving behaviour, the fuzzy logic model can obtain estimates of journey time which are closer to the average journey time of all vehicles on the same link for the same time interval.
- Speed profiles can also be used for incident detection. Estimated journey time and deceleration rate derived from speed profiles have been shown independent and normally distributed in incident-free traffic. An incident can be declared, if either journey time or deceleration rate of a journey exceeds a confidence level of the normal distribution. The method can achieve good detection rate but relatively high false alarm rate.

9.2 Potential applications

An increasing number of GPS equipped vehicles are already running in the traffic stream. Some of them have been used as probe vehicles to collect journey times. Results provided by this research should be useful for designing journey time data collection efforts and systems, performing journey time studies, and summarising journey time data. This research focused on link journey time estimation. Link journey times can be cumulated to generate route journey times. Link journey time is more flexible in use than route journey times. For example, car navigation systems consider a road network as database of links. Therefore, providing real time journey time of each link is critical for dynamic route planning.

Probe vehicle can be expected to play an important role in collecting traffic data in the near future. By mid 2002, over 1.5 million mobile phones in U.S. have been equipped with GPS to meet Federal Communications Commission (FCC) requirements of locating emergency callers (Borras, 2003). Vehicles carrying the GPS equipped mobile phones are therefore potential probe vehicles if regularly reporting locations. Location based road charging schemes have been proposed in many countries. The approach enables a real-time assessment of road user charges based on time, location and distance travelled. GPS based location devices have been considered as one of potential choices. In such a scheme, GPS devices can be used also for traffic data collection and sufficient sample size of probe vehicles is possible (Srinvivasan et al., 2002). Suppliers of vehicle navigation systems have realised that good route guidance should be based on real-time traffic information. Since traffic information should be provided based on vehicle location, two-way communication has been emerged in vehicle navigation systems. Through frequent reports of location, the vehicles can be used as probe vehicles (Li and McDonald, 2003).

The European Commission has supported a new generation satellite navigation system, known as Galileo. The fully deployed Galileo system consists of 30 satellites, providing positioning accuracy to within 45 cm on the earth surface. Galileo is scheduled to lunched the first satellites in 2005 and provide full user capabilities between 2007 and 2008 (ESA, 2003). The Galileo initiative has triggered the search for commercial and mass-market location-enabled services, including road charging and

furthermore traffic data collection. Applications of probe vehicles will be promising with Galileo.

Location technology applications within the automobile industry are dominated by GPS. However, mobile phone location technologies have also been developed. On GSM network, currently sub-50 meters accuracy has been delivered. With the introduction of third generation networks, the accuracy can be improved to 20 meters (VTT, 2003). The mobile phone location technologies enable any motorist carrying a mobile phone to be a traffic data provider. Sample size of probe vehicles is expected to be further enlarged.

9.3 Further work

As discussed above, a wide range of vehicles can be potential probe vehicles. Widespread implementation of probe vehicle data collection systems has created a pressing demand for standards and protocols to provide interoperability and compatibility between various data sources. Standards and protocols will combine all potential probe vehicle data sources, such as car navigation systems, bus and taxi management. Standards and protocols will also enable integrating location data from different technologies, such as GPS, Galileo and mobile phone. The unification would provide a sufficiently large sample size of probe vehicles to collect traffic data over an entire road network.

The traditional way of defining the state of network is by determining the relationships between the three parameters of flow, speed and density and creating journey time by combining these parameters. Probe vehicles with GPS and mobile communication systems enable direct communication between vehicles and central traffic monitoring systems, providing real-time or near real-time journey time. Although journey time is a fundamental measure in transportation, it is unable to provide a comprehensive estimation of the state of the network. For example, the AID model developed in this research can declare an incident, which has caused a significant increase in journey time. The detection time may be longer than AID algorithms using signals from inductive loops. Moreover, without comprehensive understanding of the state of all

parts in the road network, good prediction of journey time is difficult to achieve, especially in an incident period.

In future ITS applications, it is likely to be advantageous to integrate data from various sources to make more informed journey time estimations and incident detection decisions. Although there is a small number of studies with data from more than one methodology, most traffic data research has concentrated only on one form of detection. It will be desirable to further study how data from various sources can be fused to provide accurate and comprehensive traffic information direct to network operators and motorists. The study on data fusion will be helpful in design and implementation of surveillance systems.

Appendix A

Converting ellipsoidal latitude and longitude to grid eastings and northings

For National Grid TM Projection,

Map coordinate of northing origin: y_0 = -100 000m

Map coordinate of easting origin: x_0 = 400 000m

Latitude of true origin: ϕ_0 = 49° N

Longitude of true origin: λ_0 = 2° W

To convert a position from the graticule of latitude and longitude coordination (ϕ, λ) to a grid of easting and northing coordinate (x, y) using a National Grid Transverse Mercator (TM) Projection, compute the following formulae:

$$\begin{cases} y = I_0 + N \\ x = x_0 + E \end{cases}$$

where: $I_0 = y_0 - N_0$

Northing and easting coordinates are computed by the following equations. N_0 is the corresponding northing coordinate of true origin, i.e. (49° N, 2° W), calculated by Equation A-1.

$$\begin{aligned} N = {} & kS(\phi) + kn[\frac{1}{2}\cos^2\Phi(\lambda-\lambda_0)^2 + \frac{t}{24}\cos^4\Phi(5-t^2+9\eta^2+4\eta^4)(\lambda-\lambda_0)^4 \\ & + \frac{t}{720}\cos^6\phi(61-58t^2+t^4+270\eta^2-330t^2\mu^2)(\lambda-\lambda_0)^6 \\ & + \frac{t}{40320}\cos^8\phi(1385-3111t^2+543t^4-t^6)(\lambda-\lambda_0)^8+\cdots] \end{aligned} \quad \text{(A-1)}$$

$$\begin{aligned} E = {} & kN[\cos\phi(\lambda-\lambda_0) + \frac{1}{6}\cos^3\Phi(1-t^2+\eta^2)(\lambda-\lambda_0)^3 \\ & + \frac{1}{120}\cos^5\phi(5-18t^2+t^4+14\eta^2-58t^2\mu^2)(\lambda-\lambda_0)^5 \\ & + \frac{1}{5040}\cos^7\phi(61-479t^2+179t^4-t^6)(\lambda-\lambda_0)^7+\cdots] \end{aligned} \quad \text{(A-2)}$$

where:

Arc length of meridian: $S(\phi)$

Radius of curvature in prime vertical: $N = \frac{a^2}{b\sqrt{1+\eta^2}}$

Auxiliary quantity: $\eta^2 = \frac{a^2 - b^2}{b^2}\cos^2\phi$

Auxiliary quantity: $t = \tan(\phi)$

Longitude of the central meridian: λ_0

Scale factor on central meridian: $k=0.9996012717$

The Arc length of the meridian $S(\phi)$ is the ellipsoidal distance from the equator to the point to be transformed and is given by the series expansion:

$$S(\Phi) = \alpha[\phi + \beta\sin 2\phi + \gamma\sin 4\phi + \delta\sin 6\phi + \varepsilon\sin 8\phi + \cdots]$$

For WGS-84 reference system the parameters are listed here:

$\alpha = 6367449.1458m$

$\beta = -2.51882792\times10^{-3}$

$\gamma = 2.64354\times10^{-6}$

$\delta = -3.45\times10^{-9}$

$\varepsilon = 5\times10^{-12}$

Appendix B

Sample of Normal Distribution

Appendix B1

Notation of $z_{\alpha/2}$

$z_{\alpha/2}$ is the upper $\alpha/2$ point of the standard normal distribution. For a standard normal distribution, $100(1-\alpha)\%$ of the distribution is contained within the range $[-Z_{-\alpha/2}, Z_{\alpha/2}]$.

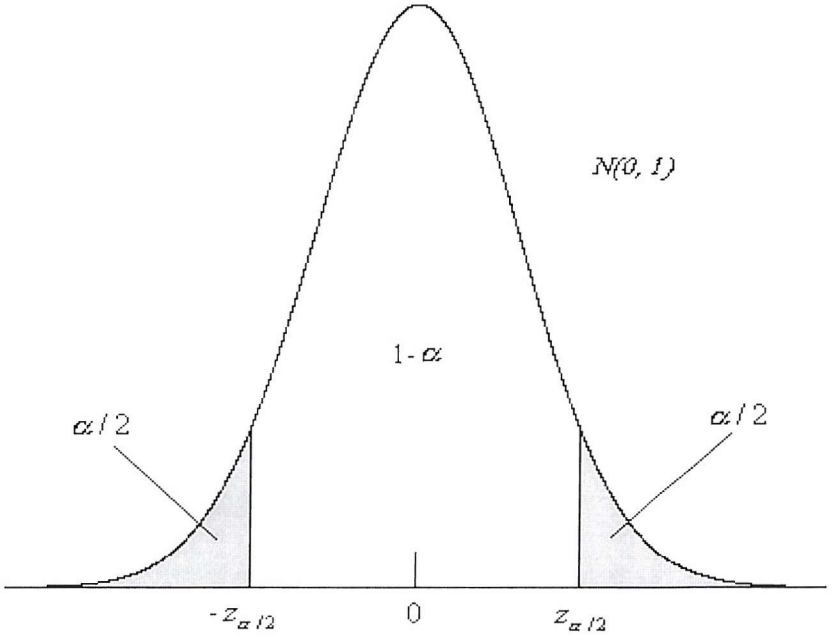


Figure B-1 the notation of $z_{\alpha/2}$

Appendix B2

Determination of Sample Size

At the planing stage of a statistical investigation, the question of sample size *n* is critical. Let parameters of a population be:

Population mean = μ

Population standard deviation = σ

Statistical inferences about the population mean are based on the sample mean

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

The distribution of the sample mean $\overline{X}$, based on a random sample size n, has:

$$E(\overline{X}) = \mu \qquad (= \text{population mean})$$

$$sd(\overline{X}) = \frac{\sigma}{\sqrt{n}} \qquad \left(= \frac{population\ standard\ deviation}{\sqrt{sample\ \ size}}\right)$$

The first result shows that the distribution of $\overline{X}$ is centred at the population mean μ in the sense that expectation serves as a measure of centre of a distribution. The second result states that the standard deviation of $\overline{X}$ equals the population standard deviation divided by the square root of the sample size. That is, the variability of the sample mean is governed by the two factors: the population variability σ and the sample size n.

In random sampling from a normal population, the sample mean $\overline{X}$ has the normal distribution. When the sample from non-normal population, if the sample size n is large, the distribution $\overline{X}$ is approximately normal. That statement is known as central limit theorem. In practice, the normal approximation is usually adequate when n is greater than 30.

For permitted error e and a confidence level $1-\alpha$:

$$Probability \left\{ \left| \frac{\overline{X} - \mu}{\mu} \right| < e \right\} \geq 1 - \alpha \qquad \text{(B-1)}$$

If $\overline{X}$ is the normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$, $\overline{X}$ can be converted to be standard normal distribution as:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \text{(B-2)}$$

Substituting Equation B2-2 in Inequality B2-1,

$$Probability\left\{\left|\frac{Z\sigma/\sqrt{n}}{\mu}\right| < e\right\} \geq 1-\alpha \qquad \text{(B-3)}$$

Rewriting Inequality B-3 as:

$$Probability\left\{|Z| < \frac{e\mu}{\sigma/\sqrt{n}}\right\} \geq 1-\alpha \qquad \text{(B-4)}$$

Since Z is standard normal distribution, according to Figure B-1, Inequality B-4 is equal to:

$$\frac{e\mu}{\sigma/\sqrt{n}} \geq z_{\alpha/2} \qquad \text{(B-5)}$$

Solving for n, it can be obtained:

$$n \geq \left[\frac{z_{\alpha/2}\sigma}{e\mu}\right]^2 \qquad \text{(B-6)}$$

Appendix C

Bivariate Data Analysis

Appendix C1

Bivariate Normal Distribution

In bivariate analysis, the relationship between the random variables X and Y is studied. The relationship between the two variables can be described by the *joint probability distribution,* which assigns probabilities to all possible outcomes (x, y). The joint distribution is usually characterized by the *joint density function f(x,y).*

A measure of the strength and direction of association between the variables X and Y is provided by the covariance which is defined by:

$$\sigma_{xy} = E[(X-\mu_x)(Y-\mu_y)] = E[XY] - \mu_x\mu_y \qquad \text{(C-1)}$$

The covariance matrix for the joint distribution is given by

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \qquad \text{(C-2)}$$

A relatively large value of σ_{xy} indicates a strong relationship between X and Y. An index of covariance between X and Y is provided by the correlation coefficient $\rho = \sigma_{xy}/\sigma_x\sigma_y$. This index has the range –1 to +1. When compared to 1, the absolute value of ρ indicates the strength of linear association between X and Y. For quantitative variables $|\rho| = 1$ is equivalent to a perfect linear relationship. The slope of the linear relationship is indicated by the sign of ρ. If $\rho = 0$, X and Y are independent. The correlation matrix for the joint distribution is given by $\rho = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$.

The bivariate normal density is given by:

$$f(x,y) = c_1^{-1}\exp\{c_2[(x-\mu_x)^2/\sigma_x^2 + (y-\mu_y)^2/\sigma_y^2 - 2\rho(x-\mu_x)(y-\mu_y)/\sigma_x\sigma_y]\} \qquad \text{(C-3)}$$

where $c_1 = 2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}$ and $c_2 = -1/[2(1-\rho^2)]$. The bivariate normal density

contains five parameters $\mu_x, \mu_y, \sigma_x, \sigma_y$ and ρ. Figure C-1 shows a bivariate normal surface. The three dimensional picture shows that the density at (x, y) is given by the height $f(x,y)$. The probability for any region of values $(x_0 \le X \le x_1, y_0 \le Y \le y_1)$ is given by the volume under the normal surface.

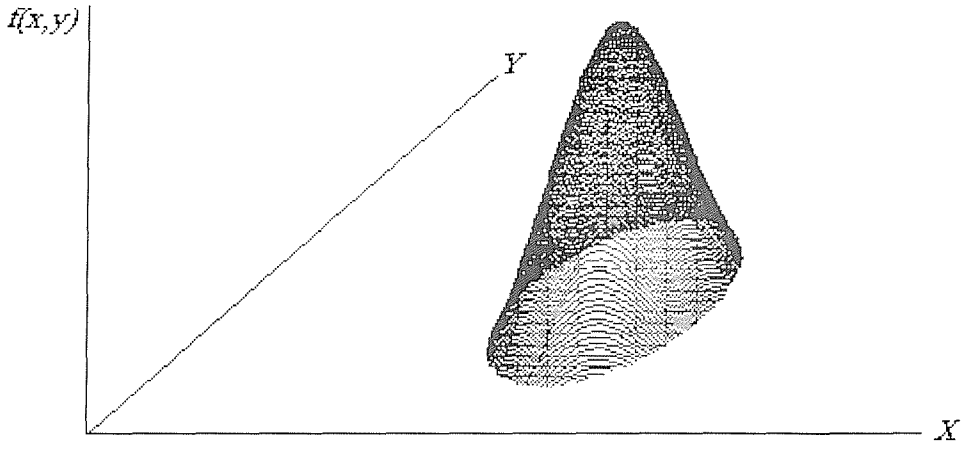


Figure C-1 Bivariate Normal Surface

The marginal density for X is obtained from the integral expression:

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp[(-1/2)(\frac{x-\mu_x}{\sigma_x})^2] \qquad \text{(C-4)}$$

and similarly for Y

$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp[(-1/2)(\frac{y-\mu_y}{\sigma_y})^2] \qquad \text{(C-5)}$$

These densities are the familiar univariate normal densities. The marginal densites for the bivariate normal therefore are also normal. It is not necessarily true however that X and Y will be bivariate normal if X and Y are both univariate normal. If $\rho = 0$ the joint density for the bivariate normal can be written as the produce of the marginal densities and hence X and Y are independent. Thus for the bivariate normal, independence and zero correlation are equivalent.

For a univariate distribution, a plot of the density of the standardized random variable $(\frac{X-\mu_x}{\sigma_x})$ is useful for making comparisons with other densities such as the normal distribution. For a normal random variable X, $100(1-\alpha)\%$ of the distribution is

contained within the range $[-Z_{-\alpha/2}, Z_{\alpha/2}]$. For a bivariate distribution, $100(1-\alpha)\%$ of the distribution contained within an elliptical contour in the X-Y plane. The equation

$$(\frac{1}{1-\rho^2})[(\frac{X-\mu_x}{\sigma_x})^2+(\frac{Y-\mu_y}{\sigma_y})^2-2\rho(\frac{X-\mu_x}{\sigma_x})(\frac{Y-\mu_y}{\sigma_y})]=\kappa \qquad \text{(C-6)}$$

describes an ellipse in the X-Y plane with centre at (μ_x, μ_y), as shown in Figure C-2. As k increases the area of the ellipse increases. The equation for the elliptical contour can also be written in matrix notation as

$$\begin{bmatrix} X-\mu_x \\ Y-\mu_y \end{bmatrix}' \Sigma^{-1} \begin{bmatrix} X-\mu_x \\ Y-\mu_y \end{bmatrix} = \kappa$$

For the bivariate normal density, the constant k on the right-hand side of Equation C-6 is equal to $x^2_{\alpha;2}$, where $x^2_{\alpha;2}$ denotes the value of a x^2 random variable with 2 degrees of freedom and a p-value of α in the upper tail. The elliptical contour will on average contain $100(1-\alpha)\%$ of the sample points.

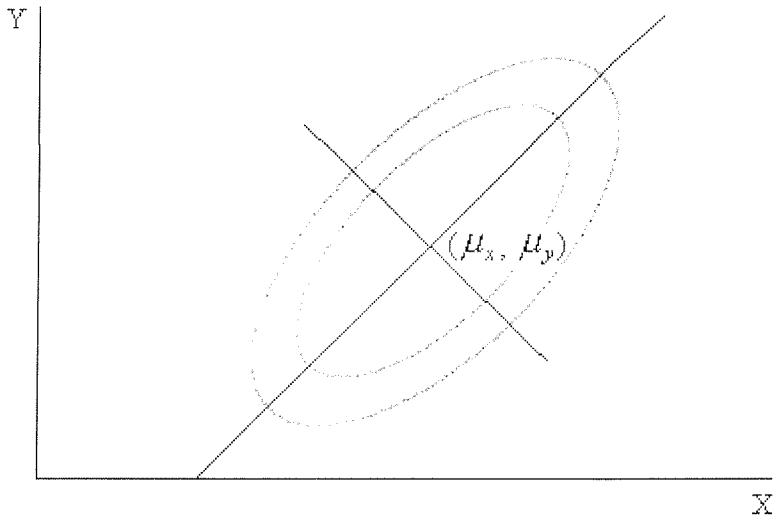


Figure C-2 Elliptical Contours for bivariate normal density

Appendix C2

Assessing Bivariate Normality

For the bivariate normal distribution, marginal distributions must be normal. However, the marginal normality is necessary but not sufficient conditions. To assess bivariate

normality, the Mahanobis Distance is used. The left-hand side of Equation C-6 measures the square of the Mahalanobis Distance, which describes the distance between any pairs of (x, y) and the center (μ_x, μ_y). For bivariate normal distribution, the squared Mahalanobis Distances are distributed as Chi-Square with 2 degrees freedom. The assessment is executed by computing the squared Mahalanobis Distances of each observation and checking to see that they are chi-square distribution by Q-Q plot:

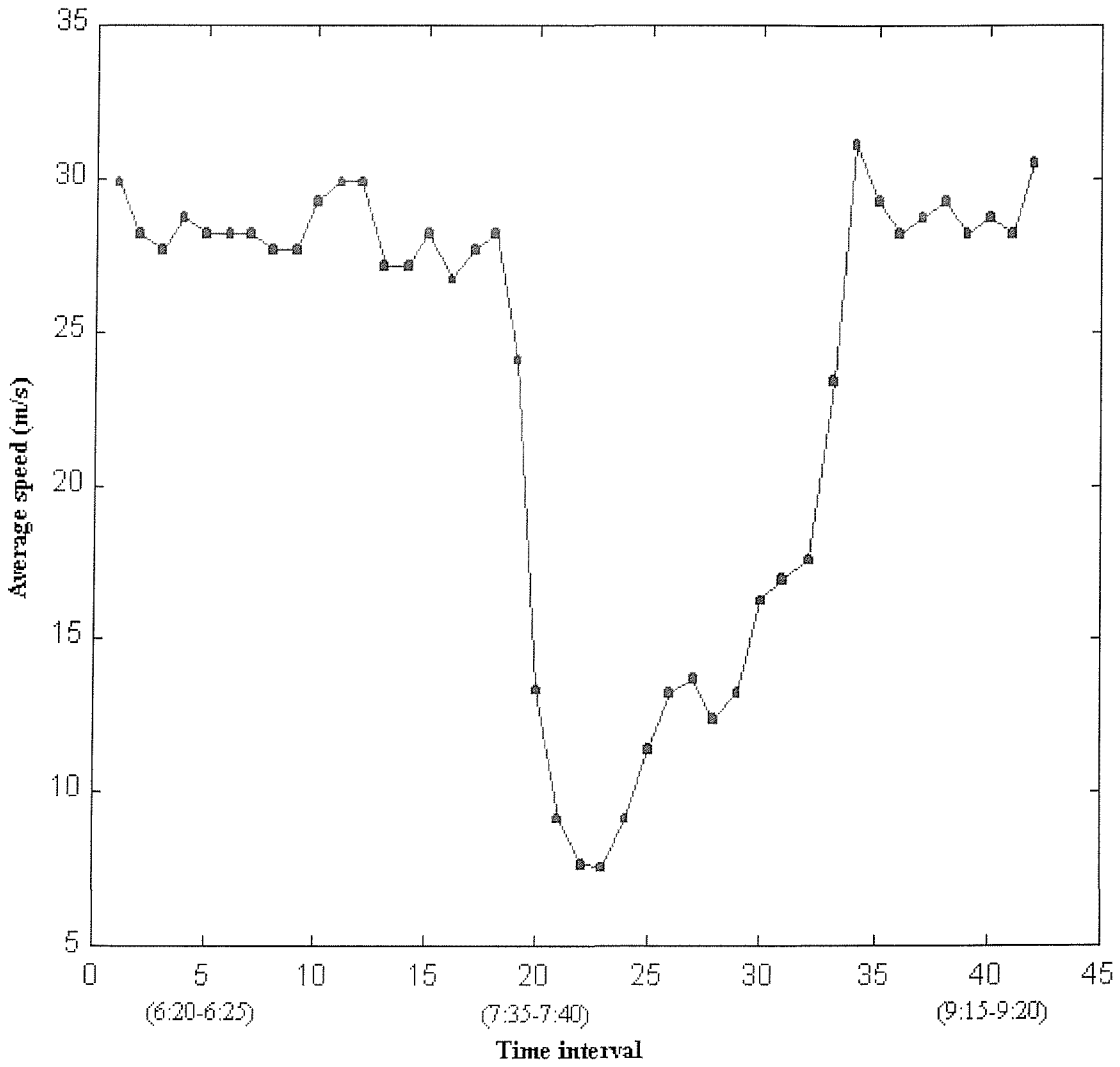
1. Compute the squared Mahalanobis Distances d_i^2, i= 1, 2, ..., n
2. Order the d_i^2 from smallest the largest to get observed quantiles of the distribution as: $d_1^2 \le d_2^2 ... \le d_n^2$
3. Compute expected quantiles from the χ_2^2 distribution, where $q_i = \chi_2^2(\frac{i-1/2}{n})$, corresponding with each d_i^2, i= 1, 2, ..., n
4. Plot d_i^2 versus q_i for i= 1, 2, ..., n and check for linearity in the plot. If the points do not form a straight line, then the observed quantiles do following from the chi-square distribution, so accept bivariate normality.

Appendix D

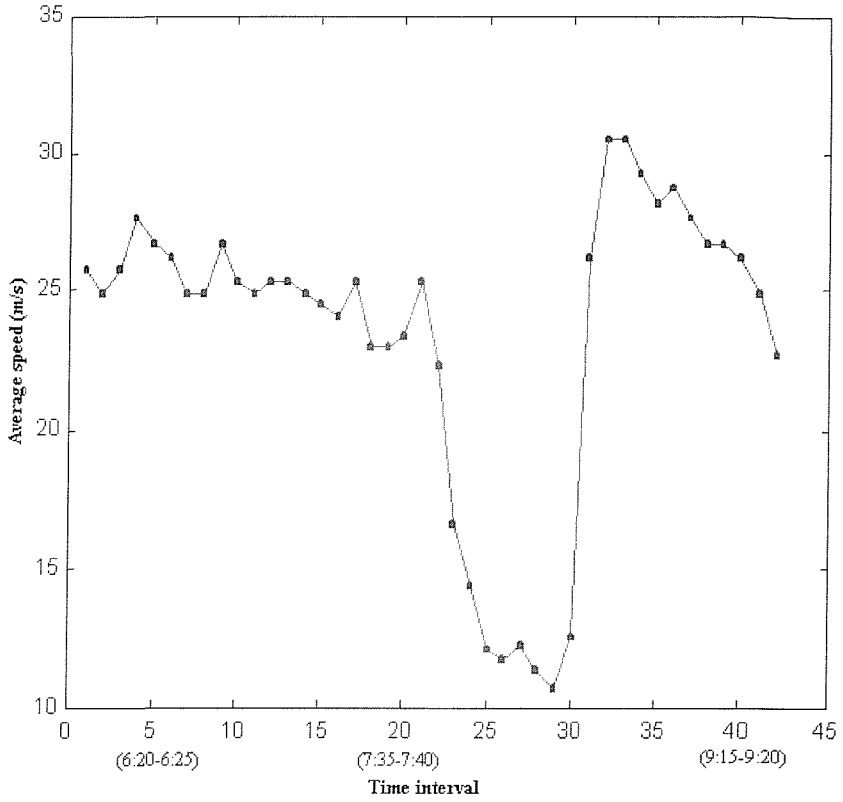
Speed Changes in Incident Periods

Changes in journey time of several incidents on link 1 and link 2 are shown in this appendix. It may be considered that changes in journey time over an entire incident period have shown substantially similar characteristics for various types of incident and various links.

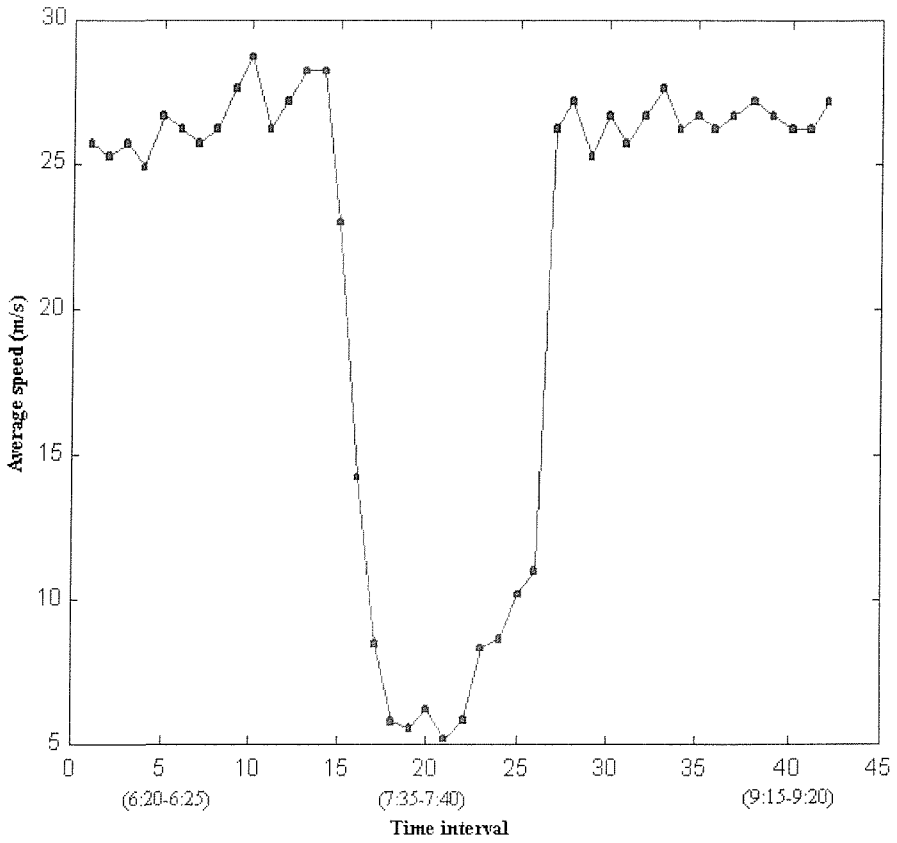
1. May 24th, 2001, on link 1



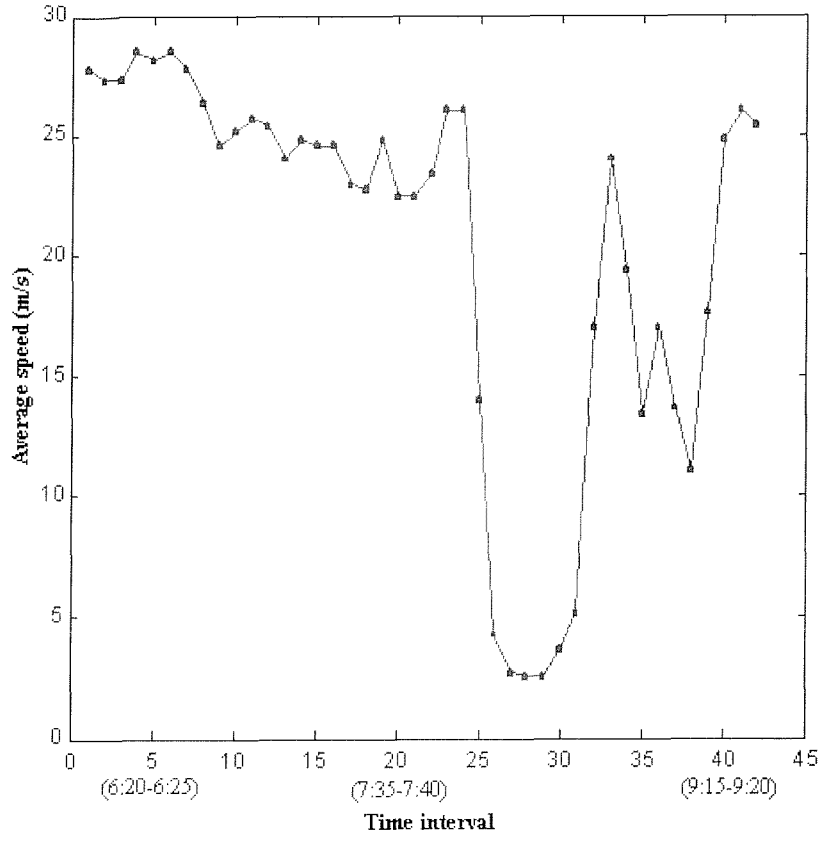
2. August 28th, 2002, on link 1



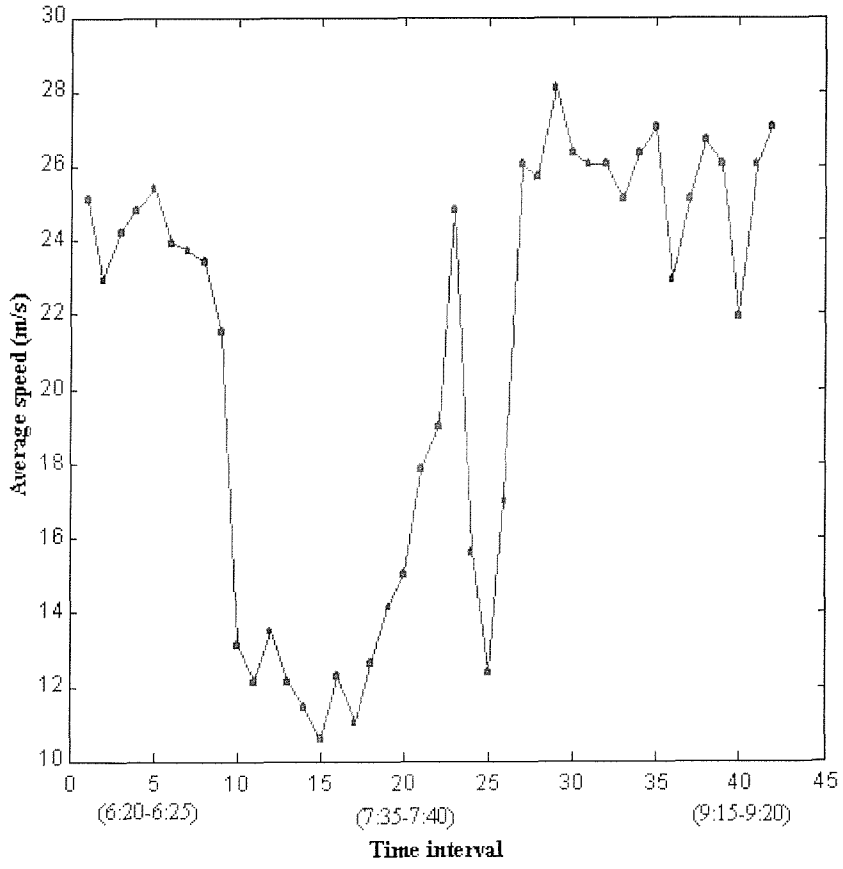
3. October 11th, 2002, on link 1



4. July 19th, 2001, on link 2



5. October 8th, 2002, on link 2



References

Abou-Rahme N., Beale S., Harbourd B. and Hardman E., 2000, Monitoring and Modelling of Controlled Motorways. In *Proceedings of the Tenth International Conference on Road Transport Information and Control*, London, 4-6 April, pp. 84-90.

Abou-Rahme N., White J. and Stewart R., 2002, Journey Time Estimation Using MIDAS Loop Data. In *Proceedings of the Eleventh International Conference on Road Transport Information and Control*, London, 19-21 March, pp. 177-181.

Adaway B., 2001, Highway Vision. *Traffic Technology International 2001: Annual Review*, pp. 119-123.

Adeli H. and Karim A., 2000, Fuzzy-Wavelet RBFNN Model for Freeway Incident Detection. *Journal of Transportation Engineering,* Vol. 126(6), pp. 464-471.

Ahmed S. A. and Cook A. R., 1982, Application of Time-Series Analysis Techniques to Freeway Incident Detection. *Transportation Research Record*, Vol. 841, pp. 19-21.

Balke K., Dudek C. L. and Mountain C. E., 1996, Using Probe-Measured Travel Time to Detect Major Freeway Incidents in Houston, Texas. *Transportation Research Record*, Vol. 1554, pp. 213-220.

Barnes J., 2000, Guiding Principal: Getting A Fix on GPS, *Traffic Technology International*, Oct/Nov, pp 56-58.

Bartlett D. and Morris P., 2002, CVB: A Technique to Improve OTDOA Positioning in 3G Networks. Technical White Paper, *Cambridge Positioning Systems Limited.*

Bibaritsch M. and Egeler C., 2002, EFC Enforcement and Automatic License Plate Reading, In *Proceedings of 9th world Congress on Intelligent Transport Systems*, Chicago, 14-17 October.

Borras K., 2003, Globaly Playing it Safe. Traffic Technology International, Feb/Mar, pp. 36-38.

Brackstone M., Fisher G. and McDonald M. 2001, The Use of Probe Vehicles on Motorways, Some Empirical Observations. In *Proceedings of the 8th World congress on Intelligent Transport Systems*, Sydney, Australia, 4-8 November.

Brackstone M., Sultan B. and McDonald M., 2002, Motorway Driver Behaviour: Studies on Car Following. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 5(1), pp31-46.

Bunn J. and Barrett D. J., 1997, Automatic Number Plate Reading: What Are The Difficulties? In *Proceedings of 4th world Congress on Intelligent Transport Systems,* Berlin, 21-24 October.

Carvell J. D. Jr., Balke K., Ullman J., Fitzpatrick K., Nowlin L. and Brehmer C., 1997, Freeway Management Handbook. Texas Transportation Institute, *Technical Report*, Report Number: FHWA-SA-97-064.

Catling I., 2000, The Prospects for Electronic Fee Collection (EFC) Using Vehicle Positioning Systems. In *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin, 6-9 November.

Chen M. and Chien S. I. .J., 2000, Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation. *Transportation Research Record*, Vol. 1719, pp. 61-68.

Chen, M. and Chien, S. I. J., 2001, Dynamic Freeway Travel-time Prediction with Probe Vehicle Data, *Transportation Research Record*, Vol. 1768, pp. 157-161.

Cheu A. R., Xie C., and Lee D-H., 2002, Probe Vehicle Population and Sample Size for Arterial Speed Estimation. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 17 (1), pp. 53-60.

Cheu R. L., Ritchie S. G., 1995, Automated Detection of Lane-Blocking Freeway Incidents Using Artificial Neural Networks. *Transportation Research Part C*, Vol. 3 (6), pp. 376-388.

Cheu R. L., Qi H., and Lee D-H, 2002, A Mobile Sensor and Sample-Based Algorihms for Freeway Incident Detection. In *Proceedings of Transportation Research Board 81st Annual Meeting*, Jan., Washington, D.C., 13-17 January.

Chien S. I. J. and Kuchipudi C. M., 2002, Dynamic Travel Time Prediction with Real-time and Historical Data. In *Proceedings of the Transportation Research Board 81st Annual Meeting,* Washington D. C., 13-17 January.

Chien S. I. J., Liu X. and Ozbay K., 2003, Predicting Travel Times for the South Jersey Real-time Motorist Information System, In *Proceedings of the Transportation Research Board 82nd Annual Meeting,* Washington D. C., 12-16 January.

Clark S. D., Grant-Muller S. and Chen H., 2002, Cleaning of Matched License Plate Data. *Transportation Research Record,* Vol. 1804, pp. 1-7.

Clark S. D. and McKimm J., 2003, Using GPS to Collect Journey Time Information. *Traffic Engineering and Control*, Vol. 44 (4), pp. 124-131.

Cowan K. W. and Gates G, 2002, Floating Vehicle Data System – A Smart Move. In *Proceedings of 9th World Congress on Intelligent Transport Systems,* Chicago, 14-17 October.

D'Angelo, M. P., Al-Deek, H. M. and Wang M. C., 1999, Travel-time Prediction for Freeway Corridors. *Transportation Research Record*, Vol. 1676, pp. 184-191.

D'Este, G. M., Zito, R. and Taylor M.A.P., 1999, Using GPS to Measure Traffic System Performance. *Computer-aided and infrastructure engineering,* Vol. 14(4), pp. 255-265.

Dia H. and Rose G., 1997, Development and Evaluation of Neural Network Freeway Incident Detection Models Using Field Data. *Transportation Research Part C*, Vol. 5(5), pp. 313-331.

Drew D. R., 1968, Traffic Flow Theory and Control. *McGraw-Hill Inc.* pp. 372-376.

Dunne D., 2003, What is 3G Technology, [Online], Available: http://www.darwinmag.com/learn/.

Ellis S., Lansdonw T. and Richardson J., 2000, The Road Traffic Advisor Project, In *Proceedings of 10th International Conference On Road Transport Information and Control*, London, 4-6 April, pp. 1-5.

Ellis T. 2002, Automatic License Plate Recognition in Law Enforcement. *Traffic Engineering and Control,* Vol. 43(11), pp. 416-419.

Eloranta T., Kokkinen M. and Rajala P., 2000, Effectiveness of Licence Plate Recognition Based on Journey Time Monitoring in The Demanding Nordic Environment. In *Proceedings of 7th world Congress on Intelligent Transport Systems,* Turing, Italy, 6-9 November.

Ericsson E., 2001, Independent Driving Pattern Factors and their Influence on Fuel-Use and Exhaust Emission Factors, *Transportation Research Part D*, Vol. 6, pp. 325-345.

European Space Agency (ESA), 2003, Galileo Becomes a Reality for Europe. [Online], Available: http://www.esa.int/export/esaCP/SEMBOBS1VED_index_0.html

Faghri A, Hamad K., Hehman D. and Henck H. (1999), Application of GPS to Travel Time and Delay Measurements--1997 Phase, *Department of Civil and Environmental Engineering, University of Delaware*, Delaware, USA.

Federal Highway Administration (FHWA), 1998, Developing Freeway and Incident Management Systems Using the National ITS Architecture. *U.S. Department of*

Transportation, Intelligent Transportation Systems Joint Program Office, Report Number: FHWA-JPO-98-032.

Feng S. and Law C. L. 2002, Assisted GPS and its Impact on Navigation in Intelligent Transportation Systems. In *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, Singapore, 3-6 Sept.

Frith B. and Pearce D., 2002, Driver Information on Journey Time Variability Generated Using ANPR Data. In *Proceedings of 11th International Conference On Road Transport Information and Control*, London, 19-21 March, pp. 190-195.

Garmin Corp., 2003, GPS 35 Series Technical Specification. [Online], Available: http://www.garmin.com/products/gps35/spec.html.

Gourld C., Umnro P. and Hardman E., 2002, M3/M27 Ramp Metering Pilot Scheme (RMPS) Implementation and Assessment. In *Proceedings of Eleventh International Conference on Road Transport Information and Control*, London, 19-21 March, pp. 161-167.

Hall R., Vyas N., Shyani C., Sabnami V. and Khetani S., 1999, Evaluation of the OCTA Transit Probe Vehicle System. *PATH Research Report*, Report Nubmer: UCB-ITS-PRR-99-39.

Harvey, A. C., 1989, Forecasting, Structural Time Series Models and the Kalman Filter. *Cambridge University Press*, pp. 100-125.

Hellinga B. and Knapp G., 2000, Automatic Vehicle Identification Technology-Based Freeway Incident Detection, *Transportation Research Record*, Vol. 1727, pp. 142-153.

Highway Capacity Manual, 1994, *Transportation Research Broad.* pp 6.8-6.16.

Hofmann-Wellenhof B., Lichtenegger H. and Collins J., 1993, GPS Theory and Practice. *Wien: Springer-Verlag*, pp. 40-70.

Hounsell N. B., McLeod F. N., Garner K., Head J. R. and Cook D., 2000. Headway-based Bus Priority in London Using AVL: First Results. In *Proceedings of 10th International Conference On Road Transport Information and Control*, London, 4-6 April, 2000, pp 218-222.

Houston TranStar, 2003, *Houston TranStar AVI traffic monitoring system.* [Online], Available: www.houstontranstar.org

Huisken G. and Berkurn E., 2002, Short-term Travel Time Prediction Using Data from Induction Loops, In *Proceedings of 9th World Congress on Intelligent Transport Systems,* Chicago, 14-17 October.

Jin X., Cheu R. L., and Srinivasan D., 2002, Development and Adaptation of Constructive Probabilistic Neural Network in Freeway Incident Detection. *Transportation Research Part C*, Vol. 10, pp. 121-147.

Jobson J. D., 1991, Applied Multivariate Data Analysis Volume I: Regression and Experimental Design, *Springer-Verlag*, pp. 89-121.

Johnson R. A. and Wichern D. W., 1992, Applied Multivariate Statistical Analysis (third Edition), *Prentice-Hall International, Inc.*, pp.152-164.

Johnson R. A. and Bhattacharyya G. K., 2001, Statistics-Principles and Methods. *John Wiley & Sons Inc.*, pp. 300-305.

Karl C. A. and Trayford R. S., 2000, Delivery of Real-time and Predictive Travel Time Information: Experiences from a Melbourne Trial. In *Proceedings of the 7th World Congress on Intelligent Transportation Systems*, 6-9 November, Turin, Italy.

Karim A. and Adeli H., 2002, Comparison of Fuzzy-Wavelet Radial Basis Function Neural Network Freeway Incident Detection Model with California Algorithm. *Journal of Transportation Engineering,* Vol. 128(1), pp. 21-30.

Khan S. I. and Thanasupsin K., 2000, Estimating Link Travel Time on I-70 Corridor: A Real-time Demonstration Prototype. Technical Report, *Department of Civil Engineering, University of Colorado*, Report Number: CDOT-DTD-R-2000-15.

Koelbl R. McDonald M., Fisher G. and Brackstone M., 2002, Probe Vehicle: A Comparison of Motorway Performance Measure with Other Motorway Flow Detection Techniques. In *Proceedings of Eleventh International Conference on Road Transport Information and Control,* London, 19-21 March, pp. 182-186.

Krikke R., 2002, Short-range Travel Time Prediction Using an Artificial Neural Network, In *Proceedings of 9th World Congress on Intelligent Transport Systems,* Chicago, 14-17 October.

Kroes E., Hagemeier F. and Linssen J., 1999, A New Probe Vehicle-based Floating Car Data System: Concept, Implementation and Pilot Study. *Traffic Engineering and Control*, Vol. 40(4), pp 200-204.

Li Y., 2003, Motorway Link Travel Time Estimation Using GPS Equipped Probe Vehicle. In *Proceedings of Universities Transport Study Group 35th Annual Conference*, Loughborough, 6-8 January.

Li Y. and McDonald M., 2002, Travel Time Estimation Using Single GPS Equipped Probe Vehicle. In *Proceedings of the IEEE 5th International Conference on Intelligent Transport Systems*, Singapore, 3-6 September.

Li Y. and McDonald M., 2003, Travel Time Information from Car Navigation Systems. In *Proceedings of VehCom 2003.* 26-27 June, Birmingham, pp. 105-109.

Lin C. K. and Chang G. L., 1998, Development of a Fuzzy-Expert System for Incident Detection and Classification. *Mathematical and Computer Modelling*, Vol. 27(9-11), pp. 9-25.

Lint J. W. C., Hoogendoom S. P. and Zuylen H. J., 2003, Toward a Robust Framework for Freeway Travel Time Prediction: Experiments with Simple Imputation and State-

Space Neural Networks. In *Proceedings of the Transportation Research Board 82nd Annual Meeting,* Washington D. C., 12-16 January.

Lu J., 1992, Spectral Analysis of Vehicle Speed Characteristics. *Transportation Research Record,* Vol. 1375, pp. 26-36.

Mahmassani H. S., Haas C., Zhou S. and Peterman J., Evaluation of incident detection methodologies, Technical Report, *The University of Texas at Austin,* Report Number: FHWA/TX-00/1795-1.

McDonald M., Allsop R. and Bonsall P., (2000), Final Report of Peer Review of Highways Agency Toolkit Projects.

Michalopoulos P. G., Jacobson R. D., Anderson C. A. and DeBruycker T. B., 1993, Automatic Incident Detection through Video Image Processing. *Traffic Engineering and Control,* Vol. 34 (2), pp. 66-75.

Middleton D. and Parker R., 2000, Initial Evaluation of Selected Detectors to Replace Inductive Loops on Freeways. Technical Report, *The Texas A&M University,* Report Number: FHWA/TX-00/1439-7.

Nanthawichit C., Nakatsuji T. and Suzuki H., 2003, Application of Probe Vehicle Data for Real-time Traffic State Estimation and Short-term Travel Time Prediction on a Freeway. In *Proceedings of the Transportation Research Board 82nd Annual Meeting,* Washington D. C., 12-16 January.

Nelson L. J., 2002, Sensors Working Overtimes. *Traffic Technology International,* Feb/Mar, pp. 40-45.

Nelson L. J., 2003, An AVId reader, *Traffic Technology International, April/May,* pp. 72-75.

Ochieng W. Y. and Sauer K., 2002, Urban Road Transport Navigation: Performance of the Global Positioning System after Selective Availability. *Transportation Research Part C*, Vol. 10, pp. 171–187.

Ordnance Survey, 2000, National GPS Network Information: A Guide to Coordinate Systems in Great Britain. [Online], Available on: www.gps.gov.uk/guidecontents.asp.

Petty K. F., Ostland M., Kwon J., Rice J. and Bickel P. J., 2002, A New Methodology for Evaluating Incident Detection Algorithms. *Transportation Research Part C*, Vol. 10, pp. 189-204.

Qualcomm CDMA Technologies Inc. (QCT), 2003, Hybrid Position Location Technology White Paper.

Quiroga C. A. and Bullock D., 1998, Travel Time Studies with GPS and GIS: An Integrated Methodology. *Transportation Research Part C*, Vol. 6(1/2), pp.101-127.

Robertson D. I., Winnett M. A. and Herrod R. T., 1992, Acceleration Signatures, *Traffic Engineering and Control*, Vol. 33(9), pp. 485-491.

Sachse T., 2002, Dual Purpose-Hong Kong's Aberdeen Tunnel, *Traffic Technology International*, April/May 2002, pp.82-111.

Sen A., Soet S. and Berka S., 1997, Probes and Detectors: Experiences Gained from ADVANCE. *The ADVANCE Project: Formal Evaluation of the Targeted Development*, Vol. 3, pp. h-19.

Sermons M. W. and Koppelman F. S., 1996, Use of Vehicle Positioning Data for Arterial Incident Detection, *Transportation Research Part C*, Vol. 4, pp. 87-96.

Sethi V., Bhandari N., Kopplelman S. and Schofer J. L., 1995, Arterial Incident Detection Using Fixed Detection and Probe Vehicle Data. *Transportation Research Part C*, Vol. 3 (2), pp. 99-112.

Shrestha B. P. 2003, Simulating Advanced Bus Priority Strategies at Traffic Signals. PhD thesis, Dept. of Civil and Environmental Eng., Southampton University. pp 14-17.

Simmons N., Gates G. and Burr J., 2002, Commercial Applications Arising from A Floating Vehicle Data System in Europe. In *Proceedings of 9th World Congress on Intelligent Transport Systems,* Chicago, 14-17 October.

Srinvivasan D., Cheu R. L. and Yeo E., 2002, Development of an Improved ERP System Using GPS and Intelligent Search Technique. In *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, 3–6 September, Singapore, pp. 944–939.

Srinivasan K. K. and Jovanis P., 1996, Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. *Transportation Research Record*, Vol. 1537, pp. 15-22.

Stephanedes Y. J., Chassiakos A. P. and Michalopoulos P. G., 1992, Comparative Performance Evaluation of Incident Detection Algorithms. *Transportation Research Record,* Vol. 1360, pp. 50-57.

Suennen M. D., 2000, A Traffic Detection Tool Kit for Travel Information Systems. Technical Report, *The Texas A&M University*, Report Number: SWUTC/00/473700-0003-1.

Sun H., Liu. H. X., Xiao H., He R. R. and Ran B., 2003, Use of Local Linear Regression Model for Short-term Traffic Forecasting, *Transportation Research Record,* Vol. 1836, pp. 143-150.

Turner S. M. and Holdener D. J., 1995, Probe Vehicle Sample Sizes for Real-time Information: The Houston Experience. In *Proceeding of the 1995 Vehicle Navigation and Information Systems Conference.* Washington, 30 July- 2 Aug.

Turner S. M., Eisele W. L. Benz R. J. and Holdener D. J., 1998, Travel Time Data Collection Handbook. Technical Report, *Texas Transportation Institute*, Report Number: FHWA-SA-97-064.

U.S. Department of Commerce (DOC), 2000, Fact Sheet Civilian Benefits of Discontinuing Selective Availability.

VTT Information Technology, 2003, Mobile Location Technology. [Online], Available: http://location.vtt.fi/

Weil R., Wootton J. R. and Garcia-Ortiz A., 1998, Traffic Incident Detection: Sensors and Algorithms, *Mathematical and Computer Modelling*, Vol. 27(9-11), pp. 257-291.

Wells C., 1996, The Kalman Filter in Finance, *Kluwer Academic Publishers*, pp. 77-84.

Wiggins A. E., 1999, Helsinki Journey Time Monitoring System, In *Proceedings of IEE Seminar CCTV and Road Surveillance*, London, pp. 7/1-7/9.

Willsky A. S., Chow E. Y., Gershwin S. B., Greene C. S., Kurkjian A. L. and Houpt P. K., 1980, Dynamic Model-Based Techniques for the Detection of Incidents on Freeway. *IEEE Transactions on Automatic control*, AC-2(3), pp. 347-360.

Wu J., McDonald M. and Chartteeji K., 2002, Impacts of Ramp Metering on Traffic on Slip Road and Motorway. In *Proceedings of Universities Transport Study Group 35th Annual Conference*, Loughborough, 6-8 January.

Xie C., Cheu R. L. and Lee D-H., 2001, Calibration-Free Arterial Link Speed Estimation Model Using Loop Data, *Journal of Transportation Engineering*, Vol. 127(6), pp. 507-514.

Ygnace. J.L., Drane C. and Yim Y. B., 2000, Cellular Phone Positioning and Travel Times Estimates. In *Proceedings of the 7th world congress on Intelligent Transport Systems*, Turin, Italy, 6-9 November.

Yim Y. B. and Cayford R., 2002, Positional Accuracy of Global Positioning System and Cellular Phone Tracking for Probe Vehicles, In *Proceedings of Transportation Research Board 81st Annual Meeting*, Jan., Washington, D.C., 13-17 January.

Zheng P., 2002, Dynamics in Merging Behaviour and its Impacts on Traffic Flow Stability. In *Proceedings of Universities Transport Study Group 34th Annual Conference*, Edinburgh, 3-5 January.