ABSTRACT

DEVELOPMENT AND APPLICATION OF DYNAMIC MODELS FOR PREDICTING TRANSIT ARRIVAL TIMES

by
Yuqing Ding

Stochastic variations in traffic conditions and ridership often have a negative impact in transit operations resulting in the deterioration of schedule/headway adherence and lengthening of passenger wait times. Providing accurate information on transit vehicle arrival times is critical to reduce the negative impacts on transit users. In this study, models for dynamically predicting transit arrival times in urban settings are developed, including a basic model, a Kalman filtering model, link-based and stop-based artificial neural networks (ANNs) and Neural/Dynamic (ND) models. The reliability of these models is assessed by enhancing the microscopic simulation program CORSIM which can calculate bus dwell and passenger wait times based on time-dependent passenger demands and vehicle inter-departure times (headways) at stops.

The proposed prediction models are integrated with the enhanced CORSIM individually to predict bus arrival times while simulating the operations of a bus transit route in New Jersey. The reliability analysis of prediction results demonstrates that ANNs are superior to the basic and Kalman filtering models. The stop-based ANN generally predicts more accurately than the link-based ANN. By integrating an ANN (either link-based or stop-based) with the Kalman filtering algorithm, two ND models (NDL and NDS) are developed to decrease prediction error. The results show that the performance

of the ND models is fairly close. The NDS model performs better than the NDL model when stop-spacing is relatively long and the number of intersections between a pair of stops is relatively large.

In the study, an application of the proposed prediction models to a real-time headway control model is also explored and experimented through simulating a high frequency light rail transit route. The results show that with the accurate prediction of vehicle arrival information from the proposed models, the regularity of headways between any pair of consecutive operating vehicles is improved, while the average passenger wait times at stops are reduced significantly.

THE DEVELOPMENT AND APPLICATION OF DYNAMIC MODELS FOR PREDICTING TRANSIT ARRIVAL TIMES

by
Yuqing Ding

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Transportation

Interdisciplinary Program in Transportation

January 2000

APPROVAL PAGE

DEVELOPMENT AND APPLICATION OF DYNAMIC MODELS FOR PREDICTING TRANSIT ARRIVAL TIMES

Yuqing Ding

Dr. Steven I-Jy Chien, Dissertation Advisor — Date
Assistant Professor of Department of Civil and Environmental Engineering, NJIT

Dr. Athanassios K. Bladikas, Committee Member — Date
Associate Professor of Industrial and Manufacturing Engineering, NJIT

Dr. Jerome M. Lutin, Committee Member — Date
Senior Director of Planning Research and Development Division,
NJ Transit Corporation

Dr. Kyriacos C. Mouskos, Committee Member — Date
Assistant Professor of Department of Civil and Environmental Engineering, NJIT

Dr. Louis J. Pignataro, Committee Member — Date
Distinguished Professor, Emeritus and Executive Director,
Institute for Transportation, NJIT

Dr. Lazar N. Spasovic, Committee Member — Date
Associate Professor of Management, NJIT

BIOGRAPHICAL SKETCH

Author: Yuqing Ding

Degree: Doctor of Philosophy in Transportation

Date: January, 2000

Undergraduate and Graduate Education:

- Doctor of Philosophy in Transportation
  New Jersey Institute of Technology, Newark, NJ, 2000

- Master of Science in Electrical Engineering,
  Shanghai Jiao Tong University, Shanghai, P. R. China, 1995

- Bachelor of Science in Electrical Engineering
  Huazhong University of Science and Technology, Wuhan, P. R. China, 1992

Major: Transportation Engineering

Presentations and Publications:

Chien, S., and Ding, Y. (1998). Microscopic transit simulation models. Project #421580, Final Report, New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Ding, Y. (1999a). "Applications of artificial neural networks in prediction of transit arrival times." Proceedings, 1999 Annual Conference of ITS America.

Chien, S., and Ding, Y. (1999b). "A dynamic headway control strategy for transit operations." Proceedings, the 6th ITS World Congress, Toronto, Canada.

Chien, S., Chowdhury, M, Mouskos, K. C., and Ding, Y. (1999c). "Enhancements of the CORSIM model in simulating transit vehicle operations." Paper accepted to the Journal of Transportation Engineering, American Society of Civil Engineers.

Chien, S., Ding, Y. and Zayas, N. A. (1999d). "Validation of enhanced CORSIM for simulating transit operations." Working Paper #98-07, New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Ding, Y. (1999e). "Evaluation of the potential for using ramp metering in the ATMS of the I-80 showcase corridor, interim report (I): Literature review." New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Ding, Y. (1999f). "Evaluation of the potential for using ramp metering in the ATMS of the I-80 showcase corridor, interim report (II): Technical/functional requirements and data collection." New Jersey Institute of Technology, Newark, NJ.

Ding, Y., Chien, S., and Zayas, N. A. (1999g). "Microscopic simulation of bus transit operations, case study: bus route #39 of New Jersey Transit." Paper accepted to the 79th Transportation Research Board Annual Meeting, Washington, D.C.

Ding, Y., and Chien, S. (1999h). "The prediction of bus arrival times with link-based artificial neural networks." Paper accepted to Joint Conference on Information Sciences (JCIS), Computational Intelligence & Neurosciences (CI&N), Atlantic City, NJ.

To my beloved Dad, Mom and Husband

ACKNOWLEDGEMENT

I am greatly indebted to my dissertation supervisor Dr. Steven I-Jy Chien, who constantly supports, inspires and encourages me with his insight, dedication and understanding during my PhD study. This dissertation is derived from a research sponsored by New Jersey Institute of Technology, the New Jersey Transportation Information Decision Engineering (TIDE) center and National Center for Transportation and Industrial Productivity (NCTIP). Special thanks are given to Dr. Louis J. Pignataro, the distinguished professor and Executive Director, for his invaluable contribution to the research and thoughtful comments when reviewing the dissertation.

I deeply appreciate Dr. Athanassios K. Bladikas, Dr. Jerome M. Lutin, Dr. Kyriacos C. Mouskos, and Dr. Lazar N. Spasovic for serving as committee members and providing precious suggestions, from both aspects of research and industrial practice. Dr. Chienhung Wei and Dr. John Tavantazis made important contribution regarding the theoretical fundamentals of neural networks and Kalman filtering models, respectively. I am grateful for receiving numerous assistance from Miss Noreen A. Zayas, who worked with me diligently for nearly one year. Other faculty and staff members as well as many of my fellow students should be recognized for their cooperation during my study.

The data collection for this dissertation involved the collaboration of many individuals with New Jersey Transit Corporation and the city governments of Newark, Kearny and Harrison. Their timely help benefited the research mostly.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

Improvement of bus transit service quality in urban settings can benefit both transit users and operators. From the users' perspective, better quality of service facilitates the trip-making at less wait, transfer or travel time. From the operators' perspective, transit demand may be stimulated because of providing reliable and accessible real-time information. However, the characteristics of real-world transit operations (e.g., travel times on links, dwell times at stops and delays at intersections) are stochastic. The combined variations in ridership (e.g., passengers surge from connecting branches) and traffic conditions (e.g., traffic congestion on roadways) further deteriorate headway/schedule adherence, thus lengthen the average passenger wait time and degrade the quality of service.

Providing timely updated transit information, such as vehicle arrival times, departure times and delays, can reduce the negative impact of schedule/headway irregularity on transit service. With the advent of Advanced Public Transportation Systems (APTS), innovative and advanced models are required for predicting accurate information to disseminate to customers. In addition, the predicted information can assist transit agencies to timely restore service disturbances.

The APTS program, one of the major components in Intelligent Transportation Systems (ITS), was initiated by the Federal Transit Administration (FTA) to encourage the applications of emerging technologies in computers, communication, and navigation for promoting the efficiency, effectiveness and safety of public transportation systems.

The APTS technologies, such as Global Position Systems (GPS), Automatic Vehicle Location Systems (AVLS) and Automatic Passenger Counters Systems (APCS), have been implemented in various public transit systems to obtain precise real-time information, including vehicle locations, speeds, and occupancies. Such information can facilitate transit passenger information systems as well as transit planning and management systems, and improve the overall service quality.

1.1 Problem Identification

Due to the stochastic nature of factors (e.g., traffic conditions and transit demand) which affect transit operations, vehicle headways and travel times fluctuate over time and space. Slight variations in headways may be amplified while propagating at downstream stops, which may even cause pairs of vehicles to bunch up. Such a phenomenon significantly influences schedule/headway adherence. Riders, relying on posted transit schedules to arrange their departure times and transfers, will experience longer average wait times than their expectation. This discourages them from using the transit system. In an urban area like the city of Newark, New Jersey, where thousands of passengers rely on public transit every day, schedule/headway deviations may increase average wait times and lengthen travel times significantly. In addition, for certain routes offering intra- or inter- modal connections (e.g., AIRLINK running between Newark Penn Station and Newark International Airport), schedule/headway deviations may also increase passenger transfer delays to other transportation modes.

Accurately predicting vehicle arrival times while considering dynamic variations affecting transit operations in urban networks is a challenging task. Transit operations are influenced by various factors attributable to vehicles (e.g., buses and other traffic), the infrastructure (e.g., roadways/tunnels, stops, terminals, vehicle control and power supply systems), and human behavior (e.g., passengers and drivers). The highly inter-correlated and time varied characteristics of these factors can hardly be well formulated by conventional modeling approaches (e.g., regression and analytical models). Therefore, the prediction models developed by those approaches may generate unreliable information for the public. Moreover, inaccurate information may confuse and render ineffective attempts by transit agencies to engage in real-time dispatching, scheduling and vehicle routing.

1.2 Motivation

Transit operations are often disrupted inevitably by various stochastic factors mentioned previously. Naturally, buses tend to bunch up on service routes. The resulting deviations in headways increase average passenger wait as well as travel times. Advanced prediction models are very critical and essential for providing accurate real-time information, such as expected vehicle arrival/departure times and delays. These models should play an important role in a variety of aspects (e.g., trip-planning, route guidance, service management and operational control) for reducing the negative effects of service disturbance. Thus, transit service quality can be improved.

The proposed prediction models in this study should be one of the most important elements in Traveler Information Systems (TIS), which are responsible for providing accurately and timely updated vehicle arrival information. The predicted information can be disseminated to travelers before trips or en-route through various media (e.g., telephones, pagers, kiosks, the Internet and cable TVs), allowing them either to arrange departure times or coordinate transfers at a lowest cost in terms of travel time. Also, it enables TIS to provide trip guidance, such as suggesting candidate routes with less travel time or number of transfers according to a given of origin-destination pair.

In addition to TIS, the proposed prediction models should have the potential to assist decision-making in transit fleet management and operational control (e.g., scheduling, dispatching, routing, control and signal priority treatment). For example, the determination of appropriate vehicle operating speeds needs a variety of information, such as the current and predicted traffic conditions (e.g., different levels of congestion), passenger demand and occurrence of incidents (e.g., vehicle breakdowns). Currently, the APTS can provide real-time information including vehicle locations and speeds. The development of a sound prediction model is extremely essential for estimating headway/schedule disruptions in advance, and timely initiating effective adjustments.

While advanced real-time prediction models may provide great opportunities for improving transit service levels, an evaluation method is required to be developed to assess the performance of the prediction models. Computer simulation is a cost-effective approach for evaluating advanced prediction models while considering time varied transit demand and traffic conditions. Various emulated real-time information and measures of effectiveness (MOEs) generated by simulation, which may be difficult to be obtained

from field studies, can provide the basis for identifying the most effective model among many competing ones. Hence, the eventual field commitment will have a high probability of success.

1.3 Objectives and Scope

This dissertation mainly focuses on developing dynamic models to accurately predict transit (buses) arrival times. Five primary objectives frame this study.

1. Evaluate contemporary prediction methodologies and models.

   Contemporary prediction models in transportation systems, developed by implementing univariate forecasting models (e.g., probabilistic estimation, time series), multivariate forecasting models (e.g., regression and state-space Kalman filtering), and Artificial Neural Networks (ANNs) approaches are thoroughly reviewed.

2. Develop dynamic models to predict transit (bus) arrival times.

   The proposed advanced prediction models, including ANNs, Kalman filtering and Neural/Dynamic (ND) models, focus on accurately predicting transit vehicle arrive times at all downstream stops in urban settings. They are able to accommodate random fluctuations (e.g., traffic conditions) in transit operations.

3. Develop a microscopic simulation program to simulate transit operations.

   The simulation program is developed to simulate transit operations on integrated transportation corridors (e.g., urban surface streets and freeways) while considering time varied passenger demand and traffic conditions. The program is calibrated and

validated through simulating an actual transit route and comparing simulation results with field observations.

4. Evaluate the developed prediction models through the proposed simulation program.

   The proposed prediction model is integrated with the simulation program individually, while the accuracy of the prediction results is evaluated by conducting reliability analysis.

5. Explore applications of the proposed prediction models.

   Applications of the proposed prediction models on advanced vehicle control systems are assessed through developing a real-time headway control model. The operational impacts (e.g., headway variations and passenger wait times) are evaluated through simulation.

1.4 Organization of the Dissertation

This dissertation is organized into seven chapters. In Chapter 1, the problem is identified and the objectives and scope of the study are addressed. In Chapter 2, an overview of current prediction methodologies and models is presented, while the review of micro-, meso-, and macroscopic simulation models is also encompassed. In Chapters 3 and 4, the development of the real-time prediction models and the applications are discussed, respectively. The proposed microscopic simulation program developed through enhancing CORSIM is described in Chapter 5, while the calibration and validation of the simulation program are also addressed in this chapter. In Chapter 6, the results generated by the proposed prediction and control models are demonstrated and finally, a conclusion

of this dissertation is stated in Chapter 7. The overview of the dissertation is shown in Figure 1.1.

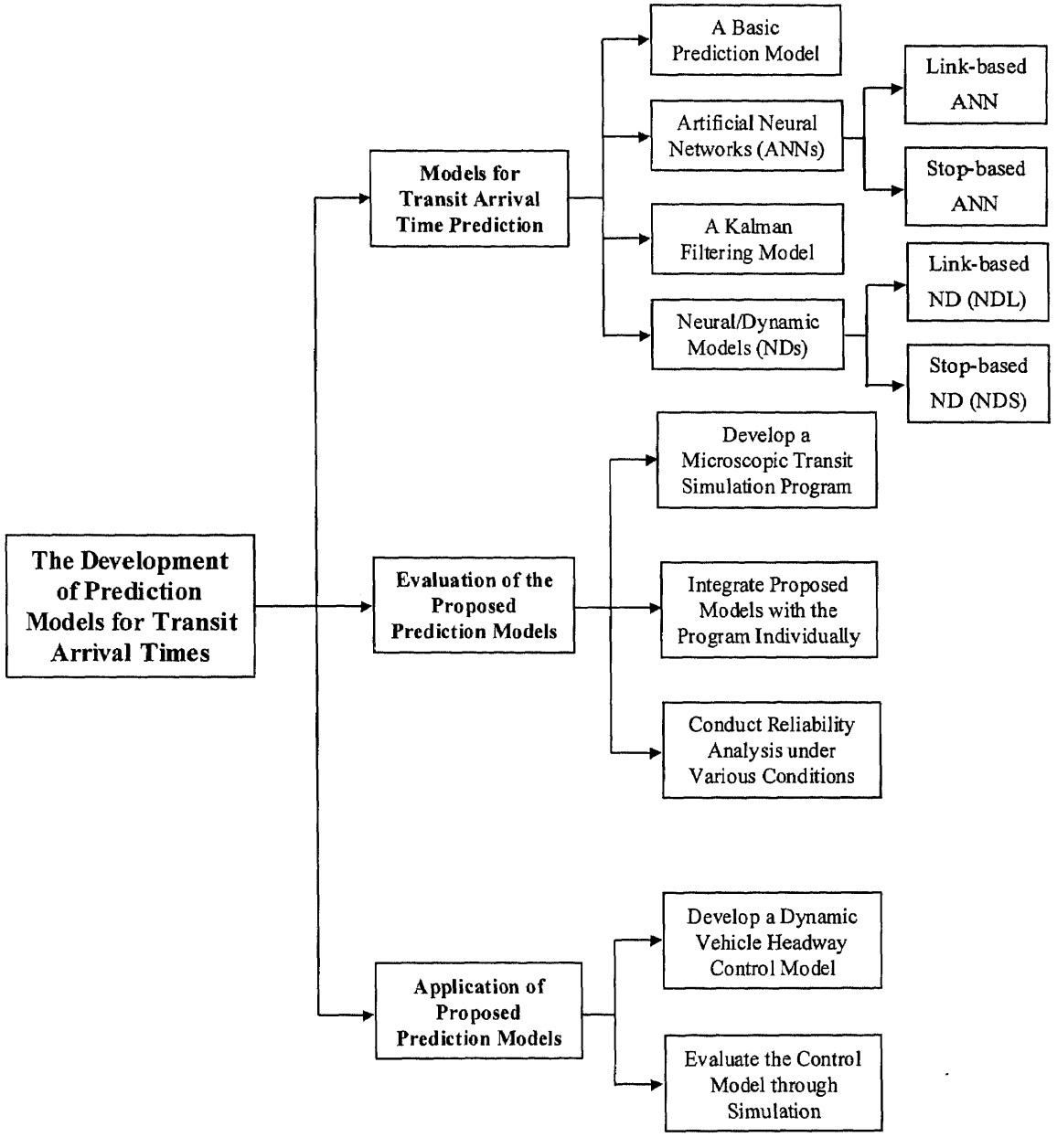


Figure 1.1 Overview of the dissertation

CHAPTER 2

LITERATURE REVIEW

The literature review investigates contemporary prediction models (e.g., univariate, multivariate and artificial neural network models) for transit operations that help to assess the potential improvement of service quality through accurately and timely updated arrival time information in the context of Advanced Public Transportation Systems (APTS). In the review, simulation models (e.g., micro-, meso-, and macroscopic simulation models) on transit operations as tools for providing various emulated real-time information and MOEs are also discussed. Such simulation models can help to evaluate the operational impacts of various traffic management and control strategies (e.g., real-time vehicle prediction and control models).

2.1 Transit Arrival Time Prediction

Transit vehicle (e.g., bus) movements along routes are disturbed frequently due to stochastic variations in vehicle travel times on links, random delays at intersections, and dwell times at stops. If a transit vehicle falls slightly behind its schedule for any reason, it will have more passengers than usual to pick up at the next stop, which increases arrival delays and passengers wait times at further downstream stops. As a result, the vehicle keeps falling further behind its schedule. The follower vehicle encounters fewer passengers than usual and thus has less dwell time at stops. This situation may even cause the follower vehicle to bunch up with its preceding vehicle (Lin, et. al., 1995; Chien and Chowdhury, 1997). Such a phenomenon deteriorates the adherence to vehicle

arrival/departure schedule and headway. The resulting large gaps between pairs of vehicles discourage riders to use the transit system because of the increased wait times. One way to reduce the wait time under this situation is to provide timely and accurate information, such as vehicle arrival/departure times and expected delays. With that, users can effectively reschedule their departure times (Federal Transit Administration, Update'98).

Currently, Automatic Vehicle Location Systems (AVLS) can provide real-time information of operating vehicles, such as locations and speeds. However, that information can not help much to alleviate user wait and transfer times without a decent prediction system. For example, AVLS can not provide the travel time and delay information twenty minutes later on a downstream link that users will be travelling on. Therefore, it is essential to develop a reliable model for providing predicted vehicle arrival and departure times. Such information can then be disseminated through Traveler Information Systems (TIS) and accessed by travelers at their homes, work places, terminal centers, wayside stops or on-board through a variety of media (e.g., TRAVELINK, Minneapolis, Minnesota; PA/CIS in New York City, NY; AZTech in Phoenix, Arizona and SMARTBUS in Atlanta, Geogia). Thus, the travelers are able to plan their trips and departure times for reducing their travel times (Abdelfattah and Khan, 1997; Federal Transit Administration, Update'98).

The predicted information can help public transportation agencies in managing and operating their systems (e.g., real-time dispatching, scheduling and control) (Dougherty, et. al., 1993; Smith and Demetsky, 1995). Moreover, the headway/schedule variations can be estimated accurately in advance. Therefore, a proper control action (e.g.,

increase/decrease operating speed or dwell longer time at some stops) can be determined. Such control can maintain a desirable level of service by dynamically restoring the disruptions in schedules or headways.

Transit vehicle arrival times are affected by many factors, such as passenger demands at stops, traffic control devices, right-of-way, roadway geometry, traffic volumes, and other unexpected factors (e.g., weather conditions, incidents, and construction activities). However, the time varied relationship between the predicted arrival time and these factors is difficult to be formulated mathematically. The previously developed models which can be potentially used for vehicle arrival prediction are classified into three categories: (1) Univariate (e.g., probabilistic estimation and time series), (2) Mutivariate (e.g., Regression and State-space), and (3) Artificial Neural Networks. Each of these categories is discussed below.

2.1.1 Univariate Forecasting Models

The univariate forecasting models are designed to predict the value of a variable (dependent variable) through describing the intrinsic relationship in historical data mathematically, without considering external factors as explanatory variables. The commonly used univariate forecasting models for the purpose of vehicle arrival prediction include probabilistic estimation and time series models.

Probabilistic Estimation Models

Some previous studies (Turquist, 1978; Tally and Becker, 1987; Guenther and Hamat, 1988) estimated the vehicle arrival time at a stop by analyzing statistics and probability

distributions. Based on field observations of vehicle arrivals at stops and the hypotheses tests on arrival time distributions, probabilistic models were thus established. These models are used to analyze the statistical characteristics of vehicle arrival times (e.g., means and variations among different stops) and the corresponding average passenger wait times. However, the probabilistic models are not very helpful to predict arrival times accurately in real-time, because they can not respond to dynamic changes in traffic conditions and transit demands.

Turquist (1978) demonstrated that the predicted vehicle arrival times followed a log normal distribution, which had a long tail to the right (late arrivals) while truncated to the left (early arrivals). According to his observation, a vehicle was considered more liably to be late than early. This is because a late vehicle usually encountered a larger number of waiting passengers and longer dwell time, which caused further delays as the vehicle traveled to downstream stops. Thus, the log-normal distribution, which was consistent with the field observation, was used for estimating vehicle arrival times in that study.

Tally and Becker (1987) suggested a symmetric exponential distribution (longer tails at both sides) for estimating vehicle arrival times. His model assumed that vehicles had equal probabilities of early and late arrivals. The data used in that study were collected from 41 transit routes in TTDC (Tidewater Transportation District Commission Virginia), which were divided into early and late arrivals. The exponential distribution of vehicle arrivals was accepted after statistical testing.

Guenther and Hamat (1988) analyzed empirical data collected from four transit routes at the Milwaukee County Transit Systems, Wisconsin. After conducting a

hypothesis test, they found that the vehicle arrival times at a stop fitted a Gamma distribution. This study indicated that the probability of early arrivals is relatively lower than that of late arrivals because of heavy traffic and passenger demands.

Time Series Models

By describing stochastic patterns of observed data (e.g., trend, cyclical and seasonal), time series models such as WMA (weighted moving average) and ARIMA (autoregressive integrated moving average) models can relate future traffic conditions to historical observations (Brockwell and Davis, 1991). Under an APTS environment, real-time information can be obtained by traffic surveillance and communication systems (e.g., AVLS) through various monitoring devices (e.g., detectors, ultrasonic beacons, video image processors). The traffic data, recorded periodically (e.g., every 5-minute), are used by time series models for short-term prediction of traffic volumes and vehicle travel times (Okutani and Stephanedes, 1984; Stephanedes, Kwon and Michalopoulos, 1990; Kwon and Stephanedes, 1994). These models with constant parameters (e.g., smoothing and autocorrelation coefficients) have a short time lag for real-time implementation. However, the accuracy of the prediction highly relies on the similarity between current and historical traffic conditions.

Urban Traffic Control Systems (UTCS), developed by FHWA, are widely used by transportation professionals for traffic demand prediction. The first generation of UTCS (UTCS-1) used a WMA model to predict traffic volume at the next time interval (ranging from 5 to 15 minutes) as the weighted average of recent measurements, while the weights, called smoothing coefficients, were determined based on historical observations

charaterized by different types of days (e.g., weekdays, weekends and holidays). The second generation of UTCS (UTCS-2) employed both recent measurements and historical observations for volume prediction. The prediction model was formulated as a function of historical average volume adjusted by a linear combination of residuals (deviations between recent measurements and the corresponding historical average) (Okutani and Stephanedes, 1984). UTCS-2 performed better than UTCS-1 because of its low mean square errors. However, the reliability of both models highly depend on the collected historical volume data. Since traffic demand may vary substantially from its historical average in real-time, UTCS can not respond accurately to such variations.

ARIMA (autoregressive integrated moving average) models are one of the most commonly used time series models for short term prediction, whose parameters (e.g., orders of autoregression and integration) were determined by Box-Jenkins, Akaike's Information Criterion or Schwarz Bayesian methods (Delurgio, 1998). Previous research on the application of the ARIMA models (e.g., seasonal ARIMA called SARIMA) for traffic volume prediction did not show great superiority over other prediction approaches (Kwon and Stephanedes, 1994; Smith and Demetsky, 1995). Sometimes these models may even cause high inaccuracy under a dynamic traffic environment (Okutani and Stephanedes, 1984). Recent studies on improving the prediction accuracy with ARIMA models include the state-space ARIMA and the integrated ARIMA and Kalman filtering models (Brockwell and Davis, 1991; Williams, et. al., 1998).

2.1.2 Multivariate Forecasting Models

Different from univariate models, multivariate models can forecast and explain a dependent variable with a mathematical function formed by a set of independent variables (include external factors). The multivariate models may be a powerful analytical tool; however, such models are still difficult to reveal the underlying cause-and-effect relationship between the vehicle arrival time and those various affecting factors (e.g., passenger demands at stops, traffic control devices, roadway geometry, traffic conditions, and other unexpected factors). The commonly used multivariate models, including regression models and state-space Kalman filtering models, are discussed below.

Regression Models

As a conventional modeling approach for analyzing vehicle arrivals (Abkowitz, et. al., 1984; Abdelfattah and Khan, 1997), regression models measured the simultaneous influences of various factors affecting vehicle arrivals through correlation and significant tests on field observations. Pre-specified functions (e.g., linear combination of explanatory variables) are established while the parameters are determined by minimizing the mean square error. However, due to the difficulty in capturing stochastic traffic conditions, the regression models can not predict accurately. Additionally, to establish a regression model, all selected explanatory variables have to be significant to the dependent variable (usually measured by t-value), and independent between one and another. Such a requirement may limit the accuracy of regression models to predict vehicle arrival times for transit systems containing various inter-correlated (multicollinear) and time varied factors.

Abkowitz and Engelstein (1984) analyzed transit operations at Queen City, Cincinnati. Factors, including traffic volumes, traffic control devices, parking restrictions, travel directions, time of day, and passenger demand, were examined as explanatory variables to the vehicle travel time. A linear regression model was developed in relation to the transit travel time and distance, boarding/alighting passengers, and the number of signalized intersections between stops. The model was tested by the data collected from Route 44 in Los Angeles, while the results showed that the average of the predicted arrival time could statistically represent the field average at a 95% confidence interval. However, very large deviations between the predicted and observed arrival times were found in that study, which deteriorated the reliability of the model.

Abdelfattah and Khan (1997) conducted a regression analysis based on the data generated by simulating a transit route in Ottawa-Carleton, Ontario. Three scenarios, including (1) general operation conditions, (2) high percentage of heavy vehicles and (3) one-lane blocked, were simulated with TRAF-NETSIM. In that study, traffic density, link length, the number of stops per link, and efficiency ratio (moving time/total travel time) were identified as factors affecting vehicle link travel time. Both linear and non-linear regression models were developed to estimate vehicle travel times at downstream stops under the three scenarios. The developed regression models were tested by using the data collected from a segment of the analyzed transit route, which only contains two intersections. The results showed that the travel times predicted by non-linear models (deviated from 0 to 60 seconds compared with field observations) had higher reliability than those predicted by linear ones.

Kalman Filtering Models

Kalman filtering is a statistical time series approach, which was originated from state-space representations in linear control theory. Unlike regression models, the Kalman filtering model can describe dynamic systems with its time dependent parameters (e.g., Kalman Gain), which can be dynamically optimized (Gelb, et. al., 1977; Stephanedes, et. al., 1990). Kalman filtering models have been used for prediction in various areas such as signal processing, numerical weather prediction, and aircraft tracking (Brockwell and Davis, 1991).

In recent years, Kalman filtering models have been applied in various transportation areas, such as predicting street traffic volumes (Okutani and Stephanedes, 1984) and ramp entering/exiting volumes (Stephanedes, Kwon and Michalopoulos, 1990; Kwon and Stephanedes, 1994; Wall, et. al., 1998). These models accommodated the dynamic variations in traffic conditions by dynamically updating the model parameters through minimizing the prediction error. The Kalman filtering models have elegant mathematical representations (e.g., linear state-space equations); however, establishing Kalman filtering models that can deal with the multicollinear relationship among multiple factors is still hard to achieve. In addition, the linearized approximation in Kalman filtering models (Gelb. et. al., 1977) may neglect some important characteristics, such as the non-linear reactions among these factors, which may cause prediction inaccuracy.

Okutani and Stephanedes (1984) developed a model for predicting short-term traffic volumes in Nagoya City, Japan. In that study, the traffic volume was predicted as a linear combination of measurements obtained from previous three time intervals. The

parameters in the prediction model can be on-line adjusted (every 5-minute interval) with a Kalman filtering model. The results were compared with those generated by UTCS-2, and showed that the accuracy of the Kalman filtering model substantially outperformed that of UTCS-2 (up to 80%).

In 1990, Stephanedes, Kwon and Michalopoulos developed a model to predict ramp entering volumes during 6:00AM to 9:00AM for the Westbound of I-35 corridor in Minneapolis, Minnesota. In that study, the ramp entering volume of the next 5-minute interval was predicted in real-time, and the parameters of the prediction model were updated every 5 minutes with a Kalman filtering algorithm. The study showed that the average of prediction errors ranged from 8.5% to 18.8%. Later, in 1994, they evaluated three prediction models: UTCS-2, a Kalman filtering model, and an artificial neural network (ANN), by comparing the predicted exit ramp volumes with the actual volume for the same site. The prediction results showed that both the Kalman filtering model and the ANN outperformed UTCS-2. Furthermore, the ANN demonstrated better performance in terms of accuracy than the Kalman filtering model.

Wall, et. al. (1998) developed a model for predicting transit arrival times in real-time by integrating two sequential components. These components were: (1) a Kalman filtering model to track the location of a target vehicle by processing the data obtained from AVL systems, and (2) a probabilistic model to predict vehicle arrival times based on an assumed normal vehicle arrival distribution and current vehicle location on the route. The prediction model was tested by data collected from two vehicle trips on a 30-minute transit route (no data source was specified), and the results showed the deviations between predicted and actual arrival times at a stop were less than 5 minutes. However,

the state-space equation of the Kalman filtering model was found erroneous (Tavantzis and Ding, 1999), which makes the prediction results presented by Wall's study spurious.

2.1.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) have been gaining popularity in transportation recently (Hua, et. al., 1994; Chin, et. al., 1994; Dougherty, 1995). ANNs, a powerful modeling approach, were motivated by attempting to simulate human brains, which can quickly identify, understand, and deduce inferences by the experience learned from the past. With versatile parallel distributed structures and adaptive learning processes, ANNs seem a promising approach to describe complex systems such as transit operations affected by various inter-correlated and time-varied factors.

Unlike the prediction models mentioned previously, ANNs do not have to specify the forms of functions (e.g., state-space equations in Kalman filtering models), which need sophisticated modeling techniques and parameter estimating processes for time-varied and non-linear systems. A well-trained ANN can capture a complex relationship between the dependent variable (output such as arrival times) and a set of explanatory variables (input such as traffic conditions and passenger demands). Thus, the ANN is extremely useful when the functional relationship between the input and output is hard to formulate mathematically.

An ANN may consist of multiple layers of processing units (named neurons), within which activated functions (either linear or non-linear) are contained. The data process in an ANN can be feedforward (with data only processed from one layer to the

next layer) or recurrent (with data processed from one layer to the next, previous, or even the same layer). The ANN with only linear activated functions is equivalent to a multivariable linear regression model (Warner and Misra, 1996). In addition, the ANN may act as a time series model if it has a recurrent data processing structure with time series input (Kalaputatu and Demetsky, 1995). It has been proved that a three-layer feedforward network with sigmoid activated functions in its hidden and output layers can represent any continuous function with a high degree of accuracy (Hagan, et. al., 1996; Warner and Misra, 1996).

The learning process of an ANN can be viewed as an optimization problem. For example, the synaptic weights of the ANN can be optimized through minimizing the performance function (e.g., prediction error function in this study) for all training examples. The learning procedure can be either off-line (trained by a complete set of examples obtained before learning starts) or on-line (trained with examples obtained from the analyzed system during operation). The selection of a proper ANN structure and learning rule, though, depends on the characteristics of the analyzed system. In general, heuristic and experimental procedures dominate the selection of these parameters. The most often used learning rules are discussed below.

Supervised Learning Rules

The supervised learning rules are usually used when the desired output is available. Early studies on supervised learning rules include Perceptron, ADALINE and MADALINE. Perceptron was used for mapping linear relationships between input and output, while

ADALINE and MADALINE can only deal with two-layer ANNs (Maren, et. al. 1989). Later, the back-propagation (BP) algorithm (Rumelhart, et. al., 1986) was demonstrated to properly describe a complex relationship with multiple-layered ANNs. The BP algorithm assumes that all neurons and synaptic weights are responsible for the difference between the actual and predicted results. With the BP algorithm, the error (e.g., the sum of squared errors (SSE), the mean squared error (MSE), the root mean squared error (RMSE), or the mean absolute error (MAE)) is propagated backwards from the output layer to its previous layers, and is adjusted through updating the synaptic weights. It is worth noting that the BP algorithm is probably the most widely applied learning rule for training ANNs today, mainly because of its excellent capability to generate correct relationships between input and output (Hua, et. al., 1994; Dougherty, 1995).

The BP algorithm can be implemented in either an off-line or an on-line training process. However, the training period is usually long because all training examples have to be entered repeatedly until the synaptic weights are optimized (Warner and Misra, 1996).

Self-organized Learning Rules

Unlike the supervised learning rules, the self-organized learning rules work only on input data (or input patterns). Self-organizing is usually applicable for classification or pattern recognition when there is no desired output known before learning begins. For example, in classification, the training data need to be categorized into different classes that are unknown beforehand. During the learning process, the performance of ANNs is computed (e.g., Euclidian distance). The winning neuron is identified (e.g., with the minimum

distance) and updated by the "winner takes all" rule (e.g., only the synaptic weights for the winning neuron are strengthened). Examples of self-organizing include Learning Vector Quantization (LVQ) and Topology-Preserving Map (TPM) (Kohonen, et. al., 1988). Both techniques are capable for pattern recognition, while TPM is specifically more applicable to deal with topological input patterns. However, LVQ and TPM can only be used for off-line training.

Unlike the aforementioned uni- and multivariable models, ANNs can be established without specifying the form of the function. Thus, the restrictions on the multicollinearity of explanatory variables can be neglected. However, ANNs never reveal the function explicitly. Instead, it is buried deeply within the networks, which may hinder the mathematical analysis (e.g., Hessian) and hypothesis tests. ANNs are often referred to as *black boxes*, lack straightforward theoretical guidelines on choosing the input variables, number of hidden layers, number of neurons on each layer, network topology (e.g., fully and partially connected) and learning parameters (e.g., learning and momentum rates). Therefore, the learning process has to be conducted on a space with great complexity in order to search for the optimal solution.

Kalaputapu and Demetsky (1995) developed ANNs with time series features for predicting vehicle schedule derivations (can be transformed to vehicle arrival times based on posted schedules). Both feedforward and recurrent ANNs (e.g., Elman and Jordan nets) were developed and evaluated based on data collected from AVL systems in Tidewater Regional Transit, Virginia. Although only historical vehicle arrival data were considered as model inputs, this pioneer work is encouraging for establishing ANNs for

vehicle arrival prediction in real-time while considering various traffic conditions (e.g., traffic volumes, speeds and delays) and passenger demands.

As a prominent tool to solve complex problems, ANNs have been applied in a variety of transportation prediction areas, which are shown in Table 2.1. It is notable that the BP algorithm was applied predominantly. In those studies, ANNs have demonstrated the potential to accurately predict traffic conditions (e.g., traffic volume, travel time and speed, O/D flow and queue length) on freeways (Hua and Faghri, 1994; Smith and Demetsky, 1994; Zhang, Ritchie and Lo, 1997) and urban streets (Chin, Hwang and Pei, 1994; Chang and Su, 1995). The results showed that the ANN is a promising approach for prediction compared with conventional approaches (e.g., regression and time series models), since ANNs can handle complex and dynamic traffic characteristics adequately under various situations (e.g., traffic congestion, lane blockage).

After recognizing the need for predicting vehicle arrival times and reflecting various dynamic factors in real-time, various models are developed and evaluated in this research, including ANNs, Kalman filtering models and other advanced prediction models.

Table 2.1 Previous prediction studies on application of ANNs in transportation

Author/Year	Predicted Subject	Learning Rule	Network Structure	Decision Variables	Objective Function
Chang and Su (1995)	queue length	BP	three-layer	flow, occupancy, speed and historical queue length	SSE
Chin, Hwang and Pei (1994)	O/D flow	BP	three-layer	Entering/existing traffic volumes	MAE
Dougherty, Kirby and Boule (1993)	traffic congestion	BP	three-layer	flow, queue length, volume ratio	RMSE
Hua and Faghri (1994)	travel time	BP	two-layer	traffic volume/blockage index	SSE
Kalaputapu and Demetsky (1995)	schedule variance	BP	three-layer	scheduled bus arrive time/ schedule variance	MSE, SSE
Smith and Demetsky (1994)	traffic volume	BP	three-layer	historical volume/speed	RMSE
Wei and Yang (1998)	Freight load	BP	three-layer	Ecnomic growth rate, industrial production index and wholesale price index	RMSE
Zhang, Ritchie and Lo (1997)	travel speed	BP	three-layer	historical speed, density, ramp entry rate	SSE

two-layer: one input and one output layers
three-layer: one input, one hidden and one output layers
SSE: sum of squared errors
MAE: mean absolute error
RMSE: root mean square error
MSE: mean square error
BP: Back-propagation learning rule

2.2 Transit Simulation Model

As a practical approach to evaluate complex systems, computer simulation is applied in transportation widely (FHWA, 1996; Prevedouros and Wang, 1998; Ho, Chien and Ting, 1999). A bus transit system operating in a large-scale network while considering time varied transit demand and traffic conditions has been simulated by Chien and Chowdhery in 1997 and Chien and Ding in 1998. In this study, the simulation approach is used for

evaluating the accuracy of developed prediction models for the following reasons. It is relatively inexpensive to obtain various real-time data to estimate MOEs needed in the study through simulation. In addition, simulation helps to experiment with different prediction models under various traffic conditions and identify the most reliable one, which facilitates the eventual field commitment.

According to the level of detail with which an analyzed transportation system can be described, simulation models can be classified into three categories: (1) microscopic, (2) mesoscopic and (3) macroscopic, which are discussed below.

Microscopic models (e.g., DRACULA, FLEXSYT II, MITSIM, CORSIM) can describe vehicle/driver behavior in different traffic environments with a high level of detail. For example, each vehicle can be identified by type (e.g., auto, carpool, truck and bus/train), and driver characteristics (e.g., cautious and aggressive), while the stochastic properties of driving behavior can be simulated individually (e.g., lane-change maneuver invoking car-following logic and driver decision process). Microscopic models can accurately describe the dynamic interactions among vehicles operating in mixed traffic. Such models, however, usually require fairly detailed geometric and traffic data and a large number of parameters (e.g., nodes, links, distributions of acceptable left-turn gaps and driver types) for calibrating real-world networks as well as generating MOEs.

Mesoscopic models (e.g., INTEGRATION, DYNASMART, TRAF-NETFLO Level 1) describe vehicles at a relatively lower level of detail than the microscopic ones, while handling driving behavior at an aggregate manner in some situations. For example, the lane-change decision of an individual vehicle is determined by an aggregate characteristic on the target lane such as lane density rather than the detailed interactions

(e.g., car-following logic) between vehicles. On the other hand, macroscopic models (e.g., TRAF-FREEFLO, TRAF-NETFLO Level 2) simulate both vehicles and driving behavior in a low level of detail, such that the real driving behavior of each vehicle (e.g., lane-change acceleration/deceleration, response lag time) can not be simulated at all. Thus, the simulated traffic stream is represented by an aggregate traffic flow rate, speed and density. Compared with microscopic models, mesoscopic and macroscopic models are of less detail in their descriptions but their simulation speeds are faster than the microscopic ones. The selection of proper simulation models is highly dependent on the purpose of the application and the underlying features of the analyzed network (Ho, Chien and Ting, 1998).

DYNASMART (University of Texas at Austin, 1992) was designed to evaluate dynamic route assignment strategies with the advent of Advanced Traveler Information Systems (ATIS). DYNASMART is a mesoscopic model which simulates vehicles individually according to aggregate speed-density relations, thus, more detailed vehicle maneuvers can not be simulated.

INTEGRATION (Queen's University and M. Van Aerde and Associates, Lts., Canada, 1997), a mesoscopic model, can simulate large-scaled transportation networks containing freeways and surface streets, while assessing effects of traffic management strategies including adaptive traffic control, route guidance and traffic assignment. INTEGRATION describes movements of each individual vehicle based on speed-flow relationships of each link. A distinct feature of INTEGRATION is that it considers the impact of ITS route guidance information in its vehicle routing logic. However, INTEGRATION can not simulate special signal timing, such as a protected left-turn.

DRACULA (University of Leeds, UK, 1993), a microscopic model, can simulate traffic variations over time, while public transit operations on reserved lanes can be enhanced. FLEXSYT II (Ministry of Transport, German, 1994), a microscopic model, was developed to analyze the effects of dynamic control strategies, such as traffic signals, HOV lanes and ramp metering. FLEXSYT II can simulate transit operations on both exclusive and shared lanes. However, users are required to specify origin-destination demand for every intersection since FLEXSYT II has no traffic assignment model.

MITSIM (Massachusetts Institute of Technology, USA, 1996), a microscopic model, was designed for evaluating innovative ITS traffic control and incident management strategies including signal pre-emptive control for transit vehicles. MITSIM has a dynamic route guidance model, allowing users to choose routes according to current traffic conditions.

TRAF (FHWA, 1992) contains a group of simulation models with different levels of detail (micro- meso- or macroscopic), which can be used to evaluate various transportation system management (TSM) strategies (e.g., ramp metering and traffic signal control). In TRAF, NETSIM and NETFLO simulate surface streets, while FRESIM and FREFLO simulate freeways. In addition, ROADSIM simulates two-lane rural roads. TRAF provides users with the flexibility to select simulation models for their specific needs.

The windows version 4.2 of TSIS/CORSIM (CORridor SIMulator), released by McTrans (June, 1997), was developed by integrating two microscopic simulation models in TRAF: NETSIM and FRESIM. CORSIM has a user-friendly graphical interface and environment distributed with the Traffic Software Integrated System (TSIS), which can

simulate traffic operations and control systems on integrated networks containing freeways and surface streets.

In CORSIM, the stochastic lane-change and operational behavior of vehicles operating in a complex and large scale roadway networks can be described. Transit vehicles will change to a proper lane and gradually slow down when they are approaching stops. CORSIM has been validated through various studies (FHWA, 1996; Prevedouros and Wang, 1998; Ding and Chien, 1999e) and widely used for developing and evaluating various advanced traffic management strategies (ATMS) (e.g., traffic assignment, signal optimization and ramp metering control) (FHWA, 1996). However, CORSIM can not properly simulate transit vehicles dwelling at stops (Chien and Chowdhery, 1997; Chien and Ding, 1998).

The vehicle dwell time is considered as one of major factors influencing the regularity of vehicle headways, especially for a long line with heavy ridership (Lin and Wilson, 1992). Most of the previous studies estimated dwell times according to the numbers of boarding and alighting passengers and average boarding/alighting time (Koffman, 1978; Vuchic, 1981; Lin and Wilson, 1992; Eberlein, et. al. 1998; Chien and Ding, 1998). Some studies (e.g., Adamski and Turnau, 1998; Chien and Schonfeld, 1997) estimated vehicle dwell time based on the mean headway.

Koffman (1978) suggested that a vehicle delayed at a stop by 4.3 seconds for each boarding passenger and 2 seconds for each alighting passenger, based on the data collected from a transit route in Cambridge, Massachusetts. Vuchic (1981) estimated dwell times for vehicles with different number of doors, while the numbers of

boarding/alighting passengers and average boarding/alighting time were considered as decision variables.

Lin and Wilson (1992) developed regression models to estimate vehicle dwell times for a high frequent LRT system based on empirical data (Boston Green Line). The effects of passenger crowdiness were analyzed while considering the numbers of boarding/alighting passengers and standees.

In CORSIM, however, the vehicle dwell time is determined by a user specified mean dwell time and a random number generated from an embedded distribution, rather than the actual number of boarding/alighting passengers arriving during vehicle inter-departure time. Thus, CORSIM can not properly model transit operations from this point of view, especially when handling vehicles dwelling at stops (Chien and Ding, 1998).

After evaluating the advantages and disadvantages of various simulation models, a simulation program is developed here by enhancing CORSIM. In this simulation program, the vehicle dwell times are determined based on the numbers of boarding and alighting passengers as well as the time-dependent vehicle inter-departure times (headways). The simulation program can provide a reliable testing environment to emulate transit operations and the surveillance and communication systems in a real-world traffic network. In addition, the models developed in this study for the prediction of vehicle arrivals are evaluated through the enhanced program.

CHAPTER 3

METHODOLOGY

Several models developed for predicting vehicle arrival times are discussed in this chapter, including a basic model, a Kalman filtering model, two Artificial Neural Networks (ANNs) (link-based and stop-based) and two Neural/Dynamic (ND) models. The ND models are developed by incorporating an ANN (either link-based or stop-based) with a Kalman filtering model. As a major component of the proposed ND models, ANNs are established to predict vehicle arrival times, which can capture the time varied relationship between the vehicle arrival time and its affecting factors. The Kalman filtering model, another component of ND models, can adjust on-line the prediction error caused by different degrees of traffic congestion. The reliability of the proposed prediction models is evaluated through the simulation program, which is discussed in Chapter 5.

In this chapter, the basic prediction model is described in Section 3.1, while the ANNs and the Kalman filtering model are discussed in Sections 3.2 and 3.3, respectively. The ND models and a brief summary are presented in Section 3.4.

3.1 Basic Prediction Model

The basic model proposed here can predict vehicle arrival times at all downstream stops by simply projecting the current average link speeds to the near future. In the basic model, the vehicle arrival time at any downstream stop is estimated by adding up the link travel times to downstream stops according to real-time information (e.g., average link

speeds). Note that a link on a transit route can be defined by a pair of upstream and downstream nodes, which can represent either a signalized or unsignalized intersection. Assume that two consecutive stops i-1 and i have n links in between and there is a vehicle k traveling towards stop i at time t, as shown in Figure 3.1. At time t, the estimated

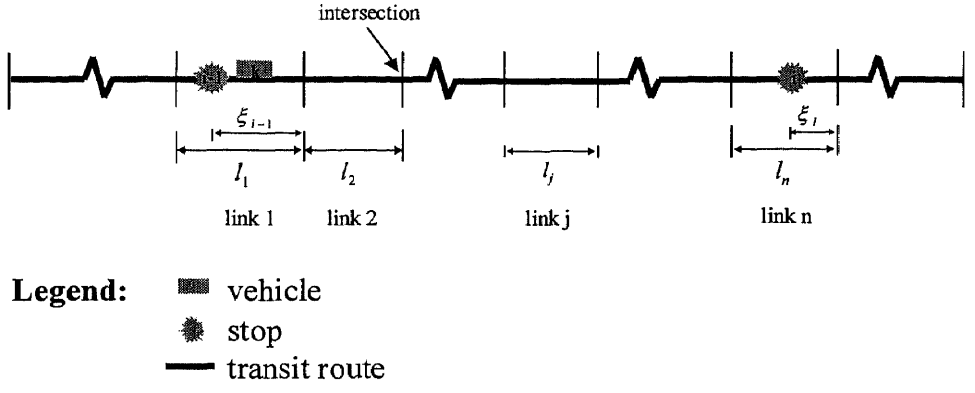


Figure 3.1 A transit route with the basic model

arrival time $E_{k,i}^{(t)}$ for vehicle k at stop i is its departure time $p_{k,i-1}$ at stop i-1 plus the travel time $\Lambda_{k,i}^{(t)}$ from stop i-1 to i.

$$E_{k,i}^{(t)} = p_{k,i-1} + \Lambda_{k,i}^{(t)} \qquad \text{for i} = 2, 3,...,S \qquad (3.1)$$

In Eq. 3.1, $p_{k,i-1}$ is either known at time t (i.e., $p_{k,i-1} \le t$, $\forall k,i$) or otherwise can be estimated ($p_{k,i-1} > t$, $\forall k,i$) (refer to Eq. 3.12), while $\Lambda_{k,i}^{(t)}$ (i = 1, 2,...,S) can be obtained based on current average link travel times as $p_{k,i-1} > t$,

$$\Lambda_{k,i}^{(t)} = \frac{\xi_{i-1}}{\overline{V}_1^t} + \sum_{j=2}^{n-1} \frac{l_j}{\overline{V}_j^t} + \frac{l_n - \xi_j}{\overline{V}_n^t} \qquad \text{for i} = 2, 3,...,S \qquad (3.2)$$

where $\overline{V}_j^t$ represents the average speed of link j at time t, l_j represents the length of link j (j=1, 2, ...,n), and ξ_i is the distance between stop i and its immediate downstream

intersection. Note that l_j (j = 1, 2,..., n) represents the links between stops i-1 and i. If $p_{k,i-1} \le t$ (vehicle k has already departed from stop i-1 at time t), the distance from the current vehicle location to stop i should be considered instead. By substituting Eq. 3.2 into Eq. 3.1, the vehicle arrival time can be estimated by Eq. 3.3

$$E_{k,i}^{(t)} = p_{k,i-1} + \frac{\xi_{i-1}}{\bar{V}_1^t} + \sum_{j=2}^{n-1} \frac{l_j}{\bar{V}_j^t} + \frac{l_n - \xi_i}{\bar{V}_n^t} \quad \text{for i} = 2, 3, ...,S \tag{3.3}$$

A reliability analysis for Eq. 3.3 is required for evaluating the accuracy of estimated arrival times. The root mean square error (RMSE) obtained from Eq. 3.4 is selected for assessing the reliability of the prediction results,

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(E_{k,i}^{(t)} - A_{k,i})^2} \tag{3.4}$$

where N represents the sample size. Eq. 3.4 shows that the closer $E_{k,i}^{(t)}$ to the actual vehicle arrival time $A_{k,i}$, the lower the prediction error RMSE, indicating higher accuracy of the predicted arrival times.

3.2 Artificial Neural Networks (ANNs)

As a powerful modeling approach, Artificial Neural Networks (ANNs) have demonstrated the potential to capture complex relationships between inputs and outputs of dynamic systems (e.g., a transit system). Therefore, the ANN approach is used here to predict vehicle arrival times at downstream stops, while time varied transit demand and traffic conditions are considered. The interconnected structure and Back-propagation (BP) learning rules used to establish the proposed ANNs are discussed below.

3.2.1 Introduction to ANNs

Of the many structures (e.g., feedforward, recurrent) available for ANNs, the multilayer feedforward network (e.g., one input layer, a number of hidden layers, and one output layer) with sigmoidal activation functions is chosen, due to its capacity to adapt to complex systems. As shown in Figure 3.2, each neuron on a layer connects to all neurons on the next layer with weighted and unidirectional links. The neurons on the input layer simply transmit the input data (factors affecting vehicle arrivals) onto the first hidden layer, while the neurons on the hidden layers contain non-linear transfer functions (e.g., sigmoid functions) to process the weighted sum of all inputs. Thus, the output of a layer will be the input of the next layer. Such a fully inter-connected and parallel distributed structure enables ANNs to describe complex systems such as transit operations within which various factors are inter-correlated and time varying.

Given the observed input/output data examples, an ANN can be trained by a properly designed learning procedure (e.g., Back-propagation) to estimate the relationship between inputs and outputs. During the training process, the synaptic weights can be optimized through minimizing the performance functions (e.g., total error functions) of all training examples. Therefore, the well-trained ANN is able to predict information such as vehicle arrivals adequately (Chien and Ding, 1999a; Ding and Chien, 1999f). The Back-propagation (BP) leaning algorithm, which has been widely applied for prediction because of its desirable reliability (Kalaputapu and Demetsky, 1995; Chang and Su, 1995; Hagan, 1996), is selected to train the proposed ANN in this study.

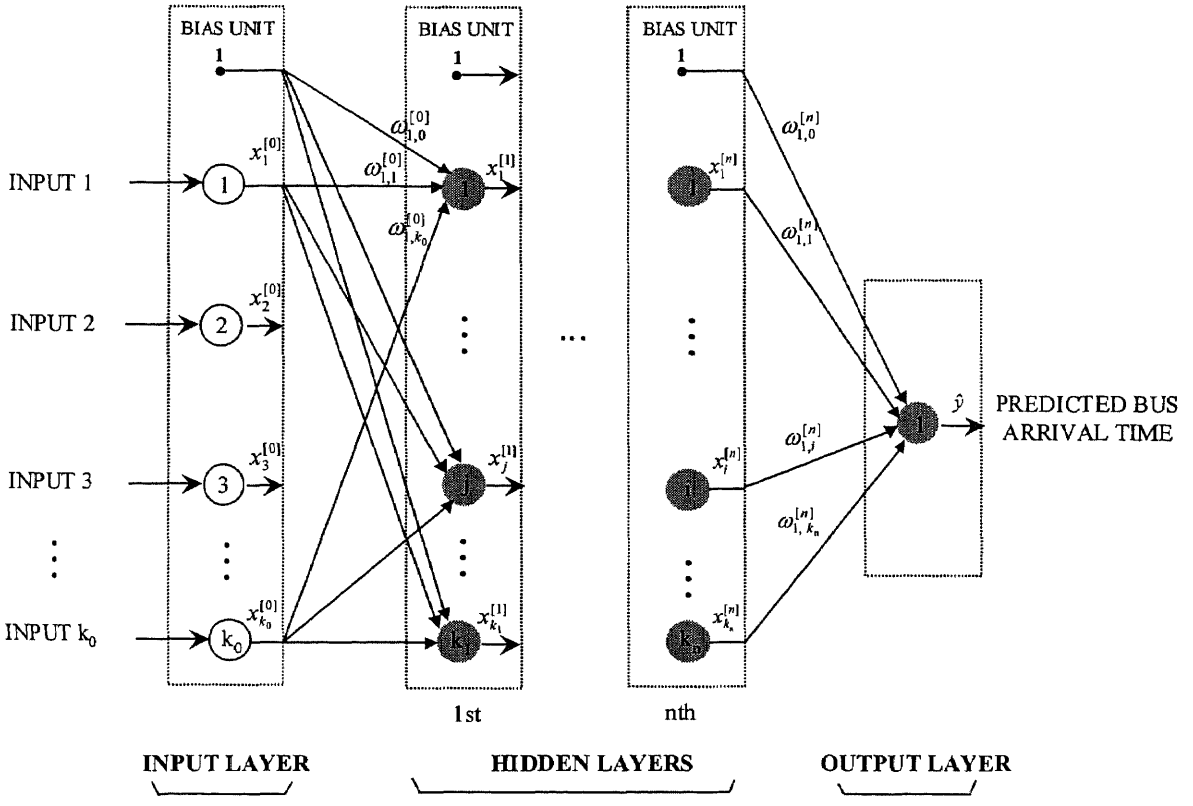


Figure 3.2 Configuration of ANNs for transit arrival time prediction

The BP algorithm coded for this research uses a steepest-decent gradient method to minimize the prediction error over all N training examples (input/output data). The error function e_r is defined by the sum of squared errors and formulated as

$$e_r = \frac{1}{2}\sum_{p=1}^{N}(y_p - \hat{y}_p)^2 \tag{3.5}$$

where y_p and $\hat{y}_p$ represent the observed and predicted values (e.g., vehicle arrival times) for the p^{th} training example, respectively.

By applying a BP algorithm for training the ANN, there are k_0 inputs for the p^{th} training example (k_0 is the number of factors affecting vehicle arrivals such as volumes,

speeds, and delays on links) transmitted from the input layer to the output layer, and the predicted value $\hat{y}_p$ can be generated. As shown in Figure 3.3, the j th neuron (j = 1, 2,..., kl) on the l^{th} hidden layer receives the weighted sum of all inputs from the neurons on

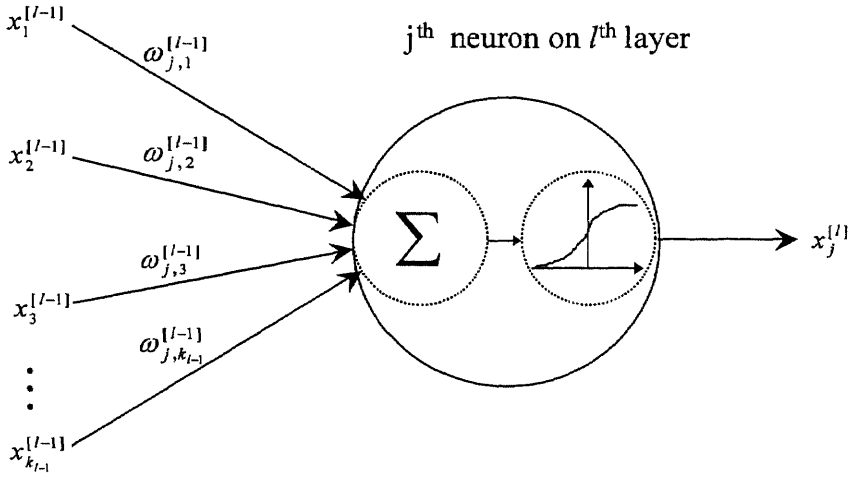


Figure 3.3 The jth neuron on the l^{th} hidden layer

the (l-1)th layer and processes them with the transfer function (e.g., sigmoid function). The output $x_j^{[l]}$ from the jth neuron on the l^{th} hidden layer can thus be obtained as (Warner, et. al., 1996)

$$x_j^{[l]} = \frac{1}{1+\exp(-\sum_{i=1}^{k_{l-1}} \omega_{j,i}^{[l-1]}\, x_i^{[l-1]})} \qquad \text{for j} = 1, 2,..., \text{k}_l \text{ and } l = 1, 2,\ldots,\text{n} \qquad (3.6)$$

Eq. 3.6 represents a sigmoid function, where $\omega_{j,i}^{[l-1]}$ is the synaptic weight between the ith neuron (i = 1, 2,..., k$_{l-1}$) on the (l-1)th layer and the jth neuron (j = 1, 2,..., kl) on the l^{th} hidden layer.

Therefore, the neuron on the output layer will process the inputs $x_i^{[n]}$ (i = 1, 2,..., k_n) from the n^{th} hidden layer, and the predicted $\hat{y}_p$ can be obtained from Eq. 3.7.

$$\hat{y}_p = \frac{1}{1+\exp(-\sum_{i=1}^{k_n} \omega_{1,i}^{[n]} x_i^{[n]})} \tag{3.7}$$

where $\omega_{1,i}^{[n]}$ is the synaptic weight between the i^{th} neuron (i = 1, 2,..., k_n) on the n^{th} hidden layer and the neuron on the output layer.

By substituting $\hat{y}_p$ (p is the index of training examples) into Eq. 3.5, the prediction error e_r can be obtained. In order to minimize e_r, the synaptic weights linking two consecutive layers are optimized by implementing the BP algorithm. The suggested adjustment of $\Delta\omega_{j,i}^{[l]}$ is (Rumelhart, et. al., 1986)

$$\Delta\omega_{j,i}^{[l]} = -\eta \frac{\partial e_r}{\partial \omega_{j,i}^{[l]}} + \gamma \Delta\overline{\omega}_{j,i}^{[l]} \qquad \text{for } l = 0, 1,\ldots, n \tag{3.8}$$

where η represents the learning rate in the BP algorithm (η>0) and scales the changes in $\omega_{j,i}^{[l]}$ during the optimization process. A large value of η can speed up the process; however, the learning process may be unstable (e.g., wild oscillations) because of the large changes in adjusted synaptic weights. On the other hand, a small value of η may lengthen the learning period (Hagan, et. al., 1996).

The purpose of introducing the momentum rate γ (usually ranging between 0 and 1) and the previously adjusted weights $\Delta\overline{\omega}_{j,i}^{[l]}$ into Eq. 3.8 is to increase the convergence speed and bound the changes in synaptic weights, thus prevent the oscillations. In order to efficiently select optimal η and γ, a heuristic procedure (Darken and Moody, 1992;

Wei and Yang, 1998) is used in this study. With Eq. 3.8, the synaptic weights $\omega_{1,i}^{[n]}$ linking the n^{th} (last) hidden layer and the output layer are the first set of synaptic weights to be adjusted using Eq. 3.9:

$$\Delta\omega_{1,i}^{[n]} = -\eta\ \delta_1\ x_i^{[n]} + \gamma\ \Delta\overline{\omega}_{1,i}^{[n]} \qquad \text{for i} = 1, 2,..., k_n \tag{3.9}$$

where $\delta_1 = \hat{y}_p\,(y_p - \hat{y}_p)(1-\hat{y}_p)$ is called the local delta responsible for the synaptic weights for the output layer. Therefore, the synaptic weight $\omega_{j,i}^{[l-1]}$ linking two layers l-1 and l (l = n, n-1,...,1) are adjusted by Eq. 3.10

$$\Delta\omega_{j,i}^{[l-1]} = -\eta\,\delta_j^{[l]}\ x_i^{[l-1]} + \gamma\ \Delta\overline{\omega}_{j,i}^{[l-1]} \qquad \text{for j} = 1, 2,..., k_l \text{ and i} = 1, 2,..., k_{l-1} \tag{3.10}$$

where $\delta_j^{[l]} = \sum_{i=1}^{k_{l+1}} \delta_i^{[l+1]}\ \omega_{i,j}^{[l]}\ x_j^{[l]}\ (1-x_j^{[l]})$ is the local delta for l^{th} layer, which is determined by the local delta $\delta_i^{[l+1]}$ (i = 1, 2,..., k_{l+1}) propagated from the $(l+1)^{th}$ layer. Note that if l is equal to n, $\delta_i^{[l+1]}$ denotes the local delta for the output layer δ_1. Based on Eqs. 3.9 and 3.10, all synaptic weights can be adjusted backwards from the output layer to the input layer, as illustrated in Table 3.1.

The process of adjusting synaptic weights can be regarded as performing the non-linear fitting over all training examples. The training examples are entered cyclically until the synaptic weights converge, while the prediction error for all N training examples is minimized. The proposed ANN with such a highly inter-connected structure and optimal synaptic weights can thus capture the time varied relationship between input and output of the analyzed system.

Table 3.1 Summary of the BP algorithm for the p^{th} training example

I. Information Processing From the Input Layer to the Output Layer	
Output from the input layer	$x_i^{[0]}$ $(i = 1, 2,..., k_0)$
Output from the l^{th} hidden layer	$x_j^{[l]} = \frac{1}{1+\exp(-\sum_{i=1}^{k_{l-1}} \omega_{j,i}^{[l-1]} x_i^{[l-1]})}$ $(j = 1, 2,..., k_l;\ l = 1, 2,...,n)$
Prediction from the output layer	$\hat{y}_p = \frac{1}{1+\exp(-\sum_{i=1}^{k_n} \omega_{1,i}^{[n]} x_i^{[n]})}$
II. Adjustment of Synaptic Weights From the Output Layer to the Input Layer	
Local delta for the output layer	$\delta_1 = \hat{y}_p (y_p - \hat{y}_p)(1 - \hat{y}_p)$
Adjustment of $\omega_{1,i}^{[n]}$ linking the n^{th} hidden layer and the output layer	$\Delta\omega_{1,i}^{[n]} = -\eta\ \delta_1\ x_i^{[n]} + \gamma\ \Delta\overline{\omega}_{1,i}^{[n]}$ $(i = 1, 2,..., k_n)$
Local Delta for the n^{th} hidden layer	$\delta_i^{[n]} = \delta_1 \omega_{1,i}^{[n]} x_i^{[n]} (1 - x_i^{[n]})$ $(i = 1, 2,..., k_n)$
Adjustment of $\omega_{j,i}^{[l-1]}$ linking the $(l\text{-}1)^{th}$ and l^{th} hidden layers	$\Delta\omega_{j,i}^{[l-1]} = -\eta\,\delta_j^{[l]}\, x_i^{[l-1]} + \gamma\ \Delta\overline{\omega}_{j,i}^{[l-1]}$ $(l = n, n\text{-}1,...,1;\ j = 1, 2,..., k_l;\ i = 1, 2,..., k_{l-1})$
Local Delta for the 1^{st} hidden layer	$\delta_j^{[l]} = \sum_{i=1}^{k_{l+1}} \delta_i^{[l+1]}\ \omega_{i,j}^{[l]}\ x_j^{[l]}\ (1 - x_j^{[l]})$ $(l = n\text{-}1, n\text{-}2,...,1;\ j = 1, 2,..., k_l;\ i = 1, 2,..., k_{l+1})$
III. Update the Synaptic Weights $\omega_{j,i}^{[l]}$ with $\omega_{j,i}^{[l]} + \Delta\omega_{j,i}^{[l]}$ $(l = 0, 1..., n)$	

The training examples could be obtained from the data collected on site or from simulation outputs. The collected data are pre-processed (e.g., sorting data into different categories such as training and testing data sets and removing redundant/erroneous data) before starting the training process. This pre-process can increase convergence speed during the learning period, and increase the reliability (accuracy) of the predicted results.

3.2.2 Prediction of Transit Arrival Times with ANNs

For predicting vehicle arrival times, a variety of real-time data are collected and processed in this study, including those affecting traffic conditions (e.g., link volumes, occupancies, speeds and delays) and transit operations (e.g., vehicle arrival and departure times, dwell times, numbers of boarding/alighting passengers). These data can be collected by traffic or transit monitoring systems (e.g., detectors, autoscopes, AVLS, APCS) or extracted from simulation results. Considering that vehicle arrival times can be captured by either accumulating the travel time at each link or estimating stop-to-stop travel times directly, both link-based and stop-based ANNs are developed. The two ANNs are designed with different structure handling different real-time data, as discussed next.

Link-based ANN

The link-based ANN is designed to predict vehicle arrival times by accumulating vehicle travel times on all traversed links between a pair of stops. Assuming that between two consecutive stops i-1 and i (Figure 3.4), there are m links numbered from 1 to m from upstream to downstream. At time t, the predicted arrival time $E_{k,i}^{(t)}$ for vehicle k at stop i can be determined by adding the total vehicle travel time $\sum_{j=1}^{m}\Gamma_{k,j}^{(t)}$ over the m links to the departure time $p_{k,i-1}$ from stop i-1, such that

$$E_{k,i}^{(t)} = p_{k,i-1} + \sum_{j=1}^{m}\Gamma_{k,j}^{(t)} = p_{k,i-1} + \Phi_L\left(\mathbf{X}_j(t)\right) \tag{3.11}$$

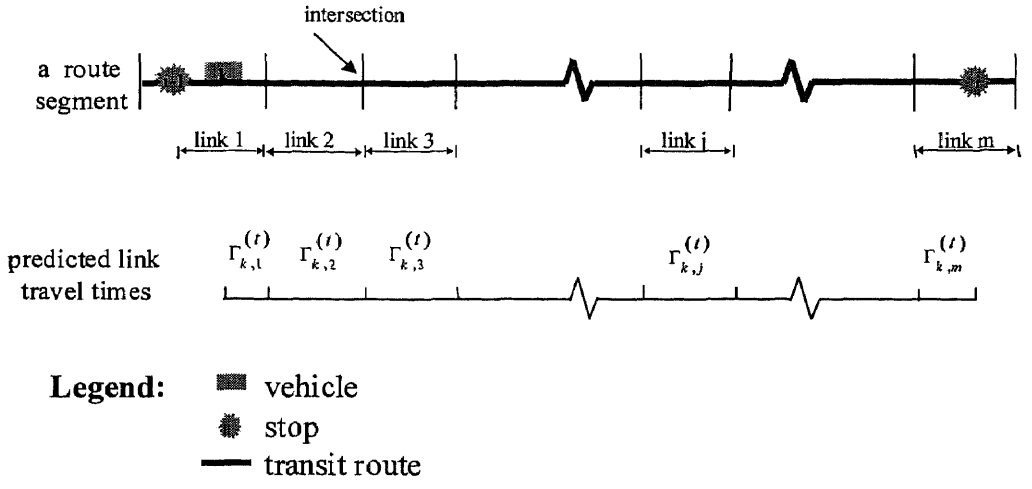


Figure 3.4 A transit route with link-based travel times

where $p_{k,i-1}$ (i.e., , $\forall k,i$) is estimated based on the predicted vehicle arrival time $E_{k,i}^{(t)}$ and dwell time $d_{k,i-1}$ of vehicle k at stop i-1,

$$p_{k,i-1} = E_{k,i-1}^{(t)} + d_{k,i-1} \tag{3.12}$$

The duration of $d_{k,i-1}$ is determined by the number of boarding/alighting passengers incurred during the departure time of vehicle k-1 and arrival time of vehicle k at stop i-1 (Chien and Ding, 1998; Ding and Chien, 1999f).

In Eq. 3.11, $\sum_{j=1}^{m} \Gamma_{k,j}^{(t)}$ can be predicted by $\Phi_L(\cdot)$, a function of $\boldsymbol{X}_j(t) = [x_{1,j}(t), x_{2,j}(t), ..., x_{n,j}(t)]^T$ (j = 1, 2, …, m) representing those factors that may affect vehicle link travel times (e.g., link volumes, speeds and delays). $\boldsymbol{X}_j(t)$ can be obtained from real or simulated traffic surveillance systems, while $\Phi_L(\cdot)$ can be captured by a well-trained ANN.

The configuration of the link-based ANN is shown in Figure 3.5. The number of boarding/alighting passengers are generated assuming that a passenger origin-destination

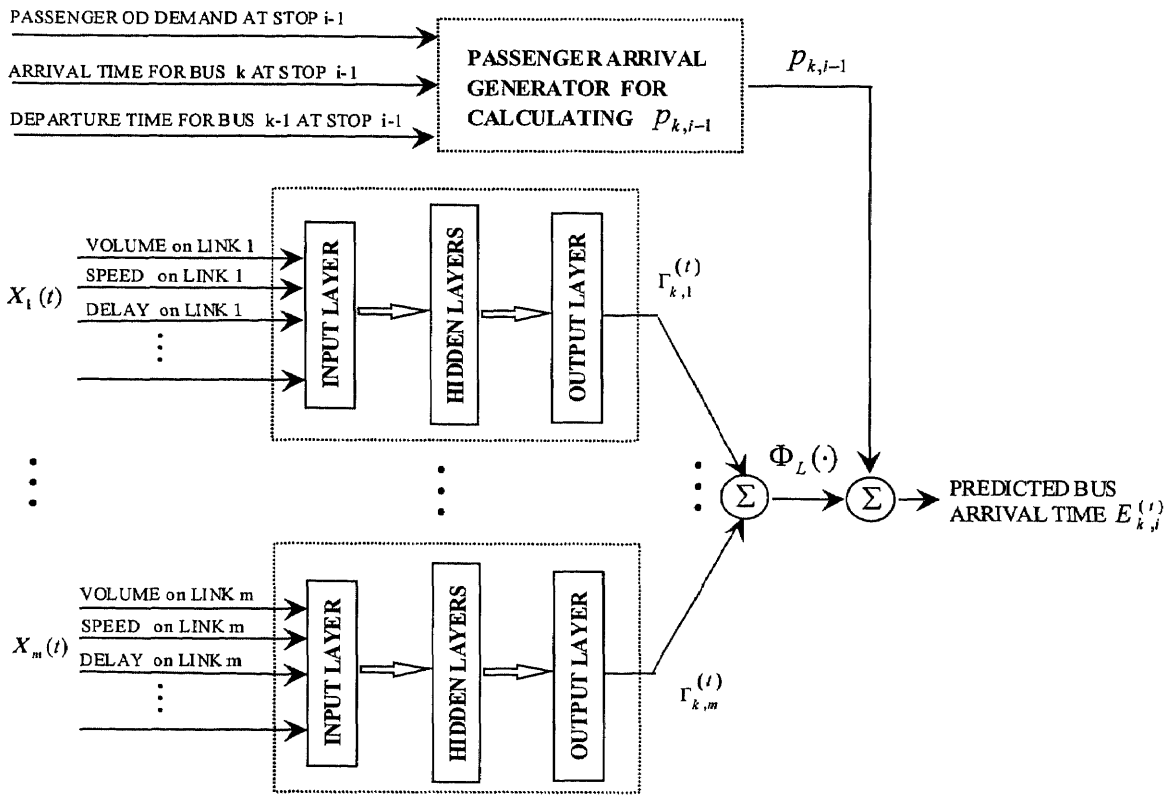


Figure 3.5 Configuration of the link-based ANN

(OD) demand matrix is given, and passenger arrivals at stops follow Poisson distributions. Since the dwell time $d_{k,i-1}$ of vehicle k at stop i-1 can be estimated, the departure time $p_{k,i-1}$ from stop i-1 can be obtained. Meanwhile, the vehicle travel time from stop i-1 to i is predicted by $\Phi_L(\cdot)$. Thus, the vehicle arrival time $E_{k,i}^{(t)}$ at stop i can thus be determined.

Stop-based ANN

Unlike the link-based ANN, the stop-based ANN estimates the vehicle stop-to-stop travel time directly, which is determined by aggregate traffic conditions (e.g., means and variances of link volumes, speeds and delays) of all traversed links between a pair of stops. Thus, the interaction of traffic conditions among links can be considered as a factor influencing vehicle travel times. As shown in Figure 3.6, at time t, the predicted

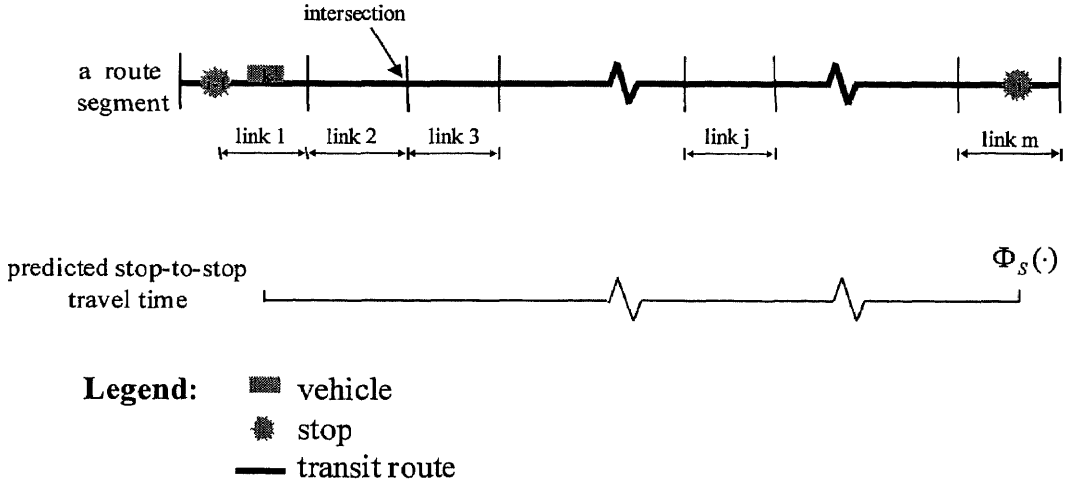


Figure 3.6 A transit route with stop-based travel times

arrival time $E_{k,i}^{(t)}$ for vehicle k at stop i can be obtained by the sum of the vehicle departure time $p_{k,i-1}$ at stop i-1 and the stop-to-stop travel time predicted by $\Phi_S(\mathbf{Z}(t))$,

$$E_{k,i}^{(t)} = p_{k,i-1} + \Phi_S(\mathbf{Z}(t)) \tag{3.13}$$

where $p_{k,i-1}$ is the departure time for vehicle k at stop i-1. In Eq. 3.13, $\Phi_S(\cdot)$ is a function of an n by 1 vector $\mathbf{Z}(t)$, which contains factors (e.g., the average and standard deviation (SD) of traffic volumes, speeds and delays among the m links) affecting stop-

to-stop travel times. $\Phi_S(\cdot)$ represents the stop-based ANN, which predicts vehicle stop-to-stop travel times.

The configuration of the stop-based ANN is shown in Figure 3.7. The dwell time

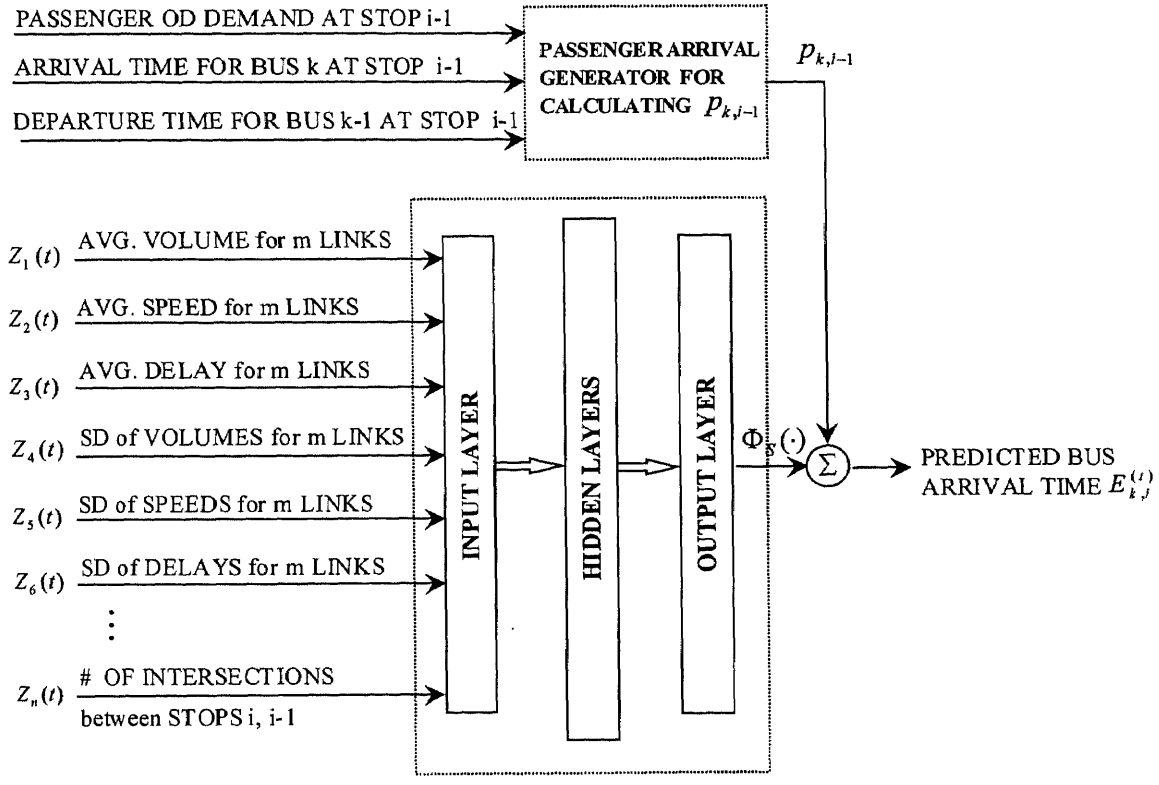


Figure 3.7 Configuration of the stop-based ANN

$d_{k,i-1}$ of vehicle k at stop i-1 is estimated based on number of boarding/alighting passengers during the time-dependent headway between vehicles k-1 and k. Thus, vehicle departure time $p_{k,i-1}$ from stop i-1 can be obtained. Since the vehicle travel time $\Phi_S(\cdot)$ from stop i-1 to i can be predicted by a well-trained stop-based ANN, the vehicle arrival time $E_{k,i}^{(t)}$ at stop i can be determined.

3.3 Kalman Filtering Model

Kalman filtering, a statistical time series modeling approach originated from state-space representations in linear control theory, has been applied in the transportation area for prediction in recent years (Stephanedes, Kwon and Michalopoulos, 1990; Kwon and Stephanedes, 1994). Such a modeling approach has the potential to accommodate traffic fluctuations adequately with its time-dependent parameters (Kalman Gain), which can be adjusted dynamically by applying the Kalman filtering algorithm.

Assume that a segment of a transit route has two stops i-1 and i, as shown in Figure 3.8, while vehicle k departs from stop i-1 at time t. The leading vehicle k-1 has departed from stop i, while the follower vehicle k+1 is still heading to stop i-1.

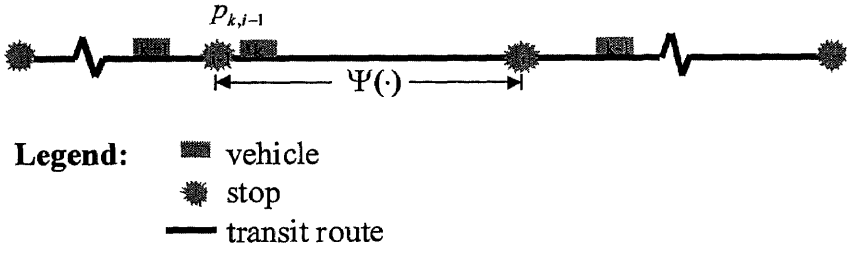


Figure 3.8 A transit route with the Kalman filtering model

The estimated arrival time $E_{k,i}^{(t)}$ of vehicle k at stop i at time t can be obtained by the sum of the vehicle departure time $p_{k,i-1}$ at stop i-1 and the stop-to-stop travel time, which can be formulated as a non-linear function $\Psi(\cdot)$ of the factors affecting vehicle operations, such as link volumes, speeds and delays.

$$E_{k,i}^{(t)} = p_{k,i-1} + \Psi[\mathbf{Y}(t)] \tag{3.14}$$

where $\Psi(\cdot)$ is a function of vector $\mathbf{Y}(t) = [y_1(t), y_2(t), \cdots, y_n(t)]^T$ representing various factors affecting vehicle travel times. By assuming that $\Psi(\cdot)$ is a linear combination of $\mathbf{Y}(t)$ (Brockwell and Davis, 1991), the prior[1] predicted arrival time for vehicle k at stop i, called $E_{k,i}(-)$, can be formulated as

$$E_{k,i}(-) = p_{k,i-1} + \boldsymbol{\Theta}_{k,i}^T(-)\mathbf{Y}(t) + v(t) \tag{3.15}$$

where $\boldsymbol{\Theta}_{k,i}(-)$, an n by 1 Jacobian vector, can be obtained when the leading vehicle k-1 arrived at stop i (refer to Eq. 3.21). The residuals of the approximation of $\Psi(\cdot)$ in Eq. 3.14 with the linear combination $\boldsymbol{\Theta}_{k,i}^T(-)\mathbf{Y}(t)$ in Eq. 3.15 are represented with $v(t)$, which can be described as uncorrelated random noise (zero mean, covariance $R(t)$) (Delurgio, 1998). Because such linearization may cause a prediction inaccuracy, the Kalman filtering algorithm is thus introduced to update $E_{k,i}(-)$ on-line by optimizing the Jacobian vector $\boldsymbol{\Theta}_{k,i}(-)$, as discussed below (Gelb, et. al., 1977).

Suppose that the updated Jacobian vector $\boldsymbol{\Theta}_{k,i}(+)$ can be represented by the prior Jacobian vector $\boldsymbol{\Theta}_{k,i}(-)$, which is disrupted by an uncorrelated noise vector (Okutani and Stephanedes, 1984)

$$\boldsymbol{\Theta}_{k,i}(+) = \boldsymbol{\Theta}_{k,i}(-) + \mathbf{w}(t) \tag{3.16}$$

where $\mathbf{w}(t)$ is a n by 1 random noise vector (zero mean, covariance matrix $\mathbf{Q}(t)$). In Eq.

[1] In the dissertation, - and + denote prior and after updating with Kalman filtering algorithm, respectively.

3.16, since the updated Jabobian vector $\boldsymbol{\Theta}_{k,i}(+)$ contains noise, the Kalman filtering algorithm is applied to obtain the optimal estimation $\hat{\boldsymbol{\Theta}}_{k,i}(+)$.

$$\hat{\boldsymbol{\Theta}}_{k,i}(+) = \hat{\boldsymbol{\Theta}}_{k,i}(-) + \mathbf{K}_{k,i}^{(t)} \rho_{k-1,i} \tag{3.17}$$

where $\rho_{k-1,i}$ is the difference between the actual and estimated stop-to-stop travel times of vehicle k-1 at stop i. In Eq. 3.17, a n by 1 Kalman Gain, $\mathbf{K}_{k,i}^{(t)}$, can be obtained in real-time based on the prior error covariance matrix $\mathbf{M}_{k,i}(-)$ for vehicle k (Gelb, et. al., 1977)

$$\mathbf{K}_{k,i}^{(t)} = \mathbf{M}_{k,i}(-)\mathbf{Y}(t)\{\mathbf{Y}(t)^T \mathbf{M}_{k,i}(-)\mathbf{Y}(t) + R(t)\}^{-1} \tag{3.18}$$

In Eq. 3.18, $R(t)$ is the covariance of noise $v(t)$, such that $R(t) = E[v(t)^2]$, while the prior n by n error covariance matrix $\mathbf{M}_{k,i}(-)$ of vehicle k can be determined, after the arrival of the leading vehicle k-1 at stop i.

By substituting Eq. 3.18 into Eq. 3.17, the optimal estimation $\hat{\boldsymbol{\Theta}}_{k,i}(+)$ can be obtained, and then the predicted vehicle arrival time $E_{k,i}(+)$ of vehicle k at stop i can be updated as

$$E_{k,i}(+) = p_{k,i-1} + \hat{\boldsymbol{\Theta}}_{k,i}^T(+)\mathbf{Y}(t) \tag{3.19}$$

The extrapolation of the algorithm for predicting the arrival time of follower vehicle k+1 at stop i includes calculating the prior error covariance matrix $\mathbf{M}_{k+1,i}(-)$ and the prior Jacobian vector $\hat{\boldsymbol{\Theta}}_{k+1,i}(-)$ (Gelb, et. al., 1977). The n by n matrix $\mathbf{M}_{k+1,i}(-)$ for vehicle k+1 can be obtained from Eq. 3.20.

$$\mathbf{M}_{k+1,i}(-) = [\mathbf{I} - \mathbf{K}_{k,i}^{(t)}\mathbf{Y}(t)]\mathbf{M}_{k,i}(-) + \mathbf{Q}(t) \tag{3.20}$$

where $\mathbf{I}$ is an n by n identity matrix, and $\mathbf{Q}(t)$ is the covariance of noise vector $\mathbf{w}(t)$ such that $\mathbf{Q}(t) = E[\mathbf{w}(t)\mathbf{w}(t)^T]$. The Jacobian vector $\hat{\mathbf{\Theta}}_{k+1,i}(-)$, used to predict the arrival time of the follower vehicle k+1, can be formulated in Eq. 3.21 (Gelb, et. al., 1977),

$$\hat{\mathbf{\Theta}}_{k+1,i}(-) = \hat{\mathbf{\Theta}}_{k,i}(+) \tag{3.21}$$

Thus, the arrival times of all the follower vehicles at stop i can be predicted based on an iterative process of Eqs. 3.15 through 3.19. Thus far, the Kalman filtering algorithm for predicting vehicle arrival times is established and summarized in Table 3.2.

The equations for the Kalman filtering algorithm listed in Table 3.2 can be regarded as optimal linear estimators (in the sense of minimizing prediction error). The mathematical abstraction of the analyzed transit system can be described by the prediction model and the Jacobian vector model in Table 3.2, whose parameters can be on-line adjusted. Thus, the vehicle arrival times can be predicted. Clearly, the accuracy of the prediction highly depends on how well the established Kalman filtering model fits into the underlying relationship of dependent and independent variables in a real world transit system.

Table 3.2 Summary of the Kalman filtering algorithm*

Prediction Model	$E_{k,i}(-) = p_{k,i-1} + \boldsymbol{\Theta}_{k,i}^T(-)\mathbf{Y}(t) + v(t) \quad v(t) \sim N(0, R(t))$
Jacobian Vector Model	$\boldsymbol{\Theta}_{k,i}(+) = \boldsymbol{\Theta}_{k,i}(-) + \mathbf{w}(t) \quad \mathbf{w}(t) \sim N(0, \mathbf{Q}(t))$
Initialization	$\hat{\boldsymbol{\Theta}}_{0,i}(-) = E[\boldsymbol{\Theta}_{0,i}(-)]$ $\mathbf{M}_{0,i}(-) = E\{[\boldsymbol{\Theta}_{0,i}(-) - \hat{\boldsymbol{\Theta}}_{0,i}(-)][\boldsymbol{\Theta}_{0,i}(-) - \hat{\boldsymbol{\Theta}}_{0,i}(-)]^T\}$
Kalman Gain Matrix	$\mathbf{K}_{k,i}^{(t)} = \mathbf{M}_{k,i}(-)\mathbf{Y}(t)\{\mathbf{Y}(t)^T\mathbf{M}_{k,i}(-)\mathbf{Y}(t) + R(t)\}^{-1}$
Jacobian Vector Estimation	$\hat{\boldsymbol{\Theta}}_{k,i}(+) = \hat{\boldsymbol{\Theta}}_{k,i}(-) + \mathbf{K}_{k,i}^{(t)}\rho_{k-1,i}$
Prediction Update	$E_{k,i}(+) = p_{k,i-1} + \hat{\boldsymbol{\Theta}}_{k,i}^T(+)\mathbf{Y}(t)$
Error Covariance Update	$\mathbf{M}_{k+1,i}(-) = [\mathbf{I} - \mathbf{K}_{k,i}^{(t)}\mathbf{Y}(t)]\mathbf{M}_{k,i}(-) + \mathbf{Q}(t)$
Jacobian Vector Update	$\hat{\boldsymbol{\Theta}}_{k+1,i}(-) = \hat{\boldsymbol{\Theta}}_{k,i}(+)$

*Note: The choice of $R(t)$, $\mathbf{Q}(t)$ and the initial values of $\hat{\boldsymbol{\Theta}}_{0,i}(-)$ and $\mathbf{M}_{0,i}(-)$ do not affect the final convergence of the Kalman filtering algorithm (Tavantzis and Ding, 1999).

3.4 Neural/Dynamic (ND) Models

Before we proceed with the formulation of Neural/Dynamic (ND) model, the characteristics of ANNs and the Kalman filtering model should be further emphasized. The ANNs have the ability to adapt to dynamic changes in a traffic environment because of the knowledge experienced from historical data. Moreover, the ANNs can be trained to recognize and ignore the spurious disturbance (noises) under adverse operating conditions, thus preventing serious degradation or even catastrophic failure in performance. However, the implementation of the BP training procedure needs extensive experiments with all training examples. Moreover, the model performance is highly

sensitive to the quality of training data and the choice of ANN structure (e.g., number of neurons on each layer) and learning/momentum rates. Because of such a lengthy and delicate training process, the parameters of ANNs (e.g., synaptic weights) are very difficult to be adjusted on-line in a changing environment, which may mitigate the performance of ANNs in real-time.

Unlike the BP training procedure for ANNs, the Kalman filtering algorithm only needs the previous observation for parameter optimization. Therefore, the model parameters (e.g., Kalman Gain and Jocobian vector) can be updated in real-time quickly, adjusting to dynamic changes in the traffic environment. Nevertheless, such rapid changes in parameters do not always lead to model performance improvement, and sometimes may even result in the very opposite. The reason is that the Kalman filtering model tends to respond to spurious disturbances with its ever-changing parameters, which may cause distinct degradation in model performance in real-time operations.

The difference between the performance of ANNs and the Kalman filtering model is often referred to as the stability-adaptivity dilemma (Hagan, 1996): *parameters of a prediction model should be stable enough to ignore the spurious disturbances while adaptive enough to respond to the dynamic changes in a traffic environment.* To achieve adaptivity and maintain relatively stable performance in predicting vehicle arrival times in real-time, Neural/Dynamic (ND) prediction models are proposed by integrating the capabilities of ANNs and Kalman filtering to

(1) adequately capture the relationship between the vehicle arrival times and various affecting factors with ANNs (either link-based or stop-based), and

(2) dynamically and gently adjust the prediction results from the ANN with the Kalman

filtering procedure thus appropriately adapting to variations in the traffic environment.

In the ND models, the estimated arrival time $E_{k,i}^{(t)}$ of vehicle k at stop i time t can be described as

$$E_{k,i}^{(t)} = p_{k,i-1} + \Theta_{k,i}\Lambda_{k,i}^{(t)} \tag{3.22}$$

where $p_{k,i-1}$, as mentioned previously, is the departure time of vehicle k at stop i-1, while $\Lambda_{k,i}^{(t)}$ is the travel time from stop i-1 to i estimated by the ANN, such that

$$\Lambda_{k,i}^{(t)} = \Phi_L(\mathbf{X}_j(t)) \text{ or} \tag{3.23a}$$

$$\Lambda_{k,i}^{(t)} = \Phi_S(\mathbf{Z}(t)) \tag{3.23b}$$

where $\Phi_L(\cdot)$ and $\Phi_S(\cdot)$ represent the predicted stop-to-stop travel times by the link-based and stop-based ANNs developed in Section 3.2 (refer to Eqs. 3.11 and 3.13), respectively. A Jacobian factor $\Theta_{k,i}$ is introduced into Eq. 3.22 to dynamically adjust $\Lambda_{k,i}^{(t)}$ with the Kalman filtering procedure. When $\Theta_{k,i}$ is equal to 1, Eq. 3.22 is equivalent to one of the aforementioned ANNs.

The Jacobian factor $\Theta_{k,i}$ can be estimated iteratively with the Kalman Gain $K_{k,i}^{(t)}$ (refer to Eq. 3.18), as shown in Eq. 3.24:

$$\hat{\Theta}_{k,i}(+) = \hat{\Theta}_{k,i}(-) + K_{k,i}^{(t)}\rho_{k-1,i} \tag{3.24}$$

where $\rho_{k-1,i}$ represents the prediction error (refer to Eq. 3.17) observed when vehicle k-1 arrived at stop i. In Eq. 3.24, the updated factor $\hat{\Theta}_{k,i}(+)$ should only slightly fluctuate around 1, which can avoid large adjustments in ANN outputs thus guaranteeing the stability of the ND model. The procedure of the proposed ND model for predicting vehicle arrival times are summarized in detail below.

Table 3.3 Summary of the ND model

Prediction Model (I)*	$\Lambda_{k,i}^{(t)} = \Phi_L(\mathbf{X}_j(t))$ or $\Lambda_{k,i}^{(t)} = \Phi_S(\mathbf{Z}(t))$
Prediction Model (II)*	$E_{k,i}(-) = p_{k,i-1}^{(t)} + \Theta_{k,i}(-)\Lambda_{k,i}^{(t)} + v(t) \qquad v(t) \sim N(0,R)$
Jacobian Factor Model	$\Theta_{k,i}(+) = \Theta_{k,i}(-)$
Initialization	$\hat{\Theta}_{0,i}(-) = E[\Theta_{0,i}(-)]$
	$M_{0,i}(-) = E\{[\Theta_{0,i}(-) - \hat{\Theta}_{0,i}(-)][\Theta_{0,i}(-) - \hat{\Theta}_{0,i}(-)]^T\}$
Kalman Gain	$K_{k,i}^{(t)} = \dfrac{M_{k,i}(-)\Lambda_{k,i}^{(t)}}{(\Lambda_{k,i}^{(t)})^2 M_{k,i}(-) + R}$
Jacobian Factor Estimation	$\hat{\Theta}_{k,i}(+) = \hat{\Theta}_{k,i}(-) + K_{k,i}^{(t)}\rho_{k-1,i}$
Prediction Update	$E_{k,i}(+) = p_{k,i-1}^{(t)} + \hat{\Theta}_{k,i}(+)\Lambda_{k,i}^{(t)}$
Error Covariance Update	$M_{k+1,i}(-) = [1 - K_{k,i}^{(t)}\Lambda_{k,i}^{(t)}]\, M_{k,i}(-)$
Jacobian Factor Update	$\hat{\Theta}_{k+1,i}(-) = \hat{\Theta}_{k,i}(+)$

*Note: Prediction Model I: Prediction of stop-to-stop travel time with ANNs
Prediction Model II: Prediction of vehicle arrival time with the ND model

Convergence of the ND Model

The performance (convergence speed) of the ND models provides the foundation for choosing model parameters, such as the noise covariance R in Prediction Model II as listed in Table 3.3. In the table, the Kalman Gain $K_{k,i}^{(t)}$ and the covariance of prediction error $M_{k+1,i}(-)$ for the ND model are obtained by

$$K_{k,i}^{(t)} = \frac{M_{k,i}(-)\Lambda_{k,i}^{(t)}}{(\Lambda_{k,i}^{(t)})^2 M_{k,i}(-) + R} \tag{3.25}$$

$$M_{k+1,i}(-) = [1 - K_{k,i}^{(t)}\Lambda_{k,i}^{(t)}]\, M_{k,i}(-) \tag{3.26}$$

The two equations can be solved iteratively, starting with $M_{0,i}(+) = M_{0,i}$. Therefore, for vehicle 1, Kalman Gain $K_{1,i}$ and the error covariance $M_{2,i}(-)$ can be obtained as

$$K_{1,i} = \frac{M_{0,i}\Lambda_{0,i}}{\Lambda_{0,i}^2 M_{0,i} + R} \tag{3.27}$$

$$M_{2,i}(-) = \frac{R\,M_{0,i}}{\Lambda_{0,i}^2 M_{0,i} + R} \tag{3.28}$$

For vehicle 2, $K_{2,i}$ and $M_{3,i}(-)$ can be similarly derived as

$$K_{2,i} = \frac{M_{0,i}\Lambda_{1,i}}{(\Lambda_{0,i}^2 + \Lambda_{1,i}^2)M_{0,i} + R} \tag{3.29}$$

$$M_{3,i}(-) = \frac{R\,M_{0,i}}{(\Lambda_{0,i}^2 + \Lambda_{1,i}^2)M_{0,i} + R} \tag{3.30}$$

Hence, for vehicle k, $K_{k,i}$ and $M_{k+1,i}(-)$ can be derived as

$$K_{k,i} = \frac{M_{0,i}\Lambda_{k-1,i}}{(\Lambda_{0,i}^2 + \cdots + \Lambda_{k-1,i}^2)M_{0,i} + R} \tag{3.31}$$

$$M_{k+1,i}(-) = \frac{R\,M_{0,i}}{(\Lambda_{0,i}^2 + \cdots + \Lambda_{k,i}^2)M_{0,i} + R} \tag{3.32}$$

Eq. 3.32 shows that, for a sufficiently large value of k, the covariance of the prediction error $M_{k+1,i}(-)$ converges to zero (Tavantzis and Ding, 1999).

As mentioned above, the prediction output from ANNs should only be updated slightly in order to ensure the stability of the ND model. Thus, the Jacobian factor $\hat{\Theta}_{k,i}(+)$ in the ND model should not fluctuate dramatically. Since the Kalman Gain $K_{k,i}^{(l)}$ scales the amount of adjustment in $\hat{\Theta}_{k,i}(+)$ (refer to Eq. 3.24), a small value of $K_{k,i}^{(l)}$ that

can avoid large changes in $\hat{\Theta}_{k,i}(+)$ is required. Since Eq. 3.31 denotes that a larger noise covariance R results in smaller $K_{k,i}^{(l)}$, large values of R are preferred by the ND model. However, R may influence the convergence speed of the ND model, as shown in Eq. 3.32, different values of R are evaluated by trial and error before choosing the proper ones (Tavantzis and Ding, 1999).

The configuration of the ND model is shown in Figure 3.9, which implements the following steps for on-line estimating/updating model parameters while predicting vehicle arrival times.

Step 1. Initialize Jacobian vector $\hat{\Theta}_{0,i}(-)$ and the error covariance $M_{0,i}(-)$ for any stop i.

Step 2. At time t when vehicle k is ready to depart from stop i-1, predict travel time $\Lambda_{k,i}^{(l)}$ from stop i-1 to i with either link-based or stop-based ANN.

Step 3. Calculate Kalman Gain $K_{k,i}^{(l)}$ based on predicted travel time $\Lambda_{k,i}^{(l)}$ and the error covariance $M_{k,i}(-)$.

Step 4. Estimate the Jacobian factor $\hat{\Theta}_{k,i}(+)$ using the Kalman Gain $K_{k,i}^{(l)}$.

Step 5. Predict vehicle arrival time $E_{k,i}(+)$ by on-line adjusting the stop-to-stop travel time $\Lambda_{k,i}^{(l)}$ predicted by the ANN with Jacobian vector $\hat{\Theta}_{k,i}(+)$.

Step 6. Calculate error covariane $M_{k+1,i}(-)$ and Jacobian vector $\hat{\Theta}_{k+1,i}(-)$ for next vehicle k+1.

Step 7. If stop i is not the last stop on the route, loop over to Step 2 for predicting arrival times at other downstream stops.

Step 8. Update the prediction error $\rho_{k,i}$ when vehicle k arrives stop i.

Summary

In this study, both link-based and stop-based ANNs are integrated with the Kalman filtering model. The link-based and stop-based ND models (e.g., NDL and NDS, respectively) will be evaluated through simulating a real transit route in Chapter 6.

The developed ND models are designed for predicting bus arrival times. They can be applied to other transit modes, such as rail transit systems (e.g., light and rapid rail transit). With exclusive or nearly exclusive (but not necessarily grade-separated) right-of-way (ROW), particularly in congested areas (e.g., in central city or on urban arterials), the operation disturbance of rail transit lines from automobile traffic is relatively small. The ND models can be trained in a similar way as discussed above for a bus system. Thus, the arrival times for rail transit vehicles can be predicted.

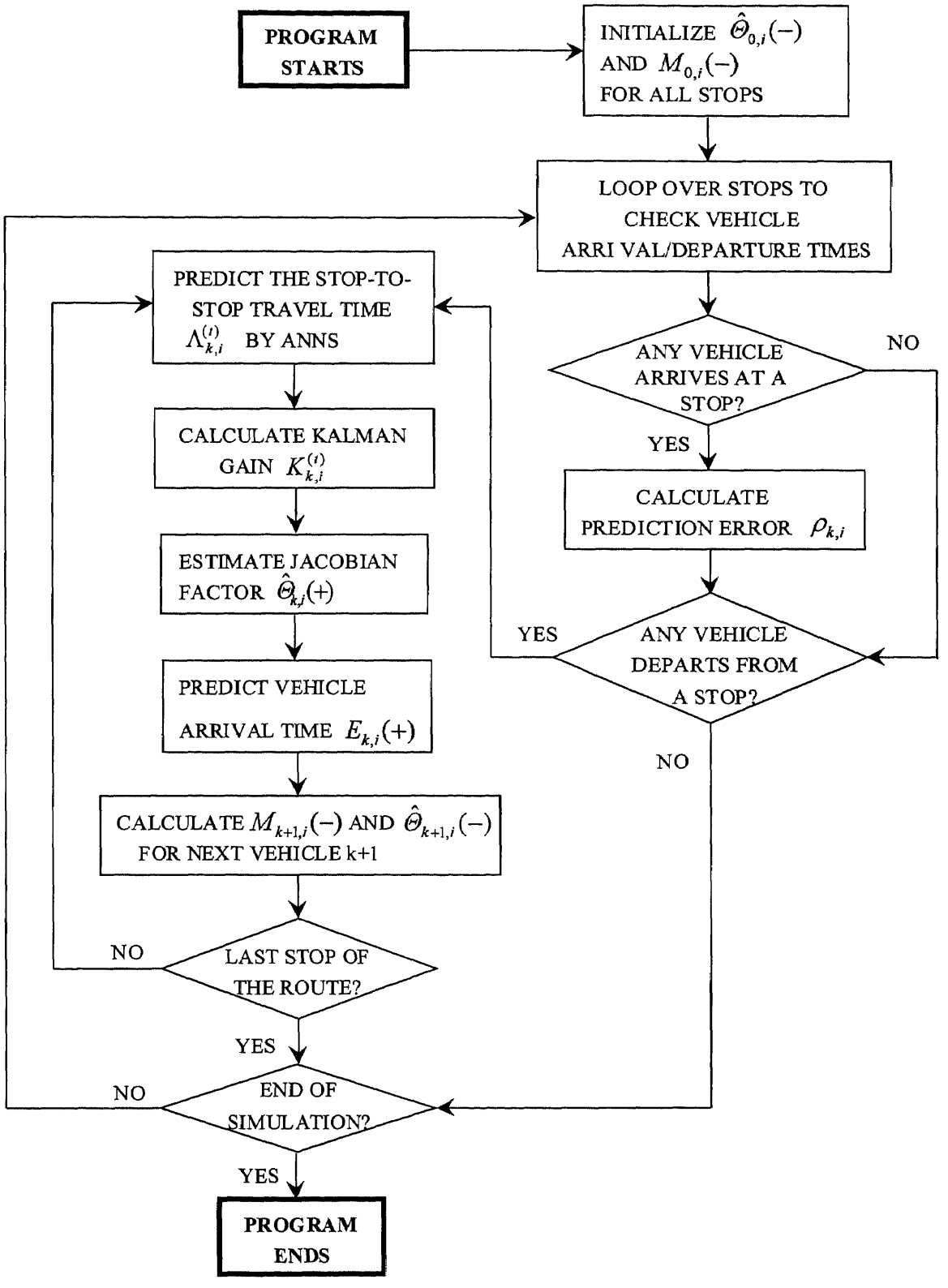


Figure 3.9 Configuration of the ND model

CHAPTER 4

APPLICATIONS OF THE PREDICTION MODELS

In the advent of Advanced Public Transportation Systems (APTS), the developed Neural/Dynamic(ND) models can be implemented in Traveler Information Systems (TIS) for disseminating the accurately predicted vehicle arrival times to passengers. The proposed link-based and stop-based ND models (NDL and NDS, respectively) are also helpful for advanced transit control systems. With reliable predicted information, various control strategies (e.g., dynamic dispatching, signal priority treatment and operational control) may be initiated timely for alleviating the severity of the headway/schedule disruptions on both transit users and operators.

This chapter addresses an application of the developed ND models to a dynamic transit operation system. The potential impacts on passenger wait times and headway variations are evaluated through the simulation program discussed in Chapter 5.

4.1 Introduction

Transit operations are frequently disturbed by many factors (e.g., random delays at signalized intersections, passenger demand fluctuations, traffic congestion, and unexpected events such as incidents and construction maintenance activities on the roadways). The joint impact of these factors on vehicle operations increases the difficulty of maitaining transit headway adherence. Riders, having longer average wait times than their expectation, will be discouraged for using the transit system. Advanced transit control strategies can modify transit operations in real-time to maintain headway

adherence, and may reduce the average wait time. The measures of effectiveness (MOEs) used for evaluating transit control strategies include stochastic variations of system performance (e.g., headway/schedule variances) and average passenger wait times (Koffman, 1981; Lin, et. al., 1995).

The operation control strategies for rail transit systems can be classified into two categories: (1) manual vehicle control and (2) automatic vehicle control (Vuchic, 1981). In the first category, drivers usually operate vehicles with the assistance of signal systems (e.g., track-based train protection systems) to keep safe spacings between pairs of successive vehicles. In the second category, operating vehicles are instructed by a centralized computer system, while drivers are only functionalized for starting automatic driving process or dealing with breakdowns.

In manual control, if a pair of consecutive vehicles running too close is detected, the follower vehicle will slow down or stop manually depending on the color of the signal observed by the driver. Therefore, a safe separation between vehicles can be guaranteed. The manual control is considered effective for light demand routes, especially when the operating speeds are lower than 45 mph (Vuchic, 1981). For high frequency routes with heavy demand, the manual control system may reduce the average vehicle speed as well as energy consumption due to frequent stoppings.

Unlike a manual control system, an automatic control system requires a centralized computer system to guide operating vehicles after processing real-time information (e.g., vehicle locations, speeds and delays) obtained from transit monitoring and communication systems (e.g., AVLS). Such a control system can determine the optimal speeds and accelerate/decelerate rates for each departure vehicle, based on the

physical and operational (e.g., length, weight and braking ability and roadway conditions such as curvature, grade percentage and superelevation) characteristics of the vehicle. The control instructions will be transmitted to each vehicle for guiding its movements, which can assist transit operators to considerably increase service frequencies without degrading safety (Vuchic, 1981; Railway Technical Research Institute, Tokyo, 1998).

Previous studies discussed various advanced control models for holding or speeding vehicles in different situations to keep uniform headways or pre-planned schedules. Several holding control models (Osuna and Newell, 1972; Turnquist and Blume, 1980; Abkowitz, et. al., 1984 and 1986) were developed to deliberately delay a vehicle ahead of the schedule when it was on the route (e.g., reduce its operating speed) or at selected stops (e.g., defer its departure time). The results showed that the implementation of holding control may cause additional delays for on-board passengers and increase vehicle travel times. Speeding control models (Koffman, 1978; Lin, et. al., 1995; Eberlein, 1998) were developed to reduce vehicle travel times by increasing operating speeds on the route or skipping stops. These models were applicable when a vehicle fell behind schedule or had a long headway from its preceding vehicle. Among various speeding control models, stop-skipping was often used to improve vehicle travel speed and saved passenger wait times at downstream stops (Turnquist, et. al., 1980 and 1981). However, it sacrificed the wait time for passengers at the skipped stops as well as frustrated the in-vehicle passengers who were destined for the skipped stops (Turnquist, 1981; Lin, et. al., 1995; Eberlein, et. al., 1998).

In the advent of APTS, great progress has been made toward developing or implementing advanced control for transit systems (Federal Transit Administration,

Update'98). The undergoing projects include Communications-based Train Control (CBTC) for the New York City subway system and Advanced Automatic Train Control (AATC) for the San Francisco Bay Area Rapid Transit (BART) District. CBTC is designed to control vehicles through a wireless communication system, which can send the instructions from the control center to the local computers installed in vehicles. CBTC aims at replacing the signaling equipment (e.g., track circuits) currently used by the system with more reliable computer-based signal systems. AATC, developed with moving block signal control technologies, is expected to provide BART with the ability to operate twice as many vehicles on their existing tracks. With AATC, higher capacity can be achieved with desirable operating speed and safety.

In this chapter, a real-time headway control model is developed to maintain desired headways between any pair of consecutive vehicles for high frequency light rail transit (LRT) systems. Note that the headway discussed in this section refers to the inter-departure time (time difference between the departure times of consecutive vehicles from a stop). The model focuses on adjusting vehicle departure times in real-time based on its optimal arrival time at the next stop, while considering the constraint of the maximum attainable operating speed and the headways to its leading and following vehicles.

4.2 Assumptions

To develop a real-time control model for minimizing the total headway variance at all stops, the following assumptions are made:

1. Passenger arrivals at stops follow Poisson distributions. This assumption has been used in various studies (Osuna and Newell, 1972; Turquist and Blume, 1980; Lin, et. al., 1995; Chien and Schonfeld, 1997; Eberlein, et. al., 1998) and is reasonable when vehicle operating frequency on the transit route is high.
2. The variation of passenger arrival rates and O/D demand distribution over stops and time are given or predictable. Therefore, the number of boarding and alighting passengers for a vehicle operating on the route can be calculated once the headway between the vehicle and its preceding vehicle is known (Chien and Ding, 1998).
3. Vehicle overtaking is not allowed, which exists in most rail transit systems.
4. The control model handles all operating vehicles, while the control instructions to vehicles are given only before departing from stops. Therefore, every vehicle is assumed to operate with its full speed.
5. Only one vehicle receives control instructions at a time. For the situation that more than one vehicle are departing from stops simultaneously (although it rarely occurs), the leading vehicle will get the control instruction first, and then the following ones will be instructed in sequence.
6. The vehicle arrival times at downstream stops can be actually predicted by prediction models developed in Sections 3.2, 3.3 and 3.4 in this research. The accuracy of the predicted arrival times is very important to estimate headway variances at stops.

Given the assumptions above, the real-time vehicle control model is formulated in the following sections.

4.3 Model Formulation

In 1957, Welding developed a model to estimate the average passenger wait time for a transit route while considering stochastic (random) vehicle/passenger arrivals at stops.

$$E(W) = E(h)/2 + \sigma^2(h)/2E(h) \tag{4.1}$$

The average wait time, $E(W)$, can be estimated if the mean headway $E(h)$ and the headway variance $\sigma^2(h)$ are known. In Eq. 4.1, $E(h)$ can be treated as the scheduled headway H (Adebisi, 1986). The average passenger wait time can thus be reduced by minimizing the headway variance $\sigma^2(h)$. The formulation of the total headway variance for a transit route is discussed below.

Headway Variance Function

Assume that a transit route (Figure 4.1) has S stops where passenger arrivals follow Poisson distributions. If there are N vehicles, numbered from 1 to N (from downstream to upstream), having been dispatched onto the route, vehicle k is ready to depart from stop i-1, whose next stop is named i.

At time t, the headway $h_{k,i}^{(t)}$ between vehicles k-1 and k at stop i (i=1, 2, ...,S) can be determined by

$$h_{k,i}^{(t)} = p_{k,i} - p_{k-1,i} \qquad \text{for k} = 2, 3, \ldots, \text{N} \tag{4.2}$$

where $p_{k-1,i}$ and $p_{k,i}$ represent the departure times of vehicles k-1 and k at stop i, respectively. Note that at time t, $p_{k,i}$ is either known if stop i has been served by vehicle k (e.g., $p_{k,i-1} \le t$, $\forall k,i$) or otherwise ($p_{k,i-1} > t$, $\forall k,i$) can be estimated by Eq. 4.3.

Legend: vehicle
station
rail transit route

Figure 4.1 A transit route with the control model

$$p_{k,i} = a_{k,i} + d_{k,i} \qquad \text{for k} = 2, 3, \ldots, \text{N and i} = 1, 2, \ldots, \text{S} \qquad (4.3)$$

where $a_{k,i}$ represents the arrival time of vehicle k to stop i (either known or predictable), and $d_{k,i}$ is the dwell time of the vehicle at stop i. Note that $d_{k,i}$ is equal to the number of boarding passengers multiplied by the average boarding time t_b,

$$d_{k,i} = t_b \lambda_i (a_{k,i} - p_{k-1,i}) \qquad \text{for k} = 2, 3, \ldots, \text{N and i} = 1, 2, \ldots, \text{S} \qquad (4.4)$$

where λ_i is the hourly passenger arrivals at stop i, while the term $(a_{k,i} - p_{k-1,i})$ denotes the headway between the previous vehicle departure and the present vehicle arrival times.

Based on the headways between N-1 pairs of operating vehicles at time t, the headway variance $\pi_i^{(t)}$ observed at stop i can be formulated as

$$\pi_i^{(t)} = \frac{\sum_{k=2}^{N} (h_{k,i}^{(t)} - H)^2}{N-1} \qquad \text{for i} = 1, 2, \ldots, \text{S} \qquad (4.5)$$

To formulate the objective total headway variance function, the impact of passenger demand at a stop on the headway variance should be considered. For example, at a stop with high passenger demand, the headway variance will result in longer average

passenger wait and vehicle dwell times than that at low demand stops (refer to Eq. 4.1). Such impact can be reflected in the objective function by imposing a higher weight factor for the stop, which can be determined by Eq. 4.6.

$$w_i = \frac{\lambda_i}{\sum_{i=1}^{S} \lambda_i} \tag{4.6}$$

If the demand at a stop increases, its weight factor w_i in Eq. 4.6, is proportionally increased. Thus, the objective function $\Pi^{(t)}$ can be formulated as a weighted sum of headway variances for all stops, as shown in Eq. 4.7.

$$\Pi^{(t)} = \sum_{i=1}^{S} w_i \pi_i^{(t)} \tag{4.7}$$

By substituting Eq. 4.5 into Eq. 4.7, the objective total headway variance function can be derived as

$$\Pi^{(t)} = \sum_{i=1}^{S} w_i \frac{\sum_{k=2}^{N} (h_{k,i}^{(t)} - H)^2}{N-1} \tag{4.8}$$

By minimizing the objective total headway variance function $\Pi^{(t)}$ whenever a vehicle departs from a stop, the headways can be regulated and average passenger wait time can thus be reduced.

All equations developed to estimate vehicle arrival and departure times at stop i are summarized in Table 4.1, in which vehicle k is assumed to depart from stop i-1. Note that the arrival time $a_{k,i}$ and the departure time $p_{k,i}$ (k=1, 2, ..., N) in every equation listed in Table 4.1 are either known or predictable. At time t, the optimal arrival time for vehicle k at stop i, denoted by $e_{k,i}^{(t)}$, is a decision variable to be determined by the

proposed vehicle headway control model. By substituting departure times $p_{k,i}$ (k=1,..., N) formulated in Table 4.1 into Eqs. 4.2 and 4.5, the headway variance at stop i can be obtained. Hence, the total headway variances for all stops can be determined by Eq. 4.8.

Table 4.1 Vehicle arrival and departure times at stop i

Vehicle No.	Arrival Time	Departure Time
1	$a_{1,i}$	$p_{1,i}$
...	...	...
k-1	$a_{k-1,i}$	$p_{k-1,i}=a_{k-1,i}+t_b\lambda_i(a_{k-1,i}-p_{k-2,i})$
k	$e_{k,i}^{(l)}$	$p_{k,i}=e_{k,i}^{(l)}+t_b\lambda_i(e_{k,i}^{(l)}-p_{k-1,i})$
k+1	$a_{k+1,i}$	$p_{k+1,i}=a_{k+1,i}+t_b\lambda_i(a_{k+1,i}-p_{k,i})$
...	...	...
N	$a_{N,i}$	$p_{N,i}=a_{N,i}+t_b\lambda_i(a_{N,i}-p_{N-1,i})$

Constraints

Since the vehicle operating speed is always bounded (Vuchic, 1981) and less than the maximum operating speed of 60 mph for most LRT systems, the optimal vehicle arrival time $e_{k,i}^{(l)}$ is anticipated to be greater than or equal to its earliest arrival time. Thus, the arrival time constraints can be formulated as

$$L_{k,i}^{(l)} \le e_{k,i}^{(l)} \qquad \text{for k = 1, 2, ...,N and i = 1, 2,..., S} \qquad (4.9)$$

where $L_{k,i}^{(l)}$ represents the earliest arrival time at stop i that vehicle k can make, which is the vehicle departure time at stop i-1 plus the minimum travel time from stop i-1 to stop i.

Real-time Headway Control Model

With the developed objective total headway variance function subject to the constraints, the proposed real-time headway control model can be formulated as follows:

$$\underset{e_{k,i}^{(l)}}{Min} \sum_{j=1}^{S} w_j \frac{\sum_{k=2}^{N} (h_{k,j}^{(l)} - H)^2}{N-1} \tag{4.10}$$

Subject to: $L_{k,i}^{(l)} \le e_{k,i}^{(l)}$ for k = 1, 2, ...,N and i = 1, 2, ..., S

From the previous discussion, we found that headway $h_{k,j}^{(l)}$ in Eq. 4.10 is correlated to headway $h_{k-1,j}^{(l)}$ and dependent on the arrival time $e_{k,i}^{(l)}$. Thus, the headway control problem formulated here is a constrained non-linear optimization problem. By minimizing the total headway variance whenever a vehicle departs from stop i-1, the optimal arrival time $e_{k,i}^{(l)}$ to the downstream stop i can be obtained, which is discussed next.

4.4 Optimization

The first and the second derivatives of the total headway variance function are derived in order to check the convexity of the objective function $\Pi^{(l)}$ and obtain the optimal solution. The first derivative of $\Pi^{(l)}$ with respect to the arrival time $e_{k,i}^{(l)}$ can be derived from Eq. 4.6 as

$$\frac{\partial \Pi^{(l)}}{\partial e_{k,i}^{(l)}} = \sum_{j=1}^{S} w_j \frac{\partial \pi_j^{(l)}}{\partial e_{k,i}^{(l)}} \qquad \text{for k = 1, 2,...,N and i = 1, 2,..., S} \tag{4.11}$$

Given a constant passenger arrival rate λ_j at stop j, $\frac{\partial \pi_j^{(l)}}{\partial e_{k,i}^{(l)}}$ can be derived below:

Situation 1

At time t, if vehicle k is ready to depart from stop i-1, its departure times from stops 1 through i-1 are known. Therefore, the first derivatives of headway variance at these stops are independent of the optimal arrival time $e_{k,i}^{(l)}$, such that

$$\frac{\partial \pi_j^{(l)}}{\partial e_{k,i}^{(l)}} = 0 \qquad \text{for j = 1, 2, ..., i-1} \tag{4.12}$$

Situation 2

The first derivative of headway variance at the next stop i with respect to $e_{k,i}^{(l)}$ can be derived by using the departure times of the N vehicles at the stop (refer to Table 4.1), such that

$$\frac{\partial \pi_j^{(l)}}{\partial e_{k,i}^{(l)}} = \frac{2}{N-1}\left\{\Delta_{k,i}\varphi_i - \Delta_{k+1,i}\varphi_i^2 + \sum_{n=k+2}^{N}[(-1)^{n-1}(t_b\lambda_i)^{n-k-1}\,\Delta_{n,i}\varphi_i^2]\right\} \quad \text{for j = i} \tag{4.13}$$

In Eq. 4.13, φ_j is a constant factor depending on passenger arrival rate λ_j, such that

$$\varphi_j = 1 + t_b\lambda_j \qquad \text{for j=1, 2, ...,S} \tag{4.14}$$

The deviation between actual headways $h_{k,j}^{(l)}$ and the scheduled headway H is represented by $\Delta_{k,j}$,

$$\Delta_{k,j} = h_{k,j}^{(l)} - H \qquad \text{for k = 1, 2,...,N and j = 1, 2, ..., S} \tag{4.15}$$

Situation 3

The headway variances at stops from i+1 to S are dependent on the optimal arrival times (e.g., $e_{k,i+1}^{(l)}$, ..., $e_{k,S}^{(l)}$, respectively) which can be predicted based on $e_{k,i}^{(l)}$ at stop i (refer to Eqs. 3.11 and 3.12). Therefore, the first derivative of headway variance at these stops with respect to $e_{k,i}^{(l)}$ can be derived as

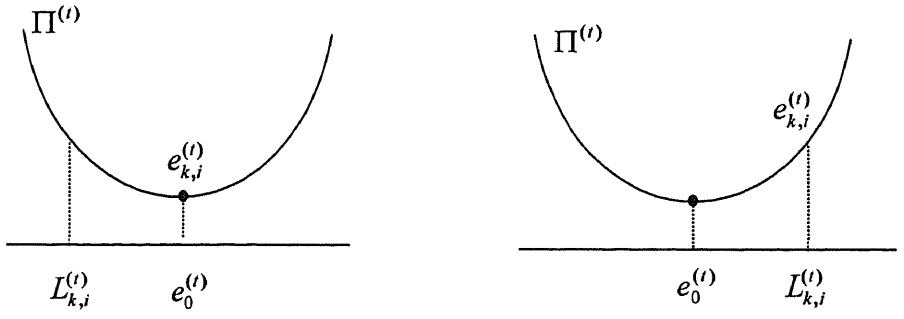
$$\frac{\partial \pi_j^{(l)}}{\partial e_{k,i}^{(l)}} = \frac{2\varphi_i \varphi_{i+1} \cdots \varphi_{j-1}}{N-1} \left\{ \Delta_{k,j}\varphi_j - \Delta_{k+1,j}\varphi_j^2 + \sum_{n=k+2}^{N} [(-1)^{n-1}(t_b \lambda_j)^{n-k-1} \Delta_{n,j}\varphi_j^2] \right\}$$

$$\text{for j} = \text{i+1,...,S} \qquad (4.16)$$

where φ_j and $\Delta_{k,j}$ can be obtained from Eqs. 14 and 15, respectively. Eq. 4.16 denotes that the passenger demands at upstream stops i+1, i+2, ..., and j-1 contribute to the headway variance $\pi_j^{(l)}$ at stop j. This implies that the impact of passenger demand at a stop on headway variance will be propagated at further downstream stops. It is one of the primary reasons that transit vehicles bunch up as they move along the transit route.

The first derivative of the objective function with respect to $e_{k,i}^{(l)}$ can be obtained by the summation of Eqs. 4.12, 4.13, and 4.16. In addition, the second derivative of the total headway variance can be formulated and found to be positive (Appendix A). Thus, the objective function $\Pi^{(l)}$ is convex, and the minimum $e_0^{(l)}$ can be found by simply using a line search algorithm (e.g., golden section or bi-section).

Considering the linear constraints formulated in Eq. 4.9, the following situations are specified to search for the optimal solution $e_{k,i}^{(l)}$ of the real-time headway control model, as shown in Figure 4.2.



(Situation 1) (Situation 2)

Figure 4.2 The optimal solution of the headway control model

Situation 1

At time t, if the minimum $e_0^{(t)}$ is greater than the earliest arrival time $L_{k,i}^{(t)}$ that vehicle k can make. This makes $e_0^{(t)}$ feasible and the optimal solution $e_{k,i}^{(t)}$ is $e_0^{(t)}$, such that

$$\text{If } e_0^{(t)} > L_{k,i}^{(t)}, \text{ then } e_{k,i}^{(t)} = e_0^{(t)} \qquad \text{for k = 1, 2, ...,N and j = 1, 2, ..., S} \quad (4.17a)$$

Situation 2

when $e_0^{(t)}$ is less than $L_{k,i}^{(t)}$, then $e_0^{(t)}$ is not feasible and the optimal solution $e_{k,i}^{(t)}$ should be $L_{k,i}^{(t)}$, denoting that vehicle k will be controlled to depart from stop i immediately at the maximum operating speed,

$$\text{If } e_0^{(t)} < L_{k,i}^{(t)}, \text{ then } e_{k,i}^{(t)} = L_{k,i}^{(t)} \qquad \text{for k = 1, 2, ...,N and j = 1, 2, ..., S} \quad (4.17b)$$

4.5 Model Configuration

The configuration of the headway control model is shown in Figure 4.3. Whenever a vehicle finishes serving a stop, the developed real-time headway control model determines whether the vehicle's departure time should be postponed based on the optimal arrival time to the next downstream stop. According to the vehicle operational characteristics (e.g., accelerate/decelerate rates and the maximum attainable speed), the vehicle can be instructed to arrive at its next stop at the optimal time. Such control is more suitable for transit routes with substantiate exclusive right-of-way (ROW). For transit vehicles operating with mixed automobile traffic such as street bus transit, the difficulty of headway control increases because of ROW competition between transit vehicles and other traffic.

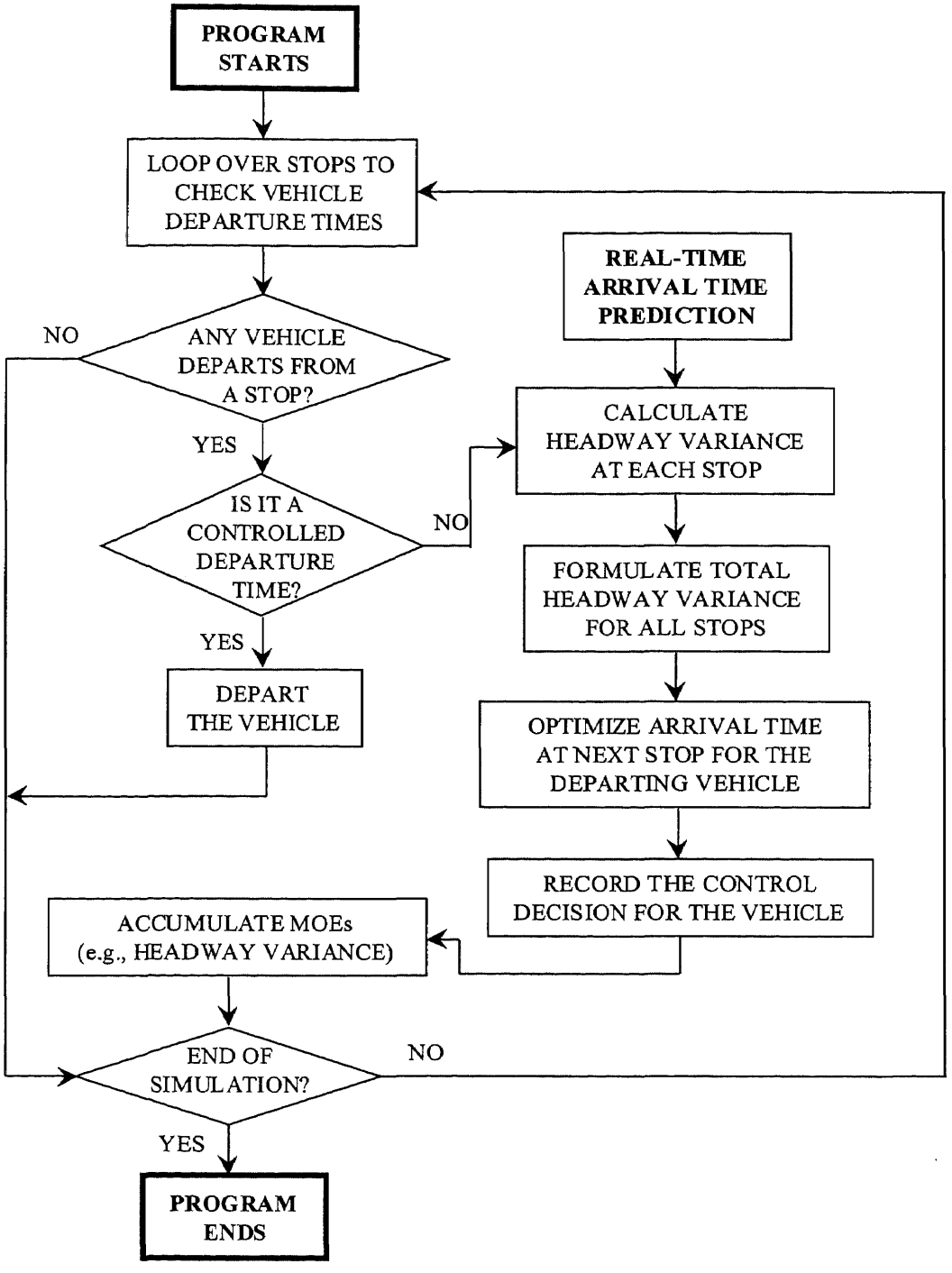


Figure 4.3 Configuration of real-time vehicle control model

CHAPTER 5

MICROSCOPIC TRANSIT SIMULATION PRPGRAM

A program is developed through enhancing CORSIM (CORridor SIMulator) for simulating transit operations. New features are built into CORSIM to appropriately determine the vehicle dwell time at a stop according to the number of boarding/alighting passengers. The number of boarding/alighting passengers is dependent on the transit O/D demand between stops and the vehicle inter-departure time at the stop. In addition, various emulated real-time data (e.g., vehicle arrival/departure times, vehicle dwell times and number of boarding/alighting passengers at stops) and transit related MOEs (e.g., average passenger wait times at stops and vehicle journey time) can be generated, collected and analyzed by the proposed simulation program. Thus, the program is also applicable for simulating a transit monitoring and communications environment (e.g., Automatic Vehicle Location Systems (AVLS) and Automatic Passenger Counting Systems (APCS)) as well as evaluating innovative strategies, such as the developed Neural/Dynamic (ND) prediction models and the application of ND models to a real-time vehicle control.

5.1 Introduction

CORSIM, a time-driven, microscopic and stochastic traffic simulation model, can simulate traffic operations on urban corridor networks containing freeways and surface streets. It has been applied extensively to a wide variety of areas by both practitioners and researchers and is perhaps the most widely used traffic simulation model. The original

CORSIM has the feature to simulate transit vehicle operations. However, it deals with vehicle dwell times purely relying on the user-specified mean dwell time at each stop and embedded distributions. Because of neglecting the impact of boarding and alighting passengers when determining the duration of vehicle dwell times, simulation results can not reflect realistic vehicle operations at stops. In this chapter, the proposed simulation program, enhanced from CORSIM, is able to determine time dependent vehicle dwell times and passenger wait times based on real-time information, such as passenger arrivals at stops and headways between vehicles.

5.2 Assumptions

The simulation program is developed in this chapter to analyze transit (i.e., buses) operations. Several assumptions are made while developing the program:

1. Passenger arrivals at stops follow Poisson distributions. Other passenger arrival distributions can be generated if a user specified arrival distribution is embedded into the program.
2. The average passenger boarding/alighting time is currently assumed to be 2 seconds. This assumption can be relaxed by allowing users to specify average boarding/alighting time for different types of stops and configurations of vehicles (e.g., the number, location and size of doors).

5.3 Features of the Simulation Program

A CORSIM based microscopic simulation program is developed to mimic conventional bus transit operations in urban corridors. The simulation program is a time-driven simulation program, in which individual vehicle movements can be simulated every second during different time periods. Therefore, users can evaluate traffic operations under a time-varied traffic environment. The program can simulate vehicle overtaking and merging maneuvers, passenger arrival distributions and the interactions of transit vehicles on other vehicles competing for the right-of-way. In addition, the program has the capabilities to simulate number of lanes and turn pockets, and a wide range of geometric and traffic flow conditions. Special information may be added into the analyzed networks including incidents and temporal events (e.g., parking activities).

In the proposed simulation program, new features are developed for capturing the fact that both the number of boarding/alighting passengers and the duration of the vehicle dwell time are dependent on the vehicle inter-departure time. Seven modules are developed and embedded with the program:

Module 1. To determine boarding/alighting rates at stops

Module 2. To dispatch buses based upon user specified transit schedule

Module 3. To calculate number of boarding/alighting passengers at stops

Module 4. To estimate bus dwell times at stops

Module 5. To calculate passenger wait times

Module 6. To predict bus arrival times by a proposed prediction model

Module 7. To generate bus operational output for the estimation of MOEs

The system configuration of the simulation program is shown in Figure 5.1. Transit and other vehicles (e.g., trucks, and automobiles) operating in networks constituted by freeways and surface streets can be simulated. Transit vehicles are dispatched onto a network according to user-specified time table (the posted schedule). They move along designated routes and serve passengers at stops, while the real-time data (e.g., vehicle arrival/departure times, vehicle dwell times, and passenger wait times) that affect transit operation performance are generated.

Equations are formulated next for determining the number of boarding/alighting passengers, the durations of vehicle dwell and passenger wait times.

Numbers of Boarding/Alighting Passengers

The numbers of boarding and alighting passengers at each stop are determined by the origin-destination (O/D) matrix of the analyzed transit route. Two matrices are established for calculating boarding and alighting passengers, including a stop O/D and a vehicle O/D matrix. The stop O/D matrix is exactly the same as the user specified O/D demand in the input file, while the vehicle O/D matrix represents the number of in-vehicle passengers destined at each downstream stop.

The stop O/D matrix of the transit route serving S stops, denoted by $\mathbf{O}$, is shown below:

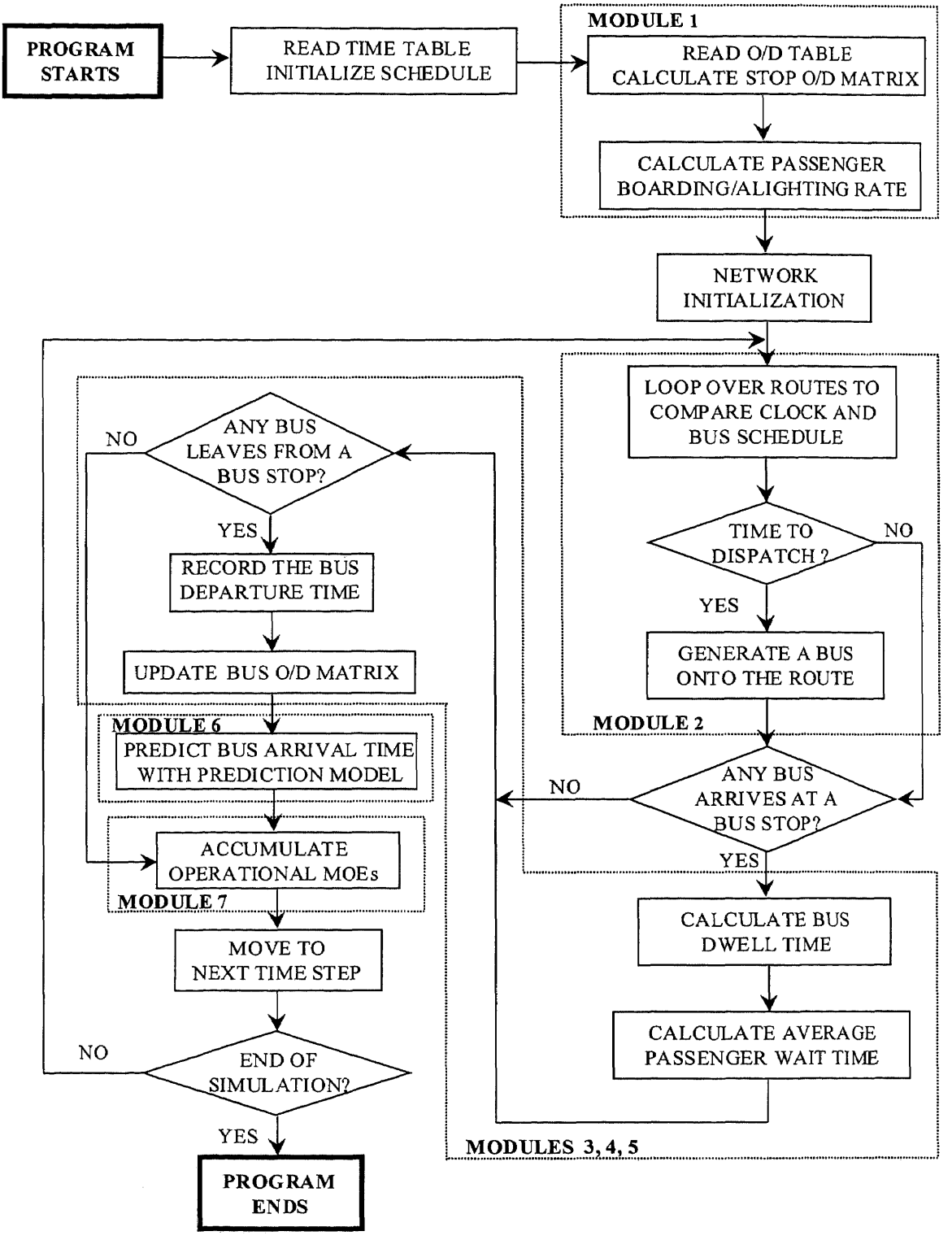


Figure 5.1 Configuration of the transit simulation program

$$\mathbf{O} = \begin{bmatrix} \overline{q}_{1,1} & \overline{q}_{1,2} & \overline{q}_{1,3} & \overline{q}_{1,4} & \cdots & \overline{q}_{1,S} \\ & \overline{q}_{2,2} & \overline{q}_{2,3} & \overline{q}_{2,4} & \cdots & \overline{q}_{2,S} \\ & & \overline{q}_{3,3} & \overline{q}_{3,4} & \cdots & \overline{q}_{3,S} \\ & & & \ddots & \cdots & \vdots \\ & & & & \overline{q}_{S-1,S-1} & \overline{q}_{S-1,S} \\ & & & & & \overline{q}_{SS} \end{bmatrix} \tag{5.1}$$

where $\overline{q}_{i,j}$ represents hourly demand originating from stop i (i=1, 2, ...,S) and destined at stop j (j = i, i+1, ...,S). If i is equal to j, $\overline{q}_{i,j}$ is zero (no passenger originating from and destined for the same stop).

The boarding rate $\overline{q}_{ib}$ denotes the hourly number of passengers boarding at stop i and destined for all downstream stops, while the alighting rate $\overline{q}_{aj}$ denoting the hourly number of passengers alighting at stop j and originating from all upstream stops. $\overline{q}_{ib}$ and $\overline{q}_{aj}$ are formulated in Eqs 5.2 and 5.3, respectively.

$$\overline{q}_{ib} = \sum_{j=i+1}^{S} \overline{q}_{i,j} \qquad \text{for i = 1, 2, ..., S-1} \tag{5.2}$$

$$\overline{q}_{aj} = \sum_{i=1}^{j-1} \overline{q}_{i,j} \qquad \text{for j = 2, 3, ...,S} \tag{5.3}$$

Once a vehicle is dispatched onto the analyzed transit route, an ID number is assigned to the vehicle to identify every vehicle in the output file. After the vehicle picks up all boarding passengers at a stop, the number of in-vehicle passengers traveling to the downstream stops may be changed. Thus, the vehicle O/D matrix should be updated once the vehicle departs from the stop. The vehicle O/D matrix $\mathbf{B}_{k,j}$, representing demand on vehicle k traveling from stop i to i+1, contains in-vehicle passengers destined for different downstream stops:

$$\mathbf{B}_{k,i} = \left(Q_{i,i+1},\ Q_{i,i+2},\ \cdots,\ Q_{i,S}\right) \tag{5.4}$$

where $Q_{i,j}$ (i = 1, 2, ..., S-1; j = i+1,..., S) is the number of in-vehicle passengers destined for downstream stop j. When vehicle k arrives at stop i+1, the number of passengers $\Delta Q_{i+1,j}$ boarding at stop i+1 and destined for stop j is

$$\Delta Q_{i+1,j} = \overline{q}_{i+1,j} \cdot (p_{k,i+1} - p_{k-1,i+1}) \qquad \text{for i = 1, 2, ..., S-2 and j=i+2,..., S} \tag{5.5}$$

where $\overline{q}_{i+1,j}$ is the passenger boarding rate from stop i+1 to j. In Eq. 5.5, the vehicle headway at stop i+1 is represented by $(p_{k,i+1} - p_{k-1,i+1})$, in which $p_{k-1,i+1}$ and $p_{k,i+1}$ represent the previous and present vehicle departure times at stop i+1, respectively.

After serving all passengers at stop i+1, the vehicle O/D matrix is updated as

$$\mathbf{B}_{k,i+1} = \left(Q_{i+1,i+2},\ Q_{i+1,i+3},\ \cdots,\ Q_{i+1,S}\right) \tag{5.6}$$

where $Q_{i+1,j}$ (i = 1, 2,..., S-2; j=i+2,..., S) is the number of in-vehicle passengers destined for stop j and can be obtained from Eq. 5.7.

$$Q_{i+1,j} = Q_{i,j} + \Delta Q_{i+1,j} \tag{5.7}$$

The total number of alighting passengers $Q_{a,i+1}$ at stop i+1, can be obtained from the vehicle O/D matrix $\mathbf{B}_{k,i}$.

$$Q_{a,i+1} = Q_{i,i+1} \qquad \text{for i = 1, 2, ..., S-1} \tag{5.8}$$

In addition, the number of boarding passengers $Q_{i+1,b}$ from stop i+1 can be obtained from Eq. 5.9.

$$Q_{i+1,b} = \sum_{j=i+2}^{S} \Delta Q_{i+1,j} = \overline{q}_{i+1,b}(p_{k,i+1} - p_{k-1,i+1}) \qquad \text{for i = 0, 1, ..., S-2} \tag{5.9}$$

Note that the number of alighting passengers $Q_{a,1}$ at the first stop and the number of boarding passengers $Q_{S,b}$ at the last stop are equal to zero.

Vehicle Dwell Time

The dwell time $d_{k,i}$ for vehicle k at stop i has been formulated by Chien and Chowdlhury (1997) as

$$d_{k,i} = \sum_{n=1}^{N} (\Delta d_{k,i,n}) \qquad \text{for i = 1, 2, ..., S} \qquad (5.10)$$

where N is the last dwell interval within which the last passenger has boarded the vehicle. In Eq. 5.10, if $n=1$, $\Delta d_{k,i,1}$ represents the dwell interval incurred by the boarding/alighting passengers who arrived prior to the arrival of vehicle k at stop i

$$\Delta d_{k,i,1} = \left[Q_{a,i} + \int_{p_{k-1,i}}^{a_{k,i}} \sum_{j=i+1}^{S} q_{i,j}(t) dt \right] t_b \qquad \text{for i = 1, 2, ..., S} \qquad (5.11)$$

where t_b is the average passenger boarding time, while $\int_{p_{k-1,i}}^{a_{k,i}} \sum_{j=i+1}^{S} q_{i,j}(t) dt$ and $Q_{a,i}$ represent the total numbers of boarding and alighting passengers at stop i, respectively. If n is greater than 1, the general form of $\Delta d_{k,i,n}$ is the j^{th} dwell interval equal to the boarding time of wait passengers who arrive from t_1 to t_2 .

$$\Delta d_{k,i,n} = \left[\int_{t_1}^{t_2} \sum_{j=i+1}^{S} q_{ij}(t) dt \right] t_b \qquad \text{for i = 1, 2, ..., S} \qquad (5.12)$$

where t_1 is equal to $a_{k,i} + \sum_{j=1}^{n-2} \Delta d_{k,j}$ and t_2 is equal to $a_{k,i} + \sum_{j=1}^{n-1} \Delta d_{k,j}$, respectively. By summation of Eqs. 5.11 and 5.12, the dwell time for vehicle k at stop i can be formulated. If the passenger arrivals follow a Poisson distribution, $d_{k,i}$ can be simplified as

$$d_{k,i} = [Q_{a,i} + \bar{q}_{i,b}(a_{k,i} - p_{k-1,i}) + \sum_{n=2}^{N} \Delta d_{k,i,n-1} \bar{q}_{i,b}] t_b \qquad \text{for i = 1, 2, ..., S} \qquad (5.13)$$

where $\bar{q}_{ib}$ represents the average passenger boarding rate at stop i.

Passenger Wait Time

If passenger O/D demand distribution $q_{i,j}(t)$ over time t between any pair of stops i and j is known, the total wait time $W_{k,i}$ for bus k at stop i incurred by the number of boarding passengers $Q_{i,b}$ (refer to Eq. 5.9) is formulated in Eq. 5.14

$$W_{k,i} = \int_{p_{k-1,i}}^{p_{k,i}} (p_{k,i} - t) \sum_{j=i+1}^{S} q_{i,j}(t) dt \qquad \text{for i = 1, 2, ...,S-1} \qquad (5.14)$$

Therefore, the average passenger wait time $\bar{w}_{k,i}$ for vehicle k at stop i is

$$\bar{w}_{k,i} = \frac{W_{k,i}}{Q_{i,b}} \qquad \text{for i = 1, 2,...,S-1} \qquad (5.15)$$

If passenger arrivals are Poisson distributed, the passenger wait time $W_{k,i}$ for vehicle k at stop i can be simplified as the average passenger wait time $\bar{w}_{k,i}$ multiplied by the number of passenger arrivals:

$$W_{k,i} = \bar{w}_{k,i}[(p_{k,i} - p_{k-1,i})\bar{q}_{i,b}] \qquad \text{for i = 1, 2,..., S-1} \qquad (5.16)$$

where the average passenger wait time $\bar{w}_{k,i}$ can be approximated by Eq. 5.17:

$$\overline{w}_{k,i} = z(p_{k,i} - p_{k-1,i}) \qquad \text{for i = 1, 2,..., S-1} \qquad (5.17)$$

where z is the ratio of the wait time to the headway between vehicles k-1 and k (e.g. z = 0.5 for random passenger arrivals). Therefore, the total wait time $W_{k,i}$ for bus k at stop i can be obtained by substituting Eq. 5.17 into Eq. 5.16

$$W_{k,i} = z\overline{q}_{i,b}(p_{k,i} - p_{k-1,i})^2 \qquad \text{for i = 1, 2,..., S-1} \qquad (5.18)$$

5.4 Required Data for Simulating Transit Operations

A series of inputs and parameters are required for describing the traffic environment and transit vehicle movement. They can be categorized as (1) supply characteristics, (2) demand characteristics, and (3) simulation control commands.

5.4.1 Supply Characteristics

The required inputs from the supply side include the descriptions of the network geometry, transit routes, transit vehicle types used in the service, prevailing traffic conditions and control devices along the analyzed routes that affect the operation of transit service.

(1) Network Geometry

The network geometry, containing one or multiple transit routes can be represented by a series of stationary nodes (e.g., entries/exits and intersections) and directional links (e.g., freeway segments, ramps, urban streets). All the stationary nodes should be identified with coordinates that facilitate the program in locating them. Links are defined by pairs of upstream and downstream nodes. The link geometry, including

number of lanes and their lengths, grade, average speed and regulation (e.g., HOV and reversible lanes of the link) should also be specified.

(2) Transit Route

A description of stationary nodes and link lengths provides the basic route characteristics required by the simulation program. The stationary nodes along the route consist of intersection nodes and entry/exit nodes. The inputs related to route characteristics can be obtained from maps provided by the public transport operator; roadway geometric, traffic and signal timing information available to local transportation/traffic authorities; and field observation. Other than the location of each stop, information related to the type of stops (e.g. protected or unprotected) should also be distinguished. A stop is "protected" when a vehicle in dwell does not block vehicles in a moving traffic stream. On the other hand, if a vehicle, while serving passengers, blocks the moving lane, the stop is identified as "unprotected". Stops can be classified according to their functional differences. A terminal can be specified with a timetable where all vehicles must stop and wait till the scheduled departure time, and vehicles must stop upon requests from wait and in-vehicle passengers along the route.

(3) Vehicle Type

The program can handle up to seven different types of vehicles in a given simulation session. Vehicles are described in terms of their lengths, capacities, passenger servicing rates and vehicle velocity-acceleration profile. Passenger service rates include mean passenger boarding and alighting rates at each stop, which are highly dependent on the number of doors of vehicles and the way to collect fare. The

construction of a vehicle velocity-acceleration profile requires the average desired cruise speed, maximum attainable speed at zero grade, desired speeds and maximum attainable acceleration/deceleration by the vehicle at zero grade. The velocity-acceleration profile is mainly responsible for generating link travel times for all vehicles.

(4) Traffic Condition

Traffic conditions in the analyzed transit network are represented by specifying different traffic volumes and vehicle compositions entering into the network and turning percentages at intersections during different time periods. The acceptable gaps for turning and lane-changing maneuvers, car-following sensitivity factor, start-up lost time and queue discharge headways can be calibrated to reflect actual driving behavior. Performance characteristics for different types of vehicles such as length, acceleration and deceleration rates can also be adjusted to simulate vehicle operations under different geometric conditions (e.g., different radius of curves and percentages of grade sections).

(5) Traffic Control Devices

Traffic control devices that affect transit operations and can be simulated by the proposed program, include signals at intersections, signs (e.g., yield and stop), and meters on ramps. Control devices can be specified by their locations, types (e.g., pretimed or actuated) and phase timings with different priority (e.g., protected or prohibited).

5.4.2 Demand Characteristics

The simulation program allows for demand variations over space and time. The program requires passenger stop based origin-destination (O/D) demand matrices for specified time periods. Again, provision is allowed to incorporate variations with the space and time of the day.

5.4.3 Simulation Control Command

Input simulation control commands are responsible for the management of the simulation process. They determine the lengths of each and total time periods of simulation, seed values for random number generation, and required input/output options from the simulation. A seed value should be provided to initiate the random number generations. The use of an external seed value ensures the repeatability of a simulation session if required. The single input seed value generates an array of secondary seed values (from a uniform distribution), which are then assigned to individual stochastic generators in the program.

The input options contain a series of cards with specific formats and identification numbers. The data describing road geometry, traffic movements, signal timings, and route features (e.g., locations of routes and stops) are required to be entered by category in a proper sequence. Some transit related information including vehicle schedules and passenger O/D demands are entered through external files linked with CORSIM (Chien and Ding, 1998). On the other hand, the output options are selected through input control

commands. A number of output options can be activated simultaneously. The output options available from the program are described in the following section.

5.5 Program Output

The program can generate results for deriving different output information as discussed below.

(1) A detailed output that provides a summary of input specifications and pre-generated vehicle departure times. The passenger O/D information created by the pregeneration session and later updated for each vehicle in simulation and the data collected for coding the network are also in this category of output. This forms the base-level output option of the simulation program. The output at this stage is amenable to further analysis.

(2) An animation option that displays the operation of the service by means of the computer animation interface TRAFVU (FHWA, 1996). In addition, this display is useful in validating the simulated behavior by visual observation.

(3) The output provided by the program can be used to generate time-velocity and time-distance trajectory diagrams of individual vehicles. This output intercepts the program at the end of each one-second interval and retrieves vehicle acceleration, speed and distance data.

As the program monitors each vehicle in the simulation, it is possible to collect much more data during simulation than through field surveys. In addition, the collection of simulation data excludes the problem of human error associated with field data collection. Although the proposed simulation makes available a large amount of clean

data, converting them into level-of-service measurements is hampered by the lack of consistent definitions for the measurements. In this study, three basic measurements, including travel time, reliability and wait time (Turquist, 1978) are calculated in the analysis section of the program.

(1) Travel Time

The vehicle travel time is an important level-of-service measure for transit operators. The mean and standard deviation of the stop-to-stop travel time are measured by processing the data of travel speeds and locations of transit vehicles, which can be obtained from an emulated AVLS environment provided by the simulation program. The mean travel time per vehicle trip can be computed by averaging all individual trip travel times over different time periods of a day.

(2) Reliability

Passengers are very concerned about the regularity of the public transport service (Turnquist, 1978). The literature suggests that reliability measures could be based on deviations from the schedule (Bly and Jackson, 1974), standard deviation of the operating headways, and measures incorporating variability of transfer delays (Polus, 1979; Skinner, 1980; Turnquist and Bowman, 1980). Vandebona and Richardson (1981) reported the application of root mean square of headway deviations from scheduled headway as a measure of reliability. The standard deviation of vehicle travel time, described under the travel time measurements, also reflects the reliability of the service because it indicates the deviations in headways/schedules.

(3) Wait Time

The wait time of the passengers also reflects the reliability of the service. Passenger wait time begins to increase when vehicles deviate from their schedule and more passengers experience long waiting periods. Vehicles with short headways can pick up a relatively smaller number of passengers than those with long headways, thus cannot compensate for the long wait time experienced by the larger number of passengers. In the program, the passenger wait time is the time interval from when a passenger joins the queue to the stop to the time at which the passenger boards the vehicle. The total passenger wait time is not meaningful for comparison purposes as the total number of passengers using the service may differ. Generally, wait time per passenger is selected as the measure of wait time.

5.6 Model Calibration and Validation

To conduct credible simulation analysis, one must be confident that the results from the proposed simulation program should represent real world transit operations reasonably well. Although the program was enhanced from CORSIM, which has been extensively validated during its development stage, procedures are taken in this study to further ensure that the program output is adequately substantiated by the data collected from field observation. Operation related data for bus Route #39 of New Jersey Transit Corporation are collected, while the procedures to assess the reliability of the simulation program include (1) the calibration of the program embedded parameters, (2) the validation of the simulation program, and (3) the analysis of the simulation results.

5.6.1 Data Collection and Model Calibration

The proposed simulation program was tested using data collected from Route #39 of New Jersey Transit and local transportation agencies to ensure that the program can reliably describe transit operations and produce outputs. The simulation experiment began by collecting various data required for the program input. Detailed data including time dependent traffic counts and signal timing at intersections along the route, were collected from the site and local governments (Cities of Newark, Kearny and Harrison). Bus operation data (e.g., arrival/departure times, stop-to-stop travel times, bus journey times, and boarding/alighting demands) during morning peak periods (7:30AM-9:30AM) in weekdays were collected on site and provided by New Jersey Transit Corporation. These data were divided into two separate sets for calibrating and validating the simulation program.

A 4.4-mile computer network containing a segment of Route #39 of New Jersey Transit Corporation was established. The network contains a total of 142 nodes and 223 links, in which the route traverses 30 intersections (of which 26 are signalized) and serves 14 stops per direction, as shown in Figure 5.2. Bus operations in a peak hour were observed with the simulation program. The default values of parameters (e.g., lane-changing time, reaction time, driver type factor, percentage of cooperative drivers, average passenger boarding time, capacity for bus stops, discharged headway and startup delay) which may affect traffic operations on streets embedded in the simulation program are identified, calibrated, and shown in Table 5.1. Other parameters, including default

start-up delay and queue discharge headway distribution for a link, were adjusted and shown in Table 5.2 (Chien and Ding, 1999d; Ding and Chien, 1999e).

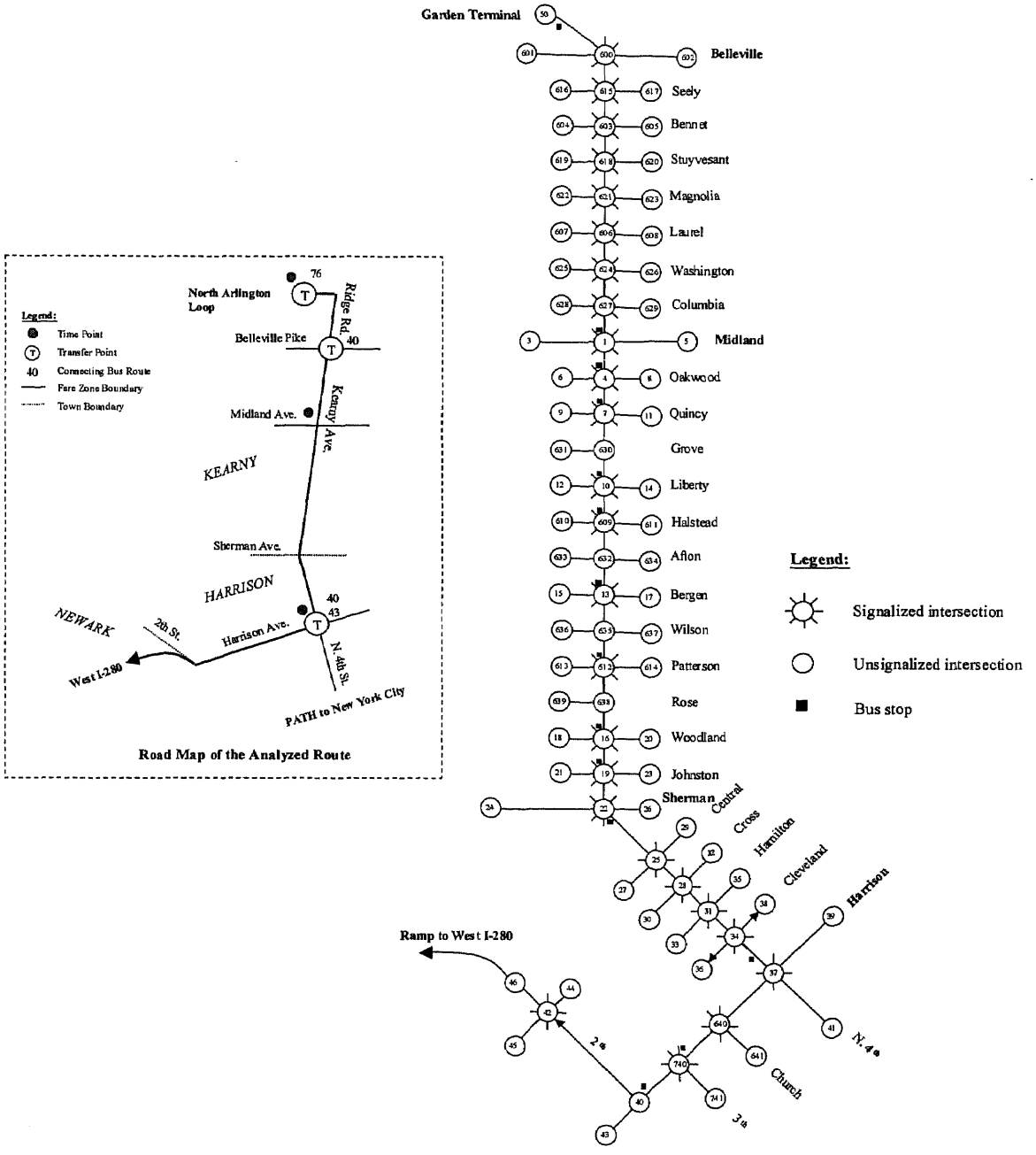


Figure 5.2 Link-node diagram of Route # 39, NJ Transit Corporation

Initial calibration of the program consists of using field-observed mean bus dwell times at stops and travel times between stops. Also repeated adjustments of link "free flow" speed and parameters identified in Tables 5.1 and 5.2 are made to fit bus arrival patterns at downstream stops observed in the field. Improvements caused by changes to the program are monitored by comparing the program output with the calibration field data set. The primary parameters causing major discrepancies appear to be the passenger boarding time, queue discharge headway, and start-up delay. The impact of these factors on the average stop-to-stop travel time between stops #4 and #5 is summarized and shown in Figure 5.3.

Table 5.1 Calibrated parameters

Parameter	Default Value	Calibrated Value
Lane-changing Time	3 s	2 s
Driver Reaction Time	1 s	1.5 s
Driver Type Factor	25	20
Cooperative Drivers*	25%	40%
Avg. Pass. Boarding Time	2.0 s/pass.	2.2 s/pass.
Capacity for Bus Stops	1 bus/stop	2 buses/stop
Start-up Delay	2.0 s	2.5 s
Discharge Headway	1.8 s	2.2 s

* Those drivers who will cooperate a lane-changing by slowing down their vehicles.

Table 5.2 Start-up delay and discharge headway distributions

Start-up Delay											
		Driver Characteristic									
	Mean	1	2	3	4	5	6	7	8	9	10
Type*1 (Default)	2.0 sec	218	140	125	118	102	86	78	63	47	23
Type 2 (Default)	2.0 sec	258	190	143	114	95	76	57	38	29	0
Type 3	2.5 sec	156	135	116	108	98	92	85	78	70	62
Type 4	2.5 sec	147	131	124	112	103	89	82	76	71	65
Discharge Headway											
		Driver Characteristic									
	Mean	1	2	3	4	5	6	7	8	9	10
Type 1 (Default)	1.8 sec	170	120	120	110	100	100	90	70	70	50
Type 2 (Default)	1.8 sec	180	140	120	110	100	90	80	70	60	50
Type 3	2.2 sec	147	131	124	112	103	89	82	76	71	65
Type 4	2.2 sec	216	136	105	98	92	81	78	72	65	63

* Each type contains an array of percentage values applied to determine the start-up delay or discharge headway for a vehicle on the specified link.

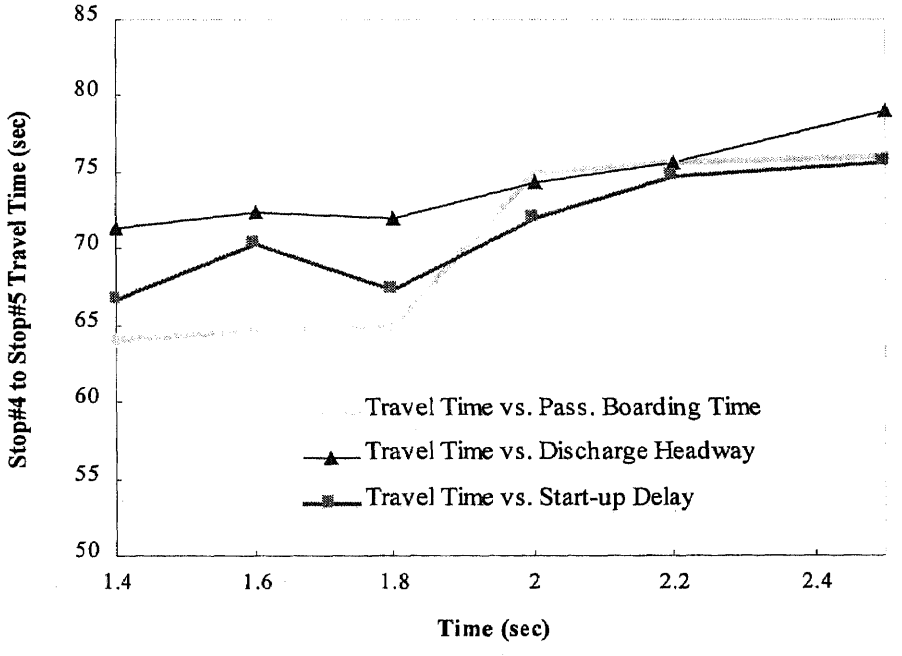


Figure 5.3 Impact of calibrated parameters on bus stop-to-stop travel time

5.6.2 Model Validation

The proposed program is validated through comparing simulation outputs generated from the calibrated program with field counterparts, such as the mean and the variance of bus stop-to-stop travel times and journey times. The simulation validation process adopted in this study is composed by the following comparisons: Graphical Comparison, Aggregate Comparison, and Statistical Comparison, which are discussed below.

(1) Graphical Comparison: The graphical comparison is a subjective validation approach. The graphical displays emphasize or reveal the aspects of data that are not easily captured by numerical summaries or tabular representations. Diagnosing the displays is especially useful for testing the results generated by the simulation program preliminarily.

(2) Aggregate Comparison: Aggregate means and standard deviations give general indication of system performances in real world and in simulation. However, they do not present an accurate trend or indication of how variables perform over time, what patterns are created, or how much individual measurements deviate. Aggregate comparison, along with the graphical comparisons of scatter plots, reveals the similarities and discrepancies of the magnitude and changing pattern for variables.

(3) Statistical Comparison: A statistical analysis is crucial for validating the proposed program based on sample data collected from the real world and simulated transit systems. It is used for assessing the accuracy of the program, testing various hypotheses and determining degree of correlation between both systems. The following indicators are used for conducting statistical analyses of the simulation outputs,

Mean Absolute Percent Error (MAPE)

MAPE measures the percentage error between simulation results and field data, which can be obtained from Eq. 5.19

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|SIMU_i - FIELD_i|}{FIELD_i} \times 100\% \tag{5.19}$$

where n is the sample size, $SIMU_i$ is the simulation output, and $FIELD_i$ is the field measurement.

Root Mean Square Error (RMSE)

RMSE denotes the error between simulation results and field data measured in time (e.g., second), which can be determined by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SIMU_i\text{-}FIELD_i)^2} \tag{5.20}$$

where n is the sample size, $SIMU_i$ is the simulation output, and $FIELD_i$ is the field measurement.

Correlation Coefficient

Correlation coefficient $CORRE$ measures numerically the degree of closeness between simulation outputs and field data, which can be obtained as

$$CORRE = \frac{\frac{1}{n}\sum_{i=1}^{n}(SIMU_i - \mu_S)\cdot(FIELD_i - \mu_f)}{\sigma_S \cdot \sigma_f} \tag{5.21}$$

where μ_S and μ_f represent the means of $SIMU_i$ and $FIELD_i$ (i = 1, 2,..., n), while σ_S and σ_f are the standard deviations of $SIMU_i$ and $FIELD_i$ (i = 1, 2,..., n),

respectively. A coefficient of zero indicates no correlation, while a coefficient of 1 denotes significant correlation.

In this study, the validity of the program is assessed by analyzing mean and standard deviation of bus stop-to-stop travel times and journey times. The preliminary test of bus stop-to-stop travel times is conducted by comparing the scattered plots of field observations and simulation results. A total of 130 samples were collected from 10 bus-trips (at 14 stops) in the field, while 312 samples were collected from 24 bus-trips from enhanced program and 312 samples were collected from 24 bus-trips from the original CORSIM, respectively. The scattered plots of field observations and the corresponding simulation results from enhanced and original CORSIM are shown in Figures 5.4(a) and 5.4(b), respectively. Both figures show that the stop-to-stop travel times tend to fluctuate because of traffic congestion, signal delays and passenger demand. Comparing the two figures, we found that the outputs generated by the enhanced CORSIM match the field observations more closely.

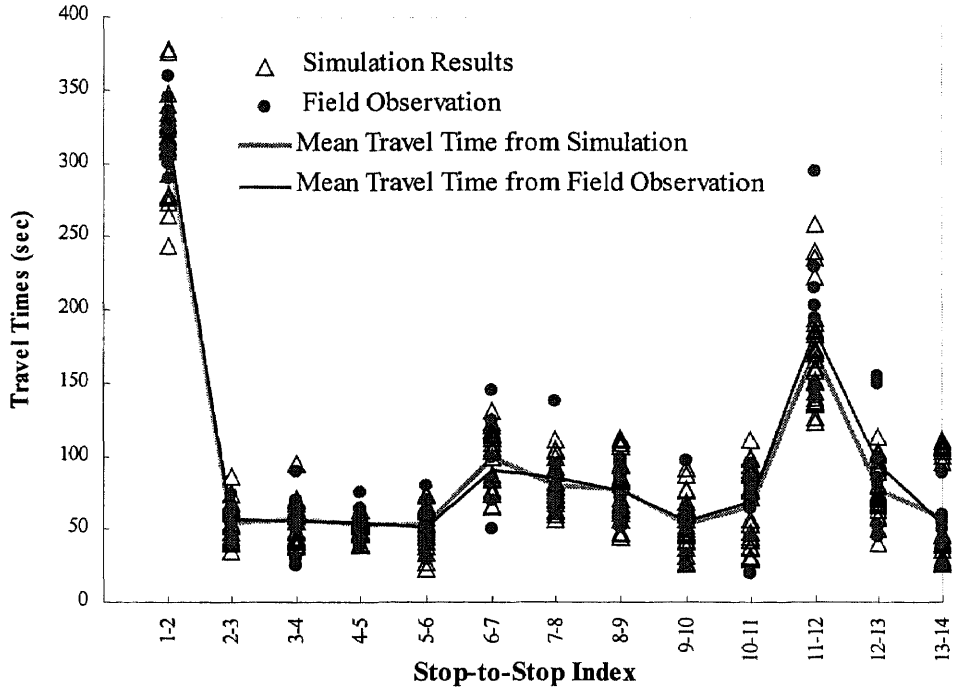


Figure 5.4(a) Field and simulated stop-to-stop travel times (enhanced CORSIM)

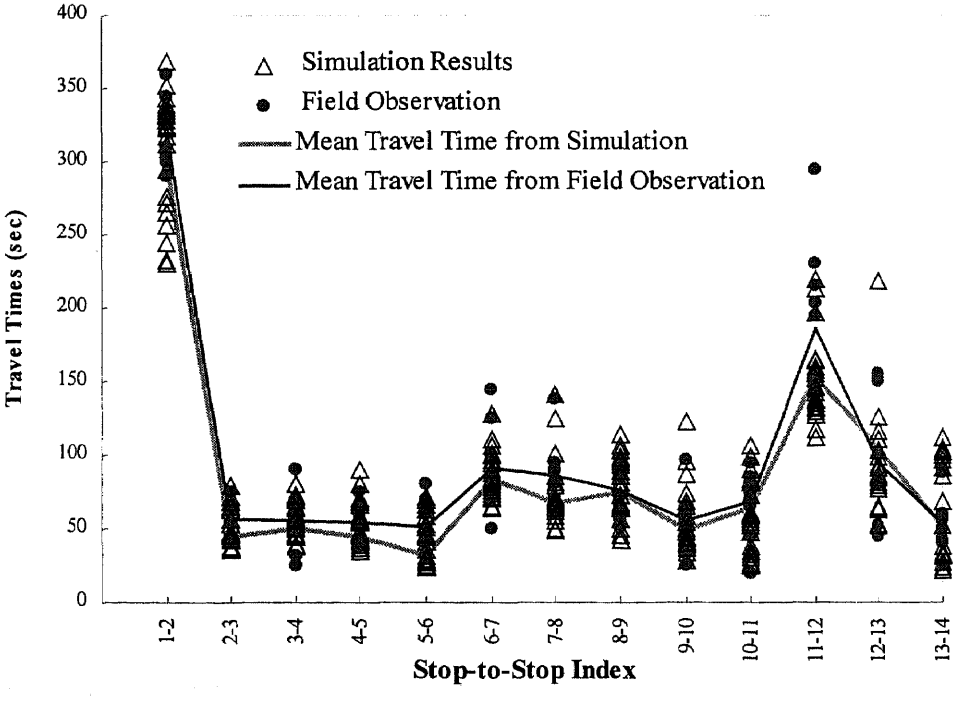


Figure 5.4(b) Field and simulated stop-to-stop travel times (original CORSIM)

The standard deviations of stop-to-stop travel times are calculated based on field observations and simulation results and shown in Figure 5.5, within which the difference between the standard deviations of the two sets of samples is large. Observing day-to-day service on the analyzed route, we found that the segments with higher standard deviation of stop-to-stop travel times are associated with more numbers of signalized intersections between a pair of stops or a short stop spacing (e.g., Stops #12-13). In order to capture the variations of stop-to-stop travel times on those segments, more field data should be collected.

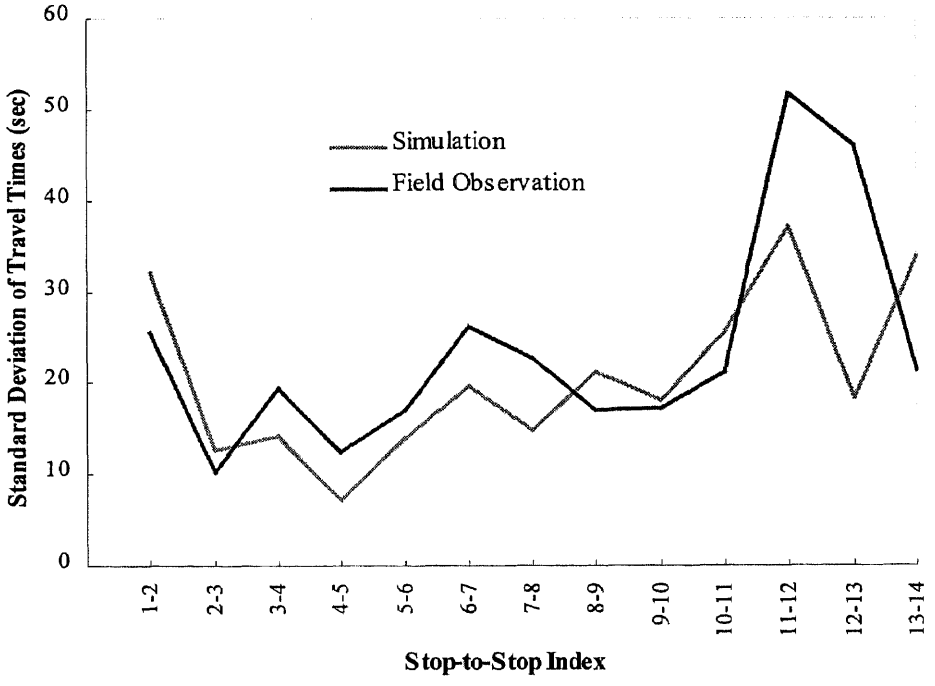


Figure 5.5 Comparison of standard deviations of stop-to-stop travel times

The mean and standard deviation of bus journey times per trip collected from the study site and simulation are analyzed. For the 130 field data and the 312 simulation outputs, the mean values of 21.03 and 20.25 minutes and the standard deviations of 2.25 and 2.52 minutes of bus journey times are derived, respectively. This indicates that the enhanced CORSIM can capture the deviation of bus journey times more accurately than that of the individual stop-to-stop travel times.

The statistical analysis of bus stop-to-stop travel times and journey times is conducted by calculating mean absolute percentage errors (MAPE) and root mean square errors (RMSE). The results of MAPE and RMSE of bus stop-to-stop travel times are shown in Figure 5.6, where the RMSE ranges from 7 to 19 seconds. The larger value of

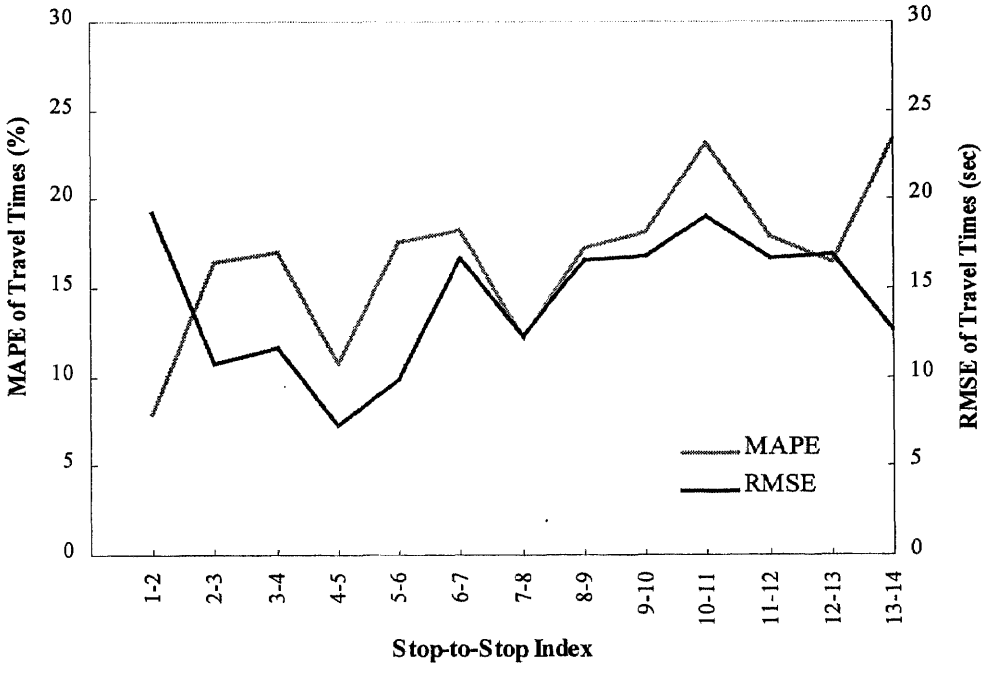


Figure 5.6 The MAPE and RMSE of predicted stop-to-stop travel times

RMSE (e.g., 19 seconds) is incurred by buses operating on a segment between stops #1 and #2. Table 5.3 gives the MAPE and RMSE of stop-to-stop travel times and the corresponding distances and the number of signalized intersections between pairs of stops. After investigating the table, we found that the segments with higher MAPE and RMSE of stop-to-stop travel times are caused by the short segment length and more number of intersections (e.g., Stops #10-11, #11-12, and #13-14). If more intersections are traversed within a short distance, buses will experience more random delays, which accounts for the deviations between the field observations and simulation results.

Table 5.3 MAPE and RMSE of predicted stop-to-stop travel times

Stop-to-Stop Index	Stop-to-Stop Distance (ft)	# of Intersections	MAPE (%)	RMSE (sec)
1-2	4466	8	7.91	19.29
2-3	1003	1	16.41	10.73
3-4	1003	1	17.00	11.63
4-5	1056	1	10.77	7.33
5-6	422	1	17.61	9.84
6-7	1478	1	18.28	16.64
7-8	1312	1	12.16	12.3
8-9	980	1	17.26	16.56
9-10	581	1	18.10	16.80
10-11	552	2	23.21	18.99
11-12	1806	5	17.95	16.63
12-13	598	2	16.40	16.91
13-14	422	1	23.43	12.66

The MAPE and RMSE of predicted bus journey times for 312 samples obtained from simulation outputs are 9.25% and 0.98 minutes, respectively. We found that the

enhanced CORSIM can capture bus journey times more accurately than the stop-to-stop travel times. The correlation analysis demonstrates strong relationship between the data collected from the field and simulation, with the correlation coefficients of 0.94 for stop-to-stop travel times and 0.98 for journey times.

The statistics (e.g., mean and variance) of stop-to-stop travel times and journey times are summarized in Table 5.4. After conducting the hypothesis tests (e.g., t-test and F-test), we concluded that the simulation results could statistically represent the field data at a 95% confidence interval.

Table 5.4 Statistical summary of stop-to-stop and total travel time

Stop-to-Stop Index	N_A	μ_A	σ_A	N_B	μ_B	σ_B	t-value	F-value
1-2	10	316.50	26.90	24	313.92	32.41	0.02	0.69
2-3	10	56.67	11.57	24	54.54	12.96	0.03	0.80
3-4	10	55.83	18.39	24	57.46	13.90	-0.03	1.75
4-5	10	54.67	11.79	24	54.08	6.81	0.01	3.00
5-6	10	51.40	16.32	24	53.67	14.14	-0.02	1.33
6-7	10	91.20	28.61	24	101.42	18.68	-0.07	2.35
7-8	10	85.40	22.29	24	79.58	15.16	0.03	2.16
8-9	10	76.20	16.87	24	79.00	20.63	-0.02	0.67
9-10	10	55.50	19.50	24	54.17	18.51	0.01	1.11
10-11	10	69.00	21.96	24	67.54	24.96	0.00	0.77
11-12	10	186.30	51.13	24	173.46	36.37	0.03	1.98
12-13	10	92.60	33.51	24	77.54	18.68	0.04	3.22
13-14	10	53.30	22.46	24	59.42	33.82	0.00	0.44
1-14	10	1244.57	123.05	24	1225.79	143.57	0.01	0.73

N_A: Sample size of field observations
μ_A: Mean travel time of field observations
σ_A: Standard deviation of travel time
$F_{0.975,9,23} = 0.366$
$t_{0.05,32} = 2.04$

N_B: Sample size from simulation output
μ_B: Mean travel time from simulation output
σ_B: Standard deviation of travel time
$F_{0.025,9,23} = 3.61$

5.6.3 Simulation Analysis

The program yields various quantitative measurements related to transit service (e.g. bus arrival/departure times, bus dwell times, average passenger wait times, headways and standard deviations of headways). The trends and variations over time and space of the measurements are analyzed. For example, the tendency for vehicle bunching initiated by abnormal ridership among stops can be readily observed in a time-distance trajectory diagram (e.g., Figure 5.8).

The validated simulation program is used to simulate bus operations at stop#7, where the boarding and alighting demand increases from 100% to 250% of the original demand. In the simulation, a 5-minute dispatching headway is assumed, while the standard deviation of the vehicle operating headways checked at every stop is determined by five consecutive buses and plotted in Figure 5.7. The figure demonstrates that, in

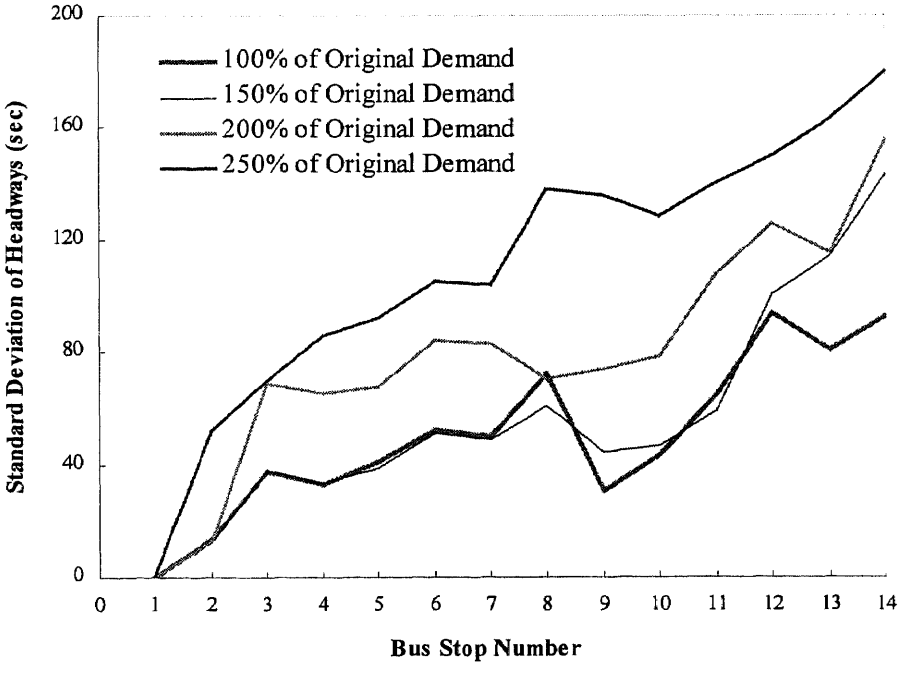


Figure 5.7 Standard deviations of headways vs. bus stops for various demands at stop #7

general, the standard deviation of headways tends to increase as the stop number increases (service route length increases) and is consistent with the real-world observation. We also found that the increasing passenger demand at stop #7 (e.g., 150%, 200%, 250%) causes larger headway variances at further downstream stops. The increase in the headway variances at stops before stop#7 is incurred by the relatively large demands at these stops in corresponding to the increase in the alighting demand at stop#7.

The time-distance trajectories of five consecutive buses are obtained based on the increase of passenger demand at stop #7 from 100% to 250% of the original demand and shown in Figure 5.8 (in simulation seconds). The situation of two buses bunching up at stop #8 was found at the 1816th second (8:30:16AM). The unevenly distributed passenger demand among stops (with unusually large demand at stop #7) contributes to the

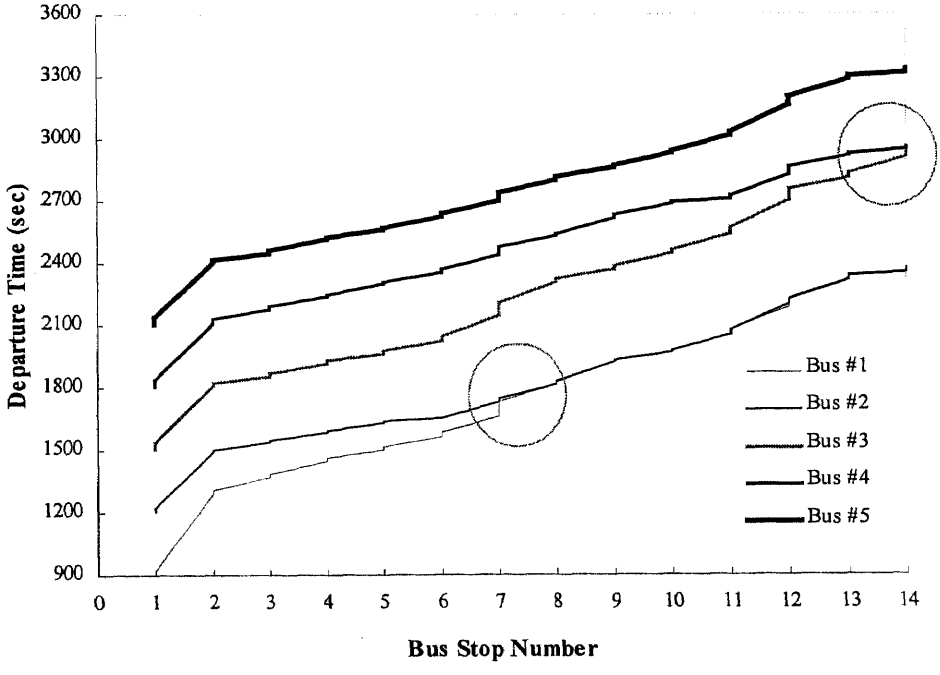


Figure 5.8 Time-space trajectories of five consecutive buses

irregularity of operating headways. Because Bus #1 was delayed by serving a large number of boarding/alighting passengers at Stop #7, the following vehicle, Bus #2 is able to catch up and eventually bunches with Bus #1 at Stop #8. The similar bunching situation between Buses #3 and #4 can also be observed at Stop #14.

The bus dwell time at every stop observed from simulation results is shown in Figure 5.9. By observing Figures 5.8 and 5.9, we found that a smaller headway between a leading bus and its follower bus may result in a relatively smaller number of waiting passengers at a downstream stop, which causes shorter bus dwell time, and vise versa. Because bus #2 operates with a gradually smaller headway to bus #1 due to continuously less boarding passengers at downstream stops, bus #2 overtakes bus #1 and arrives at stop #8 earlier than bus #1. Bus #1 keeps at a close distance following bus #2 at stops #9, #10 and #11, with a very small headway and less dwell time (mainly for unloading passengers at these stops), until it overtakes bus #2 at stop #12. Figure 5.10 shows that, at stop #8, the passenger wait time for bus #1 decreases dramatically because of a very small headway between buses #1 and the previous bus #2, while the passenger wait time for bus #2 decreases sharply because it is overtaken by bus #1 at stop#12. Such bunching phenomena have been successfully simulated and observed from these figures.

In this chapter, the simulation program was validated through calibrating a real world bus route, while simulation results were used to analyze transit related MOEs. The analysis on mean average percentage error (MAPE) and the root mean square error (RMSE) have demonstrated that simulation outputs are reliable. The study also shows that the enhanced CORSIM can simulate the disruptions of transit headways, such as the vehicle bunching phenomenon due to ridership fluctuations at stops.

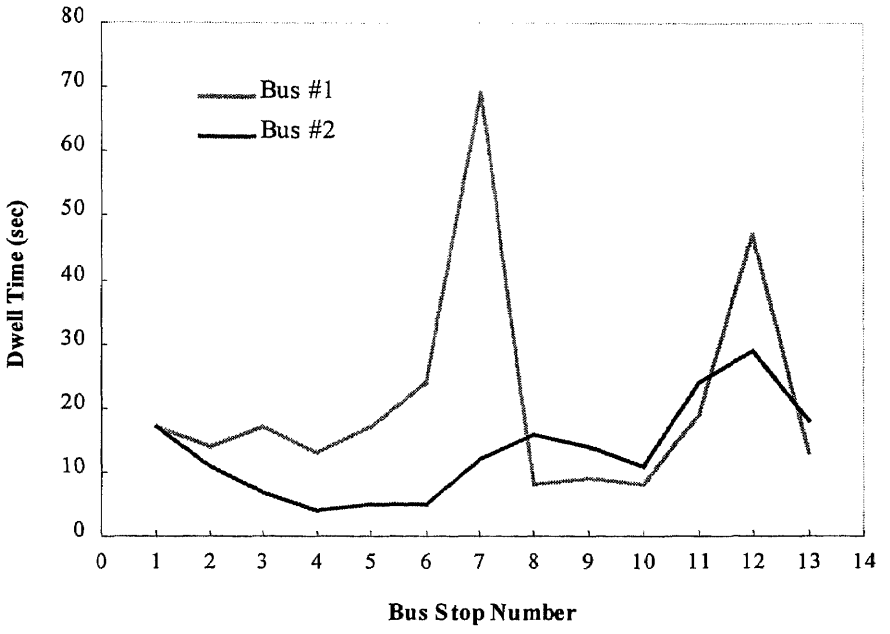


Figure 5.9 Bus dwell times at different stops

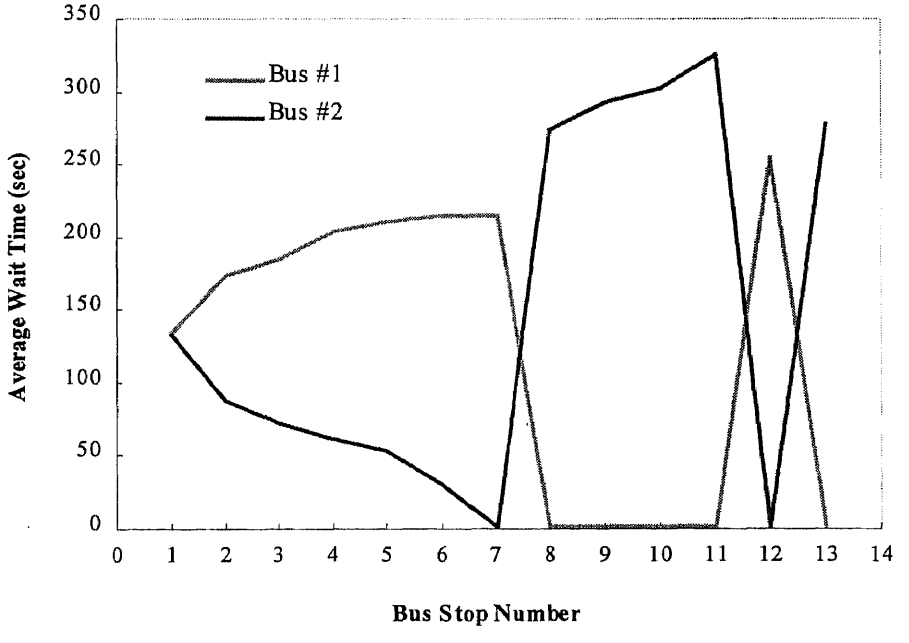


Figure 5.10 Average wait times at different stops

CHAPTER 6

MODEL EVALUATION

In this chapter, the developed transit arrival prediction models discussed in Chapter 3, including the basic, Kalman filtering, two Artificial Neural Networks (ANNs) and two Neural/Dynamic (ND) models are evaluated by simulating a real world transit route, while the reliability analysis of predicted arrival times is conducted. The testing and evaluation results are discussed and organized into the following sections: (1) Basic and Kalman filtering prediction models (2) ANNs, (3) ND prediction models and (4) the real-time vehicle control model.

6.1 Basic and Kalman Filtering Prediction Models

The basic model and a Kalman filtering-typed model developed in sections 3.1 and 3.3 are tested by simulating Route #39 of New Jersey Transit, which serves stops at the cities in North Arlington, Kearny, Harrison, Newark, and Irvington, as shown in Figure 6.1. The analyzed 4.4-mile long segment starts from North Arlington Loop and ends at the 2nd Street of Harrison, encompassing 30 intersections (of which 26 are signalized) and serving 14 stops. The basic model and the Kalman filtering model are integrated with the simulation program individually. The arrival times of 24 buses at each of the 13 stops (excluding the first stop at the dispatching terminal) of the segment are predicted during simulation.

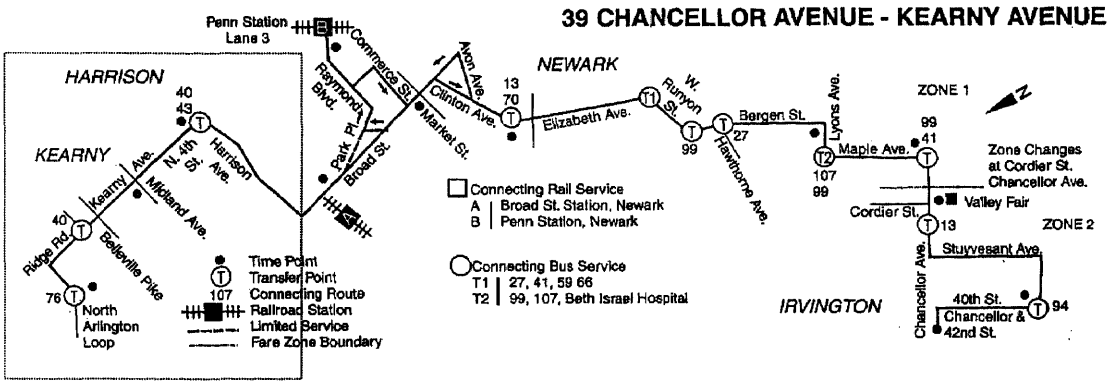


Figure 6.1 NJ Transit Route #39

The basic model predicts bus arrival times at downstream stops of the segment by simply adding up the current average link travel times between a pair of stops. Note that a link on a transit route represents either a signalized or unsignalized intersection. The inputs of the basic model include travel distances on links, average link speeds and passenger (boarding/alighting) demand, which can be obtained during simulation.

The Kalman filtering model predicts bus arrivals by adjusting its parameters (e.g., Kalman Gain) in real-time to accommodate traffic conditions, whose inputs are stop spacings and the means of link speeds between pairs of stops and passenger demands. The initial values of the Jacobian factor $\Theta_{0,i}(-)$ (i = 1, 2,..., 14) are chosen as 1 while the error covariance $M_{0,i}(-)$ (i = 1, 2, ..., 14) are chosen to be 100 sec^2. These initial values may speed convergence of the prediction error in the Kalman filtering algorithm for the analyzed system (Tavantzis and Ding, 1999).

The time-space diagram for the bus dispatched at the 1322 simulation second (8:22:02 AM) is plotted and shown in Figure 6.2, along with the arrival times predicted by the basic and the Kalman filtering models. In the figure, the Kalman filtering model appears to perform better than the basic model. In Figure 6.2(a), we found that the discrepancy between the simulated and predicted arrival times increases as the stop number increases, indicating that the prediction error of the basic model accumulates quickly along the route. Both variations of traffic conditions on downstream links and demand at stops affect the inaccuracy of the predicted arrival times.

The accuracy analysis of the stop-to-stop travel times predicted by the basic model and Kalman filtering model is conducted and shown in Figure 6.3. The Kalman filtering model appears to have better accuracy than that of the basic model at nearly all stops. This is because the Kalman filtering model can adjust its parameters on-line (e.g., Kalman Gain) to reduce the prediction error. In Figure 6.3, deviations between the predicted and simulated travel times at those pairs of stops with longer stop-spacings (e.g., stops#1-2 and #11-12) are relatively high for both models. The RMSE of the predicted stop-to-stop travel times for both models is listed in Table 6.1. We found that the Kalman filtering model outperforms the basic model at all pairs of stops. However, the RMSEs at stops #1-2 and #11-12 are fairly high for both models, indicating that neither model can predict accurately for a pair of stops because the number of intersections between these stops are relatively large.

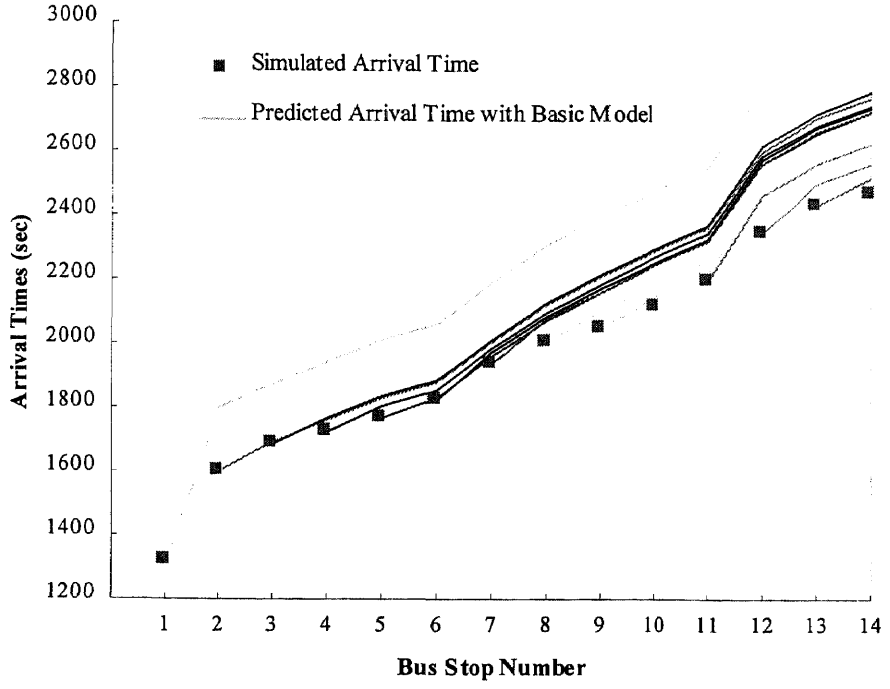


Figure 6.2(a) Predicted vs. simulated arrival times for a bus (the basic model)

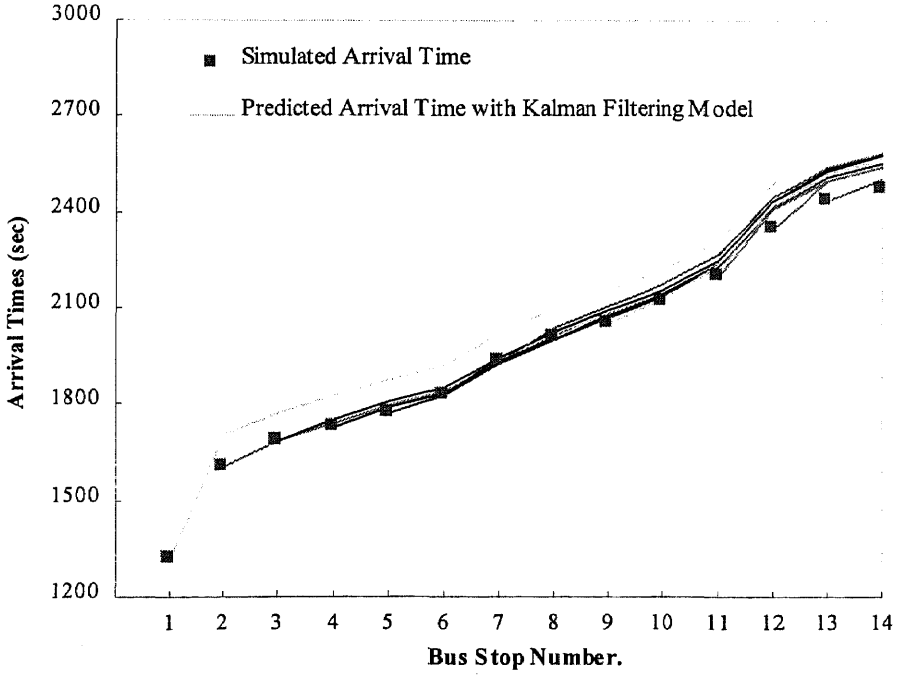


Figure 6.2(b) Predicted vs. simulated arrival times for a bus (the Kalman filtering model)

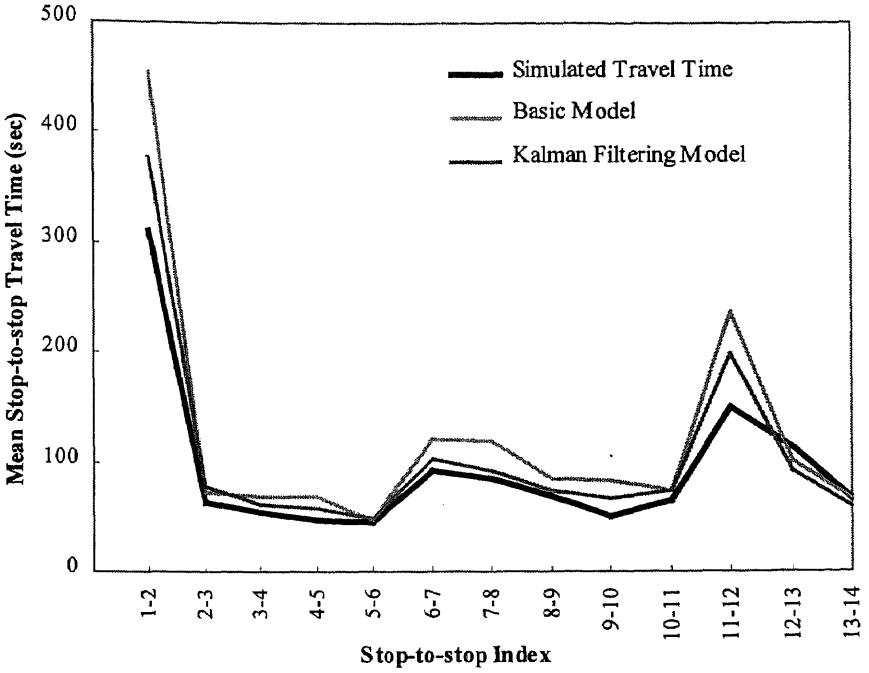


Figure 6.3 Predicted stop-to-stop travel times by the basic and Kalman filtering models

Table 6.1 RMSE of bus stop-to-stop travel times by the basic and Kalman filtering models

Stop-to-stop Index	Stop-to-stop Distance (ft)	# of Intersections	RMSE (sec)	
			Basic Model	Kalman Filtering Model
1-2	4466	8	154.77	77.69
2-3	1003	1	33.80	29.30
3-4	1003	1	28.63	20.12
4-5	1056	1	33.89	21.23
5-6	422	1	27.85	17.56
6-7	1478	1	51.07	33.50
7-8	1312	1	59.18	31.83
8-9	980	1	35.51	29.79
9-10	581	1	47.58	32.95
10-11	552	2	47.20	37.85
11-12	1806	5	118.86	84.73
12-13	598	2	47.81	35.91
13-14	422	1	55.44	35.69

6.2 Artificial Neural Networks

In this section, a link-based and a stop-based ANN for predicting bus arrival times are evaluated. The simulation analysis of a segment of NJ Transit Route #39 is conducted. Simulation outputs (e.g., link volumes and speeds, vehicle delays, bus locations, bus arrival/departure times) are collected and analyzed, while factors affecting bus travel times are identified and classified as explanatory variables for the developed ANNs. The BP algorithm is applied for training the ANNs with sufficient training examples, which are collected while simulating bus operations on the analyzed segment under various traffic conditions. Both the link-based and stop-based ANNs are integrated with the simulation program individually, while the reliability analysis of the two ANNs at various downstream stops is conducted.

6.2.1 ANN Training

Bus operations in the morning peak (7:30AM-9:30AM) on the analyzed segment of NJ Transit Route #39 are simulated. The results (e.g., link volumes and speeds, vehicle delays, bus locations, bus arrival/departure times) obtained during simulation for 24 buses are collected for further analysis. We found that many factors affect bus arrival times, such as traffic control devices, link lengths, stop spacings, traffic volumes and densities, vehicle speeds and delays, and passenger (boarding/alighting) demand.

After analyzing those factors identified above and the corresponding bus travel times on links and between stops, MOEs with the potential to be the explanatory variables for the link-based and stop-based ANNs are listed in Table 6.2. The MOEs

affecting bus link travel times include those related to traffic conditions on each link, such as bus travel distance on a link (LDIS), average link traffic volume (LVOL), average link speed (LSPD), average link delay (LDLY), average queue time on a link (LQUE) and number of boarding/alighting passengers at stops (PASS). On the other hand, the MOEs affecting bus stop-to-stop travel times include those related to aggregate traffic conditions over all the links between a pair of stops. These MOEs are stop-spacing (SDIS), mean and standard deviation of volumes (SVOL and DVOL), mean and standard deviation of speeds (SSPD and DSPD), mean and standard deviation of delays (SDLY and DDLY), number of intersections traversed between a pair of stops (SINTE) and number of boarding/alighting passengers at stops (PASS).

Different combinations of the MOEs are experimented as input variables for the link-based and stop-based ANNs, respectively, as shown in Table 6.3. According to a correlation analysis among these MOEs shown in Table 6.4, we found that for the link-based ANN, LDLY is correlated with LQUE (0.96) and LSPD (-0.52), respectively; and LSPD is correlated with LQUE (-0.67). The correlation constraints are taken into account for choosing a combination (e.g., Models#3, #4 and #5). However, some MOEs in conflict with these constraints are also grouped into one combination (e.g., Models #1 and #2), considering that the ANN performance may be enhanced when trained with more information (Hagan, et. al., 1996). For the stop-based ANN, Table 6.4 shows that SDLY is correlated with SSPD (-0.73) and DDLY is correlated with DSPD (0.83). These constraints are reflected in Models #6 and #7. Meanwhile, SINTE is found to be correlated with DVOL (0.65), DSPD (0.42) and DDLY (0.50), respectively, which is considered in Models #8, #9 and #10.

Table 6.2 MOEs related to link-based and stop-based ANNs

MOEs Related to Link-based ANN	
MOEs	**Definition**
LDIS (ft)	Bus travel distance on a link.
LVOL(vph)	Average link volume, accumulated by traffic counts on a link.
LSPD(mph)	Average link speed, calculated as the total vehicle-mile divided by the total travel time for all vehicles accumulated on a link.
LDLY(sec/veh)	Average link delay, calculated as the difference between the simulated and desirable (measured by the free-flow speed) total travel times over the total vehicle trips on a link.
LQUE(sec/veh)	Average queue time on a link, calculated as the average time for vehicles idling in a queue due to traffic control or congestions.
PASS (-)	Number of boarding/alighting passengers at stops.
MOEs Related to Stop-based ANN	
MOEs	**Definition**
SDIS(ft)	Stop-spacing.
SVOL(vph)	Mean of LVOLs on all the links between a pair of stops.
DVOL(vph)	Standard deviation of LVOLs on all the links between a pair of stops.
SSPD(mph)	Mean of LSPDs on all the links between a pair of stops.
DSPD(mph)	Standard deviation of LSPDs on all the links between a pair of stops.
SDLY(sec/veh)	Mean of LDLYs on all the links between a pair of stops.
DDLY(sec/veh)	Standard deviation of LDLYs on all the links between a pair of stops.
SINTE (-)	Number of intersections traversed between a pair of stops.
PASS (-)	Number of boarding/alighting passengers at stops.

Table 6.3 SSEs for different link-based and stop-based ANNs

Link-based ANN				
Model No.	**Input Variables**	**No. of Hidden Neurons**	**No. of Training Examples**	**SSE (sec^2)**
1	LDIS, LVOL, LSPD, LDLY, PASS	6	380	**0.0965**
2	LDIS, LVOL, LSPD, LQUE, PASS	6	380	0.1108
3	LDIS, LVOL, LSPD, PASS	5	380	0.1108
4	LDIS, LVOL, LDLY, PASS	5	380	0.1104
5	LDIS, LVOL, LQUE, PASS	5	380	0.1108
Stop-based ANN				
Model No.	**Input Variables**	**No. of Hidden Neurons**	**No. of Training Examples**	**SSE (sec^2)**
6	SDIS, SVOL, SSPD, DVOL, DSPD, PASS	7	344	0.0694
7	SDIS, SVOL, SDLY, DVOL, DDLY, PASS	7	344	0.0758
8	SDIS, SVOL, SSPD, SDLY, SINTE, PASS	7	344	**0.0410**
9	SDIS, SVOL, SSPD, SINTE, PASS	6	344	0.1103
10	SDIS, SVOL, SDLY, SINTE, PASS	6	344	0.0504

Table 6.4 Correlation of MOEs* related to link-based and stop-based ANNs

Correlation of MOEs Related to Link-based ANN				
	LVOL	**LSPD**	**LDLY**	**LQUE**
LVOL	1	-0.27	0.25	0.22
LSPD		1	**-0.52**	**-0.67**
LDLY			1	**0.96**
LQUE				1

Correlation of MOEs Related to Stop-based ANN							
	SVOL	**SSPD**	**SDLY**	**DVOL**	**DSPD**	**DDLY**	**SINTE**
SVOL	1	-0.24	0.45	0.01	-0.24	0.07	0.05
SSPD		1	**-0.73**	-0.06	-0.03	-0.28	-0.28
SDLY			1	-0.18	-0.26	0.00	0.05
DVOL				1	0.36	0.42	**0.65**
DSPD					1	**0.83**	**0.42**
DDLY						1	**0.50**
SINTE							1

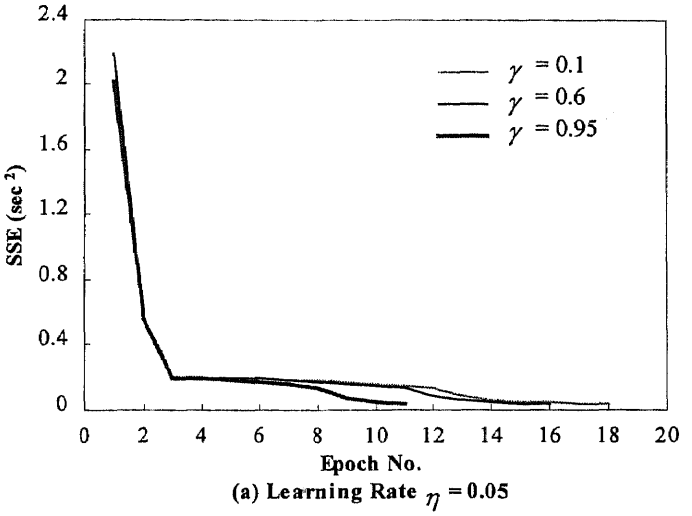
*Note: The LDIS and PASS, SDIS and PASS are not included in the correlation analysis since they are chosen as input variables for each ANN in Table 6.3.

The BP algorithm is applied to train all ANNs illustrated in Table 6.3, while the training data are collected from various buses during simulation. In ANN training,

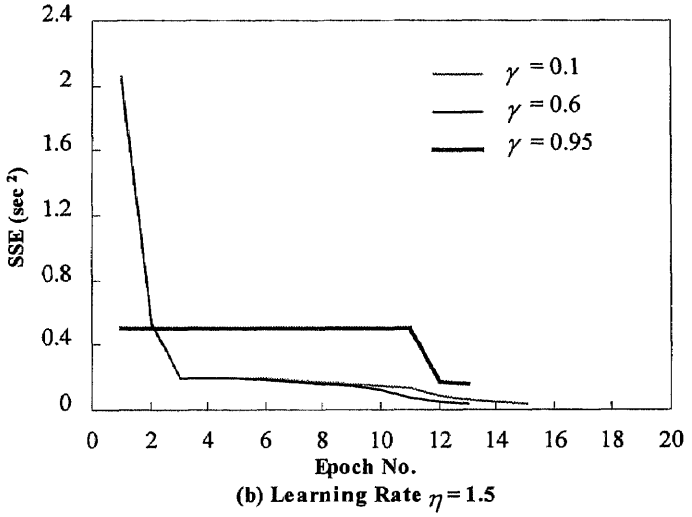
different values of parameters (the number of hidden neurons, the momentum rate γ and the learning rate η) are tested and shown in Table 6.3. For the link-based ANN, the inputs with the lowest SSE over 380 training examples are LDIS, LVOL, LSED, LDLY and PASS (Model #1). On the other hand, for the stop-based ANN, the inputs with the lowest SSE over 344 training examples are SDIS, SVOL, SSPD, SDLY, SINTE and PASS (Model #8).

The convergence process of the SSE over 380 training examples while training Model #8 with different momentum and learning rates are shown in Figure 6.4. Note that one epoch is equal to the number of iterations for all training examples to be entered in one cycle (e.g., 1 epoch = 344 iterations for a stop-based ANN). Figure 6.4(a) shows that a small learning rate (e.g., η=0.05) associated with a large momentum rate (e.g., γ=0.95) accelerates the convergent speed of SSE. Figure 6.4(b) demonstrates that for a large learning rate (e.g., η=1.5), a large momentum rate (e.g., γ=0.95) yields larger SSE compared with that obtained by using smaller momentum rates (e.g., γ=0.1 and 0.6).

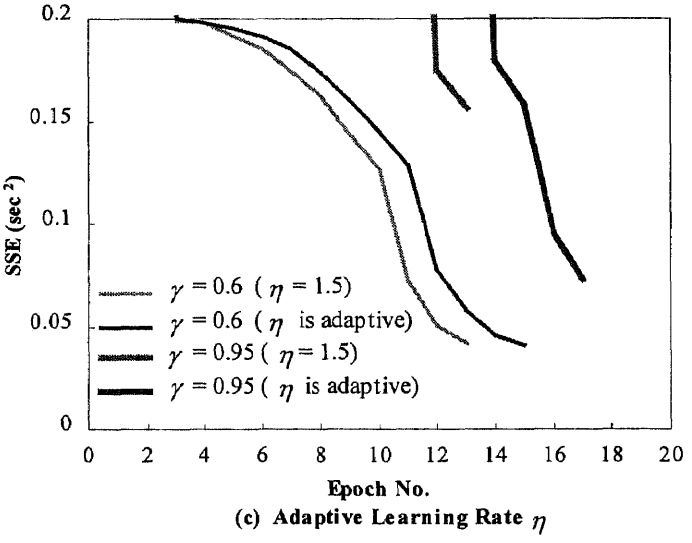
The convergence process during the last few epochs with the adaptive learning rate (Hagan, et. al, 1996) is shown in Figure 6.4(c), in which η proportionally decreases from an initial value (initialized as 1.5) with the decrease of SSE. Compared with the constant learning rate η=1.5, the adaptive learning rate may slow down the convergent speed but yield a lower SSE, especially when the momentum rate is large (e.g., γ=0.95). This indicates that the adaptive learning rate may help the BP algorithm to search efficiently thus locate a smaller local minimum.



(a) Learning Rate $\eta = 0.05$



(b) Learning Rate $\eta = 1.5$



(c) Adaptive Learning Rate η

Figure 6.4 Different learning /momentum rates for a stop-based ANN (Model #8)

6.2.2 ANN Evaluation

For evaluating the performance of two ANNs (Model #1 and #8), both of them are integrated with the simulation program individually. The arrival times of 24 buses (different from those for training purpose) at each of the 13 stops on the analyzed segment of NJ Transit Route #39 are predicted. Figure 6.5 shows the time-space diagram for the bus dispatched at 1322 simulation seconds (8:22:02AM). Both models demonstrate desirable accuracy with the predicted values. The accuracy analysis of the stop-to-stop travel times predicted by the two ANNs is plotted in Figure 6.6. The figure indicates that the stop-based ANN may capture the mean stop-to-stop travel times more accurately than the link-based one, especially at pairs of stops with longer distance (e.g., stops #1-2 and #11-12).

To analyze the accuracy of predicted arrival times at stops, the resulting RMSE for the 24 buses is shown in Figure 6.7. The prediction error for both ANNs increases steadily as the stop number increases. Moreover, the RMSE of the link-based ANN increases sharply at stop #12. The outperformance of stop-based ANN indicates that it accommodates traffic conditions at these stops better than the link-based one because of the aggregate characteristics of its inputs. In Figure 6.7, the RMSE for both ANNs decreases a little starting at downstream stop #13, indicating that prediction errors at stops may compensate one another thus decrease slightly at a certain downstream stop.

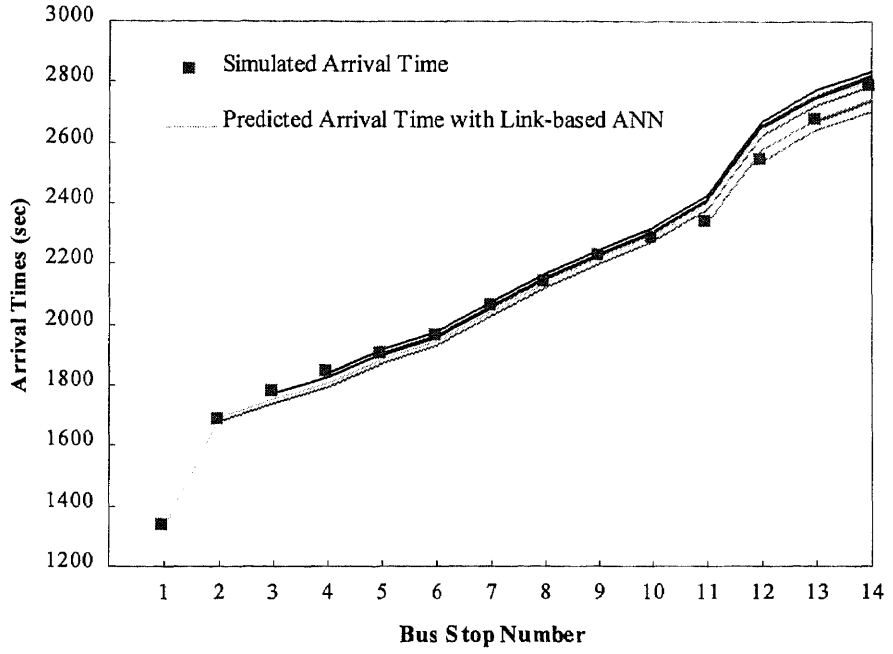


Figure 6.5(a) Predicted vs. simulated bus arrival times for a bus (link-based ANN)

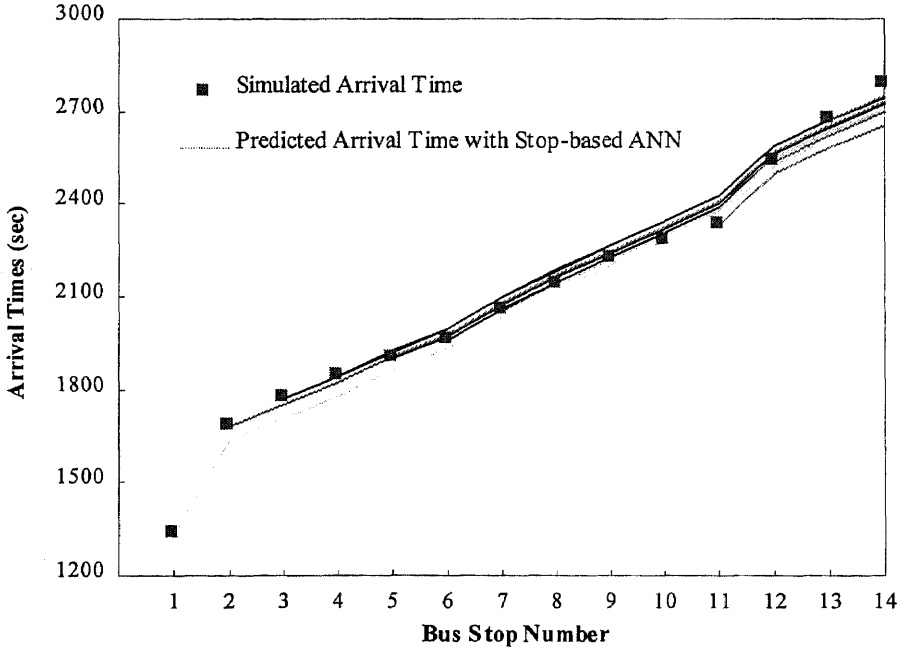


Figure 6.5(b) Predicted vs. simulated bus arrival times for a bus (stop-based ANN)

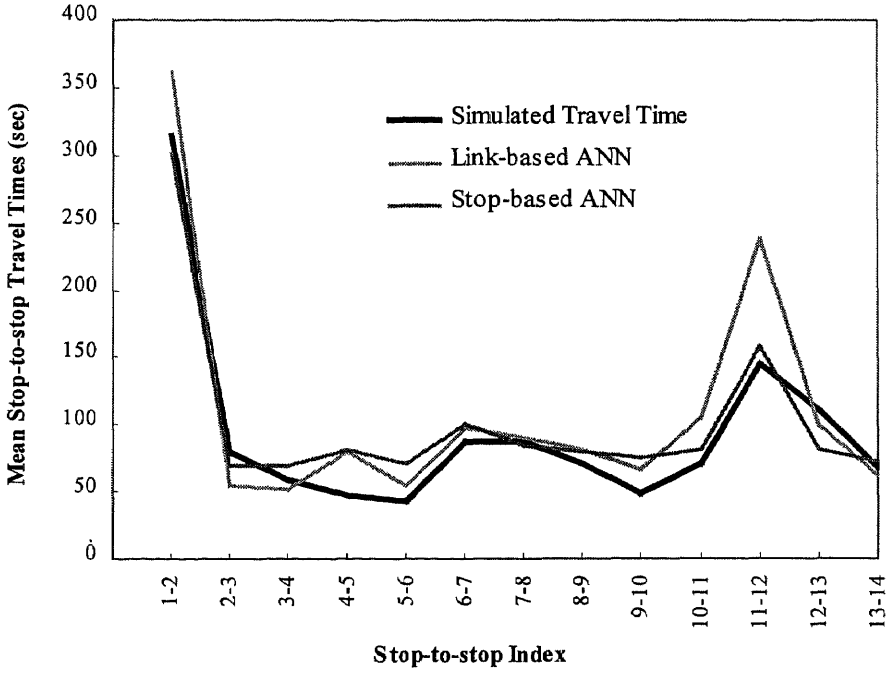


Figure 6.6 Mean stop-to-stop travel times by link-based and stop-based ANNs

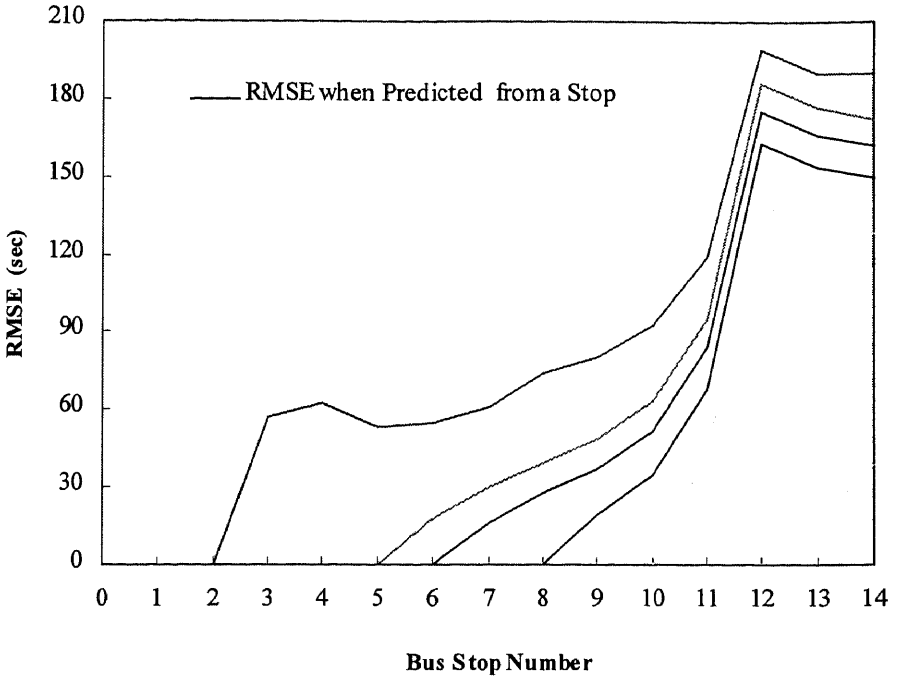


Figure 6.7(a) The RMSE for predicted arrival times at different stops (link-based ANN)

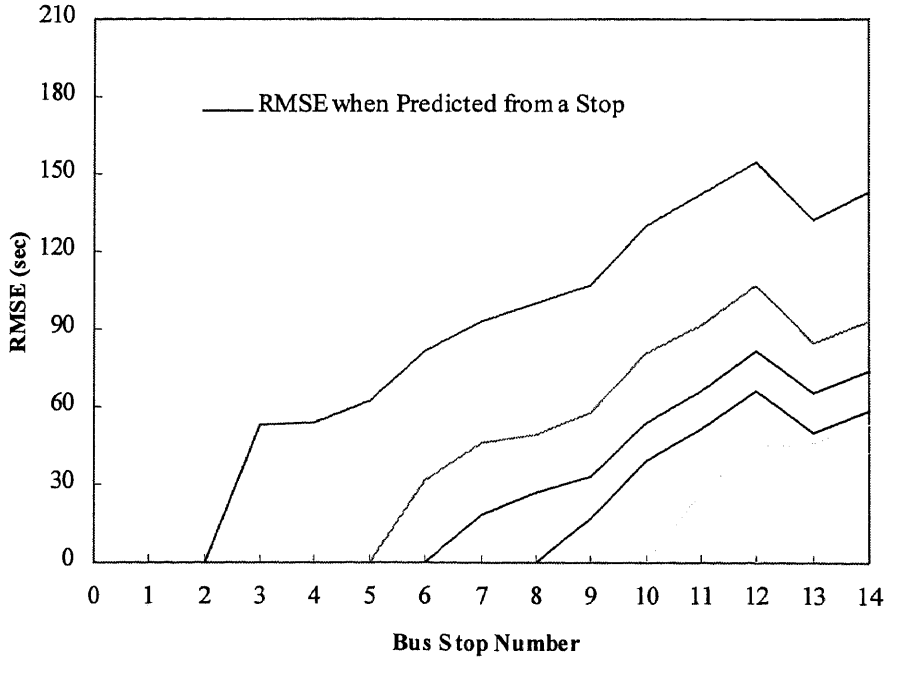


Figure 6.7(b) The RMSE for predicted arrival times at different stops (stop-based ANN)

Because the numbers of intersections between different pairs of stops may vary, the impact of intersections on predicted bus travel times is explored. Table 6.5 shows the resulting RMSE of predicted stop-to-stop travel times calculated from the simulation results of 24 buses. For a pair of stops with only one intersection in between, the link-based ANN performs better than the stop-based one (e.g., at stops#2-3, #3-4, #5-6, #9-10 and #13-14). However, as the number of intersections increases, the RMSE for the link-based ANN substantially exceeds that for the stop-based ANN, especially at stops#1-2, and #11-12. This indicates that the link-based ANN can not adapt to traffic variations at stops with a large number of intersections between them.

Table 6.5 RMSE of stop-to-stop travel times by link-based and stop-based ANNs

Stop-to-stop Index	Stop-to-stop Distance (ft)	# of Intersections	RMSE (sec)		
			Link-based ANN	Stop-based ANN	Kalman Filtering Model
1-2	4466	8	82.50	33.91	77.69
2-3	1003	1	18.06	28.84	29.30
3-4	1003	1	13.43	27.76	20.12
4-5	1056	1	43.83	44.23	21.23
5-6	422	1	23.69	39.73	17.56
6-7	1478	1	28.96	31.44	33.50
7-8	1312	1	32.83	29.00	31.83
8-9	980	1	31.75	30.27	29.79
9-10	581	1	30.68	40.87	32.95
10-11	552	2	67.67	47.55	37.85
11-12	1806	5	121.34	44.17	84.73
12-13	598	2	43.04	28.32	35.91
13-14	422	1	34.94	42.83	35.67

Table 6.5 also shows that at some stops with fewer intersections between them (e.g., stops #2-3, #3-4, #6-7, #9-10 and #13-14), the link-based ANN generally has equal or better performance than the Kalman filtering model. However, at stops with a large number of intersections (e.g., stops #1-2, #11-12), the Kalman filtering model performs better than the link-based ANN but still worse than the stop-based ANN.

It is difficult to accurately predict travel times for a pair of stops with a number of intersections because the bus travel time is affected by the traffic conditions not only on individual links, but also at intersections. Both the Kalman filtering model and the link-based ANN can not respond to such relation adequately. In the stop-based ANN, such relation is adapted while the traffic conditions aggregated over all links between a pair of stops rather than on individual links are applied for training ANNs. Thus, it can perform well when the distance and the number of intersections between stops are large. However, the stop-based ANN can not perform as well as the other two models for a pair of stops with a small number of intersections, because it is not very sensitive to travel times on individual links due to the aggregate characteristic of its inputs.

6.3 Test Neural/Dynamic (ND) Prediction Models

According to the discussion in Section 6.2, the link-based ANN only performs well for some pairs of stops with a smaller number of intersections between them, while the stop-based ANN performs better for pairs of stops with a large number of intersections in between. To improve the performance of ANNs under various conditions, the Neural/Dynamic (ND) models are thus designed by integrating the well-trained ANNs

with the Kalman filtering model, such that the prediction error can be adjusted in real-time. In the developed ND models, the ANN (either link-based or stop-based) captures the relationship between travel times and various affecting factors adequately, while the Kalman filtering model adjusts the results predicted by the ANNs in real-time. The two ND models, NDL (link-based) and NDS (stop-based), are also evaluated by simulation.

In the ND models, the initial values of the Jacobian factor $\Theta_{0,i}(-)$ (i = 1, 2,..., 14) are chosen to be 1, while the covariance error $M_{0,i}(-)$ (i = 1, 2,..., 14) ranges between 75 sec^2 and 125 sec^2 based on the analysis of the convergence speed of the prediction error (Tavantzis and Ding, 1999). Bus operations on the analyzed segment of NJ Transit Route #39 during a peak period (7:30-9:30AM) are simulated, while the prediction results for 24 buses at each stop during simulation are analyzed. The accuracy comparison of bus stop-to-stop travel times predicted by the link-based ANN and the NDL are shown in Figure 6.8. It is demonstrated that by implementing the NDL, the discrepancy between the simulated and predicted mean stop-to-stop travel times is significantly reduced, especially at stops #1-2, #10-11 and #11-12. Similarly, Figure 6.9 shows that the NDS also outperforms the stop-based ANN (e.g., at stops #3-4, #4-5, #5-6 and #9-10).

For prediction of arrival times by both ND models, the resulting RMSE are shown in Figure 6.10. Comparing with Figure 6.7 presented in Section 6.2, we found that both ND models perform better than their ANN counterparts by applying the Kalman filtering algorithm. For example, at stop#14, the RMSE of bus arrival times predicted from stop#2 decreased from 208 and 166 seconds (by the two ANNs in Figure 6.7) to 108 and 100 seconds (by NDL and NDS, as shown in Figure 6.10), respectively.

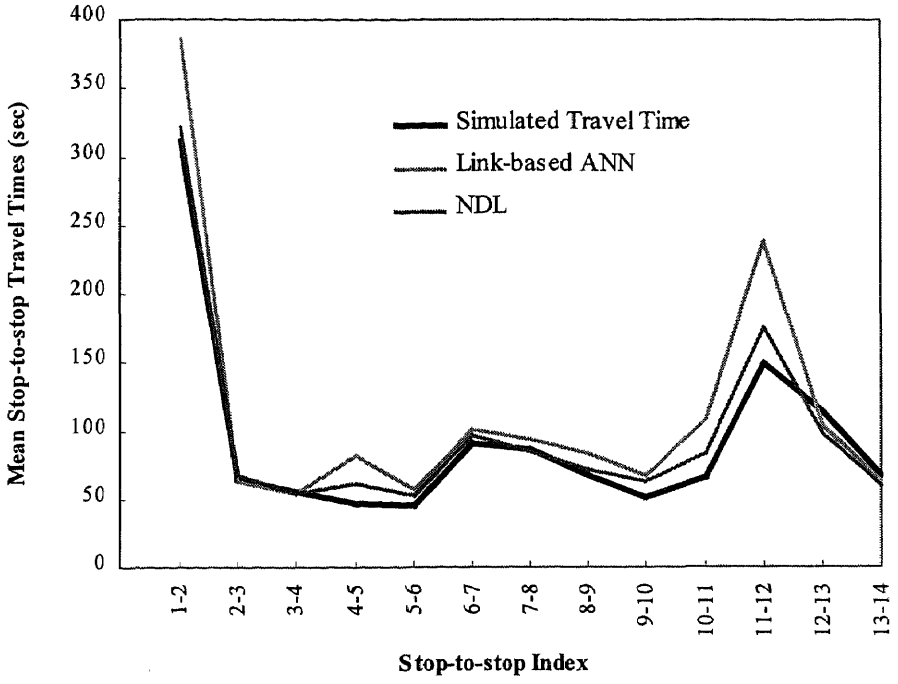


Figure 6.8 Mean stop-to-stop travel times by link-based ANN and NDL

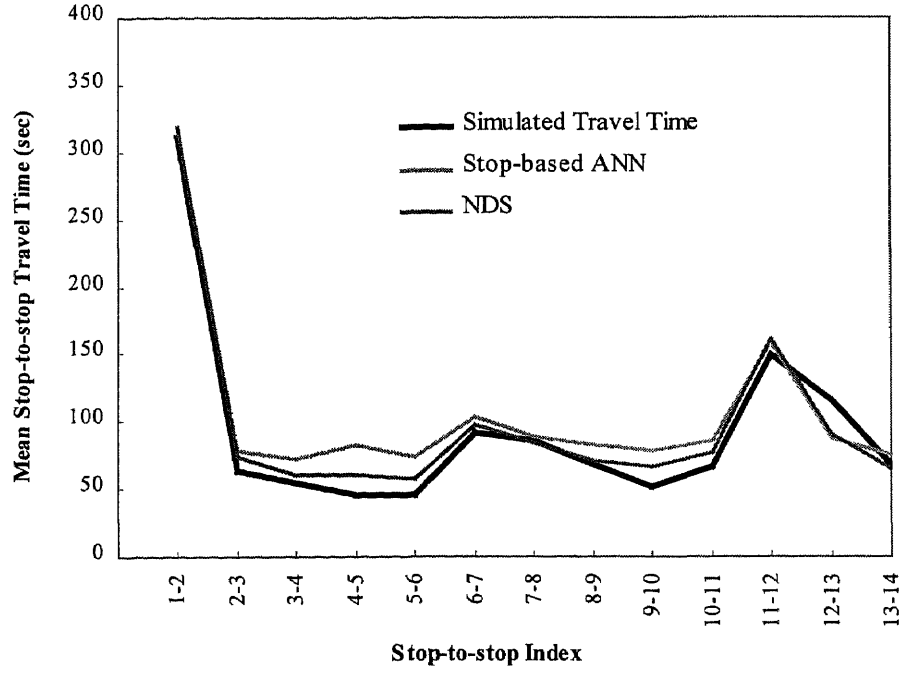


Figure 6.9 Mean stop-to-stop travel times by stop-based ANN and NDS

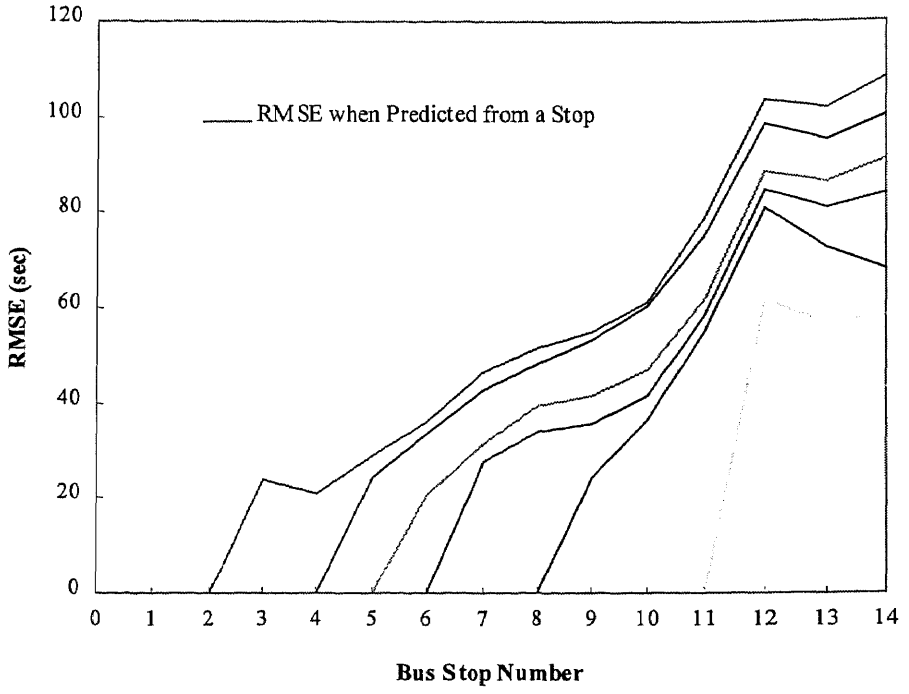


Figure 6.10(a) The RMSE for predicted arrival times at different stops (NDL)

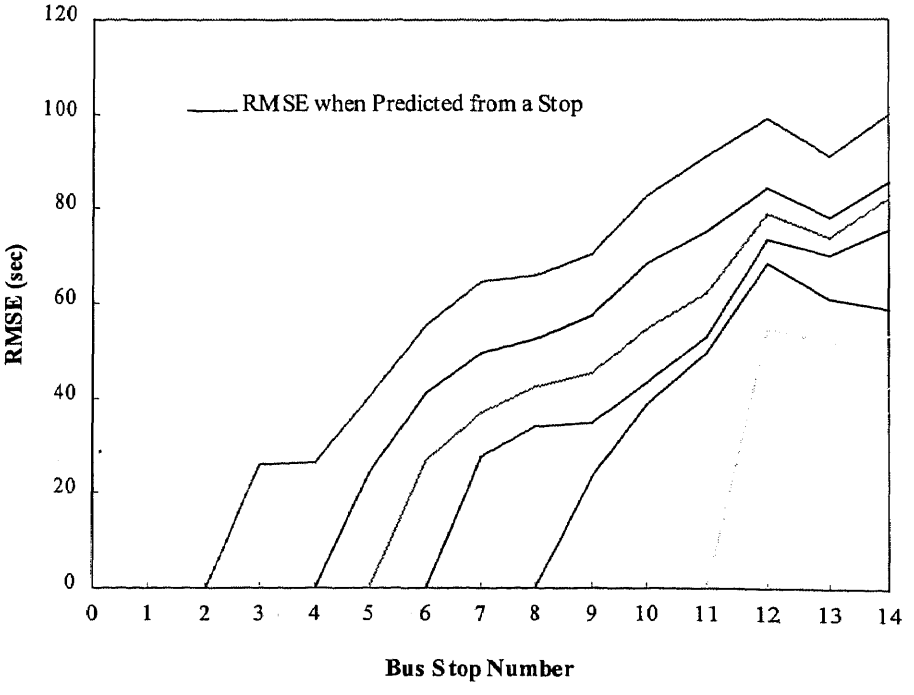


Figure 6.10(b) The RMSE for predicted arrival times at different stops (NDS)

The RMSEs of predicted travel times between different pairs of stops are calculated and listed in Table 6.6, where the impact of the number of intersections

Table 6.6 RMSE of bus stop-to-stop travel times by the ND Models

Stop-to-stop Index	Stop-to-stop Distance (ft)	# of Intersections	RMSE (sec)	
			NDL	NDS
1-2	4466	8	73.82	37.75
2-3	1003	1	24.25	26.43
3-4	1003	1	14.81	20.14
4-5	1056	1	24.97	24.79
5-6	422	1	21.06	27.56
6-7	1478	1	28.18	28.37
7-8	1312	1	26.48	27.03
8-9	980	1	24.84	23.95
9-10	581	1	28.62	31.84
10-11	552	2	46.35	39.68
11-12	1806	5	62.86	55.85
12-13	598	2	37.74	33.48
13-14	422	1	35.98	40.08

between a pair of stops on the accuracy of predicted bus travel times emerges. At stops with a large number of intersections in between (e.g., stops#1-2 and #11-12), the NDS predicts more accurately than the NDL. However, the difference between the RMSEs of the prediction arrival times by the NDL and NDS (e.g., only 7 seconds at stops#11-12) is not as significant as that between the link-based and stop-based ANNs (e.g., 73 seconds at stops#11-12) as shown in Table 6.5. This is because the RMSEs for the link-based ANN at these stops has been decreased substantially (e.g., from 121 to 63 seconds at stops #11-12 for the NDL). When the number of intersections decreases (e.g., at stops#2-3, #4-5, #6-7, #7-8 and #8-9), the NDL performs only slightly better than the NDS,

because the RMSEs for the NDS at these stops has been deceased effectively (e.g., from 44 to 25 seconds at stops #4-5).

The resulting RMSEs of the predicted bus arrival times from stop #1 to all downstream stops are shown in Figure 6.11. We found that both ND models predict more accurately than their ANN counterparts, with lower RMSE at nearly all stops. The figure also demonstrates that the RMSEs for the two ND models is very close from stops #2 to #8; however, at further downstream stops (e.g., stops #9 through #14), the RMSE of the NDS is slightly lower than that of the NDL. This indicates that, compared with NDL, NDS has equal or better capability to adapt to the variations in traffic conditions at downstream stops.

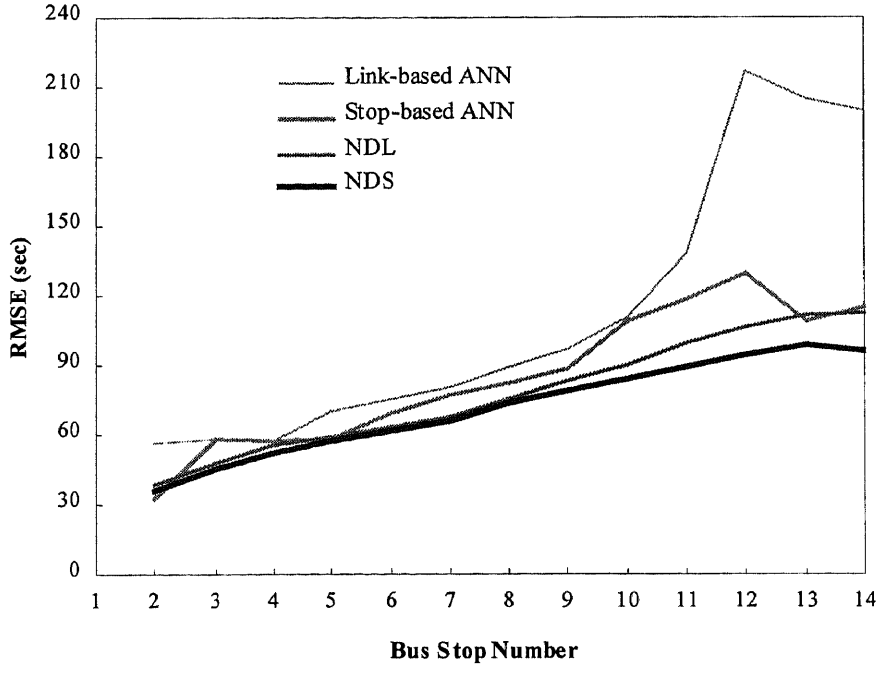


Figure 6.11 The RMSE for bus arrival times predicted from Stop# 1 to all stops

Summary of the Developed Prediction Models

The performance of the six developed transit arrival prediction models, including the basic, Kalman filtering, two ANNs (link-based and stop-based) and two ND models (NDL and NDS) are summarized in Tables 6.7 through 6.8. Table 6.8 demonstrates that by integrating the ANN with the Kalman filtering model, the prediction error for both ND models can be decreased substantially, especially at stops #10 through #14. In Table 6.9, the prediction error for the two ND models in different pairs of stops is either decreased effectively or remained similar, indicating that adjusting prediction error in real-time with the Kalman filtering model will not degrade but rather enhance the model performance. Tables 6.8 and 6.9 also demonstrate that the NDS performs a little better than the NDL at further downstream stops with large number of intersections in between.

Table 6.7 Summary of input variables used in developed prediction models

Model Name	Model No.	Input Variables
Basic	a	LDIS, LSPD, PASS
Kalman Filtering	b	SDIS, SSPD, PASS
Link-based ANN	c	LDIS, LVOL, LSPD, LDLY, PASS
Stop-based ANN	d	SDIS, SVOL, SDLY, SINTE, PASS
NDL	e	LDIS, LVOL, LSPD, LDLY, PASS
NDS	f	SDIS, SVOL, SDLY, SINTE, PASS

Table 6.8 The RMSEfor bus arrival times predicted from Stop#1 at all stops

Stop No	Model a (sec)	Model b (sec)	Model c (sec)	Model d (sec)	Model e (sec)	Model f (sec)
2	151.5081	76.05	56.49	32.33	38.30	36.12
3	156.7332	79.75	58.32	58.25	48.22	45.30
4	167.4026	82.21	57.21	57.13	55.34	52.38
5	187.0724	90.85	70.36	58.00	59.50	57.13
6	184.2975	90.49	75.61	69.17	63.29	61.53
7	208.229	102.62	80.80	76.82	68.10	66.40
8	239.7997	116.37	89.48	82.64	75.60	73.36
9	252.1499	127.33	96.46	88.59	82.91	78.97
10	281.2846	139.88	110.46	108.70	90.19	84.32
11	288.336	143.73	138.10	118.65	99.01	88.85
12	365.8051	193.35	216.48	129.27	106.55	93.93
13	350.0673	174.10	205.01	109.07	111.34	98.91
14	348.4922	175.68	199.53	114.62	112.05	95.95

Table 6.9 RMSE of bus stop-to-stop travel times

Stop-to-stop Index	Distance (ft)	# of Intersections	Model a (sec)	Model b (sec)	Model c (sec)	Model d (sec)	Model e (sec)	Model f (sec)
1-2	4466	8	154.77	77.69	82.50	33.91	73.82	37.75
2-3	1003	1	33.80	29.30	18.06	28.84	24.25	26.43
3-4	1003	1	28.63	20.12	13.43	27.76	14.81	20.14
4-5	1056	1	33.89	21.23	43.83	44.23	24.97	24.79
5-6	422	1	27.85	17.56	23.69	39.73	21.06	27.56
6-7	1478	1	51.07	33.50	28.96	31.44	28.18	28.37
7-8	1312	1	59.18	31.83	32.83	29.00	26.48	27.03
8-9	980	1	35.51	29.79	31.75	30.27	24.84	23.95
9-10	581	1	47.58	32.95	30.68	40.87	28.62	31.84
10-11	552	2	47.20	37.85	67.67	47.55	46.35	39.68
11-12	1806	5	118.86	84.73	121.34	44.17	62.86	55.85
12-13	598	2	47.81	35.91	43.04	28.32	37.74	33.48
13-14	422	1	55.44	35.69	34.94	42.83	35.98	40.08

6.4 Test the Real-time Vehicle Control Model

In Chapter 4, a real-time headway control model was developed to maintain desired headways between any pair of vehicles for high frequency light rail transit (LRT) systems. The model determines vehicle departure times in real-time based on its optimal arrival time at the next stop, while considering the constraint of the maximum attainable operating speed and the headways to its leading and following vehicles. The developed headway control model is tested by simulating a light rail system - Newark City Subway[2], which is 4.3 miles long with 11 stations, running from Newark Penn Station to Franklin Avenue Station, as shown in Figure 6.12.

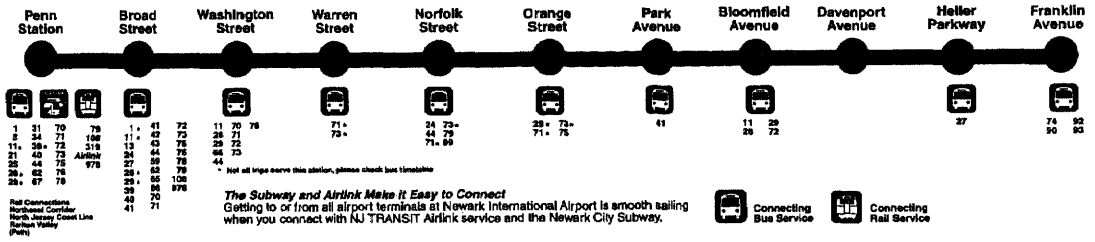


Figure 6.12 Newark City Subway and connecting bus and rail routes

Transit units (TU) in use in the subway system are PCC (Presidential Conference Committee) cars, and fares were collected on-board when passengers boarding or alighting from cars. The LRT operational characteristics include 2-minute headway in peak periods, 1.7-second average passenger boarding/alighting time, 50 mph maximum operating speed, 3.5 mphps acceleration/deceleration rate and 4.75 mphps emergency braking rate (NJ Transit, 1997). Passenger arrival rates and O/D demand at different

[2] The subway operational data in this study are collected from NJ Transit and on site during 1997-1998.

subway stations during peak periods are obtained from a field survey conducted by NJ Transit (1997). The high frequency LRT route with heavy and unevenly distributed ridership is a challenge to implement the headway control model.

The LRT operations in Newark City Subway during peak hour 5:00-6:00PM are simulated by the developed simulation program discussed in Chapter 5. The variations of passenger boarding/alighting rates in different time intervals (10minutes/interval) are assumed as shown in Figure 6.13, while the trajectories of consecutive trains with and

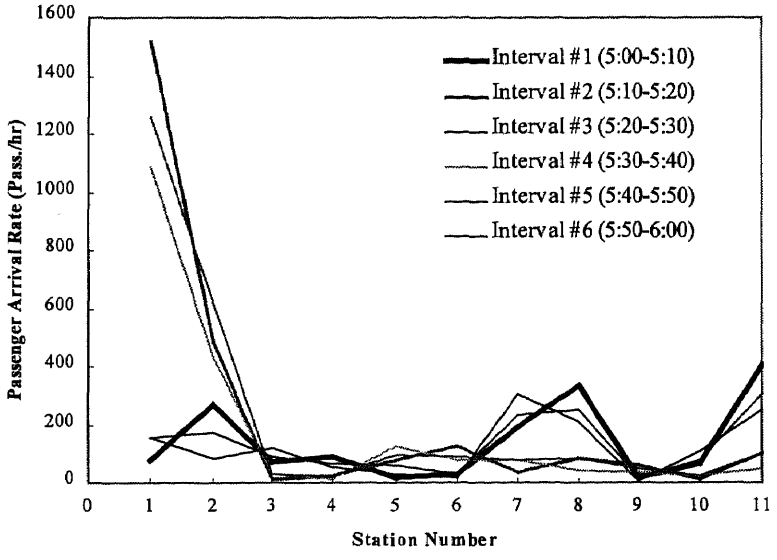


Figure 6.13 Passenger boarding/alighting rates at stations during 5:00-6:00PM

without headway control are shown in Figures 6.14(a) and 6.14(b). Figure 6.14(a) demonstrates that without dynamic headway control, the headways between trains #1 and #2, #6 and #7, #8 and #9 and #15 and #16 are irregular starting from stations #7 through #11. However, if the developed headway control is implemented, the headway irregularity at downstream stations can be alleviated by preventing the bunching of some pairs of trains, as shown in Figure 6.14(b).

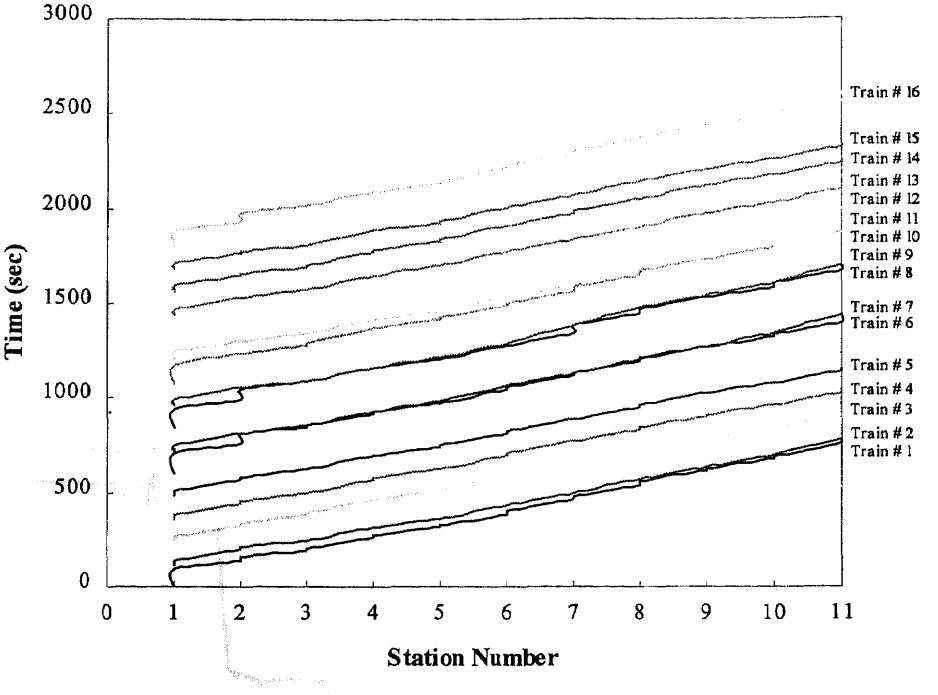


Figure 6.14(a) Trajectories of consecutive trains without headway control

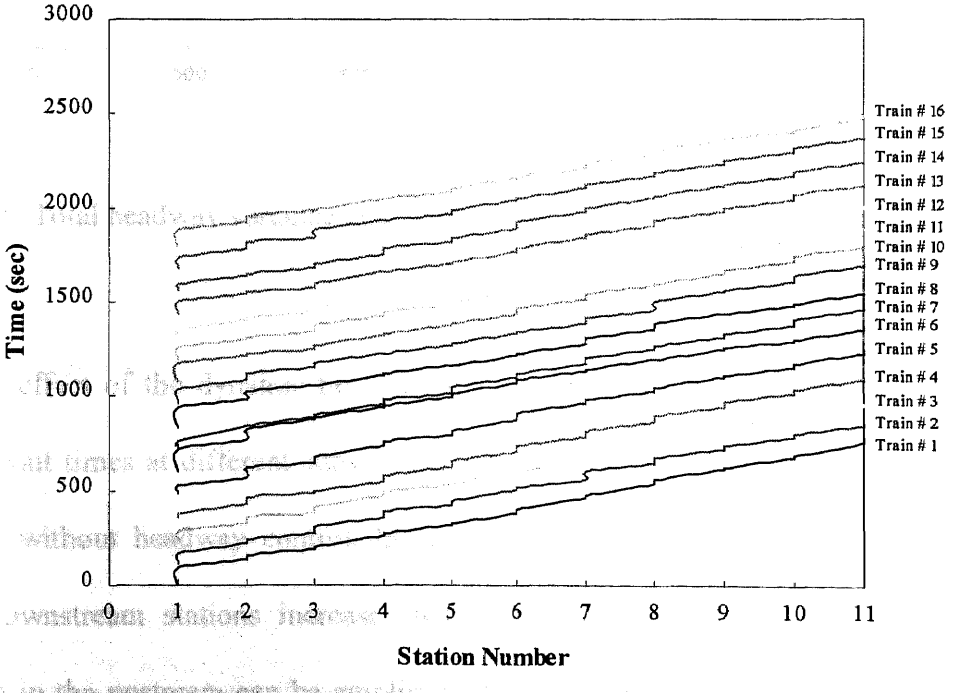


Figure 6.14 (b) Trajectories of consecutive trains with headway control

Figure 6.15 shows the variations of total headway variances and average passenger wait times during the simulation period. We found that with dynamic headway control, the total headway variance decreases from 6114 to 2368 second^2 while the average passenger wait time decreases from 83 to 66 second/passenger.

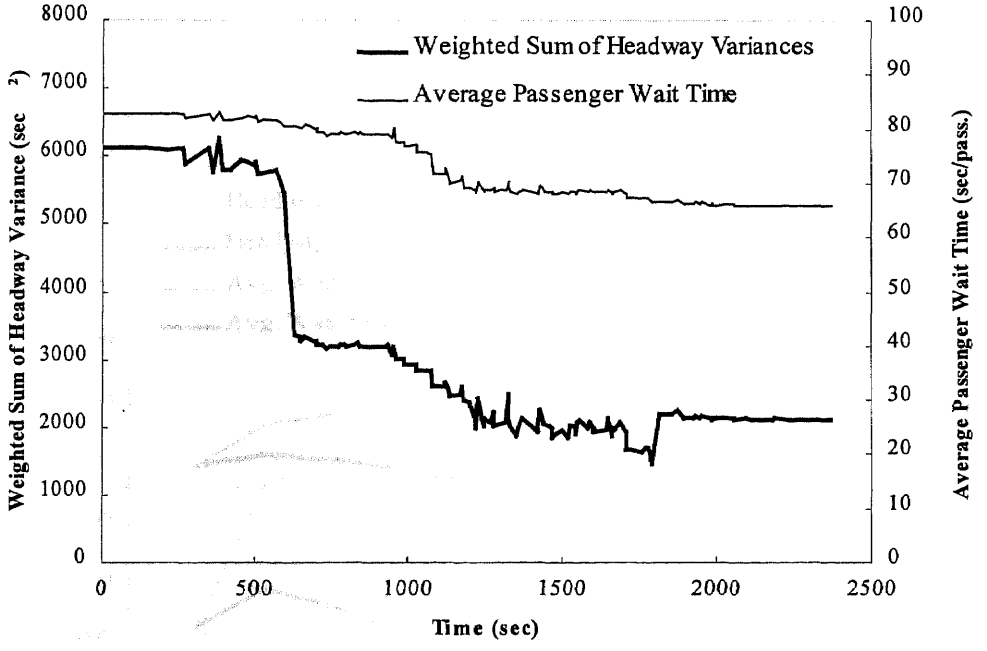


Figure 6.15 Total headway variance and average passenger wait time vs. simulation time

The effect of the dynamic headway control on headway variances and average passenger wait times at different subway stations are shown in Figure 6.16. The figure shows that without headway control, headway variances and average passenger wait times at downstream stations increase steadily. This arises because slight headway disturbance in the upstream can be amplified at a downstream segment, causing longer passenger wait times. However, after implementing headway control, the headway

variance decreases considerably, especially at downstream stations starting from #7 to #11. For example, the headway variance at station #11 decreases from 8463 to 944 second^2 (or the standard deviation of headways at the station decreases from 92 sec to 31 sec). Additionally, the average passenger wait times at further downstream stations are reduced (e.g., the average passenger wait time at station #10 decreases from 91 to 63 seconds).

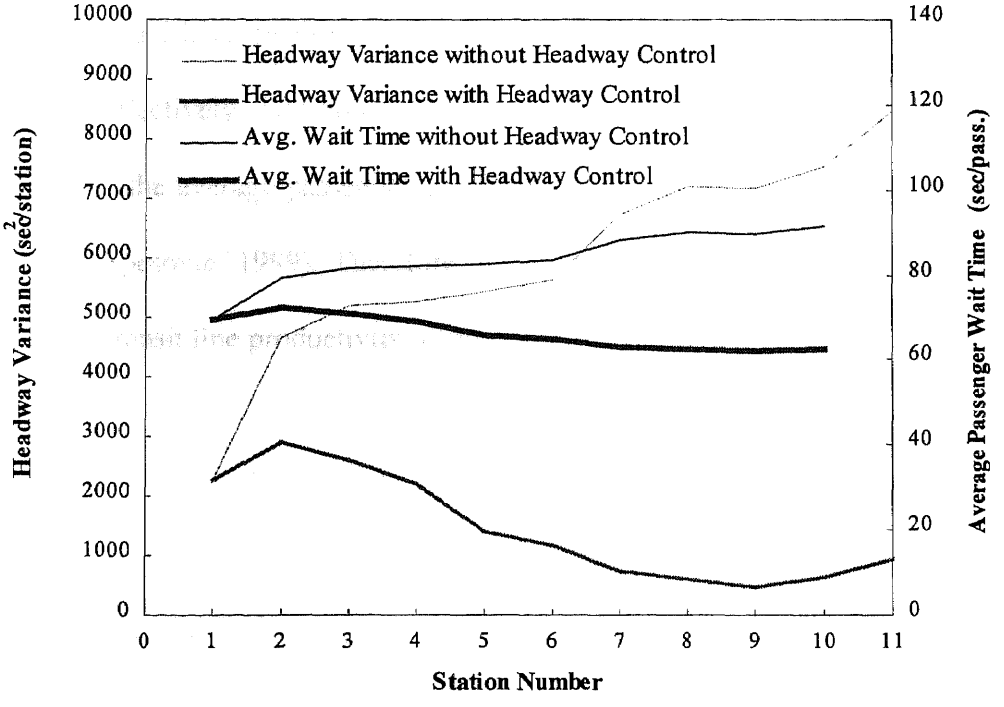


Figure 6.16 Headway variances and average passenger wait times at stations

The impact of dynamic headway control on transit productivity of the analyzed route was also studied. Transit productivity is often used to measure the passenger output carried by transit units and the distance traveled, which is highly related to the average vehicle operating speed and passenger demand (Vuchic, 1981). In this study, the product

of the number of operating vehicles (e.g., 16 trains in Figure 6.14) and their average speed is estimated (Vuchic, 1981), which is 349.44 veh-mile/hr without control and 349.48 veh-mile/hr with control. This indicates that the overall vehicle operating speed will not be influenced negatively under the proposed headway control. By analyzing Figure 6.14, we found that although some vehicles (e.g., Trains #5, #7 and #9) are slowed down deliberately by the headway control model for regulating headways, other vehicles (e.g., Trains #6, #8, #10 and #16) can be operated faster because the serious delays (e.g., signal delays and long dwell times at stations) experienced without the headway control are reduced effectively. Additionally, the decrease in the average passenger wait time will also decrease the average passenger travel time, thus, transit demand may be stimulated (Chien and Spasovic, 1999). Therefore, the proposed headway control is anticipated to increase the transit line productivity of the analyzed route.

CHAPTER 7

CONCLUSIONS

A major stochastic characteristic of transit operation is that vehicle arrivals tend to deviate from posted schedules. Poor schedule/headway adherence is undesirable for both users and operators since it increases passenger wait/transfer times and discourages passengers for using the transit system. Several prediction models developed in this study can provide vehicle arrival information. Such information can be disseminated to transit users through a variety of media, thus greatly helping them to schedule their departures and transfers at lower wait cost. Moreover, based on predicted information, real-time fleet management (e.g., on-demand fleet scheduling and vehicle routing) and vehicle control models (e.g., dispatching, headway control and signal priority control) can be developed for reducing service disturbance. Therefore, the quality of transit service can be enhanced.

Predicting bus arrival times in real-time is not an easy task, because it is affected by many stochastic factors such as travel times on links, dwell times at stops and delays at intersections. In this study, dynamic models for predicting arrival times of transit vehicles (buses) operating in urban settings are developed, including the basic model, Kalman filtering model, two artificial neural networks (ANNs) and two Neural/Dynamic (ND) models. The developed prediction models are evaluated using a simulation program which is an enhancement of CORSIM (CORridor SIMulator). An application of the prediction models on real-time vehicle control systems is also explored. Conclusions of the study are discussed in the following sections: (1) Model Development, (2) Model Evaluation, (3) Model Application, and (4) Suggestions for Further Research.

(1) Model Development

The developed basic model predicts vehicle arrival times at all downstream stops simply based on current average link speeds, while the Kalman filtering model can adjust model parameters (e.g., Kalman Gain) in real-time to reduce prediction error. Both models can not provide accurate prediction results because they lack the capability to adapt to dynamic changes in traffic conditions at downstream stops. Artificial Neural Networks (ANNs) are thus designed for accurately predicting transit arrivals. With a massive distributed structure (e.g., multilayer feedforward) and adaptive learning processes (e.g., Back-propagation (BP)), the developed ANNs (e.g., link-based and stop-based) have been well trained to capture time varied relationship between bus arrivals and various affecting factors including traffic conditions (e.g., volumes, speeds and delays) and passenger demands.

Of the two ANNs, the link-based ANN is designed to predict vehicle arrival times by accumulating vehicle travel times on all traversed links based on traffic conditions on individual links. However, the interaction of traffic conditions over different links on the vehicle arrival time is neglected. The stop-based ANN is therefore developed, using the aggregate traffic conditions (e.g., means and variances of link volumes, speeds and delays) over all links between pairs of stops as explanatory variables.

The BP training process for ANNs requires extensive experiments to search for the optimal values of model parameters (e.g., number of neurons on the hidden layer and synaptic weights linking successive layers). Such a lengthy training process is difficult to be conducted on-line. To enhance the performance of ANNs and their dynamic capability, Neural/Dynamic (ND) models have been developed by integrating both ANNs

with the Kalman filtering model individually. With the predicted information from ANNs being adjusted in real-time to reduce prediction errors, two ND models have demonstrated the capability to adapt to dynamic changes in the traffic environment over time and provide more accurate prediction results.

(2) Model Evaluation

The computer simulation approach has been extensively applied in this study to assess the performance of the developed prediction models while considering various transit demand and traffic conditions. CORSIM has been enhanced to simulate transit operations. New features have been built into the enhanced CORSIM to appropriately determine the vehicle dwell time and average passenger wait time according to a time dependent passenger boarding/alighting demand and the vehicle inter-departure time (headway). In addition, various emulated real-time data (e.g., vehicle arrival/departure times and number of boarding/alighting passengers at stops) and transit related MOEs (e.g., average passenger wait times at stops and vehicle journey times) have been generated, collected and analyzed during simulation.

To further ensure that the simulation output adequately represents real world traffic operations under existing conditions, the developed simulation program is validated through calibrating a segment of Route #39 of New Jersey Transit Corporation. Sufficient data related to transit operations collected during the simulation are used to develop link-based and stop-based ANNs. After conducting reliability analyses, we found that for the link-based ANN, the optimal input combination is the bus travel distances on individual links, average link volumes, speeds and delays, and the number of

boarding/alighting passengers at stops. On the other hand, for the stop-based ANN, the optimal input combination is stop-spacings, the number of intersections and means of volumes, speeds and delays on all the links between a pair of stops, and the number of boarding/alighting passengers at stops.

For evaluating the performance of the developed prediction models, the basic model, the Kalman filtering model, two ANN models and two ND models are tested individually by the enhanced CORSIM to compare the difference between simulated and predicted vehicle arrival times. The results demonstrate that both ANNs outperform the basic and Kalman filtering models. Moreover, the stop-based ANN performs better than the link-based one between stops with a large number of intersections. This is because in the stop-based ANN, the aggregate traffic conditions (e.g., means and variations of volumes, speeds and delays) over all links between pairs of stops are used in ANN training. Thus, the impact of intersections between a pair of stops on the vehicle travel time can be considered. The results also show that at a pair of stops with few intersections in between, the link-based ANN performs well, while the stop-based ANN does not show promising prediction results due to the aggregate characteristics of the input.

The reliability analysis of the two ND models has indicated that by integrating the ANN with the Kalman filtering model, the prediction errors (e.g., RMSE) at pairs of stops can be decreased substantially (e.g., from 121 to 63 seconds at stops #11-12 for the link-based ANN and from 44 to 25 seconds at stops #4-5 for the stop-based ANN). This indicates that the ND models will not degrade but rather enhance the ANN performance. The reliability analysis also demonstrates that the NDS performs a little better than the

NDL at a pair of stops with long stop-spacing and large number of intersections (e.g., RMSE of 56 seconds for NDS and 63 seconds for NDL at stops #11-12).

(3) Model Application

In this study, the application of the developed prediction models to a real-time headway control model has been explored and experimented through simulating a high frequency light rail transit (LRT) route. The headway control model is designed to maintain desired headways for any pair of successive vehicles through minimizing the total headway variance over all stops. The optimal departure time of a vehicle from a stop is determined by its optimal arrival time at the next stop, while the maximum attainable operating speeds, the headways to its leading and following vehicles, and the predicted arrive and departure times at downstream stops are considered.

The simulation results have demonstrated that the developed real-time headway control model regulates headways effectively (e.g., the average headway variance at station #11 decreases from 8463 to 944 second^2) and reduces average passenger wait times (e.g., the average passenger wait time at station #10 decreases from 91 to 63 seconds). Moreover, the results indicate that the average operating speed will not decrease (e.g., 349.44 veh-mile/hr without control and 349.48 veh-mile/hr with control). The ridership may be increased due to the decrease in passenger wait times (Chien and Spasovic, 1999). Therefore, the transit line productivity is anticipated to increase if the headway control model is used.

(4) Suggestions for Further Research

Further research for this study may focus on the following three aspects: (a) Prediction Technologies, (b) Control Technologies and (c) Transit Simulation. The future extensions in each aspect are itemized and discussed below.

Prediction Technologies

1. The developed ANNs and ND models are all well-trained with data collected from simulation results. With various real-time data and MOEs collected from transit surveillance and monitoring systems in the advent of APTS (e.g., GPS, APCS, AVLS), these models can be trained the same way, thus adapting to real-world transit operations in different networks.
2. In real transit systems, vehicles will experience unusual delays due to inclement weather conditions (especially during winter times), accidents or incidents (e.g., construction activities), which is undesirable to both transit operators and users. The prediction models need to be trained with data collected under these occurrences in order to incorporate the impacts of such short-term events on vehicle travel times.
3. In this study, passenger arrivals at stops are assumed to follow Poisson distributions. However, passenger arrival behavior may be altered because of the implementation of prediction models and dissemination of vehicle arrival information to the public. Thus, further analysis of the passenger arrival distribution is necessary, as discussed by Jolliffe and Hutchinson (1975) and Turquist and Bowman (1980), in order to depict more accurately the impact of passenger demands at stops on vehicle arrival times.

Control Technologies

1. A rather general headway control model has been described in this study to illustrate positive benefits of the applications of the developed prediction models. However, details of the implementation of such a model would vary from system to system.
2. The analysis of passenger demand elasticity as well as the transit line productivity should be conducted by considering both the benefits (e.g., savings in passenger wait times) and cost (e.g., on-board passenger delays) of such control.
3. In the developed vehicle control model, those stops with high passenger demand are highlighted. In further studies, alternative control strategies can be used, such as controlling vehicles before a stop with high demand, thus benefiting more passengers from the reduction of the wait time.
4. The headway control model could be enhanced to deal with bus transit by considering the impact of mixed automobile traffic on transit operations.

Transit Simulation

1. The simulation program could be extended to accept more passenger arrival distributions defined by users to fit specific transit systems.
2. The network could be enlarged while more data can be collected for further improving the accuracy of the prediction models considering various operating conditions, such as harsh weather or lane closures due to accidents.
3. The vehicle capacity (spaces/bus) could be considered, thus incorporating additional transit MOEs, such as passenger comfort level and in-vehicle times to differentiate among bus routes operating in transit networks.

APPENDIX A

SECOND DERIVATIVE OF TOTAL HEADWAY VARIANCE

The second derivative of total headway variance $\Pi^{(t)}$ is studied to determine the convexity of the objective total headway variance $\Pi^{(t)}$ with respect to the optimal vehicle arrival time $e_{k,i}^{(t)}$. The second derivative of $\Pi^{(t)}$ can be obtained from Eq. A-1 as

$$\frac{\partial^2 \Pi^{(t)}}{\partial (e_{k,i}^{(t)})^2} = \sum_{j=1}^{S} w_j \frac{\partial^2 \pi_j^{(t)}}{\partial (e_{k,i}^{(t)})^2} \tag{A-1}$$

where $\frac{\partial^2 \Pi^{(t)}}{\partial (e_{k,i}^{(t)})^2}$ is the second derivative of headway variance at stop i, which is derived from the first derivative $\frac{\partial \pi_j^{(t)}}{\partial e_{k,i}^{(t)}}$ as indicated below.

Situation 1

When vehicle k is ready to depart from stop i-1, the second derivatives of headway variance at stops j = 1, 2, ...,i-1, can be derived directly from Eq. 4.12 as

$$\frac{\partial^2 \pi_j^{(t)}}{\partial (e_{k,i}^{(t)})^2} = 0 \qquad (j = 1, 2,..., i-1) \tag{A-2}$$

Situation 2

The second derivative of headway variance at the next stop i with respect to $e_{k,i}$ can be derived from the first derivative in Eq. 4.13 as

$$\frac{\partial^2 \pi_j^{(t)}}{\partial (e_{k,i}^{(t)})^2} = \frac{2}{N-1}[\varphi_j^2 + \varphi_j^4 + \sum_{n=k+2}^{N} (t_b \lambda_i)^{2(N-k-1)} \varphi_i^4] \quad (j = i) \tag{A-3}$$

Situation 3

The second derivatives of headway variance at downstream stops j = i+1,..., S, can be derived from Eq. 4.16 as

$$\frac{\partial^2 \pi_j^{(l)}}{\partial (e_{k,i}^{(l)})^2} = \frac{2\varphi_i^2 \varphi_i^2 \cdots \varphi_{j-1}^2}{N-1}[\varphi_j^2 + \varphi_j^4 + \sum_{n=k+2}^{N}(t_b \lambda_j)^{2(N-k-1)} \varphi_j^4] \quad (j = i+1,..., S) \quad (A\text{-}4)$$

Thus, $\frac{\partial^2 \Pi^{(l)}}{\partial (e_{k,i}^{(l)})^2}$ can be obtained by the summation of Eqs. A-2, A-3 and A-4. Obviously, for any $e_{k,i}^{(l)}$, $\frac{\partial^2 \Pi^{(l)}}{\partial (e_{k,i}^{(l)})^2}$ is positive, which denotes that the total headway variance $\Pi^{(l)}$ is convex with respect to $e_{k,i}^{(l)}$.

APPENDIX B

NOTATIONS

The following notations are used in this dissertation.

Variable	Description	Unit
$a_{k,i}$	arrival time of vehicle k at stop i	hour
$A_{k,i}$	actual arrival time for bus k at stop i	hour
$d_{k,i}$	dwell time of vehicle k at stop i time t	hour
$\Delta d_{k,i,1}$	first dwell time interval of $d_{k,i}$	hour
$\Delta d_{k,i,n}$	n^{th} dwell time internal of $d_{k,i}$	hour
e_r	sum of square error in BP learning	hour^2
$e_0^{(t)}$	minimum of total headway variance $\Pi^{(t)}$	hour
$e_{k,i}^{(t)}$	optimal arrival time for vehicle k at next stop i time t	hour
$E_{k,i}^{(t)}$	predicted arrival time for bus k at stop i at time t	hour
$h_{k,i}^{(t)}$	headway between two vehicles k-1 and k at stop i time t	hour
$L_{k,i}^{(t)}$	the earliest arrival time that vehicle k can make at stop i time t	hour
l_j	length of link j between two consecutive stops	mile
$p_{k,i}$	departure time of vehicle k at stop i time t	hour
$\overline{q}_{i,j}$	average boarding rate from stop i to j	pass./hour
$\overline{q}_{aj}$	average passenger alighting rate at stop j	pass./hour
$\overline{q}_{ib}$	average passenger boarding rate at stop i	pass./hour
$Q_{i,j}$	number of in-vehicle passengers destining at downstream stop j	pass.
$\Delta Q_{i+1,j}$	number of passengers boarding at stop i+1 destining at downstream stop j	pass.
Q_{ai}	total number of alighting passengers at stop i	pass.

Variable	Description	Unit
Q_{ib}	total number of boarding passengers at stop i	pass.
t_b	average passenger boarding time	hour
$v(t)$	random noise with $v(t) \sim N(0, R(t))$	-
$\overline{V}_j^t$	average speed on link j at time t	mph
w_i	weight factor for stop i	-
$W_{k,i}$	total passenger wait time for vehicle k at stop i	hour
$\overline{w}_{k,i}$	average passenger wait time for vehicle k at stop i	hour/pass.
$x_j^{[l]}$	the output for jth neuron on the l^{th} layer	-
y_p	desired output for pth example in BP learning	hour
$\hat{y}_p$	ANN model output for pth example in BP learning	hour
z	ratio of the wait time to the vehicle inter-departure time	-
$\delta_i^{[l]}$	local delta for the jth neuron on the l^{th} layer	-
$\Delta_{k,i}$	the deviation between $h_{k,i}$ and H	hour
γ	momentum rate in BP algorithm	-
$\Gamma_{k,j}^{(t)}$	travel time for bus k on link j at time t	hour
η	learning rate in BP algorithm	-
φ_i	a factor related to passenger arrival rate at stop i	-
$\Phi_L(\cdot)$	predicted bus stop-to-stop travel time by link-based ANN	hour
$\Phi_S(\cdot)$	predicted bus stop-to-stop travel time by stop-based ANN	hour
λ_i	the passenger arrival rate at stop i	pass./hour
$\Lambda_{k,i}^{(t)}$	estimated travel time for vehicle k traveling from stop i-1 to i at time t	hour
$\pi_i^{(t)}$	headway variance at stop i time t	hour^2
$\Pi^{(t)}$	the total headway variance for all stops on the analyzed route at time t	hour^2
$\rho_{k,i}$	difference between actual and predicted travel times for bus k from stops i-1 to i	hour

Variable	Description	Unit
$\omega_{j,i}^{[l]}$	the synaptic weight between the ith neuron on l^{th} layer and the jth neuron on the $(l+1)^{th}$ layer	-
ξ_i	distance between stop i and the downstream intersection	mile
$\mathbf{B}_{k,i}$	vehicle O/D matrix for vehicle k traveling from stop i to i+1	-
$\mathbf{K}_{k,i}^{(l)}$	Kalman Gain matrix for bus k at stop i	-
$\mathbf{M}_{k,i}$	error covariance matrix for bus k at stop i in Kalman filtering model	-
$\mathbf{O}$	stop O/D matrix for a route	-
$\mathbf{w}(t)$	random noise vector with $\mathbf{w}(t) \sim N(0, \mathbf{Q}(t))$	-
$\mathbf{X}_j(t)$	vector containing factors that affect bus travel times on link j in the Link-based ANN	-
$\mathbf{Y}(t)$	vector containing factors that affect bus travel times from stop i–1 to i in the Kalman filtering model	
$\mathbf{Z}(t)$	vector containing factors that affect bus travel times from stop i–1 to i in the Stop-based ANN	-
$\mathbf{\Theta}_{k,i}$	Jacobian vector for bus k at stop i in Kalman filtering model	-
$\hat{\mathbf{\Theta}}_{k,i}$	Optimal estimation of Jacobian vector $\mathbf{\Theta}_{k,i}$	-

REFERENCES

Abdelfattah, A. M., and Khan, A. M. (1998). "Models for predicting bus delays." Transportation Research Record 1623, 8-15.

Abkowitz, M., and Engelestein, I. (1984). "Method for maintaining transit service regularity." TRB, Transportation Research Record 961, 1-8.

Abkowitz, M., and Engelstein. I. (1983). "Factors affecting running time on transit routes." Transportation Research, 17A(2), 107-113.

Abkowitz, M., Eiger, A., and Engelestein, I. (1986). "Optimal control of headway variation on transit routes." Journal of Advanced Transportation, 20(1), 73-88.

Adamski, A., and Turnau, A. (1998). "Simulation support tool for teal-time dispatching control in public transport." Transportation Research, 32A(2), 73-87.

Adebisi, O. (1986). "A mathematical model for headway variance of fixed-route buses." Transportation Research, 20B(1), 59-70.

Brockwell, P. J., and Davis, R. A. (1991). Time series: theory and methods. Springer-Verlag, New York, NY.

Chang, G., and Su, C. (1995). "Predicting intersection queue with neural network models." Transportation Research, 3C(3), 175-191.

Chien, S. (1996). Class notes for mass transportation systems. New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Chowdhury, M. (1997). Enhancement of the CORCIM model in simulating transit vehicle operations. Project #421580, Final Report, New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Ding, Y. (1998). Microscopic transit simulation models. Project #421580, Final Report, New Jersey Institute of Technology, Newark, NJ.

Chien, S., and Ding, Y. (1999a). "Applications of artificial neural networks in prediction of transit arrival times." Proceedings, 1999 Annual Conference of ITS America, Washington, D.C.

Chien, S., and Ding, Y. (1999b). "A dynamic headway control strategy for transit operations." Proceedings, the 6th ITS World Congress, Toronto, Canada.

Chien, S., Chowdhury, M, Mouskos, K. C., and Ding, Y. (1999c). "Enhancements of the CORSIM model in simulating transit vehicle operations." Accepted to the Journal of Transportation Engineering, American Society of Civil Engineers.

Chien, S., Ding, Y., and Zayas, N. A. (1999d). "Validation of enhanced CORSIM for simulating transit operations." Working Paper #98-07, New Jersey Institute of Technology, Newark, NJ.

Ding, Y., Chien, S., and Zayas, N. (1999e). "Analysis of bus transit operations with enhanced CORSIM, case study: bus route #39 of New Jersey Transit." Proceedings, the 79th Transportation Research Board Annual Meeting, Washington, D.C.

Ding, Y., and Chien, S. (1999f). "The prediction of bus arrival times with link-based artificial neural networks." Proceedings, Joint Conference on Information Sciences (JCIS), Computational Intelligence & Neurosciences (CI&N), Atlantic City, NJ.

Chien, S., and Spasovic, L. N. (1999). "Optimization of grid transit systems with elastic demand." submitted to Journal of Advanced Transportation, American Society of Civil Engineers.

Chien, S., and Schonfeld, P. (1997). "Optimization of grid transit system in heterogeneous urban environment." Journal of Transportation Engineering, American Society of Civil Engineers, 123(1), 28-35.

Chin, S., Hwang, H., and Pei, T. (1994). "Using neural networks to synthesize origin-destination flows in a traffic circle." TRB, Transportation Research Record 1457, 134-142.

Darken, C. and Moody, J. (1992). "Towards faster stochastic gradient search." Advances in Neural Information Processing Systems 4, 1009-1016.

Delurgio, S. A. (1998). Forecasting principles and applications. McGraw-Hill, New York, NY.

Dougherty, M. S. (1995). "A review of neural networks applied to transport." Transportation Research, 3C(4), 247-260.

Dougherty, M. S., Kirby, H. R., and Boule, R. D. (1993). "The use of neural networks to recognize and predict traffic congestion." Traffic Engineering Control, 34(6), 311-314.

Eberlein, X. J., Wilson, N. H. M., Barnhart, C., and Bernstein, D. (1998). "The real-time deadheading problem in transit operations control." Transportation Research, 32B(2), 77-100.

Federal Highway Administration. (1992). TRAF users' reference guide. Report No. FHWA-RD-92-103, Washington, D.C.

Federal Highway Administration. (1996). CORSIM user's manual. Contract No. DTFH61-92-Z-00074, Washington, D.C.

Federal Transit Administration. (1998). "Advanced public transportation systems: the state of the art, update'98." U. S. Department of Transportation, Washington, D.C.

Gelb, A., Kasper, J. F. Jr., Nash, R. A., Price, C. F. Jr., and Sutherland, A. A. Jr. (1977). Applied optimal estimation. MIT Press, Massachusetts Institute of Technology, Cambridge, MA.

Guenthner. R. P., and Hamat, K. (1988). "Distribution of bus transit on-time performance." TRB, Transportation Research Record 1202, 1-8.

Guenthner. R. P., and Sinha, K. C. (1983). "Maintenance, schedule reliability and transit system performance." Transportation Research, 17A(5), 355-362.

Hagan, M. T., Demuth, H. B., and Beale, M. (1996). Neural network design. PWS Publishing Company, Boston, MD.

Ho, E., Chien, S., and Ting, C. (1999). "A hybrid modeling method for the planning and evaluation of intelligent transportation systems." Accepted to the Journal of Transportation Planning and Technology, Leicestershire, UK.

Hua, J., and Faghri A. (1994). "Application of ANNs to IVHS." TRB, Transportation Research Record 1453, 83-90.

Jolliffe, J. K., and Hutchinson, T. P. (1975). "A behavioral explanation of the association between bus and passenger arrivals at a bus stop", Transportation Science, 9(4), 248-282.

Kalaputapu, R., and Demetsky, M. J. (1995). "Application of artificial neural networks and automatic vehicle location data for bus transit schedule behavior modeling." TRB, Transportation Research Record 1497, 44-52.

Kohonen, T. (1988). Self-organization and associative memory. Springer-Verlag, Berlin, Germany.

Kwon, E., and Stephanedes, Y. J. (1994). "Comparative evaluation of adaptive and neural-network exit demand prediction for freeway control." TRB, Transportation Research Record 1446, 66-76.

Lin, G., Liang, P., Schonfeld, P., and Larson, R. (1995). Adaptive control of transit operations. Report No. MD-26-7002, University of Maryland, College Park, MD.

Lin, T., and Wilson, N. H. M. (1992). "Dwell time relationships for light rail systems." TRB, Transportation Research Record 1361, 287-295.

Lin, W., and Zeng, J. (1998). "An experimental study on real time bus arrival time prediction with GPS data." Proceedings, the 77th Transportation Research Board Annual Meeting, Washington, D.C.

Maren, A. J., Harston, C. T., and Pap, R. M. (1990). Handbook of neural computing applications. Academic Press, San Diego, CA.

O'Flaherty, C. A., and Mangan, D. O. (1970). "Bus passenger waiting times in central areas." Traffic Engineering and Control, 11(9), 419-421.

Okutani, I., and Stephanedes, Y.J. (1984). "Dynamic prediction of traffic volume through Kalman filtering theory." Transportation Research, 18B(1), 1-11.

Osuna, E. E., and Newell, G. F. (1972). "Control strategies for an idealized public transportation system." Transportation Science, 6(1), 52-72.

Polus, A. (1979). "An analysis of the headway distribution of an urban bus service." Traffic Engineering and Control 20(9), 419-421.

Prevedouros, R. D., and Wang, Y. (1998). "Simulation of a large freeway/arterial network with CORSIM, INTEGRATION and WATSim." Proceedings, the 77th Transportation Research Board Annual Meeting, Washington, D.C.

Skinner, R. J. (1980). "Bus planning method: service reliability." Traffic Engineering and Control, 21(9), 554-558.

Railway Technical Research Institute. (1998). An integrated train control system based on digital ATP. Transportation Systems Development Department, Tokyo, Japan.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning internal representations by error propagation." Parallel Distributed Processing 1, MIT Press, Massachusetts Institute of Technology, Cambridge, MA, 318-362.

Saddon, P. A., and Day, G. F. (1974). "Bus passenger waiting times in greater Manchester." Traffic Engineering and Control, 6(9), 442-445.

Smith, B. L., and Demetsky, M. J. (1995). "Short-term traffic flow prediction: neural network approach." TRB, Transportation Research Record 1453, 98-104.

Stephanedes, Y. J., Kwon, E., and Michalopoulos, P. (1990). "On-line diversion prediction for dynamic control and vehicle guidance in freeway corridors." TRB, Transportation Research Record 1287, 11-19.

Talley, W. K., and Becher, A. J. (1987). "On-time performance and the exponential probability distribution." TRB, Transportation Research Record 1108, 22-26.

Tavantzis, J., and Ding, Y. (1999). Independent study: implementation of Kalman filtering models in travel time prediction. New Jersey Institute of Technology, Newark, NJ.

Turnquist, M. A. (1981). "Strategies for improving reliability of bus transit service." TRB, Transportation Research Record 818, 8-11.

Turnquist, M. A. (1978). "A model for investigating the effects of service frequency and reliability on bus passenger waiting time." TRB, Transportation Research Record 663, 70-73.

Turquist, M. A., and Bowman, L. A. (1980). "The effect of network structure on reliability of transit service." Transportation Research, 14B(1), 79-86.

Vandebona, U., and Richardson, A. J. (1981). "The effect of priority signals on transit route performance." Transportation Reference of National Conference 81, the Institution of Engineers, Australia, 56-62.

Venglar, S., Fambro, D., and Bauer, T. (1995). "Validation of simulation software for modeling light rail transit." Proceedings, the 74th Transportation Research Board Annual Meeting, Washington, D.C.

Vuchic, V. R. (1981). Urban public transportation: System and Technology. Prentice-Hall, Englewood Cliffs, NJ.

Wall, A., and Dailey, D. J. (1998). "An algorithm for predicting the arrival time of mass transit vehicles using Automatic Vehicle Location data." Proceedings, the 77th Transportation Research Board Annual Meeting, Washington, D.C.

Warner, B., and Misra, M. (1996). "Understanding neural networks as statistical tools." The American Statistician, 50(4), 284-293.

Wei, C., and Yang, Y. (1998). "Assessing input layer elements of artificial neural networks-an empirical study for transit containers forecast in Kaohsiung port." Submitted to Journal of the Chinese Institute of Transportation, Taiwan.

Wei, C., and Wu, K. (1997). "Developing intelligent freeway ramp metering control systems.", Proceedings of National Science Council, 7C(3), Taiwan, 371-389.

Welding, P. I. (1957). "The instability of close interval service." Operational Research Quarterly, 8(3), 133-148.

Williams, B. M., Durvasula, P. K., and Brown, D. E. (1998). "Urban freeway traffic flow prediction: application of seasonal ARIMA and exponential smoothing models." Proceedings, the 77th Transportation Research Board Annual Meeting, Washington, D.C.

Zhang, H., Ritchie, S. G., and Lo, Z. (1997). "Macroscopic modeling of freeway traffic using an artificial neural network", TRB, Transportation Research Record 1588, 110-119.