

**Development of Queue Length
Models at Two-way STOP
Controlled Intersection:
*A Surrogate Method***

Facility Analysis and Simulation Team
Transportation Development Division
Oregon Department of Transportation

October 2010

Table of Contents

TABLE OF CONTENTS	II
TABLE OF FIGURES	IV
TABLE OF TABLES	V
EXECUTIVE SUMMARY.....	1
1 INTRODUCTION	3
1.1 METHODOLOGIES	3
1.1.1 <i>Two-Minute Rule</i>	3
1.1.2 <i>Harmelink Curves</i>	4
1.1.3 <i>Highway Capacity Manual</i>	5
1.1.4 <i>Gard's Equation</i>	6
1.2 PROBLEM	7
1.3 PURPOSE	7
1.4 STUDY METHODOLOGY.....	7
1.5 SUMMARY	8
2 DATA COLLECTION & ANALYSIS	9
2.1 DATA COLLECTION.....	9
2.2 DATA ANALYSIS	9
2.3 SUMMARY	11
3 COMPARISON OF EXISTING METHODOLOGIES.....	12
3.1 MAJOR LEFT (MJL)	12
3.2 MINOR, SHARE LTR (MNLTR)	14
3.3 MNLR	16
3.4 MNL	19
3.5 SUMMARY	21
4 MODEL DEVELOPMENT	22
4.1 FACTORS INFLUENCING QUEUE BEHAVIOR.....	22
4.2 POISSON REGRESSION MODEL	24
4.3 MAJOR LEFT TURN (MJL) MODEL.....	24
4.4 MINOR, SHARED LTR (MNLTR).....	31
4.5 MINOR, LR (MNLR).....	38
4.6 MINOR LEFT (MNL)	44
4.7 MINOR RIGHT (MNR).....	49
4.8 SUMMARY	50
5 MODEL VALIDATION	51
5.1 DATA COLLECTION	51
5.2 MJL	52
5.3 MNLTR.....	52
5.4 MNLR	53
5.5 MNL	54
5.6 SUMMARY	54
6 MODEL SENSITIVITY ANALYSIS	55
6.1 MJL	55
6.2 MNLTR.....	56
6.3 MNLR	57

6.4	MNL.....	58
6.5	SUMMARY.....	59
7	SUMMARY, CONCLUSIONS & SCOPE FOR FUTURE STUDY	60
7.1	SUMMARY.....	60
7.2	CONCLUSIONS.....	60
7.3	FUTURE STUDY SCOPE.....	61
	APPENDIX A - HCM & MUTCD EXHIBITS.....	62
	APPENDIX B - DATA COLLECTION SITES FOR MODEL DEVELOPMENT	66
	APPENDIX C - DATA COLLECTION SITES FOR MODEL VALIDATION	69

Table of Figures

Figure 1-1 Expected maximum number of vehicles in Queue by HCM.....	5
Figure 1-2 Study Methodology.....	8
Figure 3-1 Scatter plot of all methods for MJL	13
Figure 3-2 Scattered plot of VOL, CONVOL, & Queue Lengths	15
Figure 3-3 Scatter plot of all methods for MNLTR	15
Figure 3-4 Scattered plot of VOL, CONVOL, & Queue Lengths for MNLTR.....	17
Figure 3-5 Scatter plots of all methods for MNLR.....	18
Figure 3-6 Scattered plot of VOL, CONVOL, & Queue Lengths for MNL.....	20
Figure 3-7 Scatter plot of all methods for MNL	20
Figure 4-1 Scattered matrix plot between the variables for MJL	25
Figure 4-2 Plot of Fitted Model for MJL	30
Figure 4-3 Residual Plot for MJL	30
Figure 4-4 Scattered Plot of QL Vs VOL, CONVOL, and VOL*CONVOL for MNLTR...	31
Figure 4-5 Plot of Predicted QL and Observed QL for MNLTR Model.....	37
Figure 4-6 Residual Plot of Predicted QL and Observed QL for MNLTR Model .	37
Figure 4-7 Scattered matrix plot for MNLR lane group data	39
Figure 4-8 Plot of QL between Predicted Vs Observed QL for MNLR.....	43
Figure 4-9 Residual Plot between Predicted QL and Residual for MNLR.....	43
Figure 4-10 Scattered Matrix Plot of Variables for MNLR.....	44
Figure 4-11 Plot of QL between Predicted Vs Observed QL for MNL	48
Figure 4-12 Residual Plot between Predicted QL and Residual for MNL.....	48
Figure 6-1 Plot of QL for VOL and CONVOL at Major Left Turn (MJL)	56
Figure 6-2 Plot of QL for VOL and CONVOL at Minor LTR (MNLTR)	57
Figure 6-3 Plot of QL for VOL and CONVOL at Minor LR (MNLR).....	58
Figure 6-4 Plot of QL for VOL and CONVOL at Minor Left (MNL).....	59
Figure A1 Traffic Streams at a Two-Way STOP Controlled Intersection (Source: HCM, Exhibit 17-3).....	62
Figure A2 Method for Computing Lane Group Conflicting Flow Rates (Source: HCM, Exhibit 17-4).....	63
Figure A3 Lane Groups (Source: HCM, Exhibit 16-5)	64

Table of Tables

Table 1-1 Selection of “t” values (<i>source: APM</i>)	4
Table 1- 2 Storage Length Adjustments for Trucks (<i>source: APM</i>)	4
Table 1-3 Description of Gard’s Study Intersections ⁵	6
Table 1-4 Gard’s Regression Equations ⁵	6
Table 3-1 Summary Statistics for MJL.....	12
Table 3-2 Difference between Observed and Model Outputs for MJL	13
Table 3-3 Summary Statistics for MJL.....	14
Table 3-4 Difference between Observed and Model Outputs for MNLTR	16
Table 3-5 Summary Statistics for MNLTR.....	17
Table 3-6 Difference between Observed and Model Outputs for MNLTR.....	18
Table 3-7 Summary Statistics for MNL.....	19
Table 3-8 Difference between Observed and Model Outputs for MNL	21
Table 4-1 Factors Influencing Queue Behavior	23
Table 4-2 Summary Statistics of Major Left Turn Data	24
Table 4-3 Selection of Independent Variable for MJL Model	26
Table 4-4 Spearman Rank Correlation Matrix.....	26
Table 4-5 Statistical Analysis of MJL Model	28
Table 4-6 Correlation Matrix for Estimated Coefficients for MJL Model	29
Table 4-7 Fitting Distributions for MNLTR Data.....	32
Table 4-8 Summary Statistics for MNLTR data.....	33
Table 4-9 Selection of Dependent Variables for Poisson Regression	34
Table 4-10 Correlation Matrix of Variables for MNLTR Model.....	35
Table 4-11 Poisson Regression Model for MNLTR.....	35
Table 4-12 Summary Statistics for MNLTR lane group.....	38
Table 4-14 Pearson Correlation Matrix of Variables in MNLTR Model.....	39
Table 4-13 Variables Selection for MNLTR Model.....	40
Table 4-15 Poisson Regression Model for MNLTR Data.....	41
Table 4-16 Summary Statistics of MNL Data.....	44
Table 4-17 Selection of Independent Variables for MNL Model	45
Table 4-18 Spearman Rank Correlation Matrix for MNL Variables.....	46
Table 4-19 Estimated Regression Model for MNL Data	47
Table 4-20 Summary Statistics of MNL Data.....	49
Table 5-1 Comparison of Queue Length Estimation Differences for MJL.....	52
Table 5-2 Comparison of Queue Length Estimation Differences for MNLTR.....	53
Table 5-3 Comparison of Queue Length Estimation Differences for MNLTR	53
Table 5-4 Comparison of Queue Length Estimation Differences for MNLTR	54
Table A1 MUTCD Warrant 1 Table (Source: 2009 MUTCD, section 4C.02)	65
Table B1 Description of Study Intersection Used for Model Development.....	67
Table B2 Periods of Data Collection Used for Model Development	68
Table C1 Description of Study Intersections Used for Data Validation	70
Table C2 Duration of Traffic Counts Used for Data Validation	71

Executive Summary

This study aims at developing queue length models at two-way STOP controlled intersections. A significant amount of research on the estimation of capacity, delay, and queue lengths at unsignalized intersection has resulted in a variety of models ranging from empirical to simulation models. Most agencies are following methods like the Two-Minute Rule, Highway Capacity Manual Method, and the Harmelink Curves to estimate queue lengths. But, these methods are yielding inconsistency estimates and questions often arise as to the correct method to use. This study documents the inconsistency among these methods and takes further steps to improve queue length estimates by developing surrogate models.

Data at 15 two-way STOP controlled intersections covering various functional classifications of highways, geometric configurations, and geographic regions were collected by using video tapes. Data was processed to meet the requirements of the methods. Queue length estimations from each method were noted. Later, models were compared for their performance in estimating 95th percentile queue lengths. It was shown that the Highway Capacity Manual method consistently underestimates queues. The two-minute rule estimated fairly closer queue lengths except for major left turn movements, due to not considering opposing volumes. The Harmelink curves are applicable only for major left turns. Queue length estimation equations developed by John T. Gard showed better results, but the variation among observed and estimated queue lengths was still high.

Data processed for comparison was used to estimate the models. First, looking at the data clearly indicated that random phenomena prevail among queue lengths and the associated explanatory variables. Exhaustive statistical analysis was conducted to understand queue behavior on both major and minor approach lane groups. Poisson regression models were fitted to explain the random process. A model comparison showed significantly improved performance of the new models in predicting maximum queue lengths.

Further, field data at 25 intersections was collected covering wide array of condition to test the developed model consistency. More than 70 % of the predicted queue lengths were close to observed queue length estimates. In addition, model sensitivity analysis was performed to assess the stability of the model. When the major and minor approach volumes are within limits of MUTCD signal warrant volumes, acceptable ranges of queue lengths are predicted. Beyond the MUTCD suggested volume ranges, marginal increase in input variable substantiates queue lengths.

This report is organized in seven chapters. Various methods are reviewed in Chapter 1. Problems identified in each method, objectives, scope and methodology are discussed. Chapter 2 is dedicated to data collection and analysis efforts. Chapter 3 compares the methods and highlights their differences. Chapter 4 explains the basic philosophy of developing Poisson regression models. Detailed statistical analysis, including data description, model selection, variable selection and model statistics are included. Data validation steps are presented in chapter 5. Chapter 6 deals with the stability analysis of the model through sensitivity tests. Chapter 7 gives the summary, conclusions, and scope for future study.

1 Introduction

A significant amount of research on the estimation of capacity, delay, and queue lengths at unsignalized intersection has resulted in the development of a variety of models ranging from empirical to simulation. A literature review of all available models is beyond the scope of this project, but unsignalized intersection theory chapter in the Traffic Flow Theory¹ gives a good start. In particular, the philosophies behind the methods like, Two-Minute Rule, Highway Capacity Manual Method, the Harmelink Curves, and equations given by John T. Gard are discussed.

This chapter introduces the above mentioned methods briefly. Next, problems associated with each method are explained. This leads to the definition of the problem. Finally, the step-by-step process used for the study is presented.

1.1 Methodologies

1.1.1 Two-Minute Rule²

The Two-Minute Rule is a rule of thumb methodology that estimates queue lengths for major street left turns and minor street movements by using the queue that would result from a two-minute stoppage of the turning demand volume. This method does not consider the magnitudes and impacts of the conflicting flows on the size of the queue. The calculation of the 95th percentile queue using the two-minute rule methodology shall use the following equation:

$$S = (v) (t) (L)$$

Where:

S = the 95th percentile queue storage length (feet)

v = the average left-turn volume arriving in a 2-minute interval

t = a variable representing the ability to store all vehicles; usually 1.75 to 2.0 (Use Table 1-1)

L = average length of the vehicles being stored and the gap between vehicles; 25 ft. for cars. This value can be increased where a significant number of trucks are present

¹ R.J.Troutbeck., and Brilon, W. “Unsignalised Intersection Theory”, Traffic Flow Theory, TRB Special Report 165, Washington D.C.

² Chapter 7: Intersection Analysis, “Analysis Procedure Manual”, Updated: May 2010, Pg:237

Table 1-1 Selection of "t" values (source: APM)

Exhibit 7-29 Selection of "t" Values

Minimum "t" Value	Percentile
2.0	98 %
1.85	95 %
1.75	90 %
1.0	50 %

It should also be noted that the value of 25-feet used in the equation represents the average storage length required for a passenger car. If a significant number of trucks are present in the turning volumes, the average storage length per vehicle should be increased according to Table 1-2. This adjustment is only for the manual methods; software packages may require a different adjustment.

Table 1- 2 Storage Length Adjustments for Trucks (source: APM)

Exhibit 7-30 Storage Length Adjustments for Trucks

Percent Trucks in Turning Volume	Average Vehicle Storage Length
< 2%	25 ft
5%	27 ft
10%	29 ft

While both the nomograph given in the Analysis Procedure Manual and the rule of thumb equation are intended for use in estimating vehicle queue lengths for single-lane left turn movements, the vehicle queue lengths for double left turn lanes can be estimated by dividing the results of these methods by 1.8. This value represents the assumption that queued vehicles will not be evenly distributed between the turn lanes.

1.1.2 Harmelink Curves³

M.D. Harmelink , in a paper that was published in 1967, provided the foundation for many current left-turn guidelines. Harmelink based his work on a queuing model in which arrival and service rates are assumed to follow negative exponential distributions. He stated that the probability of a through vehicle arriving behind a stopped, left-turning vehicle should not exceed 0.02 for 40 mph (64 km/h), 0.015 for 50 mph (80 km/h), and 0.01 for 60 mph (96 km/h). He presented his criteria in the form of graphs, 18 in all. To use his graphs, the advancing volume, opposing volume, operating speed, and left-turn percentage

³ M.D.Harmelink, "Volume Warrants for Left-Turn Storage Lanes at Unsignalized Grade Intersections", Highway Research Record 211, 1967

need to be known. Graphs for speeds of 40, 50, and 60 mph (64, 80, and 96 km/h) are given, as well as 5, 10, 15, 20, 30, and 40 percent left-turn volumes.

1.1.3 Highway Capacity Manual⁴

HCM 2000 relies on refined models developed in Germany based on both gap acceptance and empirical models which describe the interaction of the minor or stop controlled approach with drivers on the major street. The following figure shows the computational steps to calculate the queue lengths for two-way stop controlled approaches.

$$Q_{95} \approx 900T \left[\frac{v_x}{c_{m,x}} - 1 + \sqrt{\left(\frac{v_x}{c_{m,x}} - 1 \right)^2 + \frac{\left(\frac{3600}{c_{m,x}} \right) \left(\frac{v_x}{c_{m,x}} \right)}{150T}} \right] \left(\frac{c_{m,x}}{3600} \right) \quad (17-37)$$

where

- Q_{95} = 95th-percentile queue (veh),
- v_x = flow rate for movement x (veh/h),
- $c_{m,x}$ = capacity of movement x (veh/h), and
- T = analysis time period (h) ($T = 0.25$ for a 15-min period).

95th percentile queue lengths are calculated by the above equation 17-37 of the HCM. For varying volume-to-capacity ratios, expected maximum number of vehicles in queue are obtained from Figure 1-1.

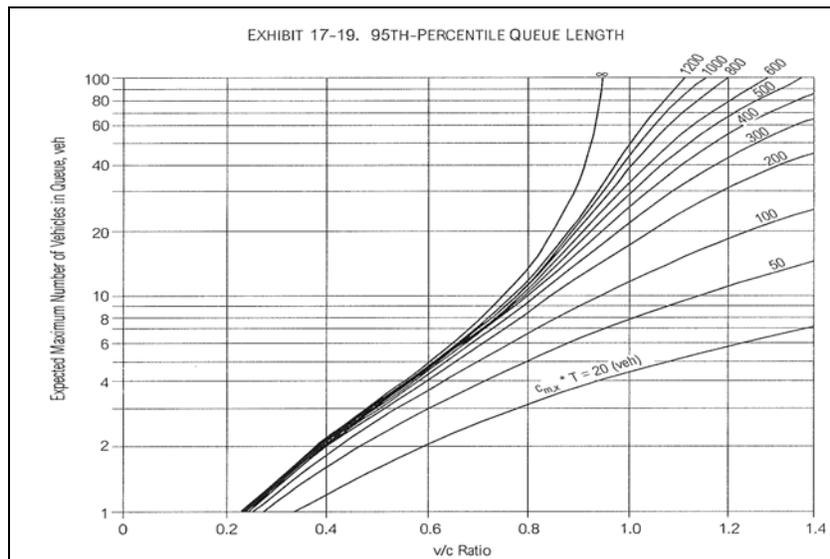


Figure 1-1 Expected maximum number of vehicles in Queue by HCM

⁴ "Chapter 17- Unsignalized intersections", 2000 Highway Capacity Manual, Transportation Research Board, Washington, D.C.

1.1.4 Gard's Equation⁵

John T. Gard developed regression equations for the prediction of queue lengths for major-street left turn, minor street left turn, minor street right turn, and minor street shared left/through/right turn configurations through a study of 15 unsignalized intersections in Sacramento, CA. Queue length represents the maximum number of vehicles in the queue. Table 1-3 describes the study intersections and Table 1-4 gives a summary of regression equations. R² values of the equations vary from 0.65 to 0.80.

Table 1-3 Description of Gard's Study Intersections⁵

Table 1. Description of study intersections.						
Intersection	Major street				Minor street	
	Roadway type	Posted speed limit	Traffic signal within one-quarter mile?	Two-way left-turn lane?	Roadway type	Lane configuration
1	2-lane collector	30 mph	Yes	No	2-lane collector	Exclusive left & right lanes
2	2-lane arterial	45 mph	Yes	Yes	2-lane collector	Shared left/right lane
3	2-lane arterial	55 mph	No	No	2-lane arterial	Shared left/right lane
4	2-lane arterial	50 mph	No	No	2-lane arterial	Exclusive left & right lanes
5	2-lane arterial	45 mph	No	No	2-lane collector	Shared left/right lane
6	2-lane highway	55 mph	No	No	2-lane collector	Shared left/right lane
7	2-lane highway	50 mph	No	No	2-lane arterial	Exclusive left & right lanes
8	4-lane arterial	45 mph	Yes	No	Hospital driveway	Exclusive left/through & right lanes
9	4-lane arterial	45 mph	Yes	No	Shopping center driveway	Shared left/right lane
10	4-lane arterial	40 mph	Yes	Yes	Office complex driveway	Exclusive left/through & right lanes
11	4-lane arterial	45 mph	No	Yes	2-lane collector	Exclusive left & right lanes
12	4-lane arterial	45 mph	Yes	Yes	Shopping center driveway	Exclusive left & right lanes
13	4-lane arterial	45 mph	Yes	Yes	2-lane collector	Shared left/through & exclusive right lanes
14	6-lane arterial	35 mph	Yes	No	Shopping center driveway	Exclusive right lane
15	6-lane arterial	45 mph	Yes	No	Shopping center driveway	Exclusive right lane

Note: All turning movements are permitted at intersections 1–13. Left-turn ingress and right-turn movements are permitted at intersections 14 and 15.

Table 1-4 Gard's Regression Equations⁵

Table 4. Regression equations.		
Movement	Condition	Equation
Major-street left turn	Approach volume ≤ 100 VPH/PHF	Max. Queue = -2.042 + 1.167 ln(AppVol) + 0.975*TS
	Approach volume > 100 VPH/PHF	Max. Queue = +4.252 - 1.23*Lanes + 0.07996*Speed + 1.412*TS - 374.028/AppVol + 0.00001144*AppVol *ConfVol
Minor-street left turn	Approach volume ≤ 60 VPH/PHF	Max. Queue = +0.958 + 0.00111*(AppVol) ² + 0.000333*(ConfVol)
	Approach volume > 60 VPH/PHF	Max. Queue = +6.174 - 2.313*TS + 0.03307*Speed - 1201.644/ConfVol + 0.00006549 (AppVol) ²
Minor-street right turn	Approach volume ≤ 100 VPH/PHF	Max. Queue = -19.822 + 0.688ln(AppVol) + 1.886*TS + 0.369*(Lanes) ² + 0.00000288*(ConfVol) ² + 0.401*Speed
	Approach volume > 100 VPH/PHF	Max. Queue = -26.23 + 0.132*Speed + 0.00000603*(ConfVol) ² + 4.909ln(AppVol)
Minor-street shared left/through/right	All conditions	Max. Queue = -12.916 + 3.225ln(AppVol) + 0.00569*(ConfVol for LTs & THs) - 0.000177*(ConfVol for RTs) - 2.109*(RT %) - 3.157*TS

Where

⁵ John T. Gard. "Estimation of Maximum Queue Lengths at Unsignalized Intersection", ITE Journal, November 2001, Pg: 26-34

AppVol = hourly traffic volume divided by peak-hour factor (PHF) for subject movement

ConflVol = hourly traffic volume divided by PHF that conflicts with subject movement (refer to the Highway Capacity Manual to identify movements that conflict with subject approach)

TS = a dummy variable with a value of 1 if a traffic signal is located on the major street within one-quarter mile of the subject intersection and 0 otherwise

Lanes = number of through lanes occupied by conflicting traffic

Speed = posted speed limit on major street (in miles per hour)

RT % = Percentage of vehicles on shared left/through/right minor street approach that turn right

In his comparison, Gard found that the 2000 HCM method showed a tendency to underestimate the queues. The Two-Minute Rule was successfully predicting 8 out of 10 cases, with in one vehicle variation. According to the author, in 49 out of 51 comparisons, the regression equation provided maximum queue-length estimates that were as accurate as or more accurate than other methods used in this study.

1.2 Problem

The problem for this work is defined as reasonably estimating the 95th percentile queue length, the length of the queue that has a probability of 5 percent or less of being exceeded during the peak hour, which is critical to the operational success and safety of the intersection.

1.3 Purpose

This study describes the maximum queue length model development and validation for two-way stop control. It further checks the consistency of queue length predictions among the widely used methods.

1.4 Study Methodology

The first step in the development of a model is data collection. Data collection requires prior effort in the form of: identifying the parameters influencing queue behavior, checking the sample size, location, season and duration of data collection. After collection, data is processed to get the required inputs. Then, model comparison is performed to identify the deficiencies among the existing models. This leads to model formulation and statistical analysis of data to predicted the maximum number of vehicles in the stopped queue. Next, model validation is conducted to check the model accuracy. Finally, sensitivity analysis will identify the limits to the value of input variables. The step-by-step procedure is shown in Figure 1-2.

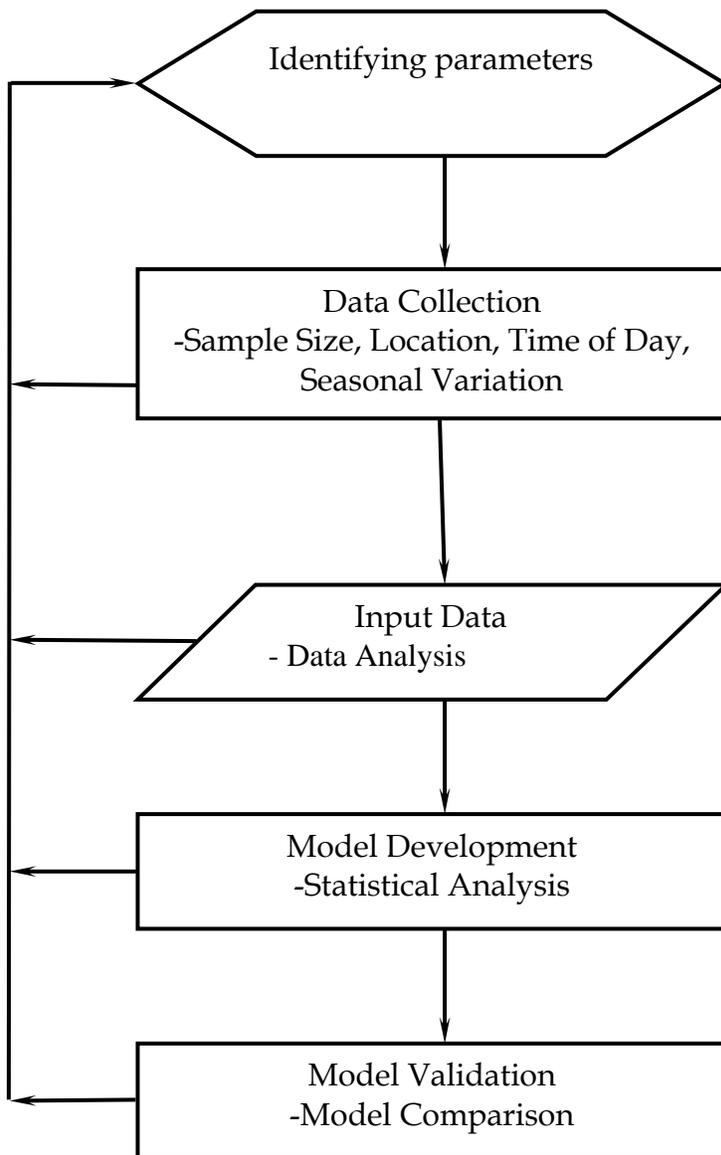


Figure 1-2 Study Methodology

1.5 Summary

This chapter briefly explains the current adopted methods to estimate the queue lengths at two-way STOP controlled intersections. Not every model is representing a true condition, so this study was directed to get a closer solution to the problem. The methodology explained above outlines the study process.

2 Data Collection & Analysis

Data plays a key role in model development, validation and comparing competitive models for consistency. This chapter focuses on the data collection procedure, data synthesis, and preparation of data sets for various lane groups. Data collection technologies are not discussed here.

2.1 Data collection

Intersections were chosen to cover a range of lane configurations, geographic regions, functional classifications, and traffic conditions. In total, 15 intersections shown in Appendix B, Table B1 were used for data collection.

Out of 15 intersections: 7 of are from region 2, 5 are from region 3, one from region 4, and the remaining 2 from region 5. 10 (67%) are within the urban growth boundary, and the remainder 5 (33%) are rural. 13 intersections have either OR or US route as the major approach. Three of the intersections have an upstream signal within 1000 ft. Six intersections have either an exclusive or two-way left turn lane. More than half of the intersections, totaling 9, have skewed approaches. Three intersection approaches are off-set from the major approach. Nearly half of the intersections are 3 legged (7 or 47%). Finally, three intersections major approaches have flaredness.

Appendix B, Table B2 represents the time frame of the data collection. All the data collected belong to either year 2010 or year 2009. Where the data is available, both AM and PM peak periods are covered.

2.2 Data Analysis

For each approach, information regarding the geometry, lane groups and associated movements, turn lane information, and traffic volume by movement is noted.

Queue data is collected through video logs provided by the Transportation Data Section. Maximum number of vehicles in the stopped queue is noted for every 15 minute interval for both peak periods by taking the maximum value of the observed queue on each lane group. Traffic volume for the same period is obtained from Traffic Count Management (TCM) program used by Transportation Data Section.

Hourly traffic volume is computed by summing the corresponding four 15 min intervals. A peak hour factor is calculated and applied to obtain hourly flow rates. Time periods having only hourly traffic volume are not considered in the analysis period. In the absence of a calculated peak hour factor, default peak hour factors can be taken to prevent data loss. This was not done due to the

importance of the study for model comparison and development in order to emulate actual traffic conditions.

The next step is to calculate the conflicting traffic flow rate according the procedure documented in Highway Capacity Manual. An excerpt from the HCM showing the two-way STOP controlled configurations are given in Appendix A, Figure A1 and the calculation of the associated conflicting flow rates for different lane movements are shown in Appendix A, Figure A2. Conflicting flow rates from individual movements in a lane group are added algebraically to obtain the lane group conflicting movements. Similarly, lane group flow rates are obtained by adding individual lane group flow rates.

This methodology adopted disaggregates approaches where individual intersection approaches and individual lane groups within approaches are treated separately. Both geometry of the intersection and the distribution of traffic movements play a key role in segmenting the intersection into lane groups. The following are the excerpts from HCM⁶ related to the definitions of various lane groups:

An exclusive left-turn lane or lanes should normally be designated as a separate lane group unless there is also a shared left-through lane present, in which case the proper lane grouping will depend upon the distribution of traffic volume between the movements. The same is true of an exclusive right-turn lane.

On approaches with exclusive left-turn or right-turn lanes, or both, all other lanes on the approach would generally be included in a single lane group. Some example lane groups according to HCM are given in Appendix A, Figure A3.

The following lane groups are considered:

MNLTR is for a minor street on a four legged intersection having a single lane for left, through and right turn movements

MNLR is for a minor street on a three legged intersection having a single lane for left and right turn movements

MJL is for a major left turn movement irrespective of exclusive/median/TWTL configuration

MNL is for an exclusive left turn lane on a minor approach of either a four or three legged intersection

MNR is for an exclusive right turn lane on a four legged intersection minor approach

⁶ "Chapter 16- signalized intersections", 2000 Highway Capacity Manual, Transportation Research Board, Washington, D.C., pg:16-6

The next step is to calculate the hourly observed queue length by taking the maximum of four 15 min interval queue lengths. Estimate queue lengths by the 2 minute rule, HCM methodology, and Gard's equation. This step completes the data set for one intersection on one approach. Repeat the process for all approaches and intersections.

2.3 Summary

The data collection methodology is explained in this chapter. Then the step by step process for data analysis and the preparation of input data are given. This step leads to the comparison of existing methodologies to check the consistency among the models.

3 Comparison of Existing Methodologies

The first part of this study is to check the performance of the 2 minute rule, HCM method and Gard’s equation for estimating queue lengths. The following sub sections analyze performance of these models for each lane configuration.

3.1 Major Left (MJL)

Summary statistics are given in Table 3-1. Observed queue length varies from a minimum of 1 vehicle to a maximum of 8 vehicles. At lower volumes, for instance a minimum volume of 1 vehicle, all methods tend to give zero queue lengths corresponding to the observed queue length of 1. If the volume reaches the maximum observed volume of 134 vph, the HCM method seems to be insensitive, only yielding a single vehicle in the stopped queue.

Table 3-1 Summary Statistics for MJL

	Observed	Two_min_Rule	HCM_Method	Gards	UOL
Count	219	219	219	219	219
Average	2.58904	0.931507	0.0182648	1.73973	29.9361
Variance	2.61933	0.843911	0.0180135	1.48699	618.436
Standard deviation	1.61843	0.918646	0.134214	1.21942	24.8684
Minimum	1.0	0.0	0.0	0.0	1.0
Maximum	8.0	5.0	1.0	4.0	134.0
Range	7.0	5.0	1.0	4.0	133.0
Std. skewness	6.76384	5.8038	43.7694	-0.987539	6.45329
Std. kurtosis	3.52603	4.16918	153.913	-3.80584	3.8392
CONVOL					
Count	219				
Average	506.95				
Variance	60516.9				
Standard deviation	246.002				
Minimum	2.0				
Maximum	972.0				
Range	970.0				
Std. skewness	-0.729985				
Std. kurtosis	-1.97501				

A scatter diagram of both observed and estimated queue lengths of the 219 data points is shown in Figure 3-1. The horizontal axis shows the observation number and the vertical axis represents the model predicted maximum number of vehicles in the stopped queue. Table 3-2 compares the relative performance of each model. The queue length given by Gard’s equations for the MJL lane configuration matches 32% of observations, which is highest among these methods. Only 8% are matched by the Two-Minute rule. If the queue length matching criterion is relaxed to predict ± 1 vehicle, 53% are matched by the Two-Minute rule and 76% by Gard’s equation. The Two-Minute rule is out-performed here due to the fact that it does not consider the opposing traffic volume.

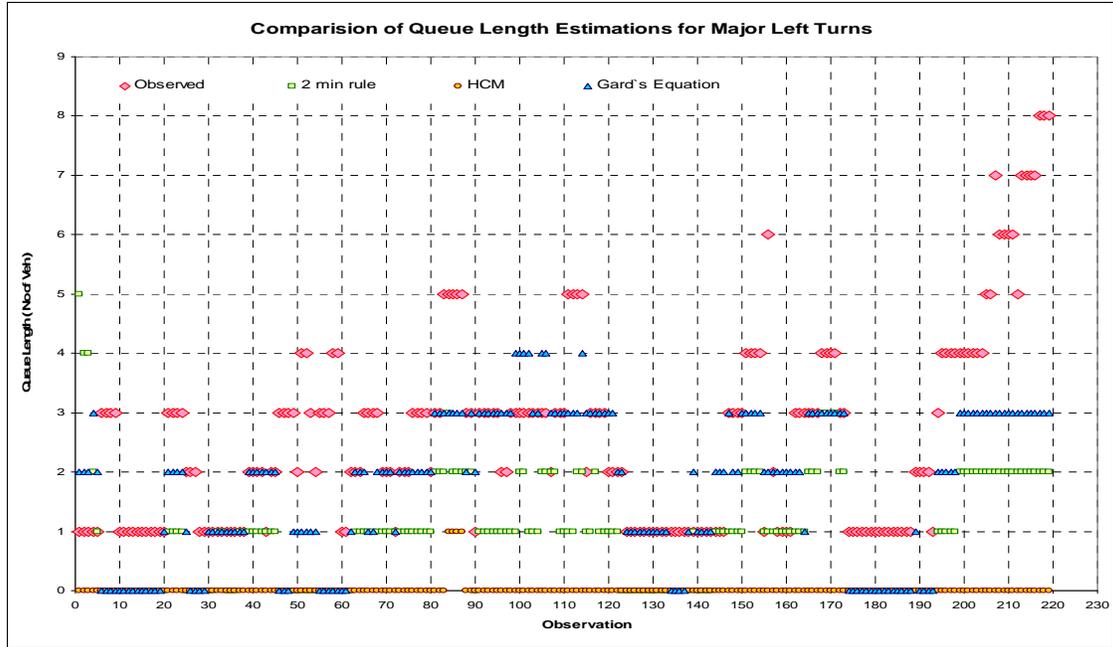


Figure 3-1 Scatter plot of all methods for MJL

Table 3-2 Difference between Observed and Model Outputs for MJL

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation	
		% of Obs	Cum. % of Obs	% of Obs	Cum. % of Obs	% of Obs	Cum. % of Obs
Over Estimated	-5	0%	1%	0%	0%	0%	0%
	-4	0%		0%		0%	
	-3	1%		0%		0%	
	-2	0%		0%		0%	
Acceptable	-1	0%	53%	0%	35%	12%	76%
	0	8%		0%		32%	
	1	44%		35%		32%	
Under Estimated	2	26%	46%	16%	65%	11%	24%
	3	11%		28%		7%	
	4	5%		12%		4%	
	5	3%		4%		1%	
	6	1%		2%		0%	
	7	0%		2%		0%	
	8	0%		1%		0%	
	9	0%		0%		0%	
	10	0%	0%	0%			

3.2 Minor, Share LTR (MNLTR)

The scatter plot matrix in Figure 3-2 clearly shows that an increase in both volume and conflicting volume has positive impact on the number of vehicles in queue. Table 3-3 shows the summary statistics for each of the selected data variables. It includes measures of central tendency, measures of variability, and measures of shape. As Volume increase, a nearly linear relation is observed between queue length and volume. A similar relation is observed for Gard's equation for increasing conflicting volume. The HCM method yields almost constant queue lengths up to a point, beyond which even a small increment in conflicting volume triggers an exponential increase in queue length.

The scatter plot of queue lengths for different models is shown in Figure 3-2 and Figure 3-3. It seems the Two-Minute Rule relatively closely follows the observed trend. To explore further, the differences between observed queue lengths and predicted queue lengths are tabulated in Table 3-4. The Two-Minute Rule shows 22% of the predictions exactly matched observations. The Two-Minute Rule attains 65% predictability within ± 1 vehicle. For the same instance, 36% are matched by the HCM method.

Table 3-3 Summary Statistics for MJL

	UOL	CONUOL	OBSERVED	Two min Rule	HCM
Count	143	143	143	143	143
Average	80.1119	2159.78	2.55944	2.67832	2.01399
Variance	9690.61	936297.0	3.19186	11.1775	25.2252
Standard deviation	98.4409	967.625	1.78658	3.34327	5.02247
Minimum	3.0	897.0	1.0	0.0	0.0
Maximum	396.0	4016.0	8.0	13.0	21.0
Range	393.0	3119.0	7.0	13.0	21.0
Std. skewness	8.73745	3.07341	6.83757	8.43334	12.4561
Std. kurtosis	5.4864	-2.00372	3.77464	5.02124	12.931
Gards					
Count	143				
Average	8.8951				
Variance	57.672				
Standard deviation	7.59421				
Minimum	0.0				
Maximum	24.0				
Range	24.0				
Std. skewness	4.16384				
Std. kurtosis	-1.2633				

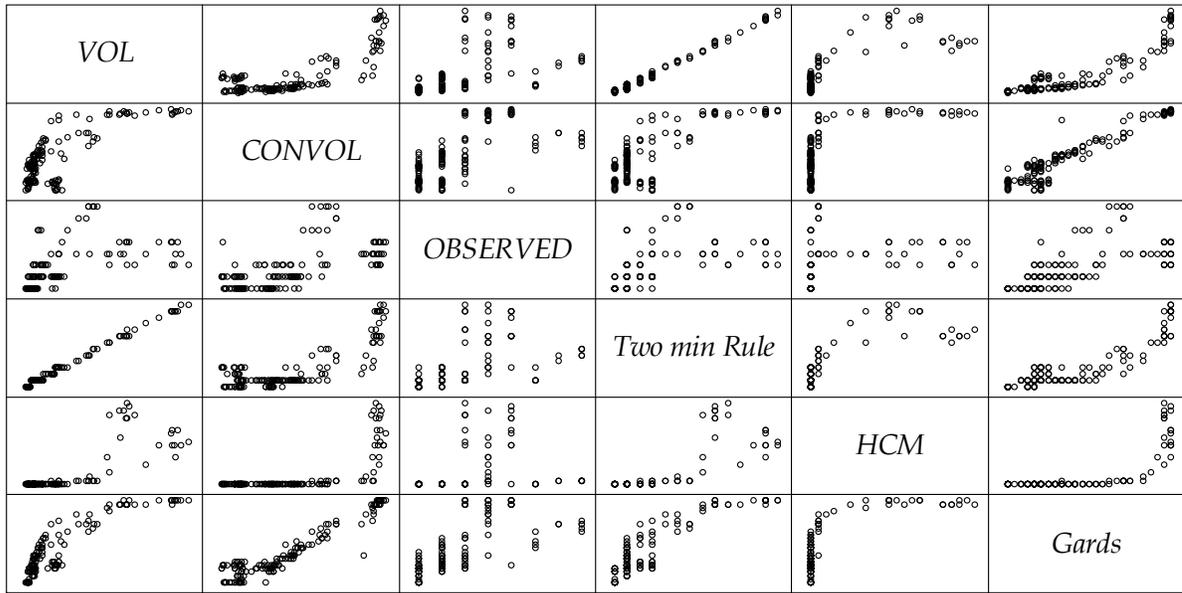


Figure 3-2 Scattered plot of VOL, CONVOL, & Queue Lengths

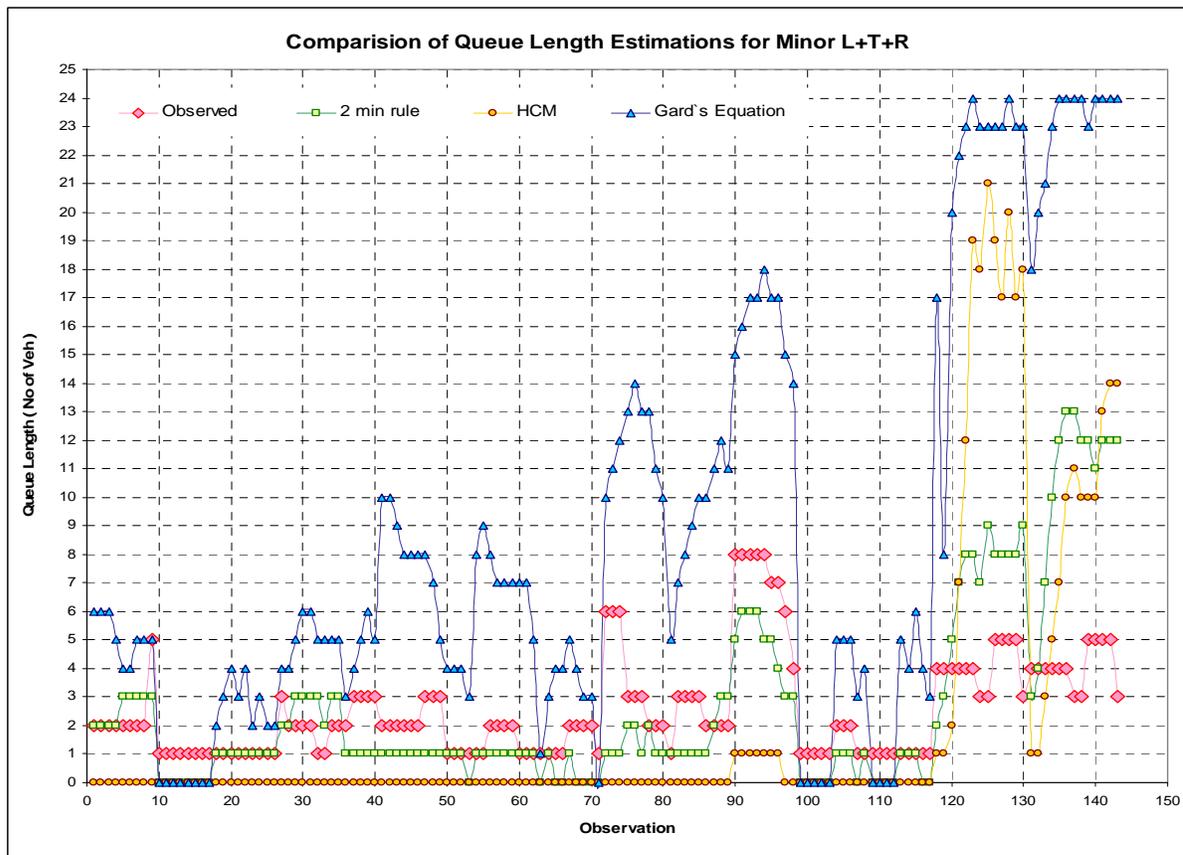


Figure 3-3 Scatter plot of all methods for MNLTR

Table 3-4 Difference between Observed and Model Outputs for MNLTR

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation	
		% of Obs	Cum. % of Obs	% of Obs	Cum. % of Obs	% of Obs	Cum. % of Obs
Over Estimated	< -10	0%	15%	6%	13%	19%	80%
	-10	1%		0%		5%	
	-9	2%		1%		4%	
	-8	1%		2%		4%	
	-7	2%		1%		3%	
	-6	3%		1%		6%	
	-5	0%		1%		5%	
	-4	2%		0%		10%	
	-3	4%		1%		15%	
-2	1%	0%	10%				
Acceptable	-1	9%	65%	1%	36%	6%	20%
	0	22%		0%		1%	
	1	34%		36%		13%	
Under Estimated	2	15%	20%	28%	50%	0%	0%
	3	3%		13%		0%	
	4	0%		1%		0%	
	5	2%		1%		0%	
	6	0%		4%		0%	
	7	0%		3%		0%	
	8	0%		0%		0%	
	9	0%		0%		0%	
	10	0%		0%		0%	

3.3 MNLTR

A minimum queue length of a single vehicle to a maximum of 7 vehicles in the queue is observed for MNLTR lane configuration. The Two-Minute Rule estimated a maximum of 7 vehicles, while the HCM predicts a maximum of 3 vehicles in queue for the similar prevailing conditions. Gard's equation tends to overestimate the queues as shown in Table 3-6. The data description is given in Table 3-5 and shown in Figure 3-4 and Figure 3-5.

Around 33% are exactly matched with observed values by the Two-Minute Rule which performs far better than other models. If the difference between the observed and estimated queue lengths is relaxed to ± 1 , 84% are matched for the Two-Minute Rule, while 46% is matched for Gard's equation.

Table 3-5 Summary Statistics for MNL

	VOL	CONVOL	Observed	Two_min_Rule	HCM
Count	81	81	81	81	81
Average	66.7654	1490.32	2.39506	2.23457	0.259259
Variance	2683.03	282928.0	1.96698	2.73179	0.519444
Standard deviation	51.798	531.909	1.40249	1.65281	0.720725
Minimum	12.0	802.0	1.0	0.0	0.0
Maximum	224.0	2489.0	7.0	7.0	3.0
Range	212.0	1687.0	6.0	7.0	3.0
Std. skewness	5.40679	1.60496	5.77836	5.02874	11.2284
Std. kurtosis	3.87235	-1.99022	5.03514	2.94678	16.0452
Gards					
Count	81				
Average	4.23457				
Variance	7.68179				
Standard deviation	2.7716				
Minimum	0.0				
Maximum	9.0				
Range	9.0				
Std. skewness	0.768588				
Std. kurtosis	-1.66468				

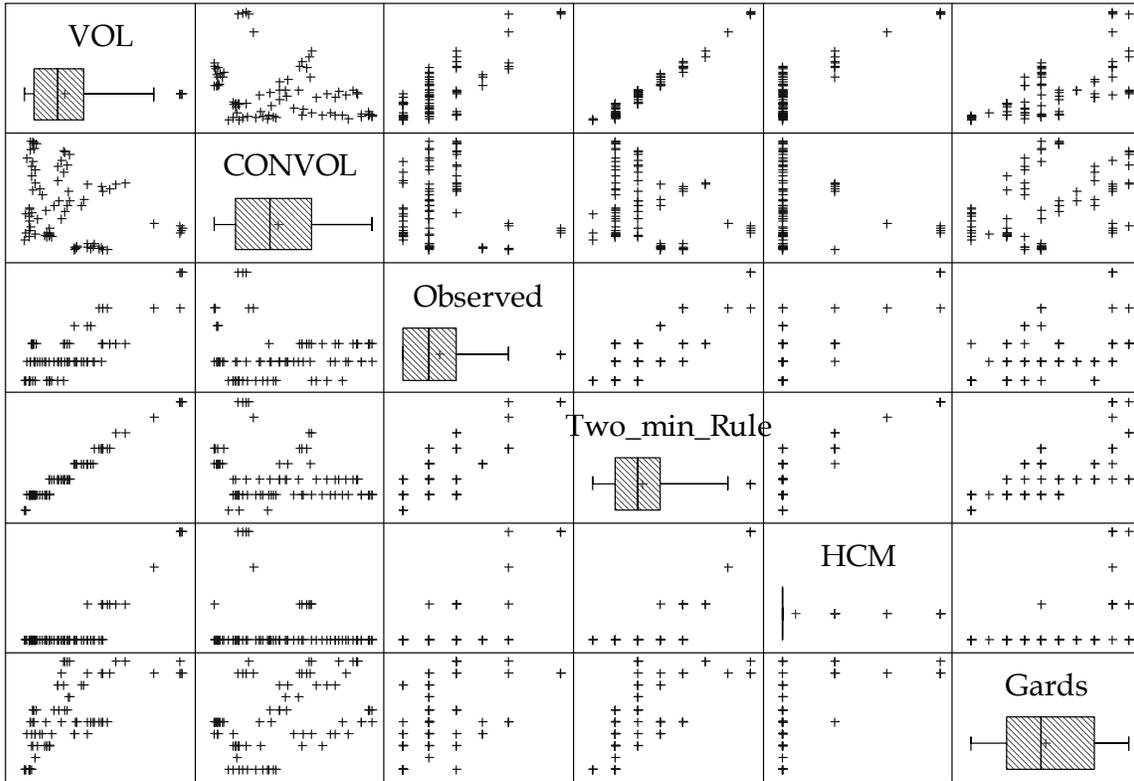


Figure 3-4 Scattered plot of VOL, CONVOL, & Queue Lengths for MNL

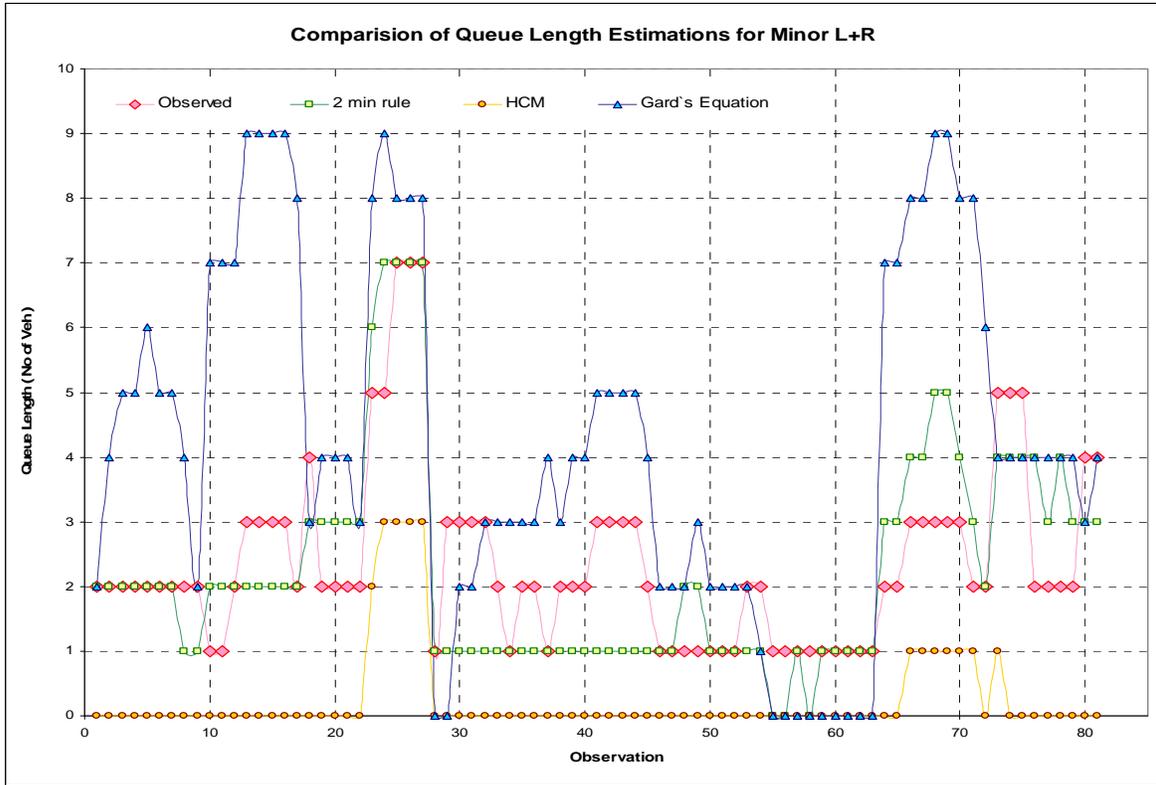


Figure 3-5 Scatter plots of all methods for MNL

Table 3-6 Difference between Observed and Model Outputs for MNL

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation	
		% Obs	Cum % of Obs	% Obs	Cum % of Obs	% Obs	Cum % of Obs
Over Estimated	-6	0%	6%	0%	0%	12%	53%
	-5	0%		0%		7%	
	-4	0%		0%		4%	
	-3	0%		0%		7%	
Acceptable	-2	6%	84%	0%	27%	22%	46%
	-1	21%		0%		17%	
	0	33%		0%		6%	
Under Estimated	1	30%	10%	27%	73%	22%	1%
	2	10%		46%		0%	
	3	0%		16%		1%	
	4	0%		9%		0%	
	5	0%		2%		0%	
	6	0%		0%		0%	
	7	0%		0%		0%	
8	0%	0%	0%	0%			
	9	0%	0%	0%	0%	0%	

3.4 MNL

All models consistently underestimate the MNL queue length. Only the Two-Minute Rule predictions show a consistent trend with increasing volumes. This group has the lowest number of observations, 38. A maximum queue length of 11 vehicles is observed in queue. The description of the data is shown in Table 3-7. Scatter plots of the observed data and the distribution of the queue length predictions are shown in Figure 3-6 and 3-7 respectively.

The Two-Minute Rule predicts 18% of the observed values. Predictions improved to nearly 40% with a tolerance of a single vehicle. Gard's equation predicts 32% as shown in Table 3-8.

Table 3-7 Summary Statistics for MNL

	UOL	CONUOL	Observed	Two_min_Rule	HCM
Count	38	38	38	38	38
Average	70.0	928.132	4.44737	2.26316	0.526316
Variance	1469.95	62556.9	7.33499	1.60455	0.472262
Standard deviation	38.3399	250.114	2.70832	1.26671	0.687213
Minimum	7.0	457.0	1.0	0.0	0.0
Maximum	132.0	1286.0	11.0	4.0	2.0
Range	125.0	829.0	10.0	4.0	2.0
Std. skewness	-0.189789	-0.58707	3.24311	-0.691012	2.4046
Std. kurtosis	-1.37725	-1.66867	1.94871	-1.15355	-0.306799
Gards					
Count	38				
Average	4.92105				
Variance	5.58819				
Standard deviation	2.36394				
Minimum	0.0				
Maximum	7.0				
Range	7.0				
Std. skewness	-2.88054				
Std. kurtosis	0.120273				

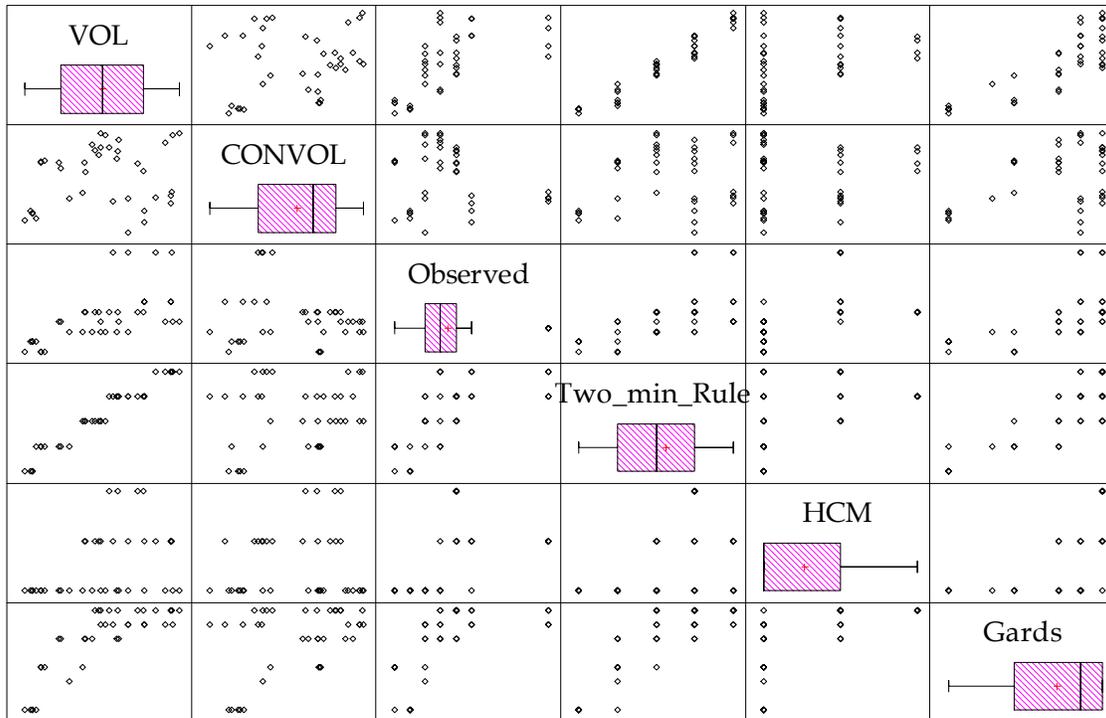


Figure 3-6 Scattered plot of VOL, CONVOL, & Queue Lengths for MNL

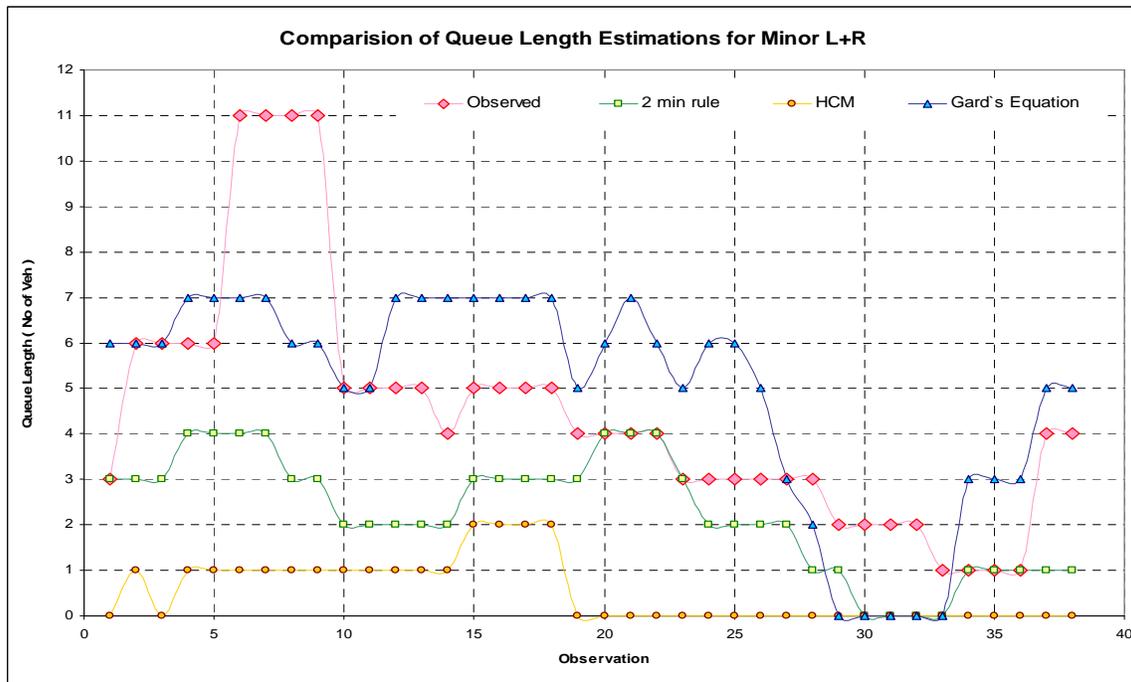


Figure 3-7 Scatter plot of all methods for MNL

Table 3-8 Difference between Observed and Model Outputs for MNL

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation	
		% Obs	Cum. % Obs	% Obs	Cum. % Obs	% Obs	Cum. % Obs
Over Estimated	-5	0%	0%	0%	0%	0%	47%
	-4	0%		0%		0%	
	-3	0%		0%		13%	
	-2	0%		0%		34%	
Acceptable	-1	0%	39%	0%	11%	13%	32%
	0	21%		0%		13%	
	1	18%		11%		5%	
Under Estimated	2	29%	61%	11%	89%	11%	21%
	3	21%		32%		0%	
	4	0%		26%		5%	
	5	0%		8%		5%	
	6	0%		3%		0%	
	7	5%		0%		0%	
	8	5%		0%		0%	
	9	0%		0%		0%	
10	0%	11%	0%				

3.5 Summary

Analysis related to the relative performances of the Two-Minute Rule, Highway Capacity Manual Method, and John T. Gard’s equations in predicting the queue lengths at two-way STOP controlled intersections are presented in the chapter. For Major Left Turn (MJL) lane configurations Gard’s equation performs well. For the remaining lane configurations, the Two-Minute Rule predicts the best. On average, the performance of matching with \pm one vehicle variation does not exceed 70 percent. This triggers the effort to build a model to assess the queue length in consistent manner. The model development procedure is presented in the next chapter.

4 Model Development

Model development plays a key role for predicting vehicles in stopped queue. A well formulated model significantly explains the changes to the dependent variable for variations in the independent variable. Though a variety of models ranging from simple regression to complex simulation exists today, this study is limited to development of a regression model. This is intuitive because of the simple model structure and easy usage. Not only are regression models straight forward to understand, but also the model user can study model sensitivity by changing the values of the independent variables.

This chapter explains the model development process. First, the factors influencing the queue behavior for different lane configuration movements at unsignalized intersections are identified. Then, a scatter diagram of the observed data is analyzed to identify the appropriate regression model. Next, for the chosen regression model the combinations of influenced variables are identified to incorporate into the model. After that the model is formulated and developed for each lane configuration movement. Finally, statistical tests are performed to check the model reasonableness.

4.1 *Factors Influencing Queue Behavior*

Primarily geometry, operations, traffic flow, and human travel characteristics influence the queue behavior. Over the past decade, numerous models have been developed taking into consideration one or a combination of these characteristics. Influencing factors are listed in Table 4-1.

It is very difficult to capture the effects of all parameters. Instead, it is assumed that lane configuration, major and minor approach volume, number of conflicting lanes, volumes and speed on the conflicting lanes, presence of turn lanes, right turn channelization, flared right turn lanes, and presence of traffic signal within 1000ft of the study intersections are the most influencing parameters.

Table 4-1 Factors Influencing Queue Behavior

Category	List of Factors
Geometry	Number of approaches Number of lanes on both major and minor approaches Lane configuration (shared/separate) Chanallization / Flared approaches Median Type Grade Sight Distance Intersection Skewness
Operations	Traffic flow speed Upstream Signal
Traffic Flow	Approach volume Conflicting volume Arrival type Turning volume Percent of heavy vehicles Gap and follow-up times Time of day / seasonal variation
Human Factors	Reaction time

It is a fact that traffic movements will behave uniquely for the given lane configuration. Lane groups are identified according the definitions provided in the 2000 Highway Capacity Manual. This chapter reiterates the lane groups considered in this study as shown below and from this point forward models are designated with the lane group codes shown below.

- *MNLTR* is for minor street on four legged intersection having a single lane for left, through and right turn movements
- *MNLR* is for minor street on three legged intersection having a single lane for left, and right turn movements
- *MJL* is for a Major left turn movement irrespective of exclusive/median/TWTL configuration
- *MNL* is for an exclusive left turn lane on minor approach of either four or three legged intersection
- *MNR* is for an exclusive right turn lane on four legged intersection minor approach

The data collection and analysis chapter explains the preparation of the data sets for modeling purposes. Scatter plots for different lane group configurations are drawn to get the outlook of the data trend and type of model to choose. As such, the trends of the queue lengths (denoted as *QL*, is the maximum number of vehicles in the stopped queue for the same time unit of volume measurement) with respect to the changes in volume (*VOL*) and conflicting volume (*CONVOL*) were evaluated. Following sub-sections describe the findings for various lane group models.

4.2 Poisson Regression Model

A general description of the Poisson Regression Model is given as it is the model which is best suited for the study conditions. Number of vehicles in the queue is of count type, often called discrete type, taking only a finite number of values. The probability distribution that is specifically suited for count data is the Poisson probability distribution. An interesting feature of the Poisson distribution: its variance is the same as its mean value.

The Poisson regression model may be written as:

$$\text{Ln}(Y) = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \dots + \beta_n \times X_n$$

Where

Ln = Natural Logarithm
Y = Dependent Variable

X₁, X₂, X_n = Independent or Explanatory Variables

β₀ = Constant

β₁, β₂, β₃ = Model coefficients corresponds X₁, X₂, X_n

4.3 Major Left Turn (MJL) Model

Table 4-2 shows summary statistics for each of the selected data variables. It includes measures of central tendency, measures of variability, and measures of shape. Of particular interest here are the standardized skewness and standardized kurtosis, which can be used to determine whether the sample comes from a normal distribution. Values of these statistics outside the range of -2 to +2 indicate significant departures from normality, which would tend to invalidate many of the statistical procedures normally applied to this data. In this case, the variables show standardized skewness values outside the expected range are QL, VOL and VOLCONVOL. QL and VOL variables show standardized kurtosis values outside the expected range. QL value varies from 1 to 8 vehicles.

Table 4-2 Summary Statistics of Major Left Turn Data

	QL	VOL	CONVOL	VOLCONVOL
Count	219	219	219	219
Average	2.58904	29.9361	506.95	17163.2
Variance	2.61933	618.436	60516.9	3.03245E8
Standard deviation	1.61843	24.8684	246.002	17413.9
Minimum	1.0	1.0	2.0	51.0
Maximum	8.0	134.0	972.0	65715.0
Range	7.0	133.0	970.0	65664.0
Std. skewness	6.76384	6.45329	-0.729985	7.04459
Std. kurtosis	3.52603	3.8392	-1.97501	1.19911

Transformations are one of the methods to make the variables more normal. A scattered matrix plot among the variables shows no definite pattern between QL and explanatory variables VOL, CONVOL, and their product.

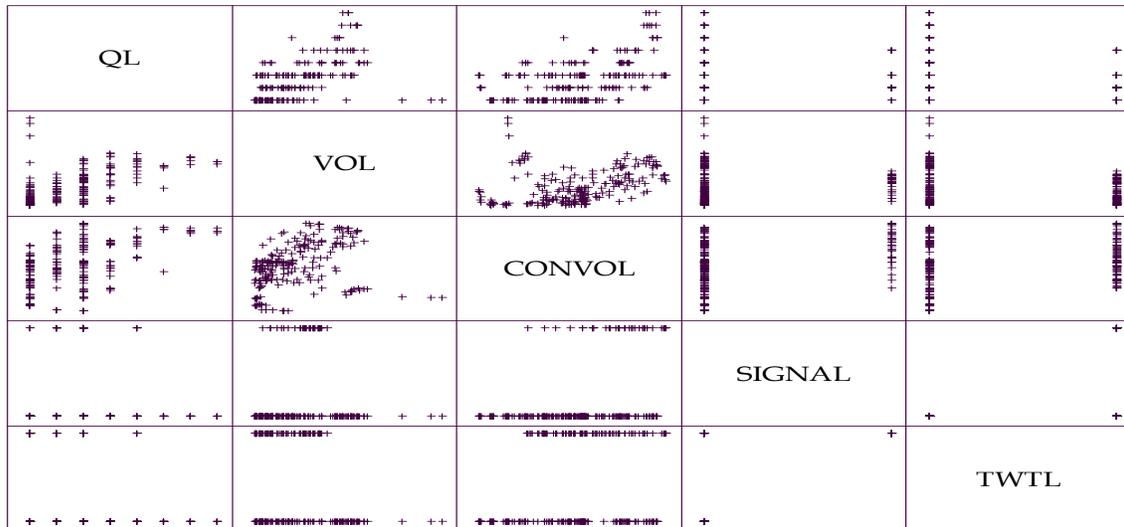


Figure 4-1 Scattered Matrix Plot between the variables for MJL

Though many variables listed in Table 4-1 adequately explain the queue behavior, only VOL, CONVOL, presence of upstream signal within 1000 ft distance from intersection (SIGNAL), and presence of exclusive left turn lane (Coded as LT, either median left turn lane or two-way left turn lane-TWTL) are selected, based upon lowest Mean Squared Error, highest R-squared value, and lowest Cp Statistics⁷. Results of this analysis are shown in Table 4-3.

After selecting the variables, analysis needs to be performed to identify the correlation among the independent variables. Table 4-4 shows Spearman rank correlations between each pair of variables. These correlation coefficients range between -1 and +1 and measure the strength of the association between the variables. In contrast to the more common Pearson correlations, the Spearman coefficients are computed from the ranks of the data values rather than from the values themselves. Consequently, they are less sensitive to outliers than the Pearson coefficients. Also shown in parentheses is the number of pairs of data values used to compute each coefficient. The third number in each location of the table is a P-value which tests the statistical significance of the estimated correlations. P-values below 0.05 indicate statistically significant “non-zero” correlations at the 95% confidence level.

⁷ Mallows' Cp statistic is a measure of the bias in a model, based on a comparison of total mean squared error to the true error variance. Unbiased models have an expected value of approximately p , where p is the number of coefficients in the fitted model including constant

Table 4-3 Selection of Independent Variable for MJL Model

Regression Model Selection				
Dependent variable: QL				
Independent variables:				
A=VOL				
B=CONVOL				
C=VOLCONVOL				
D=SIGNAL				
E=TWTL				
Number of complete cases: 219				
Number of models fit: 32				
Model Results				
MSE	R-Squared	Adjusted R-Squared	Cp	Included Variables
2.61933	0.0	0.0	502.988	
1.70135	35.3442	35.0463	250.514	A
1.97238	25.0443	24.6989	324.672	B
1.04799	60.1737	59.9902	71.7446	C
2.6145	0.642103	0.184232	500.365	D
2.15002	18.2936	17.9171	373.277	E
1.43358	45.7712	45.2691	177.441	AB
1.05279	60.1756	59.8068	73.7312	AC
1.7088	35.3602	34.7617	252.399	AD
1.53786	41.8265	41.2879	205.842	AE
1.00842	61.8539	61.5007	61.6472	BC
1.94017	26.6084	25.9288	315.411	BD
1.21427	54.0672	53.6419	117.711	BE
1.02662	61.1654	60.8058	66.6044	CD
0.822207	68.898	68.61	10.9306	CE
1.85755	29.7336	29.0829	292.91	DE
0.992793	62.619	62.0974	58.1386	ABC
1.40365	47.1491	46.4117	169.52	ABD
1.04388	60.6953	60.1469	71.9892	ABE
1.03133	61.1682	60.6264	68.5844	ACD
0.808386	69.5624	69.1377	8.14724	ACE
1.44915	45.4362	44.6748	181.853	ADE
1.0039	62.2008	61.6733	61.1501	BCD
0.824137	68.9693	68.5363	12.4173	BCE
1.16617	56.0909	55.4783	105.14	BDE
0.808903	69.5429	69.1179	8.28729	CDE
0.991387	62.8456	62.1511	58.5075	ABCD
0.811127	69.6012	69.033	9.86749	ABCE
1.02763	61.4872	60.7673	68.2879	ABDE
0.792048	70.3162	69.7614	4.7195	ACDE
0.811509	69.5869	69.0184	9.97064	BCDE
0.793088	70.4162	69.7217	6.0	ABCDE

Table 4-4 Spearman Rank Correlation Matrix

	QL	VOL	CONVOL	SIGNAL	TWTL
QL		0.6551 (219) 0.0000	0.4673 (219) 0.0000	0.1529 (219) 0.0239	-0.4519 (219) 0.0000
VOL	0.6551 (219) 0.0000		0.4411 (219) 0.0000	0.2112 (219) 0.0018	-0.2724 (219) 0.0001
CONVOL	0.4673 (219) 0.0000	0.4411 (219) 0.0000		0.3937 (219) 0.0000	0.1713 (219) 0.0114
SIGNAL	0.1529 (219) 0.0239	0.2112 (219) 0.0018	0.3937 (219) 0.0000		0.4983 (219) 0.0000
TWTL	-0.4519 (219) 0.0000	-0.2724 (219) 0.0001	0.1713 (219) 0.0114	0.4983 (219) 0.0000	
Correlation (Sample Size) P-Value					

The model for predicting the maximum number of vehicles in the stopped queue is formulated as a Poisson regression equation with volume, conflicting volume, presence of upstream signal, and presence of left turn lane as the independent or explanatory variables.

$$\text{Ln}(QL_{\text{MJL}}) = \beta_0 + \beta_1 \times \text{VOL}_{\text{MJL}} + \beta_2 \times \text{CONVOL}_{\text{MJL}} + \beta_3 \times \text{SIGNAL} + \beta_4 \times \text{LT}$$

Where

Ln = Natural Logarithm

VOL = Approach volume for MJL lane configuration

CONVOL = Conflicting volume for MJL

SIGNAL = Presence of upstream signal within 1000 ft of the intersection (1 if there is a signal, otherwise 0)

LT = Presence of left turn lane (1 if there is an exclusive left turn lane/median left turn lane/ two-way left turn lane, otherwise 0)

β_0 = Constant

$\beta_1, \beta_2, \beta_3, \beta_4$ = Model coefficients corresponds to VOL, CONVOL, SIGNAL, and LT variables

The developed model with coefficient values and corresponding statistical tests are explained below:

$$\text{Ln}(QL) = 0.3925 + 0.0059 \times \text{VOL} + 0.00104 \times \text{CONVOL} + 0.49 * \text{SIGNAL} - 0.81 * \text{LT}$$

$$QL = e^{(0.3925 + 0.0059 \times \text{VOL} + 0.00104 \times \text{CONVOL} + 0.49 * \text{SIGNAL} - 0.81 * \text{LT})}$$

Percentage of deviance explained by model = 66.959

As the volume and conflicting volume increases, QL increases. Presence of an upstream signal within 1000 feet from the intersection increases QL. Moreover, presence of a separate left turn lane decreases QL as compared to not having a turn lane. These signs indicate the reasonableness of the model. Statistical analysis of the model is shown in Table 4-5.

Table 4-5 Statistical Analysis of MJL Model

Poisson Regression			
Dependent variable: QL			
Factors:			
UOL			
CONUOL			
SIGNAL			
TWTL			
Estimated Regression Model (Maximum Likelihood)			
Parameter	Estimate	Standard Error	Estimated Rate Ratio
CONSTANT	0.392535	0.119829	
UOL	0.00586284	0.00187515	1.00588
CONUOL	0.0010416	0.000187399	1.00104
SIGNAL	0.48999	0.16482	1.6323
TWTL	-0.811879	0.140689	0.444023
Analysis of Deviance			
Source	Deviance	Df	P-Value
Model	137.162	4	0.0000
Residual	67.6827	214	1.0000
Total (corr.)	204.845	218	
Percentage of deviance explained by model = 66.959			
Adjusted percentage = 62.0772			
Likelihood Ratio Tests			
Factor	Chi-Square	Df	P-Value
UOL	9.22746	1	0.0024
CONUOL	33.0217	1	0.0000
SIGNAL	8.98112	1	0.0027
TWTL	37.2249	1	0.0000
Residual Analysis			
	Estimation	Validation	
n	219		
MSE	2.78244		
MAE	0.761339		
MAPE	33.5926		
ME	0.0189021		
MPE	-16.1152		

Because the P-value for the model in the Analysis of Deviance table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level. In addition, the P-value for the residuals is greater than or equal to 0.10, indicating that the model is not significantly worse than the best possible model for this data at the 90% or higher confidence level. The percentage of deviance in QL explained by the model equals 66.959%. This statistic is similar to the usual R-Squared statistic. The adjusted percentage, which is more suitable for comparing models with different numbers of independent variables, is 62.0772%.

In determining whether the model can be simplified, notice that the highest P-value for the likelihood ratio tests is 0.0027, belonging to SIGNAL. Because the P-value is less than 0.01, that term is statistically significant at the 99% confidence level. Consequently, it is not advisable to remove any variables from the model.

Table 4-6 shows estimated correlations between the coefficients in the fitted model. These correlations can be used to detect the presence of serious multicollinearity, i.e., correlation amongst the predictor variables. In this case, there is 1 correlation with an absolute value greater than 0.5.

Table 4-6 Correlation Matrix for Estimated Coefficients for MJL Model

Correlation matrix for coefficient estimates				
	CONSTANT	VOL	CONVOL	SIGNAL
CONSTANT	1.0000	-0.3962	-0.6333	0.3000
VOL	-0.3962	1.0000	-0.3402	-0.1638
CONVOL	-0.6333	-0.3402	1.0000	-0.2150
SIGNAL	0.3000	-0.1638	-0.2150	1.0000
TWTL	-0.3673	0.3793	-0.0517	-0.7019

	TWTL			
CONSTANT	-0.3673			
VOL	0.3793			
CONVOL	-0.0517			
SIGNAL	-0.7019			
TWTL	1.0000			

The plot of the fitted model with 95% confidence limits (shown as red lines), predicted QL Vs Observed QL, and the Residual plot for QL, are shown in Figures 4-2 and 4-3, respectively. The Residual plot for QL shows most of the predictions have an error of ± 1 vehicle, due to rounding off to the nearest integer.

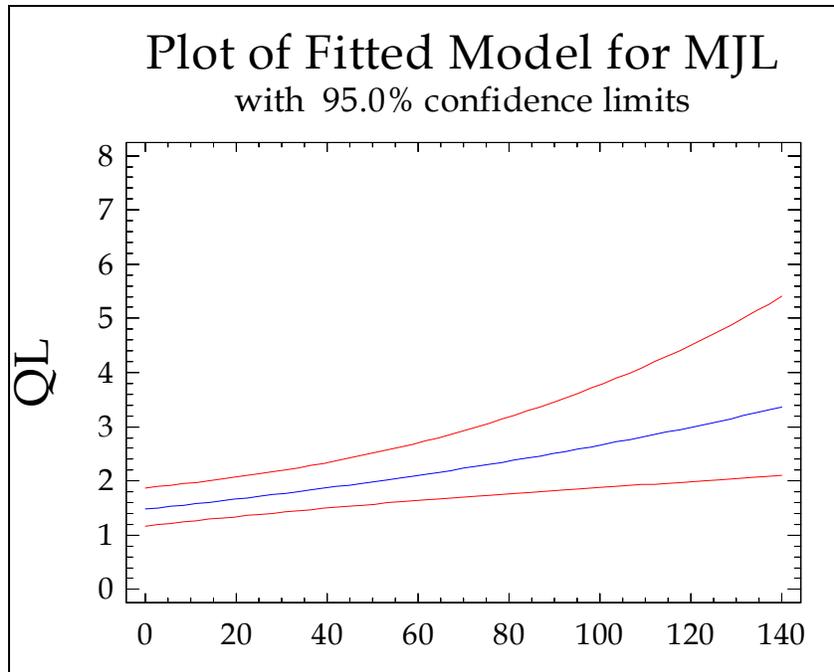


Figure 4-2 Plot of Fitted Model for MJL

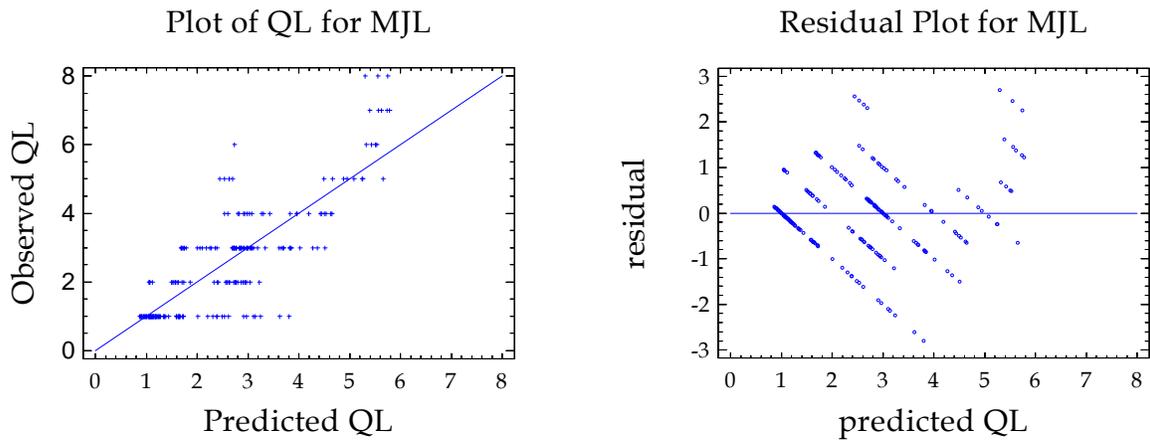


Figure 4-3 Residual Plot for MJL

4.4 Minor, Shared LTR (MNLTR)

There are 143 observations for this lane group category. Scatter plots are shown in Figure 4-4.

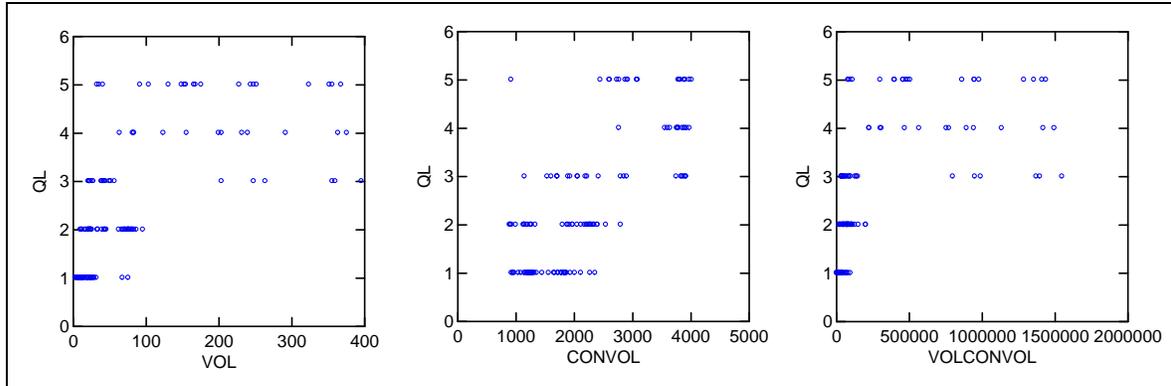


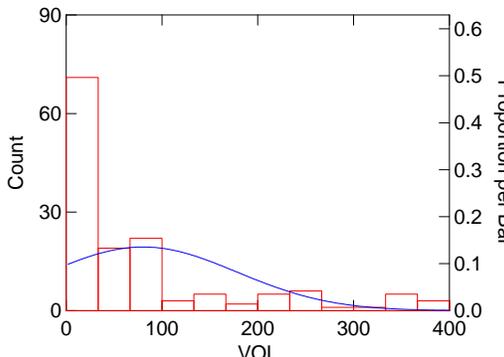
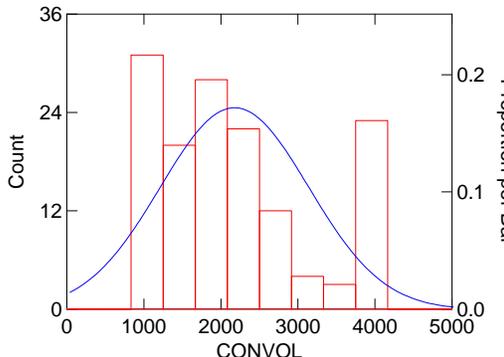
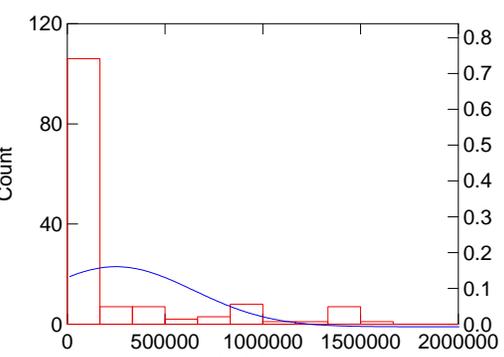
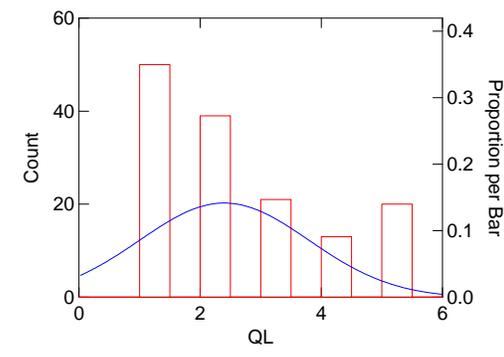
Figure 4-4 Scattered Plot of QL Vs VOL, CONVOL, and VOL*CONVOL for MNLTR

There exists no definite pattern among the variables, which shows the requirement of the transformation of the explanatory variables. This may be due to the random arrival of the minor street movements, which will influence the queue formation to be random. Various combinations of transformations are analyzed to recognize the patterns in the data, but none are explaining the queue behavior properly.

The distributions of volume, conflicting volume, product of these volumes, and queue lengths are following gamma distribution as listed in Table 4-7. In the Shapiro-Wilk test statistic for normality, the P-value for each distribution is less than the alpha value of 0.05. It is concluded that the data are not from the normally distribution population. So, VOL and CONVOL are not normally distributed.

It is assumed that the queue lengths tend to follow the random vehicle arrivals and therefore the Poisson regression model is tested. Before going further into the model development it is worthwhile to get a snap shot of the summary statistics for each variable in the model.

Table 4-7 Fitting Distributions for MNLTR Data

<p style="text-align: center;">FITTED DISTRIBUTION</p> 	<p style="text-align: center;">FITTED DISTRIBUTION</p> 
<p>Distribution: Gamma Kolmogorov-Smirnov test statistic = 0.246430 Lilliefors Probability (2-tail) = 0.000000 Shapiro-Wilk test statistic for normality = 0.714618 p-value = 0.000000 Estimated Shape (alpha) = 0.666946 Scale (beta) = 120.117511 Chi-square test statistic = 30.032026 df = 5, p-value = 0.000015</p>	<p>Distribution: Gamma Kolmogorov-Smirnov test statistic = 0.118544 Lilliefors Probability (2-tail) = 0.000041 Shapiro-Wilk test statistic for normality = 0.895070 p-value = 0.000000 Estimated: Shape (alpha) = 5.017084 Estimated: Scale (beta) = 430.484357 Chi-square test statistic = 68.746965 df = 8, p-value = 0.000000</p>
<p style="text-align: center;">FITTED DISTRIBUTION</p> 	<p style="text-align: center;">FITTED DISTRIBUTION</p> 
<p>Distribution: Gamma Kolmogorov-Smirnov test statistic = 0.339958 Lilliefors Probability (2-tail) = 0.000000 Shapiro-Wilk test statistic for normality = 0.620258 p-value = 0.000000 Variable Name: VOLCONVOL Estimated: Shape (alpha) = 0.387296 Estimated: Scale (beta) = 636926.604762 Chi-square test statistic = 61.330933 df = 4 p-value = 0.000000</p>	<p>Distribution: Gamma Kolmogorov-Smirnov test statistic = 0.234447 Lilliefors Probability (2-tail) = 0.000000 Shapiro-Wilk test statistic for normality = 0.831541 p-value = 0.000000 Variable Name: QL Estimated: Shape (alpha) = 2.935354 Estimated: Scale (beta) = 0.817142 Chi-square test statistic = 192.914128 df = 8 p-value = 0.000000</p>

As given in Table 4-8, the maximum queue length observed is 5 with minimum of 1 vehicle in the queue. Volumes range from 3 vehicles per hour to 396 vehicles with a mean arrival rate of 80 vph. A mean conflicting flow rate of 2160 vph is observed. The distribution of the product of volume and conflicting volume is skewed more to the right side, and all data sets are showing a trend that is not normally distributed.

Table 4-8 Summary Statistics for MNLTR data

Item	VOL	CONVOL	VOLCONVOL	QL
N of cases	143	143	143	143
Minimum	3.000	897.000	3600.000	1.000
Maximum	396.000	4016.000	1550149.000	5.000
Range	393.000	3119.000	1546549.000	4.000
Sum	11456.000	308848.000	3.52752E+07	343.000
Median	34.000	1935.000	69315.000	2.000
Mean	80.112	2159.776	246679.392	2.399
95% CI upper	96.385	2319.734	312434.895	2.631
95% CI lower	63.839	1999.819	180923.888	2.166
Std. Error	8.232	80.917	33263.424	0.117
Standard Dev	98.441	967.625	397772.702	1.405
Variance	9690.607	936297.414	1.58223E+11	1.974
C.V.	1.229	0.448	1.613	0.586
Skewness(G1)	1.790	0.630	2.002	0.697
SE Skewness	0.203	0.203	0.203	0.203
Kurtosis(G2)	2.248	-0.821	2.846	-0.808
SE Kurtosis	0.403	0.403	0.403	0.403
SW Statistic	0.715	0.895	0.620	0.832
SW P-Value	0.000	0.000	0.000	0.000

The next step after choosing the model is to select the appropriate independent variable(s) from the pool of identified influencing variables. The combinations of the variables are tested and chosen based upon the largest R² Value⁸, lowest Mallows' Cp statistic⁹ value, and lowest Mean Square Error (MSE). The analysis results are shown in Table 4-9. Analysis indicates that VOL, CONVOL, and the product of VOL and CONVOL may explain the queue behavior significantly. This step leads to the basic model formulation and development.

⁸ The adjusted R-Squared statistic measures the proportion of the variability in QL which is explained by the model.

⁹ Mallows' Cp statistic is a measure of the bias in a model, based on a comparison of total mean squared error to the true error variance. Unbiased models have an expected value of approximately p, where p is the number of coefficients in the fitted model including constant

Table 4-9 Selection of Dependent Variables for Poisson Regression

Regression Model Selection				
Dependent variable: QL				
Independent variables:				
A=VOL				
B=CONVOL				
C=RATIO				
D=CONLANES				
E=CONSPEED				
F=VOLCONVOL				
Number of complete cases: 143				
Number of models fit: 63				
Model Results				
MSE	R-Squared	Adjusted R-Squared	Cp	Included Variables
0.611827	69.6575	69.0026	1.439	ABF
0.614559	69.7412	68.8642	3.06236	ABDF
0.615163	69.7115	68.8336	3.19602	ABEF
0.615628	69.6886	68.81	3.29892	ABCF
0.618901	69.7483	68.6442	5.03076	ABCDF
0.618946	69.7461	68.6419	5.04066	ABDEF
0.619454	69.7212	68.6162	5.15227	ABCEF
0.76147	62.2362	61.4212	34.8098	BCE
0.762063	62.4787	61.3911	35.7194	ABCE
0.764151	61.8306	61.2853	34.6337	BC
0.765162	62.3261	61.2341	36.4056	BCDE
0.765886	62.0172	61.1974	35.7946	BCD
0.766978	61.963	61.1421	36.0382	ABC
0.767159	62.5015	61.1329	37.617	ABCDE
0.768823	61.8715	61.0486	36.4495	BCF
0.770255	62.3501	60.976	38.2975	BCDEF
0.848438	57.6204	57.015	53.5652	AB
0.877404	56.1736	55.5475	60.0711	BD
0.884581	55.4995	55.1839	61.1023	B
0.885017	55.7933	55.1618	61.7809	BF
0.88622	55.7332	55.1008	62.0512	BE
1.08969	45.1812	44.7924	107.5	A
1.16248	41.5195	41.1047	123.965	F
1.44374	27.3698	26.8546	187.591	C
1.85577	6.64192	5.9798	280.797	D
1.9738	0.0	0.0	308.663	

A correlation matrix between QL and the independent variables is given in Table 4-10. A correlation coefficient of either +1 or -1 shows perfect correlation in positive or negative manner. A 0 correlation coefficient shows that independent variables can not explain queue behavior sufficiently. A positive sign in the matrix represents a positive correlation which indicates that QL increases as the independent or explanatory variable increases. Moreover, all explanatory variables are correlated with QL. There exists correlation between VOL and CONVOL. This may be due to the fact that CONVOL and VOL are volumes on different lanes, and as VOL increase from the off-peak period to the peak period, CONVOL may increase in volume to represent the peak condition. Finally, the correlation between the product of VOL and CONVOL with VOL and CONVOL

will be expected as the product comes from both VOL and CONVOL. Though consideration of all three explanatory variables seems to dampen the model performance as a result of correlation, the comparison of models with various combinations of variables as shown in Table 4-11 indicates that these variables perform better for the observed data. The other reason to consider the product term is to capture the QL for the corresponding pair of VOL and CONVOL.

Table 4-10 Correlation Matrix of Variables for MNLTR Model

Variable	VOL	CONVOL	VOLCONVOL	QL
VOL	1.000	0.780	0.986	0.672
CONVOL	0.780	1.000	0.824	0.745
VOLCONVOL	0.986	0.824	1.000	0.644
QL	0.672	0.745	0.644	1.000

The model for predicting the maximum number of vehicles in the stopped queue is formulated as a Poisson regression equation with volume, conflicting volume, and their product as independent or explanatory variables.

$$\ln(QL_{MNLTR}) = \beta_0 + \beta_1 \times VOL_{MNLTR} + \beta_2 \times CONVOL_{MNLTR} + \beta_3 \times (VOL_{MNLTR} * CONVOL_{MNLTR})$$

Where

Ln = Natural Logarithm

VOL = Approach volume for MNLTR lane configuration

CONVOL = Conflicting volume for MNLTR

β_0 = Constant

$\beta_1, \beta_2, \beta_3$ = Model coefficients corresponds to VOL, CONVOL, and VOL*CONVOL

The Poisson Regression model with coefficient values and corresponding statistical tests are explained below:

$$\ln(QL) = -0.7844 + 0.01636 \times VOL + 0.0006 \times CONVOL - 0.0000043 \times (VOL * CONVOL)$$

$$QL = e^{(-0.7844 + 0.01636 \times VOL + 0.0006 \times CONVOL - 0.0000043 \times (VOL * CONVOL))}$$

Percentage of Deviation Explained by Model = 71.643

Signs of the independent variables show reasonableness in the models, as a positive change in the VOL and CONVOL results in increasing QL. But the occurrence of VOL and the corresponding CONVOL as a product triggers a decrease in QL, due to the fact that lower approach volume may not yield larger queue lengths, rather more waiting time in the stopped queue. As approach volumes increase to capacity, the approach volume has a greater higher impact on queue than the conflicting volume.

Table 4-11 Possion Regression Model for MNLTR

Poisson Regression			
Dependent variable: QL			
Factors: VOL, CONVOL, VOLCONVOL			
Estimated Regression Model (Maximum Likelihood)			
Parameter	Estimate	Standard Error	Estimated Rate Ratio
CONSTANT	-0.784374	0.247345	
VOL	0.0163606	0.00357521	1.0165
CONVOL	0.000598612	0.000102971	1.0006
VOLCONVOL	-0.0000043145	9.55645E-7	0.999996
Analysis of Deviance			
Source	Deviance	Df	P-Value
Model	80.5605	3	0.0000
Residual	31.8861	139	1.0000
Total (corr.)	112.447	142	
Percentage of deviance explained by model = 71.6433			
Adjusted percentage = 64.5289			
Likelihood Ratio Tests			
Factor	Chi-Square	Df	P-Value
VOL	19.6951	1	0.0000
CONVOL	33.5012	1	0.0000
VOLCONVOL	19.6016	1	0.0000
Residual Analysis			
	Estimation	Validation	
n	143		
MSE	1.53264		
MAE	0.581838		
MAPE	24.0153		
ME	-0.00423549		
MPE	-9.38096		

Statistical significance is given in Table 4-11. The model accounts for 71.6 percent of deviance¹⁰ explained in QL. Because the P-value for the model in the Analysis of Deviance table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level. In addition, the P-value for the residuals is greater than or equal to 0.10, indicating that the model is not significantly worse than the best possible model for this data at the 90% or higher confidence level.

¹⁰ The percentage of deviance statistic is similar to the usual R-Squared statistic.

The adjusted percentage, which is more suitable for comparing models with different numbers of independent variables, is 64.5289%. In determining whether the model can be simplified, notice that the highest P-value for the likelihood ratio tests is 0.0000, belonging to VOLCONVOL. Because the P-value is less than 0.01, that term is statistically significant at the 99% confidence level. Consequently, it is not advisable to remove any variables from the model. The plot of predicted QL and Observed QL is shown in Figure 4-5. The Residual plot for QL given in Figure 4-6 shows most of the predictions have an error of ± 1 vehicle, due to rounding-off error to the nearest integer.

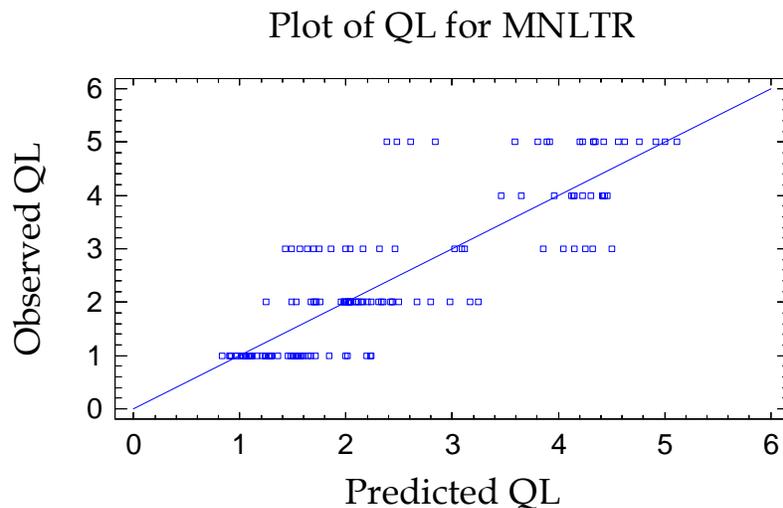


Figure 4-5 Plot of Predicted QL and Observed QL for MNLTR Model

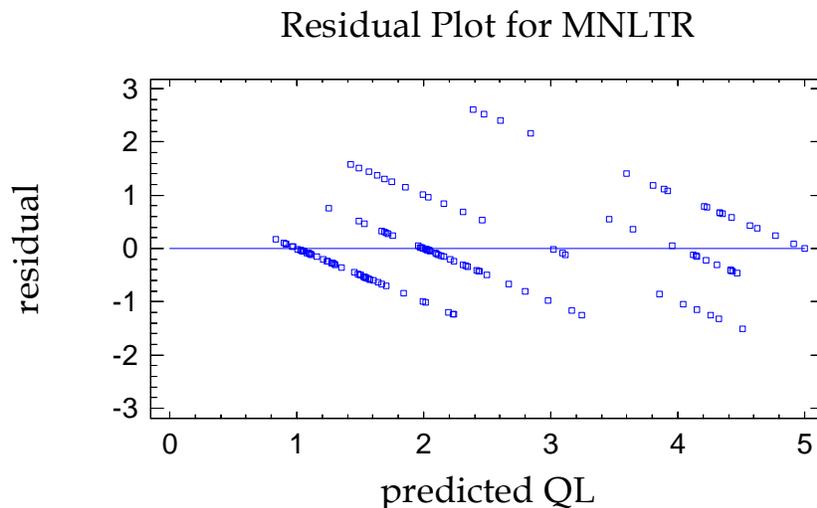


Figure 4-6 Residual Plot of Predicted QL and Observed QL for MNLTR Model

4.5 Minor, LR (MNLr)

This section explains model development for minor approach queue length estimation for a T-intersection where there exists a single lane for both left and right turn movements. There are 81 data point for model development. Summary statistics of the observed data corresponding to VOL, CONVOL, and their product is given in Table 4-12.

Table 4-12 Summary Statistics for MNLr lane group

	VOL	CONVOL	VOLCONVOL	QL
Count	81	81	81	81
Average	66.7654	1490.32	91333.1	2.39506
Variance	2683.03	282928.0	4.53467E9	1.96698
Standard deviation	51.798	531.909	67340.0	1.40249
Minimum	12.0	802.0	11250.0	1.0
Maximum	224.0	2489.0	273273.0	7.0
Range	212.0	1687.0	262023.0	6.0
Std. skewness	5.40679	1.60496	4.51837	5.77836
Std. kurtosis	3.87235	-1.99022	1.18442	5.03514

The observed maximum number of vehicles in the stopped queue (QL) varies from a minimum of 1 to a maximum of 7. Standardized skewness and standardized kurtosis statistics outside the range of -2 to +2 indicate significant departures from normality. So, QL is believed to be not from a Normal Distribution. Also VOL and VOLCONVOL show standardized skewness values outside the expected range. Standardized kurtosis values for VOL are also outside the expected range. Transformations may make these variables to be normal.

Analysis of the scatter plot between QL and VOL, and CONVOL shows the possible relation among variables. A box-and-whisker diagram or plot ¹¹ is shown as the diagonal of the matrix in Figure 4-7. No definite pattern is observed among these plots. Therefore QL is assumed to follow the random process of vehicle arrivals. A Poisson regression model is tested as the initial step in model development.

¹¹ is a convenient way of graphically depicting groups of numerical data through their five-number summaries: the smallest observation (sample minimum), lower quartile (Q1), median (Q2), upper quartile (Q3), and largest observation (sample maximum). A boxplot may also indicate which observations, if any, might be considered outliers.

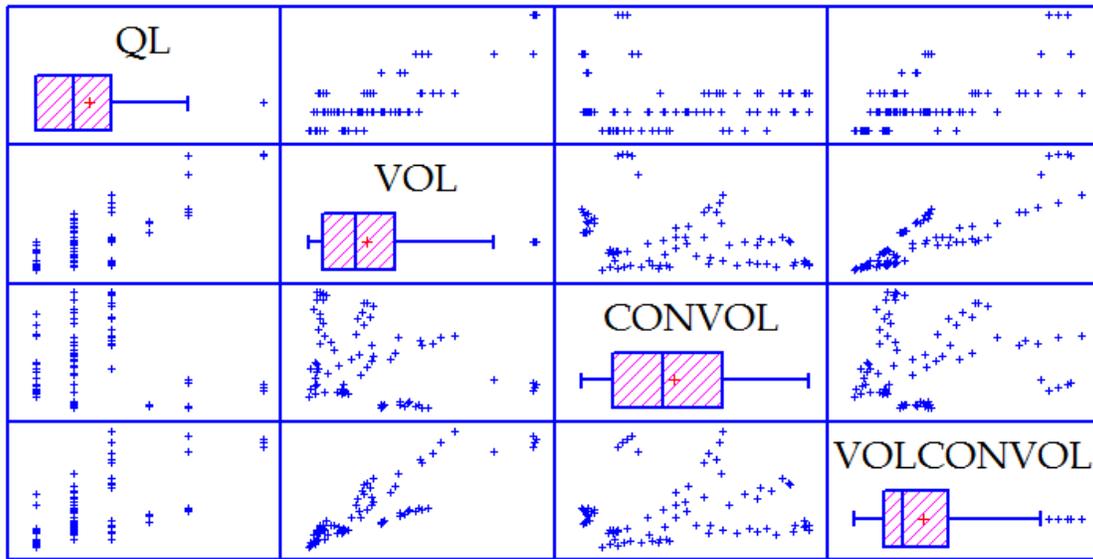


Figure 4-7 Scattered Matrix Plot for MNL Lane Group Data

The next step is to choose appropriate set of variables from the identified list of variables from Table 4-13. Based on the lowest MSE, highest R-squared, and lowest Cp Statistic, VOL, CONVOL, and the product of VOL and CONVOL are considered in the model.

Variable selection leads to the step of correlation analysis to check for the serial correlation among the independent variables. Due to the fact that VOL and CONVOL are simply volume occurring on different lanes, there may be some correlation between them. Likewise, the product term has either VOL or CONVOL which will trigger the correlation. This has an impact on the model, but they can not be excluded to capture the impact on QL. The correlation matrix is shown in Table 4-14.

Table 4-14 Pearson Correlation Matrix of Variables in MNL Model

Variable	QL	VOL	CONVOL	VOLCONVOL
QL	1.000	0.771	-0.064	0.643
VOL	0.771	1.000	-0.300	0.869
CONVOL	-0.064	-0.300	1.000	0.123
VOLCONVOL	0.643	0.869	0.123	1.000

Table 4-13 Variables Selection for MNL Model

Regression Model Selection				
Dependent variable: QL				
Independent variables:				
A=VOL				
B=CONVOL				
C=RATIO				
D=VOLCONVOL				
Number of complete cases: 81				
Number of models fit: 16				
Model Results				
MSE	R-Squared	Adjusted R-Squared	Cp	Included Variables
1.96698	0.0	0.0	220.885	
0.806398	59.5156	59.0031	44.4064	A
1.96698	1.25	0.0	221.638	B
1.77722	10.7766	9.64718	190.567	C
1.16805	41.3593	40.617	98.8545	D
0.755034	62.5741	61.6145	37.2345	AB
0.625702	68.9849	68.1896	18.0095	AC
0.810442	59.8276	58.7975	45.4708	AD
1.74266	13.6189	11.404	184.044	BC
1.14078	43.4533	42.0034	94.5747	BD
1.18154	41.433	39.9313	100.633	CD
0.633765	68.988	67.7797	20.0002	ABC
0.526761	74.224	73.2197	4.29828	ABD
0.604781	70.4063	69.2533	15.7471	ACD
1.03289	49.4574	47.4882	78.5694	BCD
0.524729	74.6569	73.3231	5.0	ABCD

The model for predicting the maximum number of vehicles in the stopped queue is formulated as a Poisson regression equation with volume, conflicting volume, and their product as independent or explanatory variables.

$$\ln(QL_{MNL}) = \beta_0 + \beta_1 \times VOL_{MNL} + \beta_2 \times CONVOL_{MNL} + \beta_3 \times (VOL_{MNL} * CONVOL_{MNL})$$

Where

Ln = Natural Logarithm

VOL = Approach volume for MNL lane configuration

CONVOL = Conflicting volume for MNL

β_0 = Constant

$\beta_1, \beta_2, \beta_3$ = Model coefficients corresponds to VOL, CONVOL, and VOL*CONVOL

The developed model with coefficient values and corresponding statistical tests are explained below:

$$\ln(QL) = -0.6319 + 0.0173 \times VOL + 0.00066 \times CONVOL - 0.000007913 (VOL * CONVOL)$$
$$QL = e^{(-0.6319 + 0.0173 \times VOL + 0.00066 \times CONVOL - 0.000007913 (VOL * CONVOL))}$$

Percentage of deviance explained by model = 69.25

The signs of the independent variables show reasonableness in the models, as a positive change in the VOL and CONVOL results in increasing QL. But the occurrence of VOL and the corresponding CONVOL as a product triggers a decrease in QL, due to the fact that lower approach volume may not yield larger queue lengths, rather increase waiting time in the stopped queue. As volumes increase to capacity, the approach volume has a higher impact on queue than the conflicting volume. The statistics of model development are given in Table 4-15.

The model accounts for 69.25 percent of deviance¹² explained in QL. Because the P-value for the model in the Analysis of Deviance table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level. In addition, the P-value for the residuals is greater than or equal to 0.10, indicating that the model is not significantly worse than the best possible model for this data at the 90% or higher confidence level.

Table4-15 Poisson Regression Model for MNL Data

¹² The percentage of deviance statistic is similar to the usual R-Squared statistic.

Poisson Regression

Dependent variable: QL

Factors:

UOL
CONVOL
UOLCONVOL

Estimated Regression Model (Maximum Likelihood)

Parameter	Estimate	Standard Error	Estimated Rate Ratio
CONSTANT	-0.631869	0.399069	
UOL	0.0172923	0.00455457	1.01744
CONVOL	0.000662669	0.000246615	1.00066
UOLCONVOL	-0.0000791327	0.0000352853	0.999992

Analysis of Deviance

Source	Deviance	Df	P-Value
Model	39.7281	3	0.0000
Residual	17.6416	77	1.0000
Total (corr.)	57.3698	80	

Percentage of deviance explained by model = 69.2492

Adjusted percentage = 55.3046

Likelihood Ratio Tests

Factor	Chi-Square	Df	P-Value
UOL	14.5754	1	0.0001
CONVOL	6.88596	1	0.0087
UOLCONVOL	5.17992	1	0.0228

Residual Analysis

	Estimation	Validation
n	81	
MSE	1.55813	
MAE	0.618471	
MAPE	26.7428	
ME	-0.0326626	
MPE	-9.84617	

The adjusted percentage of deviance, which is more suitable for comparing models with different numbers of independent variables, is 55.3%. In determining whether the model can be simplified, notice that the highest P-value for the likelihood ratio tests is 0.0228, belonging to VOLCONVOL. Because the P-value is less than 0.05, that term is statistically significant at the 95% confidence level. Consequently, it is not advisable to remove any variables from the model. The plot of predicted QL and Observed QL is shown in Figure 4-8. The Residual plot for QL given in Figure 4-9 shows most of the predictions have an error of ± 1 vehicle, due to rounding-off error to the nearest integer.

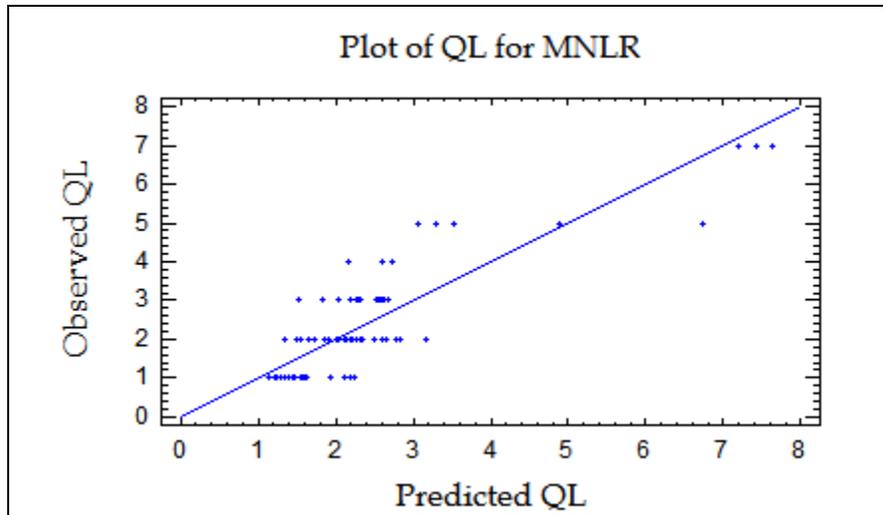


Figure 4-8 Plot of QL between Predicted Vs Observed QL for MNL

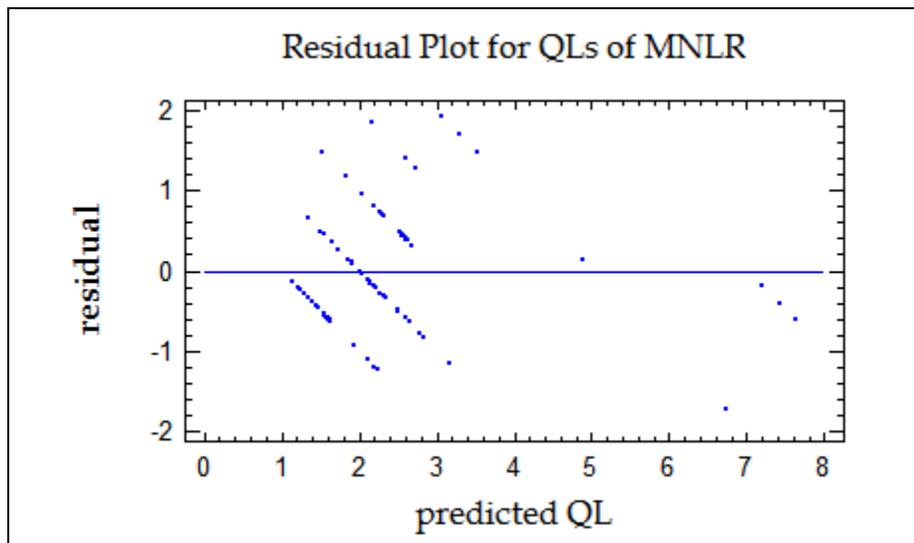


Figure 4-9 Residual Plot between Predicted QL and Residual for MNL

4.6 Minor Left (MNL)

There are 34 data point available for model development. Table 4-16 shows the maximum QL observed ranges from one vehicle to 6 vehicles. Sample QL, VOL and CONVOL come from a normal distribution except the ratio of conflicting volume to the approach volume (designated as RATIO). Intuition for taking RATIO as the explanatory variable is partly explained through the scattered diagram shown in Figure 4-10, where a pattern is observed between QL and RATIO.

Table 4-16 Summary Statistics of MNL Data

	QL	VOL	CONVOL	RATIO
Count	34	34	34	34
Average	3.67647	66.2353	949.412	23.2288
Variance	2.40731	1467.16	65609.0	350.695
Standard deviation	1.55155	38.3085	256.143	18.7268
Minimum	1.0	7.0	457.0	5.03
Maximum	6.0	132.0	1286.0	83.35
Range	5.0	125.0	829.0	78.32
Std. skewness	-0.589831	0.183785	-1.09969	3.62622
Std. kurtosis	-1.08621	-1.27444	-1.42946	2.37195

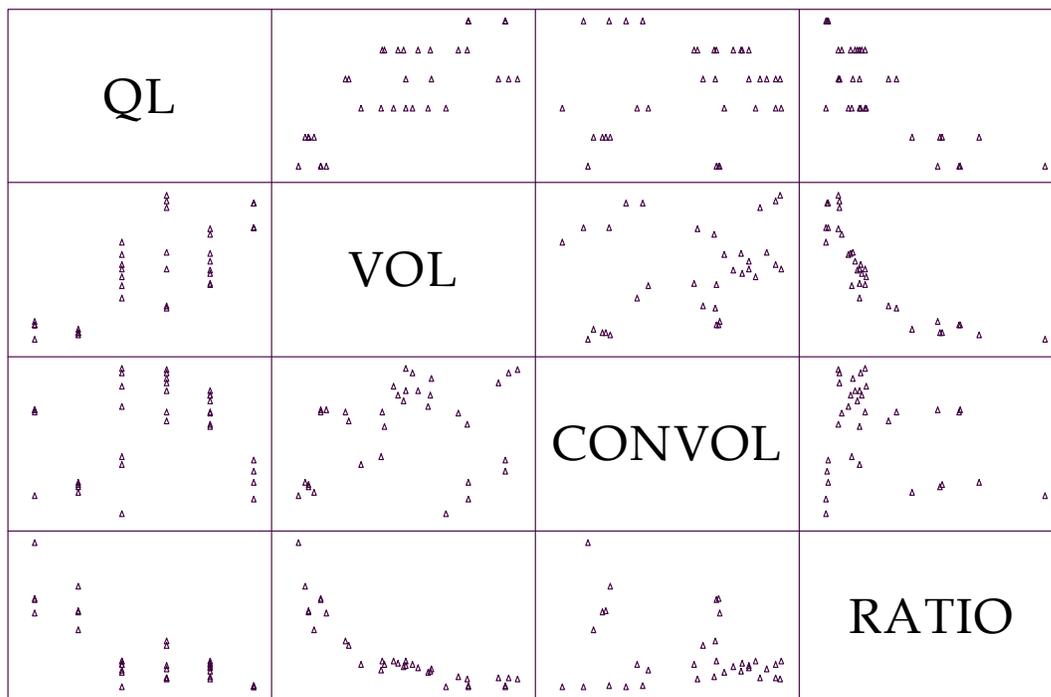


Figure 4-10 Scattered Matrix Plot of Variables for MNL

Table 4-17 shows the results of fitting various multiple regression models to describe the relationship between QL and 4 predictor variables. Models have been fit containing all combinations of from 0 to 4 variables. The statistics

tabulated include the mean squared error (MSE), the adjusted and unadjusted R-Squared values, and Mallows' Cp statistic. Models are determined to be best according to lowest values for all three criteria.

Table 4-17 Selection of Independent Variables for MNL Model

Regression Model Selection				
Dependent variable: QL				
Independent variables:				
A=VOL				
B=CONVOL				
C=VOLCONVOL				
D=RATIO				
Number of complete cases: 34				
Number of models fit: 16				
Model Results				
MSE	R-Squared	Adjusted R-Squared	Cp	Included Variables
2.40731	0.0	0.0	57.7769	
1.10936	55.3133	53.9169	10.1183	A
2.40731	3.0303	0.0	58.6512	B
1.69847	31.5832	29.4452	31.4225	C
0.91006	63.3415	62.196	2.91082	D
1.12208	56.2137	53.3888	11.31	AB
1.04489	59.2256	56.595	8.60595	AC
0.89479	65.083	62.8303	3.34741	AD
1.49105	41.8154	38.0616	24.2363	BC
0.910698	64.4622	62.1695	3.90472	BD
0.939369	63.3434	60.9784	4.90918	CD
1.01232	61.7708	57.9479	8.321	ABC
0.887672	66.4781	63.126	4.09488	ABD
0.861009	67.485	64.2336	3.19092	ACD
0.921012	65.2191	61.741	5.22521	BCD
0.884873	67.6977	63.2422	5.0	ABCD

Table 4-18 shows Spearman rank correlations between each pair of variables. These correlation coefficients range between -1 and +1 and measure the strength of the association between the variables. In contrast to the more common Pearson correlations, the Spearman coefficients are computed from the ranks of the data values rather than from the values themselves. Consequently, they are less sensitive to outliers than the Pearson coefficients. Also shown in parentheses is the number of pairs of data values used to compute each coefficient. The third number in each location of the table is a P-value which tests the statistical significance of the estimated correlations. P-values below 0.05 indicate statistically significant non-zero correlations at the 95% confidence level. The

following pairs of variables have P-values below 0.05: QL and VOL, QL and RATIO, and VOL and RATIO.

Table4-18 Spearman Rank Correlation Matrix for MNL Variables

Spearman Rank Correlations				
	QL	VOL	CONVOL	RATIO
QL		0.7117 (34) 0.0000	0.0560 (34) 0.7478	-0.7294 (34) 0.0000
VOL	0.7117 (34) 0.0000		0.3265 (34) 0.0607	-0.9276 (34) 0.0000
CONVOL	0.0560 (34) 0.7478	0.3265 (34) 0.0607		-0.0393 (34) 0.8215
RATIO	-0.7294 (34) 0.0000	-0.9276 (34) 0.0000	-0.0393 (34) 0.8215	

Correlation
(Sample Size)
P-Value

The model for predicting the maximum number of vehicles in the stopped queue is formulated as a Poisson regression equation with volume and conflicting volume as independent or explanatory variables.

$$\ln(QL_{MNL}) = \beta_0 + \beta_1 \times (CONVOL_{MNL}/VOL_{MNL})$$

Where

Ln = Natural Logarithm

QL = Maximum number of vehicles in the stopped queue

VOL = Approach volume for MNL lane configuration

CONVOL = Conflicting volume for MNL

β_0 = Constant

β_1 = Model coefficients corresponds to RATIO (CONVOL/VOL)

The Poisson regression model developed with coefficient values and corresponding statistical tests are explained below:

$\ln(QL) = 1.7934 - 0.025 \times (CONVOL/VOL)$ $QL = e^{(1.7934 - 0.025 \times (CONVOL/VOL))}$ <p>Percentage of deviance explained by model = 69.404</p>

Because the P-value for the model in the Analysis of Deviance table shown in Table 4-19 is less than 0.01, there is a statistically significant relationship between

the variables at the 99% confidence level. In addition, the P-value for the residuals is greater than or equal to 0.10, indicating that the model is not significantly worse than the best possible model for this data at the 90% or higher confidence level.

Table 4-19 Estimated Regression Model for MNL Data

Poisson Regression			
Dependent variable: QL			
Factors:			
RATIO			
Estimated Regression Model (Maximum Likelihood)			
Parameter	Estimate	Standard Error	Estimated Rate Ratio
CONSTANT	1.79343	0.14742	
RATIO	-0.0247881	0.00684342	0.975517
Analysis of Deviance			
Source	Deviance	Df	P-Value
Model	16.7513	1	0.0000
Residual	7.38449	32	1.0000
Total (corr.)	24.1357	33	
Percentage of deviance explained by model = 69.4043			
Adjusted percentage = 52.8314			
Likelihood Ratio Tests			
Factor	Chi-Square	Df	P-Value
RATIO	16.7513	1	0.0000
Residual Analysis			
	Estimation	Validation	
n	34		
MSE	3.41565		
MAE	0.820284		
MAPE	23.642		
ME	-0.00104352		
MPE	-6.81105		

The above Table also shows that the percentage of deviance in QL explained by the model equals 69.4043%. This statistic is similar to the usual R-Squared

statistic. The adjusted percentage, which is more suitable for comparing models with different numbers of independent variables, is 52.8314%.

In determining whether the model can be simplified, notice that the highest P-value for the likelihood ratio tests is 0.0000, belonging to RATIO. Because the P-value is less than 0.01, that term is statistically significant at the 99% confidence level. Consequently, it is not advisable to remove any variables from the model. Plots of the predicted vs observed QL and their residuals are shown in Figure 4-11 and 4-12.

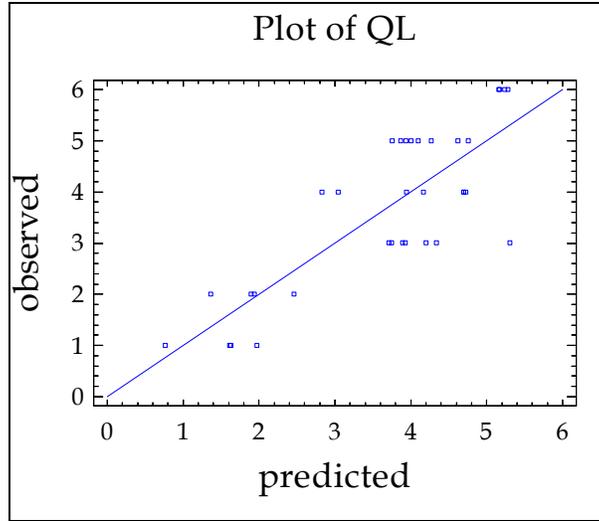


Figure 4-11 Plot of QL between Predicted Vs Observed QL for MNL

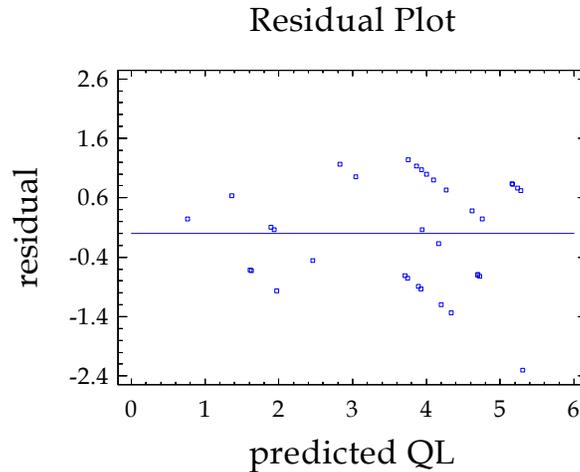


Figure 4-12 Residual Plot between Predicted QL and Residual for MNL

4.7 Minor Right (MNR)

Unfortunately, there are only 18 data points available for model development. A minimum of 2 vehicle in the queue and maximum of 5 vehicles are observed for this lane configuration. Table 4-20 shows summary statistics of data.

Table 4-20 Summary Statistics of MNR Data

	QL	VOL	CONVOL	VOLCONVOL
Count	18	18	18	18
Average	3.22222	56.1667	304.167	17124.6
Variance	1.12418	77.9118	7107.79	2.75093E7
Standard deviation	1.06027	8.82676	84.3077	5244.94
Minimum	2.0	44.0	141.0	6166.0
Maximum	5.0	71.0	412.0	23323.0
Range	3.0	27.0	271.0	17157.0
Std. skewness	0.871404	0.273118	-0.594507	-1.09394
Std. kurtosis	-0.702281	-0.991707	-0.862272	-0.426578

Only the developed model is shown with out further explanation of the model. This model is not tested for validation, as more data is required to obtain a significant model.

The model for predicting the maximum number of vehicles in the stopped queue is formulated as a Poisson regression equation with only the product of volume and conflicting volume as the independent or explanatory variable.

$$\ln(QL_{MNR}) = \beta_0 + \beta_1 \times (VOL * CONVOL)_{MNR}$$

Where

Ln = Natural Logarithm

QL = Maximum number of vehicles in the stopped queue

VOL = Approach volume for MNR lane configuration

CONVOL = Conflicting volume for MNR

β_0 = Constant

β_1 = Model coefficients corresponds to CONVOL

Poisson Regression is shown below:

$\ln(QL) = 0.225058 + 0.00005316 \times (VOL * CONVOL)$ $QL = e^{(0.225058 + 0.00005316 \times (VOL * CONVOL))}$ <p>Percentage of deviance explained by model = 64.7485</p>
--

4.8 *Summary*

This chapter summarizes the step by step procedure in the development of a model. It is not practically possible to consider all explanatory variables in the model development. Only volume, conflicting volume, either their product or ratio between them, presence of a signal, and presence of a separate turn lane are assumed to have a significant impact on queue. Scatter diagrams of these identified variables show the random phenomenon which triggers the development of Poisson regression models. Model development steps are presented in a detailed manner for each lane configuration except for minor right turn configuration (MNR), due to presence of only 18 data points. Accuracy of these models is validated through data validation presented in the next chapter.

5 Model Validation

Validation gives the estimation of the model accuracy in predicting the maximum number of vehicles in the stopped queue. Validation can be done by using the subset of data prepared for model development process but not used for model building. If there is a possibility to collect separate data under similar conditions where the model is developed, those data will be preferred.

This chapter explains the data collection efforts for model validation. Collected raw data need to be processed to be used in model. For each model category, observed queue lengths are compared with predicted queues to check the consistency. Later, this step is extended to compare other methods with the existing methodology.

5.1 Data Collection

Intersections are chosen to cover good proportions of various lane configurations, geographic regions, functional classifications, and traffic conditions. In total, 25 intersections shown in Table C1 in the Appendix C are used for data collection.

Out of 25 intersections: 17 of them are from Region 1, and 8 are from Region 2. 12 (48%) are within the urban growth boundary, and the remaining 13 (52%) are rural. 24 intersections have either OR or US route as the major approach. Ten of the intersections have an upstream signal within 1000 ft. Thirteen intersections have either an exclusive or two-way left turn lane. Only 7 of the intersections have skewed approaches. None of the intersection approaches are off-set from the major approach. 17 intersections are 3 legged (68%), and 8 of them are 4 legged intersections. Finally, two intersections major approaches have flaredness.

Table C2 represents the time frame of the data collection. All the data were collected in 2010, on typical weekdays of either last week of August or the first week of September, but before the Labor Day Weekend.

5.2 MJL

There are 41 data points available for validation. Although there are many indicators of the strength of the model in predicting the intended behavior, residual analysis is primarily to document the accuracy. The difference between the observed and model predicted value is used to assess the model performance. The following Table 5-1 shows the difference between various models. 39% of observed values are exactly predicted by the new model. Gard's equation and the Two-Minute Rule are behind the new model with 22% and 20% matching. If the error is relaxed to either +1 or -1 vehicles, 79% are matched by the new model and nearly the same percentage of match by Gard's and Two-Minute Rule. HCM consistently yields lower estimates. None of the model outputs underestimate queue length by more than 3 vehicles.

Table 5-1 Comparison of Queue Length Estimation Differences for MJL

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation		Observed - Model	
		% Obs	Cum % of Obs	% Obs	Cum % of Obs	% Obs	Cum % of Obs	% Obs	Cum % of Obs
Over Estimated	< = -5	0%	2%	0%	0%	0%	12%	5%	17%
	-4	0%		0%		0%			
	-3	0%		0%		7%		2%	
	-2	2%		0%		5%		10%	
Acceptable	-1	20%	61%	0%	29%	20%	56%	20%	78%
	0	20%		0%		22%		39%	
	1	22%		29%		15%		20%	
Under Estimated	2	17%	37%	29%	71%	10%	32%	5%	5%
	3	12%		12%		15%		0%	
	4	0%		17%		2%		0%	
	> = 5	7%		12%		5%		0%	

5.3 MNLTR

Overall 15 observations are available for this category. 33% of predicted queue lengths are exactly matched with the observed values. Two-Minute rule matched 13% of the observations. The new model matches 60% with a variation of a single vehicle on both sides. For the same variation 53% are matches for the Two-Minute Rule. Gard's equation is overestimating the queues, while the HCM is under predicting. The variation of the error is shown through both Table 5-2.

Table 5-2 Comparison of Queue Length Estimation Differences for MNLTR

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation		Observed - Model	
		% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs
Over Estimated	< = -5	0%	0%	0%	0%	60%	73%	0%	20%
	-4	0%		0%		0%			
	-3	0%		0%		7%			
	-2	0%		0%		7%			
Acceptable	-1	7%	53%	7%	7%	13%	20%	0%	60%
	0	13%		0%		0%			
	1	33%		0%		7%			
Under Estimated	2	13%	47%	27%	93%	7%	7%	13%	20%
	3	33%		47%		0%			
	4	0%		13%		0%			
	> = 5	0%		7%		0%			

5.4 MNLTR

Only 25% of predicted queue lengths are exactly matched with the observed queues, while 42% are matched by the Two-Minute Rule. With a variation of 1 vehicle, almost 92% are matched by Two-Minute Rule. For the same situation, 67% are matched by the developed model. The HCM method is matching 58% of the time and 33% are matched by Gard’s equation. The results are shown in Table 5-3. The predicted model seems to be underestimating. One reason may be only 12 data points are available for data validation purpose. As the sample size increases, there is a good chance of model convergence with the observed values.

Table 5-3 Comparison of Queue Length Estimation Differences for MNLTR

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation		Observed - Model	
		% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs
Over Estimated	< = -5	0%	0%	8%	8%	50%	58%	0%	0%
	-4	0%		0%		0%			
	-3	0%		0%		0%			
	-2	0%		0%		8%			
Acceptable	-1	0%	92%	8%	58%	17%	33%	25%	67%
	0	42%		8%		8%			
	1	50%		42%		8%			
Under Estimated	2	8%	8%	25%	33%	8%	8%	25%	33%
	3	0%		0%		0%			
	4	0%		0%		0%			
	> = 5	0%		8%		0%			

5.5 MNL

Only 10 observations are available for the MNL lane configuration. 50% are exactly matched for the developed model with 90% for one vehicle variation. 70% are matched by Two-Minute Rule, and 60% by Gard's Equation. The HCM methodology underestimates the queue lengths. Error distribution is shown in Table 5-4.

Table 5-4 Comparison of Queue Length Estimation Differences for MNR

Type of Estimation	Difference (against Observed)	Observed - 2 min rule		Observed - HCM method		Observed - Gard's Equation		Observed - Model	
		% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs	% of Obs	Cum % of Obs
Over Estimated	< = -5	0%	0%	0%	0%	0%	40%	0%	10%
	-4	0%		0%		0%			
	-3	0%		0%		40%		0%	
	-2	0%		0%		0%		10%	
Acceptable	-1	0%	70%	0%	0%	20%	60%	40%	90%
	0	10%		0%		20%		50%	
	1	60%		0%		20%		0%	
Under Estimated	2	20%	30%	30%	100%	0%	0%	0%	0%
	3	10%		50%		0%			
	4	0%		20%		0%			
	> = 5	0%		0%		0%		0%	

5.6 Summary

The developed models are predicting consistently closer values. Although error varies from one lane group model to other, the percent of matching varies from 60% to 90% with error of ± 1 vehicle. Importantly, the data set available for most of the lane groups is not sufficient for analysis. As more data is made available, the predictions should be closer.

6 Model Sensitivity Analysis

The sensitivity of the model with respect to changes in the independent or explanatory variables gives an idea how well models will perform for varying conditions. As such, one can check the model reasonableness by the sign of the change of dependent variable for the corresponding change in the explanatory variables. The magnitude of the change will also be obtained. In addition, one can check the limit(s) or range(s) of the independent variables where the model will adequately explain the queue behavior.

This chapter presents the explanations for sensitivity of the models developed in Chapter 4. For each model, limits are set for the independent variables based on the outcome of the model queue lengths which indicate the model stability.

6.1 MJL

Volume (VOL), conflicting volume (CONVOL), presence of upstream signal within 1000 feet of the intersection (SIGNAL), and presence of left turn lane (LT) are considered for modeling queue lengths for major left turns. The developed model is shown below:

$$\ln(QL) = 0.3925 + 0.0059 \times VOL + 0.00104 \times CONVOL + 0.49 * SIGNAL - 0.81 * LT$$
$$QL = e^{(0.3925 + 0.0059 \times VOL + 0.00104 \times CONVOL + 0.49 * SIGNAL - 0.81 * LT)}$$

Percentage of deviance explained by model = 66.959

As VOL and CONVOL increases, QL increases. Presence of an upstream signal increases QL due to vehicle arrivals in platoon. If there is a separate left turn lane, either exclusive or two-way left turn lane, it decreases queue length compared to a shared left turn lane. The bounds for VOL and CONVOL are set by drawing a 2-D contour map of the QL.

As shown in Figure 6-1, as the VOL and CONVOL pair reaches MUTCD 2009 edition warranted volumes given in Table A1, for condition A (Minimum vehicular volume for 2 or more lanes on major street and 2 or more lanes on minor street), a maximum of fifteen vehicles are predicted to be in the stopped queue. For condition B (interruption of continuous traffic condition for the same lane configuration), eleven vehicles at maximum are in the stopped queue condition.

Volumes exceeding these points trigger a substantial increase in queue lengths. As such, unacceptable queue lengths are obtained for volumes greater than 300 VPH and corresponding conflicting volumes of 2000 VPH. Beyond these points the model is unstable for queue length prediction. Caution - the outcomes

shown in Figure 6-1 are obtained by assuming an upstream signal and shared left turn lane, which will test the worst possible scenario.

6.2 MNLTR

QL on a single shared left, through, and right turn movement on minor approach is affected by volume, conflicting volume, and their product.

$$\ln(QL) = -0.7844 + 0.01636 \times VOL + 0.0006 \times CONVOL - 0.0000043 \times (VOL * CONVOL)$$

$$QL = e^{(-0.7844 + 0.01636 \times VOL + 0.0006 \times CONVOL - 0.0000043 \times (VOL * CONVOL))}$$

Percentage of Deviation Explained by Model = 71.643

Signs of the independent variables show reasonableness in the models, as a positive change in the VOL and CONVOL increases QL. But the occurrence of VOL and the corresponding CONVOL as a product triggers a decrease in QL, due to the fact that a lower approach volume may not yield larger queue lengths, rather triggers more waiting time in the stopped queue.

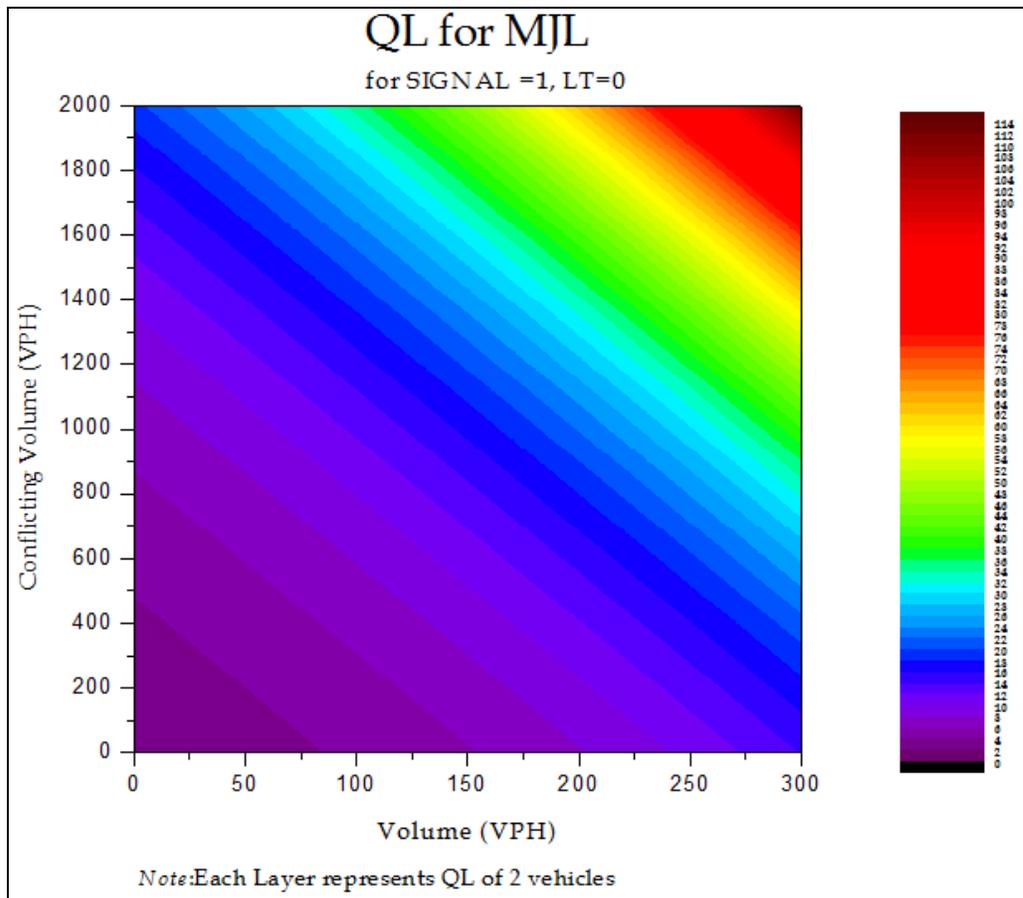


Figure 6-1 Plot of QL for VOL and CONVOL at Major Left Turn (MJL)

As approach volume increase to capacity, approach volume has a higher impact on queue than the conflicting volume. This behavior can be seen in Figure 6-2. As Highway Capacity Manual (HCM) limits the maximum 95th percentile queue lengths to be 100, the limits for VOL and CONVOL are set such that predicted queue length is not exceeding 100 vehicles. This is not a bad idea because, if the volume on the subject approach and the corresponding volume on major street reach the MUTCD Chapter 4C, section 4C.02 thresholds, it will warrant the installation of a signal.

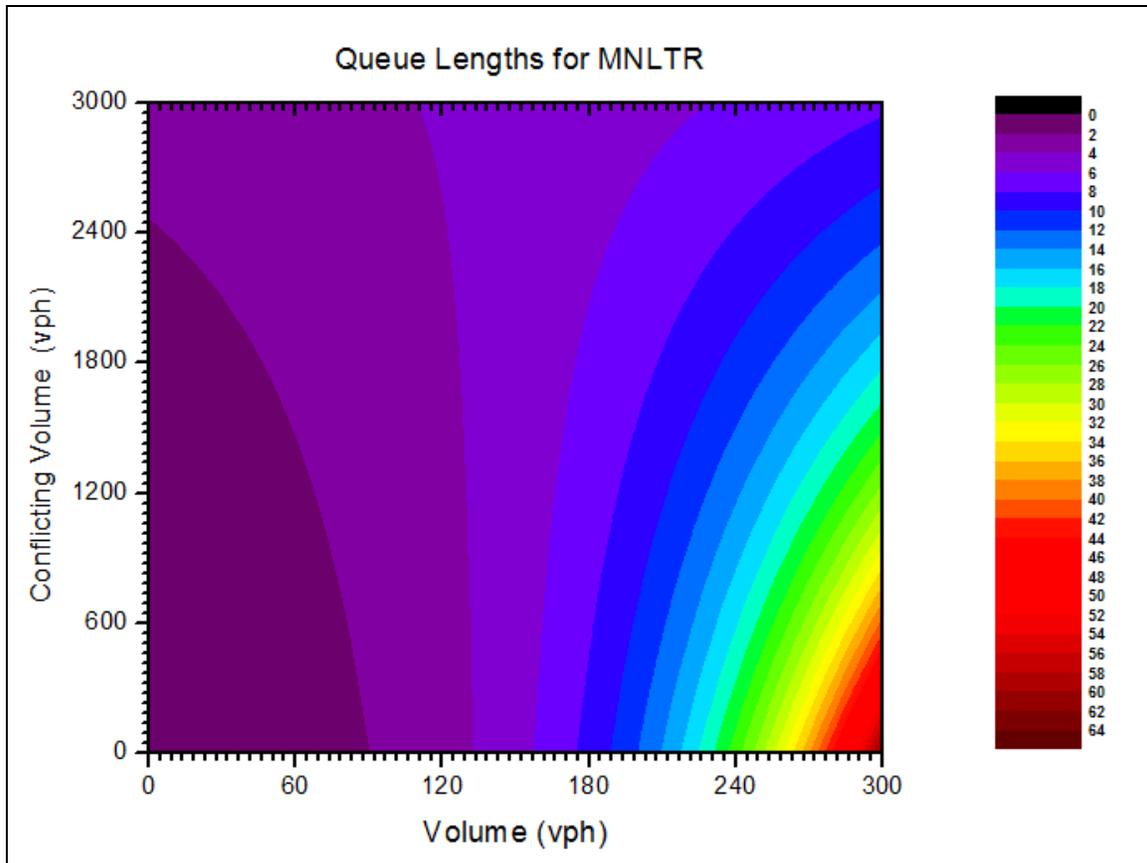


Figure 6-2 Plot of QL for VOL and CONVOL at Minor LTR (MNLTR)

6.3 MNLTR

QL on a single shared left, through, and right turn movement on minor approach is affected by volume, conflicting volume, and their product.

$$\ln(QL) = -0.6319 + 0.0173 \times VOL + 0.00066 \times CONVOL - 0.000007913 (VOL * CONVOL)$$

$$QL = e^{(-0.6319 + 0.0173 \times VOL + 0.00066 \times CONVOL - 0.000007913 (VOL * CONVOL))}$$

Percentage of deviance explained by model = 69.25

Signs of the independent variables show reasonableness in the models, as a positive change in the VOL and CONVOL increases QL. But the occurrence of VOL and the corresponding CONVOL as a product triggers a decrease in QL, due to the fact that lower approach volume may not yield larger queue lengths, rather triggers more waiting time in the stopped queue. As approach volume increase to capacity, approach volume has higher impact on queue than the conflicting volume. This behavior can be seen in Figure 6-3. As Highway Capacity Manual (HCM) limits the maximum 95th percentile queue lengths to be 100, the limits for VOL and CONVOL are set such that predicted queue length is not exceeding 100 vehicles. This is not a bad idea because, if the volume on the subject approach and the corresponding volume on major street reach the MUTCD Chapter 4C, section 4C.02 thresholds, it will warrant the installation of a signal.

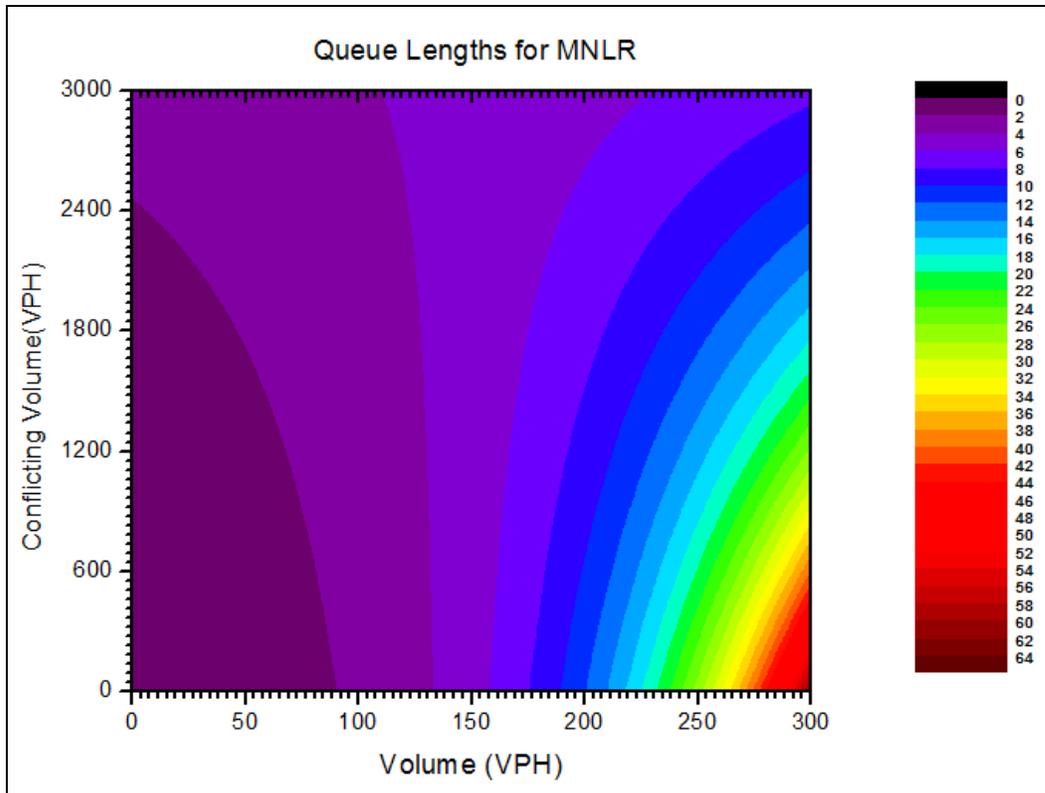


Figure 6-3 Plot of QL for VOL and CONVOL at Minor LR (MNLR)

6.4 MNL

Poisson Regression developed is shown below:

$$\begin{aligned} \ln(QL) &= 1.7934 - 0.025 \times (CONVOL/VOL) \\ QL &= e^{(1.7934 - 0.025 \times (CONVOL/VOL))} \\ \text{Percentage of deviance explained by model} &= 69.404 \end{aligned}$$

As the volume and conflicting volume increases, QL increases as shown in Figure 6-4. Beyond the volume of 300 VPH and conflicting volume of 3000 VPH, queue lengths are not realistically represented by the model. These are used as the limits for the models. MNL model is developed using only 34 data points, which may limit the strength of the model. This is evident from MNL model form, which has 1.7934¹³ as a constant. So, the output is greatly affected by a constant value rather than variation of explanatory variables. MNL model needs to be improved by collecting more data.

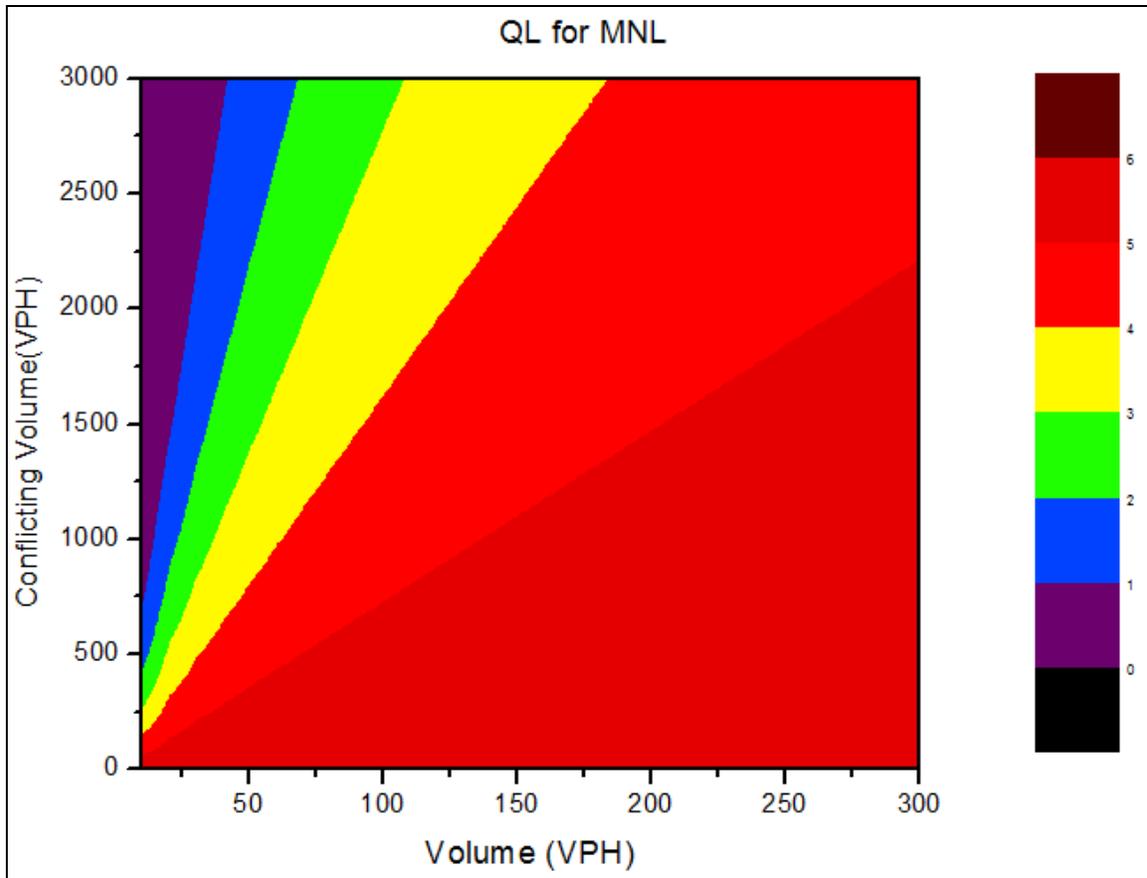


Figure 6-4 Plot of QL for VOL and CONVOL at Minor Left (MNL)

6.5 Summary

Sensitivity Analysis is used to test the model for the all possible ranges of the input variables. During the model development only certain range for input variables are represented. Sensitivity analysis gives an opportunity, as explained in the above sections, to test the model behavior for most of the combinations of inputs. Following table summaries models for each lane groups with the limitation to the input variables.

¹³ In the absence of VOL, and CONVOL, $QL = e^{1.7934} \cong 6$ vehicles

7 Summary, Conclusions & Scope for Future Study

7.1 Summary

The following table summarizes the developed models, the applicable ranges for input data, and the percentage of deviation for each model:

VOL = Traffic flow rate on the subject approach in vehicles per hour
 CONVOL = Conflicting traffic flow rate calculated according HCM methodology, expressed as vehicles per hour
 SIGNAL = Presence of Upstream Signal within 1000 ft of the intersection, Applicable for Major Left Turn only, 1 if there is a signal, otherwise 0
 LT = Presence of a separate left turn lane, Applicable for Major Left Turn only (1 if there is an exclusive left turn lane/ median left turn lane/ two-way left turn lane, otherwise 0)

Lane Group	Model Equation & Ranges	Percent of Deviation
MJL	$QL = e^{(0.3925+0.0059 \times VOL+0.00104 \times CONVOL+0.49 \times SIGNAL-0.81 \times LT)}$ <p>VOL = (0, 300] ; CONVOL = (0,2000] SIGNAL = 0 or 1 ; LT = 0 or 1</p>	66.96
MNLTR	$QL = e^{(-0.7844+0.01636 \times VOL+0.0006 \times CONVOL-0.0000043 \times (VOL \times CONVOL))}$ <p>VOL = (0, 300] ; CONVOL = (0,3000]</p>	71.64
MNLR	$QL = e^{(-0.6319+0.0173 \times VOL+0.00066 \times CONVOL-0.000007913 \times (VOL \times CONVOL))}$ <p>VOL = (0, 300] ; CONVOL = (0,3000]</p>	69.25

7.2 Conclusions

Following conclusions are made from this study:

- The Two-Minute Rule performs better than other existing methods except for the Major Left Turn (MJL) configuration where Gard’s Equation does better for the reason that the Two-Minute Rule does not include opposing volume.
- Existing methods are not exactly predicting queue lengths for more than 50 percent of the cases.
- A Poisson regression model is developed to improve the queue length estimations.
- The developed models accurately predict more than 65 percent of the cases.

- Improvements to the model predictions may be achieved by expanding the data sample size.
- As presented these models are a proto type, one may not adopt this model until they are validated for a wide range of conditions.

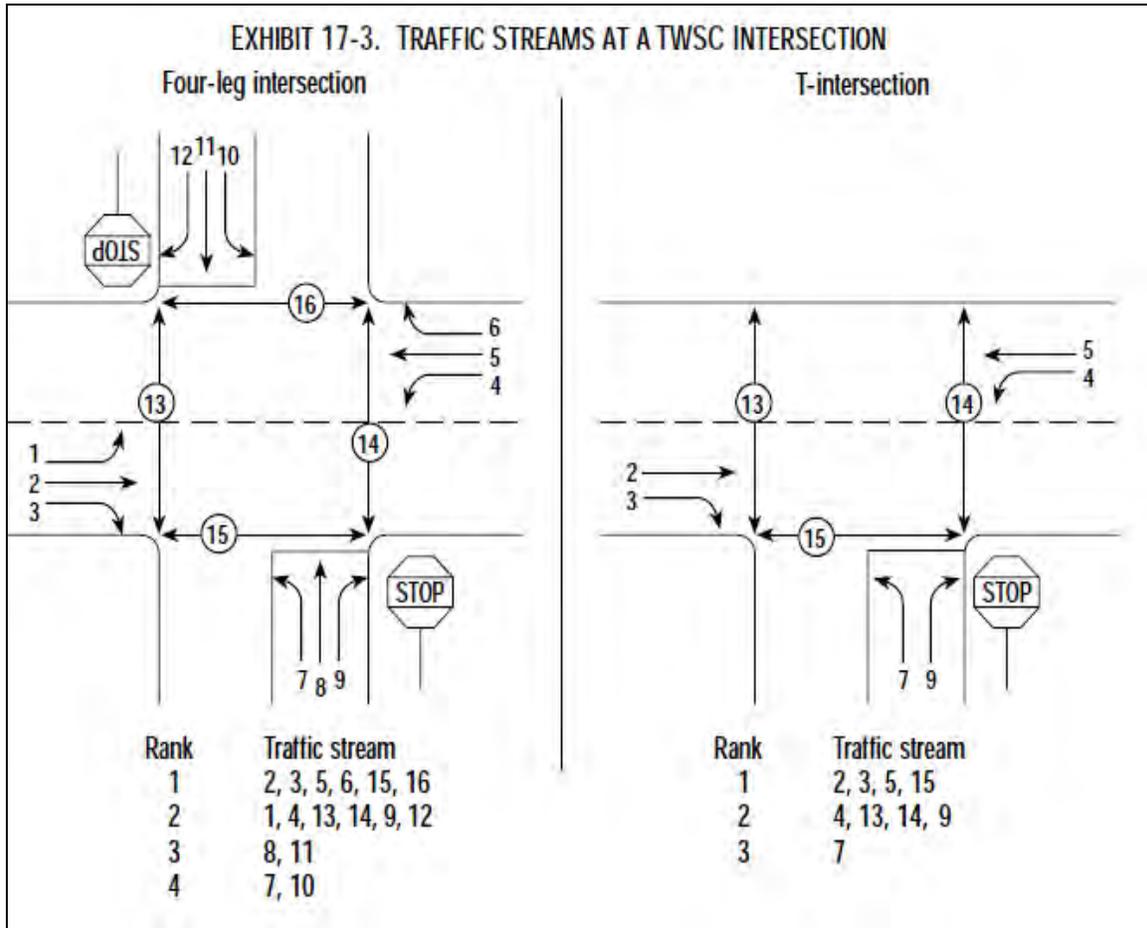
7.3 *Future Study Scope*

It is a well known fact that model development is an iterative process, as such there is always room for improvement.

- The current study may be elaborated with expanded data sets. Special attention need to be given to collect MNL lane configuration data to refine the model performance.
- A good proportion of data groups are adopted for study, but expansion to different geographic regions, highway networks, traffic loads as time of day and seasonal variations, multi-year data sets is possible
- The model form may be scoped to suit various traffic and queue conditions
- The methods of model development may be varied to capture dynamics of queues
- Even explanatory variables and their combination may be studied

Appendix A - HCM & MUTCD Exhibits

Figure A1 Traffic Streams at a Two-Way STOP Controlled Intersection
 (Source: HCM, Exhibit 17-3)



**Figure A2 Method for Computing Lane Group Conflicting Flow Rates
(Source: HCM, Exhibit 17-4)**

EXHIBIT 17-4. DEFINITION AND COMPUTATION OF CONFLICTING FLOWS	
Subject Movement	Subject and Conflicting Movements Conflicting Traffic Flows, $v_{c,x}$
Major LT	$v_{c,1} = v_5 + v_6^{[3]} + v_{16}$ $v_{c,4} = v_2 + v_3^{[3]} + v_{15}$
Minor RT	$v_{c,9} = \frac{v_2^{[2]}}{N} + 0.5v_3^{[1]} + v_{14} + v_{15}$ $v_{c,12} = \frac{v_9^{[2]}}{N} + 0.5v_6^{[1]} + v_{13} + v_{16}$
Minor TH	<p>Stage II</p> $v_{c,II,8} = 2v_4 + v_5 + v_6^{[3]} + v_{16}$ $v_{c,II,11} = 2v_1 + v_2 + v_3^{[3]} + v_{15}$
	<p>Stage I</p> $v_{c,I,8} = 2v_1 + v_2 + 0.5v_3^{[1]} + v_{15}$ $v_{c,I,11} = 2v_4 + v_5 + 0.5v_6^{[1]} + v_{16}$
Minor LT	<p>Stage II</p> $v_{c,II,7} = 2v_4 + \frac{v_5}{N} + 0.5v_6^{[6]} + 0.5v_{12}^{[4,5]} + 0.5v_{11} + v_{13}$ $v_{c,II,10} = 2v_1 + \frac{v_5}{N} + 0.5v_3^{[6]} + 0.5v_9^{[4,5]} + 0.5v_8 + v_{14}$
	<p>Stage I</p> $v_{c,I,7} = 2v_1 + v_2 + 0.5v_3^{[1]} + v_{15}$ $v_{c,I,10} = 2v_4 + v_5 + 0.5v_6^{[1]} + v_{16}$

The following footnotes apply to Exhibit 17-4:

[1] If there is a right-turn lane on the major street, v_3 or v_6 should not be considered.

[2] If there is more than one lane on the major street, the flow rates in the right lane are assumed to be v_2/N or v_6/N , where N is the number of through lanes. The user can specify a different lane distribution if field data are available.

[3] If right-turning traffic from the major road is separated by a triangular island and has to comply with a yield or stop sign, v_6 and v_3 need not be considered.

[4] If right-turning traffic from the minor road is separated by a triangular island and has to comply with a yield or stop sign, v_9 and v_{12} need not be considered.

[5] Omit v_9 and v_{12} for multilane sites, or use one-half their values if the minor approach is flared.

[6] Omit the farthest right-turn v_3 for subject movement 10 or v_6 for subject movement 7 if the major street is multilane.

Figure A3 Lane Groups (Source: HCM, Exhibit 16-5)

EXHIBIT 16-5. TYPICAL LANE GROUPS FOR ANALYSIS		
Number of Lanes	Movements by Lanes	Number of Possible Lane Groups
1	LT + TH + RT	①
2	EXC LT TH + RT	②
2	LT + TH TH + RT	① OR ②
3	EXC LT TH TH + RT	② OR ③

Table A1 MUTCD Warrant 1 Table (Source: 2009 MUTCD, section 4C.02)

Table 4C-1. Warrant 1, Eight-Hour Vehicular Volume

Condition A—Minimum Vehicular Volume

Number of lanes for moving traffic on each approach		Vehicles per hour on major street (total of both approaches)				Vehicles per hour on higher-volume minor-street approach (one direction only)			
Major Street	Minor Street	100% ^a	80% ^b	70% ^c	56% ^d	100% ^a	80% ^b	70% ^c	56% ^d
1	1	500	400	350	280	150	120	105	84
2 or more	1	600	480	420	336	150	120	105	84
2 or more	2 or more	600	480	420	336	200	160	140	112
1	2 or more	500	400	350	280	200	160	140	112

Condition B—Interruption of Continuous Traffic

Number of lanes for moving traffic on each approach		Vehicles per hour on major street (total of both approaches)				Vehicles per hour on higher-volume minor-street approach (one direction only)			
Major Street	Minor Street	100% ^a	80% ^b	70% ^c	56% ^d	100% ^a	80% ^b	70% ^c	56% ^d
1	1	750	600	525	420	75	60	53	42
2 or more	1	900	720	630	504	75	60	53	42
2 or more	2 or more	900	720	630	504	100	80	70	56
1	2 or more	750	600	525	420	100	80	70	56

^a Basic minimum hourly volume

^b Used for combination of Conditions A and B after adequate trial of other remedial measures

^c May be used when the major-street speed exceeds 40 mph or in an isolated community with a population of less than 10,000

^d May be used for combination of Conditions A and B after adequate trial of other remedial measures when the major-street speed exceeds 40 mph or in an isolated community with a population of less than 10,000

Appendix B - Data Collection Sites for Model Development

Table B1 Description of Study Intersection Used for Model Development

Sl. No	Name	City	Region	Geographic Area(R/U)	No of Legs (3/4)	Presence of Skewness/ Presence of curve at the intersection	Presence of Offset on Minor Approaches	No of Lanes on Major Street (2/3/4/5)	Functional Classification of the Major ST	Functional Classification of the Minor ST	Major Approach				
											Approach Speed	Presence of Upstream Signal	Presence of TWLT Lane/ Median Left Turn Lane	Presence of Right Turn Lane	Presence of Flared Right Turn Lane
1	Nevada ST and Oak ST	Ashland	3	U	4	0	0	2	C	L	25	0	0	0	0
2	Tolman Creek Rd. and Mistletoe Rd + Takelma Wy	Ashland	3	U	4	1	0	2	C	L	25	0	0	0	0
3	OR 99 E and NE Carl Rd	Woodburn	2	U	3	1	0	5	OR	C	45	0	1	0	0
4	US 101 and 20th ST	Reedsport	3	U	4	0	0	4	US	C	30	0	0	0	0
5	US 730 and Umatilla River Rd	Umatilla	5	U	3	0	0	3	US	C	25	0	1	0	0
6	OR 99 E and E Cleveland ST	Woodburn	2	U	4	1	1	4	OR	C	35	0	0	0	0
7	OR 214 / OR 211 and Lawson Avenue	Woodburn	2	U	3	0	0	3	OR	C	30	1	1	0	0
8	OR 99 E with Industrial Ave and Mc Laren School Rd NE	Woodburn	2	R	4	1	0	5	OR	L	45	1	1	0	1
9	US 97 vs Lakeport Blvd	Klamath Falls	4	R	3	1	0	3	US	L	50	0	1	0	0
10	OR 99 E and Food Services Road	Woodburn	2	R	3	1	0	2	OR	L	45	0	0	0	0
11	US 20 and 17th ST	Philomath	2	U	4	0	0	4	US/OR	C	30	1	0	0	0
12	OR 20 and Dead Indian Memorial Road	Ashland	3	R	3	1	0	2	OR	C	35	0	0	0	1
13	US 395 and Power City Rd	Umatilla	5	R	4	1	1	5	US/OR	C	50	0	1	0	1
14	OR 99 and (W Hersey Road & Wimer ST)	Ashland	3	U	4	1	1	4	OR	C	25	0	0	0	0
15	US 20 (OR 34) and (S 7th ST & N 7th ST)	Philomath	2	U	3	0	0	2	US	C	35	0	0	0	0

Note: U=Urban, R=Rural, OR= Oregon Route, US = US Route, C= Collector, L=Local, C/L = Collector/Local, N/A = Not available, 0/1 = Flag showing Yes/No

Table B2 Periods of Data Collection Used for Model Development

Sl. No	Name	City	Region	Data Collection Date	Data Collection Day	Data Collection Time	Data Collection Duration (Hrs)
1	Nevada ST and Oak ST	Ashland	3	9/14/2009	Monday	3 - 6 PM	3
2	Tolman Creek Rd. and Mistletoe Rd + Takelma Wy	Ashland	3	9/16/2009	Wednesday	4 - 6 PM; 7-8 AM	3
3	OR 99 E and NE Carl Rd	Woodburn	2	2/23/2010	Tuesday	3 - 6 PM ; 6-9 AM	6
4	US 101 and 20th ST	Reedsport	3	11/4/2009	Wednesday	3 - 6 PM ; 7:30-9 AM	4.5
5	US 730 and Umatilla River Rd	Umatilla	5	1/13/2010	Wednesday	3 - 6 PM ; 7:30-9 AM	4.5
6	OR 99 E and E Cleveland ST	Woodburn	2	2/23/2010	Tuesday	3 - 6 PM ; 6-9 AM	6
7	OR 214 / OR 211 and Lawson Avenue	Woodburn	2	3/1/2010	Monday	3 - 6 PM ; 6-9 AM	6
8	OR 99 E with Industrial Ave and Mc Laren School Rd NE	Woodburn	2	2/22/2010	Monday	3 - 6 PM ; 6-9 AM	6
9	US 97 vs Lakeport Blvd	Klamath Falls	4	4/7/2010	Wednesday	3 - 6 PM	3
10	OR 99 E and Food Services Road	Woodburn	2	2/22/2010	Monday	3 - 6 PM ; 6-9 AM	6
11	US 20 and 17th ST	Philomath	2	3/9/2010	Tuesday	3 - 6 PM ; 6-9 AM	6
12	OR 20 and Dead Indian Memorial Road	Ashland	3	9/15/2009	Tuesday	3 - 6 PM	3
13	US 395 and Power City Rd	Umatilla	5	1/19/2010	Tuesday	3 - 6 PM ; 6-9 AM	6
14	OR 99 and (W Hersey Road & Wimer ST)	Ashland	3	9/14/2009	Monday	2 - 6 PM	4
15	US 20 (OR 34) and (S 7th ST & N 7th ST)	Philomath	2	3/9/2009	Tuesday	3 - 6 PM ; 6-9 AM	6

Appendix C - Data Collection Sites for Model Validation

Table C1 Description of Study Intersections Used for Data Validation

Sl.No	Name	Region	No of Legs (3/4)	Presence of Skewness/ Presence of curve	Presence of Offset on Minor Approaches	Geographic Area(R/U/SU)	No of Lanes on Major Street (2/4/4/5)	Functional Classification of the Major ST	Functional Classification of the Minor ST	Major Approach				
										Approach Speed	Presence of Upstream Signal	Presence of TWLT / Exclusive Left	Presence of Right Turn Lane	Presence of Flared Right Turn Lane
1	US 30 and NE 185th Ave	1	3	1	0	U	4	US	C	40	1	0	0	0
2	NE 185th Ave and NE Portal Way	1	4	1	0	U	4	C	L	40	0	1	0	1
3	US 30 and NE 172ND PL	1	3	0	0	U	4	US	L	N/A	0	0	0	0
4	US 26 and SE 79th Ave	1	4	0	0	U	5	US	C/L	35	1	1	0	0
5	US 26 / SE 99th Ave	1	3	0	0	U	2	US	L	35	1	0	0	0
6	US 26 / SE 130th Ave	1	3	0	0	U	2	US	L	35	1	0	0	0
7	OR 99 E / SE HULL AVE	1	4	0	0	U	5	OR	C	40	1	1	0	0
8	OR 99 E / SE VINEYARD RD	1	4	0	0	U	5	OR	C	40	1	1	0	0
9	OR 99 E / SE HOLLY AVE	1	3	0	0	U	5	OR	L	40	1	1	0	0
10	OR 224 & OR 212 / SE 106th ST	1	3	0	0	U	5	OR	L	N/A	1	1	0	0
11	OR 224 & OR 212 / SE 114th AVE	1	3	0	0	U	5	OR	L	N/A	0	1	0	0
12	OR 224 & OR 212 / SE 122nd AVE	1	3	0	0	U	5	OR	C	45	1	1	0	0
13	OR 22 W / PERRYDALE RD	2	3	1	0	R	2	OR	L	N/A	0	1	0	0
14	OR 22 W / DOAKS FERRY RD	2	3	1	0	R	5	OR	C	N/A	0	1	0	0
15	OR 221 / DOAKS FERRY RD NW	2	3	1	0	R	2	OR	C	N/A	1	1	0	0
16	OR 221 / SE PALMER CREEK RD	2	3	1	0	R	2	OR	C	50	0	0	0	0
17	OR 219/ SE FARMINGTON RD	1	3	0	0	R	3	OR	OR	55	0	0	1	0
18	OR 219/ SE SCHOLLS FERRY RD(OR 210)	1	3	1	0	R	2	OR	OR	45	0	0	0	1
19	OR 219/ BELL RD / N VALLE RD	2	4	0	0	R	2	OR	C/L	N/A	0	0	0	0
20	OR 219/ SW UNGER RD	1	3	0	0	R	3	OR	C	45	0	1	0	0
21	OR 219/ Tongue Ln	1	3	0	0	R	4	OR	C	N/A	0	1	0	0
22	OR 211 / S Kropff Rd or Canby - Marquam Rd	2	4	0	0	R	2	OR	OR	N/A	0	0	1	0
23	OR 211 / S Meridian Rd	2	4	0	0	R	2	OR	C	55	0	0	0	0
24	OR 214 / Howell Prairie Rd	2	3	0	0	R	2	OR	C	40	0	0	0	0
25	OR 213 / S CARUS RD	1	4	0	0	R	2	OR	C	20	0	0	0	0

Note:

U=Urban, R=Rural, OR= Oregon Route, US = US Route, C= Collector, L=Local, C/L = Collector/Local,

N/A = Not available, 0/1 = Flag showing Yes/No

Table C2 Duration of Traffic Counts Used for Data Validation

Sl.No	Name	Region	No of Legs (3/4)	Duration	Date	Day	Time of Day
1	US 30 and NE 185th Ave	1	3	1 Hr	8/24/2010	Tuesday	8:25 AM - 9:25 AM
2	NE 185th Ave and NE Portal Way	1	4	1 Hr	8/24/2010	Tuesday	9:40 AM - 10:41 AM
3	US 30 and NE 172ND PL	1	3	1 Hr	8/24/2010	Tuesday	11:07 AM - 12:07 AM
4	US 26 and SE 79th Ave	1	4	1 Hr	8/24/2010	Tuesday	1:22 PM - 2:22 PM
5	US 26 / SE 99th Ave	1	3	1 Hr	8/24/2010	Tuesday	2:40 PM - 3:40 PM
6	US 26 / SE 130th Ave	1	3	1 Hr	8/24/2010	Tuesday	4:05 PM - 5:05 PM
7	OR 99 E / SE HULL AVE	1	4	1 Hr	8/25/2010	Wednesday	8:20 AM - 9:20 AM
8	OR 99 E / SE VINEYARD RD	1	4	1 Hr	8/25/2010	Wednesday	9:49 AM - 10:49 AM
9	OR 99 E / SE HOLLY AVE	1	3	1 Hr	8/25/2010	Wednesday	11:01 AM - 12:01 AM
10	OR 224 & OR 212 / SE 106th ST	1	3	1 Hr	8/25/2010	Wednesday	1:18 PM - 2:18 PM
11	OR 224 & OR 212 / SE 114th AVE	1	3	1 Hr	8/25/2010	Wednesday	2:49 PM - 3:49 PM
12	OR 224 & OR 212 / SE 122nd AVE	1	3	1 Hr	8/25/2010	Wednesday	4:00 PM - 5:00 PM
13	OR 22 W / PERRYDALE RD	2	3	1 Hr	8/26/2010	Thursday	8:56 AM - 9:56 AM
14	OR 22 W / DOAKS FERRY RD	2	3	1 Hr	8/26/2010	Thursday	10:14 AM - 11:14 AM
15	OR 221 / DOAKS FERRY RD NW	2	3	1 Hr	8/26/2010	Thursday	12:11 PM - 1:26 PM
16	OR 221 / SE PALMER CREEK RD	2	3	1 Hr	8/26/2010	Thursday	1:40 PM - 2:40 PM
17	OR 219/ SE FARMINGTON RD	1	3	2 Hr	8/31/2010	Tuesday	7:45 AM - 9:45 AM
18	OR 219/ SE SCHOLLS FERRY RD (OR 210)	1	3	1 Hr	8/31/2010	Tuesday	10:21 AM - 11:21AM
19	OR 219/ BELL RD / N VALLE RD	2	4	1 Hr	8/31/2010	Tuesday	12:46 PM - 1:46 PM
20	OR 219/ SW UNGER RD	1	3	1 Hr	8/31/2010	Tuesday	2:44 PM - 3:44 PM
21	OR 219/ Tongue Ln	1	3	2 Hr	8/31/2010	Tuesday	4:00 PM - 6:00 PM
22	OR 211 / S Kropff Rd	2	4	1 Hr	9/1/2010	Wednesday	7:10 AM - 8:10 AM
23	OR 211 / S Meridian Rd	2	4	1 Hr	9/1/2010	Wednesday	8:27 AM - 9:27 AM
24	OR 214 / Howell Prairie Rd	2	3	1 Hr	9/1/2010	Wednesday	10:28 AM - 11:28 AM
25	OR 213 / S CARUS RD	1	4	1 Hr	9/1/2010	Wednesday	4:40PM - 5:40 PM