Access to Destinations: Travel Time Estimation on Arterials

Report # 3 in the series
Access to Destinations Study

Report # 2007-35

Sponsored by:

University of Minnesota
Center for Transportation Studies

Minnesota Department of Transportation

Metropolitan Council

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.	
MN/RC 2007-35			
4. Title and Subtitle Access to Destinations: Travel Time Estimation on Arterials		5. Report Date August 2007 6.	
7. Author(s) Gary A. Davis and Hui Xiong		8. Performing Organization Report No.	
9. Performing Organization Name and Address Dept. of Civil Engineering University of Minnesota 500 Pillsbury Drive SE Minneapolis, MN 55455		10. Project/Task/Work Unit No. 11. Contract (C) or Grant (G) No. (c) 81655 (wo) 188	
12. Sponsoring Organization Name and Address Minnesota Department of Transportation 395 John Ireland Boulevard Mail Stop 330 St. Paul, Minnesota 55155		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code	

15. Supplementary Notes
http://www.lrrb.org/PDF/200735.pdf
Report #3 in the series: Access to Destinations Study

16. Abstract (Limit: 200 words)

The primary objective of this project was to identify and evaluate parametric models for making default estimates of travel times on arterial links. A review of the literature revealed several candidate models, including the Bureau of Public Roads (BPR) function, Spiess's conical volume delay function, the Singapore model, the Skabardonis-Dowling model, and the Highway Capacity Manual's model. A license plate method was applied to a sample of 50 arterial links located in the Twin Cities seven county metropolitan area, to obtain measurements of average travel time. Also obtained were the lengths of each link, measurements of traffic volume, and signal timing information. Default values for model parameters were obtained from the Twin Cities planning model's database. Using network default parameters, we found that the BPR and conical volume-delay models produced mean average percent errors (MAPE) of about 25%, while the Singapore and Skabardonis-Dowling models, using maximal site-specific information, produced MAPE values of around 6.5%. As site-specific information was replaced by default information the performance of the latter two models deteriorated, but even under conditions of minimal information the models produced MAPE values of around 20%. A cross-validation study of the Skabardonis-Dowling model showed essentially similar performance when predicting travel times on links not used to estimate default parameter values.

17. Document Analysis/Descriptors	18. Availability Statement
Travel time, arterials, license-plate matching, volume/delay functions	No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	83	

Access to Destinations:
Travel Time Estimation on Arterials

Final Report

Prepared by:

Gary A. Davis
Hui Xiong

Department of Civil Engineering
University of Minnesota

August 2007

Published by:

Minnesota Department of Transportation
Office of Research Services
Mail Stop 330
395 John Ireland Boulevard
St. Paul, Minnesota 55155-1899

Acknowledgements

The authors would like to thank the Minnesota Department of Transportation for funding the project and providing support, Mark Filipi of the Metropolitan Council and Steve Wilson of SRF Consulting Group for providing technical assistance, and all traffic engineers who replied to the travel time data survey. They would also like to thank HunWen Tao, Nathan Aul, and Paul Morris for assistance in collecting data.

Table of Contents

CHAPTER 1 INTRODUCTION 1

1.1 Project Motivation 1

1.2 Arterial Travel Time Estimation Models 3

1.2.1 Bureau of Public Roads (BPR) Function 3

1.2.2 Conical Volume-Delay Function 3

1.2.3 Singapore Model 4

1.2.4 Skabardonis-Dowling Model 4

1.2.5 Highway Capacity Manual (HCM) Formula 5

CHAPTER 2 PILOT STUDY I: COMPARISON OF TRAVEL TIME ESTIMATION METHODS ... 7

2.1 Problem Description 7

2.2 Pilot Study I 7

2.2.1 Site Information 7

2.2.2 Data Collection 8

2.2.3 Model Comparison 10

2.3 Travel Time Data Survey 12

CHAPTER 3 PILOT STUDY II AND MODIFIED PLATE-MATCHING METHOD 14

3.1 Pilot Study II 14

3.1.1 Site Information and Data Collection 14

3.1.2 Bias of Traditional Plate Matching Method 15

3.2 Modified Plate Matching Method 17

3.2.1 Mixture Distribution 17

3.2.2 Model Fitting Algorithm 19

3.2.3 Binomial Test of Classification 20

3.3 Estimation Result 21

CHAPTER 4 TRAVEL TIME DATA COLLECTION 24

4.1 Sampling Arterial Sites 24

4.1.1 Available Network Data Sources 24

4.1.2 Arterial Sites Sampling Procedure 25

4.2 Field Data Collection 26

4.3 Data Processing 28

4.4 Data Summary 29

CHAPTER 5 MODEL COMPARISON AND EVALUATION 31

5.1 Mean Absolute Percentage Error (MAPE) and Nonlinear Regression Model 31

5.2 BPR Function and Conical Volume-Delay Function 31

5.3 Singapore Model and Skabardonis-Dowling model 33

5.4 Cross-Validation of Skabardonis-Dowling Model 35

CHAPTER 6 SUMMARY AND CONCLUSION 37

REFERENCES 39

APPENDIX A. TRAVEL TIME RELATIVE FREQUENCY PLOTS

APPENDIX B. R CODE FOR TRAVEL TIME DATA PROCESSING

List of Tables

Table 1-1 1990 Regional Model Daily Peak Period Speed (mph) 2
Table 1-2 1990 Regional Model Lane Capacity (vehicles per-lane-per-hour) 2
Table 2-1 Travel Time Data Collected by Floating Car 8
Table 2-2 Free-flow Speed Data Collected by Laser Gun 9
Table 2-3 Traffic Data Collected by Digital Camcorder 9
Table 2-4 Stopped Delay Study Field Sheet 10
Table 2-5 Travel Time Estimation Model Comparison 11
Table 2-6 Travel Time Data Survey Results 13
Table 3-1 R Output of Estimation Results 23
Table 4-1 Column Definitions of Twin Cities 1995 and 2020 Transportation Network 25
Table 4-2 Data Requirement for Candidate Models 27
Table 4-3 Data Collection Method 28
Table 4-4 Data Summary for Model Comparison 30
Table 5-1 Travel Time Estimation Models without Calibration 32
Table 5-2 Data Availability Scenarios for Singapore Model 34
Table 5-3 Data Availability Scenarios for Skabardonis-Dowling Model 34
Table 5-4 Cases of Singapore Model 35
Table 5-5 Cases of Skabardonis-Dowling Model 35
Table 5-6 LOOCV of Skabardonis-Dowling Model 36

List of Figures

Figure 2-1 Description of Travel Time Measurement 7
Figure 2-2 Illustration of Pilot Study I 8
Figure 2-3 Travel Time Estimation Model Comparison 11
Figure 3-1 Illustration of Pilot Study II 15
Figure 3-2 Illustration of Plate Matching Method 15
Figure 3-3 Travel Time Relative Frequency Plot of Video Sample 16
Figure 3-4 Travel Time Relative Frequency Plot of License Plate Sample 17
Figure 3-5 Travel Time Relative Frequency Plot of Stopped Vehicles in Video Sample 18
Figure 3-6 Travel Time Relative Frequency Plot of Non-stopping Vehicles in Video Sample 18
Figure 3-7 Classification Error Plot of Non-stopping Vehicles 21
Figure 3-8 Classification P-value Plot of Non-stopping Vehicles 21
Figure 3-9 Fitted Curve of Video Sample 22
Figure 3-10 Fitted Curve of License Plate Sample 22
Figure 4-1 Link Attribute Table of Twin Cities 1990 Transportation Network 24
Figure 4-2 Link Attribute Table of Twin Cities 2000 Transportation Network 25
Figure 4-3 Map of Sampled Sites 26
Figure 4-4 Illustration of Field Data Collection 27
Figure 5-1 Measured Travel Time vs Estimated Travel Time (BPR Function) 32
Figure 5-2 Measured Travel Time vs Estimated Travel Time (Conical Volume-Delay Function) 33
Figure A-1 Travel Time Relative Frequency Plot of Site 2 ..... A-1
Figure A-2 Travel Time Relative Frequency Plot of Site 3 ..... A-1
Figure A-3 Travel Time Relative Frequency Plot of Site 6 ..... A-2
Figure A-4 Travel Time Relative Frequency Plot of Site 8 ..... A-2
Figure A-5 Travel Time Relative Frequency Plot of Site 14 ..... A-3
Figure A-6 Travel Time Relative Frequency Plot of Site 24 ..... A-3
Figure A-7 Travel Time Relative Frequency Plot of Site 25 ..... A-4
Figure A-8 Travel Time Relative Frequency Plot of Site 27 ..... A-4
Figure A-9 Travel Time Relative Frequency Plot of Site 29 ..... A-5
Figure A-10 Travel Time Relative Frequency Plot of Site 30 ..... A-5
Figure A-11 Travel Time Relative Frequency Plot of Site 31 ..... A-6
Figure A-12 Travel Time Relative Frequency Plot of Site 32 ..... A-6
Figure A-13 Travel Time Relative Frequency Plot of Site 33 ..... A-7

Figure A-14 Travel Time Relative Frequency Plot of Site 34 ..... A-7
Figure A-15 Travel Time Relative Frequency Plot of Site 35 ..... A-8
Figure A-16 Travel Time Relative Frequency Plot of Site 36 ..... A-8
Figure A-17 Travel Time Relative Frequency Plot of Site 38 ..... A-9
Figure A-18 Travel Time Relative Frequency Plot of Site 39 ..... A-9
Figure A-19 Travel Time Relative Frequency Plot of Site 40 ..... A-10
Figure A-20 Travel Time Relative Frequency Plot of Site 41 ..... A-10
Figure A-21 Travel Time Relative Frequency Plot of Site 50 ..... A-11
Figure A-22 Travel Time Relative Frequency Plot of Site 51 ..... A-11
Figure A-23 Travel Time Relative Frequency Plot of Site 52 ..... A-12
Figure A-24 Travel Time Relative Frequency Plot of Site 53 ..... A-12
Figure A-25 Travel Time Relative Frequency Plot of Site 54 ..... A-13
Figure A-26 Travel Time Relative Frequency Plot of Site 55 ..... A-13
Figure A-27 Travel Time Relative Frequency Plot of Site 61 ..... A-14
Figure A-28 Travel Time Relative Frequency Plot of Site 63 ..... A-14
Figure A-29 Travel Time Relative Frequency Plot of Site 64 ..... A-15
Figure A-30 Travel Time Relative Frequency Plot of Site 65 ..... A-15
Figure A-31 Travel Time Relative Frequency Plot of Site 66 ..... A-16
Figure A-32 Travel Time Relative Frequency Plot of Site 67 ..... A-16
Figure A-33 Travel Time Relative Frequency Plot of Site 68 ..... A-17
Figure A-34 Travel Time Relative Frequency Plot of Site 70 ..... A-17
Figure A-35 Travel Time Relative Frequency Plot of Site 80 ..... A-18
Figure A-36 Travel Time Relative Frequency Plot of Site 81 ..... A-18
Figure A-37 Travel Time Relative Frequency Plot of Site 83 ..... A-19
Figure A-38 Travel Time Relative Frequency Plot of Site 85 ..... A-19
Figure A-39 Travel Time Relative Frequency Plot of Site 86 ..... A-20
Figure A-40 Travel Time Relative Frequency Plot of Site 87 ..... A-20
Figure A-41 Travel Time Relative Frequency Plot of Site 96 ..... A-21
Figure A-42 Travel Time Relative Frequency Plot of Site 97 ..... A-21
Figure A-43 Travel Time Relative Frequency Plot of Site 98 ..... A-22
Figure A-44 Travel Time Relative Frequency Plot of Site 100 ..... A-22
Figure A-45 Travel Time Relative Frequency Plot of Site 104 ..... A-23
Figure A-46 Travel Time Relative Frequency Plot of Site 109 ..... A-23
Figure A-47 Travel Time Relative Frequency Plot of Site 111 ..... A-24

Figure A-48 Travel Time Relative Frequency Plot of Site 130 ..... A-24
Figure A-49 Travel Time Relative Frequency Plot of Site 131 ..... A-25
Figure A-50 Travel Time Relative Frequency Plot of Site 134 ..... A-25
Figure A-51 Travel Time Relative Frequency Plot of Site 140 ..... A-26
Figure A-52 Travel Time Relative Frequency Plot of Site 142 ..... A-26
Figure A-53 Travel Time Relative Frequency Plot of Site 146 ..... A-27
Figure A-54 Travel Time Relative Frequency Plot of Site 148 ..... A-27
Figure A-55 Travel Time Relative Frequency Plot of Site 149 ..... A-28

Executive Summary

The primary objective of this project was to identify and evaluate parametric models for making default estimates of travel times on arterial links, using information typically available from a transportation planning model. The chosen method of evaluation was to compare travel time predictions generated by the models to field measurements. A review of the literature revealed several candidate models, including the Bureau of Public Roads (BPR) function, Spiess's conical volume delay function, the Singapore model, the Skabardonis-Dowling model, and the Highway Capacity Manual's model. A survey of Twin Cities traffic agencies indicated that it would be difficult to obtain, from existing data sources, a representative sample of arterial links with all data needed to evaluate the candidate models, so it was decided to conduct a field study to collect the needed data. In a first pilot study, average travel times on an arterial link were measured using (1) a combination of spot speed data to estimate mean free-flow travel times and intersection delay measurements to estimate average waiting time, and (2) a floating car method. The floating car method turned out to be sensitive to the relative fraction of runs where the floating car was delayed by a red signal indication, and it was not possible to obtain enough runs to reliably estimate this fraction. A comparison of travel time models using the data collected with method (1) indicated best performance by the HCM model, the worst performance by the BPR and conical volume-delay models, with the remaining two models being slightly worse than the HCM model.

To maximize the number of data collection sites within the project's resource constraints it was then decided to evaluate a license plate matching method for collecting travel time data, which could be carried out by only two field personnel. A second pilot study comparing the travel times measured using the license plate method to travel times measured from video revealed a tendency for the license plate method to under-sample vehicles stopped by red signal indications, and hence to underestimate average travel times. However, by using a mixture decomposition method to estimate mean travel times for stopped and non-stopped vehicles, together with an independent estimate of proportion stopping, an estimation method that substantially eliminated this bias was developed.

The license plate method was then applied to a sample of 50 arterial links located in the Twin Cities seven county metropolitan area, to obtain measurements of average travel time. Also obtained were the lengths of each link, measurements of traffic volume, and signal timing information. Default values for model parameters were obtained from the Twin Cities planning model's database. Using network default parameters, we found that the BPR and conical volume-delay models produced mean average percent errors (MAPE) of about 25%, while the Singapore and Skabardonis-Dowling models, using maximal site-specific information, produced MAPE values of around 6.5%. As site-specific information was replaced by default information, the performance of the latter two models deteriorated, but even under conditions of minimal information the models produced MAPE values of around 20%. A cross-validation study of the Skabardonis-Dowling model showed essentially similar performance when predicting travel times on links not used to estimate default parameter values.

Chapter 1 Introduction

1.1 Project Motivation

In transportation planning, the link travel time is a key variable, which determines the route choice and the distribution of trips. Accurate travel time estimates play an important role in successful traffic prediction and evaluation of alternatives.

The link capacity and free-flow speed are two basic inputs for travel time prediction, and the Highway Capacity Manual (HCM) 2000 provides detailed procedures for estimating these. However, for transportation planning purposes, it is not generally practical to collect the site-specific information individually since there are many links in the network (for instance, there are 8754 arterial links in the Twin Cities 1990 transportation network). An alternative method is to find default values representing the average conditions describing a facility type and area type. That means that, in each category, all arterial links are assumed to have the same free-flow speed and lane capacity. The default values of the Twin Cities 1990 transportation network are shown in Tables 1-1 and 1-2.

For travel time estimation on freeways, free-flow speed, volume, and capacity are usually sufficient to meet the accuracy requirement. Compared to freeways, travel time on arterials has more variability due to the presence of signal control. For arterials, travel time estimation may thus require more characterizing data, such as effective green time, cycle length, and proportion of arrivals on green, in order to obtain reliable results.

This leads to two inter-related questions: (1) how much detail is needed to produce reasonably accurate forecasts of arterial travel time, and (2) for a given level of detail, which model functional forms perform best.

The primary objective of this project is to develop, test, and recommend methods for network wide estimation and prediction of travel time on arterials. This project will include the following tasks: (1) reviewing the availability of arterial travel time related data in the Twin Cities region, (2) determining candidate arterial travel time models and the data needed to calibrate and test them, (3) evaluating candidate arterial travel time models, and (4) recommending a model to be used in later phases of the project.

Table 1-1 1990 Regional Model Daily Peak Period Speed (mph) for Minneapolis/St. Paul: Area Type by Facility Type*

Area Type / Facility Type	Rural	Developing	Developed	Center City	Central CBD	Outlying Business Area
Metered Freeway	60	60	55	50	N/A	50
Unmetered Freeway	60	60	53(51)	49(46)	N/A	49(45)
Metered Ramp	35	30	25	25	N/A	25
Unmetered Ramp	35	30	25	25	N/A	25
Divided Arterial	**50(45)**	**47(39)**	**41(36)**	**37(23)**	**N/A**	**37(25)**
Undivided Arterial	**48(40)**	**41(39)**	**38(32)**	**23(20)**	**N/A**	**23(20)**
Collector	40	33(31)	30(28)	23(21)	15(12)	23(20)
HOV	32	26(24)	24(22)	18(16)	12(9)	18(16)
Centroid Connector	N/A	65	60	55	N/A	N/A
HOV ramp	N/A	35	30	25	N/A	N/A
Note: Minneapolis and St. Paul CBD links are coded only as collectors or centroid connectors.						

* Source: Model Calibration Technical Memo #5, 1990 Highway Network and TAZ Documentation, Metropolitan Council

Table 1-2 1990 Regional Model Lane Capacity (vehicles per-lane-per-hour) for Minneapolis/St. Paul Area: Type by Facility Type

Area Type / Facility Type	Rural	Developing	Developed	Center City	Central CBD	Outlying Business Area
Metered Freeway	1950	1950	1950	1950	N/A	1950
Unmetered Freeway	1750	1750	1750	1750	N/A	1750
Metered Ramp	750	725	675	625	N/A	600
Unmetered Ramp	1500	1450	1350	1250	N/A	1200
Divided Arterial	**1000**	**950**	**850**	**750**	**N/A**	**700**
Undivided Arterial	**900**	**850**	**750**	**650**	**N/A**	**600**
Collector	600	550	500	450	400	400
HOV	N/A	N/A	N/A	N/A	N/A	N/A
Centroid Connector	N/A	1400	1400	1400	N/A	N/A
HOV ramp	N/A	1450	1350	1250	N/A	N/A
Note: Minneapolis and St. Paul CBD links are coded only as collectors or centroid connectors.						

* Source: Model Calibration Technical Memo #5, 1990 Highway Network and TAZ Documentation, Metropolitan Council

1.2 Arterial Travel Time Estimation Models

1.2.1 Bureau of Public Roads (BPR) Function

In the BPR function, the travel time is assumed to have a nonlinear relationship with volume/capacity ratio ([1]). Its standard form and variations have been widely used in transportation planning applications. The standard BPR function is

$$TT = FFT(1+\alpha(\frac{v}{c})^{\beta}) \qquad (1\text{-}1)$$

where

TT = predicted mean travel time,
FFT = free-flow travel time,
v = volume,
c = capacity (possibly adjusted by green time/cycle length ratio),
α, β = parameters (α = 0.15, β = 4).

The BPR function was developed in the late 1950's by fitting data collected on uncongested freeways. It does not explicitly consider signals. Thus it may be of limited use on arterials even if the capacity is adjusted by green time/cycle length ratio. When the volume is very low, the predicted travel time is approximately equal to free-flow travel time.

1.2.2 Conical Volume-Delay Function

Spiess ([2]) identified necessary conditions for a well-behaved volume-delay function and proposed the conical volume-delay function:

$$TT = FFT\,(2+\sqrt{\alpha^2(1-\frac{v}{c})^2+\beta^2}-\alpha(1-\frac{v}{c})-\beta) \qquad (1\text{-}2)$$

where

TT = predicted mean travel time,
FFT = free-flow travel time,
v = volume,
c = capacity (possibly adjusted by green time/cycle length ratio),
α = positive number greater than 1, and
$\beta = \dfrac{2\alpha-1}{2\alpha-2}$.

The Metropolitan Council has proposed using the conical volume-delay function to estimate arterial travel time in planning models. α is set equal to 4 for divided arterials and 5 for undivided arterials. It can be shown that, as with the BPR function, the conical volume-delay function does not explicitly consider the effects of signalization.

1.2.3 Singapore Model

In the Singapore model proposed by Xie, Cheu, and Lee ([3]), the travel time is divided into two components:

$$TT = \text{cruise time} + \text{signal delay} \tag{1-3}$$

$$\text{cruise time} = \frac{L}{FFS} \tag{1-4}$$

$$\text{signal delay} = \frac{9}{10}\left[\frac{C(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)}\right] \tag{1-5}$$

where

TT = predicted mean travel time,
FFS = free-flow travel speed,
C = cycle length (s),
λ = effective green proportion (g/C),
x = volume/capacity ratio ($0 \le x < 1$), and
q = arrival rate (veh/s).

The signal delay in the Singapore model is the modified two-term Webster formula. A limitation of this model is its feasible domain, since if x is greater than 1, the last term for the signal delay component becomes negative. In planning models however, volume/capacity rates greater than 1.0 should be allowed for.

1.2.4 Skabardonis-Dowling Model

Skabardonis and Dowling ([4]) proposed improved speed-flow relationships for planning applications. Based upon that, a travel time prediction function is given by:

$$TT = \left(\frac{L}{FFS} + 0.5NC\left(1-\frac{g}{C}\right)^2 PF\right)\left(1 + 0.05\left(\frac{v}{c}\right)^{10}\right) \tag{1-6}$$

where

TT = predicted mean travel time,
FFS = free-flow travel speed,
N = number of signals in the link,
C = cycle length,
g = effective green time,
PF = progression adjustment factor,
v = volume, and
c = capacity (adjusted by green time/cycle length ratio).

The progression adjustment factor is in turn given by

$$PF = \frac{(1-P)\,f_{PA}}{1-\frac{g}{C}} \tag{1-7}$$

where

PF = progression adjustment factor,
P = proportion of vehicles arriving on green,
g/C = proportion of green time available, and
f_{PA} = supplemental adjustment factor for platoon arriving during green (approximately =1).

When the link has only one signalized intersection at the downstream site, Equation 1-6 can be simplified to

$$TT \approx (\frac{L}{FFS} + 0.5(1-P)(C-g))(1+0.05(\frac{v}{c})^{10}) \tag{1-8}$$

1.2.5 Highway Capacity Manual (HCM) Formula

In the HCM 2000 ([5]), the travel time on a segment is estimated by:

$$TT = T_R + d_1 \cdot PF + d_2 + d_3 \tag{1-9}$$

where

TT = predicted mean travel time,
T_R = running time (determined by street's classification and its free-flow speed),
d_1 = uniform delay,
PF = progression adjustment factor,
d_2 = incremental delay, and
d_3 = initial queue delay.

The uniform delay is calculated by

$$d_1 = \frac{0.5C(1-\frac{g}{C})^2}{1-[\min(1,x)\frac{g}{C}]} \tag{1-10}$$

where

x = volume to capacity (v/c) ratio for the lane group,
C = cycle length, and
g = effective green time for lane group.

The incremental delay is calculated by

$$d_2 = 900T[(x-1) + \sqrt{(x-1)^2 + \frac{8kIx}{cT}}] \quad (1\text{-}11)$$

where

T = duration of analysis period (typically, T = 15min = 0.25h),

k = incremental delay adjustment for the actuated control (for pretimed intersections, k = 0.5),

I = incremental delay adjustment for the filtering or metering by upstream signals (when $0 \le x \le 1$, $I \approx 1 - 0.91x^{2.68}$; when $x > 1$, $I \approx 0.09$), and

c = capacity (adjusted by green time/cycle length ratio).

Among the above five models, the HCM formula requires the most data, some of which may require field measurements.

Chapter 2 Pilot Study I: Comparison of Travel Time Estimation Methods

2.1 Problem Description

Figure 2-1 depicts a typical arterial link. It has two intersections at its ends and the downstream one is signalized. There is no other signalized intersections or stop signs (controlling through movements of the link) in between. The problem of interest is to measure the average time of vehicles traversing the link during a time period. Let t_i be the entering time of vehicle i, t_i+TT_i be the leaving time. The travel time of vehicle i is TT_i (as shown in Figure 2-1). Suppose during a time period, there are n vehicles passing through the link, then the average travel time is $\frac{1}{n}\sum_{i=1}^{n} TT_i$.

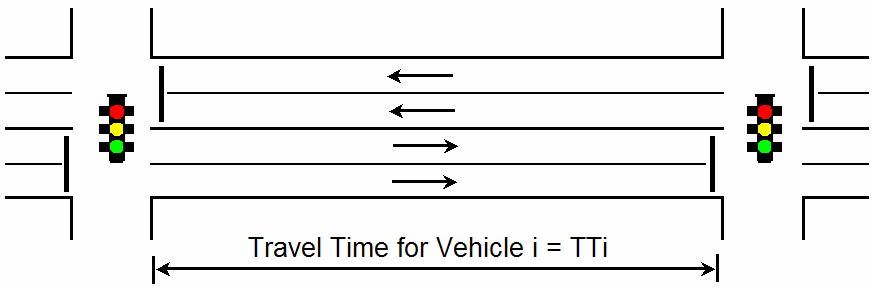

Figure 2-1 Description of Travel Time Measurement

2.2 Pilot Study I

2.2.1 Site Information

On October 13th 2005, pilot study I was performed to assess travel time measurement procedures and to do a rough comparison of candidate models. As shown in Figure 2-2, the study segment has two lanes and is located on Washington Ave (Minneapolis, MN) from the boundary of Washington Ave Bridge to the intersection at Church St. SE (east bound). The analysis periods were 4:00~4:15pm, 4:16~4:30pm, 4:31~4:45pm, and 4:46~5:00pm.

The signal controller type of the study intersection is pretimed and the phase plan is two-phase with C=100s and g/C = 0.46. The grade of the study segment is approximately 2%, and the segment length is 840ft (0.1591 mile).

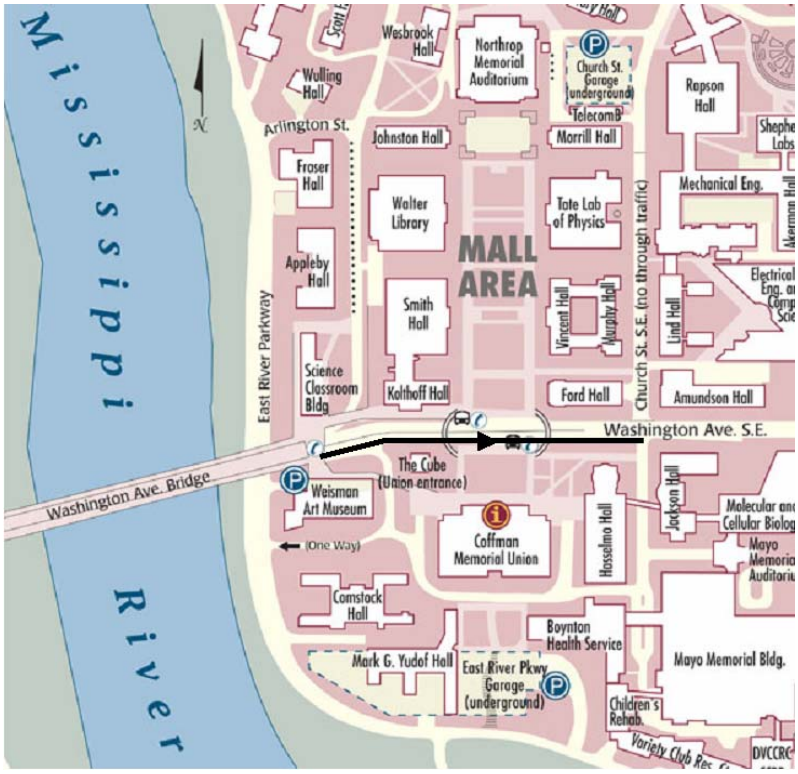

Figure 2-2 Illustration of Pilot Study I

2.2.2 Data Collection

A floating car was running on the study segment to collect the travel time and the control delay (see summary in Table 2-1). The floating car technique was to "travel according to the driver's judgment of the average speed of the traffic stream" ([6]).

Table 2-1 Travel Time Data Collected by Floating Car

Trip	Observation Time	Travel Time(s)	Control Delay (s)	Lane
1	4:01pm	87	64	R
2	4:10pm	28	0	L
3	4:16pm	71	41	R
4	4:25pm	80	54	L
5	4:32pm	25	0	R
6	4:40pm	73	47	L
7	4:46pm	42	16	R
8	4:56pm	22	0	L
Average		53.5	27.25	

The free-flow speed data were collected by a laser gun and are summarized in Table 2-2. The speeds of buses and campus shuttles were not collected because they stopped at a bus station near the downstream intersection, the distance from which to the signal was not sufficient to accelerate from 0 to the free-flow speed.

Table 2-2 Free-flow Speed Data Collected by Laser Gun

Time Period	Speed (mph)	Number of Samples	Average Speed (mph)
4:00-4:15pm	22,23,24,26,27,28,29,29,30,30,31,31	12	27.50
4:16-4:30pm	26,28,29,31,31,34	6	29.83
4:31-4:45pm	22,22,26,29,30,38	6	28.14
4:46-5:00pm	26,27,27,28,28,29,32	7	28.14
4:00-5:00pm		31	28.16

Finally, the data in Table 2-3 were extracted from video shot from the roof of Amundson Hall (located at the northeast corner of the intersection). Since the eastbound traffic is not allowed to turn left at the intersection, the sum of through movements and right turns is equal to the volume. The data listed in Table 2-3 are used to find v/c, PF, and adjusted saturation flow.

Table 2-3 Traffic Data Collected by Digital Camcorder

Time Period	Through Movements	Right Turns	Heavy Vehicles	Buses	Pedestrians	Bicycles	Arrivals on Green
4:00-4:15pm	116	10	13	10	102	9	39
4:16-4:30pm	98	8	9	8	145	16	38
4:31-4:45pm	128	11	11	9	128	13	46
4:46-5:00pm	134	9	9	8	103	16	41
Total	476	38	42	35	478	54	164

The video was also used to estimate average stopped delay. 15 seconds was chosen as the analysis interval for the stopped delay study. Measured stopped delay multiplied by 1.3 gives the average total delay per vehicle ([6]). Numbers of stopped vehicles in each 15-sec interval were counted by observing the video.

Table 2-4 Stopped Delay Study Field Sheet

Sec / Min	0-15	16-30	31-45	46-60	Sec / Min	0-15	16-30	31-45	46-60
0-1	3	4	0	0	**30-31**	6	6	0	0
1-2	1	3	6	12	**31-32**	0	2	4	5
2-3	16	8	0	0	**32-33**	6	0	0	0
3-4	1	1	3	3	**33-34**	1	2	2	0
4-5	0	0	1	7	**34-35**	0	0	0	3
5-6	9	11	0	0	**35-36**	3	6	0	2
6-7	0	0	0	4	**36-37**	0	0	2	8
7-8	8	0	0	1	**37-38**	13	0	0	0
8-9	2	4	6	5	**38-39**	2	4	8	0
9-10	0	0	2	2	**39-40**	0	0	0	5
10-11	2	5	0	0	**40-41**	6	10	0	0
11-12	0	2	6	9	**41-42**	0	1	3	5
12-13	10	0	0	0	**42-43**	6	1	0	0
13-14	1	1	4	0	**43-44**	3	6	8	0
14-15	0	0	0	3	**44-45**	0	0	0	4
15-16	6	6	0	0	**45-46**	8	10	0	0
16-17	0	1	1	2	**46-47**	0	1	7	8
17-18	4	0	0	0	**47-48**	9	0	0	0
18-19	2	4	5	0	**48-49**	2	4	7	0
19-20	0	0	0	0	**49-50**	0	0	0	0
20-21	0	0	0	0	**50-51**	1	4	0	0
21-22	0	1	3	5	**51-52**	0	1	3	12
22-23	6	0	0	0	**52-53**	15	0	0	0
23-24	3	7	8	6	**53-54**	0	3	5	0
24-25	0	0	1	6	**54-55**	0	0	0	0
25-26	11	12	0	0	**55-56**	5	8	0	0
26-27	0	0	5	8	**56-57**	0	7	11	13
27-28	9	0	0	0	**57-58**	16	0	0	0
28-29	2	7	7	7	**58-59**	0	1	4	0
29-30	0	0	0	2	**59-60**	0	0	0	4

Stopped Delay: 4:00-4:15pm 19.64s/veh; 4:16-4:30pm 21.22s/veh;

4:31-4:45pm 16.63s/veh; 4:16-4:30pm 17.73s/veh.

2.2.3 Model Comparison

Figure 2-3 and Table 2-5 provide the arterial travel time estimation model comparison. From the pilot study I, it can be seen that:

- The standard BPR function and the conical volume-delay function are not sensitive to the change of v/c especially when v/c < 0.5. Both of them underestimated the mean travel time.

- The floating car method is sensitive to whether or not the test vehicle was stopped at the signal. Accurate estimation of mean travel time then would require an estimate of the fraction of stopped runs, which in turn would require more runs that it was possible to make with a single vehicle.

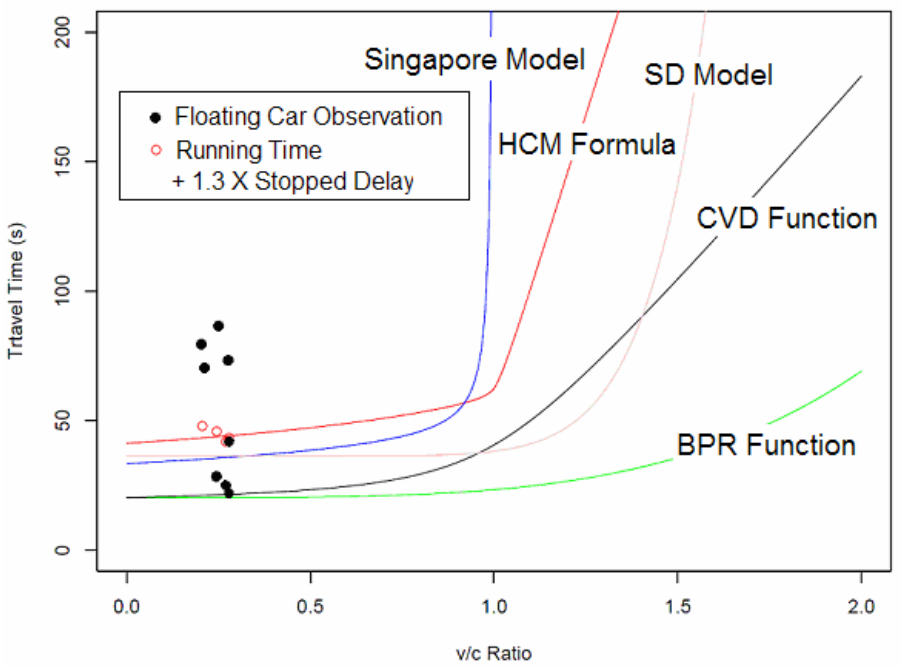

Figure 2-3 Travel Time Estimation Model Comparison

Table 2-5 Travel Time Estimation Model Comparison

Time Period	Measured Total Delay +Running Time	BPR Function	Conical Volume-Delay Function	Singapore Model	Highway Capacity Manual Formula	Skabardonis -Dowling Model
4:00-4:15pm	45.88s/veh	20.35s/veh	20.68s/veh	35.42s/veh	46.40s/veh	36.39s/veh
4:16-4:30pm	47.94 s/veh	20.35s/veh	20.62s/veh	35.03s/veh	44.27s/veh	36.39s/veh
4:31-4:45pm	41.97 s/veh	20.36s/veh	20.72s/veh	35.69s/veh	46.22s/veh	36.39s/veh
4:46-5:00pm	43.40 s/veh	20.36s/veh	20.73s/veh	35.77s/veh	46.78s/veh	36.39s/veh
Note		α=0.15, β=4	α=4			

2.3 Travel Time Data Survey

The purpose of the travel time data survey was to find out if the agencies in the Twin Cities region have travel time data helpful and sharable for this project.

The data of interest included:

- Intersection stopped delay
- Arterial running travel time (floating car studies)
- Spot speed studies
- Signal timing plan
- Traffic volume or average daily traffic
- Intersection turning movement counts
- Data from automatic loop detectors
- Data from video surveillance
- GIS database for arterial road network or traffic control
- Other data helpful to the project

The survey questions included:

- Data availability
- Data collection time
- Data resolution (30 seconds; 5 minutes interval; 15 minutes; hourly)
- Storage medium
- Sharable?

In March 2006, Survey letters were mailed to the traffic engineers of 89 agencies in the Twin Cities region, including:

- The Minnesota Department of Transportation (Mn/DOT)
- The Metropolitan Council
- 8 Counties: Anoka, Carver, Chisago, Dakota, Hennepin, Ramsey, Scott, and Washington
- 79 Cities: Andover, Anoka, Apple Valley, Arden Hills, Belle Plaine, Bloomington, Brooklyn Center, Brooklyn Park, Burnsville, Champlin, Chanhassen, Columbia Heights, Coon Rapids, Corcoran, Cottage Grove, Crystal, Dayton, Eagan, East Bethel, Edina, Falcon Heights, Farmington, Forest Lake, Fridley, Golden Valley, Ham Lake, Hastings, Hopkins, Hugo, Lakeville, Lake Elmo, Lino Lakes, Little Canada, Mahtomedi, Maple Grove, Maplewood, Mendota Heights, Minneapolis, Minnetonka, Mound, Mounds View, New Brighton, New Hope, New Prague, North Branch, North St. Paul, Oak Grove, Oakdale, Orono, Plymouth, Prior Lake, Ramsey, Richfield, Robbinsdale, Rogers, St. Anthony, St. Francis, Saint Paul, St. Paul Park, South St. Paul, Savage, Shakopee, Shorewood, Spring Lake Park, Vadnais Heights, Victoria, Chaska, West St. Paul, White Bear Lake, Woodbury, Blaine, Eden Prairie, Inver Grove Heights, Rosemount, Roseville, Stillwater, Waconia, St. Louis Park, and Shoreview

By May 2006, 59 agencies (66.3%) replied to the survey. The results are summarized in Table 2-6.

Table 2-6 Travel Time Data Survey Results

	Survey Questions	Yes	No
1	Intersection stopped delay	1	58
2	Arterial running travel time (floating car studies)	4	55
3	Spot speed studies	25	34
4	Signal timing plan	9	50
5	Traffic volume or average daily traffic	44	15
6	Intersection turning movement counts	18	41
7	Data from automatic loop detectors	2	57
8	Data from video surveillance	1	58
9	GIS database for arterial road network or traffic control	13	46
10	Other data helpful for the project	3	56

To carry out the main portion of this project we would need, at a minimum, traffic volume, spot speed and intersection delay data for a representative sample of arterial links. Although the agencies do collect these data, the number of links for which all data items might reasonably be expected to be available, for comparable time periods, is limited. Thus, it was decided to directly collect the field data needed for this project.

Chapter 3 Pilot Study II and Modified Plate-Matching Method

3.1 Pilot Study II

3.1.1 Site Information and Data Collection

For locations lacking good video camera placements, intersection delay data must be collected by two field observers at ground level. Together with an observer collecting spot speed data gives three individuals needed to collect the necessary data. Project budget constraints limited the number of data collectors to two, so we sought a data collection method satisfying this constraint. On April 4th 2006, pilot study II was performed to test a license plate matching method. As shown in Figure 3-1, the study segment had two lanes and was located on Washington Ave (Minneapolis, MN), from the intersection at Union St. SE to the intersection at Church St. SE (west bound). Data were collected from 3:50pm to 4:30pm.

Observers at each end of the link recorded the last 3 digits of license plate numbers into laptop computers (see Figure 3-2). Time stamps were added automatically when inputting the first digit. Plate numbers of 157 vehicles were matched by a MS Access program and their travel times (called as “license plate sample”) then were calculated. A video camera was installed on the roof of Weaver Densford Hall (southwest corner of the intersection at Washington Ave. SE and Harvard St. SE) to record traffic during license plate study. Travel times of 525 vehicles (called the “video sample”) were extracted from video, which are viewed as the ground truth.

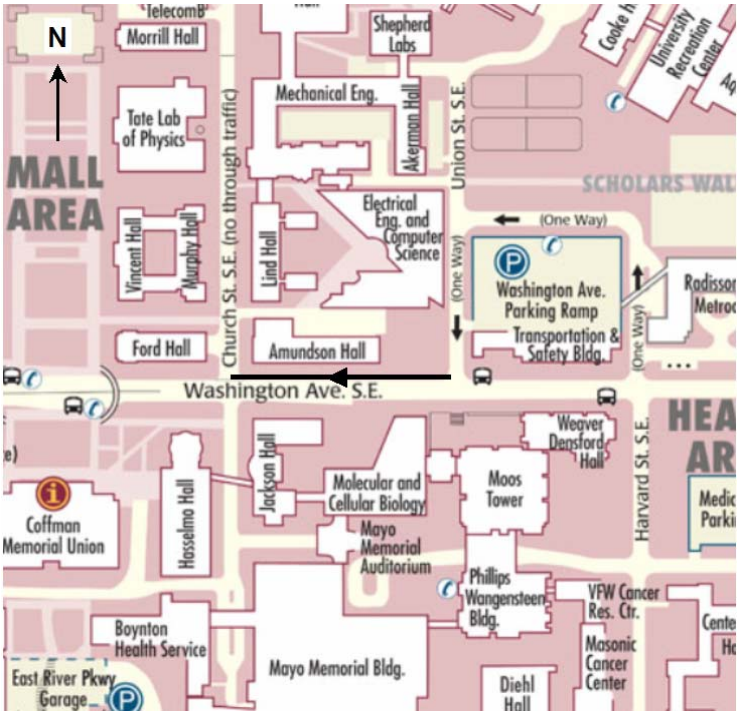

Figure 3-1 Illustration of Pilot Study II

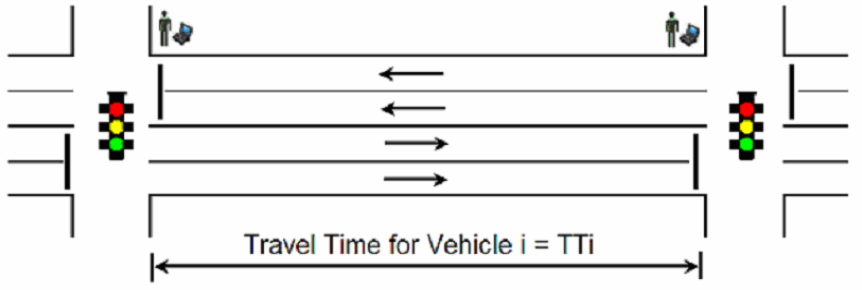

Figure 3-2 Illustration of Plate Matching Method

3.1.2 Bias of Traditional Plate Matching Method

If the data collection is unbiased (i.e. the matched vehicles have the same travel time distribution as all vehicles traversing the link during the study period), the average travel time of matched vehicles will be an unbiased estimator of the mean link travel time. Figures 3-3 and 3-4 show the travel time relative frequency plots of the video sample and the license plate sample. It can be seen that the proportion of non-stopping vehicles is higher for the license plate sample. This implies that the simple average from the license plate sample will underestimate the mean travel time. In this study, the average travel time of the video sample was 28.57s/veh, while that for the license plate sample was 25.21s/veh.

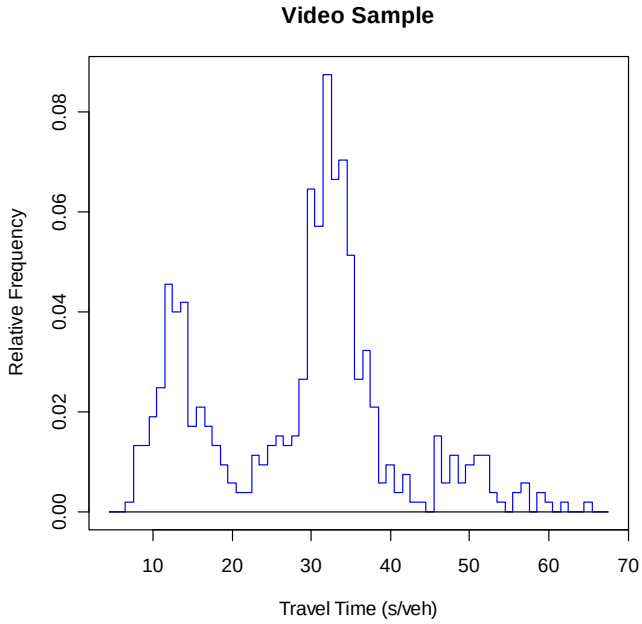

Figure 3-3 Travel Time Relative Frequency Plot of Video Sample

The reason for this difference was that, in the license plate sample, stopped and non-stopping vehicles have different sample rates. In the plate-matching method, to ensure sufficient samples, observers have to record as many vehicles as possible. This is one of the basic strategies ([9]). Observers do not have any preference on the selection of stopped or non-stopping vehicles. However, the recording speed of the observers is limited. If the time difference between two departures at the downstream intersection is shorter than the recording cycle time of observers, one vehicle has to be skipped.

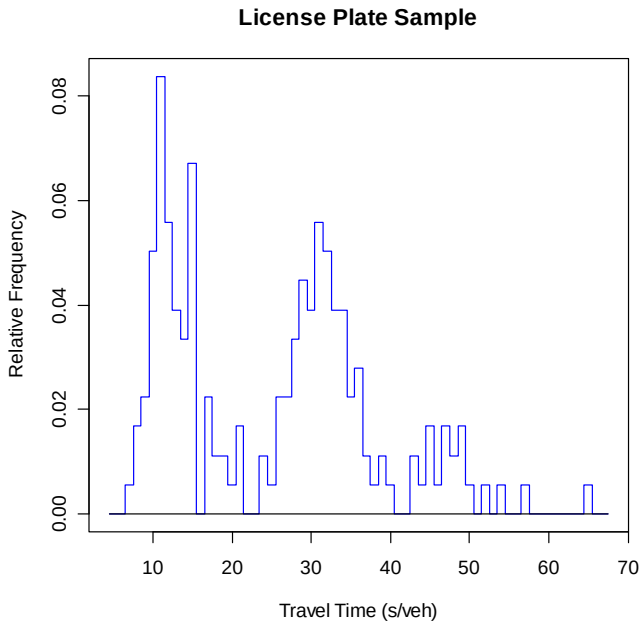

Figure 3-4 Travel Time Relative Frequency Plot of License Plate Sample

In general, the sample rate difference is determined by two factors:

- The number of lanes at the downstream intersections. The more lanes, the more likely that departure time difference is less than the recording cycle time.
- The proportion of stopped vehicles. If it is higher, relatively more non-stopping vehicles tend to be recorded, and vice versa.

3.2 Modified Plate Matching Method

3.2.1 Mixture Distribution

Note that in Figures 3-3 and 3-4, both plots have two peaks with apparently similar mean values. Tracking each vehicle in the video gives Figures 3-5 and 3-6, which show the travel time relative frequency plots of stopped vehicles and non-stopping vehicles in the video sample. Referring to Figure 3-3 leads us to consider classifying vehicles into two clusters, non-stopping vehicles and stopped vehicles. If their mean travel times can be estimated using matched travel time data, counting the non-stopping and total vehicles from a ground based video should provide an estimate of actual proportions of these clusters, which can be used to reduce the bias.

Figure 3-5 Travel Time Relative Frequency Plot of Stopped Vehicles in Video Sample

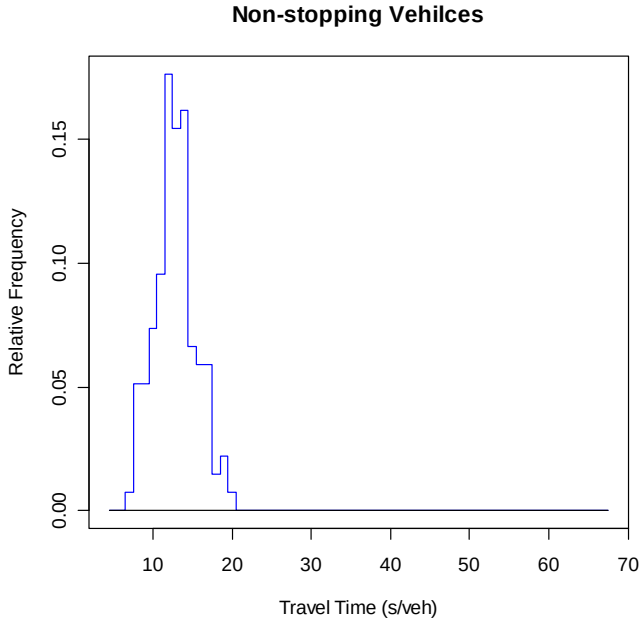

Figure 3-6 Travel Time Relative Frequency Plot of Non-stopping Vehicles in Video Sample

The shapes of two histograms in Figures 3-3 and 3-4 suggest that the link travel time distribution can be approximated by a mixture distribution with two normal components, representing non-stopping vehicles (group 1) and stopped vehicles (group 2). Then the probability density function (pdf) of travel time TT is

$$f(TT) = \pi \times f_N(TT) + (1-\pi) \times f_S(TT), \tag{3-1}$$

where f_N, f_S are pdf's of normal distributions with means μ_1, μ_2 and standard deviations σ_1, σ_2. π denotes the proportion of non-stopping vehicles.

The data collection in pilot study II can be viewed as sampling from two groups: group 1 and group 2. Within each group, vehicles have similar probabilities to be recorded. If, for matched vehicles, the travel time distributions of group 1 and group 2 are approximately symmetric, then μ_1, μ_2 could be estimated to give $\hat{\mu}_1$, $\hat{\mu}_2$.

To remove the bias of estimation, the actual proportion of non-stopping vehicles $\hat{\pi}$ has to be observed. Then the mean link travel time is estimated by

$$\hat{\pi} \times \hat{\mu}_1 + (1-\hat{\pi}) \times \hat{\mu}_2 . \tag{3-2}$$

3.2.2 Model Fitting Algorithm

In the travel time study, given the travel time of a specific vehicle, it is unknown that it belongs to group 1 or group 2. So we have an incomplete data problem. To improve the efficiency of data processing, the travel time data are grouped into *m* intervals, i.e. the input for model fitting algorithm is the number of observations n_1, …, n_m in intervals $[a_0, a_1), ..., [a_{m-1}, a_m)$. In this project, the length of intervals was chosen to be 1 second.

Let ψ denote the parameter set $(\pi, \mu_1, \mu_2, \sigma_1, \sigma_2)$. Then the probability that an individual travel time TT falls in the j^{th} interval is given by

$$P_j(\psi) = \int_{a_{j-1}}^{a_j} f(TT \mid \psi) dTT, \; j = 1,...m . \tag{3-3}$$

The grouped data follow a multinomial distribution and its likelihood function is

$$L(\psi) = \frac{n!}{n_1! \cdots n_m!} \{P_1(\psi)\}^{n_1} \cdots \{P_m(\psi)\}^{n_m} \tag{3-4}$$

and the log-likelihood is

$$\log L(\psi) = \sum_{j=1}^{m} n_j \log P_j(\psi) + \log \frac{n!}{n_1! \cdots n_m!}. \tag{3-5}$$

Maximizing $\log L(\psi)$ is equivalent to minimizing

$$Q = -\sum_{j=1}^{m} n_j \log P_j(\psi) . \tag{3-6}$$

The grouped travel time data were processed by an R software package *mixdist*([7]). This routine uses the standard maximum likelihood estimation method and combines the estimation-maximization (EM) algorithm with a Newton-type method implemented in the function *nlm*, provided by the R software ([8]). From the histograms of matched travel times, rough estimates of means of two groups were obtained and used as starting points. Applying Equation 3-3 gives $P_j(\psi)$. Minimizing Q generates a new ψ. Repeating the above two steps until convergence provided the estimates of ψ.

3.2.3 Binomial Test of Classification

A prerequisite of applying Equation 3-2 is that the algorithm is able to classify the non-stopping and stopped vehicles. A Binomial test was performed to verify this.

Let t_i be the travel time of i^{th} interval, n_i be the total number of observations in the video sample, and n_{iN} be the number of non-stopping vehicles in i^{th} interval. The probability of an observation in i^{th} interval from a non-stopping vehicle is given by

$$P_i = \frac{\pi f_N(t_i)}{\pi f_N(t_i) + (1-\pi) f_S(t_i)} \tag{3-7}$$

The classification error in i^{th} interval is defined as

$$\text{Classification Error} = \frac{|\, n_{iN} - P_i \times n_i \,|}{n_i} \times 100\% \tag{3-8}$$

Figure 3-6 shows the classification error plot. When the travel time is around 17s, the classification error is largest. This is reasonable since the two components in the mixture model overlap there.

In the i^{th} interval, the observations have a Binomial distribution with $p=P_i$, where P_i is the probability of being classified as non-stopping vehicles. We can test the hypothesis H_0: $p = P_i$ vs H_1: $p \neq P_i$ at a significance level $\alpha = 0.05$ for all intervals. We reject H_0 if

$$B(n_{iN}; n_i, P_i) \le \alpha/2 \text{ or } 1 - B(n_{iN} - 1; n_i, P_i) \le \alpha/2 . \tag{3-9}$$

where B denotes a binomial cumulative distribution function (CDF).

Figure 3-7 shows the P-value plot of this test. The minimum P-value is 0.425 (when t = 19s). That means the null hypothesis cannot be rejected and for all intervals the relative numbers of observed stopping and non-stopping vehicles are consistent with what would be expected from the mixture model.

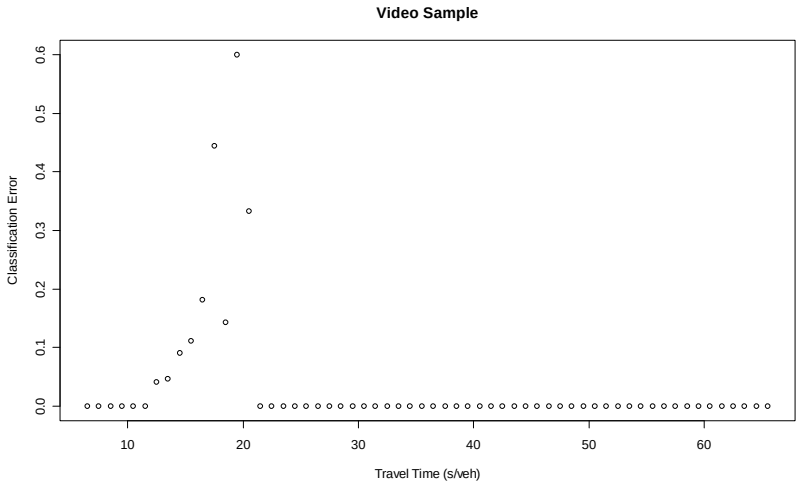

Figure 3-7 Classification Error Plot of Non-stopping Vehicles

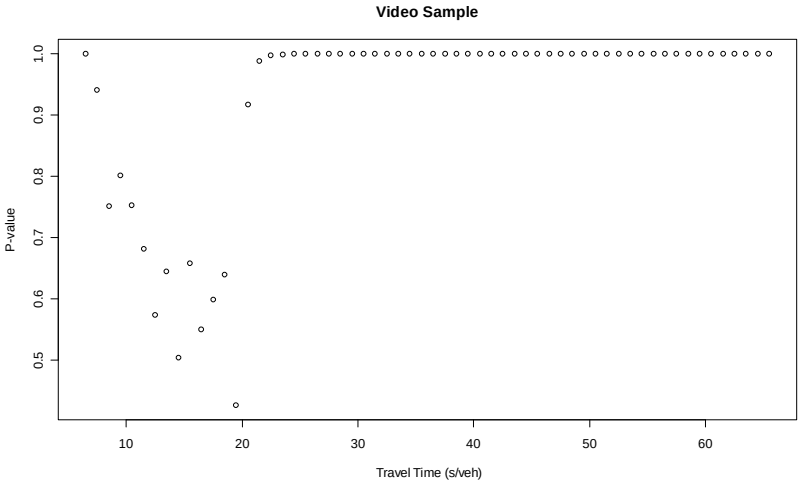

Figure 3-8 Classification P-value Plot of Non-stopping Vehicles

3.3 Estimation Result

The fitted curves of the video sample and license plate sample are shown in Figures 3-9 and 3-10 (assuming a normal mixture), and the R output of estimation is summarized in Table 3-1. They show that while the travel times of non-stopping vehicles can be approximately modeled by a normal distribution, the fit to the distribution of stopped travel times is weaker. Since the distribution of the stopped travel times is determined by many factors (the difference between cycle lengths of upstream and downstream signals, the green time of downstream intersection, the volume, the capacity, etc.), it is not surprising that it does not readily follow a single distribution. However, our objective is to estimate the mean link travel time, rather than to find a perfect fit for the matched license plate data. If the algorithm is able to classify vehicles, and estimate the means of each component correctly, it should be sufficient for our purpose.

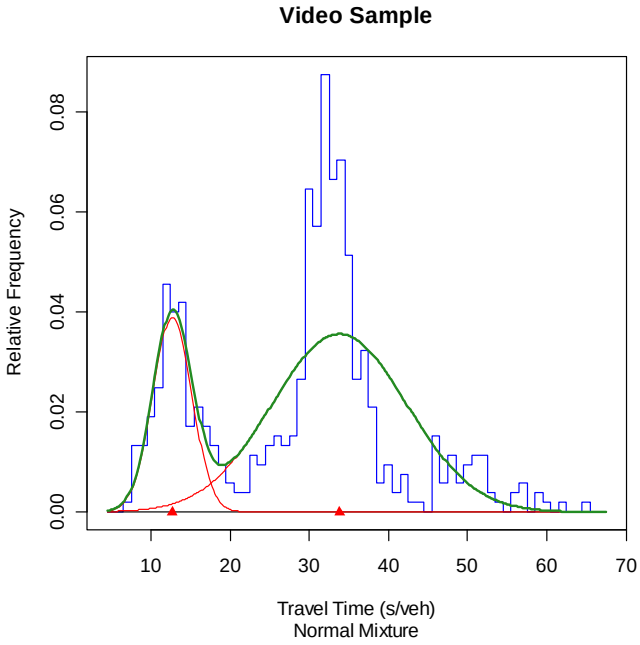

Figure 3-9 Fitted Curve of Video Sample

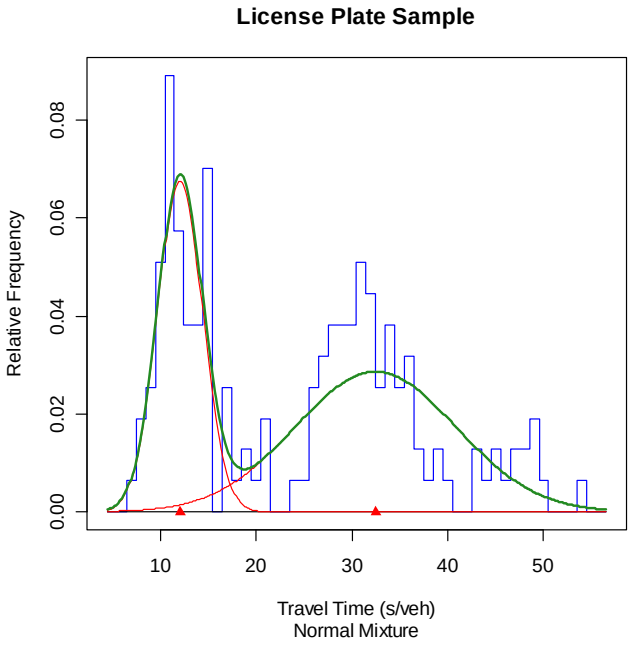

Figure 3-10 Fitted Curve of License Plate Sample

Table 3-1 R Output of Estimation Results

	π	μ_1	μ_2	σ_1	σ_2	Mean TT
Video Sample EM Estimation	0.242	12.71 [12.18, 13.24]	33.81 [32.81, 34.81]	2.48	8.49	28.83
License Plate Sample EM Estimation	0.362	12.03 [11.36, 12.70]	32.65 [30.49, 34.81]	2.27	9.45	27.78
Direct Observation from Video	0.236	12.45	33.74	2.34	8.48	28.71

In Table 3-1, it can be seen that the average travel times for the mixture components estimated from the video sample belong to the 95% confidence intervals (CI) estimated from the license plate sample. That is, $\mu_1 = 12.45 \in [12.18,13.24]$, and $\mu_2 = 33.79 \in [32.81,34.81]$.

In the video sample, $\hat{\pi} = 133/525 = 0.2533$. Substituting this into Equation 3-2, along with the component averages from the license plate sample gives:
$0.236 \times 12.71 + (1 - 0.236) \times 33.81 = 28.83$ s, which lowers the estimation error from (28.71-25.21) = 3.5s to (28.83-28.71) = 0.12s.

Chapter 4 Travel Time Data Collection

4.1 Sampling Arterial Sites

4.1.1 Available Network Data Sources

The network files helpful and available for the arterial sites sampling were as follows:

- Twin Cities 1990 transportation network Access file "metcouncil.mdb". It can be opened by the ArcGIS software to generate a network map. In the table "AM90HWY", the attributes describing the link-related information are shown in Figure 4-1, where "A" and "B" denote origin and destination nodes respectively. For divided arterials, ASGNGRP (assignment group) is 5; for undivided arterials, it is 6.

Field Name	Data Type
OBJECTID	AutoNumber
Shape	OLE Object
A	Number
B	Number
DISTANCE	Number
TIME1	Number
TIME2	Number
CAPACITY	Number
LINKGRP1	Number
LINKGRP2	Number
LINKGRP3	Number
ASGNGRP	Number
USER_	Number
COST	Number
TWOWAY	Number
VOLUME	Number
DIRCODE	Number
Shape_Length	Number

Figure 4-1 Link Attribute Table of Twin Cities 1990 Transportation Network

- Twin Cities 2000 transportation network Access file "functional roads.mdb". It can also be opened by ArcGIS software to generate a network map. In the table "TC2000_roads", the link attributes "STREETALL", "ALT_NAM1", and "ALT_NAM2" provide the link's address information.

	Field Name	Data Type
▶	OBJECTID	AutoNumber
	Shape	OLE Object
	LENGTH	Number
	STREETALL	Text
	ALT_NAM1	Text
	ALT_NAM2	Text
	TLGID	Number
	FUN_CLAS	Number
	Shape_Length	Number

Figure 4-2 Link Attribute Table of Twin Cities 2000 Transportation Network

- Twin Cities Transportation Network Excel file "1995-2020.xls", including worksheets "notes", "1995 network", and "2020 network". The "notes" worksheet gives the column definitions in the 1995 and 2020 networks (see Table 4-1).

Table 4-1 Column Definitions of Twin Cities 1995 and 2020 Transportation Network

	Column Name	Definition
1	Node 1	Origin node for the link
2	Node 2	Destination node for the link
3	Assgn Grp	Assignment group (the link type)
4	Location	Region type
5	Lanes	Number of lanes the link contains
6	Length	Length of the link in miles
7	FF Time	Free-flow travel time on the link in minutes
8	Capacity	Given in vehicles per hour
9	AM Peak	Traffic flow from 6:30 to 7:30am
10	PM Peak	Traffic flow from 3:40 to 4:40pm
11	Daily	Total daily traffic flow

4.1.2 Arterial Sites Sampling Procedure

A single network file cannot provide all information needed for test and comparison of candidate models (such as capacity, number of lanes, free-flow time, link length) and enable us to locate sampled sites in real world as well. A comparison of origin node, destination node, capacity, assignment group between the 1990 and 1995 network arterial links shows that each origin/destination pair is unique and represents the same link in these two networks. The procedure followed in arterial sites sampling is listed as below:

- Randomly select 150 sites out of 8754 arterial links in Twin Cities 1990 network (arterial site ID 1, …, 150).
- In the ArcGIS software, overlap the 1990 network, the 2000 network, as well as a Zip code GIS file ([10]) to obtain the real world addresses of 150 sites.

- Link attribute tables of the 1990 and 1995 Networks based upon one-to-one relationship of origin/destination nodes.
- Pre-observe the sites in June 2006 to cross some sites off the list if: (1) they do not have traffic signals; (2) they do not have convenient nearby parking; (3) there is very light traffic on them; (4) they are too short (less than 0.1mile).
- Determine 65 sites as initial candidates (55 of them became final sample sites).

The Figure 4-3 shows the map of sampled sites generated by the ArcGIS software.

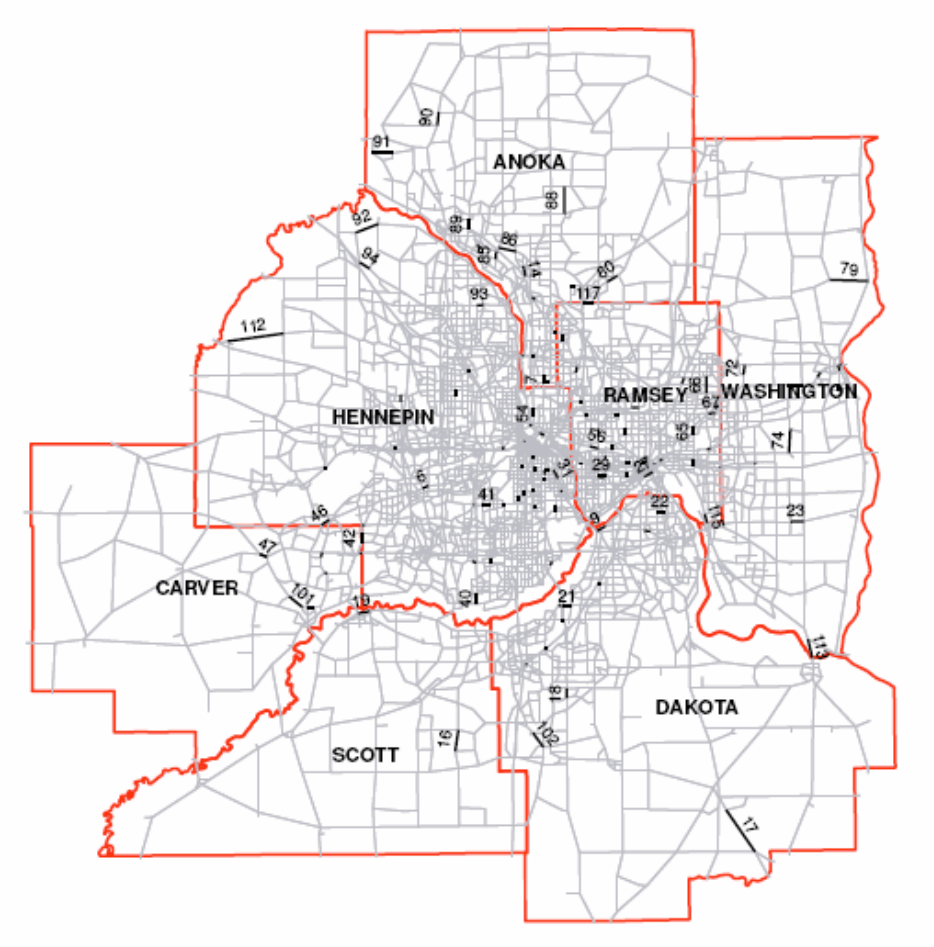

Figure 4-3 Map of Sampled Sites

4.2 Field Data Collection

In order to perform model tests, calibration, and comparison, the field study needs to accomplish two tasks: (1) collect signal timing and traffic data for travel time calculation of candidate models; (2) collect real travel time data by modified plate-matching method. Table 4-2 lists the data requirement for the four candidate models.

Table 4-2 Data Requirement for Candidate Models

Model	Data Requirement
BPR function	Volume (v), capacity (c), link length (L), free-flow speed (FFS)
Conical volume-delay function	Volume (v), capacity (c), link length (L), free-flow speed (FFS)
Singapore model	Volume (v), capacity (c), link length (L), free-flow speed (FFS), effective green time (g), cycle length (C)
Skabardonis-Dowling model	Volume (v), capacity (c), link length (L), free-flow speed (FFS), effective green time (g), cycle length (C), proportion of arrivals on green (P)

Figure 4-4 shows the illustration of field data collection. From June to September 2006, two graduate students studied the 55 sampled arterial sites. The observation periods were determined by the number of through lanes and traffic situation. For those sites that had very light traffic or only one through lane, the plate-matching method was able to record most vehicles and there was no bias in the proportion of non-stopping/stooped vehicles. In this case, 30 minutes was deemed sufficient. Otherwise, the study period was 1 hour. The last 3 digits of license plate numbers were input into laptop computers when vehicles entered/left the sampled link. Every time before the field observation, the two computers were coordinated via Internet. One camcorder was installed at the downstream intersection to record the traffic and signal. The ideal location of camcorder was beyond the maximum queue length from the downstream intersection. However, for security reasons, usually it was installed less than 150ft from the downstream observer, which made the proportion of non-stopping vehicles non-observable. In addition, some sites had a high proportion of left or right turns. Thus there were 5 observed sites not used in the final model comparison. Their ID numbers are: 6, 38, 85, 104, and 134.

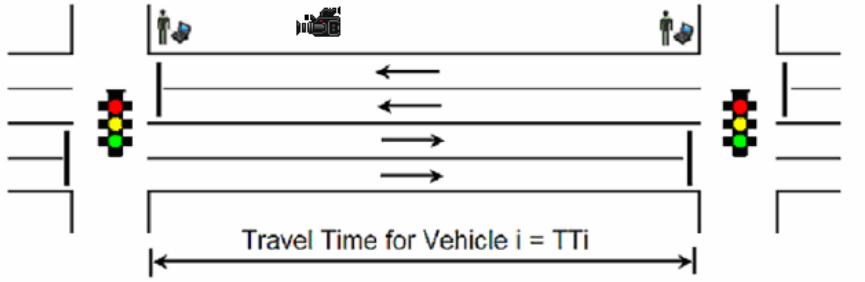

Figure 4-4 Illustration of Field Data Collection

Table 4-3 lists the data collection methods.

Table 4-3 Data Collection Method

Data	Collection Method
Volume (v)	Count from the video
Capacity (c)	Use default values of 1995 network
Link length (L)	(1) If origin and destination nodes in the 1990 Network match the sample link, use the link length of network file; (2) otherwise, estimate from Google maps
Free-flow speed (FFS)	Use default values of 1995 network
Effective green time (g)	Observe from the video
Cycle length (C)	Observe from the video
Proportion of non-stopping vehicles ($\hat{\pi}$)	Count the number of non-stopping vehicles from the video and divide it by volume (v)
Proportion of arrivals on green (P)	(1) Observe from video; (2) if direct observation is not available, estimate from the proportion of non-stopping vehicles
Measured Travel Time (TT)	Record last three characters of vehicles and process the data by modified plate-matching method

4.3 Data Processing

The field data collected by two observers were stored in two Access tables, including the plate numbers and time stamps. After one site observation, these two tables were saved to a plate-matching Access program as "site_ID_up" and "site_ID_down". In the program, a one-to-one query was designed to match all the same plate numbers of upstream and downstream sites.

However, directly generating the travel times from this query is problematic. For example, if both observers recorded the same plate number 123 at time t_1, t_2 from the downstream site and at time t_3, t_4 from the upstream site. Then from the query, we may have following matched travel times: (1) t_1-t_3, (2) t_1-t_4, (3) t_2-t_3, and (4) t_2-t_4. It is obvious that at most two of these travel times are true, i.e. the corresponding matched plate number is recorded from the same vehicle. There are three usually-used screening algorithms for the traditional plate matching method ([7]): (1) delete those matched travel times corresponding to speeds less than 5km/h and greater than 120km/h; (2) delete those matched travel times falling outside three or four standard deviations; (3) visual inspection of travel time/speed profile to identify and delete outlying matched travel times. None of the above can be directly used for our data screening in that the lower bound and upper bound of the travel time is highly site specific, which cannot be determined by global parameters or a general rule.

Given a sampled link, there should be a reasonable range for its travel time determined by its signal timing and posted speed limit. If it is known, then the spurious matched travel times could

be easily removed. It is worth noting that it is not guaranteed that all the remaining matched travel times are true. For example, suppose vehicles A and B have the same last three digits on their plate numbers. Vehicle A passed the upstream site at 9:15:15 and downstream site at 9:17:00, and vehicle B passed the downstream site at 9:15:25 and downstream site at 9:17:15. It is possible that we only get time stamps 9:15:15 and 9:17:15. Their difference may fall inside the reasonable range for that site and we accept a false match. However, in our screening algorithm, this case is ignored since it is a low probability event and dependent on the number of arrivals in a time period. If the traffic is light, it has an even lower probability to happen. If the traffic is heavy, adding a reasonable travel time does not make any significant difference in terms of estimation result. In this project, as a rule of thumb, the lower and upper bounds are chosen as

$$T_L = \frac{L \times 3600}{\text{Speed Limit} + 10}\ , \tag{4-1}$$

and

$$T_U = \frac{L \times 3600}{\text{Speed Limit} - 10} + C \tag{4-2}$$

where C denotes the signal's cycle length.

To obtain the signal timing information from the video, for pre-timed signals, at least 5 cycles were observed to take the average; for actuated signals, 30 minutes of video were observed.

4.4 Data Summary

The field data of 50 sites are summarized in Table 4-4. The columns are defined as follows:

- Site ID: index with respect to the original 150 sampled sites,
- c: capacity (veh/h/direction),
- Speed limit: posted speed limit (mph),
- L: link length (mile),
- v: volume (veh/h),
- C: cycle length (s),
- g: effective green time (s),
- TT: travel time (adjusted by the algorithm),
- N: number of through-lanes in one direction,
- P: proportion of arrivals on green,
- FFT: free-flow time (in second, estimated from travel time histogram plot), and
- FFS: free-flow speed (default value of 1995 network).

Table 4-4 Data Summary for Model Comparison

Site ID	c	Speed Limit	L	v	C	g	TT	N	P	FFT	FFS
2	650	30	0.21	388	90	45	40.64	1	0.42	32	28
3	750	35	0.51	268	93	35	84.32	1	0.22	53.69	38
8	750	30	0.35	268	146	30	91.11	1	0.17	52.47	38
14	1700	40	0.75	1487	126	55	89.85	2	0.29	67.5	41
24	1200	30	0.29	432	70	50	38.91	2	0.63	33.42	28
25	1300	30	0.32	529	80	44	49.02	2	0.46	41.36	28
27	1500	40	0.72	995	100	58	58.93	2	0.69	49.4	37
29	1500	30	0.49	522	68	28	58.13	2	0.3	48	37
30	650	30	0.37	200	65	22	73.23	1	0.5	65	28
31	1300	30	0.4	450	90	40	51	2	0.7	41.17	28
32	650	30	0.29	212	87	29	55.4	1	0.34	40	28
33	1300	30	0.24	701	45	25	29.84	2	0.79	26.29	28
34	650	30	0.24	146	90	26	51.64	1	0.44	30	28
35	1300	35	0.24	539	90	60	33.4	2	0.69	25.46	28
36	650	30	0.24	200	56	18	43.91	1	0.51	35	28
39	650	30	0.43	276	90	40	59.15	1	0.53	50	28
40	1700	45	0.76	815	50	24	74.55	2	0.57	66.06	41
41	750	30	0.5	483	90	45	60.58	1	0.47	46.99	38
50	550	40	0.19	357	120	40	63.09	1	0.14	21.1	33
51	750	35	0.34	306	100	22	55.24	2	0.18	27.28	38
52	1700	30	0.2	559	70	50	21.13	2	0.76	17.73	41
53	1300	30	0.15	274	90	40	37.59	2	0.18	21.76	28
54	1300	30	0.48	670	90	60	51.72	2	0.52	41.12	28
55	1500	30	0.45	1160	85	28	60.13	2	0.12	28.16	38
61	1300	30	0.49	650	102	24	69.71	2	0.23	45	28
63	1300	30	0.49	614	80	52	60.76	2	0.72	55.86	28
64	750	35	0.51	212	122	34	94.07	1	0.17	61.26	38
65	1300	30	0.19	422	76	33	31.84	2	0.51	20.58	28
66	750	30	0.47	124	88	28	70.24	1	0.36	58	38
67	750	30	0.37	146	76	9	67.06	1	0.41	45	38
68	750	30	0.48	214	52	12	76.05	1	0.29	55	38
70	750	35	0.85	106	54	16	95.9	1	0.35	85	38
80	900	50	0.75	362	100	75	62.3	1	0.67	56.05	48
81	1700	40	0.24	571	86	65	33.82	2	0.69	28.59	41
83	1900	50	0.45	603	80	32	44.6	2	0.69	36.08	47
86	850	45	0.83	246	112	26	115.14	1	0.12	90	41
87	600	30	0.49	257	50	24	51.69	1	0.37	50	28
96	650	30	0.14	276	90	25	39.81	1	0.26	22	28
97	650	30	0.13	111	90	44	43.8	1	0.33	25	28
98	1300	30	0.24	966	90	70	46.46	2	0.6	39.24	28
100	1300	30	0.25	733	94	50	42.1	2	0.5	26.72	28
109	650	30	0.26	301	88	54	44.97	1	0.58	40	28
111	1300	30	0.23	876	90	36	50.45	2	0.62	31.56	28
130	1500	30	0.24	157	62	14	47.34	2	0.21	30	23
131	1500	30	0.21	607	110	86	28.76	2	0.71	24.68	38
140	1900	45	0.38	513	86	30	55.5	2	0.34	35.13	47
142	1500	30	0.25	618	68	34	34.26	2	0.82	30.01	37
146	2250	35	0.49	334	90	33	63.2	3	0.66	47.91	37
148	1400	35	0.2	1098	62	44	31.27	2	0.35	21.5	37
149	1500	30	0.8	647	100	60	102.18	2	0.53	90.06	37

Chapter 5 Model Comparison and Evaluation

5.1 Mean Absolute Percentage Error (MAPE) and Nonlinear Regression Model

To compare across 4 candidate models in terms of their prediction accuracy, the mean absolute percentage error is defined as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|TT_i - \hat{TT}_i|}{TT_i} \times 100\% \ . \quad (5\text{-}1)$$

A lower value for MAPE means more accurate prediction of travel time.

When calibrating the candidate models, a regression analysis needs to be performed. The model has the form

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon} \quad (5\text{-}2)$$

where $\mathbf{y}$ = response variable, $\mathbf{x}$ = predictors, $\boldsymbol{\theta}$ = parameters to be estimated, $\boldsymbol{\varepsilon}$ = estimation errors, and f is a function of $\mathbf{x}$ (if f is linear, the regression is linear; otherwise, the regression is nonlinear). $\boldsymbol{\theta}$ is estimated to obtain the best fit of model for the data ($\mathbf{y}$, $\mathbf{x}$) such that the sum of square error (SSE) is minimized.

The estimated parameters in a nonlinear regression model do not have closed functional forms if they cannot be transformed into a linear regression problem. We chose the Bayesian WinBUGS software ([11]) instead of the R software to carry out the model calibration because: (1) WinBUGS tends to provide better estimates of parameter quantiles; (2) convergence is more tractable. The latter is important especially we have to "guess" the initial values of parameters.

5.2 BPR Function and Conical Volume-Delay Function

The Mn/DOT and the Metropolitan Council currently use the BPR function and conical volume-delay function. Their parameters are pre-determined and so unlike the comparison targets of the Singapore model and the Skabardonis-Dowling models, it is not necessary to calibrate them. For these two models, we simply plug in the 1995 network default values (free-flow speed, capacity) plus field data (volume) to obtain the estimated travel time. Table 5-1, Figures 5-1 and 5-2 give some details about these two models.

Table 5-1 Travel Time Estimation Models without Calibration

Travel Time Estimation Model	y	x	θ	MAPE
BPR function $TT = \frac{L}{FFS}(1+\alpha(\frac{v}{c})^{\beta})$	TT	v, c, L, FFS	α=0.15 β=4	28.7%
Conical volume-delay function $TT = \frac{L}{FFS}(2+\sqrt{\alpha^2(1-\frac{v}{c})^2+\beta^2}-\alpha(1-\frac{v}{c})-\beta)$	TT	v, c, L, FFS	α=4, β=1.167 divided arterial; α=5, β=1.125 undivided arterial	24.7%

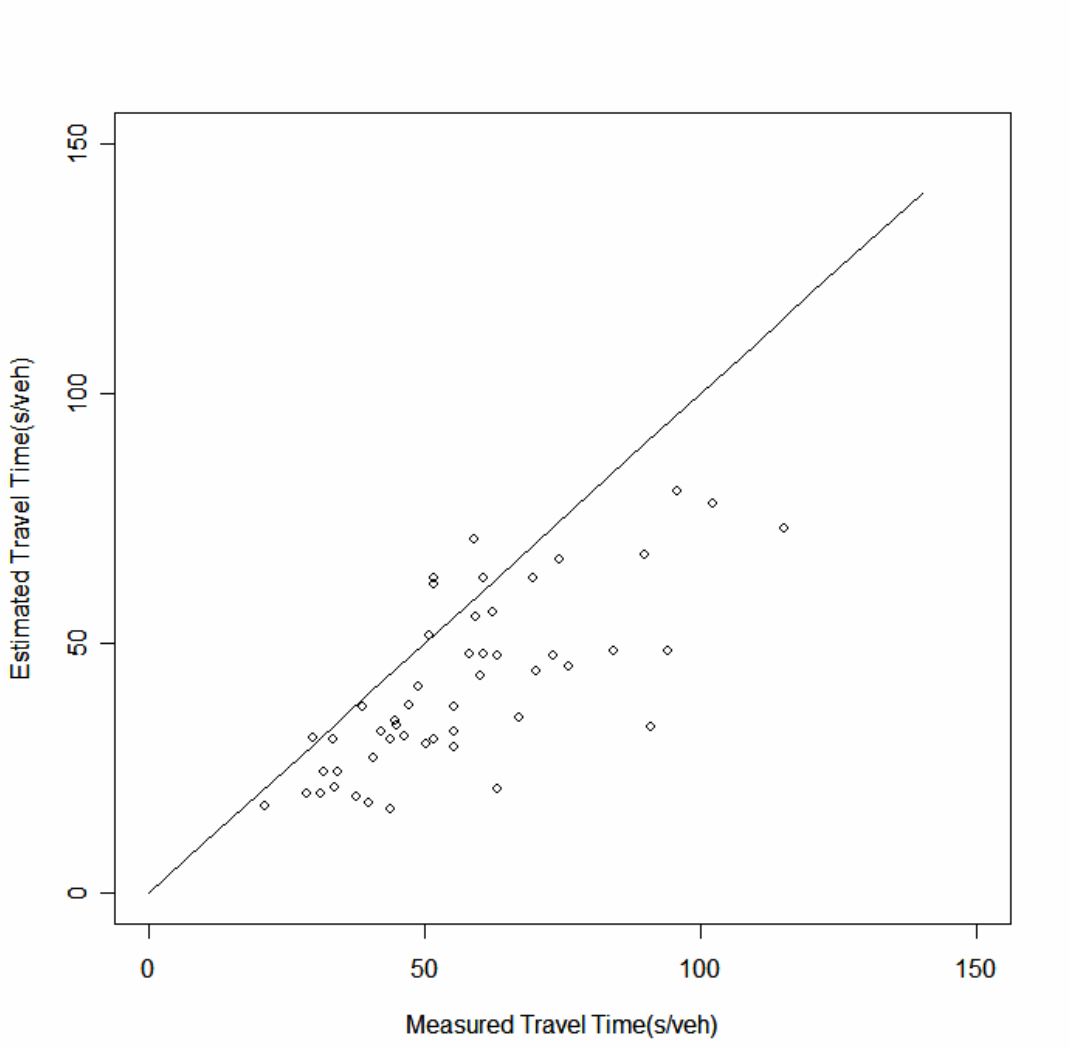

Figure 5-1 Measured Travel Time vs Estimated Travel Time (BPR Function)

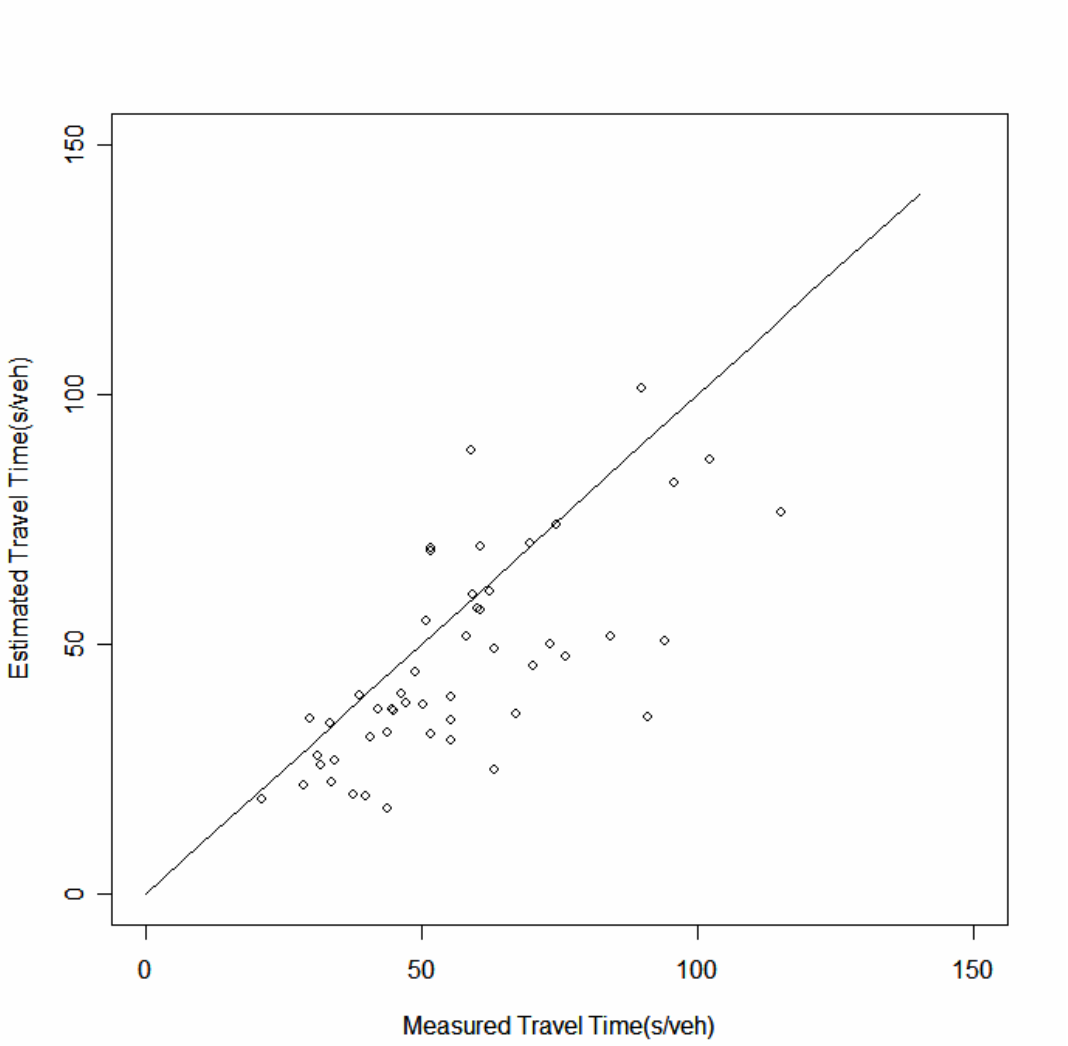

Figure 5-2 Measured Travel Time vs Estimated Travel Time

(Conical Volume-Delay Function)

Figures 5-1 and 5-2 confirm again that both the BPR function and the conical volume delay function tend to underestimate the travel time. In these two figures, we may notice some data points above the straight line where the estimated travel is equal to the measure travel time. That is because for some sites the network FFS default values appear to be underestimated.

5.3 Singapore Model and Skabardonis-Dowling model

We calibrated the Singapore model and Skabardonis-Dowling model on the basis of data availability. The assumed scenarios are shown in Tables 5-2 and 5-3 respectively, in which "√" and "N/A" indicate the availability or non-availability of site-specific field data. The corresponding models and their MAPE's are summarized in Tables 5-4 and 5-5.

Table 5-2 Data Availability Scenarios for Singapore Model

	Signal Timing	FFT	Assumption
1	√	√	The actual FFT is a function of observed FFT
2	√	N/A	(1) The actual FFS is a function of the default network FFS; (2) All sites have the same proportion of arrivals on green
3	N/A	√	(1) The actual FFT is a function of observed FFT; (2) All sites have the same signal timing
4	N/A	N/A	(1) The actual FFS is a function of the default network FFS; (2) All sites have the same signal timing

Table 5-3 Data Availability Scenarios for Skabardonis-Dowling Model

	Signal Timing	FFT	P	Assumption
1	√	√	√	The actual FFT is a function of observed FFT
2	√	√	√	(1) The actual FFT is a function of observed FFT; (2) All sites have the same saturation flow (veh/h/lane)
3	√	√	N/A	(1) The actual FFT is a function of observed FFT; (2) All sites have the same proportion of arrivals on green
4	√	N/A	N/A	(1) The actual FFS is a function of the default network FFS; (2) All sites have the same proportion of arrivals on green
5	N/A	N/A	N/A	(1) The actual FFS is a function of the default network FFS; (2) All sites have the same control delay

From Tables 5-4 and 5-5, we can see that: the site-specific free-flow travel time and the signal timing are the most important and second most important predictors for travel time; while the proportion of arrivals on green is not that important. Without it, a MAPE of less than 7% could still be achieved. The travel time does not appear to be sensitive to the capacity, in that the site-specific saturation flow (veh/h/lane) does not significantly improve the performance of the Skabardonis-Dowling model. Finally, the Skabardonis-Dowling model is more desirable than the Singapore because it has a simpler form and relatively better performance than the Singapore model.

Table 5-4 Cases of Singapore Model

	Travel Time Estimation Model	y	x	θ	MAPE(95%CI)
1	$TT=(a+b\times FFT)+\frac{9}{10}[\frac{C(1-\frac{g}{C})^2}{2(1-\frac{g}{C}\times\frac{v}{c})}+\frac{(\frac{v}{c})^2}{\frac{v}{1800}(1-\frac{v}{c})}]$	TT	FFT, v, c, g, C	a, b	6.9% ([6.7%, 7.4%])
2	$TT=\frac{L}{w\times FFS}+\frac{9}{10}[\frac{C(1-\frac{g}{C})^2}{2(1-\frac{g}{C}\times\frac{v}{c})}+\frac{(\frac{v}{c})^2}{\frac{v}{1800}(1-\frac{v}{c})}]$	TT	FFT, v, c, g, C	w	13.2% ([13.0%, 4.1%])
3	$TT=(a+b\times FFT)+\frac{9}{10}[\frac{C(1-\frac{g}{C})^2}{2(1-\frac{g}{C}\times\frac{v}{c})}+\frac{(\frac{v}{c})^2}{\frac{v}{1800}(1-\frac{v}{c})}]$	TT	FFT, v, c, g=38.48s/veh	a, b, C	15.0% ([13.6%,16.7%])
4	$TT=\frac{L}{w\times FFS}+\frac{9}{10}[\frac{C(1-\frac{g}{C})^2}{2(1-\frac{g}{C}\times\frac{v}{c})}+\frac{(\frac{v}{c})^2}{\frac{v}{1800}(1-\frac{v}{c})}]$	TT	FFT, v, c, g=38.48s/veh	w, C	20.4% ([18.5%, 3.7%])

Table 5-5 Cases of Skabardonis-Dowling Model

	Travel Time Estimation Model	y	x	θ	MAPE(95%CI)
1	$TT=[(a+b\times FFT)+0.5(1-P)(C-g)](1+0.05(\frac{v}{c})^{10})$	TT	FFT, P, C, g, v, c	a, b	6.3% ([6.1%, 6.9%])
2	$TT=[(a+b\times FFT)+0.5(1-P)(C-g)](1+0.05(\frac{v}{N\times\frac{g}{C}\times s})^{10})$	TT	FFT, P, C, g, N, v	a, b, s	6.4% ([6.1%, 7.0%])
3	$TT=[(a+b\times FFT)+0.5(1-P)(C-g)](1+0.05(\frac{v}{N\times\frac{g}{C}\times s})^{10})$	TT	FFT, C, g, N, v	a, b, s, P	6.6% ([6.2%, 7.5%])
4	$TT=[(b\times FFT+T_d](1+0.05(\frac{v}{c})^{10})$	TT	FFT, v, c	b, T_d	14.9% ([13.5%, 16.9%])
5	$TT=(\frac{L}{w\times FFS}+T_d)(1+0.05(\frac{v}{c})^{10})$	TT	L, FFS, v, c	w, T_d	20.6% ([18.4%, 24.3%])

5.4 Cross-Validation of Skabardonis-Dowling Model

In the previous section, the data used to calibrate a model were also used to assess its accuracy. A more informative assessment would involve comparing predicted and measured travel times for locations not used in calibration. Since our sample size is small, to avoid over-fitting the model, we performed the leave-one-out cross-validation

(LOOCV) of the Skabardonis-Dowling, where one site is deleted from the calibration data set, the model is calibrated using the remaining sites, and then the predicted and measured travel times for the deleted site are compared. Repeating this procedure with each site taking a turn as the deleted site then gives us an idea of how the model should perform when applied to new sites not included in our sample.

It turned out this is procedure was more easily accomplished using the R software package Brugs. As noted above, the procedure is as follows: (1) use a single observation as the validation data, and the remaining 49 observations as the training data; (2) run the WinBUGS via R to obtain the absolute percentage error of the single observation; (3) repeat 50 times and then compute MAPE. Table 5-6 lists the results of LOOCV.

Table 5-6 LOOCV of Skabardonis-Dowling Model

	Travel Time Estimation Model	y	x	θ	MAPE
1	$TT = [(a + b \times FFT) + 0.5(1 - P)(C - g)](1 + 0.05(\frac{v}{c})^{10})$	TT	FFT, P, C, g, v, c	a, b	6.4%
2	$TT = [(a + b \times FFT) + 0.5(1 - P)(C - g)](1 + 0.05(\frac{v}{N \times \frac{g}{C} \times s})^{10})$	TT	FFT, P, C, g, N, v	a, b, s	6.6%
3	$TT = [(a + b \times FFT) + 0.5(1 - P)(C - g)](1 + 0.05(\frac{v}{N \times \frac{g}{C} \times s})^{10})$	TT	FFT, C, g, N, v	a, b, s, P	7.1%
4	$TT = [(b \times FFT + T_d](1 + 0.05(\frac{v}{c})^{10})$	TT	FFT, v, c	b, T_d	15.4%
5	$TT = (\frac{L}{w \times FFS} + T_d)(1 + 0.05(\frac{v}{c})^{10})$	TT	L, FFS, v, c	w, T_d	21.3%

Chapter 6 Summary and Conclusion

The primary objective of this project was to identify and evaluate parametric models to use in making default estimates of travel times on arterial links, using information typically available from a transportation planning model. The chosen method of evaluation was to compare travel time predictions generated by the models to field measurements. A review of the literature revealed several candidates models, including the Bureau of Public Roads (BPR) function, Spiess's conical volume delay function, the Singapore model, the Skabardonis-Dowling model, and the Highway Capacity Manual's model. In a first pilot study, average travel times on an arterial link were measured using (1) a combination of spot speed data to estimate mean free-flow travel times and intersection delay measurements to estimate average waiting time, and (2) a floating car method. The floating car method turned out to be sensitive to the relative fraction of runs where the floating car was delayed by a red signal indication, and it was not possible to obtain enough runs to reliably estimate this fraction. A comparison of travel time models using the data collected with method (1) indicated best performance by the HCM model, the worst performance was by the BPR and conical volume-delay models, with the remaining two models being slightly worse than the HCM model.

The combination of spot speed/intersection delay method of measuring travel times would generally require three individuals to collect data, and project resources limited the total number of person-hours available for field data collection. To maximize the number of data collection sites within this constraint it was decided to evaluate a license plate matching method for collecting travel time data, which could be carried out by only two people. A second pilot study comparing the travel times measured using the license plate method to travel times measured from a video recording revealed a tendency for the license plate method to under-sample vehicles stopped by red signal indications, and hence to underestimate average travel times. However, by using a mixture decomposition method to estimate mean travel times for stopped and non-stopped vehicles, together with an independent estimate of proportion stopping, an estimation method that substantially eliminated this bias was developed.

The license plate method was then applied to a sample of 50 arterial links located in the Twin Cities seven county metropolitan area to obtain measurements of average travel time. Also obtained were the lengths of each link, measurements of traffic volume, and signal timing information. Default values for model parameters were obtained from the Twin Cities planning model's database. Using network default parameters, we found that the BPR and conical volume-delay models produced mean average percent errors (MAPE) of about 25%, while the Singapore and Skabardonis-Dowling models, using maximal site-specific information, produced MAPE values of between 6% and 7%. As site-specific information was replaced by default information, the performance of the latter two models deteriorated, but even under conditions of minimal information, the models produced MAPE values of around 20.5%. A cross-validation study of the Skabardonis-Dowling model showed essentially similar performance when predicting travel times on links not used to estimate default parameter values.

Overall, the Singapore and Skabardonis-Dowling models showed similar accuracies under all data availability scenarios, were able to effectively use site-specific information when it was available, and performed at least as well as the more traditional models when site-specific data were minimal. Their superiority to the more traditional models indicates the importance of including signal timing information when predicting the travel times on signalized arterial links. For travel time prediction on links containing no intervening signals either model appears to be acceptable. If the network model contains arterial links with intervening signals, then the more general form of the Skabardonis-Dowling model would be preferred, and our recommendation is that the Skabardonis-Dowling model be used in the second phase of this project.

References

[1] *Traffic Assignment Manual*. U.S. Department of Commerce, Bureau of Public Roads, June 1964.

[2] H. Spiess, "Conical volume-delay functions". *Transportation Science*, Vol. 24, No.2, 1990.

[3] C. Xie, R. Cheu, and D. Lee, "Calibration-Free Arterial Link Speed Estimation Model Using Loop Data". *ASCE J. of Transportation Engineering*, Nov/Dec 2001, pp. 507-514.

[4] A. Skabardonis and R. Dowling, "Improved Speed-Flow Relationship for Planning Applications". *Transportation Research Record: Journal of the Transportation Research Board, No. 1572*, National Research Council, Washington, D.C., 1997, pp. 18-23.

[5] Transportation Research Record, *Special Report 209: Highway Capacity Manual*. National Research Council, Washington, D.C., 2000.

[6] W. Homburger and J. Kell, *Fundamentals of Traffic Engineering, 12th Edition.* Institute of Transportation Studies, University of California, Berkeley, 1989.

[7] J. Du , *Combined Algorithms for Constrained Estimation of Finite Mixture Distributions with Grouped Data and Conditional Data*, www.math.mcmaster.ca/peter/mix/JuanDuReport.pdf. Accessed April 20, 2006.

[8] W. N. Venables, D. M. Smith and the R Development Core Team, *An Introduction to R*. http://cran.r-project.org/doc/manuals/R-intro.pdf. Accessed October 11, 2005.

[9] Office of Highway Policy Information, *Travel Time Data Collection Handbook.* www.fhwa.dot.gov/ohim/start.pdf. Accessed June 28, 2006.

[10] Metropolitan Council, ZIP Code Boundaries - 5 Digit (Twin Cities). http://www.datafinder.org/metadata/zip5_a.htm. Accessed March 10, 2006.

[11] D. Spiegelhalter, A. Thomas, N. Best, and Dave Lunn, *WinBUGS manual*. http://www.mrc-bsu.cam.ac.uk/bugs. Accessed January 23, 2006.

[12] R. Dowling, W. Kittelson, J. Zegeer, and A. Skabardonis, *NCHRP Report 387: Planning Techniques to Estimate Speeds and Service Volumes*. TRB, National Research Council, Washington, D.C., 1997.

[13] D. Branston, "Link Capacity Functions: A Review. *Transportation Research*, Vol. 10, 1976, pp. 223-236.

[14] J. Colyar and N. Rouphail, “Measured Distributions of Control Delay on Signalized Arterials”. *Transportation Research Record: Journal of the Transportation Research Board, No. 1852*, TRB, National Research Council, Washington, D.C., 2003, pp. 1-9.

Appendix A. Travel Time Relative Frequency Plots

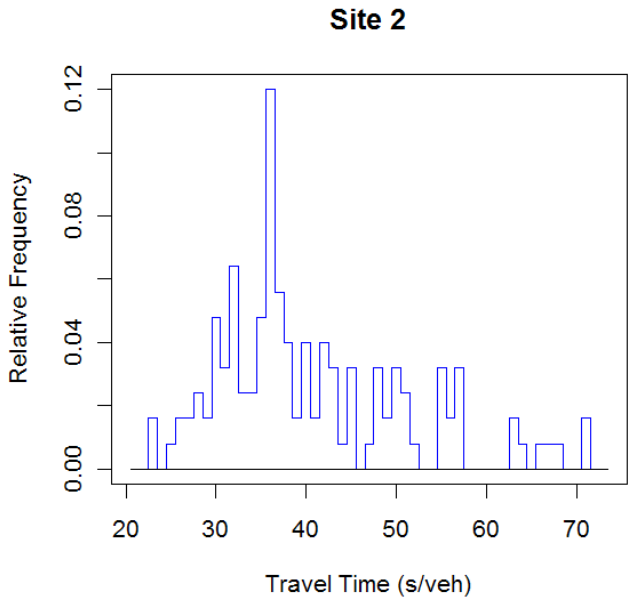

Figure A-1 Travel Time Relative Frequency Plot of Site 2

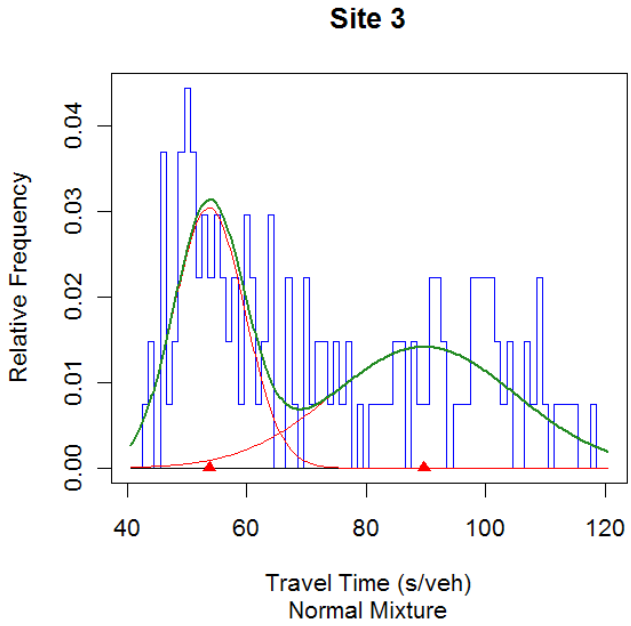

Figure A-2 Travel Time Relative Frequency Plot of Site 3

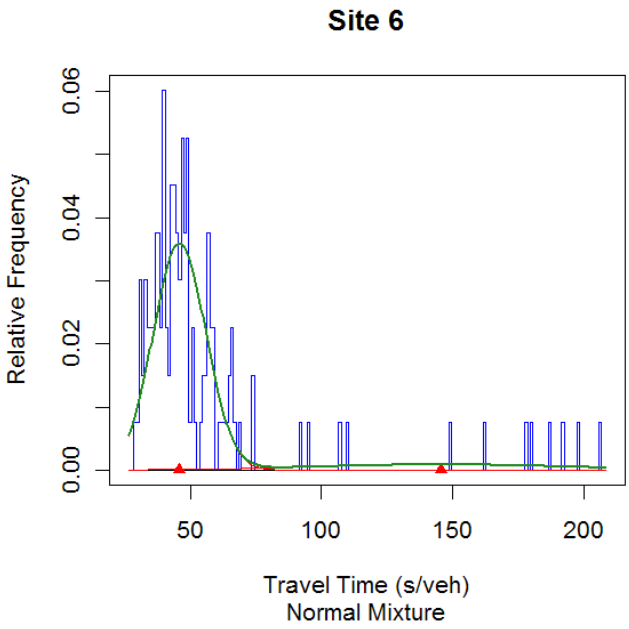

Figure A-3 Travel Time Relative Frequency Plot of Site 6

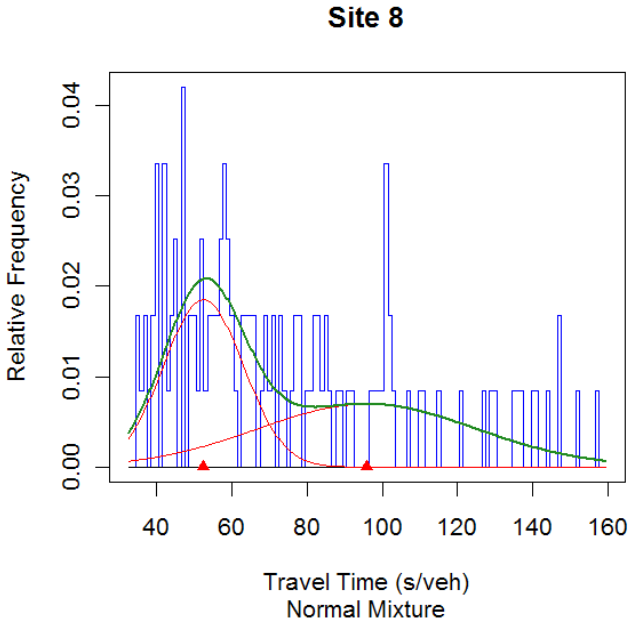

Figure A-4 Travel Time Relative Frequency Plot of Site 8

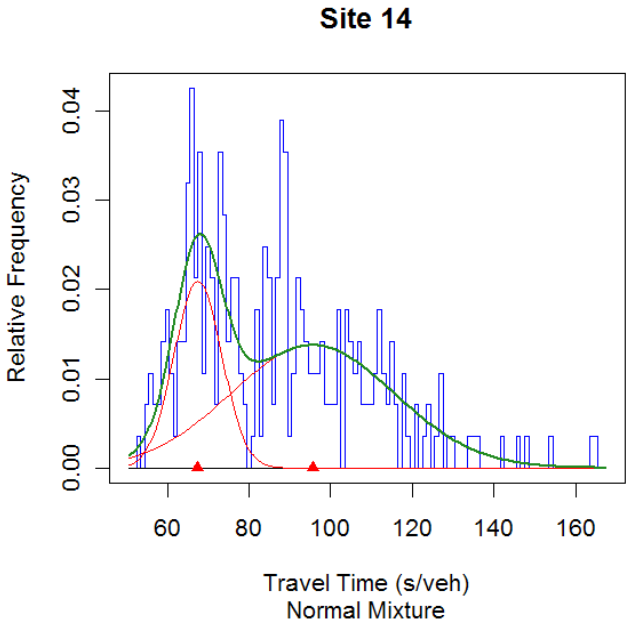

Figure A-5 Travel Time Relative Frequency Plot of Site 14

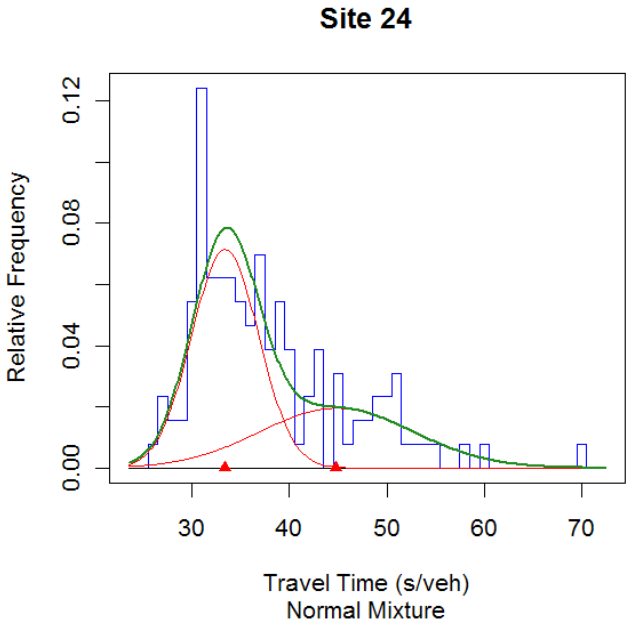

Figure A-6 Travel Time Relative Frequency Plot of Site 24

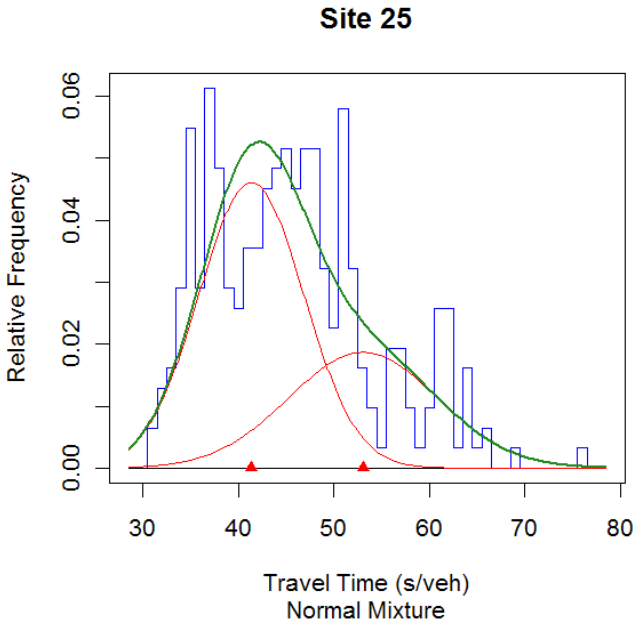

Figure A-7 Travel Time Relative Frequency Plot of Site 25

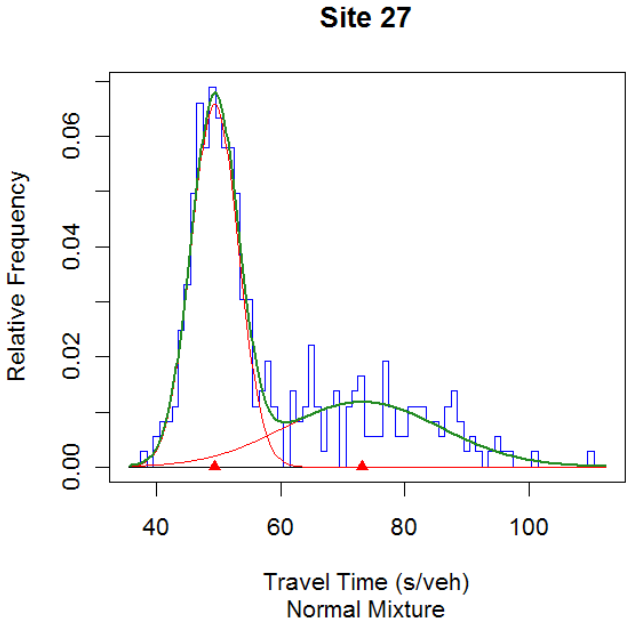

Figure A-8 Travel Time Relative Frequency Plot of Site 27

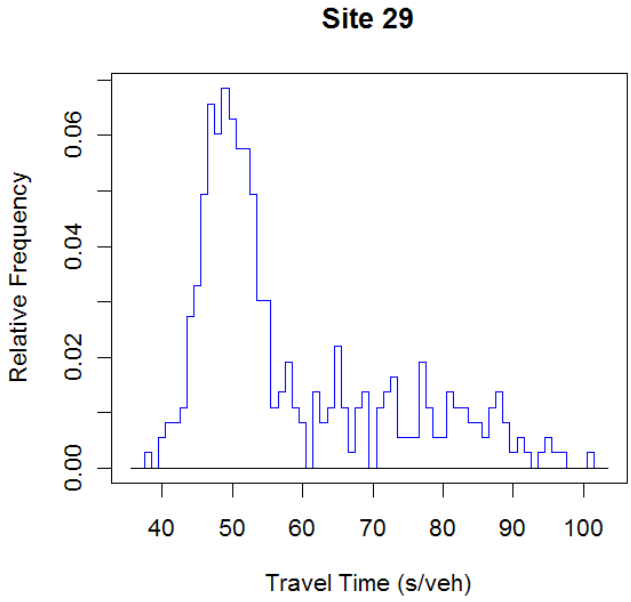

Figure A-9 Travel Time Relative Frequency Plot of Site 29

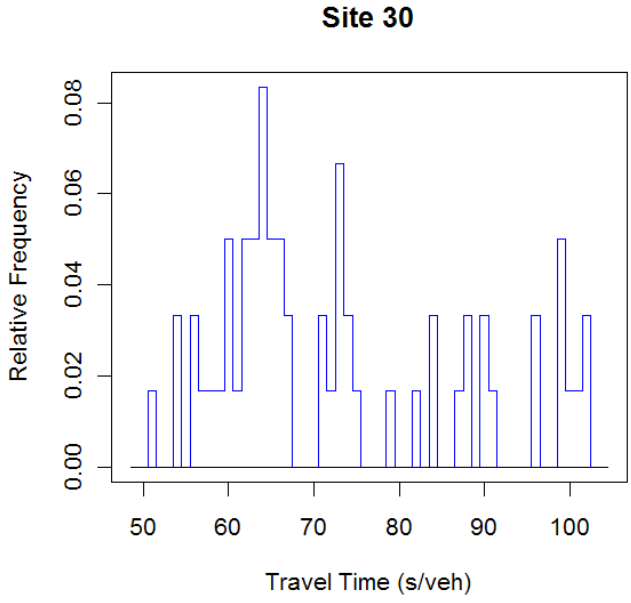

Figure A-10 Travel Time Relative Frequency Plot of Site 30

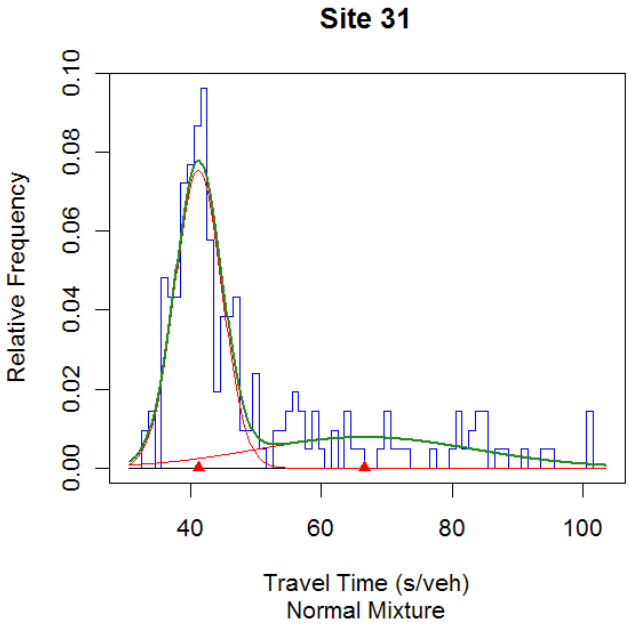

Figure A-11 Travel Time Relative Frequency Plot of Site 31

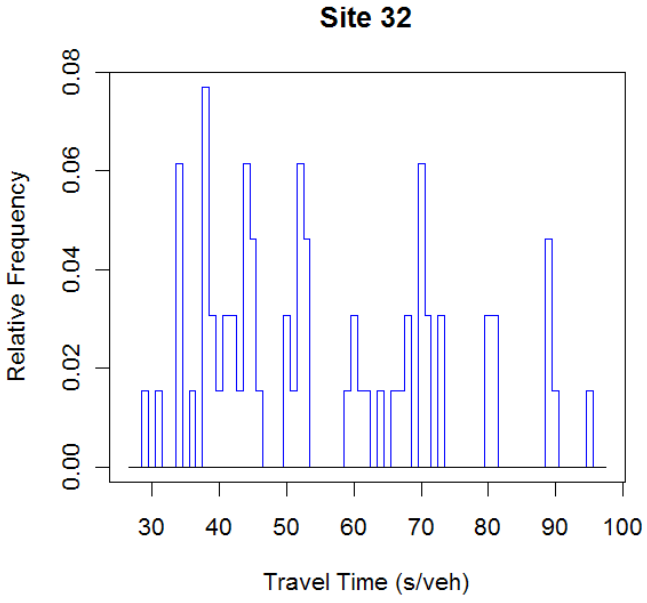

Figure A-12 Travel Time Relative Frequency Plot of Site 32

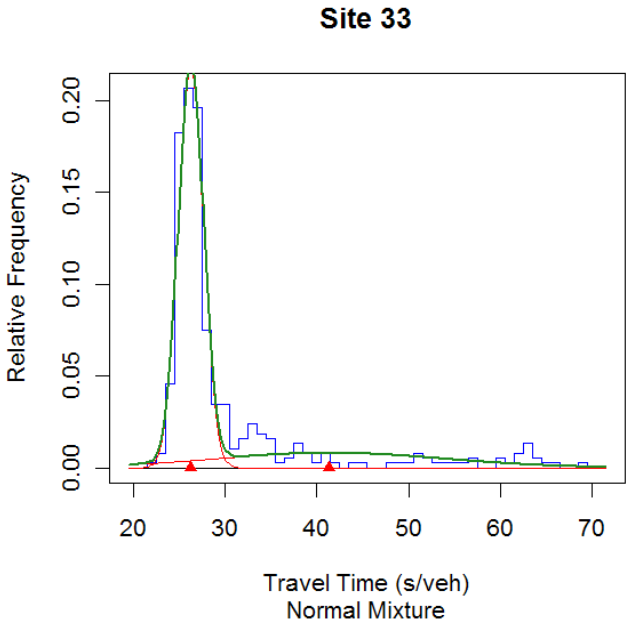

Figure A-13 Travel Time Relative Frequency Plot of Site 33

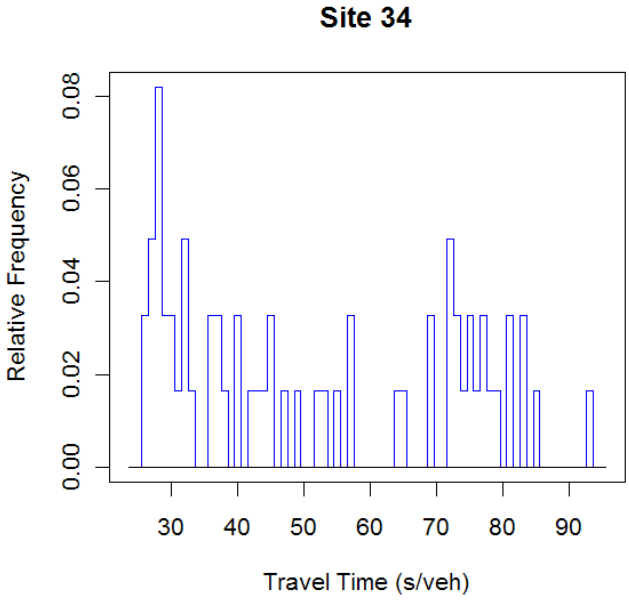

Figure A-14 Travel Time Relative Frequency Plot of Site 34

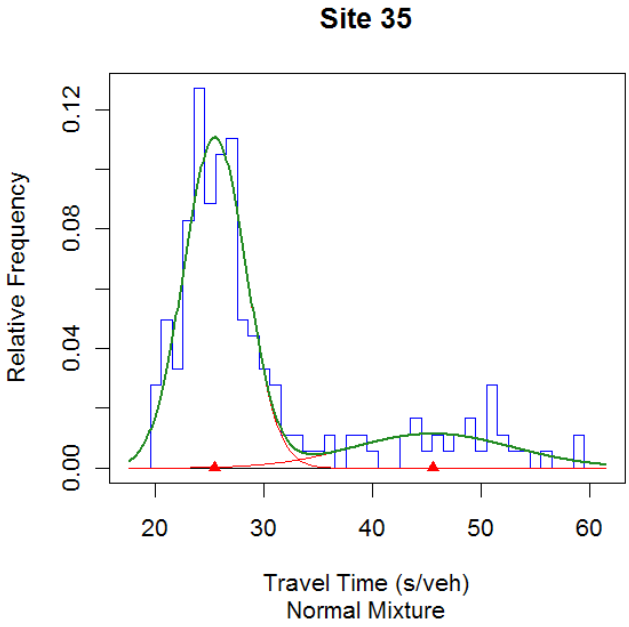

Figure A-15 Travel Time Relative Frequency Plot of Site 35

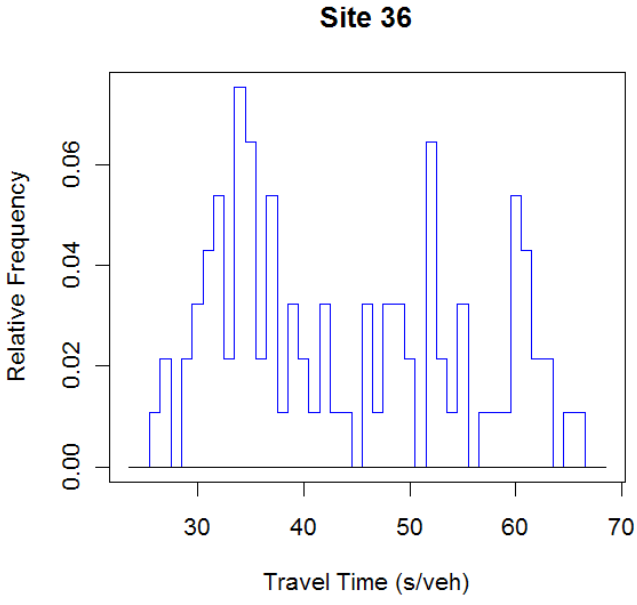

Figure A-16 Travel Time Relative Frequency Plot of Site 36

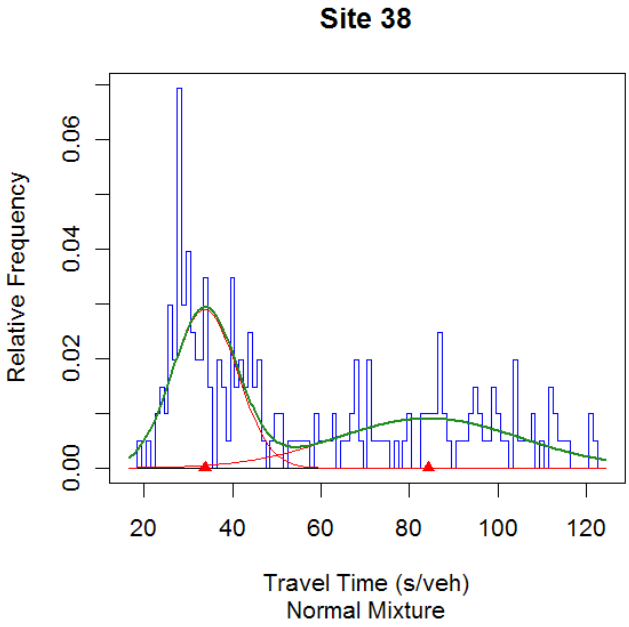

Figure A-17 Travel Time Relative Frequency Plot of Site 38

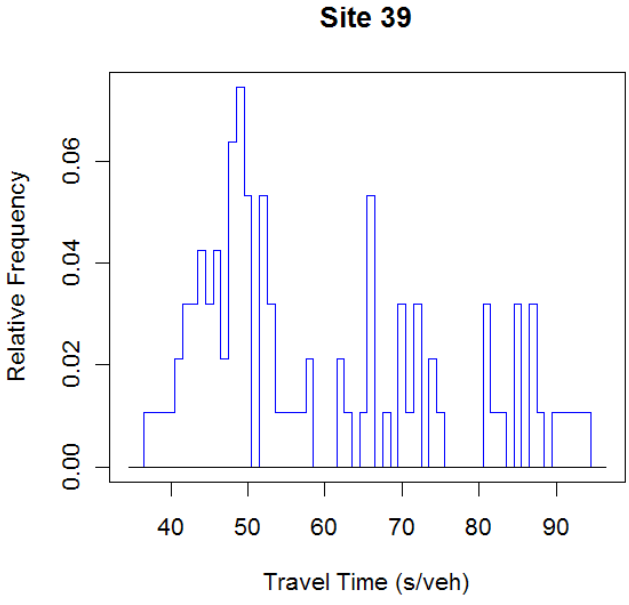

Figure A-18 Travel Time Relative Frequency Plot of Site 39

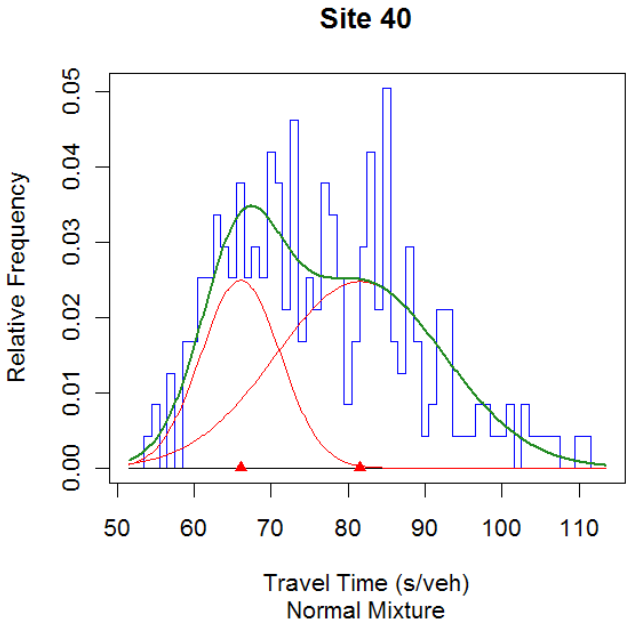

Figure A-19 Travel Time Relative Frequency Plot of Site 40

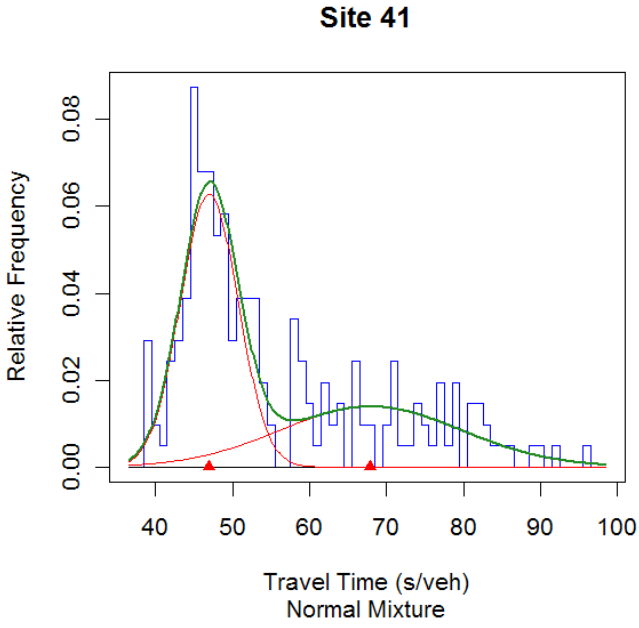

Figure A-20 Travel Time Relative Frequency Plot of Site 41

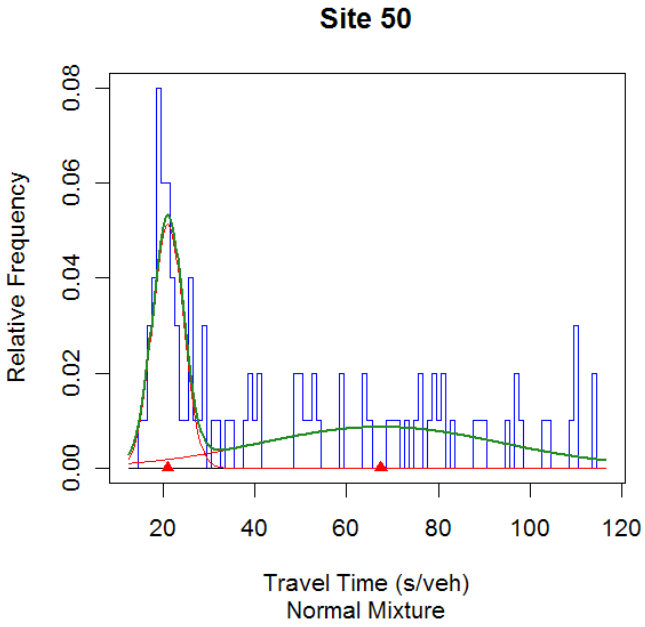

Figure A-21 Travel Time Relative Frequency Plot of Site 50

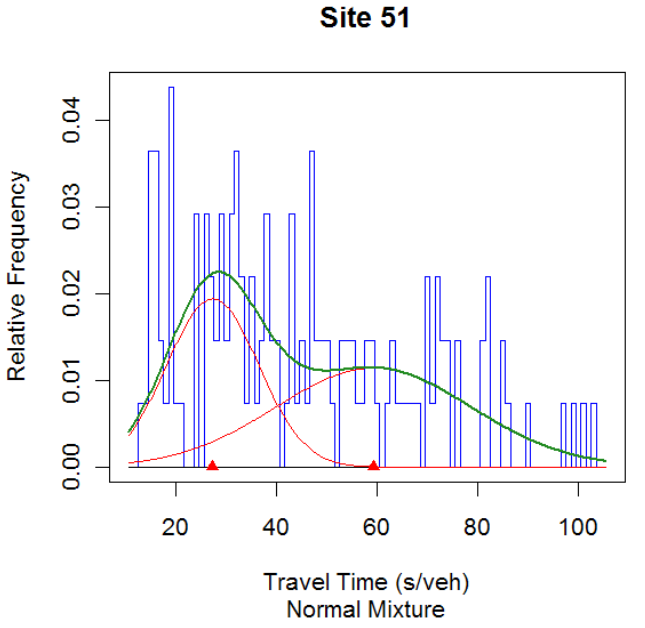

Figure A-22 Travel Time Relative Frequency Plot of Site 51

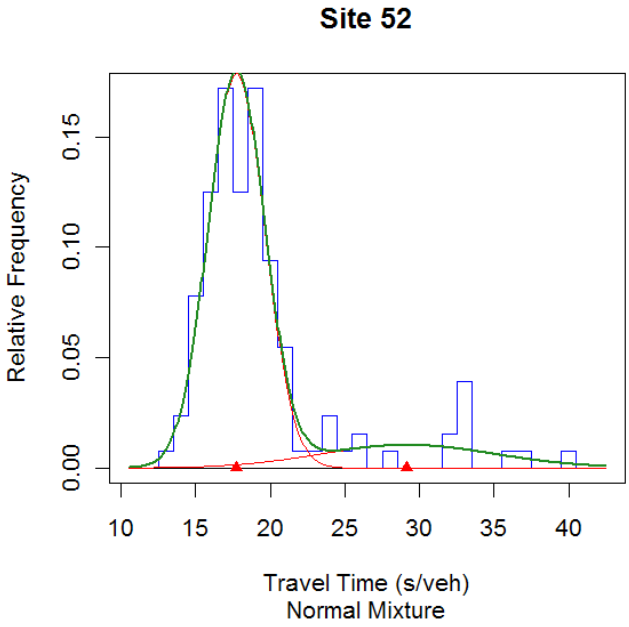

Figure A-23 Travel Time Relative Frequency Plot of Site 52

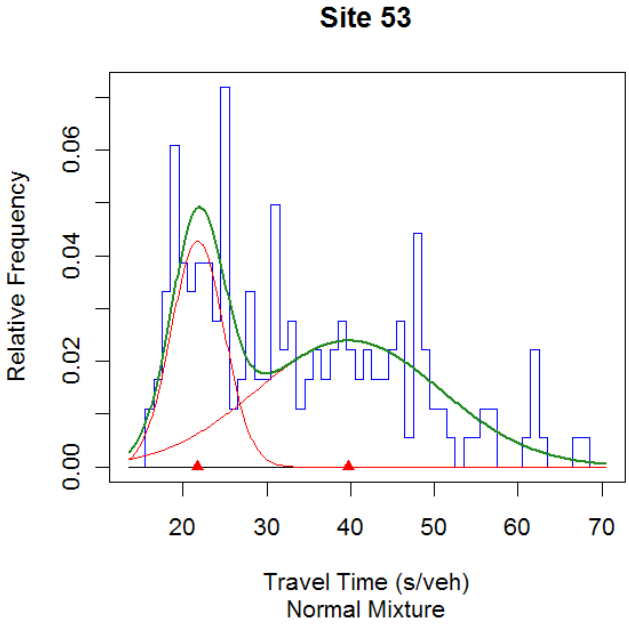

Figure A-24 Travel Time Relative Frequency Plot of Site 53

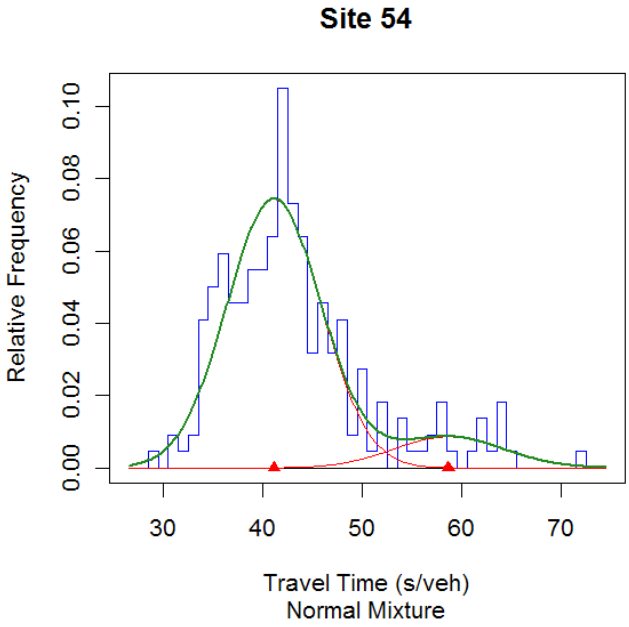

Figure A-25 Travel Time Relative Frequency Plot of Site 54

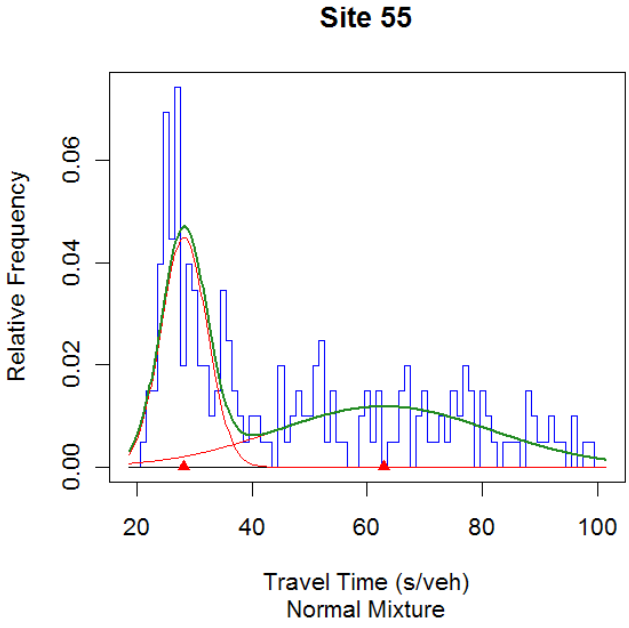

Figure A-26 Travel Time Relative Frequency Plot of Site 55

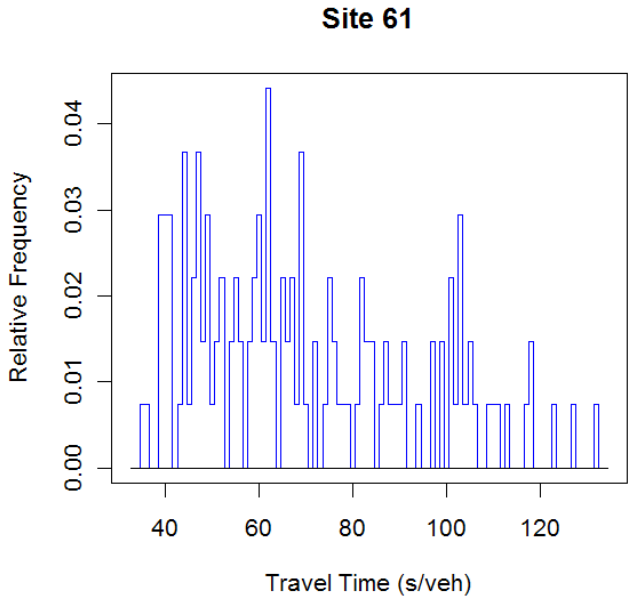

Figure A-27 Travel Time Relative Frequency Plot of Site 61

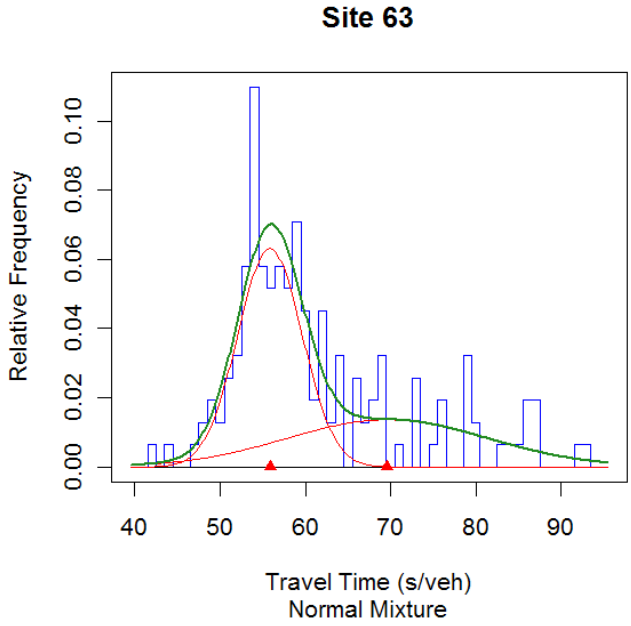

Figure A-28 Travel Time Relative Frequency Plot of Site 63

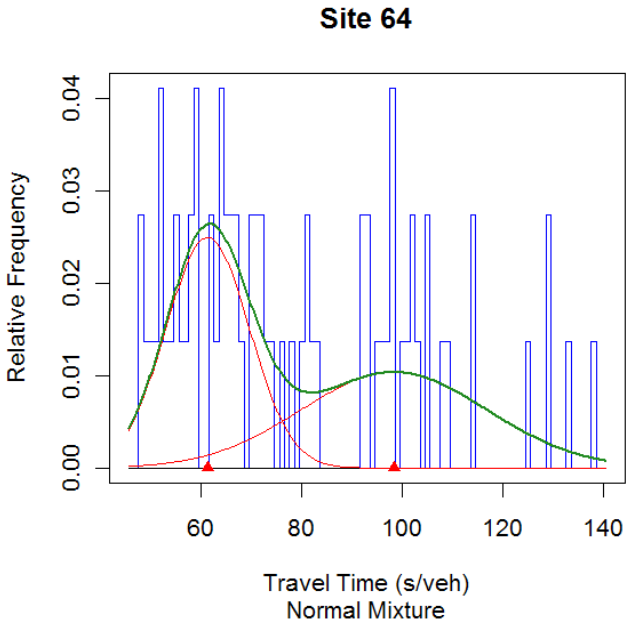

Figure A-29 Travel Time Relative Frequency Plot of Site 64

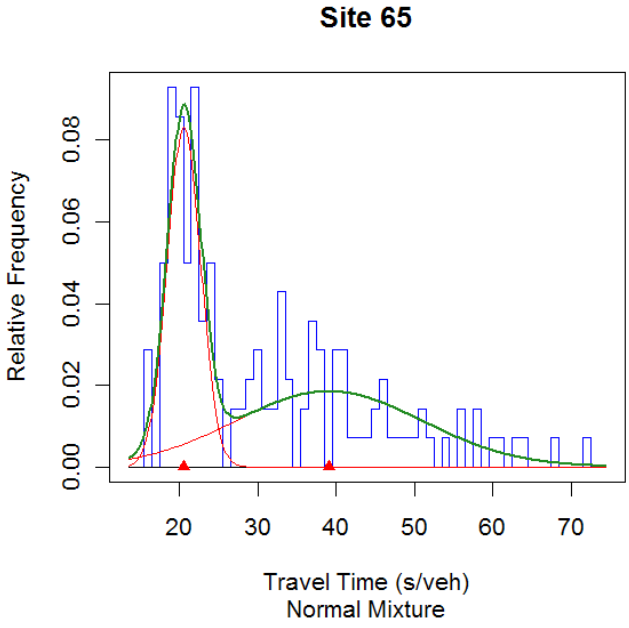

Figure A-30 Travel Time Relative Frequency Plot of Site 65

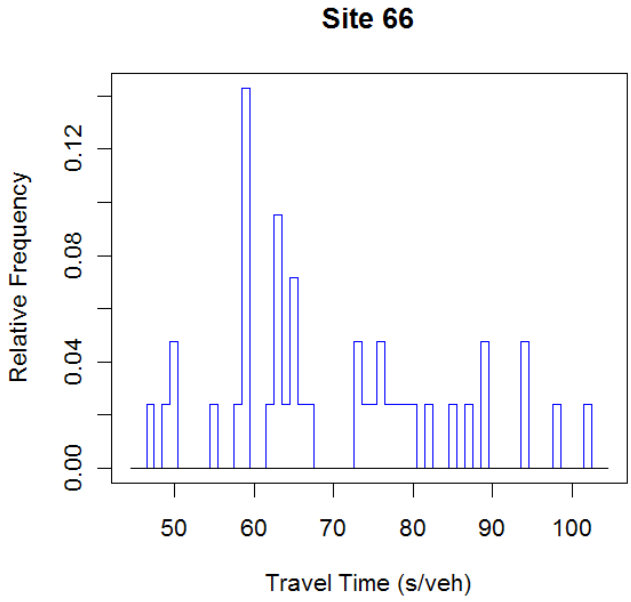

Figure A-31 Travel Time Relative Frequency Plot of Site 66

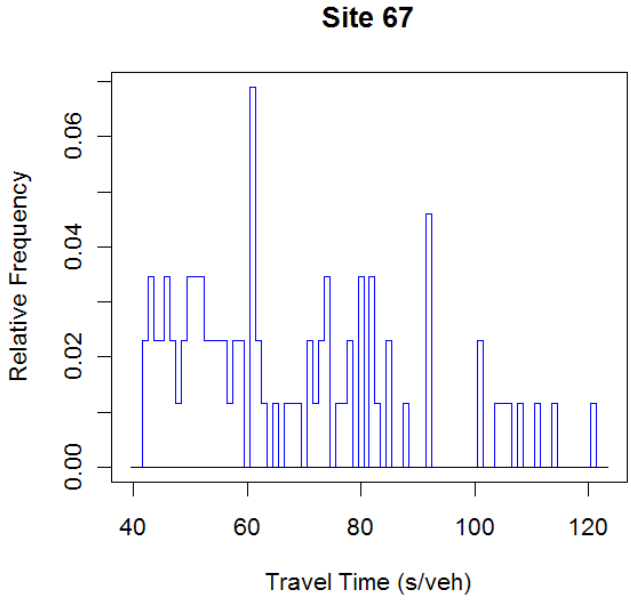

Figure A-32 Travel Time Relative Frequency Plot of Site 67

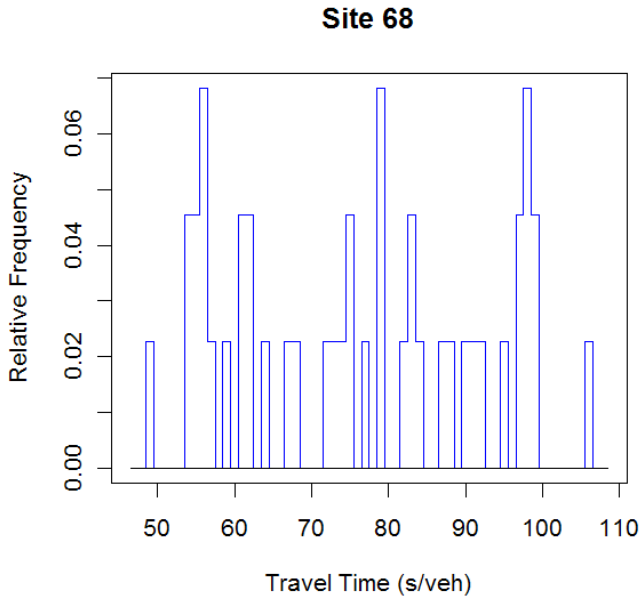

Figure A-33 Travel Time Relative Frequency Plot of Site 68

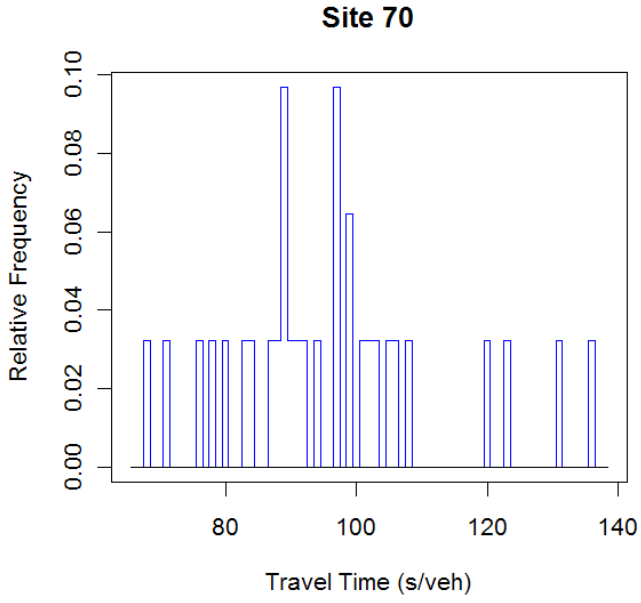

Figure A-34 Travel Time Relative Frequency Plot of Site 70

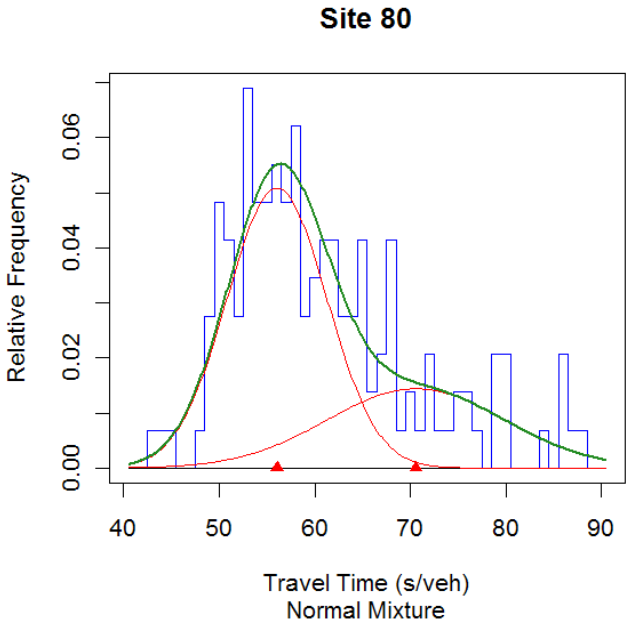

Figure A-35 Travel Time Relative Frequency Plot of Site 80

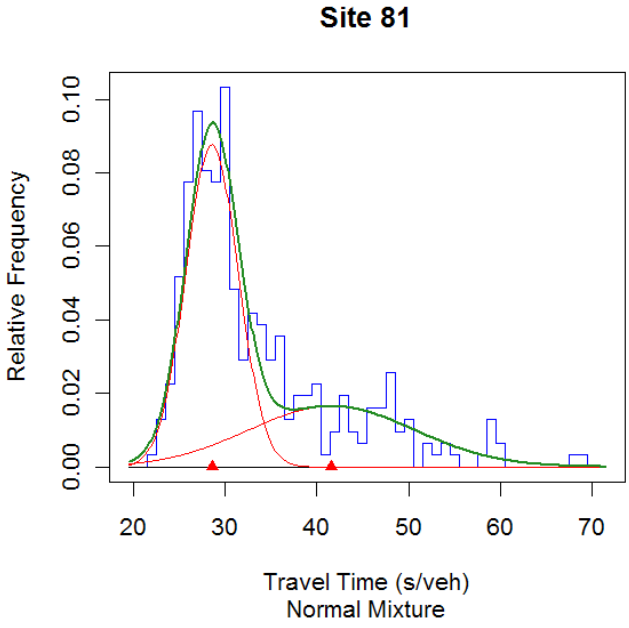

Figure A-36 Travel Time Relative Frequency Plot of Site 81

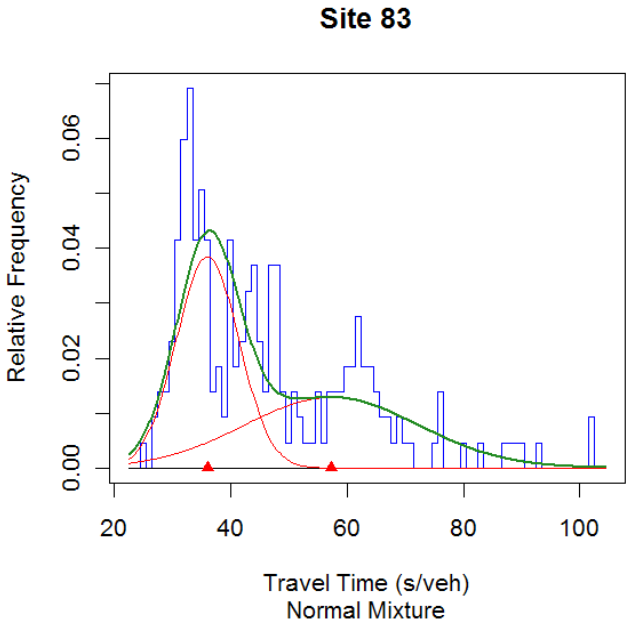

Figure A-37 Travel Time Relative Frequency Plot of Site 83

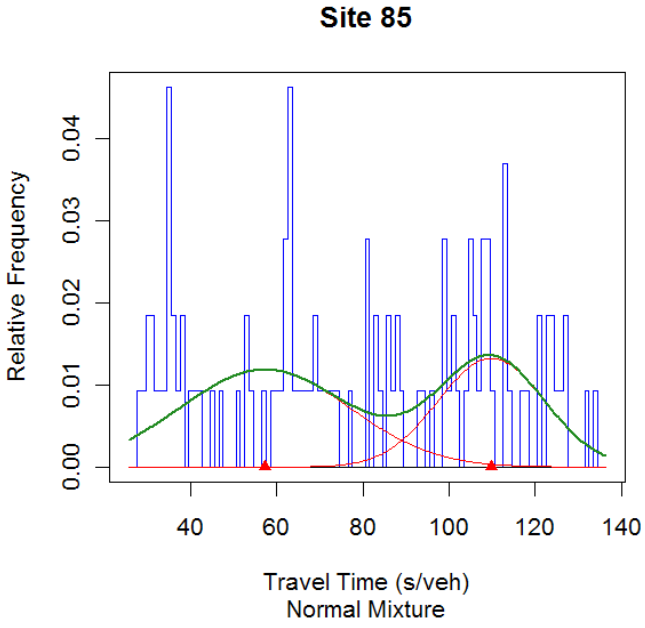

Figure A-38 Travel Time Relative Frequency Plot of Site 85

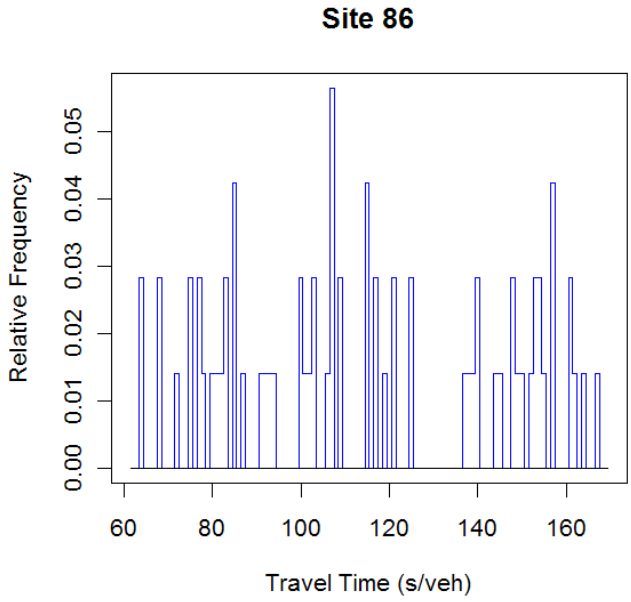

Figure A-39 Travel Time Relative Frequency Plot of Site 86

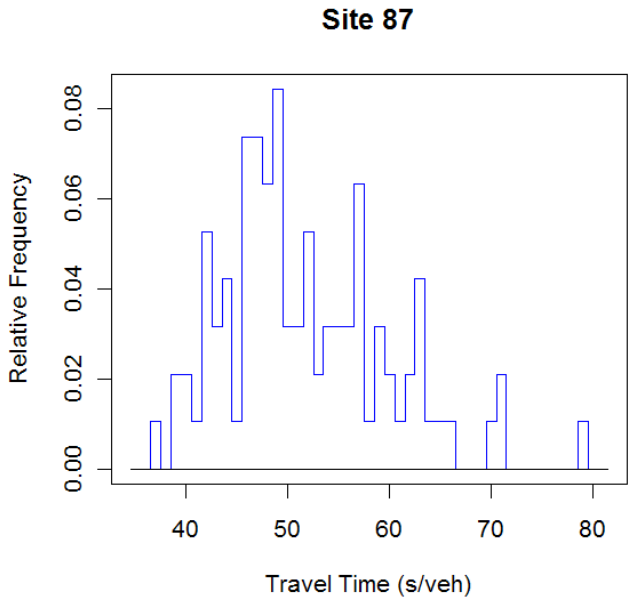

Figure A-40 Travel Time Relative Frequency Plot of Site 87

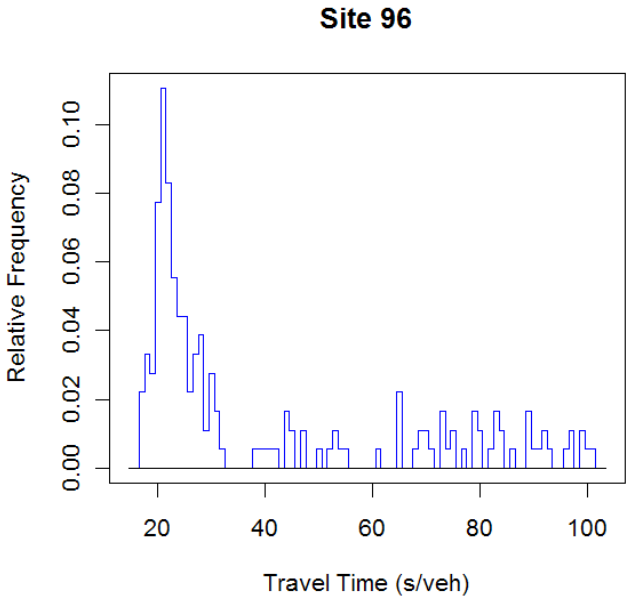

Figure A-41 Travel Time Relative Frequency Plot of Site 96

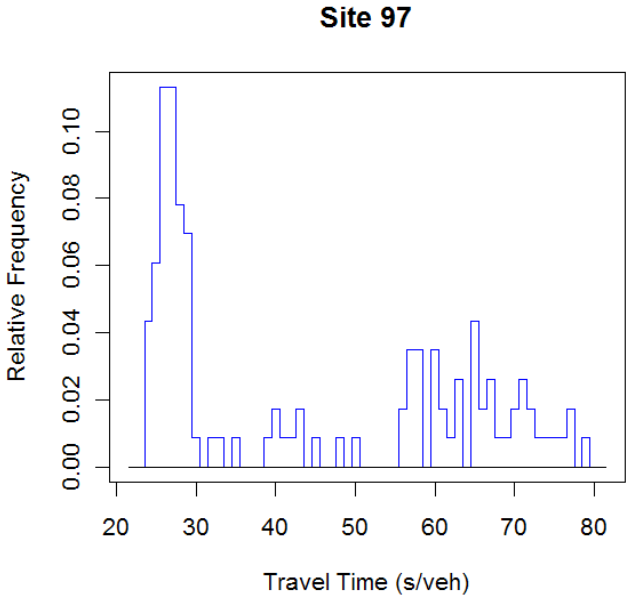

Figure A-42 Travel Time Relative Frequency Plot of Site 97

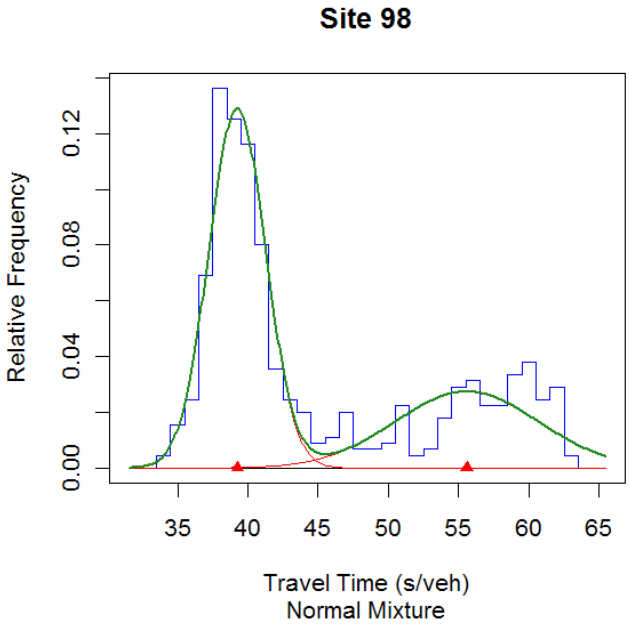

Figure A-43 Travel Time Relative Frequency Plot of Site 98

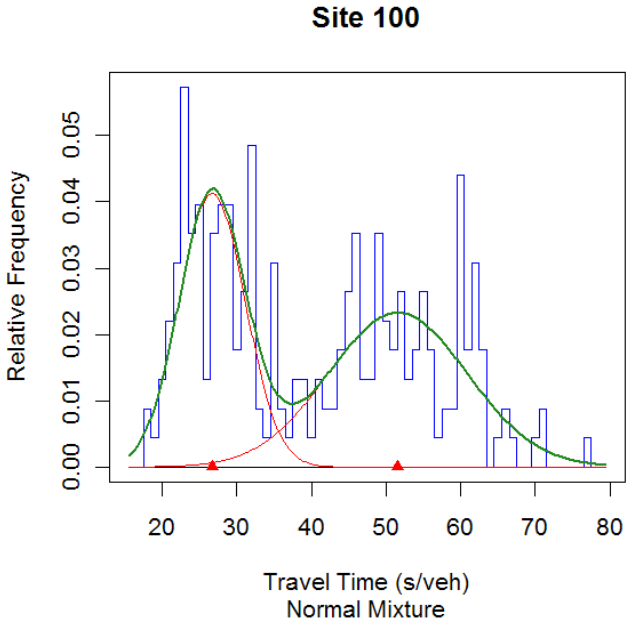

Figure A-44 Travel Time Relative Frequency Plot of Site 100

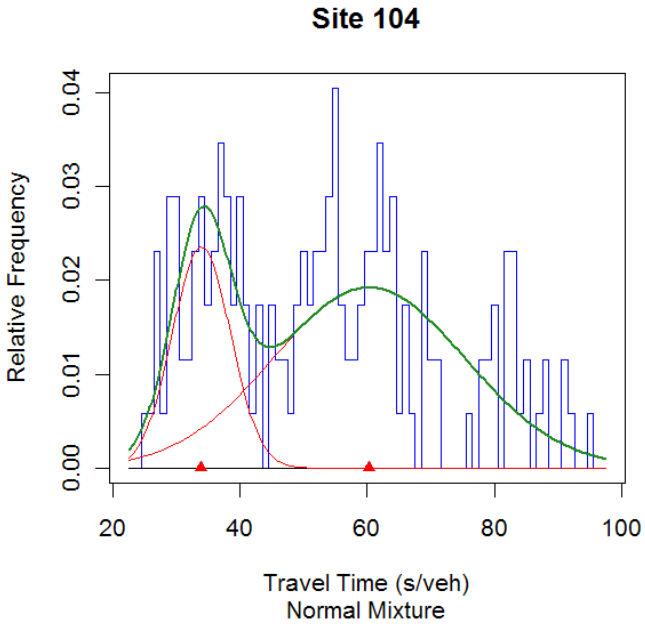

Figure A-45 Travel Time Relative Frequency Plot of Site 104

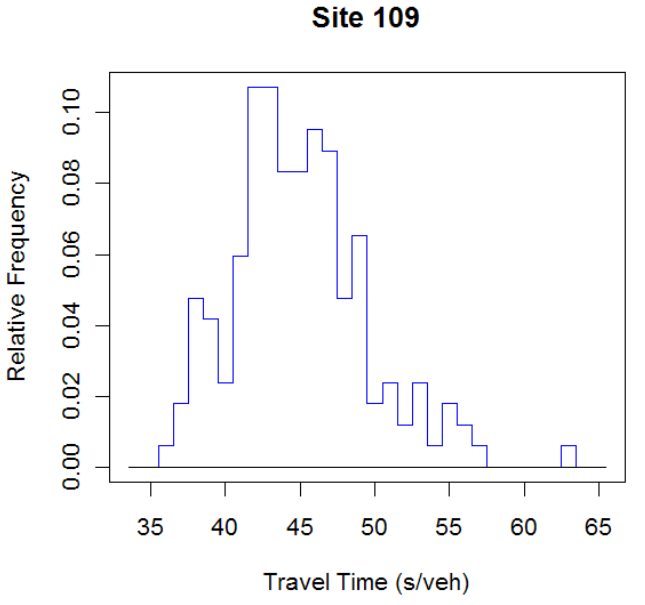

Figure A-46 Travel Time Relative Frequency Plot of Site 109

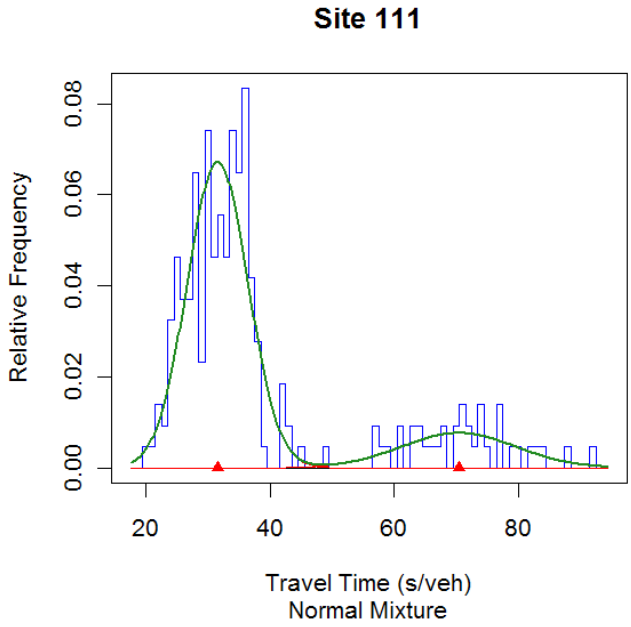

Figure A-47 Travel Time Relative Frequency Plot of Site 111

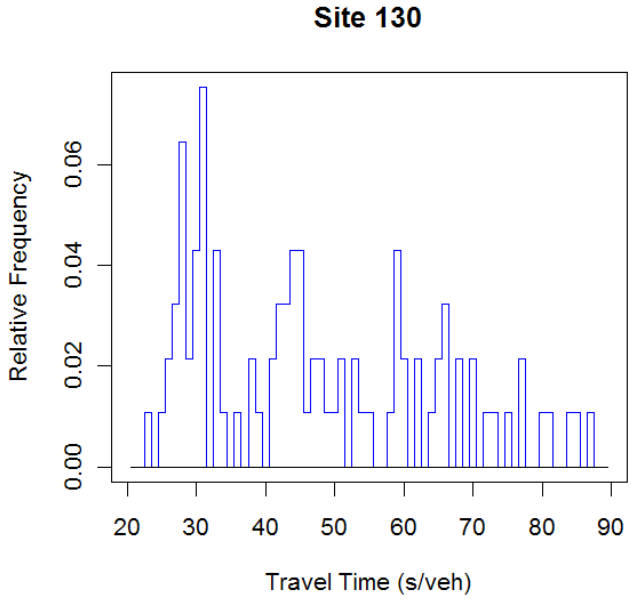

Figure A-48 Travel Time Relative Frequency Plot of Site 130

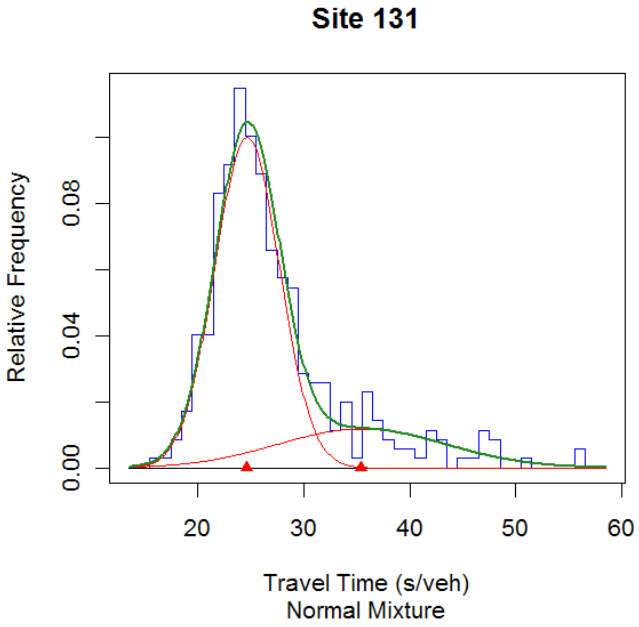

Figure A-49 Travel Time Relative Frequency Plot of Site 131

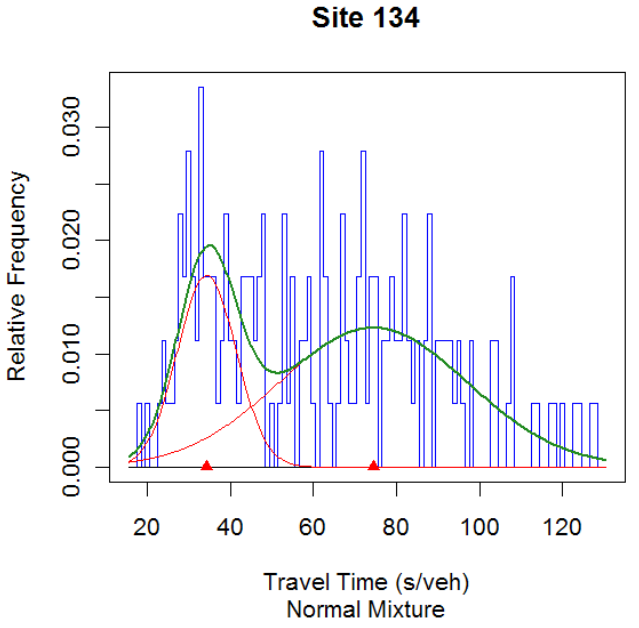

Figure A-50 Travel Time Relative Frequency Plot of Site 134

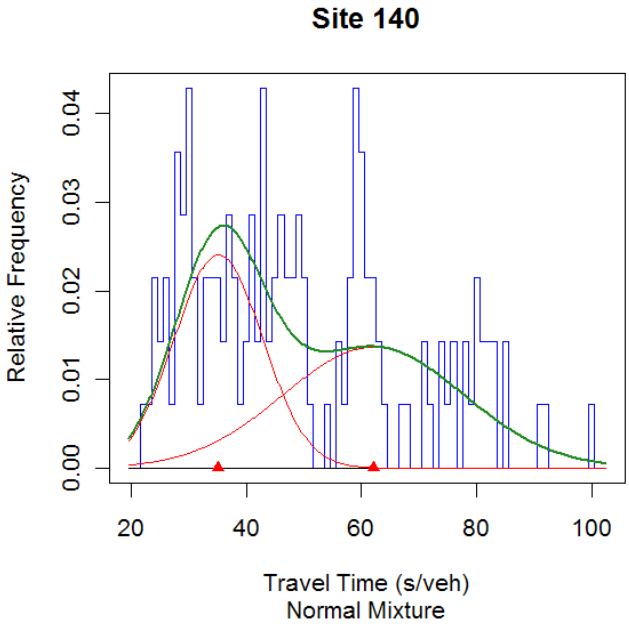

Figure A-51 Travel Time Relative Frequency Plot of Site 140

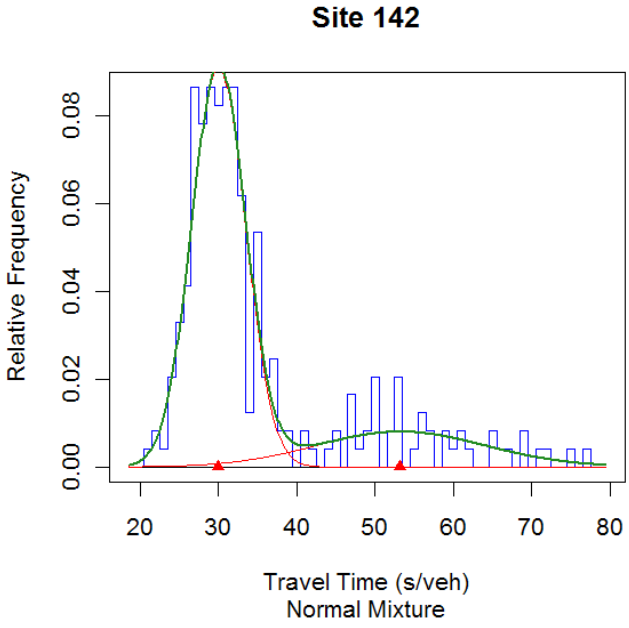

Figure A-52 Travel Time Relative Frequency Plot of Site 142

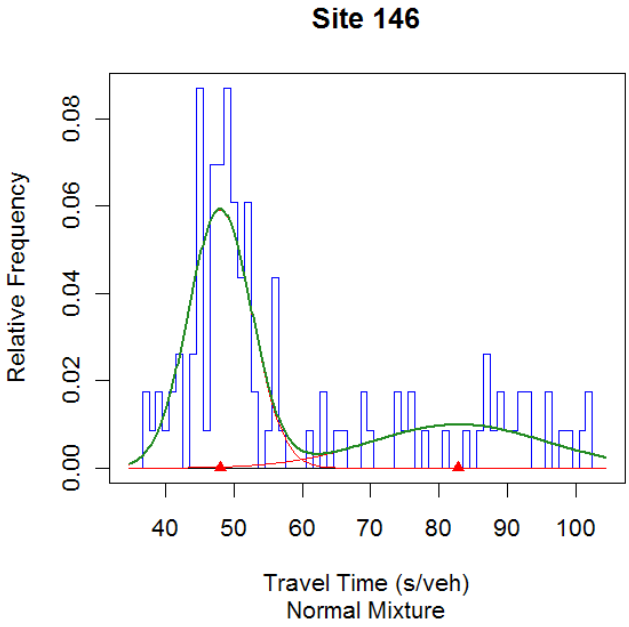

Figure A-53 Travel Time Relative Frequency Plot of Site 146

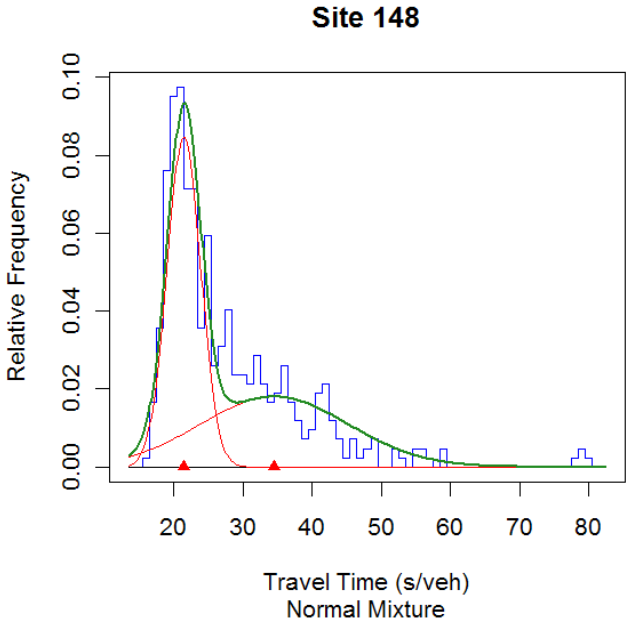

Figure A-54 Travel Time Relative Frequency Plot of Site 148

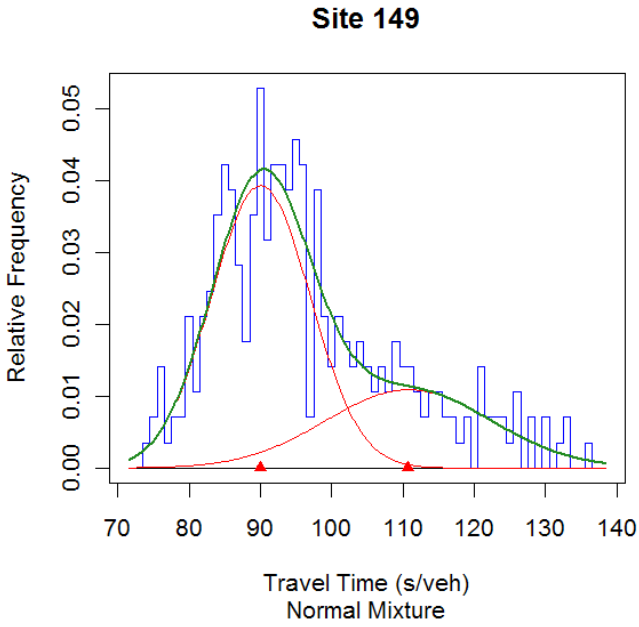

Figure A-55 Travel Time Relative Frequency Plot of Site 149

Appendix B. R Code for Travel Time Data Processing

```
# load mixdist package
library(mixdist)
# load site ID
Site_ID<-as.matrix(read.table("Site_ID.txt"))
# load plot titles
Main<-as.matrix(read.table("Main.txt"))
# load counts of total and non-stopping vehicles
count<-as.matrix(read.table(file="count.txt",header=T))
# load initial values for mu1, mu2
ini<-as.matrix(read.table(file="Mu1_Mu2.txt",header=T))
# load plot names for JPG output
JPG<-as.matrix(read.table(file="JPG.txt"))
# calculate the proportions of non-stopping vehicles
p_n<-count[,2]/count[,1]
# define MEAN as the mean travel time
MEAN<-rep(0,55)
# run estimate loop for all sites
for (w in 1:55)
{
# load matched data of one site
matched<-as.matrix(read.table(file=Site_ID[w],header=TRUE))
# generate bin intervals
breaks<-c(0,seq(min(matched)-0.5,max(matched)+0.5,1),max(matched)+1)
# generate frequency data
H_m<-hist(matched, breaks = breaks, plot = FALSE)
# generate grouped data
data_m<-as.mixdata(data.frame(X = c(H_m$breaks[c(-1, -
length(H_m$breaks))], Inf), count = H_m$counts))
# Two cases: (1) p_n=0, not apply EM; (2) p_n<>0, apply EM
if (p_n[w]==0)
{
# estimate mean travel time by taking the average
MEAN[w]<-round(mean(matched),2)
# define output file format
png(JPG[w], height=580, width=600, pointsize=16)
# plot histogram
plot(data_m, main=c(Main[w]),xlab="Travel Time (s/veh)")
dev.off()} else
{
# fit matched data by EM
fit_m<-mix(data_m,mixparam(c(ini[w,1],ini[w,2]),20),"norm")
# define output file format
png(JPG[w], height=580, width=600, pointsize=16)
# plot histogram and fitting curves
plot(fit_m, main=c(Main[w]), xlab="Travel Time (s/veh)")
dev.off()
# get estimates of mu1 and mu2 from EM
mu1<-fit_m$parameters[1,2]
mu2<-fit_m$parameters[2,2]
# estimate mean travel time by adjusted proportion of non-stopping
vehicles
MEAN[w]<-p_n[w]*mu1+(1-p_n[w])*mu2
}
```

```
}
# summarize estimation results
ID<-read.table("ID.txt")
summary<-cbind(ID, count, p_n, MEAN)
write.csv(summary,"summary.csv", row.names=FALSE, col.names=TRUE)
```