

DEVELOPMENT AND EVALUATION OF TRAFFIC PREDICTION SYSTEMS

by

Changkyun Kim

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University


in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

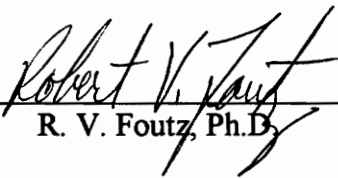
in

Civil Engineering

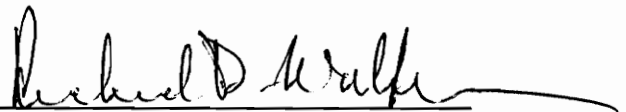
APPROVED:



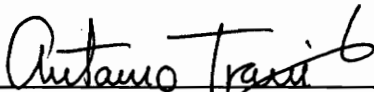
A. G. Hobeika, Ph.D., Chairman



R. V. Foutz, Ph.D.



R. D. Walker, Ph.D.



A. A. Trani, Ph.D.



R. Sivanandan, Ph.D.

April 12, 1994
Blacksburg, Virginia

C.7

LD
5655
V856
1994
K56
C.2

DEVELOPMENT AND EVALUATION OF TRAFFIC PREDICTION SYSTEMS

by

Changkyun Kim

A.G. Hobeika, Ph.D., Chairman

Civil Engineering

(ABSTRACT)

Developing real-time traffic diversion strategies is a major issue of Advanced Traffic Management Systems (ATMS), a component of Intelligent Vehicle Highway Systems (IVHS). Traffic diversion utilizes available capacity in the urban network during a congestion-causing event. If an alternative route selected for diversion is not congested at the current time, a certain part of the route may become congested by the time the diverted drivers reach that part of the network. Thus the ability to forecast future traffic variables on each link along various routes in a prompt and accurate fashion may be necessary to ensure the success of a diversion strategy. Forecasting future traffic variables would also be helpful for urban traffic control. In addition, the forecasting model may help assign the vehicles onto the alternate roads, if the information on driver destinations and the routes between a diversion point and the destinations are available.

This dissertation is aimed at developing and evaluating two prediction models : link-based model and network-based model. The link-based prediction model has two components. One component is an Auto Regressive Integrated Moving Average (ARIMA) time series model based on the latest (current) traffic data. The other

component is the smoothed historical traffic volume (historical average) for that period as obtained from previous days. These two components are combined to represent the dynamic fluctuations in the traffic flow behavior. The combined model is designed to produce the predicted traffic volumes for a look-ahead period of 30 minutes, divided into 6-minute time intervals. The results show that the combined model is promising for light to medium congested traffic conditions

The network-based prediction model combines current traffic, historical average, and upstream traffic. It is presumed that traffic volume on the upstream can be used to predict the downstream traffic in a specific time period. Three prediction models are developed for traffic prediction: a combination of historical average and upstream traffic, a combination of current traffic and upstream traffic, and a combination of all three variables. The three models are evaluated through regression analysis. The third model is found to be the most applicable while the first model was the least. In order to consider current traffic conditions, a heuristic adaptive weighting system is devised based on the relationships between the origin of prediction and the previous periods. The developed models are applied to real freeway data in 15-minute time interval measured by regular induction loop detectors. The prediction models are shown to be capable of producing reliable and accurate forecasts under congested traffic condition. The prediction systems perform better in the 15-minute range than in the ranges of 30-minute or 45-minute. It is noted that the combined models usually produce more consistent forecasts than the historical average.

ACKNOWLEDGMENTS

The author is indebted to Dr. Antoine. G. Hobeika for many helpful suggestions, guidance, support, and encouragement at all stages of his study. The author would like to extend his appreciation to the following persons, who served as members of his dissertation committee: Dr. Robert V. Foutz for educating me on the time-series analysis and his comments on the statistical analysis; Dr. Antonio A. Trani for his advice and suggestions; Dr. R. Sivanandan and Dr. Richard D. Walker for their assistance and comments on the heuristic prediction models. He is also grateful to Los Angeles Department of Transportation (DOT) and New York DOT (INFORM project) for providing the traffic data.

The author wishes to express very special thanks to his wife, Aejin, and daughter, Nakyung, and son, Minjae, for their continuous support, encouragement, patience, and love during the entire period of his study here in the United States. He extends thanks to his parents and parent-in-law who encourage and support for his study. Without the support and love from his all family members, his doctoral degree would have not been achieved.

TABLE OF CONTENTS

1. INTRODUCTION	1
1.1 Problem Statements	2
1.2 Scope and Objective of Research	5
1.3 Proposed Approach	6
1.4 Research Tasks	7
1.5 Organization of Dissertation	9
2. LITERATURE REVIEW	11
2.1 Forecasting Models	12
2.1.1 Time-series analysis	12
<u>2.1.1.1 Univariate B-J time-series approach</u>	12
<u>2.1.1.2 Multivariate time-series approach</u>	14
2.1.2 Filtering theory	17
2.1.3 Heuristic Approaches	21
2.1.4 Prediction of travel time	29
2.1.5 Other statistical approach	31
2.1.6 Neural network approach	34
2.1.7 Weather forecasting	35
2.1.8 Use of time-series model for traffic flow analysis	38
2.1.9 The advantages and weaknesses of UBJ model	38
2.2 Time-Series Analysis	41
2.2.1 Univariate ARIMA model	41
<u>2.2.1.1 When may UBJ models used ?</u>	43
2.2.1.2 A basic structure of ARIMA model	43
2.2.1.3 The UBJ modeling procedure	45
2.2.1.4 Forecasting using the ARIMA model	50

2.2.2 Combined time-series model	52
3. LINK-BASED PREDICTION MODEL	54
3.1 Introduction	54
3.2 Optimal Time Interval for Diversion Strategies	56
3.3 A Universal ARIMA Model	57
3.4 Multi-Step-Ahead Forecasting	58
3.5 Combined Time-Series Model for Multi-Step Forecasting	58
3.6 Evaluation of the Model	64
3.6.1 Links used	64
3.6.2 Criteria for evaluation of the model	66
3.6.3 Comparison of modeled values and historical averages	67
3.7 Findings and Critiques	72
4. NETWORK-BASED PREDICTION MODEL	75
4.1 Introduction	75
4.1.1 Hypothesis for the network-based prediction model	76
4.1.2 Data used	77
4.2 Identification of Upstream Traffic	81
4.3 Multiple Combined Prediction Models	82
4.3.1 Parameters(a and b) for 'upstream' and 'current' components	84
4.3.2 The 'a' and 'b' parameters estimated from regression analysis	85
4.3.3 Comparison of the three prediction models	89
4.3.4 Comparison of the <i>upstream</i> parameters in prediction models	93
4.3.5 Adaptive heuristic weighting systems for real-time application	96
4.4 Evaluation of the Multiple Combined Models	107
4.4.1 Evaluation of the combined models with peak-hour traffic data	107
4.4.2 Evaluation of the combined models with non-peak-hour traffic data	112

4.4.3 Comparison of the historical averages bet. 11 days and 22 days	117
4.4.4 A final decision rule of the models for better prediction	120
4.4.5 Theoretical weighting vs. heuristic weighting	121
4.4.6 Evaluation of the multiple combined models in the network with closer distance between the links	128
4.4.7 Comparison of the combined model with other prediction model	131
4.5 Findings	133
5. CONCLUSIONS	135
5.1 Recommendations for Further Study	140
REFERENCES	142
VITA	147

LIST OF FIGURES

Figure 1-1	Evolution of the research questions	3
Figure 1-2	Conceptual model of research question	4
Figure 2-1	Adaptive prediction system	28
Figure 2-2	A simple neural model	34
Figure 2-3	The idea of UBJ forecasting	42
Figure 2-4	Stages in the B-J iterative approach to model building	46
Figure 2-5	Examples of ARMA model in ACF and PACF	48
Figure 2-6	Plots of stationary and non-stationary data sets over time	49
Figure 3-1	Procedures for dealing with a link-based prediction model	55
Figure 3-2	Traffic flow in links '199', '152', and '252' over 6-minute period	60
Figure 3-3	Traffic flow of four variables over forecasting steps	65
Figure 3-4	Comparisons of MAPEs on six W.S. for 7 data sets with differences between 13-30 percent	68
Figure 4-1	Traffic volume over links along the network at 7 A.M.	79
Figure 4-2	Traffic volume on link '176' over time	80
Figure 4-3	Prediction comparisons of three models in 15-minute forecasting	95
Figure 4-4	Prediction comparison of three forecasting ranges in model I	97

LIST OF TABLES

Table 3-1	A pseudo-Weighting System for combined model	62
Table 3-2	MAPEs and MSEs for six data sets with 25-40 percent margin	71
Table 4-1	Configuration of network used	78
Table 4-2	Parameters comparisons of the modeled values bet. regression model and intuitive weighting	87
Table 4-3	Characteristics of the regression models in the three prediction ranges	92
Table 4-4	The errors of the three models in the 15-minute forecasting	94
Table 4-5	All possible weighting scenarios for model I and II	99
Table 4-6	Adaptive weighting system for model I and II	101
Table 4-7	Adaptive weighting systems for third combined model in 15-minute forecasting	103
Table 4-8	Adaptive weighting systems for third combined model in 30-minute forecasting	105
Table 4-9	Adaptive weighting systems for third combined model in 45-minute forecasting	105
Table 4-10	Comparison of three models and historical averages in 15-minute forecasting with peak-hour traffic data bet. 6 A.M. and 9 A.M.	108
Table 4-11	Comparison of three models and historical averages in 30-minute forecasting with peak-hour traffic data bet. 6 A.M. and 9 A.M.	110
Table 4-12	Comparison of three models and historical averages in 45-minute forecasting with peak-hour traffic data bet. 6 A.M. and 9 A.M.	111
Table 4-13	Comparison of three models and historical averages in 15-minute forecasting with peak-hour traffic data bet. 11 A.M. and 3 P.M.	113

Table 4-14	Comparison of three models and historical averages in 30-minute forecasting with peak-hour traffic data bet. 11 A.M. and 3 P.M.	115
Table 4-15	Comparison of three models and historical averages in 15-minute forecasting with peak-hour traffic data bet. 5 P.M. and 8 P.M.	116
Table 4-16	Comparison of three models and historical averages in 30-minute forecasting with peak-hour traffic data bet. 5 P.M. and 8 P.M.	118
Table 4-17	Comparison of prediction performance between 11-day based historical average and 22-days based historical average	119
Table 4-18	Evaluation of the values by final decision in 15-minute forecasting with traffic data between 6 A.M. and 9 A.M.	122
Table 4-19	Evaluation of the values by final decision in 30-minute forecasting with traffic data between 6 A.M. and 9 A.M.	123
Table 4-20	An example showing the theoretical weighting system	124
Table 4-21	The results of the prediction performance based on theoretical system, heuristic system, and hist. avg. from 6:30 A.M. to 11:30 A.M.	127
Table 4-22	Configuration of the links added	128
Table 4-23	The prediction results from two types of network : <i>Detailed</i> and <i>Approximated</i>	130
Table 4-24	Comparison of combined models and the UTCS model	132

1. INTRODUCTION

There are many traffic management systems to relieve traffic congestion in urban areas, but relieving congestion in urban areas needs a strategy that not only increases road capacity, but also manages it. A strategy for dealing with congestion has to be built within a framework for implementing specific congestion-reduction measures. Planning for the effective diversion of traffic during congestion-causing events, a major issue of Advanced Traffic Management Systems (ATMS), could maximize the use of available capacities in the roadway systems, thus saving travel time and cost. However, when areawide diversion is attempted, pre-planning of diversion routes is frequently difficult due to the presence of multiple alternates. Once the alternate routes for diversion are determined, one or more routes among multiple alternate routes must be selected at a diversion point for the assignment of the vehicles. By utilizing real-time traffic information from roadway facilities, the traffic may be appropriately assigned on each route available. However, since traffic flow patterns change all the time, a control strategy based on the previous traffic flow pattern may no longer be good a few minutes later. For example, even if an alternate route selected for diversion is not congested at the current time, one part of the chosen route may be congested by the time the driver reaches that part of the network. Thus, forecasting future traffic flow variables at each link along the diversion routes is a required process for selecting the most efficient alternate route. If a prediction system representing the traffic flow-fluctuation over time could be developed, it would help

produce reliable forecasts for solving such route guidance problems. The term forecasting has been used interchangeably to the term prediction in this dissertation.

1.1 Problem Statements

Traffic flow is unpredictable due to such variables as incidents and traffic lights at intersections. Research on the attributes of congestion problems and traffic lights might investigate the general traffic flow pattern over various periods of time. The traffic flow pattern is crucial for predicting traffic. Prediction of the traffic will be essential for effective use of the links along the routes in the context of assigning the traffic.

The aim of this research is to develop and evaluate analytically the prediction models that are combined with the real traffic system. The research question evolved from an initial question of how the traffic flow is identified (see Figure 1-1). An intermediate question of how the prediction models are designed was asked based on the first question. Finally, a more realistic problem of how the prediction models perform within the real traffic system was identified as the research problem.

The research questions are clarified by a conceptual diagram, Figure 1-2, showing their components. The task of this research is an identification of variables for short-term traffic forecasting. The variables are identified throughout the comprehensive analysis of traffic-flow data as bases for building the prediction model. Second, the prediction models are built by including the variables identified in the first frame. The methodology of appropriately combining the selected variables is crucial for building the prediction

Initial
Question

How should the traffic
flow be identified ?

Intermediate
Question

How should the prediction
model be designed ?

Research
Question

How do the prediction
models perform within
the real traffic system ?

Figure 1-1 Evolution of the research questions

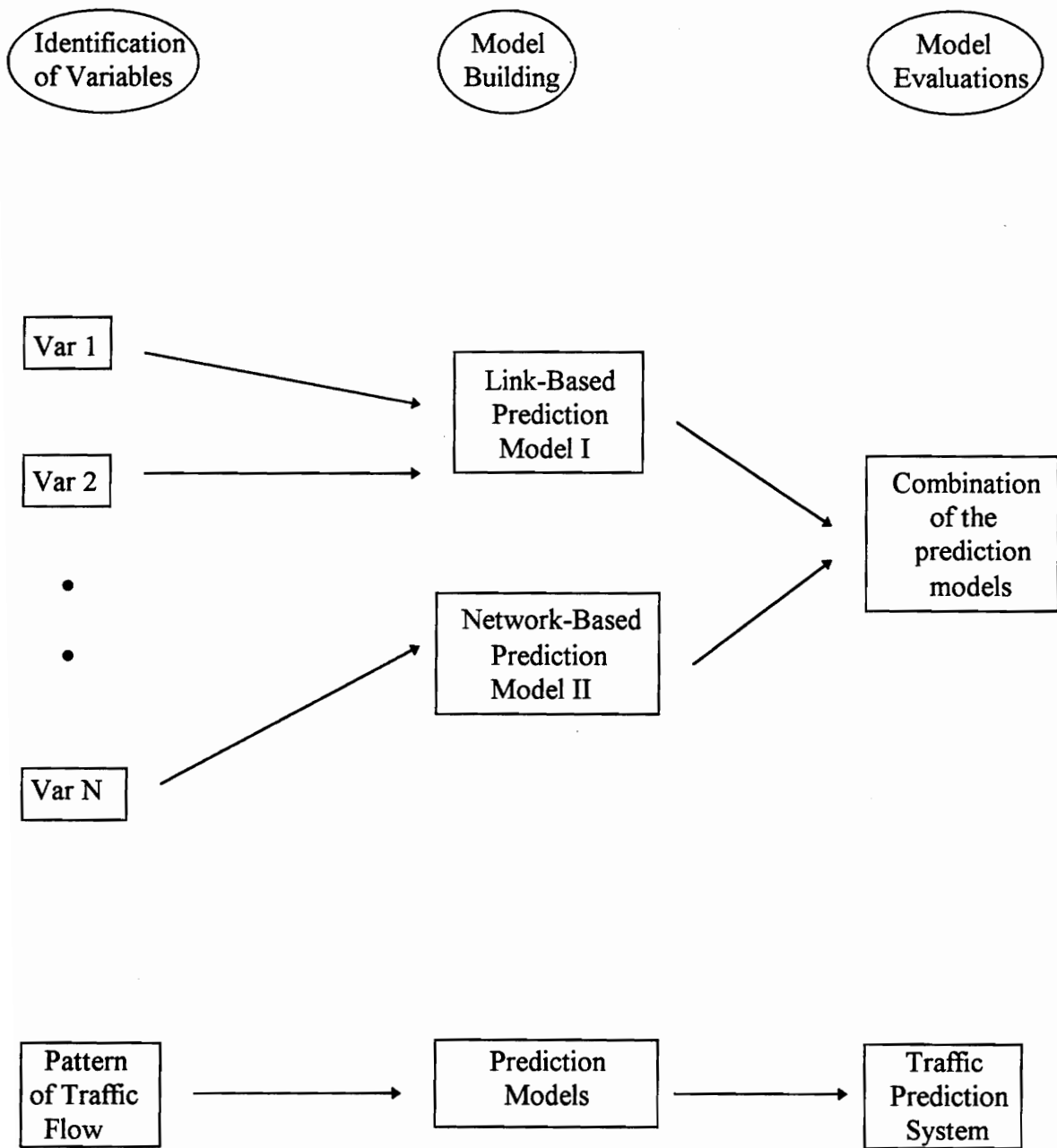


Figure 1-2 Conceptual model of research question

models. Also, the characteristics of each variable must be examined to assist that combination. The third component of this research is an evaluation of the prediction models in diverse way. By applying the prediction model, a realistic prediction system can be developed that considers diverse situations in the real-world traffic flow. From the main research question, a number of more specific subquestions can be stated as follows:

1. What approach is the most suitable for prediction models ?
2. What is the optimal time interval for multi-step forecasting ?
3. What are the appropriate measures for evaluating the effectiveness of the prediction model ?

1.2 Scope and Objective of Research

While the ultimate goal of the real-time diversion strategies is to relieve traffic congestion through the effective assignment algorithm, this work is aimed at the development and evaluation of prediction models. Two types of prediction models will be developed. The first model is based on an individual link considering historical averages and current traffic data. The second model describes the relationships between the continuous links through the network. The models developed in this research will be evaluated with the traffic data from roadway facilities. The use of those models will also be discussed. Some more specific objectives are :

- To identify variables that influence prediction of traffic flow.
- To empirically determine the optimal time interval for real-time diversion strategies.
- To determine the most appropriate periods for computing historical average value.
- To examine the quality of the adaptive weighting system of the network-based prediction model by comparing with the theoretical weighting system.

1.3 Proposed Approach

For predicting traffic along a route, the Box-Jenkins (B-J) ARIMA (Auto Regressive Integrated Moving Average) time-series model will be applied. Employing three sophisticated model-building steps produces forecasts based on past traffic data. This statistical model may be adaptable for real-time prediction due to its simplicity. However, the existing ARIMA forecasting models adapting traffic data are only efficient in respect to one-step ahead prediction values. Also, the ARIMA forecasting models focuses on individual link information. To overcome these disadvantages, heuristic approaches will be attempted.

Before developing a prediction model, an optimal time interval will be defined through the examination of traffic data. A description of the basic concept of the time-series model related to this research will also be included. In this research, the multi-step ahead predictions (i.e. for 30-minute, six-step forecasting with five time intervals) are

necessary to adapt the concept of the real-time diversion strategies and preview the dynamic flow of the traffic along the routes. That is, a prediction model has to represent the traffic flow throughout the network. It is assumed that Origin and Destination (O-D) is defined. In the first phase of the prediction model development, a model considers each link individually. This model combines the historical averages with predictions from ARIMA model. A weighting system is essential to obtain the multi-step forecasts. The weighting system is devised based on the analysis of the traffic flow. The predictions will be produced in terms of the time interval from this prediction model.

In the meantime, a second prediction model will be built considering the continuous links along the route. The basic idea in approaching this problem is to apply an idea of combined time-series model. The relationships between upstream traffic and downstream traffic will be carefully investigated. The comparative speeds on the upstream traffic may help predict the downstream traffic. Heuristic decision rules will be devised for determining the upstream traffic influencing the link in question. The prediction model then will be activated for predictions on that link. In order to examine the validity of the prediction models, they will be evaluated through traffic data collected from continuous links on the network.

1.4 Research Tasks

This research specifies several tasks : 1) the identification of the traffic pattern, 2) the development of a link-based prediction model, 3) the evaluation of the link-based

prediction model, 4) the development of a network-based prediction model, and 5) the evaluation of the network-based prediction models. Each task can be broken down as follows :

- 1) Identification of the traffic trend
 - a. identify the traffic pattern as related to time
 - b. define the optimal time interval for diverse strategies
 - c. develop a universal univariate time-series model.
- 2) Development of link-based prediction model
 - a. devise a heuristic weighting system based on the identification of traffic flow trend
 - b. develop empirically a combined time-series model considering historical averages and past traffic data.
- 3) Evaluation of the link-based prediction model
 - a. test the accuracy of the predictions from the prediction models in terms of mean square error (MSE) and mean absolute percentage error (MAPE)
 - b. compare the prediction performances of link-based model with historical average in various scenarios.
- 4) Development of network-based prediction model
 - a. investigate the relationships between upstream and downstream traffic
 - b. identify the upstream traffic influencing the downstream traffic in question

- c. test the importance of the selected variables (upstream traffic, current traffic, and historical average)
 - d. devise an adaptive weighting system considering the fluctuation of traffic
 - e. develop network-based prediction models incorporating the three variables.
- 5) Evaluation of the network-based prediction model
- a. test the accuracy of the predictions from the prediction models in terms of root mean square error (RMSE), MAPE, and accuracy ratio through various data sets
 - b. compare the adaptive weighting system to the theoretical weighting system in prediction performance
 - c. develop a final rule of selecting one of three combined prediction models in terms of the current traffic condition.

1.5 Organization of dissertation

The dissertation consists of five chapters. Chapter 2 reviews the literature on forecasting models which found in various areas. Most forecasting models have been used for predicting the traffic flow. Also, a basic concept of time-series analysis theory is addressed in the second section of this chapter. In addition, the advantages of time-series model for prediction of traffic flow are discussed. The combined time-series model is described to give an idea of developing heuristic combined model.

Chapter 3 is devoted to the development and the evaluation of link-based prediction model. A heuristic model combining the historical average and the ARIMA

model is developed based on the weighting system adapting the current traffic condition. The combined model in 6-minute time interval is evaluated through traffic data sets collected from urban arterial.

Chapter 4 deals with the three network-based prediction models which combines upstream traffic, current traffic, and historical average. The importance of the selected variables is examined in prediction performance. The prediction capability of three combined models is investigated by regression analysis. In order to obtain the forecasts in terms of current traffic conditions, an adaptive weighting system is devised. The multiple combined models with the adaptive weighting system are evaluated through various traffic data sets. The adaptive weighting system developed here is compared to the theoretical weighting system in prediction performance.

Chapter 5 concludes the dissertation with the summary of the two prediction models. The results from the prediction models provide some conclusions. The topics for further study are followed by critiques on the prediction models.

2. LITERATURE REVIEW

Most literature related to this research is found in the areas of transportation engineering, electrical engineering, industrial engineering, statistics, management science, and economics. Since development of a prediction algorithm as proposed in this research blends a forecasting model with the concept of time-series analysis, addressing the present research problem requires a survey of the literature in two categories : existing forecasting models, and the basic idea of time-series analysis. First, numerous forecasting models show the capabilities to predict traffic flow. Although, these models are restricted by the specific problems reviewing previous forecasting models aids in understanding the problems associated with building a prediction algorithm. Based on the reviews about various approaches to the forecasting model, the time-series analysis approach has been selected because of the simplicity of its structure and suitability to real-time application. Second, to help explain the conceptual structure of the prediction algorithm, the basic concepts of time-series analysis theory are addressed. Specific topics for two categories are as follows :

- Forecasting models
 - Time-series analysis
 - Filtering models
 - Heuristic approaches
 - Prediction of travel time
 - Other statistical techniques

- Neural network approach
- Weather forecasting
- Use of the forecasting model
- Why time-series analysis choice
- Time-Series Analysis
 - Univariate ARIMA model
 - Combined time-series

The descriptions and critiques of diverse approaches to the forecasting model will be done in the first phase (forecasting model). The second phase provides a better understanding how time-series analysis is used in this research to develop a prediction model.

2.1 Forecasting Models

2.1.1 Time-series analysis

2.1.1.1 Univariate Box-Jenkins time-series approach

The earliest effort to traffic forecast using Box-Jenkins techniques is found in Ahmed and Cook (1979). The objective of their research was to apply the Box-Jenkins technique for modeling traffic time-series data. An ARIMA(0,1,3) model was found to represent volume and occupancy data from three different freeway systems of varying detector configurations and data-aggregation time intervals. Mathematically,

$$Z_t = Z_{t-1} - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \theta_3 A_{t-3} + A_t$$

where $Z_{t,i}$: Time – series traffic data ($i = 0,1,2,\dots$)
 $A_{t,i}$: Random shocks ($i = 0,1,2,\dots$)
 θ_i : Coefficients

The evaluation of the ARIMA(0,1,3) model against some other ad hoc smoothing models (moving average and exponential smoothing) has shown the overall superiority of the ARIMA(0,1,3) model in providing short-term forecasts of traffic parameters. Meanwhile, Levin and Tsao (1979) found that an ARIMA(0,1,1) model to be the most statistically significant for all forecasting intervals for both volume and occupancy. Also, they attempted to determine the most-effective forecasting interval using the forecasting efficiency 'D' as follows :

$$D = \sum |Z_t - Y_t|$$

where ' Y_t ' was the observed occupancy at time t and ' Z_t ' is the ARIMA occupancy forecast at time t . The 'D' values were compared for the 20-, 40-, and 60-s forecasts. The 60-s, which has the smallest 'D' value, was found to be the most-effective interval. In their research, the efficiency of volume forecasting and occupancy forecasting were compared by the ratio of the residual sum of squares to the mean of the observations. Volume forecasting was found to be less variable than occupancy. Both studies recommended development of incident detection algorithms might that use the Box-Jenkins time-series analysis. Ahmed and Cook (1980) chose traffic occupancy as the key state variable for detection of incidents, where incident is detected if the observed occupancy value lies outside confidence limits constructed two standard deviations away from the corresponding points forecasts from ARIMA model.

Meanwhile, a forecasting method based on linear time-series analysis was again explored and found to do reasonably well at forecasting mean values, but not so well for

those extremes corresponding to the onset of congestion (Davis et al., 1989). Authors stated that the extreme values were due to the transition from smoothly flowing traffic to the stop-and-go traffic they were attempting to forecast. A recommendation from the research is that for accurate forecasts, more work must be done in the area of combining several variables to describe extreme values.

By working with a longer time interval such as monthly traffic data, seasonal B-J time series models were built for short-term forecasting. Nihan and Holmsland (1980) found that a B-J time-series technique could produce highly accurate and inexpensive short term forecasts based on monthly traffic data. Nine-year monthly traffic counts were used to forecast the next 12 months of traffic volume. The seasonal ARIMA model selected as the final model was only 2 percent off at the end of a 12-month forecast period. Moorthy and Racliffe (1988) also developed a seasonal B-J time series model for monthly traffic flow obtained from automatic traffic counters. The comparison of the observed and modeled traffic flows showed that over 90 percent of the forecasts were made with less than ± 5.0 percent error. They found that significantly different forecasts from observed data were due to inconsistent observed data.

2.1.1.2 Multivariate time-series approach

Kyte et al (1989) explored multivariate time-series analysis to model freeway traffic flow. They described the transfer function model and the use of the cross-

correlation function in identifying and estimating the model. Transfer function models relate a dependent variable Y_t to one or more independent variables X_{jt} . The cross-correlation function (CCF) is the tool that is used to identify the nature of the transfer function relating the input series (independent variable) to the output series (dependent variable). They collected traffic data from 4:15 p.m. through 6:05 p.m. on a Los Angeles freeway. Data were obtained at 20-second intervals, resulting in a total of 330 data points.

The first transfer function model was related to the occupancy at a station to the occupancies at the adjacent upstream and downstream stations. The additional information from upstream and downstream stations into the occupancy-based model were expected to improve the ability to forecast changes in traffic conditions at a given station.

This transfer function can be written as follows :

$$OCC_{1,t} = \frac{w_1(B)}{\delta_1(B)} B^b OCC_{2,t} + \frac{w_2(B)}{\delta_2(B)} B^b OCC_{3,t} + \frac{\theta(B)}{\phi(B)} a_t$$

where

$OCC_{2,t}$ = occupancy at upstream station

$OCC_{3,t}$ = occupancy at downstream station

$\frac{\theta(B)}{\phi(B)} a_t$ = ARIMA noise model

$\frac{w_i(B)}{\delta_i(B)}$ = Coefficients

The above model was evaluated in uncongested and congested regions, respectively. For the uncongested regions, the change in occupancy at the home station was a function of

what is occurring immediately upstream. The effects of the transmission for the congested regions were significantly lower than for the uncongested region.

The second form related the flow at a given point on a freeway to the occupancy at the same point. This model can be written :

$$VOL_{1,t} = \frac{w(B)}{\delta(B)} B^b OCC_{1,t} + \frac{\theta(B)}{\phi(B)} a_t$$

where

$VOL_{1,t}$ = volume on the link 1 at time t

This transfer function included an occupancy variable for estimating volume. However, the relationship between occupancy and volume is not clear. In other words, it is hard to identify which is truly the independent variable. Meanwhile, the transfer function model was compared with the linear regression model. While this model was somewhat more complex than that of the regression analysis procedure, the transfer function models offered a more accurate method for the estimation of models relating traffic speed, flow, and occupancy when using time-series data.

Based on the above comparison, it was mentioned that the two transfer function models provided more statistically sound results and a way to capture the serial correlation for effectively identifying the model form. This multivariate time-series technique holds promise for analysis of freeway traffic flow issues and in the development of strategies for traffic monitoring and incident detection.

Polhemus (1979) developed univariate, bivariate, and multivariate models for modeling dynamic fluctuations in aircraft concentration. First, he modeled arrival control (AR), departure control (DP), and local control (LC), respectively. Second, by investigating the relationship between the three variables, bivariate models were formed such as AR-LC model and LC-DP model. Finally, all parameters were simultaneously combined to estimate each parameter. The LC and AR were considered as independent parameters for DP estimate. The LC and AR parameters were also calculated by including only a random error term. The transfer functions related changes in one sector to changes in the adjacent sector, modeling implicitly the movement of aircraft through the system. In analyzing the behavior of the traffic system, the ability to construct dynamic models such as a multivariate time-series model based on data recorded over time provided the analysis with a powerful method which can be applied to many different situations.

2.1.2 Filtering Theory

In 1984, two models using Kalman filtering theory were established for predicting short-term traffic volume by Okutani and Stephanedes. While most existing demand prediction methods made use of the time series of traffic flow data on the study link only, an alternative method has been developed for predicting 15-minute volume during the day using not only the traffic flow on the study link but also the traffic flows on the other links which may feed into the study link.

For predicting $z(\tau + k)$, the traffic volume k time intervals ahead of time τ , the following linear prediction model was introduced :

$$*z(\tau + k) = H_0(\tau)X(\tau) + H_1(\tau)X(\tau - 1) + \dots + H_r(\tau)X(\tau - r) + w(\tau)$$

where $H_j(\tau)$ ($j=0,1,2,\dots,r$) is a parameter matrix of dimension $m \times n$, $X(\tau)$ is a vector of dimension 'n' composed of traffic characteristics judged to be useful for the prediction of $*z(\tau + k)$ by the results of a priori investigation or by the experience of practicing traffic engineers, and $w(\tau)$ is a noise vector of dimension m . This model was compared to UTCS and UTCS-2, which predicts the next-control-interval (on the order of 5-15 minutes) traffic volume at each detector location in real time based on the measurements from the same location only. The UTCS algorithms made use of both smoothed historical traffic data and current traffic-volume measurements from the vehicle detector. As a result of the comparison, the proposed models by Okutani and Stephanedes performed better than UTCSs in 5-minute prediction. The tests also indicated a robustness in model performance as the prediction horizon increases, thus suggesting considerable promise for longer-term prediction.

An extended Kalman filter has been used to identify the diversion model parameters recursively using the most recent prediction error in real time (Stephanedes et al., 1989). A new approach to modeling on-line demand diversion in freeway corridors is directly applicable to dynamic control and vehicle guidance. In order to develop a Kalman filter, two major hypotheses were made to model the diversion of freeway demand at entrance ramps. The first hypothesis indicated that the perceived trip time, including

waiting time on the ramp, was the dominant factor determining diversion at entrance ramps. Socioeconomic factors, such as income, do not appear to be relevant. That suggested a utility-based model structure treating traffic demand as a homogeneous flow with diversion behavior depending primarily on the perceived delay caused by congestion. The first hypothesis underlying the new model was that freeway demand could be treated as a decision-making entity that responds to ramp control and the resulting on-ramp traffic condition according to the principle of utility maximization. The second hypothesis was that for a short (5-minute) time interval, the ramp condition did not affect total freeway demand approaching the ramp.

Two prediction models were determined corresponding to the above two hypotheses. The first model predicted the ramp-entering proportion (P_k); that is, the ratio of the vehicles entering the ramp to the vehicles arriving at the ramp entrance at every time interval (k). This model stated that the ramp-entering proportion was a function of the current freeway-entering rate (C) and the number of vehicles ($X+R$) on the ramp in the previous time interval. This formulation was as follows:

$$P_k = 1 \{ 1 + \exp[\theta_{1k} + \theta_{2k} C_k + \theta_{3k} (X_{k-1} + R_{k-1})] \}$$

where X_k is the number of vehicles on the ramp in the beginning of the k th time interval, R_k is the ramp-entering rate, θ_{ik} are parameters to be updated in real time and k usually represents a 5-minute interval. The second model used historical information to predict the number of vehicles arriving at the ramp entrance (Q_k) at every time interval k :

$$Q_k = q_{4k} E(Q_k) + q_{5k} E(Q_{k-1}) + q_{6k} Q_{k-1}$$

where $E(Q_k)$ is the past average rate during the k th time interval.

The above two models could not be applied in real-time control, because they contain an error. This error could make the diversion prediction meaningless since the error is not bounded. To address this issue, an on-line predictor was developed by incorporating the extended Kalman filter algorithm into the models, thus allowing the model parameters to be adjusted dynamically. As part of the prediction formulation, the values of model parameters θ_{ik} are assumed to represent the behavioral state of the demand at the k th time interval following the nonstationary random walk process. The model parameters were continuously updated by the extended Kalman filter using the prediction models as the observation equations. The resulting filter formulation was summarized as follows:

$$\begin{aligned} \text{F-I: } & \theta_{i,k+1} = \theta_{i,k} + w_{ik} \\ \text{F-II: } & \begin{cases} P_k = 1 / \{1 + \exp[\theta_{1k} + \theta_{2k} C_k + \theta_{3k} (X_{k-1} + R_{k-1})]\} \\ Q_k = \theta_{4k} E(Q_k) + \theta_{5k} E(Q_{k-1}) + \theta_{6k} Q_{k-1} + v_{2k}, \end{cases} \end{aligned}$$

where w_k, v_k are state and observation noise vectors, respectively, assumed to be for white noise. As a result of testing the above algorithm at freeway entrance ramps of the test corridor, the entrance ramp diversion proportion was predicted with reasonable accuracy. Also, the filtering algorithm was effective in reducing real-time prediction error.

A Kalman filter was applied to traffic density estimation in a 500-metre section in Paris using presence detector data (Bhourri et al. 1990). The method applied a Kalman filter in order to reduce density observation errors included in the occupancy measurements. A particularly simple Kalman filter for traffic density estimation was tested against real density measurements extracted from a filmed record. The estimation error

variance was halved. Sensitivity of estimates with respect to filter gain and sample time interval was low. Tracking behavior in the case of a wrong initial estimate was found to be rapid, in the order of one minute.

2.1.3 Heuristic approaches

Kreer (1975) evaluated the second and third generation Urban Traffic Control System (UTCS) predictor algorithms based on extensive testing. The second generation UTCS predictor algorithm made use of smoothed historical information and current-day vehicle detector measurements of volume. The former gave good prediction of changes in volume in the near future if the volume for the day in question follows the historical average pattern; the latter was used by the algorithm to correct for deviations from the historical average values. The second generation predictor was given by the following equations:

$$v(t) = m(t) + c(t) + d(t)$$

where $v(t)$ = the predicted volume for time interval t

$$c(t) = 0.9 c(t-1) + 0.1 r(t-1)$$

$$r(t-1) = f(t-1) - m(t-1)$$

$f(t-1)$ = measured volume for time interval $t-1$

$m(t)$ = smoothed historical volume for time interval t

$$d(t) = 0.2 \{ r(t-1) - c(t-1) \}$$

Meanwhile, the third generation predictor relied solely on present-day measurements. Vehicle detector measurements were exponentially smoothed. The predicted value was an interpolation between the most recent smoothed and unsmoothed

measurements. The interpolation constant depended on how far ahead in time the prediction was made. The third generation predictor was given by the following equations:

$$v(t) = x(t-2) + a Y(t-2)$$

$$\text{where } x(t-2) = .95 x(t-3) + .05 f(t-2)$$

$$f(t) = \text{measured volume for time interval } t$$

$$Y(t-2) = f(t-2) - x(t-2)$$

$$\text{and, } a = \frac{(N-1) \sum_{t=1}^{N-2} Y(t)Y(t+2)}{(N-3) \sum_{N-1}^N Y(t)^2}$$

Kreer tested predictor errors from the above two algorithms in terms of mean square error and mean absolute value of error. Generally, the second generation predictor was more accurate than the third. This means it would be worth the extra effort involved to obtain the historical information for each detectorized location in the system. Kreer (1976) also investigated whether vehicle detectors in a computerized traffic control system provided better information on network traffic demands for control purposes than could be obtained from historical average data. With various tests, he demonstrated that the vehicle detectors and the associated predictor algorithm could give only slightly better results than historical average data. The main reason that the on-line measurements together with a good predictor were only slightly better than historical information was that the vehicles that were measured by the detectors were not the ones present in the system at the time the new control actions became effective. Thus, Kreer emphasized the

necessity of further research, especially the need for a control algorithm to sense the emergence of saturation at any intersection in the network.

A simple prediction algorithm, which requires considerably less information and fewer computations, was subsequently proposed and compared with the best of the existing algorithms by Stephanedes et al. (1981). In addition to high reliance of UTCS on historical data, they stated that traffic volume could vary substantially, depending on various external factors (e.g. weather conditions, special events, developments in other modes of transportation, and even the traffic-control change itself). However, because of this reliance on historical data, UTCS-2 was not readily transferable across systems and therefore was not practical. They also stated that time-lagging was a serious drawback of UTCS algorithms. Thus, they developed a new algorithm to overcome the drawbacks of the UTCS predictor. Their algorithm used the volume during the next time period as the predicted variable; the current volume, the difference between current volume and previous volume, and the average volume during the previous three, four, or five time periods were the independent variables. The prediction equation was

$$\hat{v}_t = a_0 + a_1 f_t + a_2 (f_t - f_{t-1}) + a_3 \left(\frac{\sum_{k=1}^N f_{t-k}}{N} \right)$$

where

\hat{v}_t = predicted volume at time t

f_i = measured volume at time i

where N was the number of time periods considered ($N=3, 4,$ and 5 were suggested values) and a_1, a_2, a_3 were control coefficients that could be found by using standard regressive techniques to best fit the measured data for the location in question. In the comparison of the proposed algorithm with UTCS predictors, the proposed algorithm performed better than UTCS-3 in all cases; however, it did not always perform as well as UTCS-2. Meanwhile, they found some disadvantages of the proposed models through comparison with the moving-average version: (a) It might need to be updated frequently, (b) it required that data be collected on at least one previous day, and (c) it performed best when used during a specified time of day making it necessary to develop more than one equation for each day (a minimum of three equations for each peak and off-peak period). Despite these criticisms, the proposed algorithm showed significant advantages over existing prediction algorithms, such that it did not need extensive historical data as UTCS-2 and the historical average algorithms/methods did.

Akahane and Koshi (1986) developed an algorithm for short-term prediction of inflow volumes of urban freeways. They stated that prediction was made based on several inflow patterns which were a series of the standard values of 5-minute inflow volumes at each time of day and were set for each on-ramp. There were three steps for the prediction algorithm: Selection from inflow patterns, correction of inflow pattern, and calibration of corrected inflow patterns. After prediction, inflow patterns and correction coefficients were updated in order to maintain the accuracy of the prediction. The prediction system was applied to the on-ramps on Tokyo Expressway. The effects of traffic control on

inflow volumes were remarkable, according to Akahane and Koshi. This system also could successfully predict inflow volumes and automatically maintained the accuracy of the prediction.

A prediction method defined by the combination of autoregressive model fitting and pattern-matching was proposed by Kawashima in 1986. It used the features extracted from historical data of traffic flow to compensate for the delay caused by the predictor itself and also real-time data combined in the form of autoregression trace the sudden changes of traffic flow. Kawashima classified traffic data into three groups according to the characteristics of traffic flow; the first group indicated the time series had a large peak at around 7:30 A.M., the second group had moderate traffic volumes with peaks not sharp but widely spread, and the last group contained the time series with small traffic volumes and only a small degree of variation. A prediction algorithm by autoregression, then, would be followed by the classification. The 'AR + δ ' forecasting model can be defined in the following manner :

$$X_{n+\tau} = a_1 P(\tau) X_n + \dots + a_p P(\tau) X_{n-p+1} + \xi_n(\tau) + \delta_n(\tau)$$

where

$X_{n+\tau}$ = traffic flow forecast at time $n + \tau$

X_i = observed traffic flow

$a_i P(\tau)$ = AR coefficient

$\xi_n(\tau)$ = pure Gaussian white noise

$\delta_n(\tau)$ = non - Gaussian impulse

The prediction algorithm would be built based on the above formulation. In the application of the 'AR + δ model' to the Tokyo highway network, the algorithm based on

the 'AR + δ model' had a considerable effect in reducing the prediction error when the lead-time of prediction relied on the group or equivalently the shape of the time series.

A heuristic prediction model developed by Garbard et al. (1986) was used for traffic control in a congested urban area. The use of their prediction model utilizing B-J ARIMA technique, fed by both historical and real-time data, enables the plan calculation program to produce on-line updated plans. Practically, the prediction of the critical links of the urban network was crucial for updating the plans. The form of the prediction model was :

$$\hat{Q}(T+1) = \alpha H(T+1) + (1-\alpha)[Q(T) + \hat{V}(T+1)]$$

where

$\hat{Q}(T+1)$ = forecasted traffic flow at time T + 1

$Q(T)$ = Actual traffic flow at time T

$H(T+1) = \frac{1}{N} \sum_{i=1}^N Q_i(T+1)$ = traffic flow prediction for time T + 1, or historic average

with $Q_i(T+1)$ = measured traffic flow at time T + 1 of day i

N = number of days used

$\hat{V}(T+1) = L_1 V(T) + L_2 V(T-1) + L_3 [\hat{V}(T) - V(T)] + L_4 [\hat{V}(T-1) - V(T-1)]$

with $V(T) = Q(T) - Q(T-1)$

$\hat{V}(T)$ = prediction of V for time T

L_1, L_2, L_3, L_4 = model coefficients

The studied time series was a series of traffic flows over 24 hours. There was no particular seasonality during this period and the number of possible models were therefore limited. The various models 'ARIMA (P,D,Q) with $0 \leq P,D,Q \leq 2$ ' have been tested to pick

up the best model. An ARIMA (2,1,2) model has been proved to be the best model for $V(T)$. The coefficient ' α ' was a weighting value between the historic mean and the immediate past: the idea was to have it adapted with respect to the 'normality' of a current situation, i.e. the difference between this current situation and a normal one, as defined by the historic mean. This ' α ' coefficient allowed the model to cope effectively with traffic fluctuations.

Lu (1989) proposed the concept of an adaptive prediction system for prediction of traffic flow. Figure 2-1 shows such a system. In this system, Z^{-1} is a one-step delay factor, and Z^{-s} is an s -step factor (where s is a positive integer). Mathematically, $q(k)Z^{-i} = q(k-i)$, and $q(k)Z^{-s} = q(k-s)$. As can be seen from Figure 2-1, the core of the system is the adaptive processor. In the adaptive processor, all of the parameters(weights) at step k are adjustable. The error of prediction $e(k)$ controls the adjustment of the system. From Figures 2-1, the following equation can be derived :

$$\hat{q}(k) = \sum_{j=0}^N W_{jk} q(k-s-j) \quad k = s+1, s+2, \dots$$

where $\hat{q}(k)$ is a linear weighted combination of $q(k-s)$, $q(k-s-1)$, ..., $q(k-s-N)$. The weights are W_{0k} , W_{1k} , ..., W_{Nk} , and the index k denotes the time step. If $\hat{q}(k)$ is used to predict $q(k)$, then the error of prediction at step k is

$$e(k) = q(k) - \hat{q}(k) \quad k = s+1, s+2, \dots$$

The purpose of using an adaptive processor is to adjust the weights at each step k so that the mean square error $E[e^2(k)]$ is minimized.

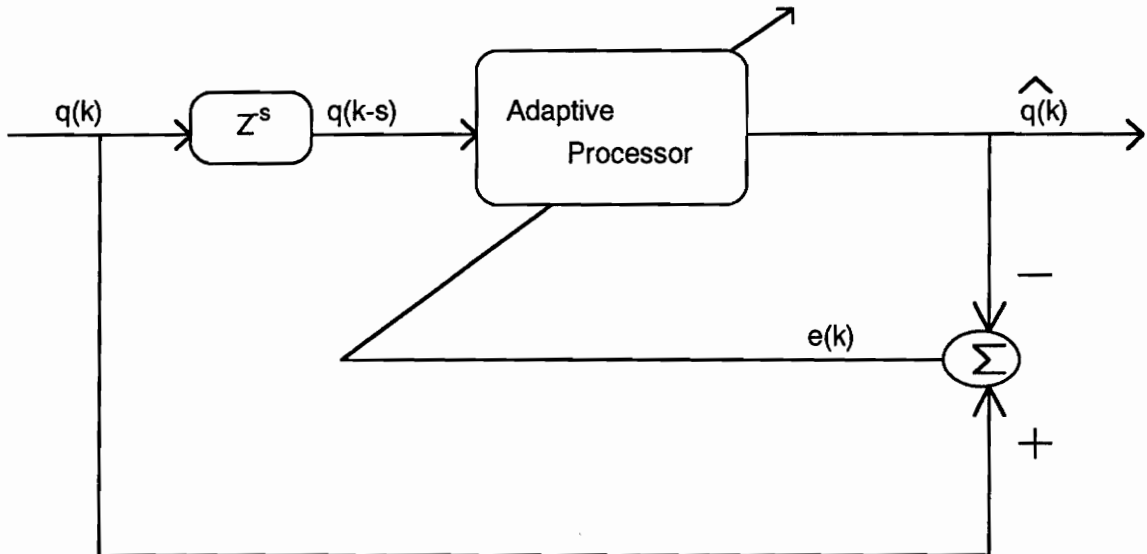


Figure 2-1 Adaptive prediction system (Lu, 1989)

As an experiment to validate the adaptive prediction system, a sine function was used to simulate traffic flow as input to the adaptive prediction system.

$$q(k) = A \sin(kT) + B \quad (\text{vehicle/hour, } k=0,1,2,\dots)$$

where A and B were positive constants and T was the sample interval. For the comparison between predicted traffic flow and observed traffic flow, the adaptive prediction system had the ability to predict precisely the future characteristics of the traffic flow by understanding the past process of $q(k)$. This ability could be the result of the

cyclic nature of the sine function input. Lu used the traffic flow of a one hour sampling interval to test the adaptive prediction system. Based on the test, he showed that the adaptive system could calculate the statistical characteristics of traffic flow using the adaptive processor and the past values to optimally predict future traffic flow.

2.1.4. Prediction of travel time

Oda (1990) proposed a prediction approach which utilized a time series model. It could predict changes in traffic conditions from origin to destination with the passage of time. In order to calculate the predicted travel time, he first, predicted traffic conditions using an auto regressive model (AR model) with lead time. Traffic conditions were based upon traffic volume and occupancy time from the past to the present.

Predicted travel time using vehicle sensor data could be calculated by predicting traffic conditions and assuming the mean vehicle length. The day's mean vehicle length h_d^r was represented as follows :

$$h_d^r = \sum_{u=1}^w \theta_u h_{d-u}^r$$

$$\sum_{u=1}^w \theta_u = 1, \quad \theta_u \geq 0$$

where 'r' was a time zone and 'd-u' was a date. Oda proved the accuracy of the algorithm to be 5-6 percent through various experiments. He also found the change in mean vehicle length had great influence on predicting travel time. The difference was found to be 51.1 percent.

Hoffman and Janko (1990) also emphasized the importance of the travel time prediction in a dynamic system. Since travel times vary with traffic volumes on the links of the network, travel times between the time of recommendation and the time of passing may change. Therefore, a recommendation of quickest routes must be changed over time. They assumed that for the estimation of the expected travel time on a downstream link the following information was available :

- the updated travel time standard profile of each link for this particular day of operation, and
- the travel time data of this day for each link up to the last 5-minutes before the prediction is started.

The travel time estimator was shown :

$$t_{i,N}^* = \delta_{i,N}^* * \bar{t}_{i,N}$$

$t_{i,N}^*$ is the best estimator for the expected travel time value on a downstream link i at a time of $T_0 + x * \Delta t$ (located in time interval N). The basis for a correction is the ratio

$$\delta_{i,n} = \bar{t}_{i,n} / \hat{t}_{i,n}$$

where $\bar{t}_{i,n}$ was again the average travel time value measured on link i in time interval n and $\hat{t}_{i,n}$ was the corresponding value of the actual travel time standard profile.

In order to obtain $\delta_{i,n}^*$ parameter, two substantial variables such as $\hat{\delta}_{i,n}$ and $\tilde{\delta}_{i,n}$ were considered :

$$\hat{\delta}_{i,n} = \alpha * \delta_{i,n} + (1 - \alpha) * \delta_{i,n-1}$$

where $\alpha=0.20$. This was developed to reflect the real situation in the link :

$$\tilde{\delta}_{i,n} = (1/a) * \sum_{r=1}^a \hat{\delta}_{r,n}$$

Considering changes in the neighborhood of a particular link a mean $\tilde{\delta}_{i,n}$ of the smoothed ratios $\hat{\delta}_{r,n}$ of all adjacent links was computed. From the these two values finally the modifying factor was built :

$$\delta_{i,n}^* = 0.5 * (\hat{\delta}_{i,n} + \tilde{\delta}_{i,n})$$

2.1.5 Other statistical approaches

The k -nearest neighbor (k -NN) approach, a non-parametric regression method, was used for freeway traffic forecasting (Davis and Nihan, 1991). The authors expected it to sidestep some of the problems inherent in parametric forecasting approaches. They assumed for the k -NN method that they had a series of observations $[x(s), y(s), s=1, \dots, n]$ of input/output pairs, which they called their learning sample, and an additional input measurement $x(t)$, from which they wanted to forecast $y(t)$. The k -NN method first ranked the input measurements in the learning sample $x(s)$ according to their distance from $x(t)$. Let s_1, \dots, s_k denote the indices of the k input vectors closest to $x(s)$. The forecast was then computed simply as

$$\hat{y}(t) = \left(\frac{1}{k}\right) \sum_i y(s_i)$$

i.e., as the average of the outputs corresponding to the k nearest neighbors of $x(t)$. Thus, the k -NN approach replaced the problem of selecting a class of models and then estimating the model parameters with the problem of maintaining and sorting an adequately large learning sample. They showed an empirical study comparing the accuracy of the k -NN method to simple univariate time-series forecasts. The k -NN method performed comparably to, but not definitely better than, the time-series approach. But it was proved that the k -NN method would be able to forecast the apparently nonlinear transitions between uncongested and congested traffic better. Possibly, a richer data base might improve the accuracy of the method.

Danech-Pajouh and Aron (1992) devised a statistical method for short-term inter-urban motorway traffic forecasting. Their method comprised three steps observations (and analysis) of historical data, modeling, and finally the choice of the proper forecast. In the observation-of-historical-data step, all the days in the historical record were divided into a number of classes (4 classes per month). The individual days were grouped into more homogeneous classes than the days as a whole. To do this, the authors used an automatic classification algorithm, the K-means method. They based their classifications on volume profile vector, which had as many components as there were half-hours in the day, each component being a half-hourly volume divided by the total volume (cumulative volumes since midnight, for example). In the modeling stage, for each half-hour and for each class, they proposed a regression model linking the traffic volume at time $T+1$ to the volumes at times T , $T-1/2$, $T-1$, and so on. At the regression stage, they did not keep the

profiles, since what concerned them was the predicting of traffic volumes. In the final traffic volume forecasts stage, the choice of volume class and the application of the regression model would be occur first. The choice was made on the basis of the last volumes available. They chose the class whose center was closest to the estimated profile for the current day. The second part--the actual forecasting--consists of applying the regression model.

In the application of their forecasting method to Nimes and Bollene in France, four models were calibrated:

- model 1: no upstream, no constraint,
- model 2: no upstream, with positive constraint,
- model 3: with upstream, without constraint,
- model 4: with upstream, with positive constraint.

The forecasting errors were generally less than 10 percent in 80 percent of cases. The forecasting method were again applied to the Beaune, which included the number of upstream points used as explanatory variables. The forecasts errors for the model adopted were less than 10 percent in 75 percent of cases (over 1395 half-hour intervals). Even though their method was efficient enough to remain below a 10 percent error threshold, it showed some problems in explaining a major accident or access restriction. To address this, a traffic flow simulation module would have to be added to the model. For correct

simulation of directional vehicle movements, additional information would be necessary (for example, vehicle counting on motorway access and exit ramps).

2.1.6 Neural network approach

An Artificial Neural Network (ANN) can "learn" and "solve" a difficult problem using complicated information, similar to a human brain. ANNs connect mathematically modeled neurons through synapses which are represented as modifiable weights. Therefore, the ANNs are adaptive, trainable, and massively parallel. The learning capabilities of a ANN are illustrated in Figure 2-2. Figure 2 show an extremely simplified neuron.

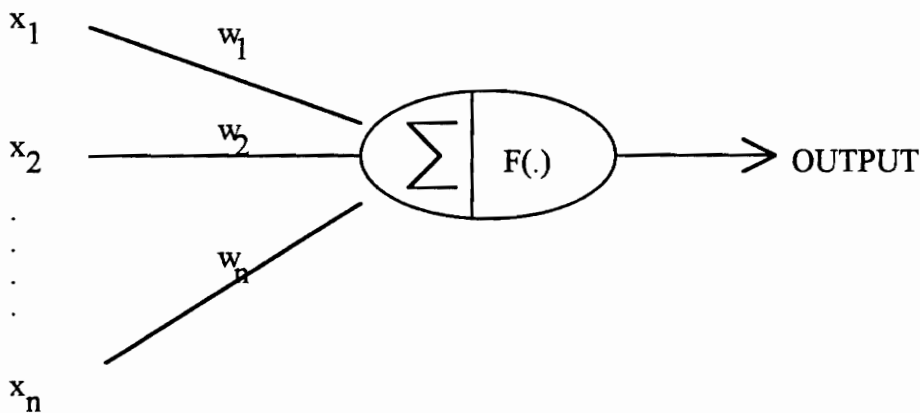


Figure 2-2 A simplified neural model (Kim, 1993)

where X 's are input data, w 's are trained weights, and $F(.)$ is the activation function that decides the output. The above idea could be applied to the various areas such as science

(image processing, pattern recognition, signal processing, and robotics), medicine, commerce, and the military (Kim, 1993). Its trainability and adaptability make the ANNs function as general prediction tools. They have been applied to areas such as stock market prediction, electric load forecasting, sunspot data prediction, prediction of chaotic systems, airline passenger data prediction, and prediction of secondary structure of proteins.

Dash et al. (1993) proposed a hybrid artificial neural network-fuzzy expert system for short-term load forecasting. The fuzzy membership values of load and other weather variables were the inputs to the neural network and the output comprises the membership value of the predicted load. The approach produced the load forecast in two steps. In the first step the inputs to the neural network were classified into overlapping regions such as low, medium, and high categories using nonlinear membership functions. The neural network produced a forecast of the power demand at the output node. In the second step, a fuzzy expert system was used to manipulate the forecasts by using a fuzzy rule base and influence mechanism pertaining to load and weather parameters to determine the final forecast. Assuming that traffic flow is fully understood, a well-developed ANN may predict the traffic more accurately.

2.1.7 Weather forecasting

The types of weather forecasts are classified into six groups. First, "persistence forecast" is a forecast that the future weather conditions will be the same as the present

weather conditions. Second, "steady-state (or trend) forecast" is based on the past movement of surface weather systems. It assumes that the systems will move in the same direction and at approximately the same speed as they have been moving. Third, "analog method of weather forecasting" is made by comparison of past large-scale synoptic weather patterns that resemble a given (usually current) situation in its essential characteristics. Fourth, "climatological forecast" is based upon the climate of a region rather than upon the current weather conditions. This is usually available for a month or more in the future. Fifth, "numerical weather prediction" is based upon the solutions of mathematical equations by high-speed computers. Sixth, "atmospheric models" are dependent on the simulations of the atmosphere's behavior by mathematical equations or by physical models (Ahrens, 1985).

Considering recent technologies for weather prediction, short-term forecasting - looking ahead only a few days - has long depended on computer predictions spruced up with some human insights, an approach that is yielding only small increases in accuracy (Kerr, 1989). Ten years of research and several generations of super computers have served to push the limits of medium-range forecasting, which is wholly dependent on computers. The attempt to peer 30 days ahead, utilizing the power of the computer, is on sketching in some detail the weather during a period of several days or a weeks. Furthermore, long-range weather forecasters are following suit by using a statistical system that considers the most recent seasons and thus seeks its 40-year memory of climates of the same season to find a close match. Until recently, the types of forecasting

tools available to the prediction branch of the weather service's Climate Analysis Center (CAC) for making 90-day forecasts had remained substantially unchanged for 25 years.

Barnston and Livezey (1988) included some additional steps to the original system. For example, after each fall season of the past 40 years has been compared with the most recent fall in terms of the distilled climate variables, the past seasons are ranked according to how good a match they make. The ten best matches are weighted according to the closeness of the analogy and then combined. Livezey and Barnston also stated an antianalogy selection in the system. Instead of assuming that what followed a similar fall in the past will follow this fall, the antianalogy method assumes the opposite : a fall that was a complete opposite would have been followed by a winter that was the opposite of the winter to be forecasted. The use of antianalogy in effect doubles the size of the sample from which comparisons can be made, a distinct advantages when seasons are so variable and the usable record so short. In addition to their works, subjective decipherings of recent atmospheric behavior around the globe, such as broad wind pattern, that are used to foreshadow future behavior have been included. Unfortunately, their long-range record is paltry 8 percent improvement over sheer chance in forecasting temperature.

Researchers in DERF (Dynamic Extended Range Forecasting) at European Center stated that Monte Carlo approach may be one of the solutions to the problems that the existing forecasting methods are not consistent enough for routine prediction of forecast skill on a day-to-day basis. In the Monte Carlo technique, the model is run at least ten times, instead of one, for each forecast. After the first run, each additional run begins with

an initial picture of the atmosphere that has been altered imperceptibly to mimic the real errors in the observations (Kerr, 1989).

2.1.8 Use of time-series model for traffic flow analysis

Maes (1982) developed B-J time-series models for three basic traffic parameters (traffic flow, space mean speed, and density). He found that none of traffic parameters considered yielded the same model for six data sets. The ARIMA models for q_1 (flow on right lane), q_2 (flow on left lane), v_1, v_2 (harmonic mean of the individual speed on right and left lane), $q_2 - q_1$, $v_2 - v_1$, q_1/v_1 (traffic density on the right lane), q_2/v_2 (traffic density on the left lane), $q_1 v_1$, $q_2 v_2$ (kinetic energy on the right and left lanes), $q_2/v_2 - q_1/v_1$, $q_2 v_2 - q_1 v_1$ were established for the analysis of traffic flow. The quasi-periodic phenomenon, damped sine wave in the autocorrelation function with a period of between 3 and 5 minutes, was detected four times for the parameter $q_2 v_2 - q_1 v_1$. This could be an indication that the difference in kinetic energy between the two lanes behaves in a periodic way in congested traffic conditions.

2.1.9 The advantages and weaknesses of UBJ model

Pankratz (1983) stated that the UBJ approach had three advantages over many other traditional single-series methods. First, the concepts associated with UBJ models are derived from a solid foundation of classical probability theory and mathematical statistics, whereas many other historically popular univariate methods are derived in ad hoc or intuitive ways. Second, ARIMA models are a family of models, not just a single

model. Box and Jenkins have developed a strategy that guides the analyst in choosing one or more appropriate models out of this larger family of models. Third, it can be shown that an appropriate ARIMA model produces 'optimal' univariate forecasts. That is, no other standard single-series model can forecast with a smaller mean-squared forecast error. Meanwhile, Nihan and Holmsland (1980) also described the advantages of B-J techniques in terms of accuracy, accessibility, cost and response time, and model domain. They showed that the B-J ARIMA model did better in forecasting than the other economic models such as models of FRB (Federal Reserve Board)-MIT-PENN (Pennsylvania) and BEA (Bureau of Economic Analysis). The ARIMA model is also friendly to beginner. The computer cost and response time are superior to the other economic models. They, finally, emphasized that it is most appropriate both for short and intermediate-range planning.

Generally, properly built UBJ models can handle a wider variety of situations due to the simplicity of the model structure. Since the model is built based on comprehensive statistical analysis with a previous data set, its predictions are reliable. Most importantly, it can be conveniently modified for multi-step ahead forecasting performance. Some heuristic models in the previous sections showed how the time-series model was efficient in adapting to diverse traffic conditions (Bennett, 1979). This is why a time-series approach is chosen for multi-step ahead forecasting.

On the negative side, the construction of proper UBJ models may require expertise and more model-identifying time than some historically popular univariate methods. But

determining a desirable ARIMA model structure in advance, an identification stage will not be necessary in the real-time application, thus reducing enormously the model-building time. Other weaknesses focus on the observation for model building. The large number of data points needed for estimation is one of the drawbacks that is common to ARIMA models. Another weakness related to observation is that the data set must be complete for all periods. If data are lacking for certain periods they must be manually created. The methodology that transfer the incomplete data set to complete one has been developed by Davis and Nihan (1984).

One of the important concepts in ARIMA modeling is a parsimony of the model. Even if a model has a very sophisticated and complex structure, there is no guarantee that the model would produce better forecasts than a single-formed model with a few parameters. Because of the above reasons, the UBJ model may help produce reliable and accurate forecasts for diversion strategies.

2.2 Time Series Analysis

2.2.1 Univariate ARIMA model

Time series analysis refers to any kind of analysis involving time-sequenced data from observations of a variable that occur in a time sequence. Time-sequenced observations in a data series ($\dots Z_{t-1}, Z_t, Z_{t+1}, \dots$) may statistically be independent. Sometimes times-series analysis is used more narrowly to describe attempts to explain behavior of time-series data using only past observations on the variable in question (Pankratz, 1983). An univariate Auto Regressive Integrated Moving Average (ARIMA) time-series analysis model was developed by Box and Jenkins in 1976. In univariate ARIMA analysis the correlation about Z at time t (Z_t) and Z at earlier time periods ($Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots$) was examined. The idea of Univariate Box-Jenkins (UBJ) forecasting is illustrated in a rough way in Figure 2-3.

Suppose that there are 80 time-sequenced observations available on a single variable. These are represented on the left-hand side of the graph in Figure 2-3, labeled "Past". By applying correlation analysis to these 80 observations, an ARIMA model is built. This ARIMA model describes how any given observation (Z_t) is related to previous observation ($Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots$). The ARIMA model could be utilized to forecast future values (for periods 81, 82, ...) of this variable.

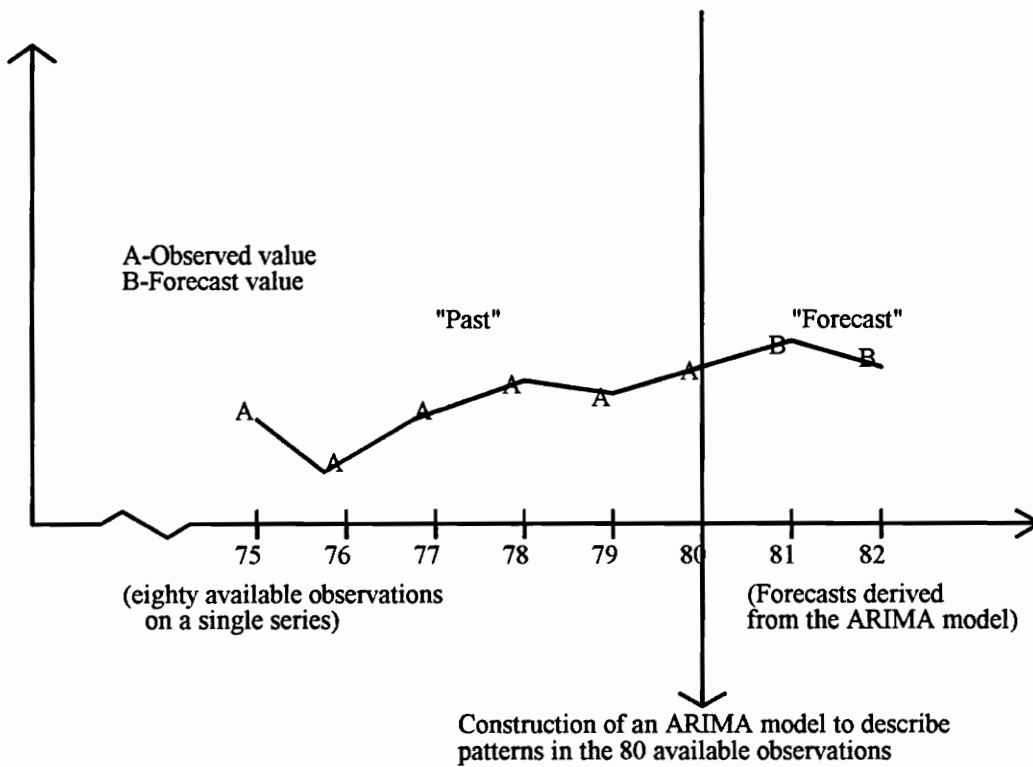


Figure 2-3. The idea of UBJ forecasting (Box and Jenkins, 1976)

2.2.1.1 When may UBJ models be used ?

Univariate Box-Jenkins (UBJ) ARIMA models are usually suitable for short-term forecasting. Most ARIMA models place heavy emphasis on the recent past rather than the distant past. Emphasis on the recent past means that long-term forecasts from ARIMA models are less reliable than short-term forecasts. In other words, the forecasts for period 'n+2' and beyond are less reliable than the forecast for period 'n+1' since they are based on less reliable information. An univariate ARIMA model usually deals with data measured at equally spaced, discrete time intervals. Also, UBJ models are particularly useful for forecasting data series that contain seasonal variation, including those with shifting seasonal patterns. Box and Jenkins (1976) suggested that about 50 observations was the minimum required sample size for building the ARIMA model. A large sample size is especially desirable when working with seasonal data.

Meanwhile, the stationary nature of the data series should be examined for univariate ARIMA modeling. A stationary time series has a mean, variance, and autocorrelation function that are essentially constant over time. The stationarity (stability) requirement may seem quite restrictive. However, most nonstationary series that arise in practice can be transformed into stationary series through relatively simple operations.

2.2.1.2 A basic structure of ARIMA model

An univariate ARIMA model is composed of two terms, which are an Auto Regressive (AR) term and a Moving Average (MA) term. The basic form of the model is shown as follows (Box and Jenkins, 1976) :

$$Z_t = C + \sum_{i=1}^k \phi_i Z_{t-i} + \sum_{i=1}^n \theta_i A_{t-i}$$

Where

ϕ_i = Auto Regressive Parameter

Z_{t-i} = Stationary Series

θ_i = Moving Average Parameter

A_{t-i} = Random Error

$$C = m \sum_{i=1}^k (1 - \phi_i)$$

m = Mean of observations

The first term 'C', in the equation (1), indicates the constant factor including ' ϕ ' values and the mean value of the observation. The second term ' $\phi_i Z_{t-i}$ ' of the model represents the Auto-Regressive function and the third term ' $\theta_i A_{t-i}$ ' of the model is the Moving Average function. The two functions are computed based on the statistical relationship between the observations available. The variable 'Z' in the second term is directly related to one or more past series values, while the variable 'A' in the third term is related to one or more past random error. The random error is determined by the difference between past predicted values and past observed values. Meanwhile, the ARIMA model can be described in simple format such as ARIMA (P,D,Q). The 'P' shows the numbers of 'AR' parameters in the model, while the 'Q' designates the number of the 'MA' parameters. If the traffic flow observation is not stationary, the 'D' term must be included in the simple format. Non-stationary data have an inconsistent trend over time, which means that the observations are not hovering around the averaged value. In order to make the data

stationary, the value on a study term of a data set is differentiated by the value of the next term or the values of the further term. The process applies to the value of every term in the data set, thereby the renewed data set by differentiation process may become another observation in stationary format to build a model. This is called a 'differencing step' for making the data stationary. For example, if there exist one hundred observed values, only 99 data would be left after one-differencing step. The 'D' in the simple format of representing the ARIMA model indicates the cumulative order of the differencing procedures implemented. If the data is not stationary after one differencing step, it goes to further differencing with the data set experienced through the first differencing. The differencing step continues until the data is obtained in stationary format.

2.2.1.3 The UBJ modeling procedure

Box and Jenkins (1976) proposed a practical three-stage procedure for finding a good model. The three-stage UBJ procedure is summarized schematically in Figure 2-4. First, the relationship between the whole observations would be investigated by the Auto-Correlation Function (ACF) and the Partial Auto-Correlation Function (PACF). These two functions provide a basis for the identification of the model. The structure of the parameters of the model are also determined in this step. In other words, the model would be classified as AR model, MA model, or ARIMA model. The 'AR' and 'MA' models are intense representations of the ARIMA model. If a 'AR' model has non-stationary data, it becomes 'ARI'. By the same token, 'MA' can be switched to 'IMA'. If the observations are stationary over time, the covariance values in the ACF would have exponentially decayed from the first observation and the covariance values in the PACF would have

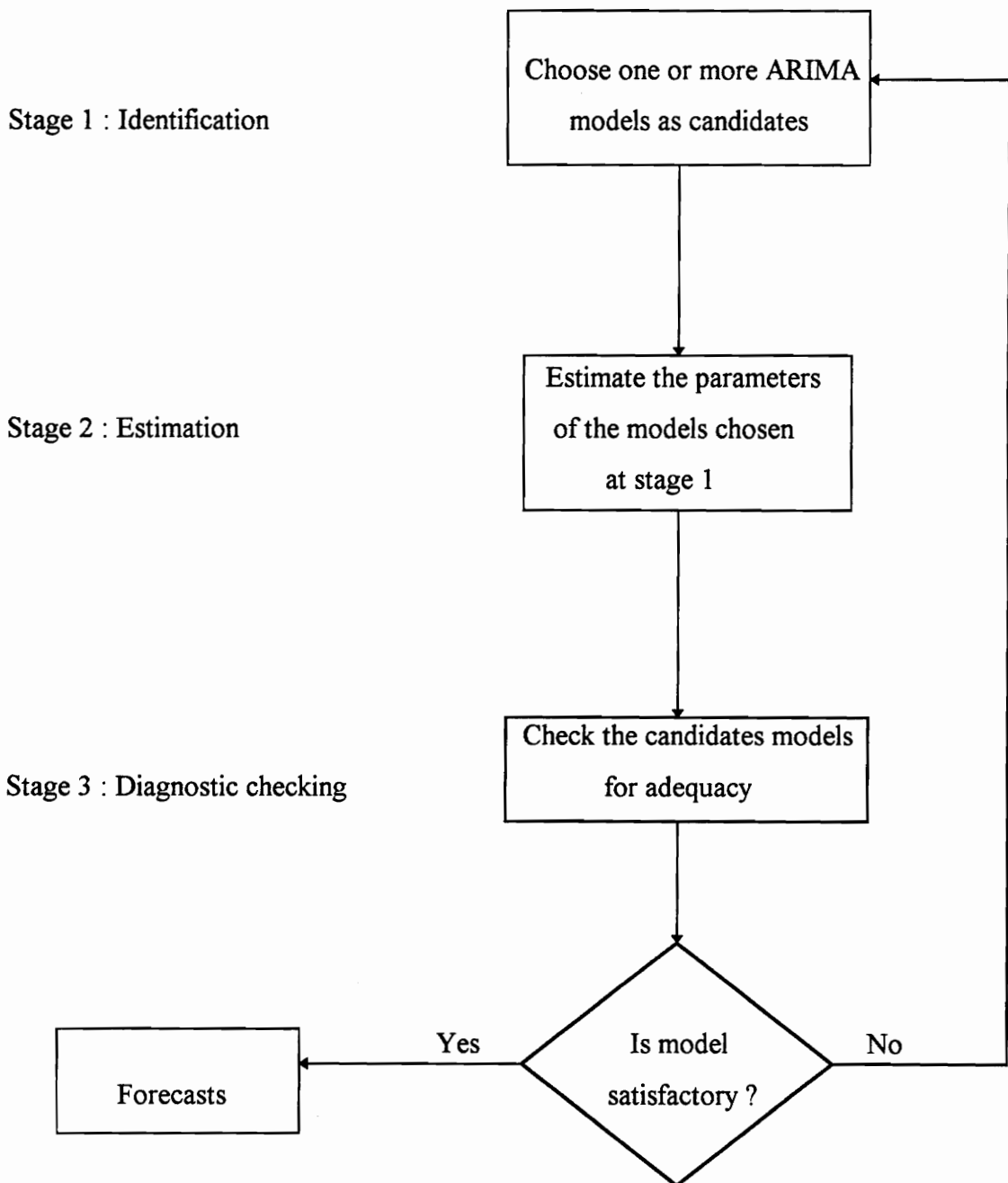
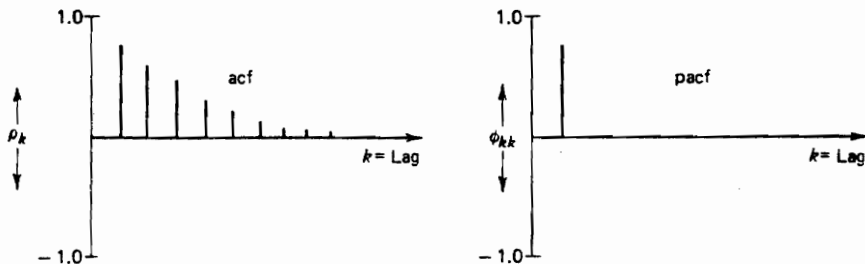


Figure 2-4. Stages in the B-J iterative approach to model building (Box and Jenkins, 1976)

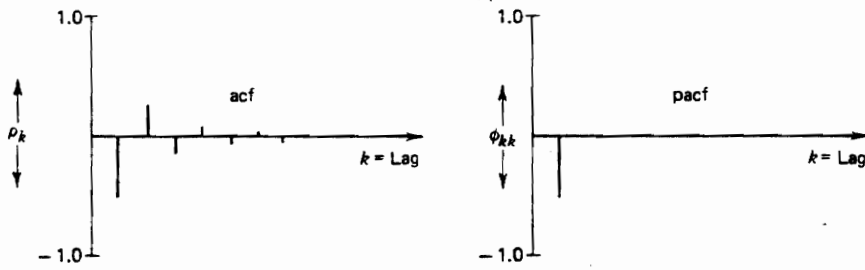
spiked at a certain lag, then cut off to zero. The above phenomenon occurs for the 'AR' model. The 'MA' model is for the reverse situation. Figure 2-5 shows a few examples of a stationary ARIMA model. The first two examples match for the 'AR(1)' model. The time-series models could also be represented as 'ARIMA (1,0,0)'. Reversibly, the third and fourth figures represent the 'MA(1)' model. These two models are also same as 'ARIMA (0,0,1)'. If the covariance values in both ACF and PACF show exponential decaying format, it becomes 'ARMA' model. Meanwhile, the non-stationarity (instability) of the observations would be tested here by looking over the ACF and PACF. Slow decaying in the ACF or PACF means that the data is not stationary, which is a required differencing step. The stationarity (stability) of the data could also be examined easily by plotting the real values against time, as shown in Figure 2-6. The first figure shows the fluctuating data hovering around the mean, whereas the next figure does not have any consistent trend over the mean of the data. Therefore the first figure is a representation of stationary data set, the next figure is an example for non-stationary data set. Second, the coefficients of the independent variables of the model, ϕ 's and θ 's, are estimated appropriately according to the model format identified. Usually, the least-square method or maximum likelihood method are used to estimate the coefficient values. In practice there are no considerable differences in the estimated values between the two methods. It is called an estimation step for building the model.

The estimation of the coefficients, the statistical test for the estimated coefficients is conducted in the diagnostic-checking step. The 't-test' and 'chi-square-test' are popular tools for testing the coefficients. When the 't' ratio of the coefficients in the 't' test is

Example I: $\phi_1 > 0$

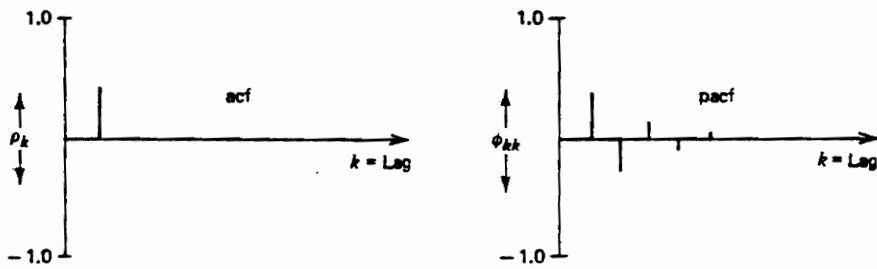


Example II: $\phi_1 < 0$

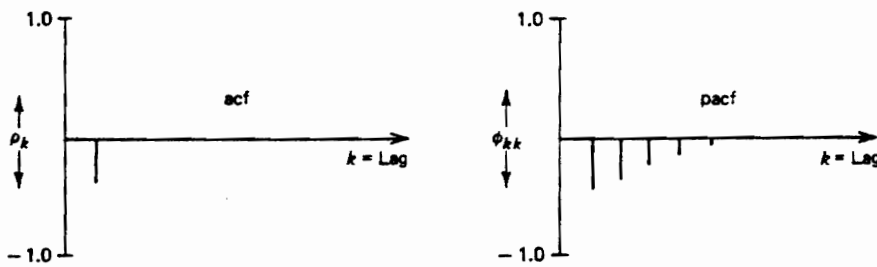


AR(1)

Example I: $\theta_1 < 0$



Example II: $\theta_1 > 0$



MA(1)

Figure 2-5 Examples of ARIMA model in ACF and PACF (Pankratz, 1983)

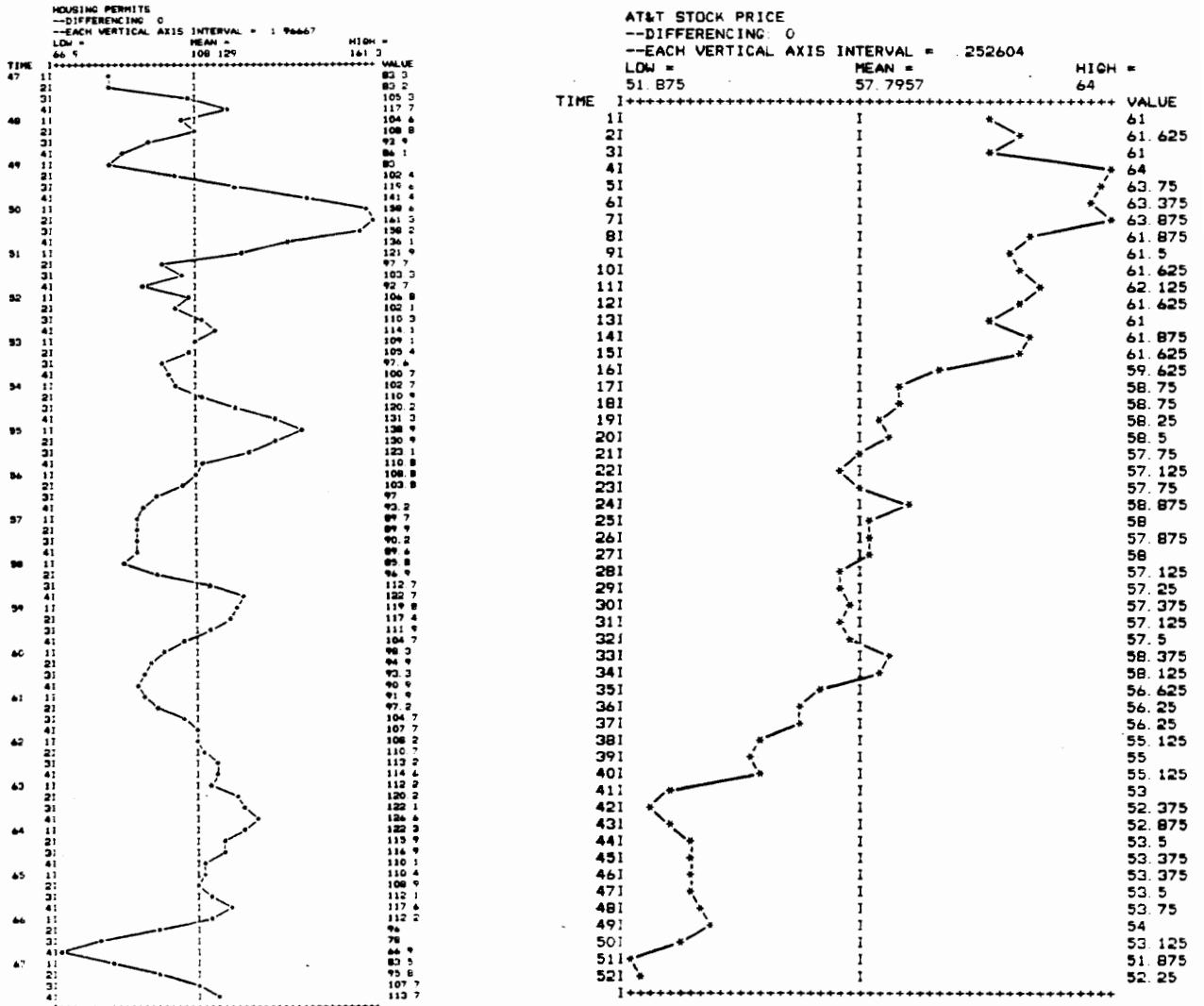


Figure 2-6 Plots of stationary and non-stationary data sets over time (Pankratz, 1983)

greater than 2.0, the coefficients are significant enough to be included in the model. The 't' ratio is computed by dividing the real coefficient value by the standard error of the coefficient. Regarding the attributes of parameters in the model, since the dependence between independent variables could result in biased predicted values, it is desirable to have low correlation between variables. The independence of the residuals, the difference between the actual values and the forecasted values, is also checked through the ACF and PACF of the residuals. If the t-values of all the residuals are kept within twice the standard error (two times the standard error of the residuals), it can be assumed that the model satisfies the condition of the diagnostic-checking step. In other words, the fact that the residuals of a desirable model are not related to one another means independence of the residuals. This final diagnostic-checking process is a critical consideration that measures the efficiency of the model for forecasting. After a model is built and checked through the three stages, the predicted values can be computed based on the model developed (Levenbach and Cleary, 1984).

2.2.1.4 Forecasting using the ARIMA model

The ultimate application of UBJ-ARIMA modeling is to forecast future values of a time series. While point forecasts can be derived algebraically from an estimated ARIMA model, interval forecasts are established based on probability limits around point forecasts. The forecasts are essentially correct in practice if the ARIMA models have been properly identified and estimated using a sufficient number of observations. The properties of

ARIMA forecasts are little affected by ordinary sampling error when the sample size is appropriate (Kendall and Ord, 1990). If two ARIMA models are equally acceptable from a statistical point-of-view through three stages, the forecasting ability of the two models must be compared to select the better one. However, if there are no considerable differences between the forecasts, the model that contains the smaller number of parameters is preferred for prediction in accordance with parsimony.

To find point forecasting from an ARIMA model using the difference equation form, write first the model in common algebraic form and solve for ' Z_t '. Insert the estimates of ' C ' and the ' ϕ ' and ' θ ' coefficients and assign ' A_t ' its expected value of zero. Then insert the appropriate values for any past observation (Z_{t-1}, Z_{t-2}, \dots) and past random shocks (A_{t-1}, A_{t-2}, \dots) (Thomopoulos, 1980). In practice estimation stage residuals may be used in place of past random shocks, or the expected value of zero if the forecast lead time ' l ' exceeds the lag length of the MA term in question. Likewise, forecast ' Z ' values must be used in place of observed ' Z ' values when the forecast lead time ' l ' exceeds the lag length of the AR term in question. In the meantime, Pankratz (1983) stated that ARIMA forecasts are optimal only if an appropriate ARIMA model has been found and only among forecasts from univariate, linear, fixed-coefficient models. It is, however, mentioned that a multivariate model, a model with a nonlinear combination of past Z 's, or a model with time varying coefficients might give forecasts with a smaller mean-squared forecast error.

2.2.2 Combined time-series model

A combination of forecasts can outperform individual forecasts. The common practice, however, is to obtain a weighted average of forecasts, with the weights adding up to unity. For instance, there are two unbiased one-step forecasts $f_{n,1}$, $g_{n,1}$ of X_{n+1} made at time n , so that $E[(X_{n+1} - f_{n,1})] = 0$ and similarly for ' $g_{n,1}$ '. The combined form was structured as follows (Bates and Granger, 1969):

$$C_{n,1} = \alpha f_{n,1} + (1 - \alpha)g_{n,1}$$

' $C_{n,1}$ ' is a superior forecast, in terms of the mean-squared error of the forecast error, than either component, ' f ' or ' g '. But Granger and Ramanathan (1984) indicated a combination of type in the above equation was the 'constrained form'. Even if it had the advantage that if ' f ' and ' g ' were unbiased then necessarily ' C ' would be unbiased, a 'unconstrained form' was introduced for optimal forecasts as follows :

$$C_{m,1}^* = \alpha_1 f_{n,1} + \alpha_2 g_{n,1} + \alpha_3 m$$

where m = average value of the variable

The unconstrained form is obviously relevant when the component forecasts are biased. They developed three combining methods based on the above equation. There were clear and obvious advantages in combining forecasts, both to better understand the generating mechanism of the series and also to pragmatically achieve better forecasts. Granger (1989) emphasized that a more intensive use of the available data was achieved by allowing the combined weight to vary with time. There are various methods available for a time-varying parameter regression. Usually the weights used to combine two different

forecasts, were based on variances and covariances of forecast error (Montgomery et al., 1988).

3. LINK-BASED PREDICTION MODEL

3.1 Introduction

It will be attempted to develop a prediction model based on all the links along the network. The forecasts on each link will be combined to represent traffic flow along the entire network. By obtaining the forecasts on each link, the traffic prediction for the entire network may be possible. In order to develop a link-based prediction model, only traffic information on the link on question will be utilized. The link-based model is designed to obtain the advantages of efficient operation through a simple-format with few traffic variables. An univariate B-J analysis is applied to build the link-based prediction model.

The prediction model used this chapter has two components. One component is an ARIMA model based on the previous traffic data. The other component is the average traffic flow for that period as obtained from previous days, called historical average. These two components are appropriately combined to represent the fluctuations in traffic-flow. The principles for a link-based prediction model are shown in Figure 3-1.

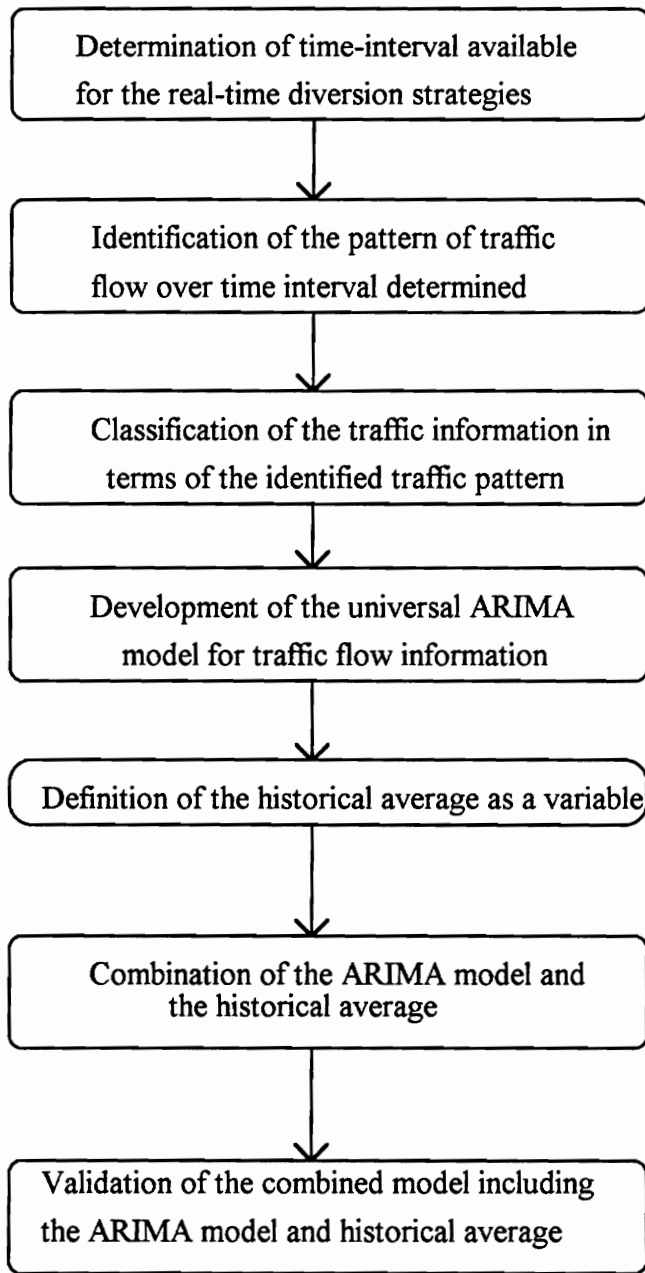


Figure 3-1 Procedures for dealing with a link-based prediction model

3.2 Optimal Time Interval for Diversion Strategies

The travel time between origin and destination in the study network developed by Hobeika et al. (1990, 1991, 1992) for real-time diversion strategies is assumed to be about 30 to 45 minutes. The prediction for 30 to 45 minutes along the routes would therefore, be necessary to provide efficient route choice to the drivers at the diversion point. The whole travel time between Origin and Destination must be divided into parts since it is too long to send real-time traffic information. A proper time interval will be determined throughout diverse experiments adjusting the time interval of the data set given.

In order to consider the most appropriate time interval, minute-by-minute traffic data has been collected from 6 a.m. to 6 p.m. on urban arterial in Los Angeles (provided by LADOT). This twelve hours of traffic information are assumed to represent a whole day of traffic flow data. The minute-by-minute data has been aggregated into six minutes to suit the diversion model application. As a matter of fact, considerable traffic changes do not occur during a one-minute time interval under normal traffic conditions. That is why six minute is appropriate time interval for prediction model. Another reason for using six-minute data is that 720 minute-by-minute data per day (12 hours a day) provides too much data for building an ARIMA model. About 100 six-minute data (600 minutes) was found appropriate for running the ARIMA model (Hobeika et al., 1992). More importantly, the predictions of 30-minutes-ahead time periods for application in diversion strategies would be difficult to perform utilizing minute-by-minute intervals because it requires a long-range prediction that can not be implemented by an ARIMA model. Thus,

the 6-minute interval has been selected for developing an ARIMA forecasting model within the real-time diversion strategies. Meanwhile, a universal ARIMA model available for traffic data in six-minute interval must be determined by operating various sets of traffic data from different locations.

3.3 Universal ARIMA Model

Identifying in advance the structure of the ARIMA model would be advantageous for real-time applications on networks. Sixty sets of data for six links for 10 days have been utilized to identify the structure of a universal model for normal traffic flow. As a result of adapting the traffic data (about 100 traffic data periods per day) with Statistical Analysis Systems (SAS), an ARIMA (2,0,0) model has been chosen using the three model-building stages. The model was tested for statistical significance through t-test, chi-square test, standard errors, and the independence of the residuals. The ARIMA (1,0,1) model may be an alternate choice for the prediction. These two models include equal number of parameters, which means that they have equal significance based on parsimony. Each model's ability to produce the forecasts must be examined when choosing between them. For one-step-ahead forecasting, the ARIMA (2,0,0) model tends to produce better forecasts than the ARIMA (1,0,1) model for limited data. Regarding the practical operation, while the ARIMA (1,0,1) model is required to know the random error, no random error is needed for the ARIMA (2,0,0) model. In other words, the observed values beyond the second step are not available in the prediction period for the ARIMA

(1,0,1). Thus, the ARIMA (2,0,0) model is chosen as a universal model for developing a link-based prediction model.

3.4 Multi-Step-Ahead Prediction

As stated earlier, prediction for a period of 30 minutes from the present is needed in order to evolve effective traffic control plans. To achieve this 30 minute prediction based on a six-minute time interval, five-term-ahead forecasting should be conducted to see how the traffic is changing in terms of time through the whole network. In other words, prediction of a dynamic movement of the vehicles is desired of the prediction model. However, the ARIMA (2,0,0) model is not a useful method to fulfill the aim of multi-steps-ahead forecasting. The prediction interval grows linearly or non-linearly as the leading times (forecasting steps) increase, which means that point prediction values in the farther term may be erroneous. Moreover, the ARIMA model, based on only current traffic information, is not enough to predict the next 30 minutes of traffic. Thus, a heuristic prediction method is definitely needed to solve the problems of multi-step-ahead forecasting.

3.5 Combined Time-Series Model for Multi-Step Forecasting

In order to improve the accuracy of the forecasts of traffic flow, the historical average of traffic flow for every six minute interval in a whole day based on nine-days of

data was combined with the forecasts resulting from the ARIMA model. A combined model is shown below.

$$\tilde{Z}_t = \alpha(\bar{Z}_t) + (1 - \alpha)(\hat{Z}_t)$$

where

\tilde{Z}_t = Forecasts at time t

\hat{Z}_t = Forecasts resulting from ARIMA model

$$\bar{Z}_t = \frac{\sum_{i=1}^n Z_{t,i}^k}{n}$$

= Historical average at time 't' on link 'k' (n = number of days)

α = Weighting factors

The ' \bar{Z}_t ' term is computed in advance for the links on the network. It may help detect abrupt increase or decrease in traffic flow, whether the changes are caused by the non-recurring or recurring congestion. The forecasts from ARIMA (2,0,0) model would be utilized as the ' \hat{Z}_t ' term in the above equation. The ' α ' coefficient is necessary for appropriate weighting values for each component.

As a preliminary work to provide weight values for each component, the traffic data throughout the network has been carefully examined to see how traffic flow changed over time. Among the 30 links of the network, links '199', '152', and '252' were classified as typical links, since most of the links throughout the network follow one of these three trends. Figure 3-2 shows the actual number of vehicles in terms of six-minute time intervals for the three links. As expected, it was found that the traffic flow has an up-and-down fluctuating trend in terms of the six-minute time interval. Thus, the forecasts are

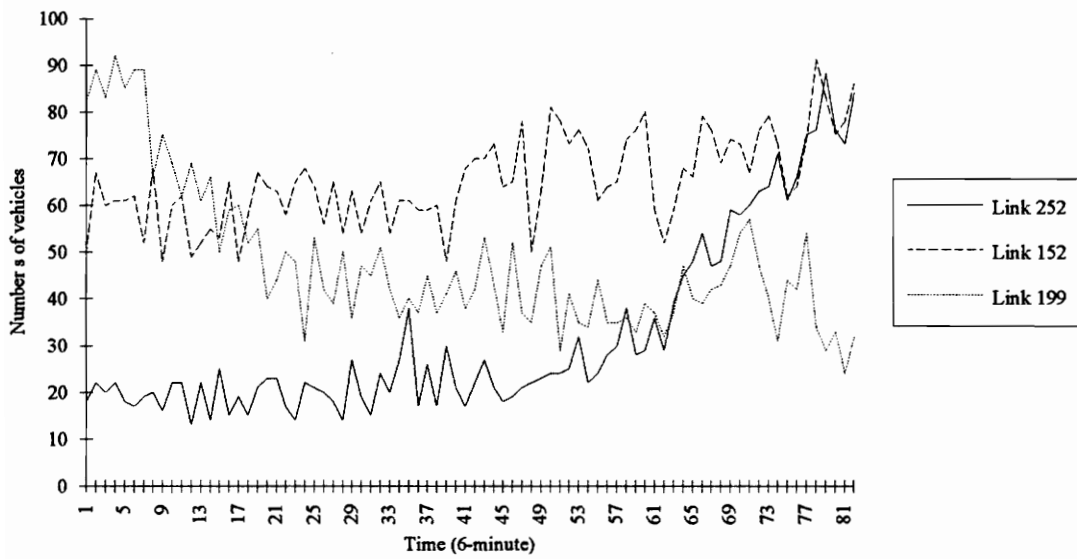


Figure 3-2 Traffic flow in links '199', '152', and '252' over six-minute period

assumed to follow the fluctuating pattern in the multi-step-ahead prediction performance. To determine the actual weighting parameters, the ' \bar{Z}_t ' term and the ' \hat{Z}_t ' term must be compared to each other. The mean values of both terms for five prediction periods are used for the evaluation of the Weighting Systems (W.S.). Table 3-1 shows the weighting systems based on the comparison between ' \bar{Z}_t ' term and the ' \hat{Z}_t ' term.

The 'DF' in column F of Table 3-1 indicates the difference between the ' \bar{Z}_t ' term and the ' \hat{Z}_t ' term during the five forecasting steps, which is defined as

$$DF = \frac{(\sum_{i=1}^5 \hat{Z}_{t+i} - \sum_{i=1}^5 \bar{Z}_{t+i})}{\sum_{i=1}^5 \hat{Z}_{t+i}}$$

The above percentile error (DF) is used as an indicator of the pattern of future traffic flow.

The case when ' \bar{Z}_t ' term is greater than ' \hat{Z}_t ' term is shown in column B and column C of Table 3-1. There are six weighting systems developed in terms of the difference between the means of both terms. The first weighting system of column B and column C ($\alpha=0.7$ and $1-\alpha=0.3$) indicates the weighting values on each term, if a one-step-behind observed value is greater than or equal to the two step-behind observed value. This model would be utilized to obtain the forecasts of every other term from the first-step forecasting (first, third, and fifth forecasting steps). This weighting values mean that the forecast is biased more or less on the ' \bar{Z}_t ' term. For the second and fourth-step-ahead forecasts, the

Table 3-1 A Pseudo-Weighting System (W.S.) for Combined Model

W.S.	average > ARIMA		average < ARIMA		DF
	B	C	D	E	
	α	$1-\alpha$	α	$1-\alpha$	
1	0.7	0.3	0.7	0.3	0 - $\pm 5\%$
	1.3	-0.3	1.3	-0.3	
2	0.6	0.4	0.6	0.4	$\pm 6 - \pm 10\%$
	1.2	-0.2	1.2	-0.2	
3	0.4	0.6	0.4	0.6	$\pm 11 - \pm 15\%$
	0.8	0.2	0.8	0.2	
4	-0.1	1.1	-0.2	1.2	$\pm 16 - \pm 20\%$
	0.4	0.6	0.3	0.7	
5	-0.3	1.3	-0.5	1.5	$\pm 21 - \pm 30\%$
	0.3	0.7	0.2	0.8	
6	-0.7	1.7	-0.9	1.9	$> \pm 30\%$
	0.2	0.8	0.1	0.9	

weighting values in the parenthesis of the first weighting system ($\alpha=1.3$ and $1-\alpha=-0.3$) would be used, which means that the forecasts are outside the range between ' \bar{Z}_t ' term and ' \hat{Z}_t ' term. This prediction would largely rely on the ' \bar{Z}_t ' term. If the two weighting factors are carefully reviewed, it can be noticed that a up-and-down fluctuating traffic flow pattern is reflected in the prediction. Conversely, if a one-step-behind observed value is less than a two step-behind observed value, the order of applying the two weighting values in the first weighting system is assumed to be switched. Regarding the rest of the weighting systems (2-6 in column A) in column B and in column C, the weighting values on the 'ARIMA' term become heavier little by little in terms of the increases in the percentile of differences between ' \bar{Z}_t ' term and ' \hat{Z}_t ' term, as shown in the column F of the Table 3-1. The last three weighting systems (4-6 in column A) in column B and column C show the weighting factors biased on ' \hat{Z}_t ' model due to the importance of the current traffic information.

When the ' \bar{Z}_t ' term is less than the ' \hat{Z}_t ' term in column D and column E, the weighting systems are almost the same as the earlier scenario (' \bar{Z}_t ' term is greater than or equal to the ' \hat{Z}_t ' term). The special feature of this scenario is in columns D and E described on the last three weighting systems, which have heavier weighting values for the ' \hat{Z}_t ' term than the systems in the earlier scenario. The combined model developed here will be evaluated with a large data set later on. Also, the comparison between the

historical average and the forecasts from the combined model will be shown in the evaluation of the model.

3.6 Evaluation of the Model

3.6.1 Links used

Six typical links for 10 days out of 30 links on the whole network were used to capture 60 sets of traffic information for evaluating the proposed model. Nine-days-data for each link were used for constructing 60 sets of historical average data. The 82 observations in each data set were analyzed to obtain the forecasts through the ARIMA model.

Figures 3-3 show the actual number of vehicles ('actual'), historical average values ('average'), the forecasts resulted from ARIMA model ('ARIMA'), and the result from the fourth weighting system of the combined model ('combined') over the five forecasting steps (see steps 3-7 in Figure 3-3) including the previous two observed values (steps 1-2). This is an example showing the performance of the combined model out of 60 data sets. The five forecasts for 'ARIMA' and 'combined' terms were computed by analyzing 82 observations. It can be observed how the 'average' term and the 'ARIMA' term are combined. In order to identify whether the first-step forecasted value is going upward or downward, the previous two traffic flow data just before the origin for the prediction are investigated in Figure 3-3. As observed, since a one-step-behind value is less than a two-step-behind actual value, the weighting values in parenthesis of the fourth

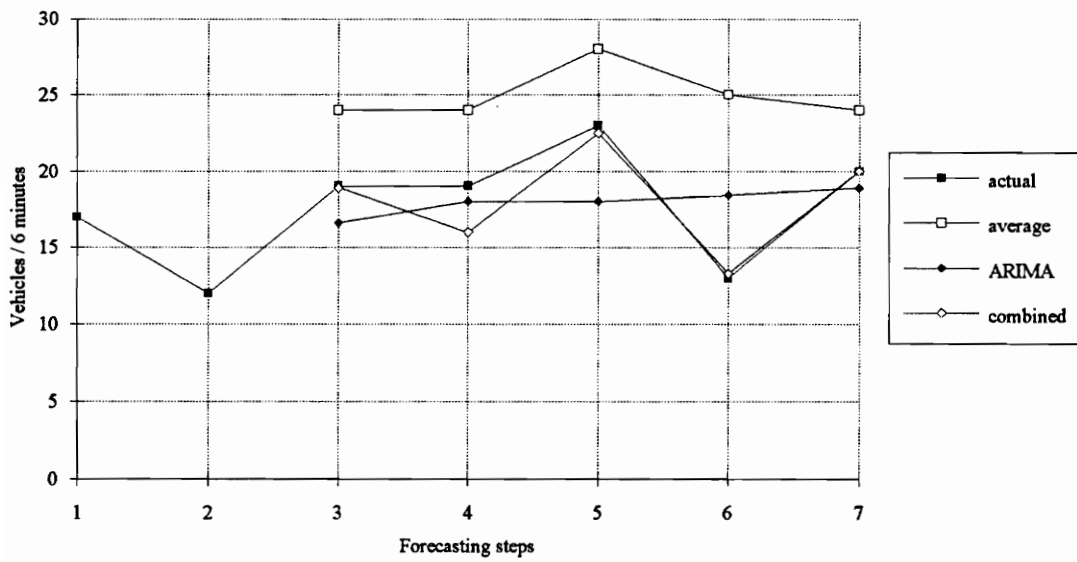


Figure 3-3 Traffic flow of four variables over forecasting steps

weighting system in Table 3-1 were activated for obtaining the 'combined' values in the first forecasting step. The weighting systems introduced make the ARIMA model realistic for multi-step-ahead forecasting by combining the historical average term. That is, the basic idea of the weighting system is to smooth the forecasts from the ARIMA model by combining forecasts from the ARIMA model with historical average values.

3.6.2 Criteria for evaluation of the model

For each prediction step, the difference between the observed values and the forecasted values is the measure used to evaluate the capability of the models developed. Square errors and absolute percentage errors are calculated for each forecasting step as tools to compare the two values. The square errors and absolute percentage errors for five terms are averaged to obtain the Mean Square Error (MSE) and the Mean Absolute Percentage Error (MAPE), which are the substantial criteria for the evaluation. The two measurements are defined as

$$MSE = \left(\frac{1}{S}\right) \sum_{t=1}^S (Z_t - \hat{Z}_t)^2, \text{ and } MAPE = \left(\frac{1}{S}\right) \sum_{t=1}^S \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right|$$

where

Z_t = observed traffic data

\hat{Z}_t = forecasts,

where S is the total number of forecasts made. The MSE detects the presence of frequent large forecast errors, while the MAPE is a measure of the expected error associated with an individual forecast (Ahmed, 1989).

3.6.3 Comparison of modeled values ('combined') and historical averages ('average')

The MAPEs for the six weighting systems are first compared to see how the pseudo-weighting systems perform in terms of the differences between 'ARIMA' term and 'average' term. As shown in Table 3-1, it is assumed that a data set with a smaller percentage of difference tends to have the lowest MAPE in the relatively lower weighting systems (first or second), while a data set with larger percentage difference has the lowest MAPE in the upper weighting systems (fourth, fifth, or sixth). In order to look at the above phenomenon, 38 data sets are randomly selected from the L.A. urban network. Thirty-one sets have a difference less than 17 percent, while seven data sets contain a difference between 13-30 percent. With respect to the first group (a difference less than 17 percent), 26 out of 31 data sets show the increasing values for MAPE over the six weighting systems, which means the lowest MAPE takes place in the first weighting system. Only five sets of traffic data out of 31 sets have the decreasing, constant, or fluctuating trend over the weighting systems, which means that the lowest MAPE is not located at the first weighting system. The MAPEs of seven sets of data with difference between 13-30 percent in terms of six weighting systems are shown in Figure 3-4. All the data sets in Figure 3-4 contain the lowest MAPE in the third or fourth weighting systems, which means that the ' \hat{Z}_t ' parameter has more weight than ' \bar{Z}_t ' parameter because of the significance of the current values for the prediction. Throughout the above tests, the use of weighting systems seems to be successful for the prediction, even though location of

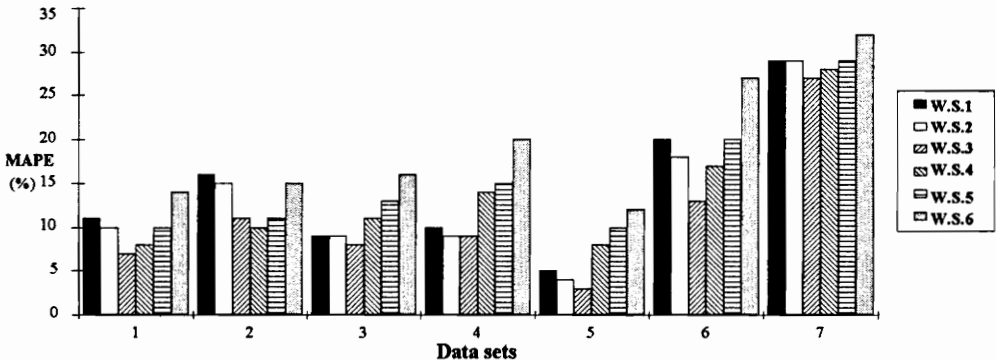


Figure 3-4 Comparisons of MAPEs on six W.S. for seven data sets with differences between 13-30 percent

the threshold for the change of the shape of MAPE over the weighting systems is not clear. In order to substantiate the assumption made above with another measure, MSE will be utilized for the evaluation.

MSEs are computed for the data sets used for MAPEs. They showed almost the same trend as described for MAPEs, even though there are some different shapes in a few data sets, which may be caused by the inconsistent nature of the traffic flow. Meanwhile, the aim of comparing the historical average and the modeled values from the six weighting systems is to see how the modeled values are better than the historical average values. The models do not perform consistently better than the pure average, since only about half of the sets of data for the models show the improved predictions for the data sets with a difference less than 17 percent. However, the data sets with a difference between 13-30 percent resulted in relatively better modeled values than the historical average values.

Meanwhile, fluctuations in traffic flow are assumed to be detected by the large margin of the difference between 'ARIMA' term and 'average' term. The large margin of that difference may mean that there are some traffic flow changes on the roads. If the data set contains values suddenly increased by one or more just before the prediction, the difference between 'ARIMA' term and 'average' term would be too large. The models' results produced by the data sets which consist of the suddenly changed values are tested to prove their superiority over the pure average values. Six data sets with 25-40 % margin are used for the test. The six data sets contain one or more suddenly changed values. Table 3-2 shows that the lowest MAPE and MSE for the six data sets occur in

around the fifth weighting system, which indicates the previous values are much more influential than the average values. There is no doubt about choosing the fourth or fifth weighting system for several-step-ahead prediction due to the much lower MAPE and MSE values than the MSE and MAPE for average values. Including very informative traffic data into the model, which are suddenly changed values, is a limitation of the model for consistently successful prediction. Since the traffic flow data naturally varies over the time frame assigned, the prediction for several terms is tremendously dependent on the origin of the prediction. In other words, the nature of the traffic data before the prediction affects much of the productions from the model. The above scenario including two informative data, which show a abrupt change in traffic flow, is a desirable situation for the prediction.

The scenario that the margin of difference between the ' \hat{Z}_t ' term and the ' \bar{Z}_t ' term is very high, at least 75 % was also investigated. Regardless of the high margin of that difference, the MAPEs and MSEs for six weighting systems indicated much higher values than the historical average. This may be caused by not including the suddenly increased values into the data set. That is, the observed data values for the prediction periods are much different from the observed data values in the data set available for prediction. Thus, it may be recognized that a high margin of that difference does not always mean better prediction from the six weighting systems. This worst case scenario is considered only under the situation without any information of sudden change in the traffic flow. This

Table 3-2 MAPEs and MSEs for six data sets with 25-40 percent margin

DATA	MAPE							
SETS	Difference	Avg.	Com -1	Com -2	Com -3	Com -4	Com -5	Com -6
1	35.31	25	26	23	15	10	13	15
2	40.84	25	27	25	19	12	14	17
3	36.17	23	28	26	19	10	10	12
4	24.3	67	69	65	52	35	33	37
5	32.36	70	71	66	49	25	18	14
6	28.16	37	39	35	23	12	11	11
DATA	MSE							
SETS	Difference	Avg.	Com -1	Com -2	Com -3	Com -4	Com -5	Com -6
1	35.31	82	98	82	41	17	24	33
2	40.84	103	100	84	44	22	32	46
3	36.17	87	113	95	45	14	19	27
4	24.3	116	122	109	74	43	39	40
5	32.36	104	110	94	51	15	9	7
6	28.16	47	52	42	20	6	5	8

scenario is another instance where the combined model is excessively dependent on the location of origin for prediction. Rather, the historical average would be more useful for prediction under that situation.

3.7 Findings and Critiques

The pseudo-weighting systems for the combined model were evaluated by comparing the accuracy of the modeled values and the historical average values in the various scenarios. The comparisons provided the following conclusions:

1. If the difference between ' \hat{Z}_t ' and ' \bar{Z}_t ' is less than 20 percent, use the first weighting system for prediction of traffic flow.
2. If the difference between ' \hat{Z}_t ' and ' \bar{Z}_t ' is more than 20 percent, use the fourth weighting system for prediction of traffic flow.
3. If the difference is very high, over 70 percent, use the pure average for prediction of traffic flow instead of using any weighting system.

It was noted that forecasts from the combined model were at least as good as the historical average. The combined model is simple to use for real-time applications, even though the historical average may need extensive historical traffic data. The above rules are applicable only for data sets under normal traffic conditions such as recurring

congestion, since the combined model is only validated through the data sets under normal traffic conditions.

If the combined time series model is utilized for the prediction of traffic flow under non-recurring congestion, a modification of the combined model may be required to handle that situation. The search for a threshold or standard to identify whether it is caused by non-recurring congestion will be helpful for prompt and accurate predictions. In order to evaluate the combined model under non-recurring congestion, the data sets under specific incident situations are required, and new weighting systems might be obtained. Comparisons between the combined model with a new weighting system, and other simple prediction methods such as the exponential smoothing method or the moving average analysis must be conducted to validate the capability of the combined model under non-recurring congestion.

Further research on this subject focuses on the improvement of the combined model for multi-steps-ahead prediction. Current traffic data and the historical averages alone are not sufficient to make for several-steps-ahead predictions. More information on traffic flow is definitely needed to accommodate diverse traffic conditions. The multi-dimensional time series analysis may be used to satisfy that need. In such a method, the prediction during each forecasting step relies on different spatial sources from upstream or downstream of the study link. If the adjacent links which are affecting the study link are found through a supplemental algorithm, a multi-dimensional time-series model including the traffic information from the upstream or downstream links can be developed for the

prediction of the traffic flow. In the multi-dimensional method, the one-step-ahead forecasting only will be produced. The algorithm that can identify the links influencing the study link will be required. The success of the application of the multi-dimensional time series analysis depends on the careful investigation of traffic flow between the links at various points-of-view. The investigation of relationships between adjacent traffic flow means the processes that find out the effective explanatory variables. The occupancy of the roadways and the speed of the vehicles might be candidates. The other weakness of the combined model is associated with the nature of traffic flow data, which tends to decrease after reaching the maximum capacity of the link. Occupancy data may be used as the variable instead of traffic flow. However, occupancy data may not be available for regular traffic due to the relatively low occupancy values for most of the day.

4. NETWORK-BASED PREDICTION MODELS

4.1 Introduction

Traffic conditions in a network at a particular time are not always the same as future conditions. In order to predict of traffic conditions, an algorithm that describes the time-varying traffic conditions throughout the network is needed. Here, three variables are investigated to construct the network-based prediction algorithm. The current traffic information (called *current*) obtained from roadway sensors for the network-based prediction model is one of the essential means of surmising the traffic trend and is a component of the prediction model. Effective communication between central traffic control site and roadway detectors will produce reliable traffic information. Practically, data from four previous time-series with 15-minute interval serve as current traffic information. Next, the historical average (called *hist. avg.*), fixed time system, represents time-of-day pattern. This variable is used to smooth the abrupt changes of the current traffic flow and avoid extreme forecasts. Finally, the upstream traffic, called *upstream*, is added as a new variable of the network-based prediction model. It is introduced to represent the dynamic nature of traffic flow. Like the flow of water, vehicles on the upstream move to the downstream along the network.

The quality of the network-based prediction model depends on the way these three variables (*current*, *hist. avg.*, and *upstream*) are combined. Based on an adaptive

weighting system, the prediction model will be evaluated by comparison with the historical averages. The adaptive weighting system combines the variables in terms of the traffic condition at the present time. If the difference between traffic conditions observed and traffic conditions estimated is large, the prediction models cannot be expected to result in improvements over *hist. avg.* (Kim and Hobeika, 1993). The quality of the traffic data collected from the detectors is another variable affecting the performance of the prediction model. If traffic data such as volume, speed, or occupancy is wrongly measured due to malfunction of detectors or weather, the prediction system is useless. Nikisch and Piotrowski (1984) stated that the quality of control in urban traffic control system depends on many variables of which the most important are

- a. Control equipment and induced by it constraints
- b. Control algorithms and their sensitive to the change of traffic parameters
- c. Quality of traffic data.

4.1.1 Hypothesis for the network-based prediction model

Traffic flow must be investigated throughout the whole network for more prompt and efficient control of traffic under the congestion conditions. A traffic prediction model that only relies on traffic information from a specific link may not detect changes in traffic elsewhere in the network, thus resulting in wrong traffic-control decisions. Therefore, a study of the relationships between the segments along the network is required to help forecast the traffic more accurately. A main uni-direction freeway is chosen to devise the

new prediction systems for several reasons. The traffic on the upstream tends to move to the downstream along that main freeway. Even though some minor parallel highways are available to divert the traffic into nearby alternates, most of the traffic on the main freeway stays on that road, especially traffic in morning peak-hours directed to the Central Business District (CBD). It has been assumed that upstream traffic for short-term prediction largely affects the concerned segment on the downstream, called *study link*. Traffic entering on or exiting from the ramp will be ignored for the prediction on the downstream. Traffic volume is used to represent the traffic conditions on the roadways. Speed data at the present time is used to identify the upstream traffic affecting the prediction for the study link.

4.1.2 Data used

Traffic data sets collected by the INformation FOR Motorists (INFORM) are utilized for developing and evaluating the prediction model, and examining the upstream traffic for purposes of the prediction. The traffic volume, speed, and occupancy at the Long Island Expressway (L.I.E./Rt. 495) in New York from Hauppauge to Kew Gardens (about 30 miles) were collected at 15-minute intervals for 22 weekdays during June, 1993. Electronic sensors are embedded in roadways at half-mile intervals throughout the corridor. The links separated by two miles were selected for study, because of the approximated data collection (aggregation of three lanes) and the long interval chosen (15 minutes). The configuration of the network is as follows:

Table 4-1 Configuration of the network used

Link	150	154	163	166	173	176	180	187	195	201	205	211	214	219	220	225	231
D		2.2	3.8	1.1	1.6	0.8	1.9	3.5	3.7	2.9	1	1.5	1	2.9	1.3	2.4	1.6

The first row (Link) in Table 4-1 is the number of links, and the second row (D) in Table 4-1 is the distance between the links selected. For example, the distance between link '150' and '154' is 2.2 miles. Link 231 is the CBD area. In order to detect the movement of vehicles on the upstream along the freeway, the long-range network was selected. Because the data-collection interval is quite long (15 minutes), only sensors located approximately two miles apart were activated for the development of the prediction models and its evaluations. Given the 15-minute interval, investigation of the *upstream* component had to be approximated. The 17 links (in Table 4-1) chosen from the network have all the reliable traffic data without being interrupted by frequent malfunction of the sensors. The 15-minute intervals for traffic volume were transformed to the vehicle per hour and given for three lanes. The unit used for calculation of the travel times was miles--per-hour. The values of the historical averages needed for the network-based model were calculated using 21 days of data for each location. Each day has different historical average because only 21 days data were used for each day. Ten-day historical averages were compared to the 21-day historical averages. A typical traffic volume over the links along the network at 7 A.M. is shown in Figure 4-1. Also, in order to show the typical traffic flow over time, traffic volume on link '176' over time is depicted in Figure 4-2.

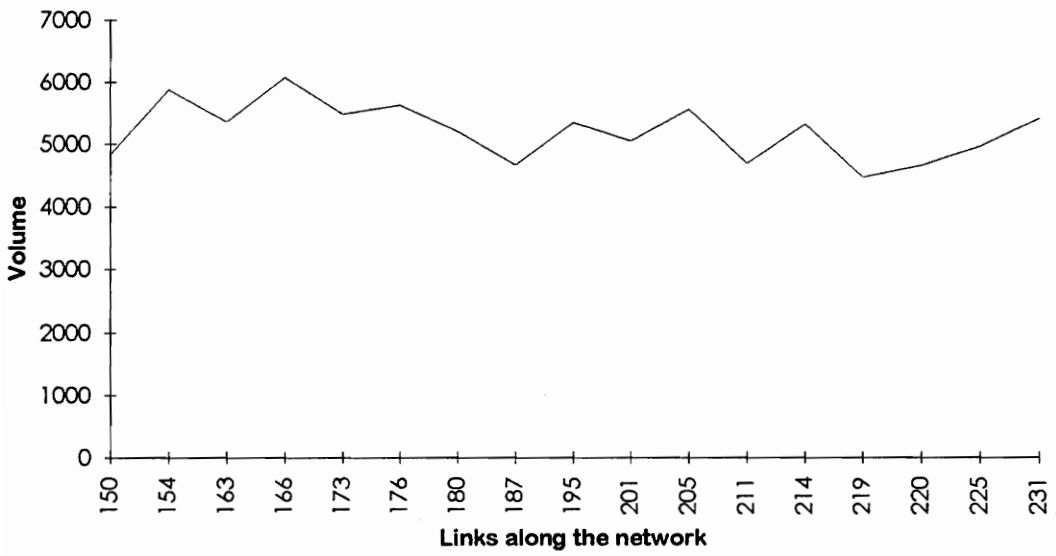


Figure 4-1 Traffic volume over links along the network at 7 A.M.

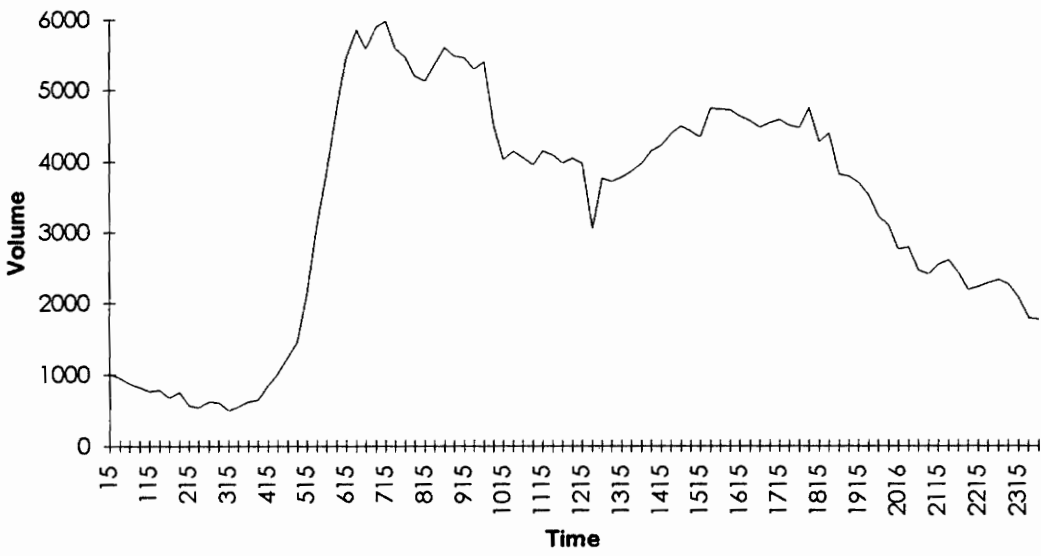


Figure 4-2 Traffic volume on link '176' over time

4.2 Identification of Upstream Traffic

The upstream segments that affect the *study link* on the downstream must be first identified for the development of the prediction model, based on current travel times between each segment. That is, if the upstream link (called *origin link*) that includes the traffic (vehicles) affecting the *study link* on the downstream after 15 minutes needs to be determined during a specific interval, the current travel times from the *study link* to the *origin link* are considered. Thus, a link at a location that takes 15 minutes to reach is selected as the *origin link*. The same procedures apply to the farther predictions. The travel times for the usual average speed on the link were calculated for comparison to the travel times based on the current traffic flow to see how the prediction models performs in terms of the traffic conditions. The 'usual average speed' that tells the daily average speed on the link at every 15-minute interval were given by the INFORM. In the case of June first's set of data, suppose the link '231' is the *study link* for the prediction performance. Let us look at Table 4-1 to see how to identify the upstream based on the current travel times. The link that takes 15 minutes to reach is selected as upstream traffic for the 15-minute-ahead prediction. Link '219' is the *origin link* because it takes about 15 minutes from link '231'--*study link*. Thus, the link '219' is selected for the 15-minute-ahead prediction. Also, link '205' and a link '187' are selections for the 30- and 45-minutes-prediction, respectively. Since the speed of vehicles (travel times) changes in each time interval, the location of the *upstream* traffic for the *study link* varies over time interval.

4.3 Multiple Combined Prediction Models

In order to conduct the short-term prediction, three combined models are developed here. The first model combines two variables such as upstream and historical average. The model is shown as follows:

$$Z_{t+1}^i = \alpha(a_1 U_t^{i+k} + a_2 U_t^{i+k-1} + a_3 U_t^{i+k+1}) + \gamma H_{t+1}^i$$

where

Z_{t+1}^i = Forecasts at time $t + 1$ on link i

U_t^{i+k} = Observed traffic at time t on link $i + k$ (origin link)

U_t^{i+k-1} = Observed traffic at time t on link $i + k - 1$ (adjacent link to origin link)

U_t^{i+k+1} = Observed traffic at time t on link $i + k + 1$ (adjacent link to origin link)

H_{t+1}^i = Historical average in time $t + 1$ at link i

$a_1, a_2, a_3, \alpha,$ and γ = Weighting parameters

For the upstream component, the three upstream sub-variables are chosen in the above model. Since fluctuation of traffic may affect the links around the *origin link* at the location where it takes 15 minutes to reach the study link, the adjacent two links of the origin link are also chosen as upstream traffic components. The historical average is obtained from the previous days data. For instance, if the historical average for 8 A.M. in a specific day is to be obtained, the values at 8 A.M. of previous days are averaged. In this manner, the historical averages are predetermined and updated.

In the second model, upstream and current traffic is combined. Four previous terms during one hour (15-minutes interval) are utilized as the current traffic component. Considering the study of the link-based prediction model, the earlier periods are supposed unnecessary for the short-term prediction. The amount of the previous intervals required

for the current component is really a matter of the sampling rate. The second model is shown as follows:

$$Z_{t+1}^i = \alpha(a_1 U_t^{i+k} + a_2 U_t^{i+k-1} + a_3 U_t^{i+k+1}) + \beta(b_1 C_t^i + b_2 C_{t-1}^i + b_3 C_{t-2}^i + b_4 C_{t-3}^i)$$

where

Z_{t+1}^i = Forecasts at time $t + 1$ on link i

U_t^{i+k} = Observed traffic at time t on link $i + k$ (origin link)

U_t^{i+k-1} = Observed traffic at time t on link $i + k - 1$ (adjacent link to the origin link)

U_t^{i+k+1} = Observed traffic at time t on link $i + k + 1$ (adjacent link to the origin link)

C_t^i = Current traffic in time t at link i

C_{t-1}^i = Current traffic in time $t - 1$ at link i

C_{t-2}^i = Current traffic in time $t - 2$ at link i

C_{t-3}^i = Current traffic in time $t - 3$ at link i

$a_1, a_2, a_3, b_1, b_2, b_3, \alpha,$ and β = Weighting parameters

The third model contains all three variables, which are upstream traffic, historical average, and current traffic. The model is as follows:

$$Z_{t+1}^i = \alpha(a_1 U_t^{i+k} + a_2 U_t^{i+k-1} + a_3 U_t^{i+k+1}) + \beta(b_1 C_t^i + b_2 C_{t-1}^i + b_3 C_{t-2}^i + b_4 C_{t-3}^i) + \gamma H_{t+1}^i$$

where

Z_{t+1}^i = Forecasts at time $t + 1$ on link i

U_t^{i+k} = Observed traffic at time t on link $i + k$ (origin link)

U_t^{i+k-1} = Observed traffic at time t on link $i + k - 1$ (adjacent link to the origin link)

U_t^{i+k+1} = Observed traffic at time t on link $i + k + 1$ (adjacent link to the origin link)

C_t^i = Current traffic in time t at link i

C_{t-1}^i = Current traffic in time $t - 1$ at link i

C_{t-2}^i = Current traffic in time $t - 2$ at link i

C_{t-3}^i = Current traffic in time $t - 3$ at link i

H_{t+1}^i = Historical average in time $t + 1$ at link i

$a_1, a_2, a_3, b_1, b_2, b_3, \alpha, \beta,$ and γ = Weighting parameters

4.3.1 Parameters (a and b) for 'upstream' and 'current' components

The 'a' parameters in the 'upstream' component of the three prediction models are determined intuitively. The applied values for the 'a' parameters are

$$a_1=3/6, a_2=2/6, \text{ and } a_3 =1/6.$$

Links that include 'a1' are *origin links*. The links that contain 'a2' and 'a3' are the adjacent links to the origin link. Since the vehicles on *origin link* have the most possibility that can be reached the study link in 15, 30, or 45 minutes according to the current travel times, it has the largest weights (3/6) among three parameters (a_1 , a_2 , and a_3). The link 'a2' is an adjacent link closer to the study link, while a link 'a3' is the link farther than the *origin link* 'a1' from the *study link*.

For the 'current' component in the second and third prediction models, the 'b' parameters are also intuitively obtained for the current component. The applied values for the 'b' parameters are

$$b_1=4/10, b_2=3/10, b_3=2/10, \text{ and } b_4=1/10.$$

The these values were selected for the same reasons as the a's. Since the 'b₁' component contains the most recent time-series data at the study link, it has the largest portion among four parameters. The weighting values about the remaining 'b' parameters are determined in order over time. The 'a' and 'b' parameters will be compared through regression analysis.

4.3.2 The 'a' and 'b' parameters estimated from the regression analysis

Two hundred observed upstream data sets were randomly selected to develop a regression model as follows:

$$Y_{k+i}^{t+1} = 1719.4 + 0.287 X_k^t + 0.337 X_{k-1}^t + 0.025 X_{k+1}^t$$

where

Y_{k+i}^{t+1} = forecasts at time $t + 1$ in link $k + i$

X_k^t = *upstream (origin link)* at time t in link k

X_{k-1}^t = *upstream (adjacent to origin link)* at time t in link $k - 1$

X_{k+1}^t = *upstream (adjacent to origin link)* at time t in link $k + 1$

This regression model, including a constant term, is obtained for comparison with intuitive parameters in the last section. Its R square value is about 0.35, which means that the *upstream* component alone is not a valuable variable for the prediction. However, P-values for the parameters are all less than 0.05, except for the third variable. The standard error for the model is 550. Meanwhile, the current traffic data produces a better regression model as follows:

$$Y_k^{t+1} = 635.9 + 1.133X_k^t - 0.298X_k^{t-1} + 0.091X_k^{t-2} - 0.06X_k^{t-3}$$

where

Y_k^{t+1} = Forecasts at time $t + 1$ in link k

X_k^t = *Current* at time t in link k

X_k^{t-1} = *Current* at time $t - 1$ in link k

X_k^{t-2} = *Current* at time $t - 2$ in link k

X_k^{t-3} = *Current* at time $t - 3$ in link k

The R square value of the above regression model is 0.86, which represents a very strong relationships between Y and X's. The P-values for each parameter are less than 0.05, excepting the third and the fourth terms. The standard error is 258, which is lower than that of the regression model calculated from upstream traffic.

The comparisons of the modeled values from intuitive parameters and from regressive parameters are presented in Table 4-2. The two regression models will occupy the places of the *upstream* and the *current* components of the three combined models.

There is a new criteria--accuracy ratio (Q-ratio)--to measure the quality of the modeled values (Saha, 1990). The rules applied for that ratio are

If the observed value is greater than the modeled value, then $Q = \text{observed} / \text{modeled}$.

If the modeled value is greater than the observed value, then $Q = \text{modeled} / \text{observed}$.

The RMSE (Ott, 1988) is calculated from

$$\text{RMSE} = \left[\frac{\sum_{i=1}^n (F_i - O_i)^2}{n} \right]^{1/2}$$

where

F_i = Forecasts at i

O_i = Observed values at i

Table 4-2. Parameters Comparisons of the modeled values between regression model and intuitive weighting

Combined Models	Regression Model				Intuitive Weighting			
	M1	M2	M3	H	M1	M2	M3	H
MAPE	.047	.036	.047	.054	.040	.033	.030	.054
Accuracy Ratio	1.028	1.011	1.034	1.031	1.028	1.015	1.009	1.031
RMSE	253.9	206.9	305.1	321.5	217.7	208.9	178.8	321.5

The values shown in Table 4-2 are the mean accuracy ratios for the data sets used. It is obvious that the best accuracy ratio is one. The closer a value is to one, the better prediction performance of the models.

Let us first look at the MAPE in the regression model division. The MAPEs of all three modeled values are lower than historical averages(H)' MAPE. Since current traffic has the highest relationship with observed value (shown in regression model for current traffic), the model combined (M2) with regressive parameters results in the lowest error rate. The MAPE derived from the first model is much worse than that of the second model's MAPE owing to the low relationship between the observed values and the upstream traffic. The third model performs as poorly as the first one, also due to the low relationship of the upstream traffic. By combining three variables in the third model, the quality of modeled values was not improved at all, which means that the fixed regressive parameters do not respond efficiently to the fluctuation of the traffic flow. The RMSE and Q-ratio show about the same results as MAPE. The Q-ratio of the third model is even worse than that derived from historical averages. Once again, the combination of the three variables for the fixed regressive parameters does not improve the quality of the model.

Meanwhile, the modeled values using intuitive weighting values present a different trend. The third model has the lowest error rates in all three criteria. Even though the upstream has low relationship with the observed values, the combinations of the variables

by incorporating the intuitive parameters provide much better forecasts than the first and second models. All three modeled values are, of course, much better than the historical averages. Comparing the second model's performance between the regression model and the intuitive weighting values in three categories, the error rates in both cases are almost identical. The prediction performances of the third model for the intuitive weighting are much better than the ones of the regression model. Thus, the intuitive weighting value is better in considering the parameters of the upstream traffic and the current traffic. It will be used for further study from now on.

4.3.3 Comparisons of the three look-ahead models

The three models are evaluated through regression analysis. Since the number of independent variables contained in the three models differ, the reliability of the each model has to be examined. The current traffic and upstream traffic of the models are determined based on the intuitive parameters identified in the previous section. Three regression models calculated from 300 peak-hour traffic data sets are:

For 15 - minute forecasting,

$$Y_1 = -3135 + 0.499X_1 + 1.141X_3$$

$$Y_2 = -579 + 0.171X_1 + 0.905X_2$$

$$Y_3 = -2269 + 0.258X_1 + 0.636X_2 + 0.538X_3$$

For 30 - minute forecasting,

$$Y_1 = -2960 + 0.649X_1 + 0.890X_3$$

$$Y_2 = -774.7 + 0.331X_1 + 0.773X_2$$

$$Y_3 = -2063 + 0.426X_1 + 0.549X_2 + 0.387X_3$$

For 45 - minutes forecasting,

$$Y_1 = -2842 + 0.393X_1 + 1.140X_3$$

$$Y_2 = -123 + 0.069X_1 + 0.866X_2$$

$$Y_3 = -2303 + 0.268X_1 + 0.329X_2 + 0.830X_3$$

Where Y_1 = Forecast from the first model
 Y_2 = Forecast from the second model
 Y_3 = Forecast from the third model
 X_1 = *Upstream* traffic
 X_2 = *Current* traffic
 X_3 = *Hist. avg.*

It is interesting to compare the parameters of the upstream component among the three models in the three prediction ranges. The parameter for the upstream component of the first model in the 15-minute forecasting is the highest, which means it has the largest influence on the prediction performance. There is a very small amount of weighting value portioned in the upstream component of the second model. It is due to high reliance on the current traffic component. The upstream parameter has been somewhat increased in the third model. It indicates that combining the three variables plays some role in providing more weights on the upstream traffic component. The trend of the upstream parameters in the three prediction ranges is almost the same.

The characteristics of the regression models are compared in Table 4-3. Table 4-3 presents the results of the regression analysis for the three look-ahead models in the three prediction ranges. The R-square values increase as the variables are incorporated into the models matching that the third model is most applicable to obtain the forecasts. Standard errors of the second and third models are less than the standard error of the first model. Comparing the three models in the three prediction ranges, the third model in the 15 minute forecasting range shows the best R-square and standard error. The worst prediction occurs at the second model of 45-minute forecasting. The two categories of the second model in the 45-minute prediction are worse than the first model in that range. It may be due to the low reliability of current traffic in the farther forecasting performance. The values are improved in the third model by combining three variables. Also, the low P-values of the models shows that the variables are significant enough to be included in the model. The exceptions that are greater than 0.05 happens in the second and third model of the 45-minute prediction range. Since the selected upstream values are too erratic (caused by the long trip of 45 minutes up to the study link), the regression models are unreliable. The current traffic in the 45-minute prediction is based on the estimates from the 15- and 30-minute forecasting performance. In other words, the estimates of the 15- and 30-minutes forecasting replace the observed values that are not available. Thus, the lack of observed values in predicting more than two-steps-ahead deteriorates the whole forecasting performance compared to the first-step-ahead forecasting.

Table 4-3 Characteristics of the regression models in the three prediction ranges

		15-minutes	30-minutes	45-minutes
Model I	R^2	0.666	0.603	0.637
	Adjusted R^2	0.654	0.587	0.611
	Standard Error	353.68	363.22	345.83
	P-value of X_1	4.45E-08	2.57E-07	0.00052
	P-value of X_3	1.7E-11	1.15E-07	6.6E-07
Model II	R^2	0.739	0.668	0.556
	Adjusted R^2	0.730	0.653	0.524
	Standard Error	312.52	332.83	382.76
	P-value of X_1	0.036	0.00032	0.666
	P-value of X_2	9.13E-15	1.6E-09	1.36E-05
Model III	R^2	0.784	0.695	0.661
	Adjusted R^2	0.772	0.676	0.624
	Standard Error	286.85	321.53	340.3
	P-value of X_1	0.0014	0.00044	0.098
	P-value of X_2	4.84E-07	0.00048	0.175
	P-value of X_3	0.00096	0.04	0.0068

At this point, forecasts were produced from the regression models. The forecasts are compared one another in terms of MAPE, RMSE, and Q-ratio in Table 4-4.

A comparison of the quality of the forecasts calculated through the regression analysis was conducted in the Table 4-4. As expected, the third model has the lowest mean error, variance, and standard error. The second model including the 'upstream' and 'current' variables are also better than the first model including 'upstream' and 'historical average'. The distribution of mean, that is, Standard Error (S.E.) of mean tells that models II and III are almost always better than model I. Considering the above comparisons, it is noted that the second and third models are more applicable to the forecasting performance. However, the regression analysis based on the fixed parameters may not be available for real-time application due to the incapability of updating the model structure according to the change of the current traffic flow. Thus, it is necessary to build an adaptation that accommodates changes over time and traffic conditions.

4.3.4 Comparison of the *upstream* parameters in prediction models

In order to compare the prediction performance of three models, the traffic data between 7 A.M. and 9 A.M. were used. Since historical average and observed values at a specific time interval are the same for the three models, the error rates (i.e. MAPE) of the historical average are the same in the three models. Figure 4-3 shows the MAPEs between the three models and historical average in 15-minute forecasting. The forecasts from the models are better than the historical average. The third model produced a little

Table 4-4. The errors of the three models in the 15-minute forecasting

Criteria		Model I	Model II	Model III
MAPE	Mean (%)	7.64	5.87	5.84
	S.E.(mean)	0.01	0.0061	0.0071
	Variance	0.012	0.0046	0.0061
RMSE	Mean	498.4	352.7	367.2
Q-ratio	Mean	1.0788	1.0609	1.0602

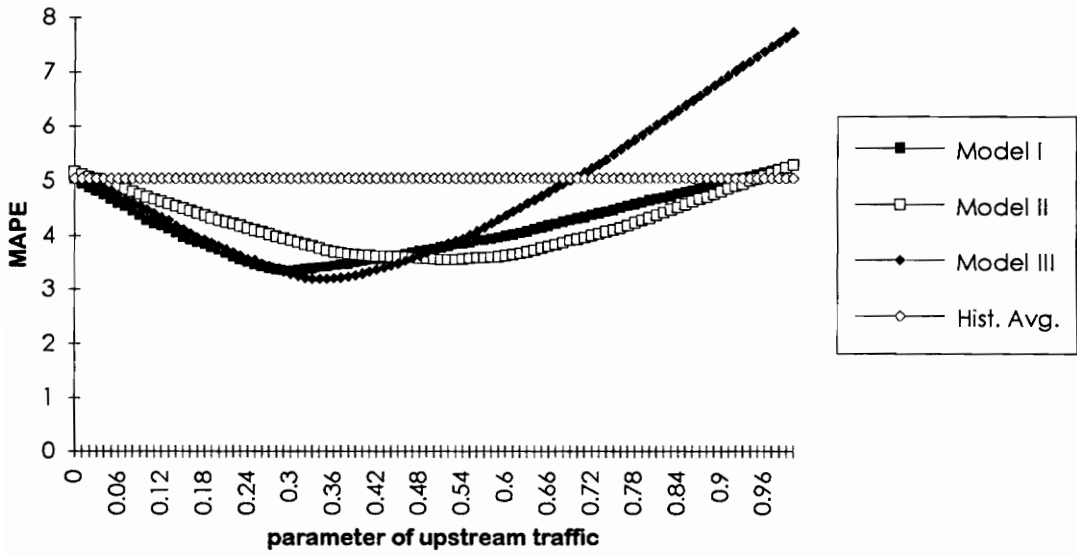


Figure 4-3 Prediction comparisons of three models in 15-minutes forecasting

better forecasts than the others. The MAPE were investigated in terms of the *upstream* parameter (α) of the prediction models. As shown in Figure 4-3, model II has the largest α value in obtaining an optimal forecast. While the *upstream* parameter of model II is about 0.6 in obtaining the optimal forecast, the one of model I is about 0.24. The *upstream* component in model II played a bigger role in obtaining the optimal forecast. Since the third model includes the three variables to combine, the α in the third model is relatively lower than the α in the second model. Figure 4-4 presents the comparison of predictions between the three forecasting ranges in model I. Since the same data sets were used, the error rate for historical average were same as the comparison in Figure 4-3. It is clear that the 15-minute forecasting produces the best forecasts. The predictions in other two ranges were worse than the historical average. Considering the predictions in terms of the *upstream* parameters, the 15-minute forecasting also has the largest α value in obtaining the optimal forecast. The influence of *upstream* traffic in 30- and 45-minute forecasting are much lower than the one in 15-minute forecasting. It is noted that the *upstream* traffic in the farther forecasting are not reliable in prediction performance.

4.3.5 Adaptive heuristic weighting systems for real-time application

Time-responsive weighting systems are developed to obtain the most appropriate parameters for real-time application of the prediction. Since three variables (*upstream*, *current*, and *hist. avg.*) are all time-dependent data, they will be influenced by the

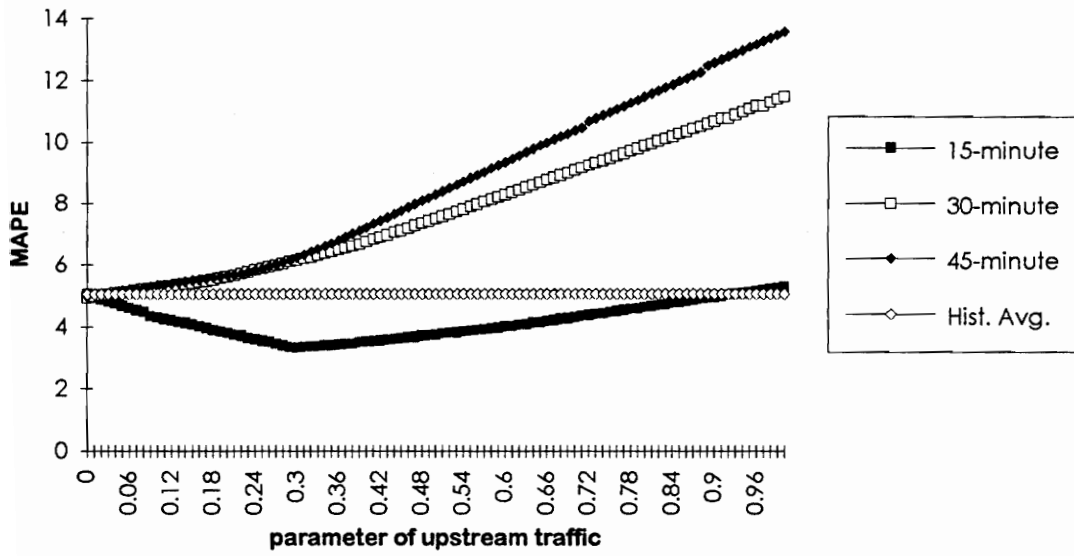


Figure 4-4 Prediction comparison of three forecasting ranges in model I

adaptive/time-responsive weighting system. Possible weighting scenarios for model I and II are shown in Table 4-5.

The two combined models generate nine sets of scenarios in terms of the fluctuation of traffic condition over time. The next step is to devise a methodology for selecting the best scenario for the traffic condition at the current time interval. Two decision factors emerge:

$$f_1 = \frac{|C_t^i - C_{t-1}^i| + |C_{t-1}^i - C_{t-2}^i| + |C_{t-2}^i - C_{t-3}^i|}{3}$$

$$f_2 = |H_{t+1}^i - C_t^i|$$

$$L = \left| \frac{f_1 - f_2}{f_1} \right|$$

$$M = \text{Standard deviation of } [|C_t^i - C_{t-1}^i|, |C_{t-1}^i - C_{t-2}^i|, |C_{t-2}^i - C_{t-3}^i|]$$

The first factor is the percentile error (L) between the differentiate of the origin of the prediction and the historical average (f_2), and the average of the differentiate of the four previous terms (f_1). The basic concept of the percentile error (L) is to measure how the traffic volume has fluctuated over the interval assigned. It is based on the supposition that reliable forecasts can be determined by carefully investigating the fluctuation or trend of the traffic flow over time. If the differentiates of the traffic volume data in each interval are small then, assuming there is no sudden change in the roadways at this current time, the historical average term must be heavily weighted compared to the other terms. The scenario 1 in Table 4-5 presents the case where historical averages are heavily weighted

Table 4-5 All possible weighting scenarios for model I and II

Scenarios	<i>Current (β) / Hist. avg (γ)</i>	<i>Upstream (α)</i>
1	0.9	0.1
2	0.8	0.2
3	0.7	0.3
4	0.6	0.4
5	0.5	0.5
6	0.4	0.6
7	0.3	0.7
8	0.2	0.8
9	0.1	0.9

($\gamma=0.9$ and $\alpha=0.1$). Conversely, if there exist huge percentile errors (as in scenario 9 in Table 4-5) a weight heavier on the upstream component than the historical average/current components is the best for forecasting. The selection numbers in Table 4-5 have been incremented as the percentile errors increase.

The second decision factor (M) is the standard deviation of the differentiates between the previous four intervals including the origin of the prediction. It is devised to detect unusual trends in the recent past. If there are huge variances in recent traffic flow, the upstream traffic will have greater influence than the historical averages/current traffic in terms of the prediction performance. The parameters (α , β , and γ) are determined according to the combination of decision factors as shown in Table 4-6.

The two decision factors are empirically harmonized in Table 4-6. While the percentile error ranges have been diagonally incremented by the model number in order, the standard deviation ranges has been vertically incremented. The combinations of the two decision factors are diagonally deployed by the order of selection number. The combination sets of parameters were found as follows: Suppose that there are traffic data sets including *upstream*, *current*, and *hist. avg.* First, the calculations for decision factors (L and M) must be conducted, then selection of cells containing the values of the two decision factors. After a scenario is selected, the combination of the parameters in the scenario is employed to obtain the forecasts. Such is the procedure for obtaining the forecasts from the data sets given. The difference between models I and II in applying this procedure is that the current component in the second model replaces the historical

Table 4-6. Adaptive weighting system for model I and II

Decision Factors	Combination of the decision factors					Scenarios
L (%)	0-49					1
M	0-49					
L (%)	0-49	50-99				2
M	50-99	0-49				
L (%)	0-49	50-99	100-199			3
M	100-149	50-99	0-49			
L (%)	0-49	50-99	100-199	200-299		4
M	150-199	100-149	50-99	0-49		
L (%)	0-49	50-99	100-199	200-299	>300	5
M	200-299	150-199	100-149	50-99	0-49	
L (%)	0-49	50-99	100-199	200-299	>300	6
M	>300	200-299	150-199	100-149	50-99	
L (%)		50-99	100-199	200-299	>300	7
M		>300	200-299	150-199	100-199	
L (%)			100-199	200-299	>300	8
M			>300	200-299	200-299	
L (%)				200-299	>300	9
M				>300	>300	

average in the first model. Meanwhile, the third model, with three variables to combine, has twelve sets of possible scenarios in Table 4-7.

The developing order of this heuristic weighting system begins with the determination of all possible parameters for three variables. The historical averages are the bases for parameter selection. In the 15-minute prediction, the ranges of 0.6 to 0.2 are selected. Ranges greater than 0.7 and 0.1 were ignored to avoid the extreme combination, which renders the three variable combinations meaningless. Since reliance on the upstream and the current traffic is relatively high in this 15-minute forecasting, lower parameters (0.4-0.2) for historical averages dominated selection four through twelve. The *upstream* and *current* parameters are determined based on correspondence to the historical average parameter. The *current* is more weighted in a lower numbered scenario than in higher model within a specific historical average parameter chosen. For example, look at the model 4, 5, and 6 from the Table 4-7. The fourth model (lower-numbered model) has 0.5 for the *current*, which is much larger than 0.1 for the *upstream*. Next, the percentile error (L) and standard deviations (M) are computed. The most appropriate values of two decision factors are fitted into the set of scenarios through trial-and-error. That is how the adaptive weighting system is developed in Table 4-7. The order of obtaining the best combination of the three parameters is the computation of the two decision factors, scenario selection, and determination of the parameters. The H.A. in Table 4-7 represents historical average.

Table 4-7 Adaptive weighting systems for third combined model in 15-minutes forecasting

L (%)	M	Scenarios	Upstream	H. A	Current
0-24		1	0.2	0.6	0.2
25-49	0-149	2	0.1	0.5	0.4
25-49	>150	3	0.4	0.5	0.1
50-99	0-99	4	0.1	0.4	0.5
50-99	100-199	5	0.3	0.4	0.3
50-99	>200	6	0.5	0.4	0.1
100-299	0-99	7	0.1	0.3	0.6
100-299	100-199	8	0.3	0.3	0.4
100-299	>200	9	0.5	0.3	0.2
>300	0-99	10	0.2	0.2	0.6
>300	100-199	11	0.4	0.2	0.4
>300	>200	12	0.6	0.2	0.2

In the farther forecasting (30- and 45-minute intervals), the method of developing the systems and of obtaining the combination are exactly the same as the ones in the 15-minute forecasting as well as in the structure of the systems (Table 4-8 and Table 4-9).

There, however, is a small difference in combining the parameters and the decision factors. Since the *upstream* and *current* variables lose their reliability in that farther prediction, the historical average parameters ranging from 0.4 to 0.2 occupy only five out of 12 scenarios (In comparison to nine out of 12 scenarios in the 15-minute forecasting). In other words, reliance on the historical average becomes larger in this range. There is a little modification in combining two decision factors corresponding to the new combination sets of parameters. The weighting systems in 45-minute forecasting show a little different scenarios from the systems in 30-minute forecasting. They are also based on the importance of historical averages in performing prediction. The remarkable change in 45-minute forecasting is the disappearance of a combination set with 0.2 of historical average parameters and the expansion of the sets with 0.4 of historical average parameters. Correspondingly, the order of decision factors are determined.

Table 4-8 Adaptive weighting systems for third combined model in 30-minute forecasting

L (%)	M	Scenarios	Upstream	H.A	Current
0-24	0-99	1	0.1	0.6	0.3
0-24	100-199	2	0.2	0.6	0.2
0-24	>200	3	0.3	0.6	0.1
25-49	0-49	4	0.1	0.5	0.4
25-49	50-99	5	0.2	0.5	0.3
25-49	100-199	6	0.3	0.5	0.2
25-49	>200	7	0.4	0.5	0.1
50-99	0-149	8	0.2	0.4	0.4
50-99	>150	9	0.4	0.4	0.2
100-299	0-149	10	0.2	0.3	0.5
100-299	>150	11	0.4	0.3	0.3
>300		12	0.4	0.2	0.4

Table 4-9 Adaptive weighting systems for third combined model in 45-minute forecasting

L(%)	M	Scenarios	Upstream	H.A	Current
0-24	0-99	1	0.1	0.6	0.3
0-24	100-199	2	0.2	0.6	0.2
0-24	>200	3	0.3	0.6	0.1
25-49	0-49	4	0.1	0.5	0.4
25-49	50-99	5	0.2	0.5	0.3
25-49	100-199	6	0.3	0.5	0.2
25-49	>200	7	0.4	0.5	0.1
50-199	0-99	8	0.2	0.4	0.4
50-199	100-199	9	0.3	0.4	0.3
50-199	>200	10	0.4	0.4	0.2
>200	0-149	11	0.2	0.3	0.5
>200	>150	12	0.4	0.3	0.3

4.4 Evaluation of the Multiple Combined Models

The three multiple combined models were comprehensively evaluated by means of traffic data collected during June 1993, from the L.I.E. in New York. The forecasts were calculated by the multiple combined models with adaptive heuristic weighting systems. The three upstream traffic, four current traffic, and/or historical averages at the concerned time period are the variables for the prediction. The resultant forecasts were compared to one another, as well as to the historical averages. The MAPE, RMSE, and Accuracy Ratio (Q-ratio) are used as measures of effectiveness for the prediction performance. The predictions in Figures 4-3 and 4-4 have used only three links on the downstream to compare the three forecasting ranges. In order to evaluate comprehensively the models, all the links will be tested here.

4.4.1 Evaluations of the combined models with peak-hour traffic data

The peak-hour traffic data between 6 A.M. and 9 A.M. were selected for evaluation of the multiple models. One-hundred forty data sets randomly selected from the entire network were used.

The mean values of three models in the three criteria are all lower than the historical averages only (Table 4-10). As expected, the third model has the lowest value among three models. Consider the distribution of mean value of MAPE between model III and historical average. In 95 percent accuracy range, the mean of model III is between 0.043 and 0.073, while the mean of historical average is between 0.0661 and 0.1. So, there are some areas of conflict between the two distributions. The portion greater than

Table 4-10 Comparison of three models and historical averages in 15-minute forecasting with peak-hour traffic data between 6 A.M. and 9 A.M.

		Model I	Model II	Model III	Hist. Avg.	Best model
MAPE	Mean (%)	6.8	6.34	5.86	8.51	2.50
	S.E.(mean)	0.0076	0.0070	0.0075	0.0096	0.0055
	Variance	0.0081	0.0068	0.0078	0.013	0.004
Q-ratio	mean	1.0695	1.064	1.0593	1.0874	1.031
RMSE	mean	414.55	385.01	376.95	518.48	239.4

0.0661 in model III is 15.8 percent, while the portion less than 0.073 in historical average is 11.7 percent. Intuitively, model III in MAPE has beaten 99 times out of 140 chances (70 percent). MAPE at model I and II have beaten 91 times (65 percent) and 97 times (68 percent) out of 140 chances, respectively. By the distribution of means and the intuitive comparison, the second and third model present almost equal accuracy of prediction performance. Meanwhile, the best combination out of 30 sets of scenarios combining the variables (9 for model I, 9 for model II, 12 and for model III) was selected from each set of data and the mean error rates out of those 140 randomly chosen data sets were computed. The mean error values of the best model are extremely low compared to the others. It is clear that the mean values of three criteria in the best model are statistically different from the other models. Unfortunately, it seems impossible to pick up the best combination from 30 sets of scenarios. Accuracy test (Q-ratio) and RMSE also shows almost same trend as MAPE.

There is no distinct difference between 30-minute forecasting and 15-minute forecasting as shown in Table 4-11. Indications are that 30-minute forecasting performs as accurately as 15-minute forecasting. Intuitively, MAPE of model III prevailed 75 percent over historical average, while the models I and II prevailed 60 percent and 64 percent over historical average, respectively.

Some different results emerged from the 45-minute forecasting in Table 4-12. The first model is the best among the three combined models. The second and third models are even worse than the historical average. This is due to the low reliability of current and

Table 4-11 Comparison of three models and historical averages in 30-minute forecasting with peak-hour traffic data between 6 A.M. and 9 A.M.

		Model I	Model II	Model III	Hist. Avg.	Best model
MAPE	Mean (%)	6.3	6.2	5.4	7.4	3.4
	S.E.(mean)	0.0065	0.0059	0.0064	0.008	0.006
	Variance	0.006	0.005	0.006	0.01	0.0056
Q-ratio	mean	1.0639	1.0541	1.0432	1.0683	1.0328
RMSE	mean	396.32	377.54	359.09	438.30	249.4

Table 4-12 Comparison of three models and historical averages in 45-minute forecasting with peak-hour traffic data between 6 A.M. and 9 A.M.

		Model I	Model II	Model III	Hist. Avg.	Best model
MAPE	Mean (%)	5.68	6.42	6.03	5.70	4.05
	S.E.(Mean)	0.0138	0.0143	0.0141	0.0130	0.0186
	Variance	0.0082	0.0089	0.0086	0.0080	0.0149
Q-ratio	mean	1.0577	1.0596	1.0593	1.0601	1.0265
RMSE	mean	357.95	383.30	375.68	371.74	265.00

upstream traffic values used in this farther prediction. The two variables are vulnerable because of unexpected ramp traffic and other interruptions between the *study link* and the *origin link*. The estimated, not observed, current traffic values from 15- and 30-minute forecasting also interfere with good prediction performance. Even though the best model has the lowest values among the above models, there is no remarkable difference between the historical average and the first model. It is noted that the first model includes the historical average. Probably, the historical average alone is good enough to predict the traffic in that range. The accuracy (Q-ratio) test shows a little different trend from the MAPE and RMSE. All the Q-ratio values of three combined models are lower than the historical average values. Also, the ratio of the best model is much lower than historical average.

4.4.2 Evaluations of the combined models with non-peak-hour traffic data

Next, the multiple combined models are evaluated using non-peak hour traffic data. First, the traffic data between 11 A.M. and 3 P.M. were randomly selected for evaluation of prediction performance. The results from 15-minute forecasting are shown in the Table 4-13.

All three criteria brought more accurate forecasts than historical average alone. The third model, of course, resulted in the best forecasts out of three models. The unique finding of Table 4-13 was that the first model predicted better than of the second model,

Table 4-13 Comparison of three models and historical averages in 15-minute forecasting
with traffic data between 11 A.M. and 3 P.M.

		Model I	Model II	Model III	Hist. Avg.
MAPE	Mean (%)	6.79	7.30	6.09	9.53
	S.E.(mean)	0.0044	0.0033	0.0036	0.0049
	Variance	0.0027	0.0016	0.0018	0.0035
Q-ratio	mean	1.0691	1.0767	1.0625	1.1028
RMSE	mean	322.18	333.96	280.59	452.95

which is in variation to the peak-hour test. It may be due to the non-peak hour traffic without any major change in traffic compared to the usual traffic flow. The historical average component of the first model plays a bigger role in prediction than the current traffic of the second model. Regardless of the invaluable current traffic component, the prediction performances have been improved by combining current traffic with the other two variables (*hist. avg.* and *upstream*).

In the MAPEs and RMSEs of Table 4-14, the forecasts of historical averages are better than the ones of the first and second models. It is noted that the combination of two variables in the 30-minutes range with non-peak hour data does not improve prediction performance. The variances of the MAPE in three models are also poorer than the values of historical average. The Q-ratios of the first and second models has equal or less values than historical average. Combining three variables has again improved the prediction performance.

Because the traffic data used in this study are collected only from westbound lanes of L.I.E. (CBD located at the east of the westbound), it is expected that there is not much traffic coming into this westbound freeway network in the evening peak-hour. In order to make comparisons with the non-peak-hour data corrected between 11 A.M. and 3 P.M., the models are evaluated with data collected between 5 P.M. and 8 P.M. The results of the 15-minutes forecasting performances are shown in the Table 4-15.

Table 4-14 Comparison of three models and historical averages in 30-minute forecasting
with traffic data between 11 A.M. and 3 P.M.

		Model I	Model II	Model III	Hist. Avg.
MAPE	Mean (%)	8.09	9.74	6.65	7.90
	S.E.(mean)	0.007	0.0088	0.0078	0.0059
	Variance	0.0076	0.0109	0.0087	0.005
Q-ratio	mean	1.0815	1.0974	1.0665	1.0816
RMSE	mean	304.86	347.81	270.96	336.75

Table 4-15 Comparison of three models and historical averages in 15-minute forecasting
with traffic data between 5 P.M. and 8 P.M.

		Model I	Model II	Model III	Hist. Avg.
MAPE	Mean (%)	5.86	5.47	4.48	4.69
	S.E.(mean)	0.005	0.004	0.0038	0.0051
	Variance	0.0035	0.0023	0.0021	0.0037
Q-ratio	mean	1.0625	1.0587	1.0466	1.0482
RMSE	mean	369.37	344.73	269.17	281.31

As shown in Table 4-15, the first and second models are much poorer at prediction than the historical averages. The third model also does not show any difference from the historical average in error rates. This proves that the models do not perform well under the uncongested traffic condition. The 30-minute performance is shown in Table 4-16.

Generally, the values of the error rates are higher than the values in the 15-minutes performance. For example, while the finest MAPE in the 30-minutes forecasting is 6.79 of the third model, the best one in the 15-minutes forecasting is 4.48 of the third model. Compared to the three models, the historical average error rate is equal or less.

4.4.3 Comparisons of the historical averages between 11 days and 22 days

All the precedent evaluations are based on the historical averages from 22 days traffic data (21 days data used). The 22-days based historical averages will be compared to the 11-days (10 days data used) based historical averages for prediction performance. Peak-hour traffic data (6 A.M. to 9 A.M.) is analyzed.

There are small differences between the two cases (Table 4-17). The addition or deletion of the number of days used for calculated of historical averages does not influence prediction performance.

Table 4-16 Comparison of three models and historical averages in 30-minute forecasting with traffic data between 5 P.M. and 8 P.M.

		Model I	Model II	Model III	Hist. Avg.
MAPE	Mean (%)	9.77	7.68	6.79	7.20
	S.E.(mean)	0.0046	0.0047	0.0040	0.0026
	Variance	0.003	0.0031	0.0023	0.001
Q-ratio	mean	1.1066	1.0842	1.0748	1.0773
RMSE	mean	523.54	457.28	405.81	386.18

Table 4-17 Comparison of prediction performance between 11 day-based historical average and 22-days-based historical average

		Days	H.A.	Model I	Model II	Model III
MAPE	Mean (%)	11	6.9	6.47	6.18	5.64
		22	6.6	6.29	6.18	5.45
	S.E.(mean)	11	0.011	0.009	0.008	0.009
		22	0.011	0.009	0.008	0.009
	Variance	11	0.009	0.006	0.004	0.005
		22	0.008	0.006	0.004	0.005
Q-ratio	mean	11	1.071	1.066	1.062	1.056
		22	1.068	1.063	1.062	1.054
RMSE	mean	11	470.98	410.54	377.54	365.71
		22	438.3	396.33	377.55	359.09

4.4.4 A final decision rule of the models for better prediction

Throughout the evaluations and comparisons of the three combined models, the third model generally seems to result in the best forecasts due to the many variables it comprised. However, the first or second model is often better than the third model. Current traffic conditions must be considered to select the best model. In order to accomplish that, monthly average travel times along the network were compared to the travel times under current traffic conditions. The comparison of the travel times was based on the entire network under consideration. An individual link could not be treated alone because some of congested link data were unfortunately not available due to detectors' malfunction. All the links along the network will be equally treated, regardless of the difference in degree of congestion of each link. If a specific day shows a traffic condition different from the monthly traffic condition, the traffic on that day has to be considered in a different way for better prediction. The rule of selecting a model in terms of current traffic conditions is as follows :

“If current travel time along the network is 25 percent longer than the monthly average travel times, use the second model for the prediction performance. Otherwise, use the third model for the prediction performance.”

This rule is based in the idea that under congested traffic conditions current traffic plays a heavier role than upstream and historical average traffic. The threshold value--25 percent-

-is determined by looking over the predictions from the evaluation of the models throughout various sets of data. The peak-hour traffic data is applied to this rule for 15-minutes forecasting in Table 4-18.

The error rate from final decision rule ("Final Dec." in Table 4-18) presents the best forecasts among five models including historical average, which means that the rule was effective in improving the prediction performance. The 30-minute forecasting range is also evaluated in Table 4-19. This range also resulted in a little improvement by applying the final decision rules.

4.4.5 Theoretical weighting vs. heuristic (adaptive) weighting

The combined models evaluated so far are based on an heuristic adaptive weighting system. The weighting values of the models were totally empirically obtained. A theoretical weighting technique (Bates and Granger, 1969) was used to compare performance with the heuristic weighting system developed in this study. The comparison of two techniques considered the combination of two alternative forecasting methods. The second model that included the upstream and current traffic was chosen for the comparison between the theoretical and heuristic weighting systems. The *upstream* and *current* traffic functioned as alternative forecasting models.

The purpose of the theoretical technique is to give more weight to the forecast model that produces the best results. Table 4-20 is an example of how this technique

Table 4-18 Evaluations of the values by final decision in 15-minute forecasting with traffic data between 6 A.M. and 9 A.M.

		Model I	Model II	Model III	Hist. Avg.	Final Dec.
MAPE	Mean (%)	9.03	8.45	8.12	10.4	7.00
	S.E.(mean)	0.010	0.009	0.010	0.0125	0.009
	Variance	0.015	0.012	0.014	0.022	0.0122
Q-ratio	mean	1.09	1.08	1.08	1.11	1.07
RMSE	mean	511.81	464.94	471.36	625.46	433.46

Table 4-19 Evaluation of the values by final decision in 30-minute forecasting with traffic data between 6 A.M. and 9 A.M.

		Model I	Model II	Model III	Hist. Avg.	Final Dec.
MAPE	Mean (%)	6.90	7.20	6.20	7.01	5.88
	S.E.(mean)	0.007	0.006	0.007	0.008	0.006
	Variance	0.008	0.006	0.007	0.01	0.005
Q-ratio	mean	1.069	1.072	1.062	1.071	1.059
RMSE	mean	429.18	419.19	397.84	474.31	365.50

Table 4-20 An example showing the theoretical weighting system

t	X_t	$\bar{X}_{1t}(1)$	$\bar{X}_{2t}(1)$	e_{1t}	e_{2t}	E_{1t}	E_{2t}	W_t	$\bar{X}_t(1)$
630	5927	5307	5504					0.5	5405
645	5949	5273	5805	-642	-446	412344	198470	0.325	5632
700	5854	5230	5874	-581	-49.4	749719	200911	0.211	5738
715	5554	5009	5760	-324	320.3	854708	303503	0.262	5563
730	5643	4993	5689	-634	117.3	1256360	317262	0.202	5549
745	5684	4988	5663	-691	5.1	1322174	118818	0.082	5607
800	5517	4924	5596	-529	145.7	1264926	137606	0.098	5530
815	5246	4662	5455	-322	350	1263892	157514	0.111	5367
830	4928	4484	5217	-266	526.6	933033	421062	0.311	4989
845	5276	4281	5190	-792	-59.2	1082598	424541	0.282	4934
900	5154	4724	5155	-873	35.7	1564408	404587	0.205	5066
915	4784	5019	5008	-59.9	370.6	1464039	419431	0.223	5010
930	5020	4875	5002	-1.06	-12.2	1393247	142272	0.093	4990
945	4494	4529	4776	380.7	507.6	910441	396425	0.303	4701
1000	4414	4073	4596	114.7	361.8	161656	526050	0.765	4196
1015	4409	3824	4489	-336	187.2	271053	423750	0.61	4083
1030	4330	4578	4387	-506	158.6	526734	448755	0.46	4475
1045	4161	4344	4287	416.7	225.9	555439	242128	0.304	4304
1100	4015	3571	4161	329.4	271.6	650772	184995	0.221	4031
1115	4526	4299	4280	-955	-365	1449890	283030	0.163	4283
1130	4379	4305	4329	-80.1	-98.9	1200629	267658	0.182	4324

applies when the first model is upstream traffic and the second model is the current traffic in time-series. X_t is observed traffic volume at time t . $\bar{X}_{1t}(1)$ and $\bar{X}_{2t}(1)$ are upstream traffic and current traffic at time t , respectively. These two values are the alternate forecasts at time $t+1$. $\bar{X}_t(1)$ is a forecast estimated by combining $\bar{X}_{1t}(1)$ and $\bar{X}_{2t}(1)$. W_t is a combining parameter that lies within $(0,1)$ and is varied at each interval. The W_t is related within the combined model as follows:

$$\bar{X}_t(\tau) = W_t \bar{X}_{1t}(\tau) + (1 - W_t) \bar{X}_{2t}(\tau)$$

The parameter W_t is selected so as to minimize the forecast error of the combined forecast is minimized. Bates and Granger (1969) present a simple option in seeking W_t . It is

$$W_t = \frac{E_2}{E_1 + E_2}$$

where

$$E_1 = \sum_{t=T-5}^T e_{1t}^2$$

$$E_2 = \sum_{t=T-5}^T e_{2t}^2$$

$$e_{1t} = \bar{X}_{1,t-1}(1) - X_t$$

$$e_{2t} = \bar{X}_{2,t-1}(1) - X_t$$

At time 630, W_t is assumed 0.5. $\bar{X}_{630}(1)$ is calculated

$$\bar{X}_{630}(1) = 0.5(5307) + 0.5(5504) = 5405$$

At time 645, the values of $\bar{X}_{1t}(1)$ and $\bar{X}_{2t}(1)$ are given by some other formula. The values of e_{1t} , and e_{2t} , are computed as follows :

$$e_{1t} = 5307 - 5949 = -642, \quad e_{2t} = 5504 - 5949 = -446$$

E_{1t} , and E_{2t} are squared values of e_{1t} , and e_{2t} , and hence W_t is

$$W_t = \frac{198470}{412344 + 198470} = 0.325$$

$$\bar{X}_t(1) = 0.325 (5273) + 0.675 (5805) = 5632$$

The system continues in this manner. The randomly selected traffic data are used to compare the theoretical and heuristic weighting system. The second combined model is activated for that comparison in Table 4-21.

The theoretical weighting systems described above do not produce a lower error rate than the heuristic weighting system, even though they are better than the historical average. This theoretical technique is limited to combining only two alternative forecasting methods--one of its major demerits. Since the third combined model, which includes *upstream*, *current*, and *hist. avg.*, usually performs better than the other two combined models, the forecasts resulting from the theoretical weighting system may be worse than the third model. The other problem of this theoretical weighting system is that it is not appropriate for several-step-ahead forecasting performance. Since the theoretical weighting system only gives more weight to the forecasting method with better results, it lacks the ability to adapt to changes in current traffic conditions. The adaptive heuristic weighting system investigates the trend of the traffic flow on the previous periods, thus it is more flexible in responding to sudden changes in traffic.

Table 4-21 The results of the prediction performance based on theoretical system, heuristic system, and historical average from 6:30 A.M. to 11:30 A.M.

		Theoretical	Heuristic	Historical average
MAPE	Mean (%)	5.55	4.87	6.77
	S.E.(mean)	0.0026	0.0024	0.0039
	variance	0.00095	0.00086	0.00213
RMSE	mean	297.20	279.78	356.78
Q-ratio	mean	1.0544	1.0508	1.0678

4.4.6 Evaluations of the multiple combined models in the network with closer distance between the links

To select more reliable upstream traffic values, the multiple combined models are now evaluated throughout the network with closer distance between the links. The network that has more links between the existing links is configured in Table 4-22.

Table 4-22 Configuration of the links added

Link	187	190	191	194	195	197	199	200	201	204	205	206	209	211
D		0.5	0.38	1.12	0.13	0.87	0.63	0.37	0.38	1	0.12	0.12	1	0.37

The first row is the number of the links, and the second row shows the distance between the links. The distance between link '187' and '190' is 0.5 miles. The link '187', '195', '201', '205', and '211' are the links already used for the evaluation of the models. The links between the existing links are chosen for another evaluation of the models. There are three links chosen between link '187' and link '195'. The three links also include the uninterrupted reliable traffic data. Notice that the three links are not the only links between the link '187' and the link '195'. In a way choosing the three upstream components, if the link '195' is called an *origin link* over the *study link*, the link '187' and '201' were the substantial upstream components for the prediction performance in the previous evaluation. However, based on the above configuration, another two link set will be chosen as upstream components out of the links '190', '191', '194', '197', '199', and

'200'. The first choice is a set of links '197' and '194'. The links '199' and '191' are the second choice as substantial upstream components. The third choice is the links '200' and '199'. The selection is determined based on the availability of the traffic data on the links selected.

The fifty peak-hour traffic data are used for evaluating the combined models under that network. The results from 15-minute forecasting performance are shown in Table 4-23. While *Approximated* mean the network shown in Table 4-1, *Detailed* mean the network shown in Table 4-22. The forecasts from both networks are statistically better than historical average in three criteria. Intuitively, in the comparison of model I the *detailed* prevailed 32 times of 50 chances (64 %) over the *approximated*. The model II and III of the *Detailed* have beaten 62 % and 42 % over the *Approximated*'s models, respectively. In the comparison of model III, the *Detailed* tied 14 times of 50 chances (28 %) with the *Approximated*. The model I and II of the *Detailed* tied only 6 % with the two models of the *Approximated*. The prediction from the *Detailed* seems a little better than the one from the *Approximated*. However, considering the distribution of the standard error of mean in MAPE, it is hard to say the forecasts from the *Detailed* are always better than the ones from the *Approximated*. That is, the upper side of the *Detailed* has large conflicted area with the lower side of the *Approximated*.

Table 4-23 The prediction results from two types of network: *Detailed* and *Approximated*

		Networks	H.A.	Model I	Model II	Model III
MAPE	Mean (%)	Approximated	9.49	6.65	6.58	6.11
		Detailed	9.49	5.56	5.68	5.61
	S.E.(mean)	Approximated	0.0119	0.0076	0.0073	0.0087
		Detailed	0.0119	0.0067	0.0063	0.0081
	Variance	Approximated	0.0200	0.0081	0.0076	0.0106
		Detailed	0.0200	0.0063	0.0055	0.0093
Q-ratio	mean	Approximated	1.0993	1.0679	1.0662	1.0616
		Detailed	1.0993	1.0571	1.0567	1.0565
RMSE	mean	Approximated	595.88	397.67	394.73	404.56
		Detailed	595.88	342.69	338.78	376.78

4.4.7 Comparison of the heuristic combined model with other forecasting model

The forecasts from the combined model developed here have been compared with historical average only. It is due to the finding that the historical average is as good as traditional forecasting model in terms of prediction performance. Urban traffic control is also dependent upon the historical traffic data. Attributes of the links or network used are a consideration for the comparison between the models. While the heuristic combined model focuses on the mainstream traffic on the freeway, some other models study the ramp traffic. Different time interval is another interruption for that comparison. Beyond the above reasons, there is a practical difficulty to compare the different two or more models. It is a lack of generally agreeable prediction model. Since there was no outstanding prediction model, the combined model was compared with historical average only. However, UTCS (Urban Traffic Control System) model was selected to compare with the models in the dissertation.

The UTCS-2 model is as follows:

$$\hat{V}_t = m_t + \gamma(m_{t-1} - f_{t-1}) + (1 - \alpha) \sum_{s=0}^{t-1} \alpha^s (f_{t-s-1} - m_{t-s-1}) \\ + \gamma(1 - \alpha) \sum_{s=0}^{t-2} \alpha^s (f_{t-s-2} - m_{t-s-2})$$

where

\hat{V}_t = predicted volume at time t ;
 m_t = historical volume at time t ;
 f_t = measured volume at time t ;
 α = constant (0.2);
 γ = smoothing coefficient (0.9).

The UTCS-2 model rely much on the difference of predicted volume and historical volume. The parameters such as α and γ are very important in combining the values as do in the combined model developed in the dissertation. The values (0.2 and 0.9) of the parameters were captured from traffic data in Washington D.C. The area-specific parameters are not adaptive for current traffic flow. Let us look at the prediction performances between the UTCS-2 model and the heuristic combined models. The peak-hour traffic data were used for that comparison.

Table 4-24 Comparison of combined models and the UTCS model

	M1	M2	M3	H.A.	UTCS
MAPE(%)	8.57	7.93	7.58	10.45	14.22
Q-ratio	1.087	1.080	1.076	1.106	1.143
RMSE	511.8	464.9	471.4	625.5	760.2

It is obvious that the combined models are superior to the UTCS model in prediction performance. In the three criteria, the UTCS model is even worse than the historical average. The results indicate that the adaptive weighting system performs overall better than the fixed weighting parameters.

4.5 Findings

By incorporating *upstream* traffic into the three combined models for prediction, the better forecasts than historical average are usually produced. Especially, the combined models performed well in the 15- and 30-minutes forecasting under peak-hours traffic conditions. However, the predictions by the combined models with non-peak-hours traffic data resulted in no improvements. Since there are not much changes in traffic on non-peak-hours compared to the daily average traffic, the models including two or three variables have not produced forecasts better than the historical average. Also, in the 45-minutes forecasting, the forecasts from the models were worse than the historical average. It is due to the variables used for the models in that farther forecasting. The *upstream* that were selected based on the current travel times are not reliable because of many factors caused by long trip from *origin link* to *study link*.

Whenever the MAPE rates are greater than 20 percent, the models have mostly prevailed the historical average in prediction performance. It represents that the combined models performs efficiently under the congested traffic condition. The historical average is not able to detect the changes in traffic on the roadways caused by an incident.

In the comparison of the forecasts between a heuristic (adaptive) weighting system and a theoretical weighting technique, the adaptive weighting system showed a little better results than the theoretical weighting system. The theoretical system consider only the

error values between the observed and the forecasts. Instead, the heuristic adaptive weighting system is based on more detailed investigations of the current traffic condition such as the variance of the previous periods and the differentiates between the historical average and the average value from the previous periods. Thus, the heuristic system is a more flexible and usable system for combining alternative forecasting methods.

5. CONCLUSIONS

Two types of prediction models have been developed through this research. One was based on link traffic information (link-based prediction model), the other on the network traffic information (network-based prediction model). The purpose of the two prediction models was to obtain forecasts for controlling traffic under congested conditions. The forecasts from the prediction models will be helpful for the assignments of the traffic on the congested roadways caused by incidents. The 6-minute traffic data collected from an urban arterial were used for developing and evaluating the link-based model, and the 15-minute traffic data collected from a suburban freeway were used for developing and evaluating the network-based model. In the development and evaluation of the two prediction models, traffic volume was used to represent traffic flow for prediction performance. Speed data was also used for the network-based model.

In the link-based model, the Box-Jenkins time-series technique was utilized to represent the traffic flow for prediction. The basic idea of the Box-Jenkins time-series technique is that previous time-series data play a major role in short-term forecasting. Actually, the time-series model and historical average were the components of the link-based prediction model. While the time-series data collected from the study link were used for considering the current traffic flow, the historical averages were used for including the usual traffic trend at a specific interval. In order to combine two

components, a heuristic weighting system was devised. The heuristic weighting system that adapts the fluctuated movements (up-and-down) of vehicles in 6-minute intervals produced reliable forecasts. The link-based model was generally successful at forecasting arterial traffic in 6-minute intervals. Especially, when there was a sudden change in current traffic such as 25-40 percent margin of 'DF' in Section 3-4, the model has performed efficiently for short-term prediction. The performance with the data under abrupt changes in traffic indicated that the heuristic weighting system was well developed to represent the movements of the vehicles in that conditions. A real-time diversion strategy can be an application to use that model. The link-based prediction model has raised a question--what if the 'DF' is greater than 40 percent. Thus, another approach has emerged to predict the movements of the vehicles.

In the network-based model, a new variable--*upstream*--has been added to investigate the vehicle flow along a continuous freeway network. The selection of upstream as a variable for the prediction model was based on the supposition that the upstream vehicle on main highways moves into downstream roadways without taking any alternate roads for diversion, especially in peak-hours. That is, the traffic on the upstream can be considered for the prediction of downstream traffic. The importance of *upstream* component was examined through the multiple combined models (network-based prediction model). The *upstream* traffic for the *study link* on the downstream was identified by investigating the current travel times along the network. The first prediction

model combined the *upstream* and the *historical average*, and the second prediction model combined the *upstream* and the *current*. The third model combined the three variables--*upstream*, *current*, and *historical average*. The relationships between the three variables were examined through three regression models incorporating three variables. The third regression model that has all the three variables was the best for prediction in terms of the error rates and their distributions. By adding more variables into the model, the quality of the model has been improved. It indicated that each of the three variables was necessary for the prediction performance. Three heuristic combined models were developed with the three variables selected.

The three combined models with the adaptive weighting system were evaluated through various data sets. The evaluations were classified into peak-hours and non-peak-hours. The predictions were performed in 15-, 30-, and 45-minutes ranges. The models with peak-hour data in the 15-minute forecasting were best in obtaining the forecasts. The forecasts from the three combined models were superior to the historical average in the 15- and 30-minutes prediction performance. The third model, as expected, produced the best forecasts among three models. The prediction performance in the 45-minutes range were equal or worse than historical average, because it was hard to select the proper *upstream* component for the *study link* on the downstream in the far upstream traffic. It might be caused by the interruptions between the *study link* and the *origin link* (location of the *upstream*) such as ramp traffic and incidents. In the comparison between the peak-

hour data and the non-peak-hour data, the performances with peak-hour data were better than the others. It indicated that the combined models were appropriate to predict traffic under congested conditions.

Meanwhile, in order to select the best among three combined models with an adaptive weighting system, the travel time through the entire network has been used again. The travel time at the present time is much longer than the monthly average travel times, the second model is selected rather than the third model for prediction performance. By applying the above rule, 14 percent for 15-minutes forecasting and 5 percent for 30-minutes forecasting (in MAPE) have been improved from the simple prediction strategy being selected the third model all the time. If the travel times are considered link by link, the prediction performance may be much more improved. The travel times on all the links throughout the entire network have been accumulated in those comparisons because of the missing speed data on some links. The second model including the *upstream* and the *current* has been used to compare the adaptive weighting system with the theoretical weighting system. The adaptive system was 12 percent better than the theoretical system in prediction performance. It concluded that the adaptive systems respond more promptly to the changes in traffic than the theoretical system. In the comparison of models with UTCS model, the adaptive system were better than the fixed weighting parameters in prediction performance. It is also noticed that the combined models including three variables are superior to the UTCS model including two variables. Considering all the

evaluations conducted in this research, the network-based models performed generally better than the historical average in short-range traffic prediction on the freeway. All the conclusions are based only to the INFORM data sets from New York state.

The prediction models developed in this study relied only on the current traffic information, not the future information. Even though a new variable--*upstream*--from an extended dimension was used for the network-based prediction model, three variables were also limited to current traffic information. As far as the components of a prediction model is dependent upon the current traffic information, the model might not produce much better forecasts than the ones produced by the combined models in this research. In the farther forecasting it was especially difficult to obtain the reliable forecasts from the prediction models.

The prediction models in the dissertation have been developed based on the limited data sets. Thus, it is hard to say that the parameters of the models are adaptable to any other data sets. The link-based model was built for traffic prediction on the urban arterial, and the network-based prediction model was built for traffic prediction on the freeway. This research does not specifies the link or network attributes for the prediction. The time interval of the data sets is another factor for consideration of the parameters in the models. The adjusted time interval requires modification of the network size.

5.1 Recommendations for Further Study

In the network-based models, a network (*Detail*) that has more links was used to obtain the forecasts. The forecasts from the *Detailed* network were compared to the ones from the *Approximated* network. Model I was 16 percent improved in the *Detailed* network. Model II and model III were improved 14 percent and 8 percent, respectively. If all the link data throughout the network are available, it will be possible to evaluate more deeply the network-based prediction model. In order to conduct such an evaluation, an advanced traffic data collection system is required. There is no doubt that reliable and prompt traffic data through the advanced data collection system are necessary for more efficient forecasting. The automatic vehicle identifications (AVI) system (Franklin, 1984) and video system (Dickinson and Waterfall, 1984 and Takaba et al., 1984) will be needed to collect better travel time data. In judging current traffic condition and identifying upstream traffic, speed data plays a big role. If data sets from various places are available in the future, the combined models developed can be evaluated in diverse ways. The availability of data collected from congested area caused by incidents will allow study of the vehicle movements under conditions caused by incidents (non-recurring congestion).

The 15-minute interval data were only used to develop and evaluate the network-based prediction model. The shorter interval data must be used to compare the prediction performance between the link-based model and the network-based model. In addition to that comparison, the traffic flow of the arterial and the freeway can be differentiated by

prediction performances. Although the ramp traffic along the main highways was ignored in this research, a complicated network including the ramps should be considered for better performance. The prediction models with the current traffic information may have a limitation to further improve the prediction performance. Virtually, another variables are required for the more reliable and accurate prediction. A new technique such as neural network analysis may be needed to approach this problem from a different direction.

The use of the prediction models is another topic for further study. The prediction models may help assign efficiently the vehicles onto the alternate roads, in case that the main highway is interfered with an incident. In order to realize that, an algorithm incorporating the prediction system and assignment technique is required.

REFERENCES

- Ahmed, M.S. and Cook, A.B. (1980), *Application of Time-Series Analysis Techniques to Freeway Incident Detection*, Transportation Research Record, 841, 19-21.
- Ahmed, M. and Cook, A (1979). *Analysis of Freeway Traffic Time Series Data Using .Box and Jenkins Techniques*, Transportation Research Record, 722, 1-8.
- Ahmed, S.A.(1989), *Operational Filtering of Traffic Data*, Journal of Forecasting, Vol. 8, 19-32.
- Ahrens, C.D. (1985), *Meteorology Today : An Introduction to Weather, Climate, and the Environment--Weather Forecasting (chapter 17)*, 2nd edition, West Pub. Co. 403-406.
- Akahane, H. and Koshi, M. (1986), *Short-Term Prediction of Inflow Volumes of Urban Freeways*, Second International Conference on Road Traffic Control, April, Institution of Electrical Engineers, 35-38.
- Barnston, A. G. and Livezey, R.E. (1988), *An Operational Multifield Analogy Antianalogy Prediction System for the United States Seasonal Temperature*, J. Geophys. res. ,93, 10, 953
- Bates J.M. and Granger, C.W.J. (1969), *The Combination of Forecasts*, Operation Research Quarterly, Vol. 20, No. 4, 319-325 and 451-468.
- Bennett, R.J. (1979), *Spatial Time Series (Analysis-Forecasting-Control)*, Pion Limited, London, 533-610.
- Bhouri, N., Hadj-Salem, H., Papageorgiou, M., and Blosseville, J-M. (1990), *Estimation of Traffic Density on Motorways*, Traffic Engineering and Control, 580-583.
- Box, G.E.P. and Jenkins, G.M.(1976). *Time Series Analysis : Forecasting and Control*, San Francisco : Holden-Day.
- Danech-Pajouh M. and Aron, M. (1992), *ATHENA : A method for Short-Term Inter-Urban Motorway Traffic Forecasting*, Recherche Transports Securite, No. 6, 11-16.

- Dash, P.K., Dash, S., and Rahman, S. (1993) *A Hybrid Artificial Neural Network-Fuzzy Expert System for Short Term Load Forecasting*, Expert System Application to Power Systems IV, La Trobe University, Melbourne, Australia, T.S. Dillon (editor), 175-180.
- Davis, G.A. and Nihan, N.L. (1984), *Using Time-Series Designs to Estimate Changes in Freeway Level of Services, Despite Missing Data.*, Transportation Research A., Vol. 18A, No. 5/6, 431-438.
- Davis, G.A. and Nihan, N.L. (1991), *Nonparametric Regression and Short-Term Freeway Traffic Forecasting*, Journal of Transportation in American Society Civil Engineers (ASCE), 178-187.
- Davis, G.A., Nihan, N.L., Hamed, M.M., and Jacobson, L.N.(1990). *Adaptive Forecasting of Freeway Traffic Congestion*, Transportation Research Record, 1287, 29-33.
- Dickinson, K.W. and Waterfall, R.C. (1984), *Video Image Processing for Monitoring Road Traffic*, Institution of Electrical Engineers, International Conference on Road Traffic Data Collection 105-109.
- Flankin, P.(1984), *Automatic Vehicle Identification Using a Profile Detector*, Institution of Electrical Engineers, International Conference on Road Traffic Data Collection, West Yorkshire Metropolitan County Council, U.K. 27-31.
- Garbard, J.F., Henry, J.J., Abours, S., and Lesort, J.B. (1986), *CALIFE: On-Line Calculation of Fixed Time Plans*, IFAC Control in Transportation Systems, Vienna, Austria, 233-237.
- Granger C.W.J. (1989), *Invited Review Combining Forecasts-Twenty Years Later*, Journal of Forecasting, Vol. 8, 167-173.
- Granger C.W.J. and Ramanathan, R. (1984), *Improved Methods of Combining Forecasts*, Journal of Forecasting, Vol. 3, 197-204.
- Hobeika, A.G., Sivanandan, R., Kim, C., Subramaniam, R., and Ozbay, K.(1992). *Real-Time Diversion Strategies for Congested Urban Networks*, U.S. DOT/VDOT

- Centers Program, Mid-Atlantic Universities Transportation Center (MAUTC), Phase 4 working paper, University Center for Transportation Research, Virginia Polytechnic Institute & State University.
- Hobeika, A.G., Sivanandan, R., Zhang, Y., Subramaniam, R., and Ozbay, K.(1991). *Real-Time Diversion Strategies for Congested Urban Networks*, U.S. DOT/VDOT Centers Program, Mid-Atlantic Universities Transportation Center (MAUTC), Phase 3 Project Report, University Center for Transportation Research, Virginia Polytechnic Institute & State University, October.
- Hobeika, A.G., Sivanandan, R., Zhang, Y., and Ozbay, K.(1990). *Optimal Diversion Strategies for a Modified Urban Network*, U.S. DOT/VDOT Centers Program, Mid-Atlantic Universities Transportation Center (MAUTC), Phase 2 Project Report, University Center for Transportation Research, Virginia Polytechnic Institute & State University, October.
- Hobeika, A.G., Sivanandan, R., Subramaniam, R., Ozbay, K. and Zhang, Y (1992). *Real-Time Traffic Diversion Model : A Conceptual Approach*, accepted for ASCE Journal of Transportation Engineering.
- Hoffman, G. and Janko, J. (1990), *Travel Times as a Basic Part of the LISB Guidance Strategy*, Third International Conference on Road Traffic Control, May, Institution of Electrical Engineers 6-10.
- Kawashima, H. (1986), *Long Term Prediction of Traffic Flow*, IFAC Control in Transportation Systems, Vienna, Austria, 75-82.
- Kendall, S.M. and Ord, J.K. (1990), *Time Series*-Third Edition, Oxford University Press.
- Kerr, R.A. (1989), *A New Way to Forecast Next Season's Climate*, Science, Vol 244 April, 30-31.
- Kerr, R.A. (1989), *Telling Weatherman When to Worry*, Science, Vol 244, June, 1137-1140.
- Kim, C. and Hobeika, A.G. (1993), *A Short-Term Demand Forecasting Model from Real-Time Traffic Data*, Proceedings of Two Parallel Conference : Infrastructure

- Planning and Management, American Society of Civil Engineers (ASCE), Denver, Colorado, 540-549.
- Kim, M.W.(1993), *Neural Networks : what they are and what they do*, KSEA (Korean Scientists and Engineers Association) Letters, Vol. 22, No.1, 19-22.
- Kreer, J.B. (1975), *A comparison of Prediction Algorithms for Computerized Traffic Control Systems*, Traffic Engineering, April, 51-56.
- Kreer, J.B. (1976), *Factors Affecting the Relative Performance of Traffic Responsive and Time-of-Day Traffic Signal Control*, Transportation Research, Vol. 10, 75-81.
- Kyte, M., Marek, J, and Frith, D.(1989) *Using multivariate Time Series Analysis to Model Freeway Traffic Flow*, Presented at the 1989 Annual Meeting, Transportation Research Board, Washington D.C., 1-32.
- Levenbach, H. and Cleary, J.P. (1984) *The Modern Forecaster-The Forecasting Process through Data Analysis*, Lifetime Learning Publications, Belmont, California.
- Levin, M. and Tsao, Y.D.(1980). *On Forecasting Freeway Occupancies and Volumes*, Transportation Research Record 773, 47-49.
- Lu, J. (1990), *Prediction of Traffic Flow by an Adaptive Prediction System*, Transportation Research Record, 1287, 54-61.
- Maes, W. (1982), *Box-Jenkins Time Series Analysis of Saturated Freeway Traffic Data*, *Time Series Analysis : Theory and Practices 1*, O.D. Anderson (editor), 287-296.
- Montgomery, D.C., Johnson, L.A., and Gardiner, J.S. (1990), *Forecasting & Time Series Analysis (second edition)*, McGraw-Hill, Inc.
- Moorth, C.K. and Ratcliffe, B.G.(1988). *Short Term Traffic Forecasting Using Time Series Methods*, Transportation Planning and Technologies, Vol. 12, 45-56.
- Nihan, N.L. and Holmesland K.O. (1980), *Use of the Box and Jenkins Time Series Technique in Traffic Forecasting*, Transportation, 9, 125-143.
- Nikisch, J.A. and Piotrowski, J (1984), *Errors in Automatic Data Collection System SCR-5*, Institution of Electrical Engineers, International Conference on Road Traffic Data Collection, Technical University of Poznan, Poland, 32-34.

- Oda, T. (1990), *An Algorithm for Prediction of Travel Time Using Vehicle Sensor Data*, Third International Conference on Road Traffic Control, May, Institution of Electrical Engineers.
- Okutani, I. and Stephanedes, Y.J. (1984), *Dynamic Prediction of Traffic Volume through Kalman Filtering Theory*, Transportation Research B, Vol. 18B, 1-11.
- Ott, L., (1988), *An Introduction to Statistical Methods and Data Analysis*, Third edition, PWS-Kent Publishing Company, Boston.
- Pankratz, A. (1983), *Forecasting with Univariate Box-Jenkins Models-Concept and Cases*, John Wiley and Sons.
- Polhemus, N.W. (1980), *The Construction and Use of Continuous Auto Regressive Model for Traffic Indices*, Transportation Research B., Vol. 14B, 271-279.
- Saha, S.K. (1990), *The Development of a Procedure to Forecast Traffic Volumes on Urban Segments of the State and Interstate Highway Systems*, Ph.D. Dissertation, Graduate School of Purdue University.
- Stephanedes, Y.J., Kwon, E., and Michalopoulos, P. (1989), *On-Line Diversion Prediction for Dynamic Control and Vehicle Guidance in Freeway Corridors*, Transportation Research Record, 1287, 11-19.
- Stephanedes, T.J., Michalopoulos, P., and Plum, R.A. (1981), *Improved Estimation of Traffic Flow for Real-Time Control*, Transportation Research Record, 795, 28-39.
- Takaba, S., Sekine, T., and Hwang (1984), B.W., *A Traffic Flow Measuring System Using A Solid-State Image Sensor*, Institution of Electrical Engineers, International Conference on Road Traffic Data Collection, University of Tokyo, Japan, 110-114.
- Thomopoulos, N.T. (1980), *Applied Forecasting Methods*, Prentice-Hall, Inc., Englewood, N.J.

VITA

Changkyun Kim was born in Seoul, Korea, on January 5, 1963. He received his Bachelor of Science degree in Civil Engineering at Polytechnic University in 1989. On Fall of 1989, he began his graduate study at the Virginia Polytechnic Institute and State University. He received Master of Science degree in Civil Engineering in 1991. During his stay at the Virginia Polytechnic Institute and State University, he was working for the Center for Transportation Research (IVHS Research Center of Excellence) as a graduate research assistant. He has been awarded by ITE southern district due to best student paper and presented a couple of papers to the major transportation conferences. He plans to return to his country and to work in the area of traffic control and IVHS. His research interests are IVHS, traffic control, and transportation system management.

김 창 준

ChangKyun Kim