Improving Highway Travel Time Estimation in FSUTMS by Considering Intersection Delays

Contract No. BD015-15
Final Report

Submitted to

The Florida Department of Transportation
Research Center
605 Suwannee Street, MS 30
Tallahassee FL 32399

Submitted by

Dr. Fang Zhao, Ph.D., P.E.
Zhen Ding, M.S., E.I.
Lehman Center for Transportation Research
Department of Civil and Environmental Engineering
College of Engineering and Computing
Florida International University
Miami, FL 33199
(305) 348-3821, (305) 348-2802 (fax)
zhaof@fiu.edu

August 2006

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Final Report for BD015-15		

4. Title and Subtitle	5. Report Date
Improving Highway Travel Time Estimation in FSUTMS by Considering Intersection Delays	August 2006
	6. Performing Organization Code
7. Author(s) Fang Zhao and Zhen Ding	8. Performing Organization Report No.
9. Performing Organization Name and Address Lehman Center for Transportation Research, Department of Civil and Environmental Engineering, Florida International University, Miami, Florida 33199	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Research Office Florida Department of Transportation 650 Suwannee Street, MS 30 Tallahassee, Florida 32399-0450	13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Planning models do not usually consider intersection delays when calculating path travel times during traffic assignment. However, since intersection delays often comprise a considerable portion of total travel time in urban areas, ignoring them will result in inaccuracy in model results. This study develops delay estimation models for five major types of intersections based on artificial neural networks (ANNs), which take all movement volumes at an intersection as input to estimate the delay for each of the movements. Assuming signal plans are optimized for all intersections based on traffic conditions, TRANSYT-7F is employed to both optimize the signal plan and simulate traffic delays at an intersection to generate the data for training the ANN models. The results show that the models are able to predict intersection delays with an accuracy level suitable for planning purposes.

17. Key Word	18. Distribution Statement

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

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

The contents of this report reflect the views of the authors and do not necessarily reflect the official views or policies of the Florida Department of Transportation. This report does not constitute a standard, specification, or regulation.

ACKNOWELDGEMENTS

This study was sponsored by a research grant from the Florida Department of Transportation Systems Planning Office through the Research Center. This support is gratefully acknowledged. We would like to thank Mr. Yongqiang Wu, the project manager, for his support and guidance; Mr. Shi-Chiang Li, Systems Planning Manager, Office of Planning and Environmental Management, FDOT, District IV, for his contributions; and Dr. Lee-Fang Chow for her helpful suggestions.

TABLE OF CONTENTS

EXECUTIVE SUMMARY viii

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

(%RMSE)	Percent Root Mean Squared Error
AADT	Annual Average Daily Traffic
ADT	Average Daily Traffic
ANN	Artificial Neural Network
FDOT	Florida Department of Transportation
FSUTMS	Florida Standard Urban Transportation Modeling Structure
FT	Facility Type
FTI	Florida Traffic Information
GIS	Geographical Information Systems
MAE	Mean Absolute Error
PTMS	Portable Traffic Monitoring Sites
RMSE	Root-Mean-Square Error
v/c	Ratios of Traffic Volume to Capacity

EXECUTIVE SUMMARY

Advances in computer hardware and software have led to an increasing level of sophistication in travel demand forecasting. Many researchers have taken advantage of the improved computing power to refine travel demand models. Traditional travel demand models rarely consider intersection delays when estimating travel time. This is because modeling intersection delays is challenging due to the complexity in roadway geometry, signal plans, and vehicle movements. The Highway Capacity Manual (HCM) method is commonly used for calculating intersection delays in planning studies. However, a regional travel model, typically having information only on the number of lanes and link capacity, often lacks detailed information for either a base year or forecast year to apply the HCM method. For instance, the HCM method requires the availability of signal plans for estimating delays. However, while signal plans may be coded in a base-year model, this information is unavailable for a future year. Because intersection delays constitute a considerable portion of total travel time in urban areas, especially on arterials during peak hours, ignoring them will introduce inaccuracies in model results.

Currently in FSUTMS, trips are assigned to links based on link impedance and penalties for turning movements. A generalized method that appropriately incorporates intersection delays is desired for short-term model improvement. This study is designed for this purpose. Twenty seven types of intersections are identified based on the street network from the Gainesville 2000 model. Five typical intersection types are selected, for which delay models are developed as a proof of concept. The artificial neural network (ANN) technology is employed to build the models, which estimate delays based on limited inputs including movement volumes and capacities.

To build the ANN delay models, a large number of examples are required to allow the ANN models to "learn" to associate a given set of intersection movements with the corresponding delays. These examples are generated by TRANSYT-7F, which optimizes signal plans for given traffic conditions and produces delay estimates. For each of the five intersection types, two ANN delay models are designed to estimate delays for left-turning and through movements, respectively. The ANN delay models are able to predict delays with the %RMSE statistic lower than 26%. Larger errors tend to occur for intersections of higher class facility types. For instance, errors in delay estimates for an intersection of two divided arterials tend to be larger than those for one with cross roads being an undivided arterial and a local collector. Moreover, the through traffic delay models have a higher accuracy than the left-turning movement delay models. The ANN models are compared to multiple linear regression models. The results show that the ANN models are superior to the regression models.

Future research may be directed at combining the ANN delay models with a FSUTMS model. It is expected that the accuracy of traffic assignment may be improved with intersection delays considered. Several issues also need to be further investigated. For instance, signal coordination is widely used in urban areas on arterials where traffic volumes are high. The model performance under oversaturated conditions also needs improvement by possibly considering the degree of situation. The issue of upstream congestion effects on down stream intersections may be addressed by incorporating intersection spacing and area type into the models.

1. INTRODUCTION

A recently completed project, titled "Calibration of Highway/Transit Speed Relationships for Improved Transit Network Modeling in FSUTMS," developed a transit link travel time estimation method (Zhao and Li, 2005). This method calculates the transit link travel time based on highway link travel time and the additional time needed to load and discharge passengers at bus stops on a link based on the passenger boarding and alighting activities. However, one problem with the current highway travel time estimation in FSUTMS is that delays at intersections are not accounted for. On urban arterials, signalized intersections contribute to a significant portion of total travel time, especially under congested conditions during peak hours. Ignoring the intersection delays results in inaccuracy in both highway and transit travel times, causes inconsistencies between the two, and affects the accuracy of model results. To further improve the accuracy of highway and transit travel time estimation, delays at intersections need to be considered.

Modeling intersection delays is challenging because of the complexity of roadway geometry, signal plans, and movements. The Highway Capacity Manual (HCM) method is a commonly used method for calculating intersection delays in panning studies. However, the detailed information required by the HCM method is costly to prepare for a regional travel model, which only has information on the number of lanes and capacity of a network link. The detailed information required by the HCM method is also unavailable for a forecast year. A generalized method that appropriately considers intersection delays without overburdening transportation planners with the task of detailed intersection coding or requiring the use of a micro-simulation based assignment procedure is desired for short-term model improvement.

This study aims at improving link travel time estimation by developing intersection delay models that may be incorporated into FSUTMS to estimate highway and transit travel times. The objectives to be achieved in this research include:

1. Understanding the state-of-the-art in intersection modeling in travel demand models.
2. Evaluating the appropriateness of different traffic engineering software for the purpose of determining intersection delays under different combinations of traffic conditions and intersection geometry.
3. Developing models that are capable of predicting intersection delays based on traffic conditions and simplified intersection geometry.
4. Evaluating model performance and developing recommendations for subsequent research.

In this research, twenty seven types of intersections are identified based on the street network from the Gainesville 2000 travel model. Five typical intersection types are selected, for which delay models are developed. The artificial neural network (ANN) technology is employed to build the models, which estimate delays based on limited inputs including movement volumes and link capacities.

To build the ANN delay models, a large number of examples are required to allow the ANN models to "learn" to correctly associate a given set of intersection movements with the corresponding delays. These examples are generated by TRANSYT-7F, which optimizes signal plans for given traffic conditions and produces delay estimates. For each of the five intersection types, two ANN delay models are designed to estimate delays for left-turning and through movements, respectively. The models are evaluated and found to produce reasonable estimates of intersection delays for demand modeling purposes.

In the remainder of this report, literature in the areas of intersection delay modeling is first reviewed in Chapter 2. An overview of the methodology is provided in Chapter 3. Chapter 4 introduces the methodology and the classification of intersections. Chapter 5 presents a procedure for generating the needed data for model development. Chapter 6 describes the ANN model architecture and calibration, and discusses the performance of the ANN delay models. Finally, Chapter 7 provides conclusions and suggests directions for future research.

2. LITERATURE REVIEW

2.1 Introduction

Many techniques are available for estimating delays at intersection approaches. However, little research has been performed to assess the consistency of estimates (Dion *et al*., 2004). Moreover, the applicability of the delay models needs to be determined based on their different data requirements and algorithms. For a transportation planning model, a balance between simplicity and accuracy is essential when choosing a delay modeling technique.

Various intersection delay estimating techniques often have different accuracies and their own limitations. A major technical difficulty for delay models arises when the v/c ratio is around one. For example, when the v/c ratio approaches one, steady-state delay models tend to produce unrealistically large delay estimates, while those of oversaturation delay models yield close to zero delays. Therefore, the v/c ratio is often the determining factor in choosing a delay model.

Intersection delays may include two components: queue delay (or stop delay) and signal delay (or acceleration/deceleration delay). Queue delay is difficult to quantify due to its stochastic nature. Sophisticated techniques may work better in estimating queue delays, but are often impractical for planning models due to intense data requirements. It is often difficult to find a well-balanced queue delay model to integrate into a planning model. Signal delay is determined from signal setting, volume, and geometry conditions of an intersection. If considered in a planning model, it needs to be updated repeatedly due to the varying volumes resulting from each traffic assignment iteration. Most problematic may be that in a planning model, of which the main purpose is to forecast future traffic conditions of a transportation network, signal timing plans and intersection geometry are unknown for a forecast year. It is impractical to perform signal phasing design for every intersection using the standard traffic analysis procedures, which are both time-consuming and data intensive. Therefore, it is necessary to facilitate signal design and optimization procedures through simplifying assumptions. For instance, one cycle length may be solely designated to arterials of the same functional classification, which is often not the actual case but is a reasonable approximation (Aashtiani and Iravani, 1999). Another issue is intersection geometry. The geometry of an intersection is not fixed and may be changed to relieve congestion or improve the intersection performance. A planning model capable of accommodating possible geometry changes is desirable for more accurate intersection delay estimates. For this purpose, it is necessary to establish a set of criteria governing changes in geometry conditions.

The achievable accuracy of a planning model also depends on realistic objectives of an intersection delay model. Nowadays, adaptive signal settings and signal coordination are becoming more common in urban areas. As a result, the platoon effects of traffic progression are often significant and cannot be ignored in delay estimation. However, generic delay models are often inadequate in reflecting progression conditions. For example, the delay model of the HCM merely uses a progression adjustment factor to account for progression while treating a studied intersection as isolated. Therefore, an intersection delay estimation model needs to be able to produce accurate results for urban streets with signal coordination.

The most frequently used delay models are based on the work by Webster (1958), expressed in the following form:

$$d = \frac{C(1-\frac{g}{C})^2}{2(1-\frac{v}{s})} + \frac{(\frac{v}{c})^2}{2v(1-\frac{v}{c})} - 0.65(\frac{c}{v^2})^{1/3}(\frac{v}{c})^{(2+\frac{g}{C})} \qquad \text{Eq. 1}$$

where

- d = control delay per vehicle (s)
- c = lane group capacity (veh/h)
- C = cycle length (s)
- g = effective green time (s)
- s = saturation flow rate (veh/h)
- v = demand for subject lane group or approach (veh/h)
- X = v/c ratio or degree of saturation for lane group

The delay models in the highway capacity manuals that are currently used in the United State, Australia, and Canada all originated from the above formula (Dion *et al.*, 2004). This formula includes three terms. The first term estimates the average approach delay assuming uniform arrivals, which is consistent with deterministic queuing models mentioned earlier. The second term considers the additional delays attributed to the randomness in vehicle arrivals. The third term is an empirical correction factor that reduces the estimated delay by 5~15% to be consistent with simulation results. Eq. 1 is among the most fundamental and is frequently referenced. There have been many efforts in determining various parameters based on local conditions or developing theoretical modifications. As a result, many delay models often have a similar form as the Webster's formula.

In the remainder of this chapter studies on delay models and their applications are reviewed. In Section 2.2, discussions are provided of the fundamental theories of delay models for a signalized intersection where the traffic is under conditions ranging from under-saturation to oversaturation. The most important applications, such as the 2000 version of Highway Capacity Manual (HCM) and the 2002 version of the Florida Quality/Level of Service Handbook, are also introduced. Section 2.3 further discusses research efforts to improve delay models. Section 2.4 is focused on issues related to combined models of signal optimization and traffic assignment, including methods dealing with the effect of signal coordination. Some relevant software programs are also briefly described.

2.2 Generic Intersection Delay Models and Applications

This section introduces the theory and history of generic delay models, which have been comprehensively studied for their characteristics and weaknesses. Applications such as Highway Capacity Manual 2000 (HCM) and Florida Quality/Level of Service Handbook 2002 are also described.

2.2.1 Intersection Delay Models

Hurdle (1984) and Dion *et al*. (2004) provide excellent reviews of major delay models. They also studied the basic principles, as well as simplifying assumptions that are not well-tailored to the real world. Although some improvements on methodologies and assumptions have been made, the theoretical core of delay models has remained basically unchanged. Hurdle's summary, which is based on a comparison of steady-state models and deterministic models, is still instructive to this date.

Most signal intersection delay models fall into two categories, steady-state models and deterministic queuing models. The former are usually considered useful only for predicting delays at intersections with light loads, while the latter do well only in the analysis of heavily loaded intersections where the volume overwhelms the capacity (i.e., v/c > 1). These models ignore the effect of random arrivals on queue lengths when intersections are slightly saturated. Because their assumptions are based on different v/c values, these two types of models are incompatible. However, when the load is heavy but v/c is still less than one, some good models are expected to produce excellent estimates. In TRANSYT developed by Transport and Road Research Laboratory, an algorithm based on a compromise between these two types of models is employed. The algorithm, while not a solid and realistic model, is able to illustrate some intuitive ideas. The algorithm may be represented by an approximated formula (Robertson, 1977):

$$D = 15\frac{T}{c}((v-c) + ((v-c)^2 + 240\frac{v}{T})^{1/2} \qquad \text{Eq. 2}$$

where

D = total delay on intersection approach (veh/s)
c = intersection approach capacity (veh/h)
v = demand for subject lane group or approach (veh/h)
T = duration of analysis period

A derivation of the TRANSYT random delay equation is presented by Kimber and Hollis (1979). The basic idea is to achieve a smooth transition between the steady-state and oversaturation models in the v/c range around 1. However, the smooth transition between the two types of models is not the result of any detailed analysis. Instead, it is based on an intuitive understanding of what happens. As Hurdle points out, to improve the delay estimates, more refined queue behavior models are required. Unfortunately, such models tend to be too complicated and demanding in data input.

As a continued effort to study steady-state versus deterministic models, Dion *et al*. (2004) compare the delay estimates at undersaturated and oversaturated pre-timed signalized intersections. Deterministic queuing models are classic applications to predicting delays for signalized intersections. These models view traffic on each intersection approach as a uniform stream of arriving vehicles seeking service from a control device that provides a high service rate. However, when the v/c ratio is much lower than one, the random effect is too evident to be ignored. This may be partly why such models have been applied mainly at intersections with far more arrivals per cycle than those can be served during a green interval (v/c > 1). In such cases, the random effect may be negligible, and model performance is adequate. Equations for calculating the average uniform vehicle delays during a cycle are presented below (Dion *et al*., 2004). Note that Eq. 4 is in fact identical to the formula in the HCM.

$$d_1 = \frac{r^2}{2C}\left(\frac{s}{s-v}\right) \qquad \text{Eq. 3}$$

$$d_1 = \frac{C\left(1-\frac{g}{C}\right)^2}{2\left(1-X\frac{c}{g}\right)} \qquad \text{Eq. 4}$$

where

d_1 = uniform delay (s)
c = lane group capacity (veh/h)
C = cycle length(s)
g = effective green time (s)
s = saturation flow rate (veh/h)
r = red time of the traffic signal (min)
v = demand for subject lane group or approach (veh/h)
X = v/c ratio or degree of saturation for lane group

Lighthill and Whitham (1955) state that traffic flow may be characterized using flow, density, and speed through an analogy to fluid dynamics. By demonstrating the existence of traffic shock waves, Lighthill and Whitham propose a theory of one-dimensional waves that could be applied to the prediction of highway traffic flow behavior. Their work is the origin of shock wave delay models.

The major difference between shock wave and deterministic queuing models is the manner in which vehicles are assumed to queue at an intersection. While queuing analysis based on deterministic models assumes vertical queuing, shock wave analysis assumes that vehicles queue horizontally. Vertical queuing assumes that the queuing vehicles at an intersection approach are stacked vertically, which is convenient for calculating delays but does not represent reality well. Horizontal queuing assigns waiting vehicles one by one to each of the lanes. The horizontal span of a queue enables the capture of more realistic queuing behavior and the determination of the maximum queue accumulated, which is not possible with deterministic queuing models, as they do not track spatial locations of queuing vehicles.

Similar to deterministic queuing models, shock wave delay models assume a non-random arrival pattern and that all vehicles accelerate and decelerate instantaneously. The consequences of these two assumptions are uniform delay estimates and that all delays are incurred on the approach side of an intersection. Because of these assumptions, shock wave theories are best used to estimate approach delays in oversaturated conditions (Hurdle, 1984).

Steady-state stochastic delay models are one type of stochastic delay models that attempt to account for the randomness in vehicle arrivals. One fundamental and most often referred example is the Webster's model (Webster, 1958). These models all consider that the number of arrivals in a given time interval follows a known distribution, typically a Poisson distribution, and that this distribution does not change over time. It is also assumed that the system remains undersaturated over an analysis period. Although temporary oversaturation may occur due to the randomness of arrivals, the system is assumed to have been running long enough to settle into a steady state.

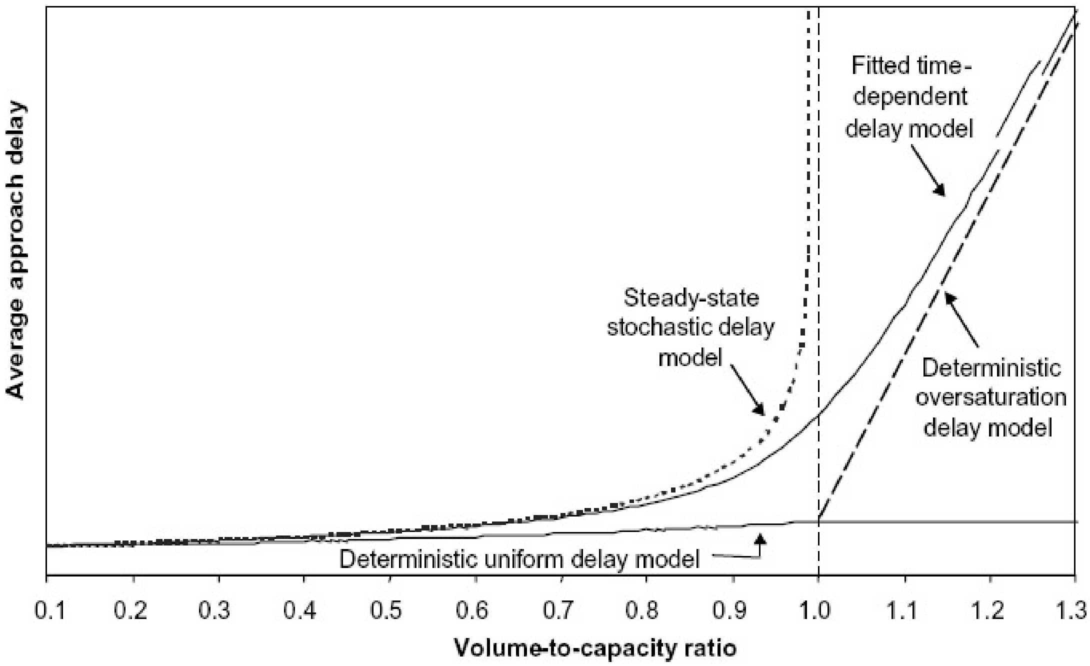

Figure 2.1 Steady-State Stochastic Models versus Deterministic Oversaturation Models (Dion *et al*., 2004)

To improve the performance of steady-state stochastic delay models and deterministic queuing models, the concept of a general time-dependent delay model is introduced by Kimber and Hollis (1979) using the coordinate transformation technique. This technique transforms the equation defining a steady-state stochastic delay model so that it becomes asymptotic to a deterministic oversaturation model. Although there is no rigorous theoretical basis for this approach according to Hurdle (1984), empirical evidence confirms that the results are reasonable. Therefore, the capacity guide delay models of the U.S., Australia, and Canada, which are similar to each other, are all based on the coordinate transformation technique. All these models assume steady-state traffic conditions. Under stochastic equilibrium conditions, the arrival and departure flow rates remain stationary for an indefinite period of time. The number of arrivals is also assumed to follow a Poisson distribution, which remains constant over time, and the headways between departures have a known distribution with a constant mean value.

Microscopic traffic simulation model is used to track individual vehicle movements in simulated networks, which allows such models to consider virtually any traffic conditions, ranging from undersaturation to severe oversaturation. The models determine the delay incurred to an individual vehicle traveling in a network with different network conditions by comparing simulated and ideal travel times.

Dion *et al*. (2004) also employ the INTEGRATION microscopic traffic simulation software to estimate delays. The simulation model integrates dynamic traffic simulation and traffic assignment. Delay is estimated for each vehicle by calculating, for each traveled link, the difference between the vehicle's simulated travel time and the travel time that the vehicle would have experienced on the link at free-flow speed.

Consistent with the conclusions by Hurdle (1984), Dion *et al*. find the same trend in the results from stochastic and deterministic models. All the analytical delay models generate similar results when the v/c ratios are low. Deterministic queuing and shock wave models always make the lowest estimates since these two types of models consider only uniform arrivals. Therefore, they are unable to account for the potential additional delays that arise from the random oversaturation delays caused by platooning of arriving vehicles.

The average delay estimates from the INTEGRATION simulation model are in general agreement with the estimates from various models such as the 1981 Australian Capacity Guide, the 1995 Canadian Capacity Guide, and the 1997 HCM delay models. Dion *et al*. (2004) point out a strong consistency in the delays estimated by the time-dependent stochastic delay models and by the INTEGRATION microscopic traffic simulation model.

To simply and effectively consider intersection delays, almost every real-world model of delays at a signalized intersection begins with Webster's delay model, which is based on a simple analytical approach to the computation of the uniform delay component. The equation has been given in Eq. 4 and is again shown below:

$$d_1 = \frac{C(1-\frac{g}{C})^2}{2(1-X\frac{c}{g})}$$

This equation is identical to that in the HCM for uniform delays. It makes the simplifying assumption that the arrival function is uniform, i.e., arrivals are at a constant rate, v (veh/s). The uniform delay formula, however, does not consider random arrivals. At isolated intersections, vehicle arrivals are more likely to be randomly distributed. The assumption of uniform arrival implies that the queue of vehicles at an intersection operating under undersaturated conditions is always cleared before the next red signal.

As mentioned before, Webster (1958) proposed a stochastic model that assumes arrivals being Poisson distributed with an average rate v (veh/s). Following Webster's work, some other stochastic models have been proposed, including, for instance, the models by McNeil (1968) and Heidemann (1994). These models all share several basic assumptions. First, the number of arrivals of a fixed time interval follows a known distribution, usually a Poisson distribution. This distribution does not change over time, which implies that these models may not be applied to estimating delays of intersections with coordinated signals, where arrivals are platooned as a result of the upstream traffic signals. Second, while it is recognized that temporary oversaturation may occur due to random arrivals, it is assumed that the system remains undersaturated throughout an analysis period. A main consequence of such steady-state stochastic delay modeling is that the estimated delays tend to infinity as traffic demand approaches saturation (v/c ratio = 1.0), which is considered a weakness by many (Roess *et al*., 1998). A proper delay estimation model should, theoretically, perform better under different demand levels. For low v/c ratios, the model is expected to produce delay estimates similar to those from deterministic queuing delay models assuming constant uniform arrivals. As the demand increases, a growing proportion of delays are attributed to random vehicle arrivals and the failure of all queued vehicles to clear in previous cycles. As the v/c ratio approaches 1.0, the model should not approach infinity, but instead should generate estimates tangent to the deterministic oversaturation model as Eq. 2 does. The concept of a time-dependent delay model is originally proposed and enhanced by Kimber and Hollis (1979). Numerous time-dependent delay formulas have been proposed and incorporated into a number of capacity guides, such as the 1994 and 1997 Highway Capacity Manuals and those of Australia and Canada. Details of the delay models employed in the HCM will be discussed in the next subsection.

2.2.2 Applications of Generic Delay Models

In the U.S., the HCM is the most comprehensively used reference (Troutbeck and Blogg, 1998). However, the HCM's methodology comes with limitations, which have been criticized widely. The intersection delay methodology of the HCM does not detect and adjust for the impacts of turn-pocket overflows on through traffic and intersection operations, nor does it take into account the potential impact of downstream congestion on intersection operations. In other words, the intersection is analyzed as an isolated facility. Therefore, the delay represents the average control delay experienced by all vehicles that arrive in an analysis period, including delays incurred beyond the analysis period when the lane group is oversaturated. Control delay includes movements at slower speeds and stops on intersection approaches as vehicles move up in queue position or slow down upstream of an intersection.

The average control delay per vehicle for a given lane group is given by

$$d = d_1\,(PF) + d_2 + d_3 \qquad \text{Eq. 5}$$

where

- d = control delay per vehicle (s/veh)
- d_1 = uniform control delay assuming uniform arrivals (s/veh)
- PF = uniform delay progression adjustment factor accounting for the effects of signal progression
- d_2 = incremental delay to account for the effect of random arrivals and oversaturation queues
- d_3 = initial queue delay accounting for the delay to all vehicles in the analysis period due to initial queue at the start of the analysis period (s/veh)

and

$$d_1 = \frac{C \times (1 - \frac{g}{C})^2}{2\left(1 - \min(1, X) \times \frac{g}{C}\right)} \qquad \text{Eq. 6}$$

where

C = cycle length(s)
g = effective green time
X = v/c ratio or degree of saturation for the lane group

and

$$d_2 = 900T\,[(X-1) + \sqrt{((X-1)^2 + 8kl\frac{X}{c}T}\,] \qquad \text{Eq. 7}$$

where

T = duration of the analysis period
k = incremental delay factor dependent on controller settings
l = upstream filtering/metering adjustment factor
c = lane group capacity (veh/h)
X = lane group v/c ratio or degree of saturation

Both the calculations of d_1 and d_2 assume that there is no initial queue at the beginning of the analysis period of duration T.

$$d_3 = \frac{1800Q_b(1+u)t}{cT} \qquad \text{Eq. 8}$$

where

c = lane group capacity (veh/h)
Q_b = initial queue at the start of period T (veh)
T = duration of analysis period
t = duration of unmet demand in $T(h)$
u = delay parameter

These delay terms are estimated from variables or parameters that are related to operations at upstream of the subject intersection. They include six vehicle arrival types (HCM 2000), green time ratio (g/C), percentage of vehicles arriving during the green time, degree of saturation (v/c), lane capacity, length of the analysis period, and size of the queue at the start of each cycle. Conditions of the downstream segments and intersections are usually ignored. As the HCM 2000 indicates, "The potential impact of downstream intersection on the upstream intersections is not taken into account." When a downstream intersection influences an upstream one, additional parameters/variables need to be considered other than those in the HCM 2000. The other major limitation is that random overflow at the downstream link is not considered.

The HCM does recommend procedures for computing and predicting performance measures for area-wide analyses. The methodology simplifies and approximates the HCM's Part III procedure, which is designed for use on the freeway, arterial, and rural highway subsystems of a regional transportation system. The procedure estimates space mean speed for vehicles on a link. However, the analyst may optionally study the mean vehicle delay for each link approach to node intersection. The node delay is then added to the link traversal times to obtain the total link travel time. The link travel times are summed up over all links of area facilities to obtain the total travel time for the entire subsystem. The procedure determines the free-flow speed (FFS), link capacity, and link speed. When computing link speed, the node delay for signal or stop-sign-controlled intersection needs to be estimated. With link speed, link length, and vehicle demand of the link, the total person-hours of delay caused by traffic congestion and intersection controls is then computed using Eq. 9 and taken as the performance measures:

$$PHD = PHT - AVO \times v \times L/S_o \qquad \text{Eq. 9}$$

where

PHD = total person-hours of delay
PHT = person-hours of travel in corridor
AVO = average vehicle occupancy
L = distance traveled
v = directional demand
S_o = FFS on link i (mi/h)

It should be noted that the impact of congestion in one time period on the delays in the following time periods is not considered. However, the HCM does provide a procedure for multi-period analysis for more advanced applications, which is described in detail below.

The procedure is called the node delay estimation procedure. It estimates the most likely type of intersection control and the delay caused by the intersection control under forecast traffic demands. The adaptive procedure is to estimate the average node delay (s/veh) for traffic on a single link approaching a signalized intersection. The procedure requires demand and geometric data of the signalized node. Additionally, a reasonable phasing plan needs to be developed. The lane groups, their saturation flow rates, and critical phases will then be determined using the HCM's corresponding functions. First, the cycle length is estimated as

$$C = \frac{TL}{1.0 - \min(1, 0.9 \times CVS)} \qquad (60\text{s} < \text{C} < 150\text{s}) \qquad \text{Eq. 10}$$

where

0.90 = target volume to capacity ratio for the intersection
CVS = sum of critical v/s ratios for the intersection
C = estimated signal cycle length (s)
TL = total intersection lost time (s)

Next, the effective green time is estimated using the following equation:

$$g/C = \text{Max}\ (EL, ET, WL, WT) / (CVS + TL/C) \qquad \text{Eq. 11}$$

where

EL, WL = v/s ratios for eastbound and westbound left-turning movements, respectively
ET, WT = v/s ratios for the eastbound and westbound through movements, respectively (if separate right-turn lane is present, these become maximum of through or right-turn v/s ratios for each compass direction)

Finally, the control delay is estimated using Equation 12:

$$D_a = [(D_l V_l) + (D_t V_t) + (D_r V_r)] / (V_l + V_t + V_r) \qquad \text{Eq. 12}$$

where

D_a = mean delay for the subject approach (s/veh)

D_l, D_t, D_r = estimated delay for the left turns, through vehicles, and right turns, respectively (s/veh)

V_l, V_t, V_r = volumes (demand) for the left-turning lane group (if present), through lane group, and right-turn lane group, respectively (veh/h)

The delay for each lane group is computed as follows:

$$d = \frac{C(1-\frac{g}{C})^2}{2(1-\min(1,X)\frac{g}{C})} + 900T[(X-1) + \sqrt{((X-1)^2 + \frac{4X}{Ts\frac{g}{C}}}] \qquad \text{Eq. 13}$$

where

d = average control delay for subject lane group (s)
c = lane group capacity (veh/h)
C = cycle length(s)
g = effective green time (s)
s = saturation flow rate (veh/h)
T = duration of analysis period (h)
v = demand for subject lane group or approach (veh/h)
X = *(v/s)/(g/C)* ratio of volume to capacity for lane group

The HCM also provides procedures for calculating delays at two-way stop control and all-way stop control. These procedures neglect the impacts of grade, heavy vehicles, wide medians, upstream signals, and flared approaches on the capacity of a stop-controlled intersection. To simplify the calculations, it is assumed that left-turning lanes are always present on the major street.

For large study areas involving multiple facilities and corridors, the HCM also covers corridor analysis, which is a part of a larger demand forecasting procedure that will reevaluate the demands for each parallel facility in response to the predicted operating conditions. These procedures assume that the demand is temporarily fixed, allowing the computation of facility operations. If traffic queues are likely to affect upstream facility operations significantly, microsimualtion analysis instead of corridor analysis is recommended.

In the HCM 2000, the node delay estimating procedure assumes that the demand is already known and has been assigned to the appropriate segments within the corridor, based on the relative travel time via each segment. In a corridor, it is normal for vehicles to change routes based on changes in the travel times on parallel facilities. At this point, an O-D matrix from planning model should be available to assist in re-computing the segment demand. The O-D matrix will then be updated based on the corridor analysis results.

The corridor analysis procedure should be used only if congestion is considered likely on one or more of the corridor facilities. The procedure then may track some of the inter-facility effects. The input data require additional variables than Part III methodology does, such as:

- length and number of lanes on each segment
- cycle length, phasing, and green times for each signal
- turning movement counts

However, for large areas involving thousands of segments and points, the input data requirements of corridor analysis tend to become intractable. Thus, the area-wide analysis of the HCM mentioned above should be applied in such circumstances.

In Florida, an important application of the HCM methodologies is the Quality/Level of Service Handbook of the Florida Department of Transportation (FDOT) (2002), referred to as FDOT Q/LOS herein, and its software, which are nationally recognized as the leading planning application of the HCM for the evaluation of automobile/truck Levels of Service (LOS). According to Figure 2.2, both control delays and LOS criteria employ the HCM procedures. While operational analyses, such as intersection signal timing, are sometimes conducted at the planning level, the handbook does not provide the necessary tools for actual design or operation of facilities or services where more appropriate resource documents or analysis methods are available.

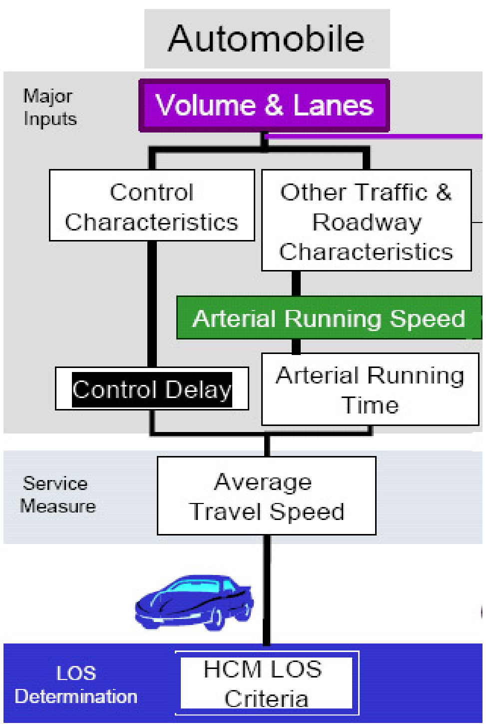

Figure 2.2 Control Delay in the Q/LOS Procedure of Florida (Quality/Level of Service Handbook, 2002)

The handbook includes two levels of analysis: (1) generalized planning and (2) conceptual planning. Generalized planning is intended for broad applications such as statewide analyses, initial problem identification, and future year analyses. Conceptual planning is increasingly more detailed than generalized planning, but does not involve comprehensive operational analyses.

Planning level analyses make extensive use of simplifying assumptions of Q/LOS evaluation techniques and default values in operational models. For example, a major simplifying assumption, which is essential to the development of the generalized tables in the FDOT Q/LOS, is the selection of a single effective green ratio (g/C) for all the intersections on an arterial.

FDOT has determined that for generalized planning analyses, the "weighted effective green ratio" yields the closest results to actual conditions. The weighted effective g/C of an arterial is the average of the critical intersection through g/C and the average of the other intersections' through g/C. Another significant planning assumption is that mainline non-through movements are adequately accommodated. Typically, the through movement is the straight movement. However, occasionally the "through" movement is a right or left turning movement, with the straight ahead movement being considered a non-through movement. Most analyses of through movements in the HCM are relatively straightforward. Complications arise with the treatment of turning/merging movements, especially for signalized intersections and arterials. By handling non-through arterial movements (i.e., turns from the arterial and side street movements) in a general way, Q/LOS analyses are greatly simplified.

As there are no software packages that are completely satisfactory to the FDOT for planning purposes, the FDOT developed planning applications based on the HCM's delay and LOS evaluating methodologies. FDOT recommends the use of HIGHPLAN for highways and ARTPLAN for signalized intersections and arterials. The assumed free flow speed is 5 mph above the posted speed.

For arterial planning, calculations of segment running time in the HCM do not include traffic volume as a variable. Based on the results from research conducted for FDOT (Quality/Level of Service Handbook, 2002), the national subcommittee overseeing the HCM approved changes to the HCM 1994. However, due to time constraints, these changes are not included in the HCM 2000. Results from that research effort including equations are included in the current 2002 version of FDOT Q/LOS Handbook and the accompanying software. Specifically, traffic volume is included as a variable in calculating running speeds to better reflect running speeds of through vehicles, as opposed to the total mix of through and turning vehicles.

The HCM LOS measures of effectiveness for urban streets are essentially for arterials. LOS is based on the average travel speed of the traffic. Generally, on major non-state roadways, motorists tend to perceive the service quality based on average travel speed. It is generally assumed that the LOS for local unsignalized roadways is acceptable. However, for roads that have signalized intersections, the methodology of the FDOT Q/LOS Handbook recommends that the HCM intersection LOS criteria (delay at the intersection) be used to determine the LOS. With this procedure, such facilities are evaluated by the delay at the signalized intersection and not the average travel speed of the roadway. In the generalized tables of the FDOT Q/LOS Handbook, these roadways are labeled as "other signalized roadways". Previous editions of the FDOT's software include a program called SIG-TAB to evaluate a roadway. To simplify and reduce the number of software programs, these facilities may now be evaluated using ARTPLAN by selecting "signal" as the roadway class. The Florida's LOS planning software, including ARTPLAN, FREEPLAN, and HIGHPLAN, is the major tool in conducting Q/LOS analysis in

Florida. The software, as well as the handbook, is based on the HCM 2000 techniques. The ARTPLAN is primarily for urban signalized roadway analyses, FREEPLAN is mainly for freeway systems analyses, and HIGHPLAN is for two-lane or multilane uninterrupted flow analyses. Like the other two programs, the ARTPLAN first calculates the LOS for the facility being analyzed and provides the service measure (e.g., average travel speed). It then calculates three service volume tables: hourly volumes in the peak direction, hourly volumes in both directions, and annual average daily traffic volumes.

The Quality/Level of Service Handbook and its accompanying software are designed for the evaluation of roadway users' quality/level of service (Q/LOS) at planning and preliminary engineering levels. Q/LOS analyses are based on three types of input variables: roadway, traffic, and control. For an urban arterial, ten variables having a significant impact on volume calculation in LOS analysis are (Quality/Level of Service Handbook, 2002):

- Number of through lanes
- Left turn lanes
- Paved shoulder/bicycle lane/outside lane width
- Sidewalk
- Average annual daily traffic (AADT)
- Planning analysis hour factor (K)
- Directional distribution factor (D)
- Bus frequency
- Signalized intersection spacing
- Effective green ratio (g/C)

Most of these variables are required and are used in the standard HCM 2000 procedures.

2.3 Research Efforts to Improve Generic Delay Models

Although widely applied, the HCM delay model does have limitations. For instance, under oversaturated traffic conditions, Benekohal and Kim (2005) found counterintuitive results, because progression adjustment factor (PF) is not applied to signalized delay models when there is an initial queue, as recommended in the HCM. In some cases, the delay under an initial queue condition is less than the delay with zero initial queue. Under oversaturated conditions, when there is an initial queue, the HCM 2000 delay model yields the same uniform delay values for all arrival types, which does not seem reasonable since platooning affects delay. Benekohal and Kim propose a new uniform delay model considering platoon impact under oversaturated traffic conditions when progression is poor. This approach directly quantifies the platooning effects in delay, eliminating the need for applying the progression adjustment factor. Like the HCM 2000, the proposed model is applicable with or without an initial queue, as given below:

$$d_1 = 0.5sg\,[Q_1C + Q2(C\text{-}t_1) - q_oC^2 - sg^2] \qquad \text{Eq. 14}$$

where

q_{av} = average arrival rate (veh/s)
q_{pl} = platoon arrival rate (veh/s)

q_n = non-platoon arrival rate (veh/s)
t_1 = platoon duration time (s)
q_o = overflow rate, which is the difference between q_{av} and c (veh/s)
Q_1 = number of arrivals when queue increase rate changes for the first time (= $q_p\, t_1$)
Q_2 = number of arrivals at the end of the cycle (= $q_{av}C$)

Compared to inputs in the HCM, this arrival based model also requires platoon duration time (t_1), platoon flow rate (q_{pl}), and non-platoon flow rate (q_n) to calculate platoon and non-platoon arrival rate to compute delays. The additional input may be difficult to obtain from the perspective of transportation planning. However, the authors claim that this arrival-based approach is more accurate than the HCM approach.

To avoid the level of details that are overburdening for planning models, Aashtiani and Iravani (1999) made two assumptions. First, it is assumed that all turning movements corresponding to links entering an intersection have the same delay. Therefore, the delay at the intersection is included as a part of the delay function for the entering links. Second, the delay associated with the intersection for each entering link depends on the physical characteristics and control policy of the intersection and the traffic volume on the link. Although these assumptions seem unrealistic at the operation level, especially for unsignalized intersections, they are justified in the planning stage when the objectives are to make strategic decisions. In comparison to exclusion of delay at intersections, they are more realistic. The control delay of signalized intersections based on Webster's formula is simplified as (Aashtiani and Iravani 1999)

$$d(x) = \frac{r^2}{2C(1 - \frac{x}{\mu w})} \qquad \text{Eq. 15}$$

where
x = traffic volume on entering link (passenger car equivalent) (vph)
μ = exiting rate of traffic volume (passenger car equivalent) (vph)
r = red time of the traffic light (min)
C = cycle length of the traffic light (min)
d = average delay at intersection (min)
w = width of the link (foot)

Because at the planning level, cycle length and red time might not be available for most intersections, they need to be estimated when considering signalized delay Assuming node j is a signalized intersection, s_j is the set of links ending at node j. The following weights, for each link entering node j, could be defined based upon their functional class as follows (Aashtiani and Iravani 1999):

$$w_{ij} = \begin{cases} 2, & \text{if the link is a local road or collector} \\ 3, & \text{if the link is a minor arterial} \\ 4, & \text{if the link is a major arterial} \\ 5, & \text{if the link is an expressway} \end{cases}$$

and the total cycle time in minutes at node j is

$$c_j = 1 + (\Sigma(w_{ij})/8)|s_j|/4 \qquad \text{Eq. 16}$$

To calculate the red cycle:

$$r_{ij} = 1.2 \times c_j \times [1 - |s_j| \times w_{ij} / 2\Sigma(w_{ij})] \qquad \text{Eq. 17}$$

The functional form of this equation has the following desirable characteristics:

- Each link of any intersection takes into account the effect of all other links connecting to the same intersection.
- Intersections with more legs will experience more delay time when the rest of explanatory variables remain the same.
- Intersections with more turn prohibitions will have with less delay.
- Delay at intersections without traffic signals, in any size of network, could be calculated using the equation.

Despite the limitations related to the exclusion of opposing traffic volume, Aashtiani and Iravani claim that the methodology is statistically valid and that it improves the calibration process.

Another major limitation of the HCM methodologies is that its delay model only deals with isolated intersections. At present, most delay models deal with congestion delays without giving considerations to the impact of downstream congestion and traffic disturbances that may include waiting queues at downstream signalized approaches (Ahmed and Abu-Lebdeh, 2005). Closely spaced intersections are frequently seen in the U.S. urban areas. Even when intersections are distantly spaced, with heavy traffic flows, they may become bottlenecks with downstream congestion causing unacceptable delays at upstream intersections.

The control delay from the HCM 2000 is a combination of three delays with a progression adjustment factor (PF), as shown below (Eq. 5):

$$d = d_1(PF) + d_2 + d_3$$

These three delay terms may be computed based on the following information: offsets, green phase at downstream intersection, distance between intersections, link traveling speed of vehicles, queue lengths, queue spillovers, speed of shockwaves, and so on. A new delay term may be needed to capture the influence of traffic operations at a downstream intersection and/or link on the neighboring upstream intersection.

To estimate the delay due to a downstream disturbance, Ahmed and Abu-Lebdeh (2005) introduce a fourth delay term (d_4). This term is determined by the geometry and traffic operational characteristics of both upstream and downstream intersections. Traffic disturbances at a downstream intersection may cause interruption in the flow on the link between this and its upstream neighboring intersections. Consequently, a number of shock waves are generated. Shockwave analysis is applied to evaluate the significance of a downstream disturbance to an upstream intersection. The average speed of traffic will be a function of space that is not occupied by traffic. This fourth delay term is shown below (Ahmed and Abu-Lebdeh, 2005):

$$d_4 = \frac{n}{2}[2d_{(4)_1} + (n-1)h_v(\frac{1}{v_1} - \frac{1}{\lambda_1})] \qquad \text{Eq. 18}$$

where

n = total number of vehicles queued at the upstream intersection

h_v = effective space headway (m)

v_1 = speed of mid block stopping wave (m/s)

λ_1 = speed of mid block starting wave (m/s)

$d_{(4)_1}$ = portion of d_4 incurred to the first vehicle at upstream intersection, computed as

$$d_{(4)_1} = (\frac{L_1}{v_2} + \frac{L_2}{v_1} + off - \frac{L_2}{v_a} - \frac{L_2}{\lambda_1}) \qquad \text{Eq. 19}$$

where

Off = offset (s)

L_1 = queue length measured from the downstream intersection stop line to the tail of the queue (m)

L_2 = remaining space on link (not occupied by vehicles) (m)

v_1 = speed of mid-block starting wave (m/s)

v_2 = speed of starting wave at downstream intersection (m/s).

v_a = average link speed (m/s)

Because the queue length at a downstream approach directly impacts the magnitude of d_4, the model needs to include parameters such as offsets, incoming volume from the upstream intersection, and other traffic control variables. Due to the many variables involved, including green/red phase, offsets, and average link speed, data requirement at this level of detail may be overburdening for a transportation planning model.

Another direction of research is the queuing theory. Troutbeck and Blogg (1998) study queue accumulation and decay for a high definition approach given random arrival and departures. The approximation of queue length and delay has been commonly called “coordinate transformation technique” following the publication of Kimber and Hollis (1979). Kimber and Hollis’ theory is fairly similar to what is described by Hurdle (1984) on signal delay, which is a mathematical representation of the steady-state queue length versus oversatuation (deterministic) curve. As shown in Figure 2.3, the transformed equation by Kimber and Hollis (1979) produces a modified curve that transitions from steady-state models to deterministic ones.

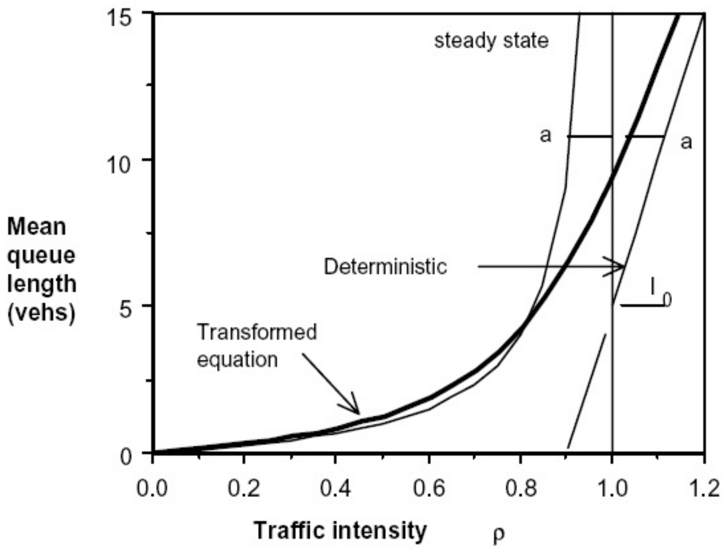

Figure 2.3 Modified Curve of Transformation Technique (Troutbeck and Blogg, 1998)

As Hurdle (1984) points out, Kimber and Hollis' approach is simply using mathematical expressions that fit the curve shown in Figure 2.3. Kimber and Hollis admit that in the limiting cases (the two ends of the curve) their results are correct, and that in the intermediate regions the behavior of their functions is sensible. Kimber and Hollis' method provides little understanding of the system, particularly when a system reaches a critical point or as the demand approaches the capacity. Troutbeck and Blogg (1998) compared Kimber and Hollis' approach to that of Newell (1982). Newell's method is versatile when dealing with arrival profiles, which allows one to estimate queues or delays for more general systems such as platooned arrivals and gap acceptance service with time dependent, parabolic arrival profiles. Newell developed a diffusion equation to describe the time dependent queuing system, shown in Eq. 20. His equation reflects non-dimensional relationships, which greatly improves the method's applicability to new situations.

$$\frac{\partial f_Q(l;t)}{\partial t} = -(\beta_A - \beta_D)\frac{\partial f_Q(l;t)}{\partial l} + \frac{i(\beta_A + \beta_D)}{2}\frac{\partial^2 f_Q(l;t)}{\partial l^2} \qquad \text{Eq. 20}$$

where

- l = queue length
- t = current time
- f_Q = joint probability density of the cumulative queue length l, at time t
- β_A = service rate of an approach
- β_B = service rate of an approach
- i = constant of Poisson processes

Troutbeck and Blogg draw conclusions from the comparison that Newell's method provides not only accurate results but also an estimate of the variance as well. The variance estimate makes it easy for practitioners to further analyze the performance of an intersection, a feature unavailable if the coordinate transformation technique is used.

2.4 The Combined Model of Signal Delay and Traffic Assignment

This section summarizes the research efforts in incorporating signal delays into traffic assignment process. Signal delay estimation models need to be reasonably simplified before being employed to improve the accuracy of traffic assignment.

2.4.1 Simultaneous Optimization of Signal Settings and Traffic Assignment

Signal timing design for an isolated intersection has been covered in the HCM and in many standard textbooks such as that by Roess *et al*. (1998). In this section, the attention is given to signal timing design with the objective of optimizing a signalized network.

For real-world applications, researchers often need to find an appropriate accuracy level for strategic planning purposes. Dunn and Johnson (1999) use EMME/2 to develop a model for the Auckland City, New Zealand. The model incorporates detailed local traffic operations, including intersection delay, and has a well defined hierarchy. At the highest level there is the Auckland Regional Transport (ART) model, which is a four-step EMME/2 model capable of assessing transport strategies at the regional level. One of the sub-models assesses the impacts of regional transport strategies on the road network to a level where appropriate mid-block lane requirements, intersection control, and signal timing strategies may be identified. Lower level models employ delay functions that relate to link characteristics. Some of the submodels incorporate link based functions that reflect, to some degree, intersection approach capacity and delay characteristics. Many urban models have been satisfactorily calibrated using link-based functions to estimate intersection delays as well as link delays.

There have been several attempts to model intersection delays in EMME/2. All efforts have resulted in unstable results on congested networks unless based on functions that strongly relate to the traffic flow along a link. Dunn and Johnson's approach uses only one signal delay function for a given progression condition, which is consistent with the methodology used to assess intersection operations in Australia.

To consider a regional model with large zones and a relatively coarse network with delay functions for links, Hill (1999) implements delay functions based on selected analytic models for priority, roundabout, and signal controlled intersections. To cope with traffic loads exceeding capacity, Hill employs a signal delay function as follows

$$D(x) = D_u(x) + D_o(x) \qquad \text{Eq. 21}$$

$$D_u(x) = \begin{cases} \dfrac{(1+Rx^{0.1})(C(1-u)^2)}{120(1-ux)} & \text{if } x < 1 \\ \dfrac{C(1-u)}{120} & \text{if } x \geq 1 \end{cases} \qquad \text{Eq. 22}$$

$$D_o(x) = \begin{cases} 15T[(x-1)+\sqrt{\dfrac{(x-1)^2+4.4(x-x_0)}{suC}}] & \text{if } x \geq x_0 \\ 0 & \text{if } x \leq x_0 \end{cases} \qquad \text{Eq. 23}$$

$$R = 0.1\varphi\,(suC/3600)^{0.25}\,(u^{0.1}) \qquad \text{Eq. 24}$$

$$x_o = \text{MIN}\,[0.95,\ 0.4 \times (suC/3600)^{0.2}] \qquad \text{Eq. 25}$$

$$x = \rho KP \,/\, suT \qquad \text{Eq. 26}$$

where

- D = average control delay
- D_u = uniform delay
- D_o = overflow delay
- φ = proportion of un-bunched traffic ρ (ratio)
- K = turning volume factor (ratio)
- P = flow peaking factor for ρ (ratio)
- s = saturation flow available to ρ (vphpl)
- u = effective green time for ρ (ratio to cycle time)
- ρ = EMME/2 assigned turn volume in vehicle per analysis period of T hours
- Q = lane capacity in vehicle/hour (= $su \times T$)
- Z = delay adjustment factor for signal coordination (ratio)
- C = cycle length
- T = analysis period

These input variables are required attributes to describe a signaled turn. C, u, s, and K, are used by a process controlled by EMME/2 macros in the feedback procedure. A three-attribute vector of φ, Z, and P is filled initially with, respectively, default values of 0.5, 0.85, and 1.1, and then modified interactively in accordance with guidelines.

To cope with traffic loads well in excess of capacity, another priority delay function, $D(x)$, is applied. The capacity module is chosen because it has been shown to produce results that correlate well with the results of detailed simulations.

$$D(x) = \frac{60+15\sqrt{(2+QT(1-x))^2+8QTx}-(2+QT(1-x))}{Q} \qquad \text{Eq. 27}$$

$$Q = \text{MAX}(Q_{\text{min},}\ \frac{\varphi(q_{0+0.1})\mathrm{e}^{-(A+d-H)q_1}}{1-e^{-F_{q1}}}) \qquad \text{Eq. 28}$$

$$q_1 = \begin{cases} \dfrac{\varphi(q_0+0.1)/3600}{1-H(q_0+0.1)/3600} & \text{if } q_0 \leq (3600/H)-1 \\ \dfrac{\varphi(3600/H+0.1)/3600}{1-H(3600/H+0.1)/3600} & \text{if } q_0 > (3600/H)-1 \end{cases} \qquad \text{Eq. 29}$$

$$x = \frac{\rho P}{QTn} \qquad \text{Eq. 30}$$

where

Q = absorption capacity of the opposing flow(s) (vehicle per hour)
Q_{min} = minimum value for Q (vehicle per hour)
q_0 = aggregate of factored opposing flows (vehicle per hour)
A = acceptance gap (s)
d = 0.35× standard deviation of A (s)
F = follow-up headway (s)
φ = un-bunched traffic in the opposing flow as a ratio
H = headway in platoons (s)
T = analysis period (hour)
P = flow peaking factor for the analysis period as a ratio
n = number of lanes for the turning movement

These input variables are required attributes to describe a priority turn. The EMME/2 network calculator computes Q. Q_{min} is the value of a required extra attribute to allow for the priority reversal that occurs with severe congestion. The factors for opposing flow components may also have large effects and need to be applied with discretion in the absence of survey data.

Based on the EMME/2 platform, Zhou and Vaughan (1999) perform intersection modeling by treating complicated intersection situations using the macro capabilities of EMME/2 other than the normal assignment methods. EMME/2 has network calculation modules to calculate the capacity and effective green time of turning movements. Their general approach of the new strategic highway assignment module involves calculating the capacity and effective green time for each movement in the network, which are used by a turn penalty function to calculate the movement delay. The equilibrium assignment process adds the movement delay to link delay to assign link and turning flows that are used in turn to calculate the effective green time and link capacity in the next iteration. The model requires more input variables than traditional travel demand models. Much more efforts are needed to prepare the input data, which include the number of lanes, shared lane existence, signal control availability, opposed flow information, green time and cycle time. Turn penalty function is used to calculate delay for each movement at an intersection. This type of turn penalty functions is actually derived from a more general function form, which embraces the delay function used in the Highway Capacity Manual, Canadian and Australian methods, and is in the same form as Eq. 21:

$$D = D_u(x) + D_o(x)$$

where

D = total delay for a turning movement (s)
D_u = uniform delay (s)
D_o = overflow delay (s)

The uniform delay D_u is, however, defined differently as

$$D_u = \frac{c(1-u)^2}{2(1-ux)} \qquad \text{Eq. 31}$$

where

c = cycle time (s)
u = green ratio (equal to the movement green time divided by cycle time)
x = degree of saturation, which is the ratio of arrival flow to capacity
x = $(qc)/(sg)$, where q is the movement arrival flow, and s is the base saturation flow

Different from that defined in Eq. 23, the overflow delay D_o has the following form:

$$D_o = 900\, T_f \left[z + \sqrt{z^2 + \frac{4x}{QT_f}} \right] \qquad \text{Eq. 32}$$

where

T_f = simulation period in hours, currently set as 1
Q = movement capacity (veh/h)
x = degree of saturation of the movement
z = $x - 1$

The turn penalty function given by Eq. 21 is expected to estimate realistically the delay when the degree of saturation, x, is closer to one. The model of Zhou and Vaughan (1999) is able to effectively represent various conditions at signalized intersections. Its iterative approach with new turn penalty function is also able to achieve relatively quick convergence.

A great deal of efforts has been spent to incorporate accurate signal delays into transportation planning. Geylan and Bell (2004) use the genetic algorithm (GA) approach to solve traffic signal control and traffic assignment problems by optimizing signal timings based on stochastic user equilibrium link flows for an entire network. They define a weighted linear combination of delay and the number of stops as the system performance index, which is evaluated by a TRANSYT traffic model. Stochastic user equilibrium assignment is formulated as an equivalent minimization problem and solved by the traffic assignment model. The objective function is the inversion of a network performance index, PI. PI is a function of signal setting variables $\psi(c,\theta,\varphi)$ and equilibrium link flows $q(\psi)$, where c is the network cycle time, θ is the feasible range of offsets, φ is the duration of green time, and q is the stochastic user equilibrium link flows. The objective function, which is the "fitness function" in the GA, minimizes PI with respect to equilibrium link flows $q(\psi)$, subject to signal setting constraints.

The GATRANSPFE (GA, TRANSYT, and the PFE) integrates genetic algorithms, traffic assignment, and traffic control to deal with the equilibrium network design problem. A bi-level approach is used. The lower level finds equilibrium link flows based on the stochastic effects of driver routing, while the upper level deals with signal setting. The procedure, however, often encounters local optima. Since equilibrium signal setting is generally a non-convex optimization process, the GA approach is used to globally optimize signal setting at the upper level by calling TRANSYT traffic model to evaluate the objective function.

For the lower level, a stochastic user equilibrium assignment, called the Path Flow Estimator (PFE), is employed. The input to PFE, such as cycle time and green duration, are results obtained from genetic algorithm.

The solution procedure in the GATRANSPFE includes the following steps:

1. Initialization: Set the user-specified GA parameters. Represent the decision variables as binary strings to form a chromosome *x*.
2. Set an initial random population of signal timing solutions. Apply the GA to produce signal timing solutions.
3. Decode signal timing parameters of timing solutions to map the chromosomes to the corresponding real numbers.
4. Solve the lower level problem by way of the PFE. This gives a SUE (stochastic user equilibrium) link flow to each link of the network.
5. Obtain the network performance index (PI) for signal timing in Step 2 and the corresponding equilibrium link flows from Step 3 by running TRANSYT.
6. Calculate the fitness functions of PI for each chromosome of the timing solution population.
7. Reproduce the timing solution population according to the distribution of the fitness function values.
8. Perform crossover operations randomly.
9. Perform mutation operations randomly to obtain a new population.
10. If the difference between the population average fitness and population best fitness index is less than 5%, re-initialize population and go to Step 1. Otherwise go to step 11.
11. If maximal generation number is reached, the chromosome with the highest fitness is adopted as the optimal solution of the problem. Otherwise return to Step 2 for the next iteration.

The performance of the GATRANSPFE shows that the GA approach is efficient and simple. Furthermore, the performance index results are significantly improved when the GATRANSPFE is applied.

In demand models, it is a common practice to consider intersection signals in isolation if signal delays need to be considered. A more rigorous approach should optimize signals on an arterial basis as TRANSYT does. Although non-signal traffic control devices such as stop and yield signs may also be included in a delay model, it is a reasonable assumption that delays due to such devices are relatively insignificant when they are placed on arterials.

Levinson and Kumar (1994) also develop a delay model that estimates the cycle time and green time using the methodologies suggested by Roess *et al*. (1998). The exact steps of the model are shown in Figure 2.4. The output of the intersection model is the average delay of a turning movement. The delay model is actually an application of Webster's formula. One important finding by Levinson and Kumar is that loading from highly aggregate zones to a single point will oversaturate the network at that point and seriously disrupt signal timings. The authors suggest that one zone per link with signal control at its head is necessary to accurately model intersections in a signal network.

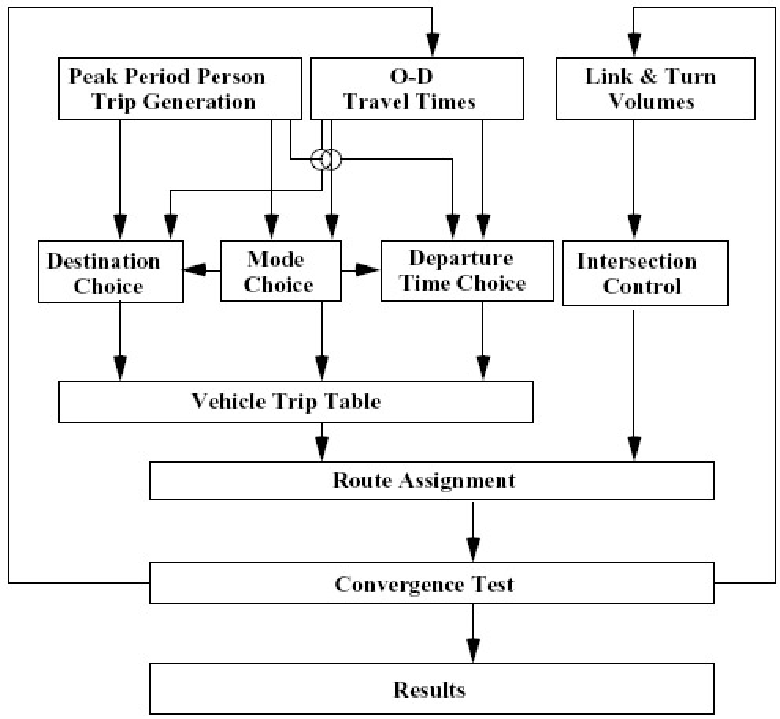

Figure 2.4 Transportation Planning Model with Feedback (Levinson and Kumar, 1994)

Ziliaskopoulos and Mahmassani (1996) apply the well known forward star structure (FSS) proposed by Dial *et al*. (1979) to decompose an intersection to a sub-network. The FSS makes it convenient to turn each intersection into as many nodes as approaches at the intersection, and as many arcs as all the possible movements in the intersection. The arc costs are either from the original network or the delay costs at the intersection. Thus, a four-leg intersection, as illustrated in Figure 2.5, is expanded into a sub-network of eight nodes, which is the number of all possible approaches to the intersection plus the original node, and 16 links, which are the arcs of the original network plus one arc for each possible intersection movement (including a U-turn for every approach). The sub-network does not have turning delays and prohibitions, and may be solved with any standard shortest path algorithm.

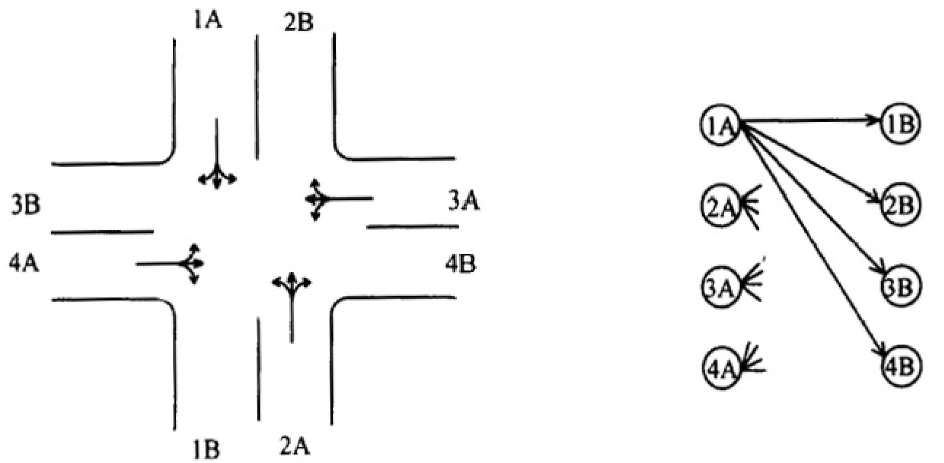

Figure 2.5 Decomposed Structure of a Four-Legged Intersection (Ziliaskopoulos and Mahmassani, 1996)

This approach has three main disadvantages: (1) it increases computational time and demand on computer memory due to the network expansion; (2) it takes extensive coding to change the network structure; and (3) it is inflexible to modify if the network topology and movement structure change. Therefore, the method may be impractical for practical applications considering its high requirement on memory and low computational efficiency. However, it provides a way to simplify network structure.

A more intuitive and practical method by Gartner and Al-Malik (1996) is promising for real-world applications. This method employs a solution procedure that enables the simultaneous optimization of the two problems: signal setting and link volume estimation. That is, the signal settings produce link costs that determine a flow pattern such that these settings are optimal for it. Signal settings are determined by a network optimization procedure, for example, MAXBAND or TRANSYT-7F, based on the traffic volume data previously estimated under the existing signal settings. The key of an efficient control strategy is to measure the effect of new signal timings, since drivers adjust to them thus resulting in new user-optimized traffic flow patterns. Gartner and Al-Malik's model simultaneously evaluates the route choice behavior of the motorist and determines the corresponding optimal signal settings, both are essential to rerouting traffic for the purpose of reducing congestion and avoiding bottlenecks.

Gartner and Al-Malik's model is among the first to introduce a way to expressing signal controls as flow variables in a deterministic manner. The procedure considers an individual signalized intersection with the following simplifications: (1) only two conflicting streams, (2) two-phase operation, (3) fixed cycle length given, and (4) one isolated intersection (offset is not considered). Therefore, it is easy to conclude that

$$g_1 + g_2 + L/C = 1$$

where

g_1, g_2 = splits of two phases
L = total intersection lost time (s)

C = signal cycle (s)

According to the Webster's formula, the optimal signal timing should be:

$$\frac{g_1}{g_2} = \frac{q_1 / s_1}{q_2 / s_2} \qquad \text{Eq. 33}$$

where
q = flows of two links
s = saturation flows of two signal phases

Eq. 33 is an approximate expression of the optimal division of the green time of a two-phase signal. This is known as the Webster's optimality ratio. According to the conditions, $(1 - L/C)$ is a fixed value. Therefore, g_1 and g_2 may be expressed as

$$g_1 = \frac{(1 - L/C)q_1 s_2}{q_1 s_2 + q_2 s_1} \qquad \text{Eq. 34}$$

$$g_2 = \frac{(1 - L/C)q_2 s_1}{q_1 s_2 + q_2 s_1} \qquad \text{Eq. 35}$$

According to these equations, the signal timings may be determined based on the flows at an intersection, and the signal settings and the flows are mutually consistent. Gartner and Al-Malik cite Webster's average delay function for a signalized approach:

$$d = 0.45[\frac{C(1-g)^2}{1-\frac{q}{s}} + \frac{3600x^2}{q(1-x)}] \qquad \text{Eq. 36}$$

where
g = green time split of the cycle
q = approach arrival flow (veh/h)
s = approach saturation flow rate (veh/h)
$x = q/gs$, the degree of saturation

Substituting g_1 and g_2 in Eq. 34 and Eq. 35, respectively, for the green split g in Eq. 36, the following average delay function of the signalized approach is obtained:

$$d = 0.45\left\{\frac{C[1-(\frac{\lambda q_1 s_2}{q_1 s_2 + q_2 s_1})]^2}{(1-\frac{q_1}{s_1})} + \frac{3{,}600(\frac{q_1 s_2 + q_2 s_1}{\lambda s_1 s_2})^2}{q_1[1-\frac{q_1 s_2 + q_2 s_1}{\lambda s_1 s_2}]}\right\} \qquad \text{Eq. 37}$$

Eq. 37 shows that the delay on a link is a function of not only the link flow but also the conflicting link flow. The difference between Eq. 36 and Eq. 37 is that the green times are not fixed but adjusted optimally in response to the traffic flows on the conflicting approaches at the intersection.

Having developed a flow-dependent signal control model, signal setting and traffic assignment procedures are ready to be combined into one inclusive model. The equilibrium is reached only when the necessary conditions of both aspects are met.

The traffic assignment aspect of Gartner and Al-Malik's application follows the generic methodology and assumptions. In other words, the users are always making wise and informed decision and the network's traveling cost cannot be reduced further.

Their proposed optimization function is

Min $Z(Q)$

Subject to the following conservation and non-negativity constraints:

$$\sum_k P_k^{mn} = f^{mn}$$

$$\sum_k P_k^{mn} \geq 0$$

The flow on a single link may be calculated as

$$q_a = \sum_m \sum_n \sum_k P_k^{mn} \delta_{ak}^{mn} \qquad \text{Eq. 38}$$

where

δ_{ak}^{mn} = 1 if link a is on path k and 0 otherwise

P_k^{mn} = flow on route k connecting O-D pair (m, n)

f^{mn} = trip demand rate between origin m and destination n

If $t_a(q_a, q_b)$ denotes the average travel time on link a (q_b denotes the conflicting flow on link b), the user equilibrium objective function is

$$Z(Q) = \frac{1}{2} \sum_a [\int_0^{q_a} t_a(w, q_b)dw + \int_0^{q_a} t_a(w,0)dw] \qquad \text{Eq. 39}$$

and the corresponding system optimization function is

$$Z(Q) = \sum_a q_a t_a(q_a, q_b) \qquad \text{Eq. 40}$$

Sheffi (1985) establishes two conditions that are required for the user-equilibrium problem to have a unique solution:

1. $\dfrac{\partial t_a(q_a,q_b)}{\partial q_a} > 0$

2. $\dfrac{\partial t_a(q_a,q_b)}{\partial q_a} > \dfrac{\partial t_a(q_a,q_b)}{\partial q_b}$

The objective functions of Eq. 39 and Eq. 40, satisfying neither condition, are hence non-convex. Although two simple network examples are presented with good performance, when confronting a complex network, it has to be determined to either search for the best in the multiple solutions or modify the network to converge to a single solution.

Compared to the study performed by Gartner and Al-Malik (1996), Lee and Machemehl (1999) made a further attempt to optimize the combined signal control and assignment problem. An iterative procedure is applied to a network with more realistic intersections than the two-phase ones discussed previously.

Because Wardrop's two principles define the user equilibrium (UE) and the system-optimized (SO) assignments, Lee and Machemehl suggest an iterative procedure to solve the combined problem of signal optimization and traffic assignment, which are treated as two sub-problems as shown in Figure 2.6.

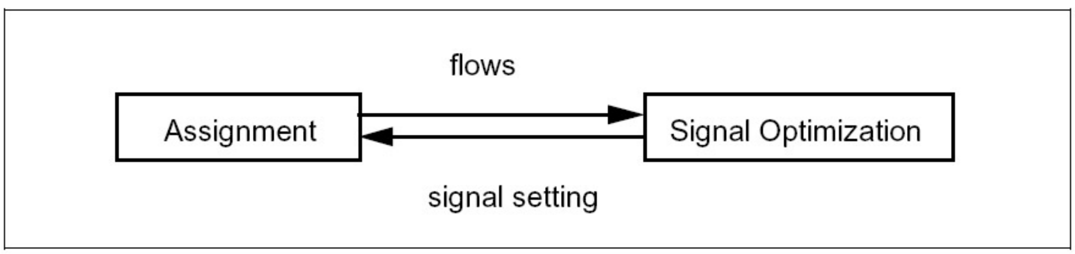

Figure 2.6 Iterative Optimization and Assignment Procedure

The assignment uses link performance functions resulted from signal optimization. Signal optimization is done with flow patterns provided by the assignment sub-problem. This is so called the Iterative Optimization and Assignment Procedure, or simply Iterative Approach (Cantarella, 1987). The procedure continues until it converges to one solution, which is called mutually consistent since the flow is at UE and the signal setting is optimal at the same time.

Similar to the study of Gartner and Al-Malik (1996), Lee and Machemehl (1999) utilize Webster's delay function in traffic assignment. Since the equilibrium network-traffic signal optimization problem may not be convex, it may have multiple local optimal solutions. Therefore, it is possible that some local and mutually consistent solutions will show poor performance compared to global solutions. The limitation of the methodology is that traffic assignments are steady-state with a fixed OD. Thus driver route choice rule is minimum time path selection so that drivers follow deterministic user equilibrium. The objective function of the model is to minimize the total travel time of the equilibrium network, that is

$$z = \sum_a t_a(x_a, \lambda_a) x_a \qquad \text{Eq. 41}$$

Subject to

$\lambda_{min} < \lambda < \lambda_{max}$

$x \sim UE$

$x > 0$ and $\lambda > 0$

where

λ_a = green time ratio for intersection link or movement a

t_a = travel time on link or movement a

x_a = flow on link or movement a

There are two difficulties in solving the above objective function (Lee and Machemehl, 1999). First, due to the problem of nonconvexity, z may have many local minima. Thus, any gradient based search will find only a local minimum. Second, z requires the knowledge of the OD pattern, which is difficult to develop for large, realistic networks. The iterative approach has been a practical alternative strategy.

To solve z, Lee and Machemehl use two approaches, namely local search and iterative approaches, to compare the optimal solutions. It is found that when the network is small, there may be only a few distinct local solutions, which may be found easily by local searches. Although the mutually consistent solution is intrinsically suboptimal, it is close to the local solutions for a small network when demand level is low. For a large network, there may be many local or quasi-local solutions. Thus, any local search may be easily trapped to worse solutions if the initial solution is not in a good domain neighborhood. Lee and Machemehl use a simplified method based on a gradient approximation suggested by Sheffi and Powell (1983). Since the iterative approach includes a signal optimization procedure, it finds a good solution showing small total travel time, which may not be mutually consistent until convergence, regardless whether the initial solution is in a good neighborhood or not. Then the search drifts to find a mutually consistent point. For a large network with high demand, there may be many mutually consistent points so that it is likely to find one close to the signal optimal point.

Simplified gradient estimation local searches show promising performance as well as computational efficiency. However, for large network with high demand, the iterative approach tends to find better solutions, and is more valuable to real-world applications. Note that nonconvexity makes it difficult to find a global optimum using any local search in large networks.

2.4.2 Signal Coordination

As Hurdle (1984) indicates, signal coordination often poses great difficulty to delay estimation, since coordinated intersections make it impossible to keep uniform arrival patterns for the downstream intersections. To cope with such difficulty, TRANSYT model provides an approach in which the arrival curve is predicted. The estimate of the nonrandom (or uniform) delay is based on this predicted arrival curve rather than on the HCM formula. Another simpler approach is proposed by Reilly (1976), which multiplies the estimate of the sum of uniform and overflow delays with an adjustment factor based on the coordination degree represented by five categories from poor to excellent. The sum of uniform and overflow delays are still obtained using the HCM formula. For travel demand models, the latter approach is more suitable due to its less demanding requirement for input data. However, many researchers have made great efforts to look for better models that consider signal coordination effects on area network performance.

The TRANSYT's traffic model may not be flexible enough sometimes because its limited signal settings cannot match well with the equilibrium flow pattern of a network. It becomes gradually accepted that traffic will adapt itself in a user optimal manner to changing control strategies. An individual driver tends to react to a new control strategy by selecting paths to minimize travel cost. Such redistribution effect on equilibrium traffic flow patterns will affect the performance of the network. Allsop (1974) is among the first who suggests that signal control may be explored to affect the traffic distribution and assignment on an equilibrium network. He also gives a rigorous mathematical framework for the problem. This issue has been known as equilibrium traffic signal settings in the literature.

For conventional network assignment problems, the cost function of a link is usually assumed to be influenced only by the attributes of that link, i.e., the travel time on a link is not affected by flows on other links in the network. However, to explicitly treat the effects of intersections in traffic assignment, the cost functions in traffic assignment problems are generally asymmetric and non-convex, making assignment problems difficult to solve. Considering signal coordination in the asymmetric assignment problems, Hall *et al*. (1980) develop a simulation-assignment model, SATURN. The major advantage of the SATURN model is that the difficult problem is transformed, at each assignment iteration, into a more manageable link-based traffic assignment problem with separable cost functions, which are solvable by an efficient algorithm. However, the SATURN is outperformed by the path based algorithm presented by Wong *et al*. (2001). The method of Wong *et al*. determines an equilibrium traffic pattern by taking into account the signal coordination effects and platoon dispersion on the streets. In their paper, a numerical example is given to demonstrate the effectiveness of the proposed methodology, in which details of a comparison with the SATURN diagonalization method are provided. The path-based methodology provides a promising approach to solving the difficult asymmetric traffic assignment problem based on TRANSYT models.

2.4.3 Applicable Software

According to Traffic Analysis Tools Primer (Alexiadis *et al*., 2004), most analytical/deterministic tools implement the procedures of the HCM. These tools quickly predict capacity, density, speed, delay, and queuing for a variety of transportation facilities and are validated with field data, laboratory test beds, or small-scale experiments. Analytical/deterministic tools are good for analyzing the performance of isolated or small-scale transportation facilities. However, they are limited in their ability to analyze network or system effects.

For many applications, the HCM is the most widely used and accepted traffic analysis technique in the U.S. The HCM procedures are good for analyzing the performance of isolated facilities with relatively moderate congestion problems. They are quick and reliable for predicting whether a facility will be operating above or below capacity, and they have been well tested through significant field-validation efforts. However, the HCM procedures are generally limited in their ability to evaluate system effects. Most of the HCM methods and models assume that the operation of one intersection or road segment is not adversely affected by conditions on adjacent roadways. Long queues at one location that interferes with another location would violate this assumption. The HCM procedures are of limited value in analyzing queues and their effects. The HCM are limited in the following situations:

- Multi-lane or two-lane rural roads where traffic signals or stop signs significantly impact capacity and/or operations
- Climbing lanes for trucks
- Short through-lane added or dropped at a signal
- Two-way left-turning lanes
- Roundabouts of more than a single lane
- Tight diamond interchanges

If the HCM procedures do not meet the needs of an intended analysis, it needs to be determined whether microscopic, mesoscopic, or macroscopic simulation is required or not. If it is not necessary to microscopically trace individual vehicle movements, then the analyst may take advantage of the simpler data entry and control optimization features available in many mesoscopic or macroscopic simulation models.

For comprehensive traffic analysis functions including signal timing optimization and signal coordination, Synchro and TRANSYT-7F have been widely applied. TRANSYT-7F has been popular since the 1980s, and many extensions have been produced for various customized applications. Some research (Wong, 2001) has indicated that TRANSYT-7F is usable to model intersection delays while considering coordination effects.

Synchro has a friendly user interface for most traffic analysis of signals. Moreover, Synchro has a versatile utility for data transfers between many mainstream traffic software packages including HCS, CORSIM, and TRANSYT-7F. Synchro produces results in Universal Traffic Data Format (UTDF), which is convenient for use for other software packages.

Synchro uses two methods for calculating delays, one is based on the Webster's formula, and the other, a newer one, is called the Percentile Delay Method. Assuming that traffic arrivals will vary according to a Poisson distribution, the Percentile Delay Method calculates vehicle delays for five different scenarios and takes a volume weighted average of the five delays. The five scenarios are the 10th, 30th, 50th, 70th, and 90th percentiles of the service volume. It is assumed that each of these scenarios will be representative of 20% of the possible cycles of signal phases. For each scenario, traffic for each approach is adjusted to that percentile. Delays are calculated using the adjusted volumes, and green times are calculated. If the intersection is near-saturation or oversaturation, additional time will be added to account for vehicles carried over between cycles. In the Webster's method, the Progression Factor (PF) is used by the HCM to account for the effects of coordination. Since Synchro has already explicitly calculated the delay with coordination effects, the true PF may be determined with the following formula:

$PF = DelayCoord / DelayUnCoord$ Eq. 42

where

DelayCoord = uniform delay calculated by Synchro with coordination

DelayUnCoord = uniform delay calculated by Synchro assuming random arrivals

The accuracy of delay calculation is expected to be greatly improved.

CUBE, a travel demand model software package by Citilabs, is capable of considering intersection delays. The delay estimates are by default based on the HCM algorithm. It requires the coding of intersection geometry, as well as signal timing plans.

2.5 Summary

Although the intersection delay estimation technique of the HCM has been widely applied, it is based on curve fitting rather than a sound mathematical model of signal systems. Therefore, when signal systems operate under oversaturated conditions (v/c = 1), many traffic conditions are not well modeled. However, reasonable results are possible under the condition that the users are aware of the model limitations. None of the deterministic or steady-state models are able to produce fully consistent or accurate results. Although not always correct, it has been generally agreed that most steady-state delay models and deterministic models described here generate relatively consistent delay estimates when employed for undersaturated signalized intersections with v/c ratios below 0.6 (Dion *et al.*, 2004). To develop a new generation of models that reasonably consider variations of travel demand over time, more information on traffic patterns is needed. At present, it may be unrealistic to expect the availability of such information.

Further research needs to be directed at evaluating more complex situations such as multi-lane approaches, intersections with actuated controllers, and intersections where non-random vehicle arrivals occur as a result of signal coordination with upstream intersections. The HCM delay model is unable to consider progression effects well as it assumes that intersections are isolated. Therefore, it is a common practice for a traffic assignment model to assume that an intersection is isolated if the estimates are made based on the HCM's delay model (Gartner and Al-Malik, 1996; Lee and Machemehl, 1999). The prediction curve based on the TRANSYT traffic model developed by Robertson (1977) has been widely accepted as an effective tool for evaluating queues and delays on links in a network. For a more flexible specification of signal timings, Wong (1996) uses a group-based method that allows the traffic redistribute itself in a user optimal manner (individual cost minimized). The methodology may be a tool potentially applicable for planning purposes.

3. OVERVIEW OF THE METHODOLOGY

The methodology developed in this study is illustrated in Figure 3.1. The box in the upper left corner is the process during which a dataset of traffic volumes and delays is created. The dataset is stored in a table format, as shown to the right. Based on this dataset, intersection delay models will be calibrated, which will predict the intersection delay for given volumes at an intersection. Finally, this model is applied during the FSUTMS traffic assignment iteration process. At each iteration of traffic assignment, the current volumes will be provided to the delay models, which will in turn produce the delay times for nodes representing intersections. The delay times will then be used to update the path costs in the next traffic assignment iteration.

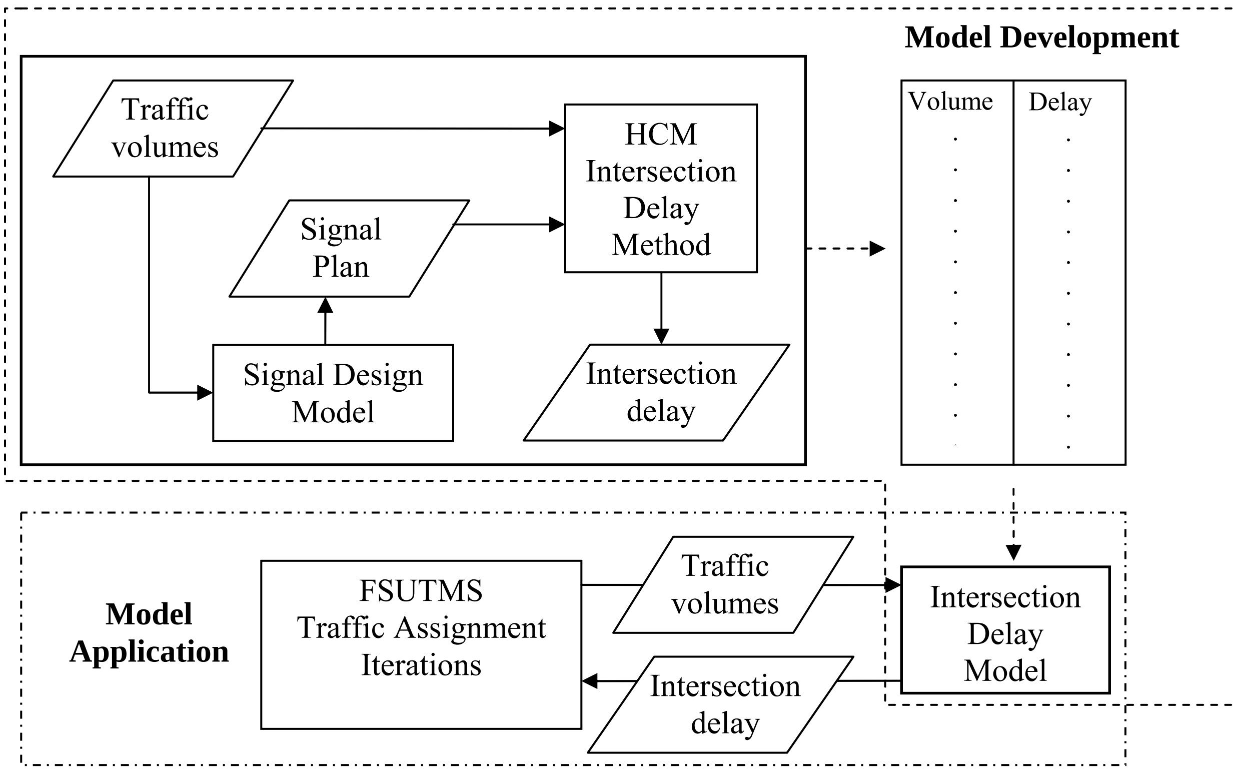

Figure 3.1 Conceptual Process of the Proposed Methodology

There are several research issues that need to be considered and addressed in the methodology:

1. Simplifying assumptions. Actual intersections may be quite complex. It is neither realistic nor necessary to attempt to model every possible roadway geometry type. Therefore, typical intersections need to be identified. These "classes" of intersections may be formed based on intersection geometry, the facility types of approaches, parking (e.g., on-street parking), etc. The groupings of different intersections into typical classes need to be based on statistical analysis to determine if the volume-delay relationships are similar statistically within a group. Criteria for classification also need to be developed. A balance needs to be maintained between the number of typical classes (complexity) and delay model accuracy.

2. Assumptions about signal coordination. At many locations in an urban area, signal plans are not designed in isolation. Signal coordination is common when there are closely spaced signalized intersections. The simultaneous consideration of multiple signals or signal progression will increase the model complexity. Similar to classifying intersections into typical ones as discussed above, criteria also need to be established regarding signal coordination. These criteria will allow assumptions to be made for model application as whether coordinated signals should be assumed or not.

3. Delays for different movements. In the FDOT LOS manual, the LOS is established based on the estimated delay for the dominant movement. In a travel model, to build minimum path trees, delays for all possible movements at an intersection will be necessary.

4. Time-of-day delay model. Although the proposed delay model will predict delays for typical intersections mostly based on volume assuming an optimal signal timing plan, in practice, signal timing plans are likely to be designed to vary during different time of day such as morning peak hours, evening peak hours, midday off-peak hours, evening and morning off-peak hours, and night. These different signal timing plans will need to be considered for peak and daily traffic assignment at a minimum, and for different model periods ideally. Research is necessary to identify the relationship between the actual signal timing plans and signal plans produced based on simulation since the latter will be based on aggregate traffic conditions.

5. Use of field data. While optimized signal timing plans and the HCM method will provide delay estimates for typical intersections, it may be useful to study the actual signal timing plans in different urban areas to evaluate possible differences in traffic characteristics thus delays. To evaluate the model performance, it is also desirable to compare the model predictions to actual delays. However, this will require extensive data collection.

6. LOS Estimation. The FSUTMS model output is used to estimate LOS of facilities to assess the adequacy of facilities and identify needed improvement projects. LOS is determined based on the methodologies described in the FDOT's Quality/Level of Service Handbook. The delay model developed in this research may potentially be used to assess the suitability and accuracy of the methodologies in the FDOT's Quality/Level of Service Handbook.

Due to the constraints on time and resources, this research scope focuses on the development of a delay model for isolated intersections as a first step and as a proof of concept. The methodology is designed to produce delay models that are capable of estimating intersection delays with adequate accuracy based on inputs directly available from FSUTMS models. Several possible types of modeling techniques have been considered for developing the delay models. These include Artificial Neural Network (ANN) models and multiple linear regression models, although the results show that the ANN models are superior to the regression models. ANN models are, therefore, chosen to be the delay models. The delay models were developed with the following procedures:

1. Identify "Typical Intersections"

All the intersections in the urban area of Gainesville are extracted using from the 2000 Gainesville model. The five most common types of intersections based on facility type and number of lanes are identified.

2. Generate Scenarios for All Types of Intersection

   The ANN delay models need to "learn" from various volume conditions to be able to predict delays accurately. Therefore, a large number of scenarios of traffic volumes and delays are created through simulation using TRANSYT-7F. A scenario simulates an intersection with certain volumes on every approach. There are eight movement delay estimates for one scenario. Four delays are for left-turning movements for each of the four legs, and the other four are for the through movements. These scenarios include most possible traffic conditions to ensure the ANN delay models acquire the ability to generalize. To create valid scenarios, the short term traffic counts at Gainesville urban area from Florida Traffic Information 2004 are examined. It has been found that the daily traffic counts approximate a normal distribution. Therefore, the traffic volumes are assumed to follow the same normal distribution. A total of 150 scenarios are created for each intersection type. Combined with 12 randomly selected turning ratios, 1,800 scenarios for ANN models are created for every intersection type.

3. Develop ANN Delay Models

   Five major intersection types are tested in this study. For each of the five types of intersections, two ANN delay models are developed. One is for the left-turning traffic, and the other is for the through traffic. The models estimate the control delay for each movement using simulated volumes of all approaches at an intersection. The models are tested using separately created datasets that are not used in the training of the ANN delay models.

4. Evaluate the ANN Delay Models

   The performance of the ANN delay models is analyzed. Statistics are calculated to evaluate the performance of the ANN delay models. The limitations of the models are discussed.

The procedures and results for each aforementioned task are described in detail in the following chapters.

4. CLASSIFICATION OF INTERSECTIONS

To establish the unique types of intersections and identify most common intersections, information on intersections are extracted from the 2000 Gainesville planning model, which includes facility type, area type, and number of lanes. However, information about turning bays or other geometry conditions is not available. Such data are not required considering the applications of the models will be for long-range planning.

Table 4.1 summarizes all the intersections in the Gainesville network model based on the link facility type, area type, and number of lanes. The definitions of one-digit facility types and area types are given in Appendix A. Each row in the table represents one combination of these three link properties. For instance, the second record has an intersection type 212322, indicating that one road (represented by two links for two opposing traffic) is of facility type 2 (divided arterial) and area type 1 (CBD) with two lanes and that the cross road is of facility type 3 (undivided arterial) and area type 2 (CBD fringe) with two lanes. Columns three to eight explain the type by giving the facility type, area type, and lane number, respectively.

Table 4.1 Types of Intersections Based on Combinations of Facility Type, Area Type, and Number of Lanes in the Gainesville Network

Type	COUNT	Road 1			Road 2		
		FT	AT	Lanes	FT	AT	Lanes
212212	1	2	1	2	2	1	2
212322	1	2	1	2	3	2	2
212411	6	2	1	2	4	1	1
212412	1	2	1	2	4	1	2
212421	2	2	1	1	4	2	1
222421	3	2	2	2	4	2	1
231231	3	2	3	1	2	3	1
231232	1	2	3	1	2	3	2
231331	4	2	3	1	3	3	1
231332	1	2	3	1	3	3	2
231431	7	2	3	1	4	3	1
231432	3	2	3	1	4	3	2
232222	1	2	3	2	2	2	2
232231	4	2	3	2	2	3	1
232232	15	2	3	2	2	3	2
232233	3	2	3	2	2	3	3
232331	8	2	3	2	3	3	1
232332	2	2	3	2	3	3	2
232421	7	2	3	2	4	2	1
232431	53	2	3	2	4	3	1
232432	8	2	3	2	4	3	2
232731	2	2	3	2	7	3	1
233232	3	2	3	3	2	3	2
233233	1	2	3	3	2	3	3
233331	2	2	3	3	3	3	1
233431	12	2	3	3	4	3	1

233432	2	2	3	3	4	3	2
233731	6	2	3	3	7	3	1
312612	2	3	1	2	6	1	2
322421	1	3	2	2	4	2	1
322432	1	3	2	2	4	3	2
331331	1	3	3	1	3	3	1
331431	19	3	3	1	4	3	1
331432	1	3	3	1	4	3	2
331731	1	3	3	1	7	3	1
332431	6	3	3	2	4	3	1
332432	1	3	3	2	4	3	2
332731	2	3	3	2	7	3	1
411411	4	4	1	1	4	1	1
411421	4	4	1	1	4	2	1
411422	1	4	1	1	4	2	2
411612	2	4	1	1	6	1	2
421411	1	4	2	1	4	1	1
421421	23	4	2	1	4	2	1
421431	2	4	2	1	4	3	1
421622	3	4	2	1	6	2	2
422421	1	4	2	2	4	2	1
431421	2	4	3	1	4	2	1
431431	64	4	3	1	4	3	1
431432	1	4	3	1	4	3	2
431632	1	4	3	1	6	3	2
432431	8	4	3	2	4	3	1
432432	1	4	3	2	4	3	2
Total	315						

There are a total of 54 records in the table, indicating 54 unique types of intersections that may be modeled. The effects of area types on intersection delay mainly come from intersection spacing, operating speed (e.g., due to link length), presence of pedestrian traffic, etc. Currently, these factors cannot be explicitly considered in FSUTMS. Similarly, a signal optimizer such as TRANSYT-7F and Synchro is unable to differentiate among area types except for central business districts (CBD). Therefore, area types are not considered at present and intersection types are defined solely based on facility types and number of lanes. In the future, when signal spacing and signal progression are considered, the area types may be included in intersection delay calculations.

Table 4.2 gives the nine types of intersections found in the Gainesville network based on the unique combinations of the facility types of two cross roads. The most common types of intersections are those involving divided arterials and local collectors and those of undivided arterials and local collectors.

Table 4.2 Types of Intersections Based on Combinations of Facility Type in the Gainesville Network

Type	Road 1 Facility	Road 2 Facility	Count
22	2	2	32
23	2	3	18
24	2	4	112
33	3	3	1
34	3	4	29
36	3	6	2
37	3	7	3
44	4	4	112
46	4	6	6
Total			315

When considering the number of lanes, the number of unique types of intersections increases to 27, as shown in Table 4.3. The most common types of intersections are two local collectors with one lane (100), divided two-lane arterial with one-lane local collector (71), one-lane undivided arterial with one-lane local collector (19), and two divided two-lane arterials (17). The six intersections of three-lane divided arterial with two-lane divided arterial are in the busiest urban area of Gainesville. These five intersection categories are the focus of this study, for which ANN delay models are developed.

To learn about the local traffic characteristics and geometry conditions, a site visit was made in Gainesville. Some unexpected characteristics are found. For instance, frequent pedestrian crossing activities are often seen along a few arterials surrounding the University of Florida. Relatively lower speed limit (30 or 35 mph) is often posted for major arterials. The signal progression is widely employed in arterials surrounding the university. However, currently, these findings contribute little to the default setup of TRANSYT-7F's simulations. Pedestrians phasing and signal progression are not considered in this study.

Table 4.3 Types of Intersections Based on Combinations of Facility Types and Numbers of Lanes in the Gainesville Network

Type	Road 1		Road 2		Count
	Facility	Lane	Facility	Lane	
2121	2	1	2	1	3
2221	2	2	2	1	5
2222	2	2	2	2	17
2322	2	3	2	2	6
2323	2	3	2	3	1
2131	2	1	3	1	4
2132	2	1	3	2	1
2231	2	2	3	1	8
2232	2	2	3	2	3
2331	2	3	3	1	2
2141	2	1	4	1	7
2142	2	1	4	2	3
2241	2	2	4	1	71
2242	2	2	4	2	9
2341	2	3	4	1	12
2342	2	3	4	2	2
2271	2	2	7	1	8
3131	3	1	3	1	1
3141	3	1	4	1	19
3142	3	1	4	2	1
3241	3	2	4	1	7
3242	3	2	4	2	2
3262	3	2	6	2	2
3171	3	1	7	1	1
3271	3	2	7	1	2
4141	41	1	4	1	100
4241	4	2	4	1	11
4242	4	2	4	2	1
4162	4	1	6	2	6
Total					330

5. GENERATION OF INTERSECTION SCENARIOS

To build a model of intersection delays, a traffic signal optimizer and a traffic simulator are needed to obtain the delay for each movement for different combinations of approach volumes. In this study, the traffic signal optimizer is employed to determine the optimal signal-phasing and timing plans for isolated signal intersections, and the macroscopic traffic simulator is for delay estimation. There are many programs capable of both optimizing signal plans and estimating control delays. For instance, PASSER, TEAPAC/Signal2000, Synchro, and TRANSYT-7F, to name a few, are commonly used. Passer, Synchro, and TRANSYT-7F are also macroscopic simulation models. In other words, they are able to obtain traffic characteristics based on optimization results. They allow signal timing optimization and simulation (delay estimation). TRANSYT-7F is chosen to be the signal optimizer and simulator for this study based on the criteria shown in Table 5.1. The main reasons for choosing TRANSYT-7F are its ability to run in batch mode, produce good results, and enjoy good technical support.

Table 5.1 Selection Criteria for Signal Optimizer/Simulator

Software	Windows/DOS Interface	Batch Mode	Technical Support	Cost-Effectiveness
PASSER	DOS	Yes	Normal	N/A
Synchro	Windows	No	Poor	Low
TRANSYT-7F	Both	Yes	Excellent	High

To validate the output from the TRANSYT-7F simulations, selected delay results based on medium to saturate conditions are compared to those from HCS2000-Signals. The results indicate that there is a close match between the delay estimates from TRANSYT-7F and HCS2000-Signals.

TRANSYT-7F has evolved into a powerful signal optimizer. However, to achieve its full potential in optimization of signal phasing, a genetic algorithm must be applied, which requires a large population size and large number of generations, incurring considerable computing time. Another problem is that TRANSYT-7F may not always produce delay estimates similar to those from the HCM due to inherent difficulties in modeling oversaturated intersections. The delay for an oversaturated approach sometimes reaches over 700~800 seconds when the corresponding v/c ratio of the lane group is in the range of 1.5~1.6, which is unrealistic. According to the TRANSYT-7F User’s Guide (McTrans 2006), TRANSYT-7F is capable of modeling oversaturated situations better if multi-period analysis is selected. In this study, single-period analysis instead of multi-period analysis is performed with the genetic algorithm. Single-period analysis refers to signal optimization that only considers travel conditions within the studied period, while multi-period analysis incorporates time-varying conditions including traffic volume and timing plan variations. Multi-period analysis provides improved results than single-period analysis does when dealing with oversaturated intersections, because residual queues and residual delays across multiple time periods are explicitly considered. However, multi-period analysis takes three to four times longer than single-period does, and it is important only in oversaturated situations. Therefore, single period analysis technique is chosen primarily based on the computing resource constraint since the number of the simulated intersection scenarios analyzed is in the order of thousands instead of a few.

After TRANSYT-7F optimizes the signal timing and estimates the corresponding delays at an intersection, the inputs for the ANN delay models are extracted from the TRANSYT-7F outputs. Developing an ANN requires large amounts of training data to obtain a well-trained model. To obtain control delays at intersections of different geometry conditions, data sets are developed to represent most traffic conditions at an intersection. Because a generic four-leg intersection has 12 movements, for which there are an infinite number of possible volume conditions, it is impractical to enumerate all possible volume conditions of all the movements at an intersection. One simplification used in most signal optimizers such as Synchro and TRANSYT-7F is to combine right-turn traffic with through volume if they share the same lane. This reduces the 12 movements to eight movements. In other words, a simulation scenario of an intersection has eight samples of movements (four left-turning and four through movements), therefore eight samples of delay estimates.

In addition to the volumes and delays, the capacities of the four cross roads are also needed as input to the ANN models. The capacities are assumed based on the approach's facility type, area type, and lane number defined in the user manual of FSUTMS (FDOT, 1997), and are given in Table 5.2 below.

Table 5.2 Network Link Capacity for Roads of Different Facility Types and Lanes

Intersection Type	**Base Capacity (vphpl)**
2322	755
2222	750
2241	750/530
3141	592/530
4141	530

To create the training datasets, a total of 150 combinations of through volumes are created. Because ANN models are sensitive to the distribution of traffic volumes, to ensure the volumes reflect the real traffic conditions, historical traffic counts from the Florida Traffic Information (FTI) 2004 CD are examined. The peak-hour traffic counts are extracted from 88 Portable Traffic Monitoring Sites (PTMS) for intersections of facility type 2 in the Gainesville urban area (Figure 5.1). The peak-hour traffic counts per lane are found to approximately follow a normal distribution curve, with a mean μ = 462.72 and a standard deviation σ = 135.83, as shown in Figure 5.2. Therefore, the normal distribution is assumed for intersections of facility type 2 when creating the simulation scenarios. Because no PTMSs are available for undivided arterials (facility type 3) and local collectors (facility type 4), it is assumed that the peak-hour traffic for these facility types also follow a similar normal distribution. The μ and σ is slightly adjusted, so that [μ-3σ, μ+3σ] contains the assumed base capacities. This range represents the 99% confidence interval of a normal distribution. In other words, the volumes selected based on this normal distribution will fall within this range with a 99% probability. Table 5.3 gives the normal distribution parameters for different intersection types.

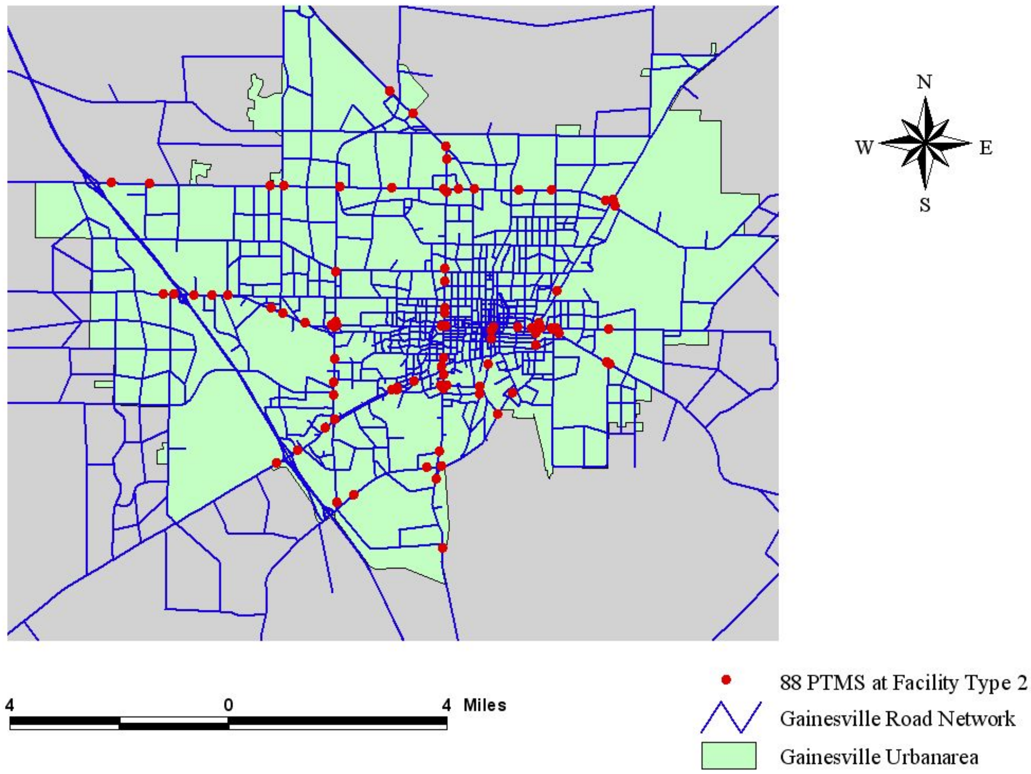

Figure 5.1 Locations of 88 PTMS for Divided Arterials in the Gainesville Urban Area

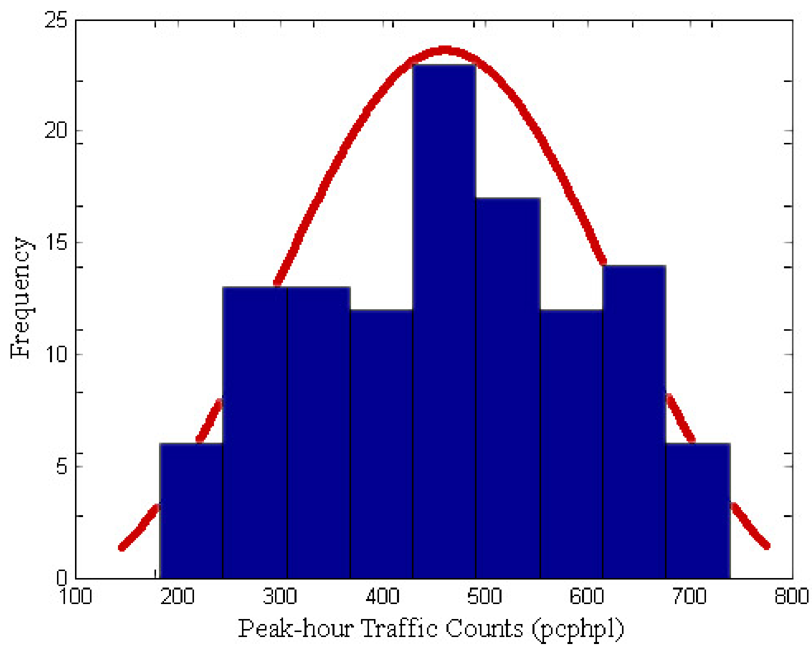

Figure 5.2 Distribution of Peak-hour Traffic Counts of 88 PTMS for Divided Arterials

Table 5.3 Normal Distribution Parameters for Volumes by Intersection Type

Intersection Type	**Normal Distribution**
2322	μ = 462.72, σ = 135.83
2222	μ = 462.72, σ = 135.83
2241	μ = 462.72, σ = 135.83
3141	μ = 400, σ = 125
4141	μ = 400, σ = 125

The 150 combinations of through traffic volumes are created by sampling from normally distributed volumes within the range of [μ-3σ, μ+3σ] using random numbers. For all types of intersections in this study, the heaviest through volume simulated is approximately 1.4 times of the corresponding approach's base capacity.

The simulation scenarios also require left-turning volumes. Assuming seven turning ratios for generating different combinations of turning traffic, which are 10%, 15%, 20%, 25%, 30%, 35%, and 40%, the number of combinations of four turning ratios is 7^4 = 2,401. However, because an ANN model does not need to be trained with all possible situations to make predictions, not all of the 2,401 turning ratio combinations have to be considered in training. It is found that ANN models do not seem to be sensitive to small variations in turning ratios under fixed through traffic conditions. For instance, for an intersection with a turning ratio combination of [0.1, 0.1, 0.25, 0.3] the movement delays are not significantly different from those based on a turning ratio combination of [0.1, 0.1, 0.2, 0.3] when the four through volumes are fixed. Therefore, only 12 turning ratio combinations are randomly selected from the 2,401 to combine with the through volumes to create the ANN training scenarios. The performance of ANN delay models has proven that these turning ratios are adequate for the analysis.

To verify that 150 through volumes and 12 turning ration combinations are adequate, a number of the ANN delay models are developed for intersection type 2241 by varying the numbers of through volume combinations and turning ration combinations in a training data set. Figure 5.3 shows the ANN model performance versus the number of through volume combinations in the training data set. It may be seen that the ANN model's performance improves little after 100 volume combinations. Figure 5.4 shows that the ANN model's performance does not change significantly after the number of turning ratio combinations exceeds eight. Therefore, in this study, 150 through volume combinations with 12 turning ratio combinations are adequate for training ANN models.

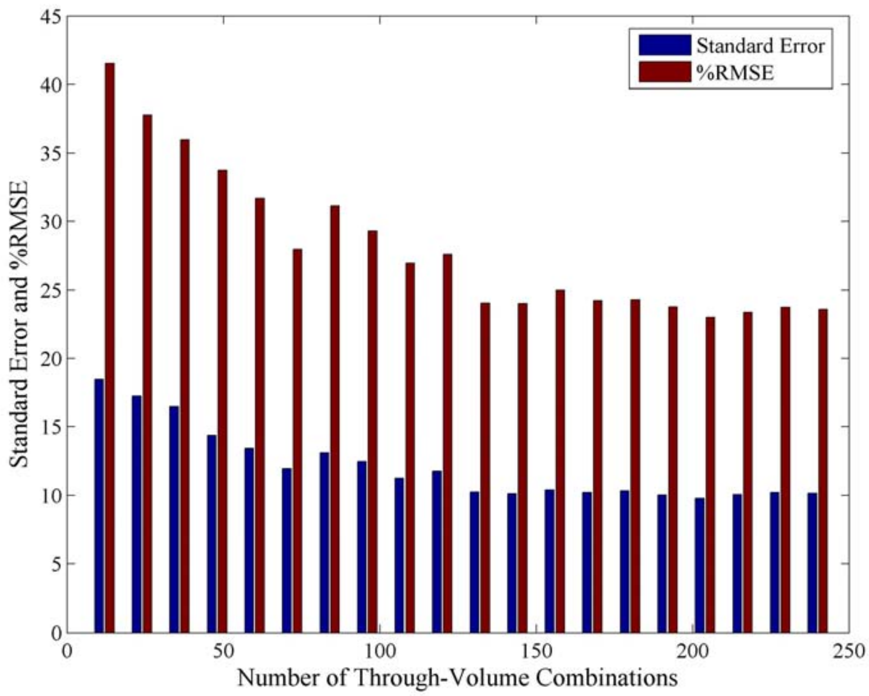

Figure 5.3 ANN Performance vs. Number of Through Volume Combinations

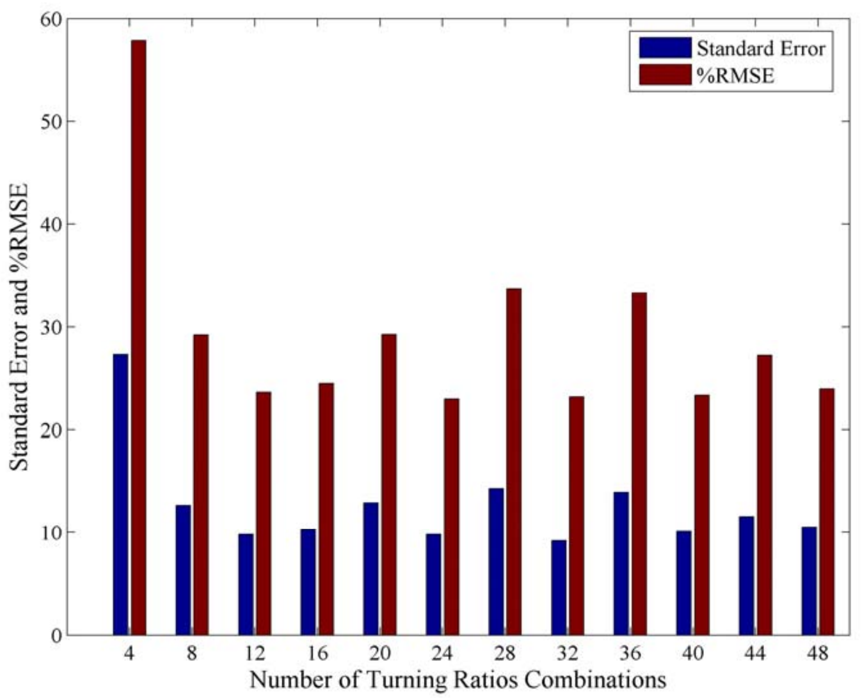

Figure 5.4 ANN Performance vs. Number of Turning Ratios Combinations

The left-turning volumes are therefore obtained by multiplying the through volumes with a chosen combination of turning ratios. To create the traffic volumes for a scenario, each of the 150 through volume combinations is combined with one of the 12 left-turning volumes computed based on the 12 turning ratios. Thus, for each intersection type, the number of through and left-turning volume combinations is 150 x 12 = 1,800. Since each through volume for an intersection approach is independently created based on a normal distribution, the total number of simulation scenarios for each intersection type is 1,800 x 4 = 7,200. These 7,200 scenarios are simulated by TRANSYT-7F, and the output delays for the eight movements, along with the input that defined the scenarios, form the training data sets for the ANN models.

6. DEVELOPMENT OF ANN DELAY MODELS

This chapter first introduces the architecture of the ANN delay models. The performances of the ANN models are then examined. A series of multiple regression models are calibrated to compare with the ANN models.

6.1 Architecture of ANN Delay Models

Learning rules, which determine the training algorithm of ANN models, are important to ANN model performance. Among the commonly used learning rules, backpropagation trains a multilayer feed-forward network with differentiable transfer functions to perform function approximation, pattern association, pattern classification, as well as a number of optimization strategies (Demuth *et al*., 2006). The term *backpropagation* refers to the process by which derivatives of network errors with respect to network weights and biases may be computed. The architecture of a multilayer network is not completely constrained by the problem to be solved. It has been suggested that a two-layer (sigmoid/linear) network may represent any functional relationship between inputs and outputs, provided that enough neurons are used (Demuth *et al*., 2006). In this study, the ANN models have two layers of neurons. As Figure 6.1 shows, one layer using sigmoid transfer function (see Figure 6.2) handles inputs vector **p** that is weighted by vector **w.** The second layer is the output layer that produces result A, which follows a linear relation as Figure 6.1 indicates. Thus, the network models an approximate mathematical relation:

$$A = \mathrm{f}\,(wp + b) \qquad \text{Eq. 43}$$

where
A = ANN output
w = weight assigned to inputs
p = inputs
b = adjusting bias

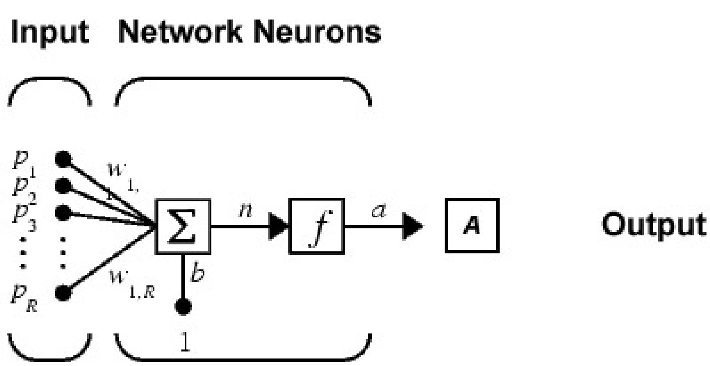

Figure 6.1 A Typical Feed-forward ANN Architecture (Demuth *et al*., 2006)

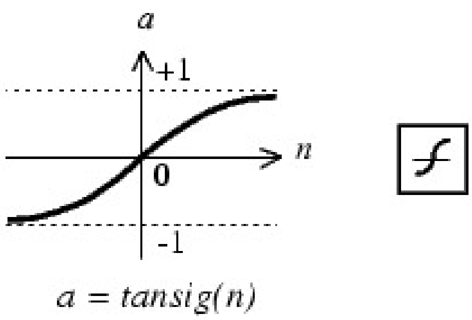

Figure 6.2 A Sigmoid Transfer Function of ANN Layer (Demuth *et al*., 2006)

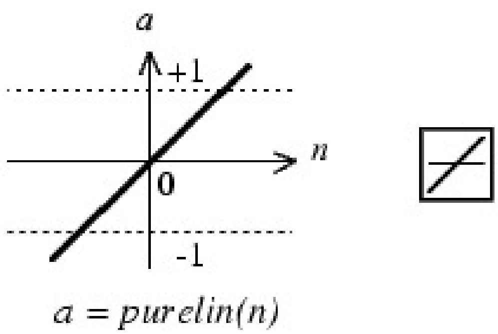

Figure 6.3 A Linear Transfer Function of ANN Output Layer (Demuth *et al*., 2006)

The number of neurons in the sigmoid layer is required to exceed that of the inputs (Demuth *et al*., 2006). By trial-and-error, one sigmoid layer with 50 neurons is determined as the best for the ANN models. The number of inputs to the network is determined by the problem at hand, and the number of neurons in the output layer is based on the number of outputs required. However, the number of layers between network inputs and the output layer, as well as the sizes of the layers, is to be determined by the analyst.

From FSUTMS models, only movement volumes may be extracted to use as the ANN inputs to predict delays. Since TRANSYT-7F considers through and right-turning volumes in the same lane group, there are eight movement volumes and two link capacities (assuming symmetric intersection configurations) that may be used as the ANN inputs for a four-leg intersection:

- through volume of the approach for which the delay is to be estimated (v_{11})
- left-turning volume of the approach for which the delay is to be estimated (v_{12})
- through volume of the opposing link (v_{21})
- left-turning volume of the opposing link (v_{22})
- through volume of the right crossing link 1(v_{31})
- left-turning volume of the right crossing link 1(v_{32})
- through volume of the left crossing link 1(v_{41})
- left-turning volume of the left crossing link 1(v_{42})
- link capacity of the approach for which the delay is to be estimated (c_1)
- link capacity of the crossing road (c_2)

Figure 6.4 illustrates the spatial relationship of movement volumes, link capacities, and the corresponding delay estimate.

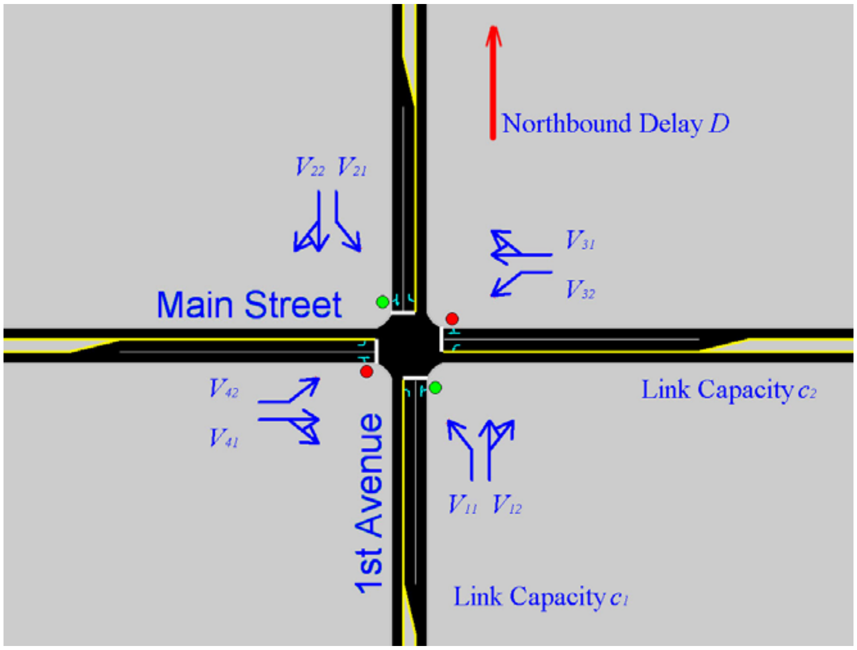

Figure 6.4 Spatial Relationship of Input Variables for the ANN Delay Models

To determine the significance of the variables, a correlation analysis is performed between the delay for one movement and the rest of the movement volumes. Table 6.1 shows the linear correlation coefficients of northbound delays and the eight volumes at significance level of 99%. It cannot be concluded that the eight volumes are uncorrelated with the delay.

Table 6.1 Correlation Coefficients of Delays and Eight Movement Volumes

Correlation Coefficients	v_{11}	v_{12}	v_{21}	v_{22}	v_{31}	v_{32}	v_{41}	v_{42}
Through Delay	0.523	0.109	0.306	0.176	0.325	0.330	0.276	0.255
Left-turning Delay	0.462	0.123	0.314	0.195	0.192	0.175	0.259	0.272

As mentioned before, the backpropagation learning rule is used in this study to train the ANN models. Several training algorithms implement the backpropagation learning rule. Levenberg-Marquardt algorithm is usually the fastest. It provides a memory reduction feature when the training dataset is large. Several other training algorithms are also considered. Requiring no line search, scaled conjugate gradient algorithm is a good general purpose training algorithm. Bayesian regularization algorithm is a modification of the Levenberg-Marquardt training algorithm to produce networks that generalize better. It reduces the difficulty of determining the optimum network architecture.

The performances of ANN models using the above three training algorithms are examined for all intersection types. It is found that scaled conjugate gradient usually produce relatively better fit compared to the other two, based on three evaluation criteria: the regression R-squared value, which is a statistical measure of how well predictions approximate real data points; root-mean-square error (RMSE); and percent root-mean-square error (%RMSE). The RMSE, also known as the standard error, represents the average error of model predictions. The %RMSE is a statistic indicating the percentage of the average expected error in the actual value, and is adopted by the FDOT as a criterion in calibrating travel models. The formulas for computing the RMSE and the %RMSE are given below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i^c - x_i^v\right)^2}{n-1}} \qquad \text{Eq. 44}$$

$$\%RMSE = \frac{RMSE}{\sum_{i=1}^{n} x_i^c \Big/ n} \qquad \text{Eq. 45}$$

where

- x^c = delay estimates from TRANSYT-7F simulation (s/veh)
- x^v = delay predictions by ANN delay models (s/veh)
- n = number of ANN inputs (i.e., simulated scenarios)

For all types of intersections, the ANN delay models that apply scaled conjugate gradient algorithm usually demonstrate relatively better performance than the other two. Table 6.2 shows the performance statistics of the ANN delay model for intersection type 2241. Based on the performance statistics, multilayer feed-forward ANN architecture with scaled conjugate gradient algorithm is applied to develop the ANN delay models.

Table 6.2 Performance Statistics of Three Training Algorithms for 2241 Type of Intersection

Training Algorithm	Regression R-squared	RMSE	%RMSE
Scaled Conjugate Gradient	0.855	8.34	18.77
Bayesian Regularization	0.767	10.47	22.23
Levenberg-Marquardt	0.807	9.55	20.27

6.2 Performance of the ANN Delay Models

In training the ANN delay models, the input data are divided into three groups: training data, testing data, and validation data. The ANN training employs "supervised learning," that is, the training process is simultaneously supervised by the scaled conjugate gradient training algorithm that utilize the validation data to the trained ANNs to correct potential overfitting. After the training of the ANN models, the testing data are used to evaluate the ANN models' performance. Theoretically, the ANN models should not have encountered the testing data during the training process so that the performance of the ANN is reliable. In this study, to measure the ANN model's capability precisely, around 10% of 7,200 input samples for ANN are randomly selected as testing data without duplicating the training or validation data. There are two intuitive methods to judge the performance of ANN models. One is to plot the error of predictions. The other is to fit a linear regression analysis between the predictions and the estimates. Figure 6.5 to

6.24 illustrate the linear regression fit and the prediction errors for through and left-turning movements of all five intersection types. The linear fit, taking Figure 6.5 as an example, is between the ANN delay outputs and TRANSYT-7F simulated delay estimates (targets). A perfect fit will results in all predictions falling on the diagonal line and the R-squared value will be 1.0. For this example, the R-squared value is 0.835, which represent the best fit between the ANN outputs and the simulated delays of all ANN delay models. Figure 6.6 exhibits the ANN delay outputs and the simulated delays in different color so that the accuracy of the ANN predictions, delays in this case, may be easily identified. In the figure captions, a code is used to indicate the type of an intersection, as well as the movement studied, as follows:

2322TR: Through traffic at an intersection of a 3-lane FT 2 main road crossing a 3-lane FT 2 road
2322LT: Left-turning traffic at an intersection of a 3-lane FT 2 road crossing a 2-lane FT 2 road
2222TR: Through traffic at an intersection of two 2-lane FT 2 roads
2222LT: Left-turning traffic at an intersection of two 2-lane FT 2 roads
2241TR: Through traffic at an intersection of a 2-lane FT 2 road crossing a 1-lane FT 4 road
2241LT: Left-turning traffic at an intersection of a 2-lane FT 2 road crossing a 1-lane FT 4 road
3141TR: Through traffic at an intersection of a 1-lane FT 3 road crossing a 1-lane FT 4 road
3141LT: Left-turning traffic at an intersection of a 1-lane FT 3 road crossing a 1-lane FT 4 road
4141TR: Through traffic at an intersection of two 1-lane FT 4 roads 4141LT: Left-turning traffic at an intersection of two 1-lane FT 4 roads

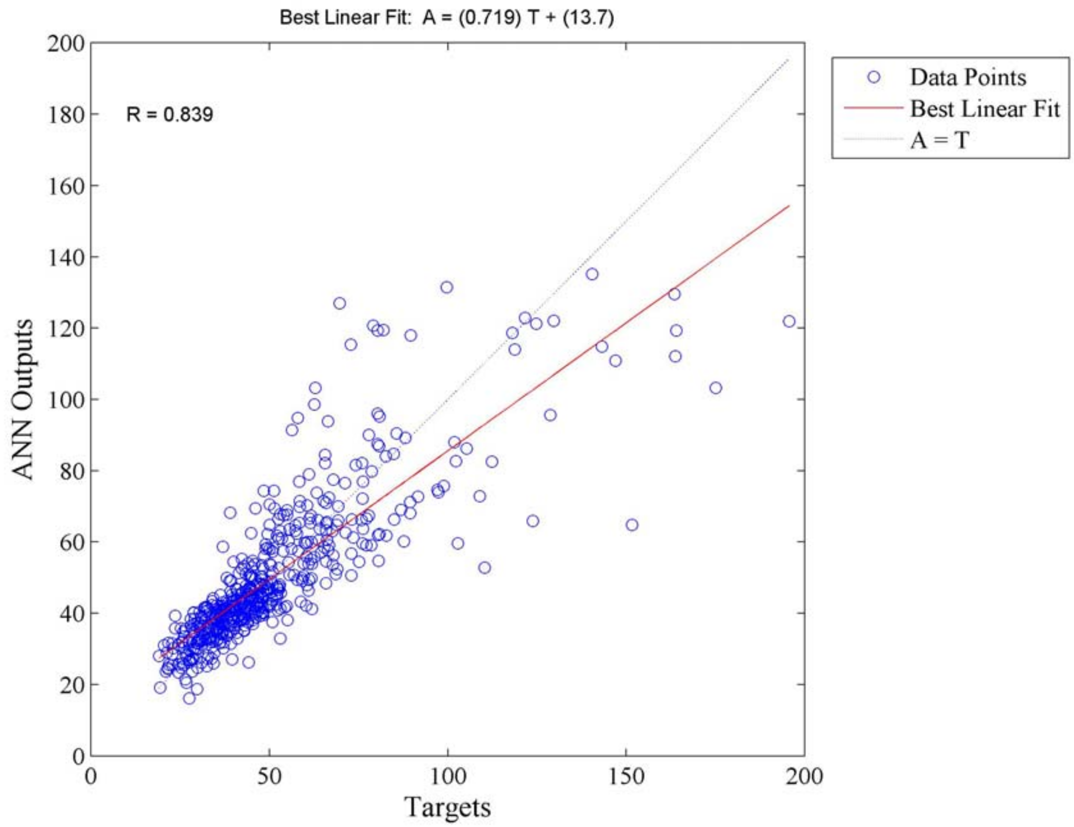

Figure 6.5 Linear Fit of ANN Delay Estimates and Targets (2322TR)

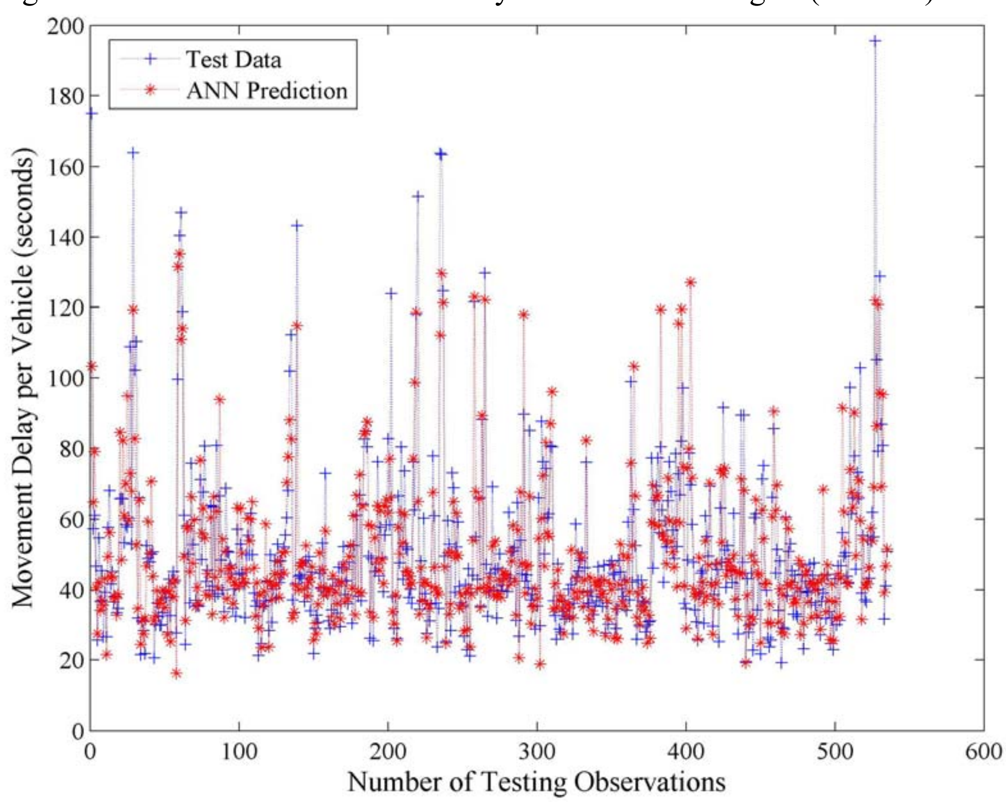

Figure 6.6 Fitting of ANN Delay Estimates to Targets (2322TR)

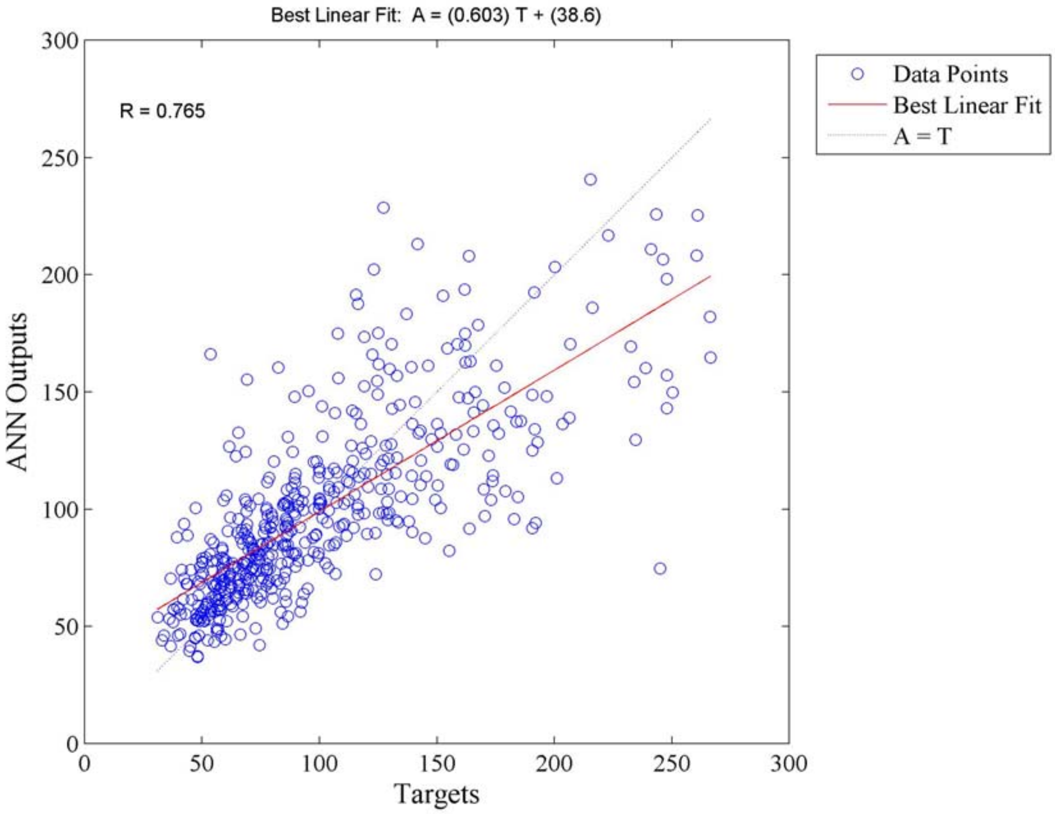

Figure 6.7 Linear Fit of ANN Delay Estimates and Targets (2322LT)

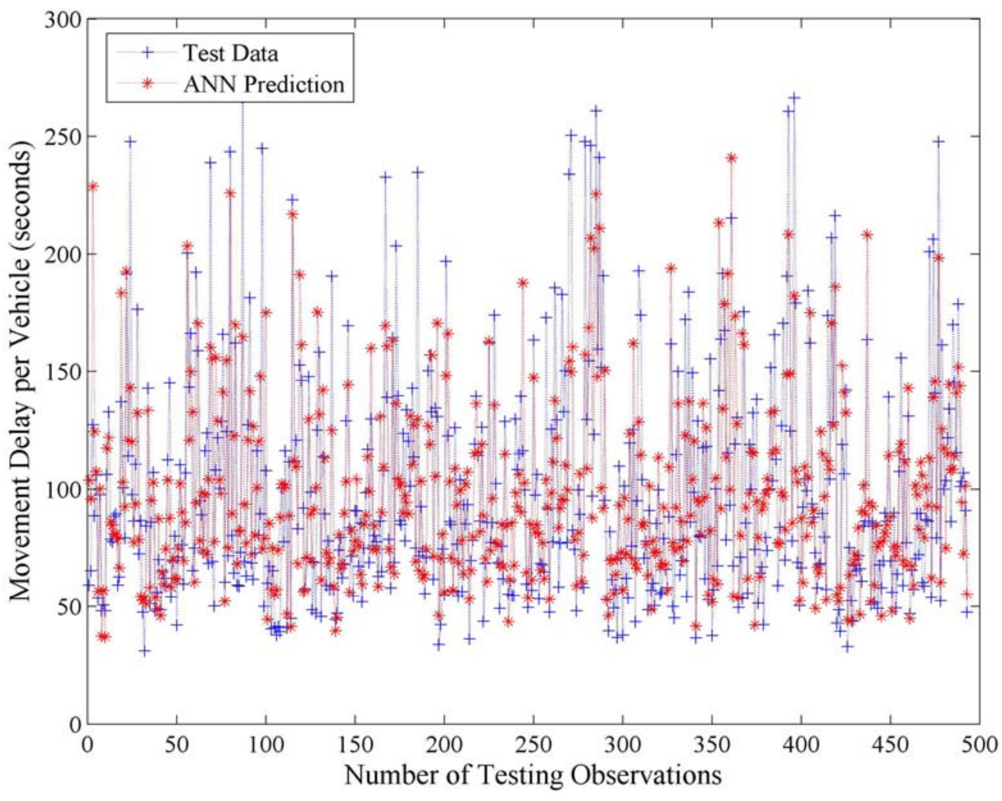

Figure 6.8 Fitting of ANN Delay Estimates to Targets (2322LT)

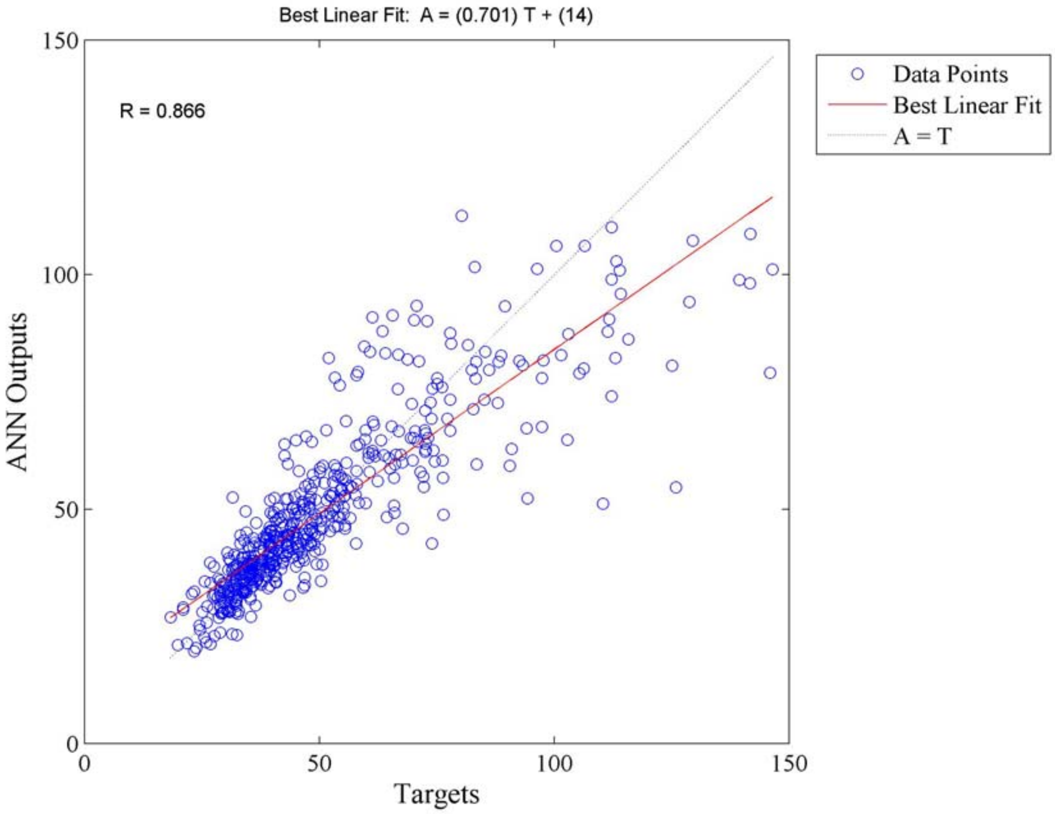

Figure 6.9 Linear Fit of ANN Delay Estimates and Targets (2222TR)

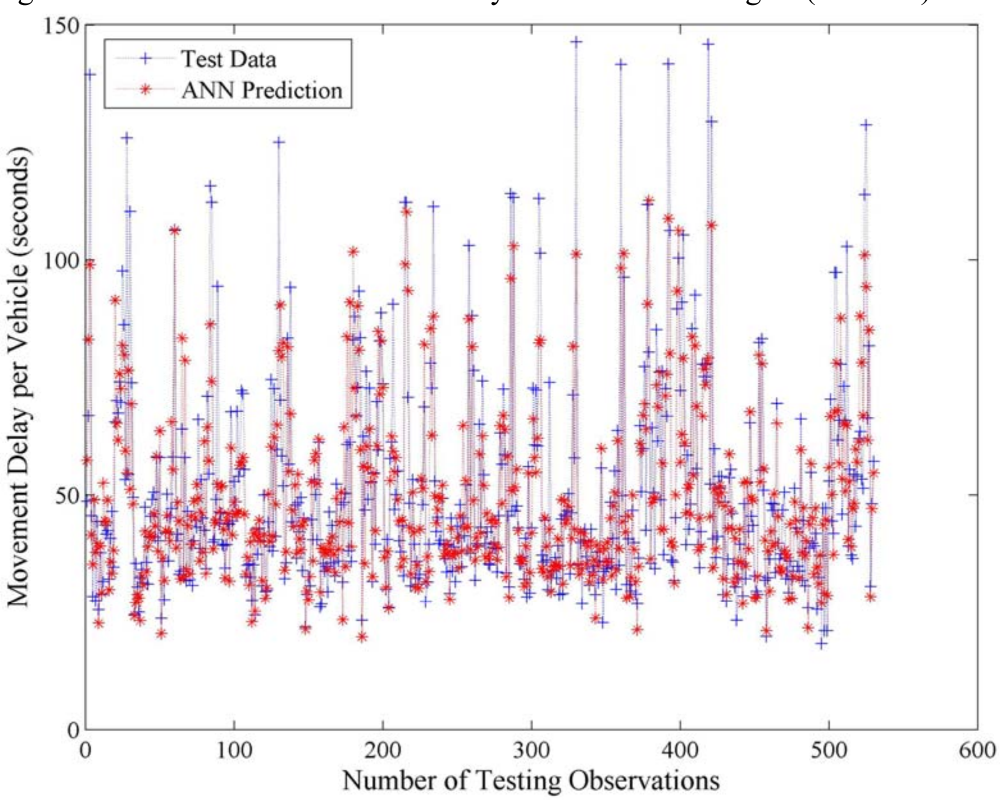

Figure 6.10 Fitting of ANN Delay Estimates to Targets (2222TR)

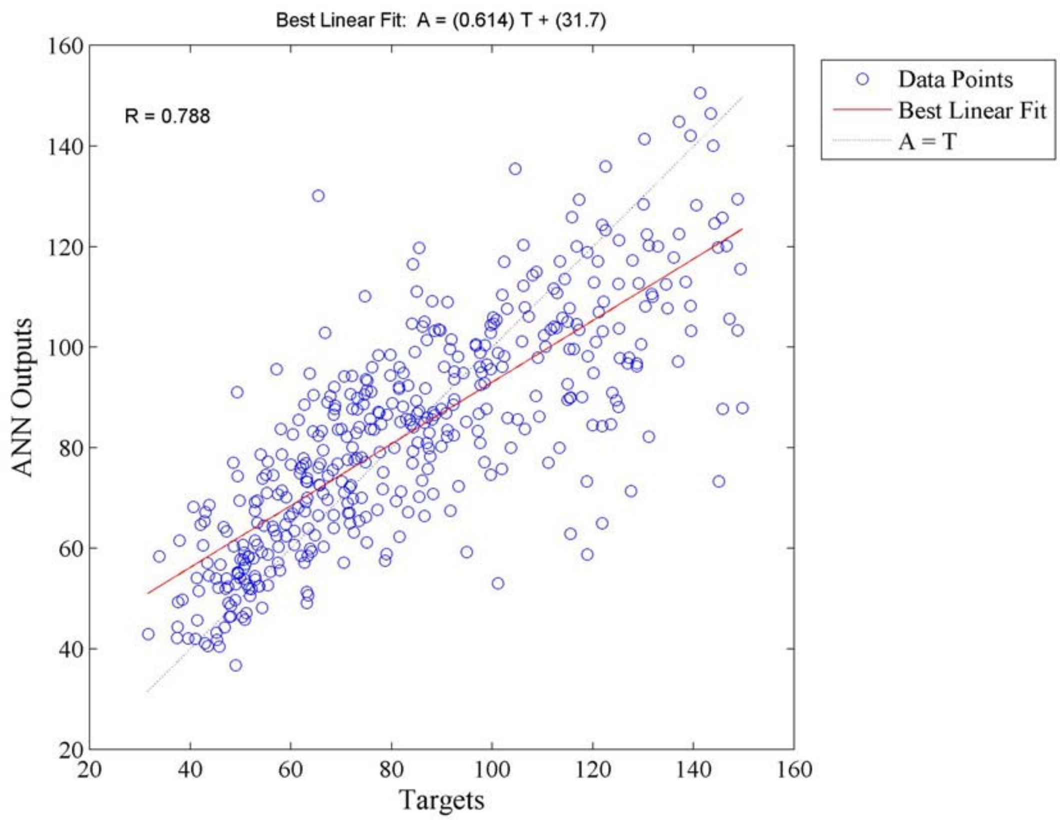

Figure 6.11 Linear Fit of ANN Delay Estimates and Targets (2222LT)

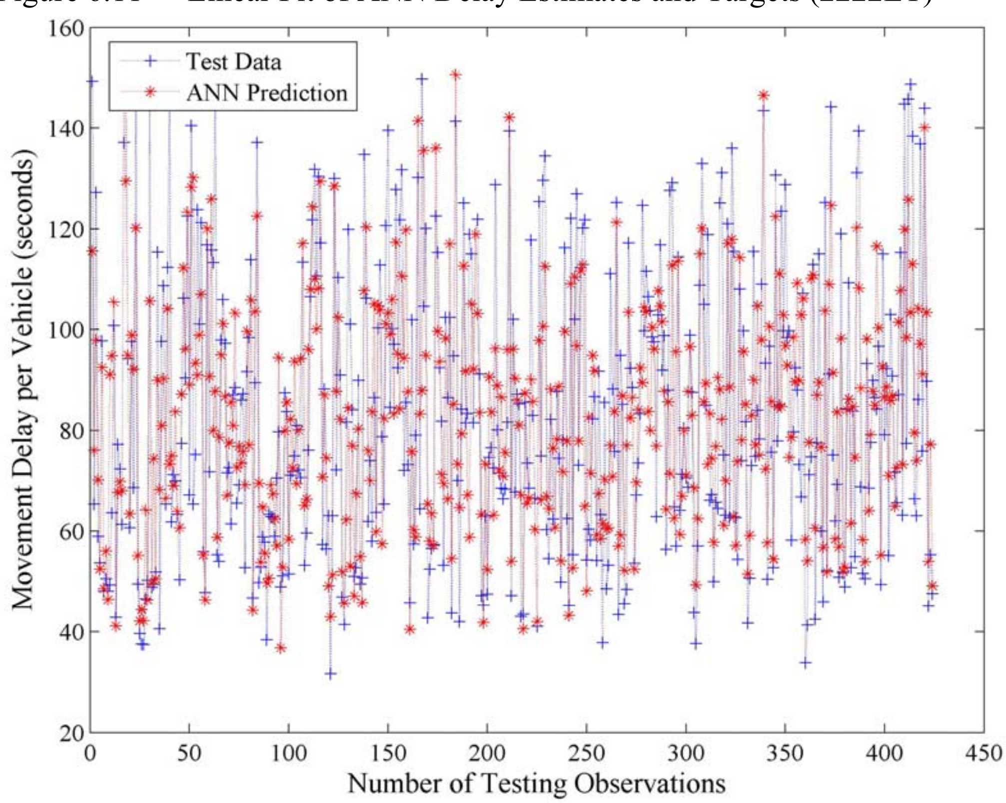

Figure 6.12 Fitting of ANN Delay Estimates to Targets (2222LT)

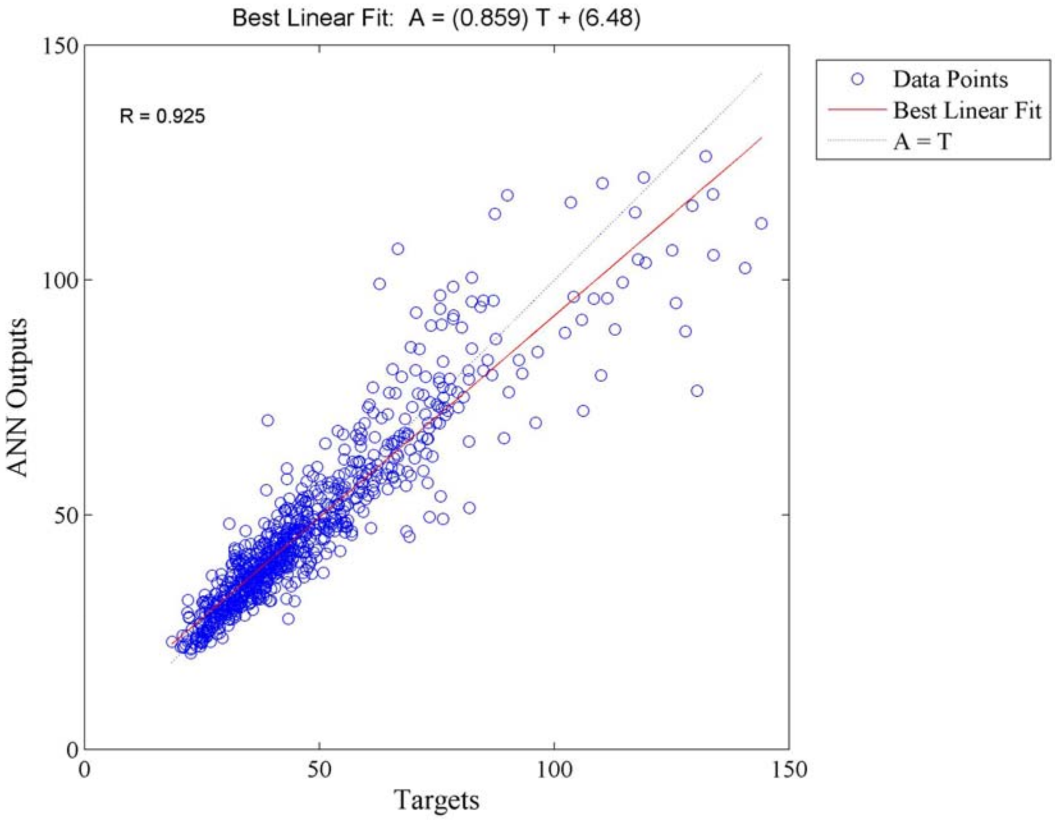

Figure 6.13 Linear Fit of ANN Delay Estimates and Targets (2241TR)

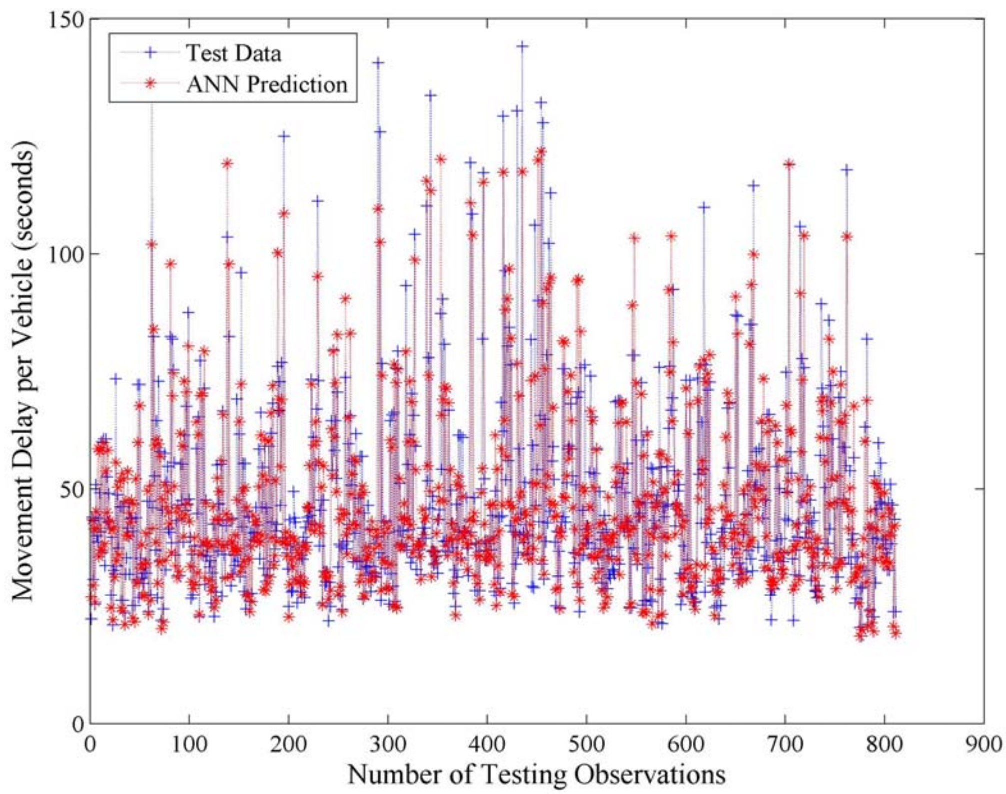

Figure 6.14 Fitting of ANN Delay Estimates to Targets (2241TR)

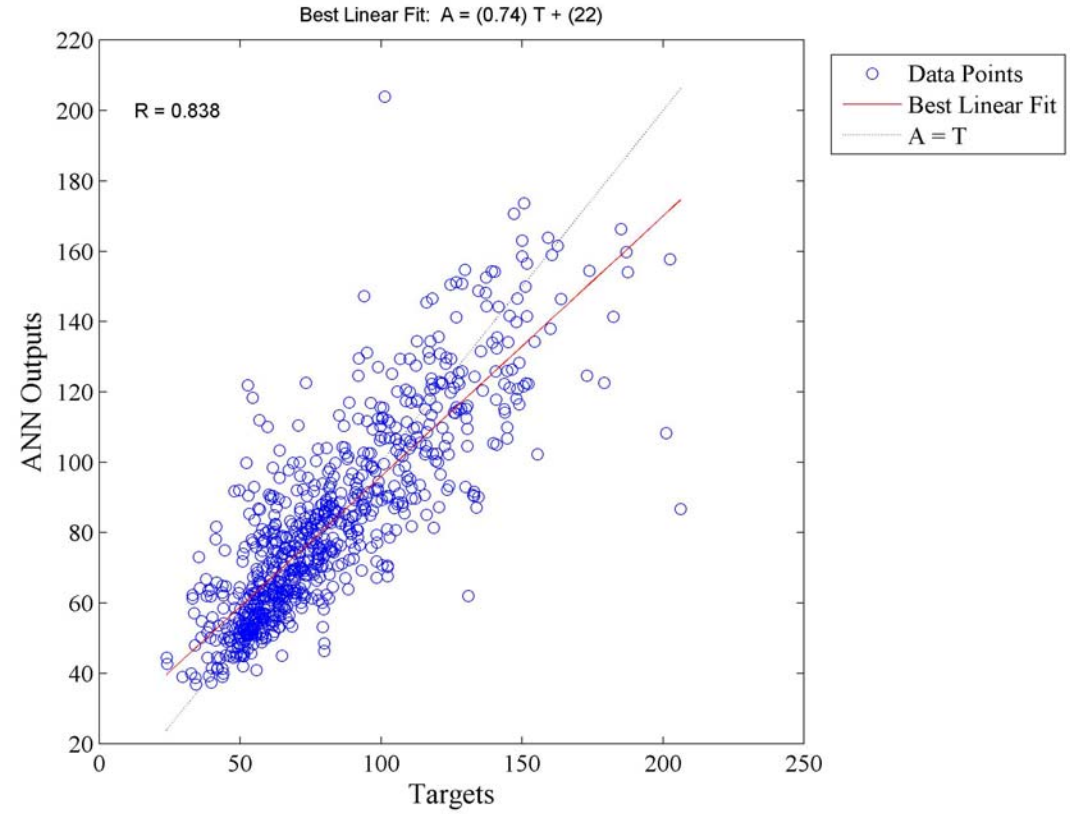

Figure 6.15 Linear Fit of ANN Delay Estimates and Targets (2241LT)

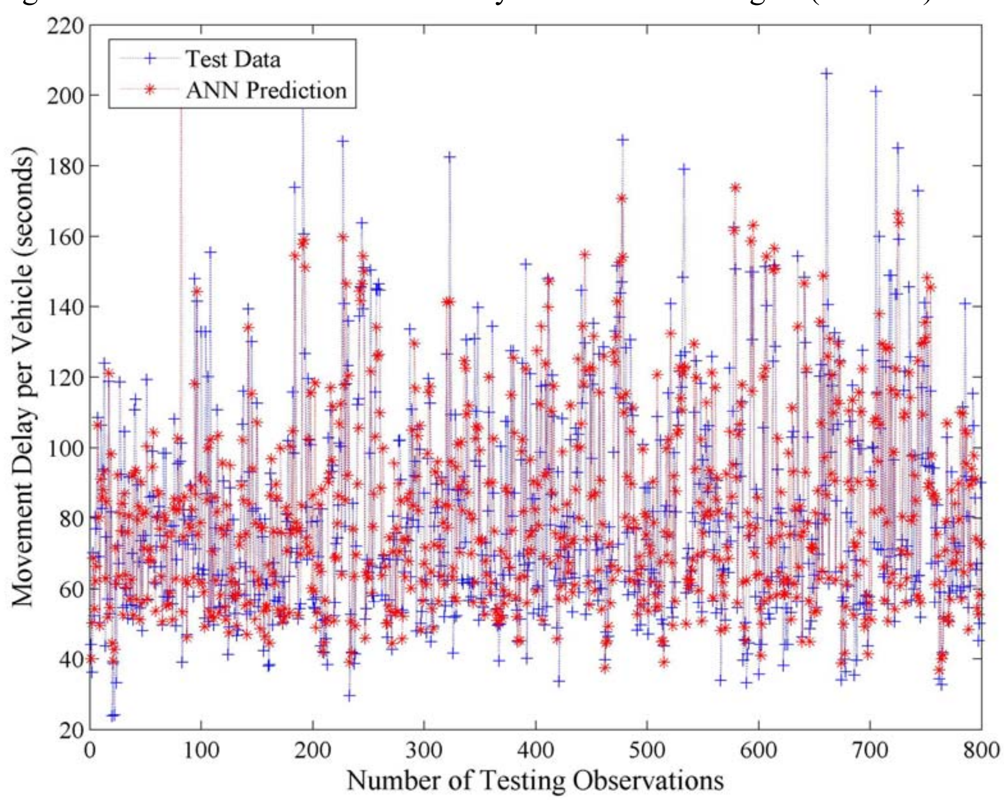

Figure 6.16 Fitting of ANN Delay Estimates to Targets (2241LT)

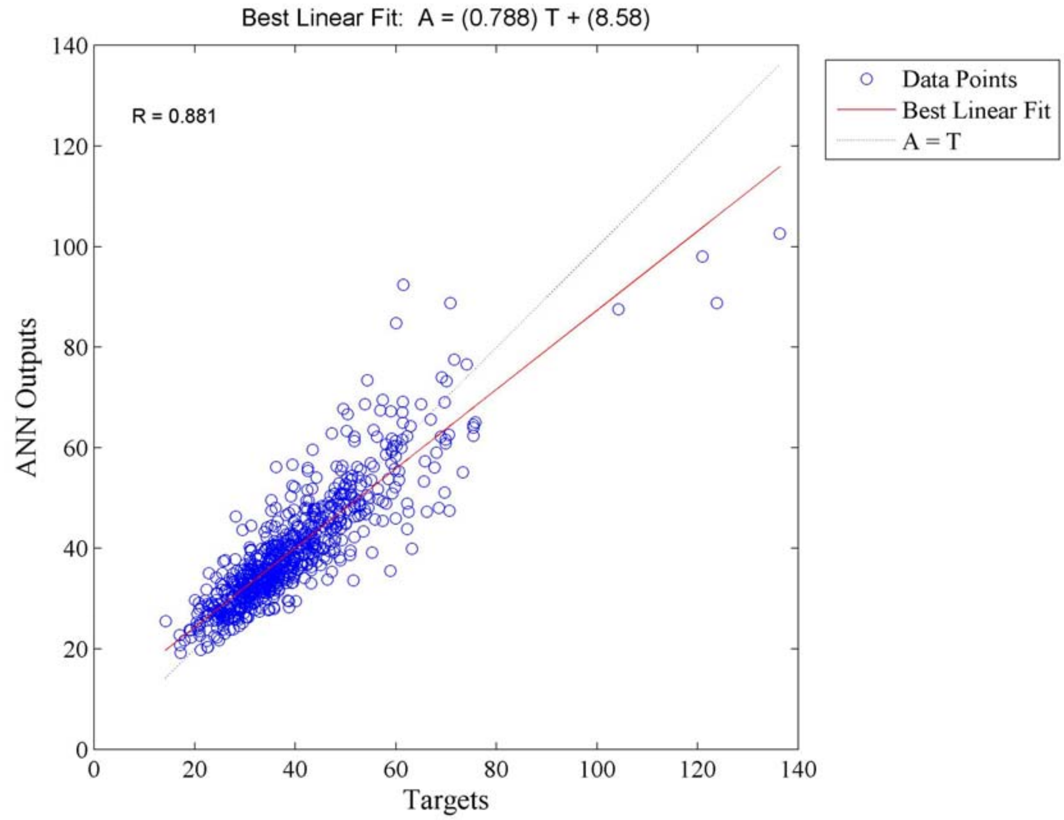

Figure 6.17 Linear Fit of ANN Delay Estimates and Targets (3141TR)

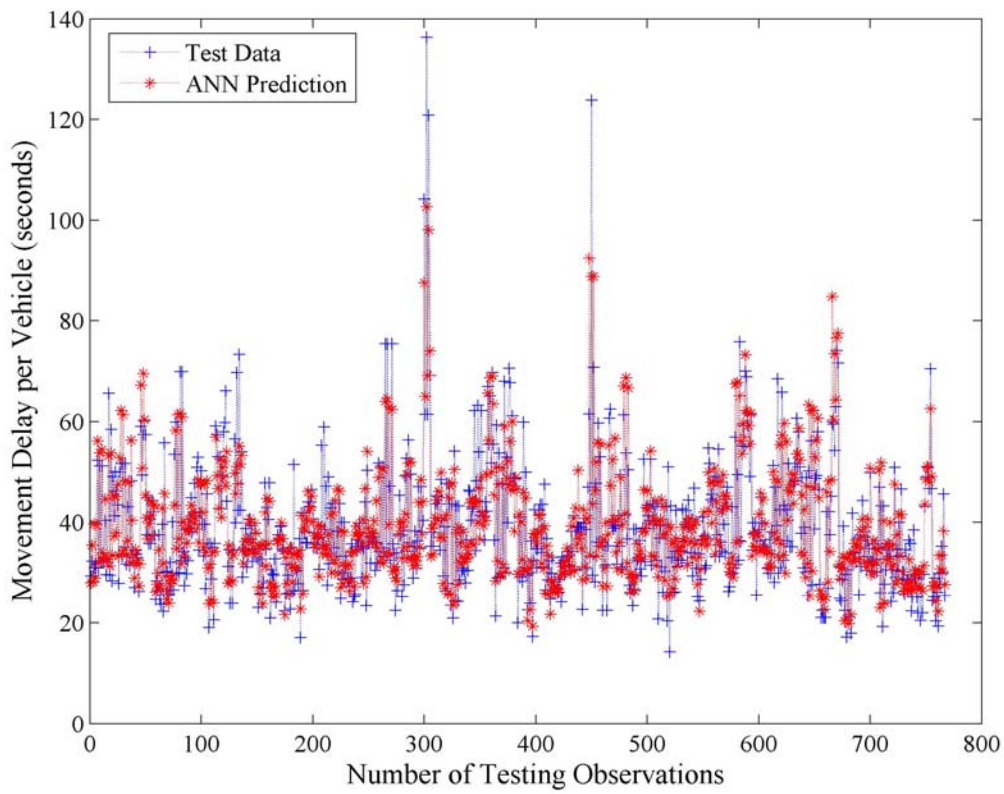

Figure 6.18 Fitting of ANN Delay Estimates to Targets (3141TR)

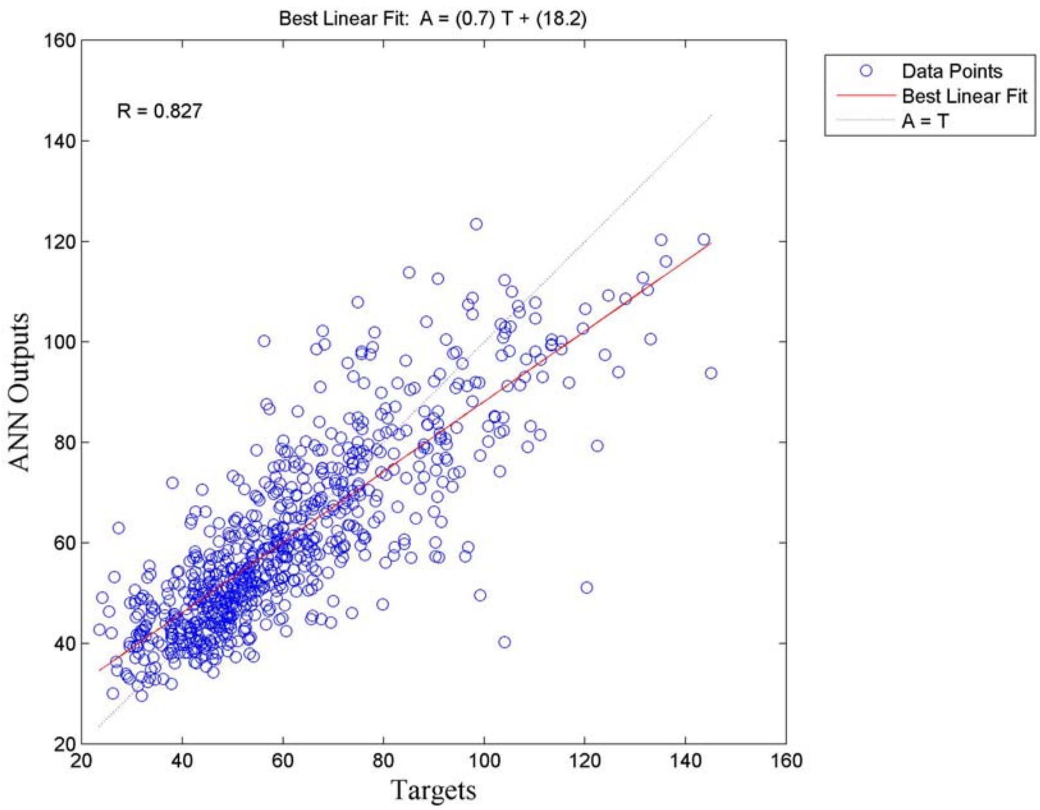

Figure 6.19 Linear Fit of ANN Delay Estimates and Targets (3141LT)

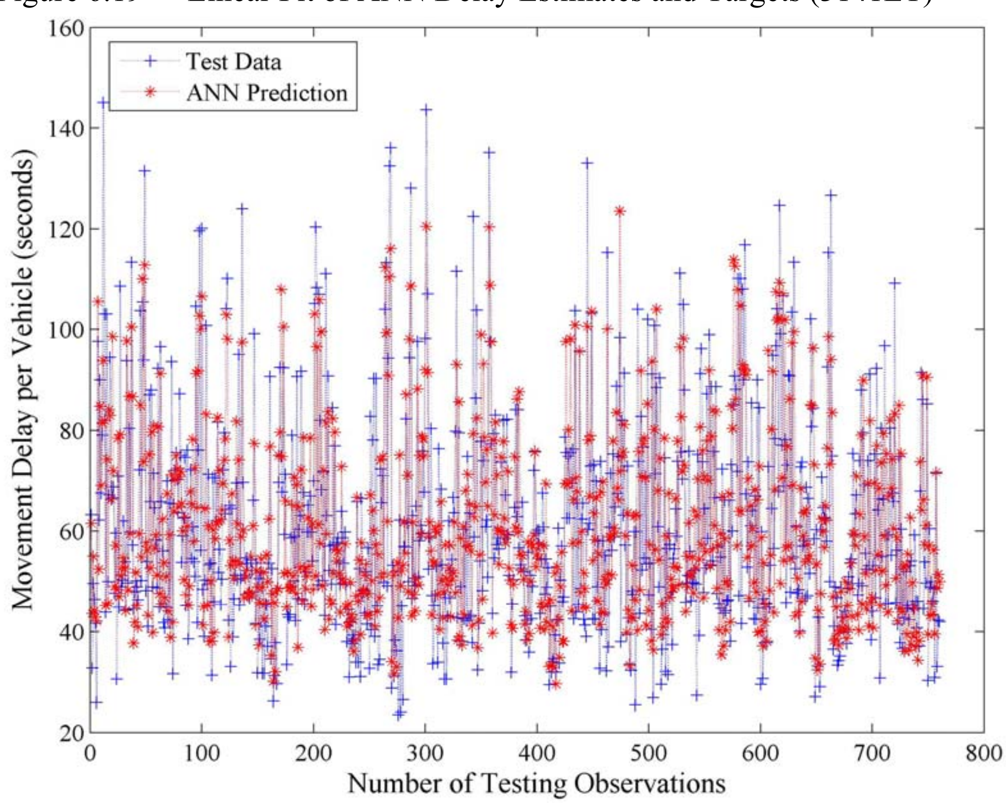

Figure 6.20 Fitting of ANN Delay Estimates to Targets (3141LT)

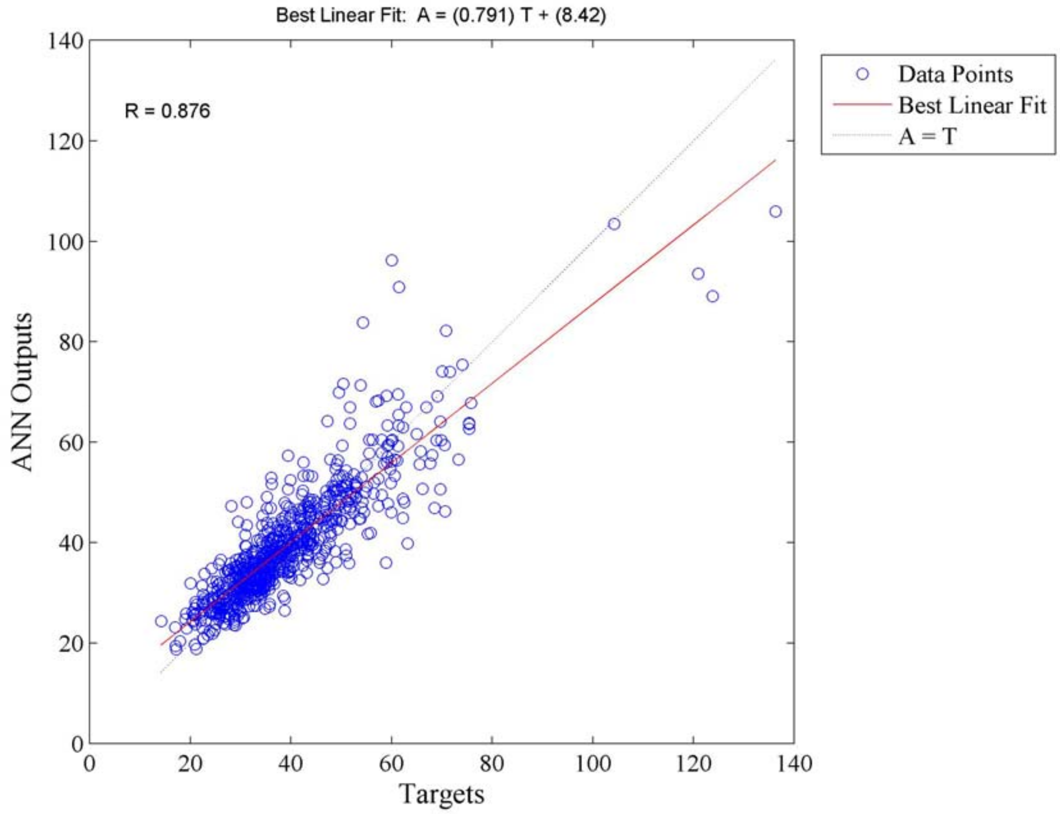

Figure 6.21 Linear Fit of ANN Delay Estimates and Targets (4141TR)

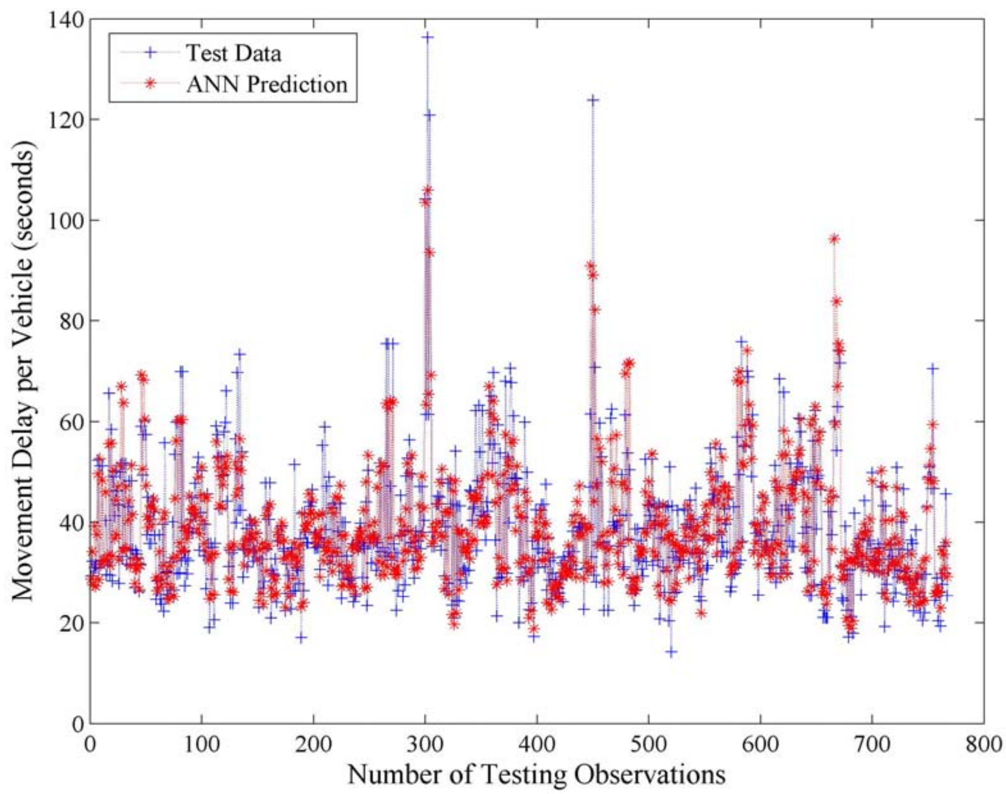

Figure 6.22 Fitting of ANN Delay Estimates to Targets (4141TR)

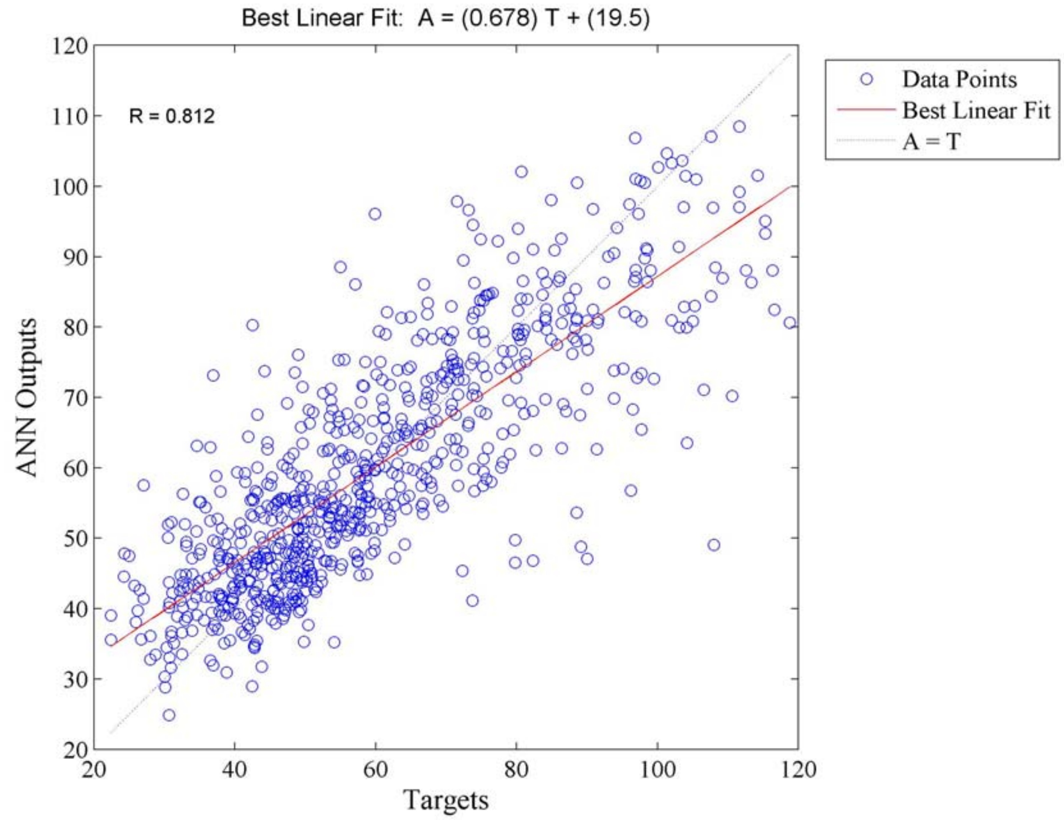

Figure 6.23 Linear Fit of ANN Delay Estimates and Targets (4141LT)

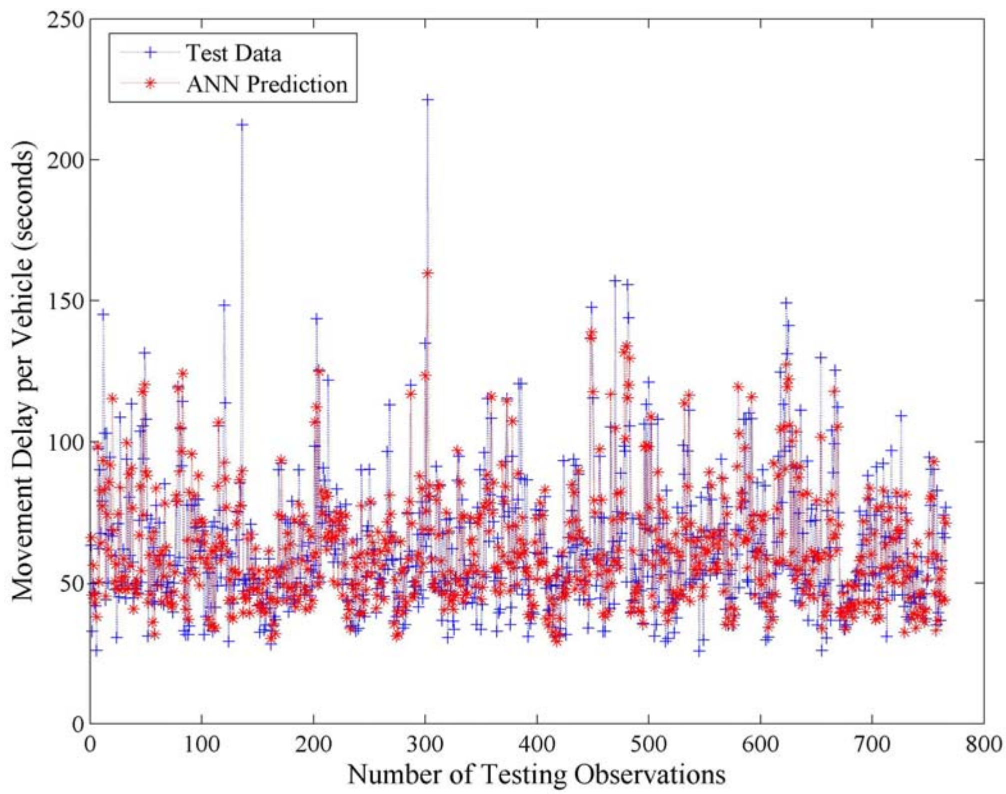

Figure 6.24 Fitting of ANN Delay Estimates to Targets (4141LT)

From these figures, some characteristics of the ANN delay models may be observed. The delay estimates for the larger intersections (facility type 2) are less accurate than those for the smaller intersections. Regardless of intersection types, the ANN models for left-turning delays always perform less well than those that predict through movement delays. That is, left-turning delays are more difficult to estimate than through movement delays.

To evaluate the performance of the ANN delay models, descriptive statistics, such as regression R-squared, RMSE, and %RMSE, are calculated, which are given in Table 6.3. Generally speaking, the ANN delay models demonstrate acceptable level of performance. The %RMSE of delay estimates are controlled at less than 25.6%.

The ANN models are also compared to multiple linear regression models. The regression models take the form as shown in Eq. 45. They are estimated using the same training date sets that are used to train the ANN delay models. The test data for the ANN models are used to evaluate the regression models. The Regression R-squared, RMSE, and %RMSE are compared between the ANN delay models, and the regression models are given in Table 6.3, which shows that the ANN delay models are superior to the regression models. Appendix B provides further details for these regression models.

$$D_c = b_0 + b_1.v_{11} + b_2.v_{12} + b_3.v_{21} + b_4.v_{22} + b_5.\ v_{31} + b_6.v_{32} + b_7.v_{41} + b_8.v_{42} \qquad \text{Eq. 45}$$

where

$D_c =$ movement delay of the studied lane group (seconds per vehicle)
$b_{0\sim 8} =$ regression coefficients
$v_{11} =$ through volume of this approach (vph)
$v_{12} =$ left-turning volume of this approach (vph)
$v_{21} =$ through volume of the opposing link (vph)
$v_{22} =$ left-turning volume of the opposing link (vph)
$v_{31} =$ through volume of the right crossing link (vph)
$v_{32} =$ left-turning volume of the right crossing link (vph)
$v_{41} =$ through volume of the left crossing link (vph)
$v_{42} =$ left-turning volume of the left crossing link (vph)

Further analysis is conducted on the characteristics of the prediction errors of the ANN delay models. Figure 6.25 shows the distributions of the mean absolute errors (MAE). The MAE (Eq. 46) is calculated to examine the size of forecast errors. MAE assumes that the severity of a prediction error increases in a linear manner (e.g. a 2% error is twice as serious as a 1% error). It may be observed that the MAE approximately follows a rising trend, which implies that the larger MAE are encountered more frequently when the intersection is seriously oversaturated. An exception is the intersection type 4141, which has a large MAE error also under light traffic conditions.

$$\text{MAE} = \frac{\sum_{i=1}^{n} \left| x_i^c - x_i^v \right|}{n} \qquad \text{Eq. 46}$$

where

x^c = delay estimates from TRANSYT-7F simulation (seconds per vehicle)

x^v = delay predictions by ANN delay models (seconds per vehicle)

n = number of testing inputs

Table 6.3 Comparison of Performance Statistics of Two Categories of Models

Intersection Category	**Statistics**	**ANN Models**	**Regression Models**	**ANN Models**	**Regression Models**
		Through	**Through**	**Left-turning**	**Left-turning**
2322	R-Squared	0.715	0.621	0.582	0.327
	RMSE	14.36	15.58	24.78	22.98
	%RMSE	22.95	24.95	25.58	37.94
2222	R-Squared	0.751	0.556	0.629	0.451
	RMSE	12.39	18.64	22.58	70.18
	%RMSE	20.73	35.29	24.97	57.91
2241	R-Squared	0.855	0.555	0.701	0.403
	RMSE	8.17	15.41	17.67	33.36
	%RMSE	18.51	31.91	21.73	39.28
3141	R-Squared	0.771	0.614	0.688	0.3916
	RMSE	6.14	7.72	13.75	19.66
	%RMSE	14.84	21.05	22.75	31.14
4141	R-Squared	0.767	0.610	0.656	0.390
	RMSE	5.67	7.89	13.46	18.58
	%RMSE	14.87	20.14	22.06	29.03

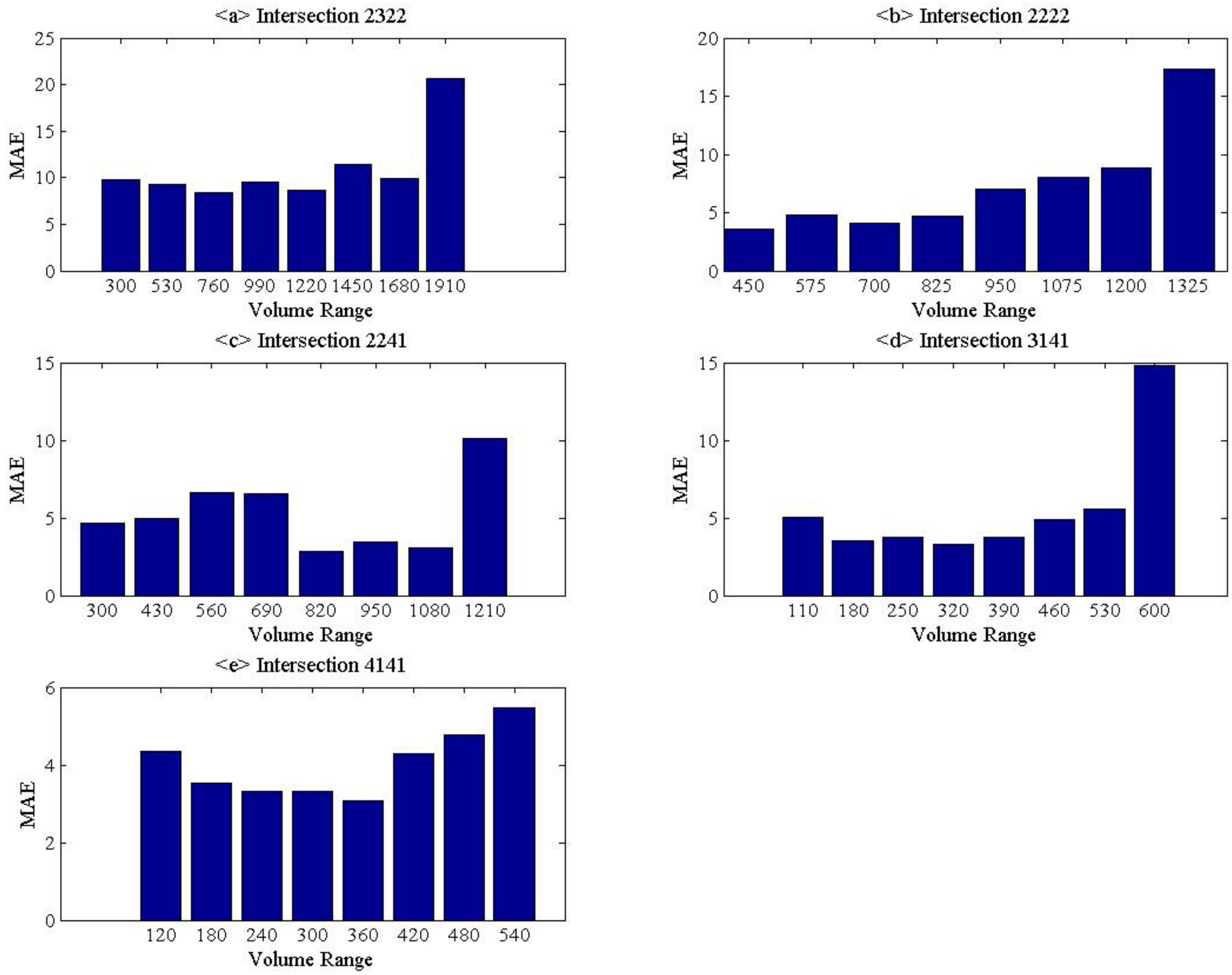

Figure 6.25 Distribution of Sum of Absolute Errors in Different Volume Ranges

7. CONCLUSIONS

This study has developed ANN delay models for five common intersection types based on the network of the 2000 Gainesville FSUTMS model. All the intersections are assumed to have a typical configuration with four-legs, 12-foot lanes, no exclusive right-turn lanes, and exclusive left-turning lanes. The ANN delay models are trained with inputs of volumes and capacities and targets of simulation results from TRANSYT-7F to predict movement delays for given input of volumes. According to the literature, optimization of signal timings in the traffic assignment process to calculate intersection delays often involves bi-level optimization. The advantage of the ANN models is that the delay estimates adapt to varied volumes, thus the signal optimizing procedure is unnecessary. As a result, traffic assignment will not require complicate algorithms to deal with bi-level optimization.

There are three main findings from this study. First, the ANN delay models are able to estimate intersection delays more efficiently than a regular complicated micro-simulation. Second, the ANN delay models are able to attain accuracy adequate for long-range planning purposes. The errors of the ANN delay models, measured as %RMSE, range from 14.8% to 25.6%. Analyses of the distributions of the prediction errors indicate that the ANN delay models estimate delays with less accuracy when traffic is severely oversaturated. This may be partly due to that TRANSYT-7F simulations are unable to deal with oversaturated conditions adequately with limited parameter settings, which may result in inaccurate portrayal of the relationships between traffic volumes and delays. As a result, the ANN models may have been unable to learn accurately these underlying relationships under seriously oversaturated conditions due to the noises in the input. Third and finally, the ANN delay models appear to produce more accurate delay estimates for smaller intersections (e.g., facility types 3 and 4) than for larger ones. This may be due to the fact that smaller intersections of local collectors have less complicated traffic conditions than larger intersections of arterials.

Recommendations regarding future research are provided below.

(1) To further test the proposed methodology and to determine its applicability to travel demand forecast, additional models may be developed in a similar manner for other types of intersections in Gainesville. The models may be applied during traffic assignment to measure the benefits of incorporating intersection delay models in terms of accuracy in link volumes and link speed.

(2) Improvement of model performance under oversaturation conditions may be achieved by employing another approach in TRANSYT-7F. TRANSYT-7F provides a different objective function that focuses on minimizing queues, because it is always "possible to measure queue lengths and vehicle trips relatively accurately" (McTrans, 2006). This new objective function may significantly help improvement results of single-period optimization. Another alternative is utilizing the multi-period optimization. Multi-period analysis requires significantly more computing resources. However, while computing resources may be a concern during model development, once the models are developed, they do not incur extra costs during model applications.

(3) Signal coordination is widely used in urban areas on arterials where traffic volumes are high. For instance, signal coordination is deployed around the University of Florida and downtown area in Gainesville. To consider signal coordination will require corridor-wide signal plan optimization, of which TRANSYT-7F is capable. Research is needed to develop generic models that represent a variety of configurations of intersections within a corridor, which may be applied at various locations and urban areas.

(4) In this study, isolated intersections are modeled in isolation. However, downstream congestion may also affect the delays at an upstream intersection. One approach to addressing this issue may be to incorporate additional input into the models that describe the nearby intersections (e.g., distance from the intersection being considered and their volumes). Another may be to perform small-area signal optimization and predict delays for the intersections in the area.

(5) Street parking and pedestrians may be common in central business districts. Since street parking and pedestrians essentially affect the speed on a segment, in future studies, these two issues may be addressed by assuming a speed lower than the posted speed. Issues related to street parking, pedestrians may also be considered through introducing area type as an input. Microscopic simulation of intersection operations may be necessary for measuring delays caused by street parking and pedestrian traffic volumes. However, simpler methods may also be developed by establishing default values based on area types. New area types may also be introduced for this purpose.

(6) In this research, symmetric intersections are investigated. Additional research is needed to determine the effect of asymmetric intersections where link properties change for through traffic.

(7) Finally, with more volume-delay data, the possibility of simplifying the models needs to be investigated. One issue is related to the number of models required. Generalization of models, e.g., through combining models, will reduce the number of models. Research may help to determine the relationship between accuracy and level of generalization. Another opportunity is to look for analytical forms of the volume-delay relationships. Although the accuracy of such simplified models may be reduced, the benefit may outweigh the drawback due to the ease of application and general applicability.

REFERENCES

Allsop, R. E. “Some Possibilities for Using Traffic Control to Influence Trip Distribution and Route Choice.” *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*. Elsevier, Amsterdam. 1996.

Aashtiani, H. Z., and Iravani, H. “Use of Intersection Delay Functions to Improve Reliability of Traffic Assignment Model.” Presented at the 14th Annual International EMME/2 Conference, Chicago, Illinois, October 1999.

Ahmed, K., and Abu-Lebdeh, G. “Modeling of Delay Induced by Downstream Traffic Disturbances at Signalized Intersections.” Presented at the 84th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C., 2005.

Alexiadis, V., Jeannotte, K., and Chandra, A. *Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer*, FHWA-HRT-04-038 Report, Turner-Fairbank Highway Research Center, McLean, VA, July 2004.

Benekohal, R. F., and Kim, S. “Arrival Based Uniform Delay Model for Oversaturated Signalized Intersections with Poor Progression.” Presented at the 84th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C., 2005.

Cantarella, G. E., and Sforza, A. “Methods for Equilibrium Network Traffic Signal Setting.” in *Flow Control of Congested Networks* (Odoni, A. R., Bianco, L. and Szego, G., eds.), Springer-Verlag, 1987, pp. 69-89.

Ceylan, H., and Bell, M. G. “Traffic Signal Timing Optimisation Based on Genetic Algorithm Approach, including Drivers' Routing.” *Transportation Research Part B.* Vol. 38, No. 4, 2004, pp. 329–342.

Demuth, H., Beale, M., and Hagan, M. (2006). *Neural Network Toolbox User’s Guide.* The MathWorks, Inc., Natick, Massachusetts.

Dion, F., Rakha, H., and Kang, Y. “Comparison of Delay Estimates at Under-saturated and Oversaturated Pre-timed Signalized Intersections”, *Transportation Research Part B.* Vol. 38, No. 2, 2004, pp. 99–122.

Dunn, P., and Johnson, B. “Assignment Techniques for Networks with Junction Modeling.” Presented at the 14th Annual International EMME/2 Conference, Chicago, Illinois, October 1999.

Florida Department of Transportation, *Documentation and Procedural Updates to the Florida Standard Urban Transportation Model Structure*, Tallahassee, Florida, 1997.

Florida Department of Transportation, *Quality/Level of Service Handbook*, Tallahassee, Florida, 2002.

Gartner, N., and Al-Malik, M. "Combined Model for Signal Control and Route Choice in Urban Traffic Networks", *Transportation Research Record 1554*, Transportation Research Board, National Research Council, Washington, D.C., 1996, pp. 27-35.

Hall, M. D., Van Vliet, D., and Willumsen, L. G. "SATURN – A Simulation /Assignment Model for the Evaluation of Traffic Management Schemes." *Traffic Engineering and Control,* Vol. 21, No. 4, 1980, pp. 168-176.

Heidemann, D. "Queue Length and Delay Distributions at Traffic Signals", *Transportation Research Part B.* Vol. 28B, No. 4, 1994, pp. 377–389.

Hill, R O'N. "An Application of EMME/2 Auto-assignment with Detailed Modeling of Activity at Nodes", *Proceedings of the 13th International EMME/2 Users Group Conference,* Houston, Texas, October 1998.

Hurdle, V. "Signalized Intersection Delay Models: A Primer for the Uninitiated." *Transportation Research Record 971*, Transportation Research Board, National Research Council, Washington, D.C., 1984, pp. 96-105.

Jeannotte, K., Chandra, A., Alexiadis, V., and Skabardonis, A. *Traffic Analysis Toolbox Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools, Final Report.* FHWA-HRT-04-039, Turner-Fairbank Highway Research Center, McLean, VA, July 2004.

Kimber, R. M., and Hollis, E. M. *Traffic Queues and Delays at Road Junctions*. Laboratory Report 909, Transport and Road Research Laboratory, Crowthorne, UK, 1979.

Lee, C., and Machemehl, R. B. "Local and Iterative Searches for the Combined Signal Control and Assignment Problem: Implementation and Numerical Examples." *Transportation Research Record No. 1683*, Transportation Research Board, Washington, D.C., 1999, pp. 102-109.

Levinson, D., and Kumar, A. "Integrating Feedback into the Transportation Planning Model: Structure and Application", *Transportation Research Record 1413*, Transportation Research Board, Washington, D.C., 1994, pp. 78-86.

Lighthill, M. J., and Whitham, G. B. "On Kinematic Waves: II. A Theory of Traffic Flow on Long Crowded Roads", *Proceedings of the Royal Society: A229*, London, 1955, pp. 317-347.

McNeil, D.R. "A Solution to the Fixed-cycle Traffic Light Problem for Compound Poisson Arrivals." *Journal of Applied Probability,* Vol. 5, No. 5, 1968, pp. 624–635.

McTrans. *Traffic Network Study Tool-TRANSYT-7F.* McTrans Center, University of Florida, Gainesville, Florida, 2006.

Newell, G. F. *Application of Queuing Theory,* 2nd Edition. Chapman and Hall, London England, 1982.

Robertson, D. I., and Gower, P. *User Guide to TRANSYT*, version 6. Supplementary Report LR 255, Transport and Road Research Laboratory, Crowthorne, England, 1977.

Roess, R. P., McShane, W. R., and Prassas, E. S. *Traffic Engineering*. Prentice Hall, Upper Saddle River, New Jersey, 1998.

SCAG (2003). *Year 2000 Model Validation and Summary: Regional Transportation Model.* Technical Report. Southern California Association of Governments, Los Angeles, California. *http://www.scag.ca.gov/modeling/2000mv.htm*

Sheffi, Y. *Urban Transportation Networks*. Prentice Hall, Englewood Cliffs, N.J., 1985.

Sheffi, Y. and Powell, W. "Optimal Signal Setting Subject to Equilibrium Constraints over Transportation Networks." American Society of Civil Engineers *Transportation Journal*, Vol. 109, No. 6, November 1983, pp. 824-839.

Troutbeck, R., and Blogg, M. "Queuing at Congested Intersections." Presented at the 77th Annual Meeting of the Transportation Research Board 1998, National Research Council, Washington, D.C., 1998.

Transportation Research Board, *Highway Capacity Manual*, Washington, D.C., 2000.

Webster, F. *Traffic Signal Settings*. Road Research Technical Paper No. 39, Road Research Laboratory, Her Majesty's Stationery Office, London, U.K., 1958.

Wong, S. C. "Group-based Optimization of Signal Timings Using the TRANSYT Traffic Model." *Transportation Research Part B*, Vol. 30, No. 3, 1996, pp. 217-244.

Wong, S. C., Yang, C., and Lo, H. K. "A Path-Based Traffic Assignment Algorithm Based on the Transyt Traffic Model." *Transportation Research Part B*, Vol. 35, No. 2, 2001, pp. 163-181.

Zhou, J., and Vaughan, B. "Junction Modeling in EMME/2." Presented at *the 14th Annual International EMME/2 Conference*, Chicago, Illinois, October 1999.

Ziliaskopoulos, A. K., and Mahmassani, H. S. "A Note on Least Time Path Computation Considering Delays and Prohibitions for Intersection Movements." *Transportation Research Part B*, Vol. 30, No. 5, 1996, pp. 359-367.

Zhao, F. and M.-T. Li, *Calibration of Highway/transit Speed Relationships for Improved Transit Network Modeling in FSUTMS.* Technical Report, prepared for the Florida Department of Transportation, Tallahassee, Florida, 2005.

APPENDIX A. AREA TYPE AND FACILITY TYPE DEFINTION

Table A.1 One-Digit Area Type Codes

Area Type	Description
1	Central Business District (CBD)
2	Fringe
3	Residential
4	Outlying Business District (OBD)
5	Rural

* FDOT FSUTMS Technical Reports (1997-1998)

Table A.2 One-Digit Facility Type Codes

Facility Type	Description
1(10)	Freeway
2(20)	Divided Arterial
3(30)	Undivided Arterial
4(40)	Collector
5(50)	Centroid Collector
6(60)	One-Way Streets
7(70)	Ramp
8(80)	HOV lane
9(90)	Tolls

* FDOT FSUTMS Technical Reports (1997-1998)

Table A.3 Two-Digit Area Type Codes

Area Type	Description
1x	**Central Business District (CBD) Areas (AT 10 is the default)**
11	Urbanized Area (over 500,000) Primary City CBD
12	Urbanized Area (under 500,000) Primary City CBD
13	Other Urbanized Area CBD and Small City Downtown
14	Non-Urbanized Area Small City Downtown
2x	**Central Business District (CBD) Fringe Areas (AT 20 is the default)**
21	All CBD Fringe Areas
3x	**Residential Areas (AT 30 is the default)**
31	Residential Area of Urbanized Areas
32	Undeveloped Portions of Urbanized Areas
33	Transitioning Areas/ Urban Areas over 5,000 Population
34	Beach Residential (per Southeast Regional Planning Model - SERPM)
4x	**Outlying Business District (OBD) Areas (AT 40 is the default)**
41	High Density OBD
42	Other OBD
43	Beach OBD (per Southeast Regional Planning Model - SEPRM)
5x	**Rural Area (AT 50 is the default)**
51	Developed Rural Areas/ Small Cities Under 5,000 Population
52	Undeveloped Rural Areas

* FDOT FSUTMS Highway Network (HNET) Procedural Enhancements Study: Final User's Manual (March 1998).

Table A.4 Two-Digit Facility Type Codes

Facility Type	Description
1x	**Freeways and Expressways (FT 10 is the default)**
11	Urban Freeway Group 1 (cities of 500,000 or more)
12	Urban Freeway Group 2 (within urbanized area and not in Group 1)
15	Collector/Distributor Lane
16	Controlled Access Expressway
17	Controlled Access Parkway
2x	**Divided Arterials (FT 20 is the default)**
21	Divided Arterial Unsignalized (55 mph)
22	Divided Arterial Unsignalized (45 mph)
23	Divided Arterial Class 1a (> 0.00 to 2.49 signalized intersections per mile)
24	Divided Arterial Class 1b (2.50 to 4.50 signalized intersections per mile)
25	Divided Arterial Class II/III (> 4.50 signalized intersections per mile)
3x	**Undivided Arterials (FT 30 is the default)**
31	Undivided Arterial Unsignalized with Turn Bays
32	Undivided Arterial Class 1a (> 0.00 to 2.49 signalized intersections per mile) with Turn Bays
33	Undivided Arterial Class 1b (2.50 to 4.50 signalized intersections per mile) with Turn Bays
34	Undivided Arterial Class II/III (> 4.50 signalized intersections per mile) with Turn Bays
35	Undivided Arterial Unsignalized without Turn Bays
36	Undivided Arterial Class 1a (> 0.00 to 2.49 signalized intersections per mile) without Turn Bays
37	Undivided Arterial Class 1b (2.50 to 4.50 signalized intersections per mile) without Turn Bays
38	Undivided Arterial Class II/III (> 4.50 signalized intersections per mile) without Turn Bays
4x	**Collectors (FT 40 is the default)**
41	Major Local Divided Roadway
42	Major Local Undivided Roadway with Turn Bays
43	Major Local Undivided Roadway without Turn Bays
44	Other Local Divided Roadway
45	Other Local Undivided Roadway with Turn Bays
46	Other Local Undivided Roadway without Turn Bays
47	Low Speed Local Collector
48	Very Low Speed Local Collector
5x	**Centroid Connectors (FT 50 is the default)**
51	Basic Centroid Connector
52	External Station Centroid Connector
6x	**One-Way Facilities (FT 60 is the default)**
61	One-Way Facility Unsignalized
62	One-Way Facility Class Ia (> 0.00 to 2.49 signalized intersections per mile)

63	One-Way Facility Class Ib (2.50 to 4.50 signalized intersections per mile)
64	One-Way Facility Class II/III (> 4.50 signalized intersections per mile)
65	Frontage Road Unsignalized
66	Frontage Road Class Ia (> 0.00 to 2.49 signalized intersections per mile)
67	Frontage Road Class Ib (2.50 to 4.50 signalized intersections per mile)
68	Frontage Road Class II/III (> 4.50 signalized intersections per mile)
7x	**Ramps**
71	Freeway On-Ramp
72	Freeway Loop On-Ramp
73	Other On-Ramp
74	Other Loop On-Ramp
75	Freeway Off-Ramp
76	Freeway Loop Off-Ramp
77	Other Off-Ramp
78	Other Loop Off-Ramp
79	Freeway-Freeway High-Speed Ramp
8x	**HOV Facilities (FT 80 is the default)**
81	Urban Freeway Group 1 (cities of 500,000 or more) 1 HOV Lane (Barrier Separated)
82	Urban Freeway Group 2 (within urbanized area and not in Group 1) HOV Lane (Barrier Separated)
83	Freeway Group 1 HOV Lane (Non-Barrier Separated)
84	Other Freeway HOV Lane (Non-Barrier Separated)
85	Non Freeway HOV Lane
86	AM&PM Peak HOV Ramp
87	AM Peak Only HOV Ramp
88	PM Peak Only HOV Ramp
89	All Day HOV Ramp
9x	**Toll Facilities**
91	Urban Freeway Group 1 (cities of 500,000 or more) Toll Facility
92	Urban Freeway Group 2 (within urbanized area and not in Group 1) Toll Facility
93	Expressway/Parkway Toll Facility
94	Divided Arterial Toll Facility
95	Undivided Arterial Toll Facility
97	Toll On-Ramp
98	Toll Off-Ramp
99	Toll Plaza

* FDOT 1995 LOS Manual.

APPENDIX B. REGRESSION MODELS OF ALL INTERSECTIONS TYPES

In this appendix, the regression equations for the five intersections are given, along with the statistics on the models. The statistics are at 0.05% significance level.

The models take the form as given in Eq. 45:

$$D_c = b_0 + b_1.v_{11} + b_2.v_{12} + b_3.v_{21} + b_4.v_{22} + b_5.\ v_{31} + b_6.v_{32} + b_7.v_{41} + b_8.v_{42}$$

where

- D_c = movement delay of the studied lane group (seconds per vehicle)
- $b_{0\sim8}$ = regression coefficients
- v_{11} = through volume of this approach (vph)
- v_{12} = left-turning volume of this approach (vph)
- v_{21} = through volume of the opposing link (vph)
- v_{22} = left-turning volume of the opposing link (vph)
- v_{31} = through volume of the right crossing link (vph)
- v_{32} = left-turning volume of the right crossing link (vph)
- v_{41} = through volume of the left crossing link (vph)
- v_{42} = left-turning volume of the left crossing link (vph)

B.1 Regression Model of 2322TR

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	0.64530	0.36090	1.79	0.0738		0
v11	0.00669	0.00026163	25.57	<.0001	0.0534	1.64752
v12	-0.00686	0.00026163	-26.22	<.0001	0.0295	1.64752
v21	0.00362	0.00147	2.47	0.0134	0.0003	1.64447
v22	0.04842	0.00147	33.04	<.0001	0.0198	1.64447
v31	0.01412	0.00026163	53.98	<.0001	0.2225	1.64752
v32	0.01400	0.00026163	53.50	<.0001	0.2864	1.64752
v41	0.01287	0.00147	8.78	<.0001	0.0034	1.64447
v42	0.01875	0.00147	12.80	<.0001	0.0058	1.64447

R-Square 0.6212
Adj R-Sq 0.6208

B.2 Regression Model of 2322LT

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-22.36240	1.48382	-15.07	<.0001		0
v11	0.13834	0.00602	22.96	<.0001	0.0811	1.64447
v12	-0.10875	0.00602	-18.05	<.0001	0.0247	1.64447
v21	0.00598	0.00108	5.56	<.0001	0.0026	1.64752
v22	0.04335	0.00108	40.30	<.0001	0.0721	1.64752
v31	0.01660	0.00602	2.75	0.0059	0.0006	1.64447
v32	0.02514	0.00602	4.17	<.0001	0.0011	1.64447
v41	0.02284	0.00108	21.23	<.0001	0.0658	1.64752
v42	0.02341	0.00108	21.76	<.0001	0.0786	1.64752

R-Square	0.3267
Adj R-Sq	0.3261

B.3 Regression Model of 2222TR

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-70.04541	1.85363	-37.79	<.0001		0
v11	0.05225	0.00116	44.88	<.0001	0.2633	1.32543
v12	-0.01414	0.00116	-12.14	<.0001	0.0102	1.32543
v21	0.03978	0.00530	7.50	<.0001	0.0046	1.35423
v22	0.09253	0.00530	17.44	<.0001	0.0116	1.35423
v31	0.02977	0.00116	25.56	<.0001	0.1005	1.32543
v32	0.03277	0.00116	28.14	<.0001	0.1486	1.32543
v41	0.05181	0.00530	9.77	<.0001	0.0077	1.35423
v42	0.06901	0.00530	13.01	<.0001	0.0098	1.35423

R-Square	0.5564
Adj R-Sq	0.5558

B.4 Regression Model of 2222LT

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-250.84125	6.98078	-35.93	<.0001		0
v11	0.57981	0.01998	29.02	<.0001	0.1682	1.35423
v12	-0.37741	0.01998	-18.89	<.0001	0.0378	1.35423
v21	0.02330	0.00439	5.31	<.0001	0.0032	1.32543
v22	0.15643	0.00439	35.67	<.0001	0.0980	1.32543
v31	0.12226	0.01998	6.12	<.0001	0.0030	1.35423
v32	0.10153	0.01998	5.08	<.0001	0.0026	1.35423
v41	0.07916	0.00439	18.05	<.0001	0.0588	1.32543
v42	0.08845	0.00439	20.17	<.0001	0.0798	1.32543

R-Square	0.4513
Adj R-Sq	0.4505

B.5 Regression Model of 2241TR

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-32.52576	1.34112	-24.25	<.0001		0
v11	0.04598	0.00098032	46.91	<.0001	0.0700	2.61454
v12	-0.01758	0.00098032	-17.94	<.0001	0.0088	2.61454
v21	0.01884	0.00460	4.10	<.0001	0.0009	1.73312
v22	0.06267	0.00460	13.63	<.0001	0.0095	1.73312
v31	0.03477	0.00098032	35.47	<.0001	0.2878	2.61454
v32	0.03224	0.00098032	32.89	<.0001	0.1554	2.61454
v41	0.06829	0.00460	14.85	<.0001	0.0105	1.73312
v42	0.07741	0.00460	16.83	<.0001	0.0119	1.73312

R-Square	0.5548
Adj R-Sq	0.5544

B.6 Regression Model of 2241LT

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-70.95439	2.90344	-24.44	<.0001		0
v11	0.38648	0.00996	38.82	<.0001	0.1834	1.73312
v12	-0.23940	0.00996	-24.05	<.0001	0.0406	1.73312
v21	0.03137	0.00212	14.78	<.0001	0.0162	2.61454
v22	0.07491	0.00212	35.29	<.0001	0.0275	2.61454
v31	0.02393	0.00996	2.40	0.0163	0.0003	1.73312
v32	0.02112	0.00996	2.12	0.0339	0.0003	1.73312
v41	0.05022	0.00212	23.66	<.0001	0.0540	2.61454
v42	0.05648	0.00212	26.61	<.0001	0.0806	2.61454

R-Square 0.4030
Adj R-Sq 0.4024

B.7 Regression Model of 3141TR

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-19.23849	0.72188	-26.65	<.0001		0
v11	0.06651	0.00109	61.03	<.0001	0.2681	1.54832
v12	-0.02740	0.00109	-25.14	<.0001	0.0169	1.54832
v21	0.02133	0.00354	6.03	<.0001	0.0018	1.54763
v22	0.06265	0.00354	17.71	<.0001	0.0186	1.54763
v31	0.03924	0.00109	36.00	<.0001	0.1332	1.54832
v32	0.03998	0.00109	36.68	<.0001	0.1640	1.54832
v41	0.03476	0.00354	9.83	<.0001	0.0053	1.54763
v42	0.03905	0.00354	11.04	<.0001	0.0065	1.54763

R-Square 0.6143
Adj R-Sq 0.6139

B.8 Regression Model of 3141LT

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-44.11740	1.83822	-24.00	<.0001		0
v11	0.29164	0.00901	32.38	<.0001	0.1801	1.54763
v12	-0.17767	0.00901	-19.73	<.0001	0.0299	1.54763
v21	0.03050	0.00278	10.99	<.0001	0.0090	1.54832
v22	0.09841	0.00278	35.46	<.0001	0.0562	1.54832
v31	0.01620	0.00901	1.80	0.0721	0.0003	1.54763
v32	0.04511	0.00901	5.01	<.0001	0.0019	1.54763
v41	0.05367	0.00278	19.34	<.0001	0.0470	1.54832
v42	0.05603	0.00278	20.19	<.0001	0.0672	1.54832

R-Square 0.3916
Adj R-Sq 0.3909

B.9 Regression Model of 4141TR

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-20.21426	0.80118	-25.23	<.0001		0
v11	0.06920	0.00123	56.40	<.0001	0.2565	1.56260
v12	-0.02751	0.00123	-22.43	<.0001	0.0167	1.56146
v21	0.01789	0.00396	4.52	<.0001	0.0012	1.54896
v22	0.06101	0.00396	15.40	<.0001	0.0168	1.54790
v31	0.04151	0.00122	33.89	<.0001	0.1397	1.55958
v32	0.03811	0.00123	31.09	<.0001	0.1675	1.56063
v41	0.03458	0.00396	8.74	<.0001	0.0049	1.54646
v42	0.04164	0.00396	10.52	<.0001	0.0070	1.54733

R-Square 0.6103
Adj R-Sq 0.6098

B.10 Regression Model of 4141LT

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Partial R-Square	Variance Inflation
Intercept	-38.37727	1.88868	-20.32	<.0001		0
v11	0.26838	0.00930	28.87	<.0001	0.1726	1.54652
v12	-0.17430	0.00930	-18.75	<.0001	0.0323	1.54664
v21	0.02744	0.00290	9.47	<.0001	0.0080	1.55063
v22	0.09571	0.00290	33.02	<.0001	0.0600	1.55113
v31	0.01715	0.00930	1.84	0.0651	0.0003	1.54501
v32	0.03396	0.00930	3.65	0.0003	0.0012	1.54462
v41	0.05163	0.00290	17.83	<.0001	0.0492	1.54928
v42	0.05366	0.00289	18.54	<.0001	0.0670	1.54867

R-Square 0.3907
Adj R-Sq 0.3900