

# **A HYBRID MODEL-BASED AND MEMORY-BASED SHORT-TERM TRAFFIC PREDICTION SYSTEM**

A Thesis

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Master of Science in  
Civil Engineering

in

The Department of  
Civil and Environmental Engineering

By

Ciprian-Danut Alecsandru  
B. Engr., University "Politehnica" of Bucharest, 1997  
August 2003

## **ACKNOWLEDGEMENTS**

I am grateful to my advisor, Dr. Sherif Ishak for the ideas that led to this work, for his valuable comments, guidance, support and patience, throughout the course of this work. I thank Dr. Chester Wilmot and Dr. Brian Wolshon for their thoughtful comments along the initial proposal for this thesis topic and for being on my defense committee. Also, I want to acknowledge the successful collaboration with Prashanth Kotha on some preliminary research studies. Last but not the least I thank everyone who has remotely helped in the successful completion of this work.

I dedicate this thesis to my dearest lovely and supportive wife, Anca-Alexandra, and to my family who gave me the early and the most important education.

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
ABSTRACT.....	viii
1 INTRODUCTION AND PROBLEM STATEMENT .....	1
2 OBJECTIVES .....	3
3 LITERATURE REVIEW .....	4
3.1 MODEL-BASED APPROACHES .....	4
3.2 MEMORY-BASED APPROACHES .....	10
4 BACKGROUND.....	13
4.1 ARTIFICIAL NEURAL NETWORKS .....	14
4.2 CASE-BASED REASONING (CBR) .....	26
5 METHODOLOGY .....	29
5.1 STUDY AREA AND DATA COLLECTION .....	30
5.2 TRAINING .....	33
5.3 TESTING.....	35
5.4 HYBRID TRAFFIC PREDICTION SYSTEM .....	36
5.5 MODEL-BASED APPROACH.....	38
5.6 MODEL-BASED CBR APPROACH.....	40
5.7 MEMORY-BASED CBR APPROACH.....	41
5.8 INTEGRATION OF MODEL-BASED AND MEMORY-BASED PREDICTIONS....	44
5.9 APPLYING THE PROPOSED METHODLOGY .....	48
6 INTERPRETATION OF THE RESULTS.....	50
6.1 OPTIMAL SETTINGS OF THE MODEL-BASED APPROACH .....	50
6.2 MEMORY-BASED PREDICTION AND OPTIMAL SETTINGS .....	51
6.3 PERFORMANCE ENVELOPES .....	62
6.4 EFFECT OF PREDICTION METHOD AND LTM COMPONENT .....	65
6.5 EFFECT OF CBR ACROSS TRAFFIC CONDITIONS .....	68
6.6 COMPARISON BETWEEN OPTIMAL AND NON-OPTIMAL PERFORMANCE ..	69
7 SUMMARY AND CONCLUSIONS.....	72
8 FUTURE WORK .....	77
REFERENCES .....	78

APPENDIX A – *AARE*-BASED WILCOXON SIGNED RANK TEST RESULTS ..... 82

APPENDIX B – *RMSE*-BASED WILCOXON SIGNED RANK TEST RESULTS ..... 84

APPENDIX C – *AARE*-BASED FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE  
RESULTS ..... 86

APPENDIX D – *RMSE*-BASED FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE  
RESULTS ..... 87

VITA..... 88

## LIST OF TABLES

TABLE 1 NN OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON <i>AARE</i> .....	53
TABLE 2 NN OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON <i>AARE</i> .....	54
TABLE 3 NN OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON <i>RMSE</i> .....	55
TABLE 4 NN OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON <i>RMSE</i> .....	56
TABLE 5 HYBRID OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON <i>AARE</i> .....	57
TABLE 6 HYBRID OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON <i>AARE</i> .....	58
TABLE 7 HYBRID OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON <i>RMSE</i> .....	59
TABLE 8 HYBRID OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON <i>RMSE</i> .....	60

## LIST OF FIGURES

FIGURE 1: EXAMPLE OF ARTIFICIAL NEURAL NETWORK.....	14
FIGURE 2: EXAMPLE OF MLP NETWORK TOPOLOGY .....	18
FIGURE 3: EXAMPLE OF THE MODULAR NETWORK TOPOLOGY .....	19
FIGURE 4: EXAMPLE OF THE PCA NETWORK TOPOLOGY .....	20
FIGURE 5: EXAMPLE OF CANFIS NETWORK TOPOLOGY .....	21
FIGURE 6: EXAMPLE OF JORDAN/ELMAN NETWORK TOPOLOGY .....	23
FIGURE 7: EXAMPLE OF PRN NETWORK TOPOLOGY.....	24
FIGURE 8: EXAMPLE OF TLFN NETWORK TOPOLOGY .....	25
FIGURE 9: ILLUSTRATION OF A TYPICAL CBR SYSTEM .....	27
FIGURE 10: MAP OF I4 IN ORLANDO, FLORIDA.....	32
FIGURE 11: HYBRID TRAFFIC PREDICTION SYSTEM OUTLINE .....	38
FIGURE 12: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 5-MINUTE PREDICTION HORIZON .....	45
FIGURE 13: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 10-MINUTE PREDICTION HORIZON .....	45
FIGURE 14: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 15-MINUTE PREDICTION HORIZON .....	45
FIGURE 15: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 20-MINUTE PREDICTION HORIZON .....	46
FIGURE 16: PREDICTION QUERY MANAGER DECISION MECHANISM .....	47
FIGURE 17: <i>AARE</i> -BASED PERFORMANCE OPTIMIZATION BY PREDICTION METHOD.....	63
FIGURE 18: <i>RMSE</i> -BASED PERFORMANCE OPTIMIZATION BY PREDICTION METHOD.....	63
FIGURE 19: PERFORMANCE ENVELOPES BASED ON <i>AARE</i> .....	64
FIGURE 20: PERFORMANCE ENVELOPES BASED ON <i>RMSE</i> .....	64

FIGURE 21: FREQUENCY OF OPTIMAL PREDICTION METHODS WITH PREDICTION HORIZONS BASED ON <i>AARE</i> .....	66
FIGURE 22: EFFECT OF TIME FACTOR ON OPTIMAL SETTINGS .....	67
FIGURE 23: <i>AARE</i> -BASED OPTIMAL CBR PREDICTION ACROSS TRAFFIC CONDITIONS .....	68
FIGURE 24: <i>RMSE</i> -BASED OPTIMAL CBR PREDICTION ACROSS TRAFFIC CONDITIONS .....	69
FIGURE 25: PERCENTAGE REDUCTIONS IN <i>AARE</i> .....	70
FIGURE 26: REDUCTIONS IN <i>AARE</i> BY OPTIMIZATION WITH TRAFFIC CONDITIONS .....	71

## **ABSTRACT**

Short-term traffic forecasting capabilities on freeways and major arterials have received special attention in the past decade due primarily to their vital role in supporting various travelers' trip decisions and traffic management functions. This research presents a hybrid model-based and memory-based methodology to improve freeway traffic prediction performance. The proposed methodology integrates both approaches to strengthen predictions under both recurrent and non-recurrent conditions. The model-based approach relies on a combination of static and dynamic neural network architectures to achieve optimal prediction performance under various input and traffic condition settings. Concurrently, the memory-based component is derived from the data archival system that encodes the commuters' travel experience in the past. The outcomes of the two approaches are two prediction values for each query case. The two values are subsequently processed by a prediction query manager, which ultimately produces one final prediction value using an error-based decision algorithm. It was found that the hybrid approach produces speed estimates with smaller errors than if the two approaches employed separately. The proposed prediction approach could be used in deriving travel times more reliable as the Traffic Management Centers move towards implementing Advanced Traveler Information Systems (ATIS) applications.

# **1 INTRODUCTION AND PROBLEM STATEMENT**

One of the critical components contributing to the success of Advanced Traffic Management Systems (ATMS) is the dissemination of real-time and predictive information of traffic conditions. With the remarkable advancement made in data collection and dissemination technology, the traveling public now has the ability to receive real-time traffic information at both pre-trip planning stage and en-route. With accurate and reliable information, travelers can make appropriate decisions to bypass congested segments of the network and change departure times and/or destination, whenever appropriate. Such decisions are likely to affect the travel demand at various points of the network and provide opportunities for better utilization of the existing transportation infrastructure capacity. Moreover, Traffic Management Centers need traffic information to support their primary real-time management and control functions.

As a primary source of mobility, urban freeways are heavily traveled by both commuters and non-commuters who continuously experience excessive delays and queuing conditions on a daily basis. Freeway travelers often seek information on traffic conditions in the form of travel times and delays along their trips. Such information provides an indirect measure of travel cost, which most travelers seek to minimize. Since travel decisions are often made in advance, prior to the onset of the trip, travelers often seek predictive information on traffic conditions along their selected routes and within the expected duration of the trip. Research in the past few years has addressed this need with a variety of short-term traffic prediction models that attempt to capture the dynamic nature of traffic conditions. A review of the literature on previously developed models and research activities in this area reveals that there is still a need to improve the short-

term traffic prediction performance and the capability of existing models to predict the evolution of traffic conditions under both recurrent and non-recurrent conditions.

The last decade has witnessed a surge of traffic prediction models that were primarily developed to forecast traffic conditions in short-term horizons (usually 5 to 30 minutes). The dissemination of forecasted traffic conditions information will essentially impact travelers' decisions at the pre-trip planning stage and en-route. Such decisions include destination selection, route choice, mode choice, and departure times. Accuracy and reliability of forecasted information are of utmost importance to achieve better redistribution of demand and maximum utilization of the existing infrastructure capacity. Theoretically, user-equilibrium assignment, where travelers seek to minimize their average cost (travel time), cannot be achieved without a robust predictive traffic information system.

The extensive literature on this subject reveals that current traffic prediction models have not achieved a satisfactory level of performance yet due to their inability to cope with predictions under recurrent and non-recurrent conditions. No study has attempted so far to integrate the predictive capabilities of models under both recurrent and non-recurrent conditions. Furthermore, no research study has attempted to optimize the prediction performance under different traffic conditions. To address both performance optimization and predictions under recurrent and non-recurrent conditions, this research developed a hybrid model-based and memory-based approach to optimize the performance of short-term traffic prediction systems under a wide spectrum of traffic conditions. The model-based approach was optimized to improve prediction performance under non-recurrent conditions, while the memory-based approach was used to improve predictions under recurrent conditions.

## **2 OBJECTIVES**

The main goal of this research is to improve the performance of a freeway traffic prediction system by accomplishing four major objectives:

1. Develop a model-based system to predict traffic conditions from recent past information using a combination of artificial neural networks.
2. Optimize the prediction performance of the model-based approach under different topological and traffic condition settings.
3. Develop a memory-based system using a Case-Based Reasoning (CBR) approach to predict recurrent traffic conditions.
4. Combine both approaches to develop a hybrid system to predict recurrent and non-recurrent traffic conditions.

The usefulness of this research resides in the fact that a more reliable travel time estimation procedure can be derived based on it, since the point estimates error are from the beginning much smaller than by employing each approach separately.

### **3 LITERATURE REVIEW**

The concept of Intelligent Transportation Systems (ITS) has introduced several functions and user services that have the potential to improve the efficiency, safety and productivity of the surface transportation system. Among them is our ability to make better travel decisions when current and predicted information becomes available. This concept has motivated researchers to seek traffic prediction models that are capable of forecasting traffic flow, speed, and travel times in short-term horizons (e.g. 5 to 30 minutes). In the remaining sections of this chapter a wide-ranging literature review on the topic of traffic prediction is presented. It was found that most of the research studies have used model-based approaches, while there are few attempts to employ memory-based approaches, such as CBR, to traffic prediction problems but rather to planning problems.

#### **3.1 MODEL-BASED APPROACHES**

Several research efforts have been conducted in the past few years to support ITS applications and provide travelers with travel time information at the pre-trip planning stage and en-route. Kaysi et al. (1996) and Ben-Akiva et al. (1991) recommended that traffic routing strategies under recurring and non-recurring congestion to be based on forecasting future traffic conditions rather than historical and/or current traffic conditions. This is because travelers' decisions are affected by anticipated traffic conditions rather than current traffic conditions.

Several prediction methods have been implemented in research in the past two decades. Ben Akiva et al. (1991) grouped those methods into three categories: (a) statistical models, (b) macroscopic models, and (c) route choice models based on

dynamic traffic assignment. One of the statistical techniques that has strong potential for on-line implementation is the time series approach. Time series models have been extensively used in traffic forecasting for their simplicity and strong potential for on-line implementation.

For example, a time series approach has been used by Stamatiadis and Taylor (1994) to predict network travel times for implementing dynamic route guidance. However, the authors have used a simulation model to produce travel times on a simulated network and there was no realistic travel time data to compare with the predicted values. Another time-series approach for short-term traffic prediction was investigated by Ishak and Al-Deek (2002). They investigated the factors that have a significant impact on the forecasting accuracy of travel times using a nonlinear time series traffic prediction model.

In 1982 Ahmed and Cook used Auto-Regressive Integrated Moving-Average time series (ARIMA) algorithm to analyze freeway traffic data. Chang and Miaou (1999) used simulated data from CORSIM and real data collected from signalized intersections, to examine the potential of using the generalized linear model to predict traffic flows and provide prediction bounds. They derived recursive algorithms based on the quasi-likelihood principle to perform on-line, multiple-step-ahead predictions of short-term arrival flows for signalized intersection.

Gazis and Knapp (1971) introduced a procedure for estimating density by first estimating travel time. The density estimate is then obtained from the travel time. The method is complicated because it requires the solution of a two point boundary value problem. Furthermore, extensive lane changing or accidents may cause significant errors in the travel time algorithm. In another study by Hamed et al. (1995), the Box-Jenkins

techniques were used to develop time-series models for forecasting traffic volumes on urban arterials. The algorithm they proposed uses the last forecasted value and the current traffic observation to predict traffic volume through the use of a 1-minute data collected on five major arterials in Amman, Jordan's capital city.

Jiang (1999) found in his study that a Kalman predictor in combination with the first-order autoregressive time series provides satisfactory dynamic predictions of work zone traffic flow. Jiang's approach can be used to trigger the decision of taking appropriate measures to avoid congestion when the predicted traffic flow rate is equal to or greater than the traffic capacity. Lee and Fambro (1999) studied the use of four time-series models of one-step-ahead volume forecasting using a 5-minute interval data. Okutani and Stephanedes (1984) proposed two models employing Kalman filtering theory for predicting short-term traffic volume. In their approach the prediction parameters are improved using the most recent prediction error and the volume prediction on a link is achieved by taking into account data from a certain number of links.

Nihan and Davis (1989) examined the use of prediction error and maximum likelihood techniques to estimate intersection turning and through movement probabilities from entering and exiting counts. For situations where full information on turning movement counts is available, a maximum likelihood estimator is derived and used as a component for a maximum likelihood algorithm. Several algorithms based on minimizing the error between observed and predicted exiting counts are also developed. The study used real traffic data collected to develop simulations for evaluating various estimators. The authors found that the maximum likelihood algorithm produced biased but more efficient estimates, while prediction error minimization approaches produced unbiased but less efficient estimates.

In a study by Lu (1990) a traffic control system model for traffic flow is described, and the importance of the accuracy of the prediction model is emphasized. In order to validate the adaptive prediction system, a sine function is used to simulate traffic flow as input to the adaptive prediction system. The adaptive prediction system is applied to actual traffic flow data collected from a highway network. As part of the effort to improve traffic control at construction zones, this study applied the time series theory and Kalman filtering theory to adaptively predict traffic flow at the construction zones on Indiana's freeways with real-time data. It was found that using the Kalman predictor in combination with the autoregressive process of time series could provide satisfactory dynamic predictions of construction zone traffic flow.

Recently, Chen and Chien (2001) conducted a study using probe vehicle data to compare the prediction accuracy under direct measurements of path-based travel time versus link-based travel times. The study showed that under recurrent traffic conditions, path-based prediction is more accurate than link-based prediction. Chien and Kuchipudi (2002) presented the results of using real-time and historical data for travel time prediction. Another study by Kwon et al. (2000) used an approach to estimate travel time on freeways derived from flow and occupancy data from single loop detectors and historical travel time information. Forecasting ranged from a few minutes into the future up to an hour ahead. The study showed that current traffic conditions are good predictors for the near future, up to 20 minutes, while long-range predictions need the use of historical data.

Lately, several studies have investigated the use of artificial neural networks to model short-term traffic prediction. For instance, Park and Rilett (1998) proposed two modular Artificial Neural Networks (ANN) models for forecasting multiple-period

freeway link travel times. One model used a Kohonen Self Organizing Feature Map (SOFM) while the other utilized a fuzzy c-means clustering technique for traffic patterns classification. Rilett and Park (1999) proposed a one-step approach for freeway corridor travel time forecasting rather than link travel time forecasting. They examined the use of a spectral basis neural network with actual travel times from Houston, Texas.

Another study by Abdulhai et al. (1999) used an advanced time delay neural network (TDNN) model, optimized using a Genetic Algorithm, for traffic flow prediction. The results of the study indicated that prediction errors were affected by the variables pertinent to traffic flow prediction such as spatial contribution, the extent of the loop-back interval, resolution of data, and others. Lint et al. (2002) presented an approach for freeway travel time prediction with state-space neural networks. Using data from simulation models, they showed that prediction accuracy was acceptable and favorable to traditional models. Several other studies applied neural networks for predicting speed, flows, or travel times. For instance, Park et al. (1999) used a spectral basis artificial neural network (SNN) to predict link travel times for one to five time periods ahead (of 5-minute duration). They used traffic data collected from the TransStar system implemented in Huston, TX. They found that the NN approach outperformed other statistical and heuristic approaches like Kalman filtering model, exponential smoothing model and historical profile.

In a study by Maschavan Der Voort et al. (1996) a hybrid method of short-term traffic forecasting is introduced. The technique uses a Kohonen self-organizing map as an initial classifier and each class has an individually tuned ARIMA model associated with it, therefore was called KARIMA. It is believed that the explicit separation of the tasks of classification and functional approximation improves the forecasting

performance, as compared to either a single ARIMA model or a backpropagation neural network. The model is tested with data from a French motorway, by forecasting traffic flow at horizons of 30 and 60 minutes.

Zhang et al. (1997) have trained a multilayer feed-forward artificial neural network to address the freeway traffic system identification problem. For this purpose the authors have used simulated traffic data from an artificially generated freeway. Several scenarios have been generated, such as different demand patterns and randomly generated incidents. The speed has been predicted at one time-step prediction horizon of 15 seconds duration. The solution has been developed with the purpose of building an improved freeway traffic model that could be used for developing real-time predictive control strategies for dynamic traffic systems.

Recently, Zhang (2000) developed a recursive traffic flow prediction algorithm using neural networks. The system prediction model is specified based on the understanding of how disturbances in traffic flow are propagated. Although the methodology presented has the advantage of its applicability to other linear and nonlinear function approximation predictors than neural networks, it also has a shortcoming. The prediction is made at one time step horizon of duration of 30 seconds. The practicability of using such short prediction horizons or the effect of increasing the time step size was not considered.

In a study by Yasdi (1999) the effectiveness of a neural network model for prediction of traffic volume based on time series data is presented. A dynamic artificial neural network, namely a Jordan-Elman recurrent network, was employed in this study to predict weekly-, daily-, and hourly-based traffic volume. Fu and Rilett (2000) presented an artificial neural network based method for estimating route travel times between

individual localities in an urban traffic network. The methodology developed in this study assumes that route travel times are time-dependent and stochastic and their means and standard deviations have to be estimated.

In a recent study by Ishak et al. (2003A and 2003B) an optimized neural network-based methodology for short-term prediction horizons of traffic conditions is presented. They found that the performance of different neural networks families can be improved if traffic conditions and the number and type of the inputs are considered. Up to 20-minute point speed predictions are performed using the real traffic data and significant improvements are demonstrated.

### **3.2 MEMORY-BASED APPROACHES**

In the transportation field, CBR approaches have been employed in a few research studies in the past as a planning tool for intelligent transportation systems. For instance; Khattak and Kanafani (1996) developed a Case-Based Reasoning (CBR) methodology for PLANITS (Planning and Analysis Integration for Intelligent Transportation Systems). In their research study, in order to address a current transportation planning situation, CBR presents similar historical cases. Specifically, it estimates the impacts of proposed transportation improvement actions, including Intelligent Transportation Systems, based on previous experiences with similar actions. In this paper, a hierarchical structure for representing historical cases is developed. All historical cases consist of transportation improvement actions, performance measures and environments defined in terms of their spatial, temporal and user/traveler dimensions. The authors found that overall the structure for the CBR is flexible and incorporates different stakeholder preferences for alternative transportation improvement actions and evaluation criteria.

Sadek et. al. (1999 and 2001) presented the use of CBR in real-time freeway traffic routing and traffic flow management applications. These studies investigated the potential of using case-based reasoning to overcome the limitations of the existing approaches to developing real-time routing strategies. The authors have tested the feasibility of this approach by developing and evaluating a prototype CBR routing system for the interstate network in Hampton Roads, Virginia. Cases for building the system's case-base are generated using a heuristic dynamic traffic assignment (DTA) model designed for the region. Using a second set of cases, the study evaluates the performance of the prototype system by comparing its solutions to those of the DTA model. The evaluation results demonstrate that the prototype system is capable of running in real-time, and of producing good quality solutions using case-bases of reasonable size.

In a study by Capus and Tourigny (1998) an approach for road safety analysis with CBR is presented. The proposed system has two main functions. The first one allows the management of a case base as a conventional database. The second one, based on concepts inherent to CBR, permits the system to search its base for cases most similar to the situation encountered and, as needed, to re-use, adapt, and save them in the form of new cases. The authors found that the proposed approach has been proven technically feasible. However, due to limitation of the case base size and its hypothetical case, expert and end-user validation is necessary to improve the proposed methodology. This approach could be useful in supporting analysts to recall and re-use successful past solutions in connection with an analogy-based problem solving method.

Comparatively, the CBR approaches are employed mostly as planning tools, versus the model-based approaches that are more appealing to traffic flow researchers. Hybrid approaches have been found in only a limited number of research studies. One example

is cited from Sadek (2001), where a hybrid approach of simulated annealing and CBR was proposed and discussed in terms of rationale and design issues. To the author's knowledge no study has attempted to engage CBR in traffic prediction systems explicitly at the operational level. Therefore, this thesis presents a practical and operational methodology that explains how CBR, as a memory-based approach, can be integrated with model-based approaches, to develop a more robust traffic prediction system.

In summary, the existing research in the area of traffic prediction mostly uses model-based prediction approaches. The advantages of memory based methods are not fully exploited yet, nor the potential of merging the two approaches in a hybrid one. It is believed that the integration of the approaches should lead to a more reliable traffic systems prediction, which could be useful in many ATIS applications, such as travel time prediction.

## 4 BACKGROUND

In the past few years the computational intelligence and soft computing field has witnessed an intensive research interest towards integrating different computing paradigms such as fuzzy set theory, genetic algorithms, and neural networks to generate more efficient hybrid systems. The major characteristic of the soft computing area is the synergistic, and not competitive, way the individual tools act to enhance each others the application domain. The purpose is to provide flexible information processing systems that can exploit the tolerance for imprecision, uncertainty, approximate reasoning and partial information to achieve tractability, robustness, low solution cost and close resemblance with human like decision making (Pal et al. 2001).

There are two major techniques frequently applied in the computational area; the more traditional model-based approach, i.e. artificial neural networks (ANNs), genetic algorithms, rough set theory etc. and the less model aware or memory-based approaches, like fuzzy based computing or CBR. The ANNs are concerned with processing the information by a learning process and by adaptively responding to inputs in accordance with a learning rule. The usual application of ANNs is in the area of learning and generalization of knowledge and patterns. They are appropriate in explicit and well defined model-based problems. They are not suitable for expert reasoning and they have poor explanation capabilities. In contrast to the traditional knowledge-based system, a CBR system may be defined as a model of reasoning that incorporates problem solving, understanding and learning, and integrates all of them with a memory process (see Kolodner, 1993).

## 4.1 ARTIFICIAL NEURAL NETWORKS

One of the paradigms employed most often in the computational intelligence area is the artificial neural networks. There are several definitions for ANNs and the following emphasizes the key features of such models. An ANN can be defined as a distributed, adaptive, generally nonlinear learning machine built from interconnecting different processing elements (PEs) (Principe et al., 2000). The functionality of ANNs is based on the interconnectivity between the PEs. Each PE receives connections from other PEs and/or itself. The connectivity defines the topology of the ANN and it plays a role at least as important as the PEs in the ANNs' functionality. The signals flowing on the connections are controlled by adjustable parameters called weights,  $w_{ij}$ . A typical PE structure is detailed in FIGURE 1 as a non-linear (static) function applied the sum of all the PE's inputs. Due to the fact that ANNs' knowledge is stored in a distributed fashion through the connection weights between PEs and also the fact that the knowledge is acquired through a learning process that involves modification of the connection strengths between PEs, the ANNs tend to resemble in functionality the human brain.

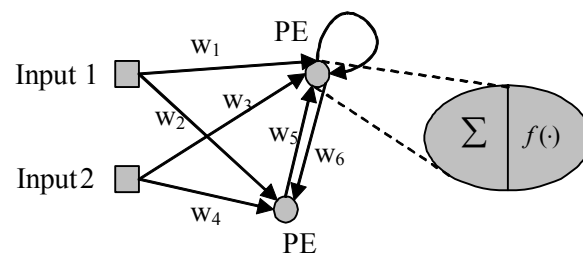


FIGURE 1: EXAMPLE OF ARTIFICIAL NEURAL NETWORK

There are many classifications of ANNs' architectures, for example one may distinguish them by the appropriateness of using different topologies to solve different problems. Hence, specialized architectures for applications like system identification,

function approximation, non-linear prediction, control, pattern recognition, clustering, feature extraction etc, have been developed. Another classification of neural networks divides them into two categories: static ANNs and dynamic ANNs. The former represent good function approximators that have the capability of building long-term memory into their synaptic weights during training. On the other hand, the dynamic networks have a built-in mechanism to produce an output based on more than one time instant in the past, establishing what is commonly referred to as short-term memory.

For most ANNs' architectures there are two stages in preparing them to be implemented in different applications. A training stage, in which the network learns from an existing dataset, and a testing stage that uses a different dataset to check the effectiveness of the learning phase. Because the knowledge incorporated in the ANNs is extracted from a given dataset during the training stage the ANNs are model-based and data driven systems. Usually the learning phase uses a back-propagation algorithm to adjust the connections' weights, based on the known data of input-output pairs. This means, after each presentation of all the data samples, which are the input-output pairs, the weights are adjusted such that the overall error output of the network is minimized. This complete cycle of running the data through the network and the weights adjustment process is called an *epoch*. A training stage of an ANN is completed based on different criteria. For example, the learning stops after a certain number of epochs or if the error reaches a certain limit or if the network's performance does not improve after a consecutive number of epochs. Another method to control the efficiency of the training stage is to test in parallel the error of the network performance on a cross-validation (CV) dataset, usually smaller than the learning dataset. The role of CV is to test for the network's generalization capabilities during the training process. If the network is over-

trained a sudden degradation of the network based on the CV data will trigger the training process to stop.

The training process is to achieve a good network performance when employed in on-line implementations. Recall that ANNs are data-driven systems and if the training process is not done properly the network may suffer of insufficient representation of the data or from overtraining. Insufficient data representation means that the dataset does not cover the complete solution space of the problem, such that when the network is tested with ‘unseen’ data it may not be able to perform satisfactorily. Conversely, overtraining occurs when the data is presented to the network in the learning stage for too many epochs. One solution to have a better control of the training issue is the use of a CV dataset. The cross-validation approach doesn’t allow the network to ‘memorize’, because this would impair dramatically the network performance on data with different characteristics from the training dataset.

One may easily notice by observing the available data of traffic conditions in any traffic monitoring system that there is much randomness and dynamics in the traffic conditions during the daytime and across weekdays. On the other hand, one may also assume the existence of a periodic component in the real traffic data. This may be due to the high likelihood of observing the onset of recurrent traffic conditions around the same time for the peak periods during the day. Based on these assumptions, the hypothesis that predicting traffic conditions on freeways, for example, becomes a non-linear function approximation problem seems reasonable.

In this respect, predicting traffic conditions in terms of speed and travel time estimates may take the form of a non-linear function over time. Moreover, it is possible that a periodic component may be encountered in this arbitrary function. This component may

describe the recurrent traffic conditions, like regularly congested conditions encountered at morning and afternoon peak periods, for example. There are several ANNs topologies used in function approximation applications. In this research, two categories of ANNs are employed, in order to handle the traffic prediction problem. The reason for this approach is derived by the assumptions that general static function approximation ANNs will be able to capture certain traffic conditions features like recurrent traffic conditions, while dynamic ANNs will be able to better cope with non-recurrent conditions, because their output depends on more than the current time instant. The following section presents details about four static neural networks and three dynamic neural networks, which have been used in this study.

#### **4.1.1 MULTI-LAYER PERCEPTRON (MLP)**

The MLP is a static ANN that has been extensively used in many transportation applications for its simplicity and ability to perform nonlinear pattern classification and function approximation. It is, therefore, considered the most widely implemented network topology by many researchers (Duda et al., 2001; Ham and Kostanic, 2001). Its mapping capabilities are believed to approximate any arbitrary function. An MLP consists of three types of layers: input, hidden, and output. It is normally trained with the backpropagation algorithm, which is based on minimizing the sum of squared errors between the desired and actual outputs. In FIGURE 2 an example of MLP topology is depicted. The example presented represents the topology used in this research study. It has two hidden layers. The number of neurons in the first hidden layer is double the number of neurons in the second one, as a general practice in the NNs topology.

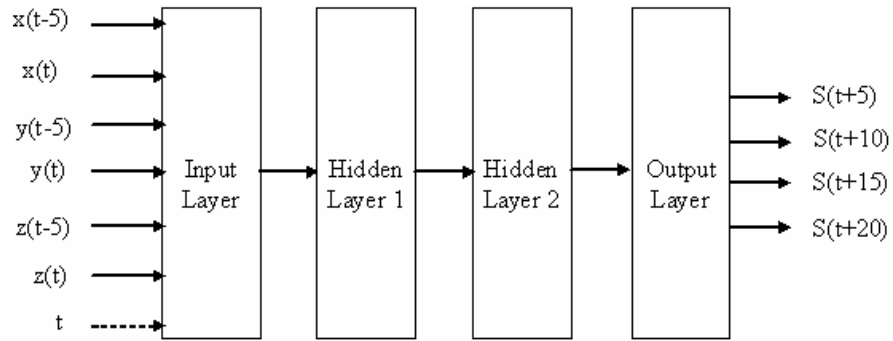


FIGURE 2: EXAMPLE OF MLP NETWORK TOPOLOGY

#### 4.1.2 MODULAR NETWORK

Modular networks are a special class of multiple parallel feed-forward MLPs. The input is processed with several MLPs and then the results are recombined. This NN type offers specialization of function in each sub-module and does not require full interconnectivity between the MLP's layers. Modular networks are often faster to train due to the smaller number of weights for the same size network. The topology used specifically for this application is composed of two primary components: local expert networks and a gating network (Jang et al., 1997, Principe et al., 2000). The basic idea is linked to the concept of "divide-and-conquer", where a complex system is better attacked when divided into smaller problems, whose solutions lead to the solution of the entire system.

Using a modular network, a given task will be split up among some local expert networks, thus reducing the load on each in comparison with one single network that must learn to generalize from the entire input space. A gating network eventually combines the output from the local experts to produce an overall output. FIGURE 3 shows the topology of the modular network.

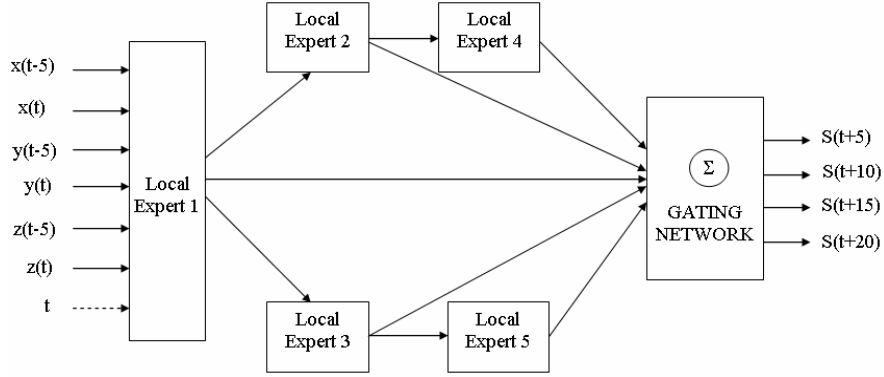


FIGURE 3: EXAMPLE OF THE MODULAR NETWORK TOPOLOGY

### 4.1.3 HYBRID PRINCIPAL COMPONENT ANALYSIS (PCA) NETWORK

PCA is a technique that finds an orthogonal set of directions in the input space and provides a way to find the projections into these directions in an ordered fashion. The orthogonal directions are called eigenvectors of the correlation matrix of the input vector and the projections of the corresponding eigenvalues. PCA has the ability to reduce the dimensionality of the input vectors, and therefore, can be used for data compression. When used in conjunction with MLP, the PCA can reduce the number of inputs to the MLP and improve its performance. In this research study a hybrid PCA/MLP network is used, combining both unsupervised and supervised learning paradigms in one topology. The PCA projects the input vector onto a smaller dimensional space, and thus, compressing the input for the MLP network.

It should be emphasized that PCA is a well known statistical procedure that is used in feature extraction from high-dimensional space (see Duda et al., 2001; Ham and Kostanic, 2001; Jang et al., 1997). The topology of the hybrid PCA network employed in this research study is illustrated in FIGURE 4.

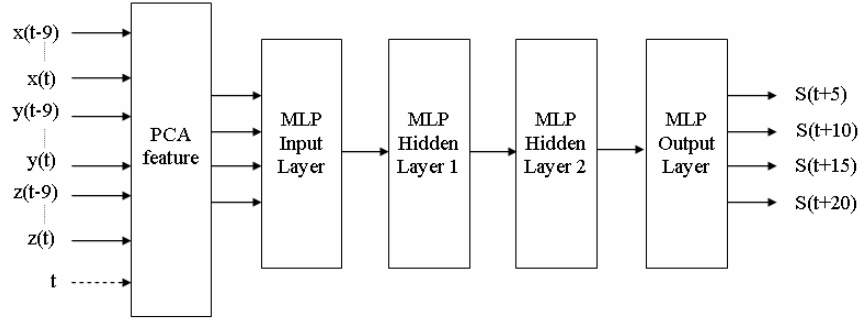


FIGURE 4: EXAMPLE OF THE PCA NETWORK TOPOLOGY

#### 4.1.4 CO-ACTIVE NEURO-FUZZY INFERENCE SYSTEM (CANFIS)

CANFIS belongs to a more general class of adaptive neuro-fuzzy inference systems (ANFIS) – see Jang et al. (1997). In the context of this research study, CANFIS is used as a universal approximator of any nonlinear function. The characteristics of CANFIS are emphasized by the advantages of integrating neural networks with fuzzy inference systems (FIS) in the same topology. The powerful capability of CANFIS stems from pattern-dependent weights between the consequent layer and the fuzzy association layer. Like the radial-basis function network (RBFN), CANFIS is locally tuned. The architecture of CANFIS is illustrated in FIGURE 4.

The fundamental component for CANFIS is a fuzzy neuron that applies membership functions (MF) to the inputs. Two membership functions are commonly used: general bell and Gaussian (see Lefebvre, 2001). The network also contains a normalization axon to expand the output into the range of 0 to 1. The second major component in the type of CANFIS used in this study is a modular network that applies functional rules to the inputs. The number of modular networks matches the number of network outputs and the number of processing elements in each network corresponds to the number of MFs. CANFIS also has a combiner axon that applies the MF outputs to

the modular network outputs. Finally, the combined outputs are channeled through a final output layer and the error is back-propagated to both the MFs and the modular networks.

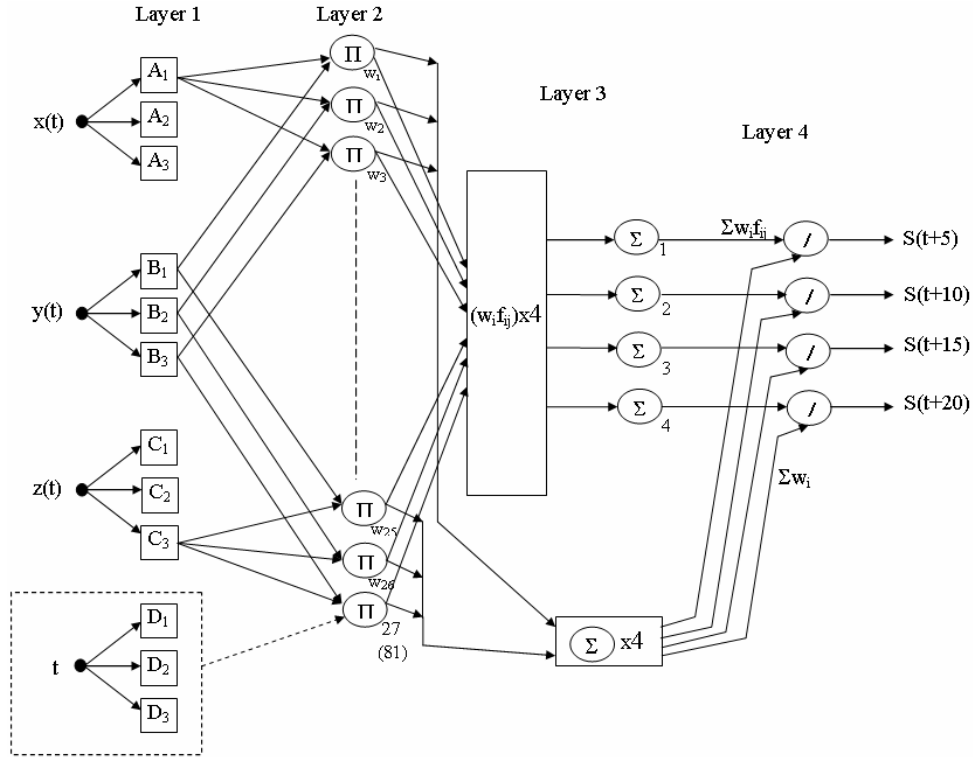


FIGURE 5: EXAMPLE OF CANFIS NETWORK TOPOLOGY

The CANFIS architecture used in this study is composed of four layers as shown in FIGURE 5. The function of each layer is described as follows. Each node in layer 1 is the membership grade of a fuzzy set (A, B, C, or D) and specifies the degree to which the given input belongs to one of the fuzzy sets. The fuzzy sets are defined by three membership functions. Layer 2 receives input in the form of the product of all output pairs from the first layer. The third layer has two components. The upper component applies the membership functions to each of the inputs, while the lower component is a representation of the modular network that computes, for each output, the sum of all the firing strengths. The fourth layer calculates the weight normalization of the output of the

two components from the third layer and produces the final predictions of speed at different prediction horizons

The rationale for which the selected four static neural networks detailed in this section were used in a fusing approach was based on the assumption that no particular NN is expected to outperform all others under all traffic conditions and all prediction horizons. While the static ANNs are recognized as good function approximators, another family of ANNs was identified as having good prediction capabilities. This category includes the dynamic neural networks, which are distinguished from previous ones as having the capability of using internal mechanisms to build short-term memory feature. This feature is possible because of the fact that the one-time instant output of the network integrates information from more than one-time instant of the input set presented to the network. In the following section details about three dynamic neural networks used in this study are provided.

#### **4.1.5 JORDAN/ELMAN NETWORK**

The Jordan/Elman network is also referred to as the Simple Recurrent Network (SRN) (Ham and Kostanic, 2001). It is a single hidden-layer feedforward network with feedback connections from the outputs of the hidden-layer neuron to the input of the hidden layer (Principe et. al, 2000). It was originally developed to learn temporal sequences or time-varying patterns. As shown in FIGURE 6, the network contains context units located in the upper portion and used to replicate the hidden-layer output signals at the previous time step.

The context units are introduced to resolve conflicts arising from patterns that are similar, yet result in similar outputs. The feedback provides a mechanism to discriminate

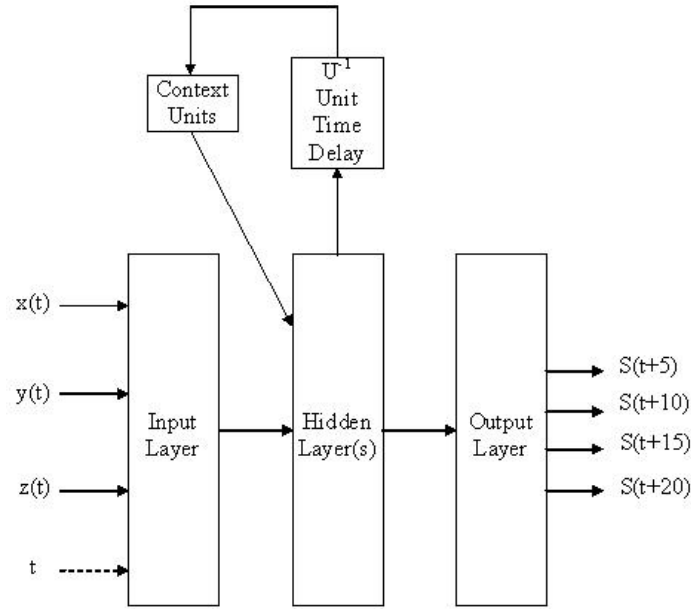


FIGURE 6: EXAMPLE OF JORDAN/ELMAN NETWORK TOPOLOGY

between identical patterns occurring at different times. The context unit is also referred to as a low-pass filter that creates a weighted average output of some of the more recent past inputs. Therefore, the context units are also called “memory units” since they tend to remember information from the past events. The training phase of this network is achieved by adapting all the weights using standard backpropagation procedures. More details on this topology can be found in (Ham and Kostanic, 2001; Lefebvre, 2001).

#### 4.1.6 PARTIALLY RECURRENT NETWORK (PRN)

PRN is considered a simplified version of the Jordan/Elman network without hidden neurons. It is composed of an input layer of source and feedback nodes, and an output layer, which is composed of two types of computation nodes: output neurons and context neurons. The output neurons produce the overall output, while the context neurons provide feedback to the input layer after a time delay. The topological structure

of the network is illustrated in FIGURE 7. More details can be found in (Haykin, 1998; Lefebvre, 2001).

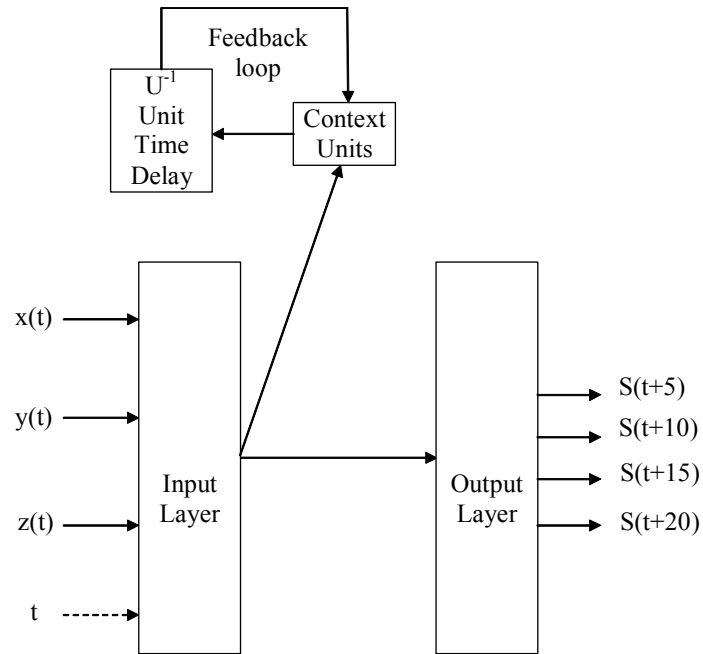


FIGURE 7: EXAMPLE OF PRN NETWORK TOPOLOGY

#### 4.1.7 TIME LAGGED FEED-FORWARD NETWORK (TLFN)

In dynamic neural networks time is explicitly included in mapping input-output relationships. As a special type, TLFN extends nonlinear mapping capabilities with time representation by integrating linear filter structures in a feedforward network. The type of topology used in this study is also called focused TLFN and has memory only at the input layer. The TLFN is composed of feedforward arrangement of memory and nonlinear processing elements. It has some of the advantages of feedforward networks such as stability, and can also capture information in input time signals. FIGURE 8 shows a simplified topological structure of the focused TLFN. The figure shows that memory PE (processing elements) are attached in the input layer only. The input-output

mapping is performed in two stages: a linear time-representation stage at the memory PE layer and a nonlinear static stage between the representation layer and the output layer. Further details underlying the mathematical operations of TLFN can be found in (Ham and Kostanic, 2001; Principe et al., 2000; Lefebvre, 2001).

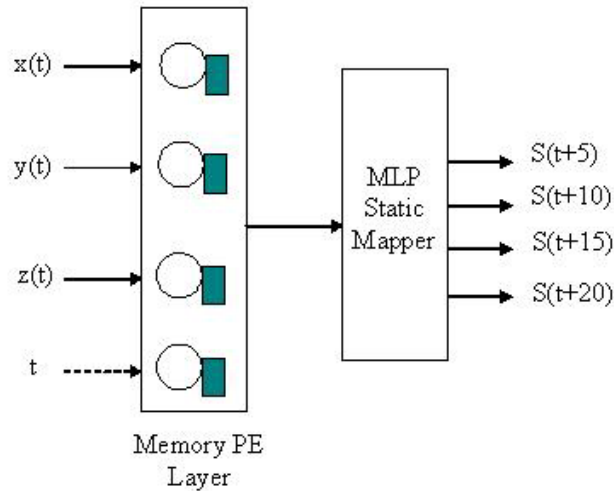


FIGURE 8: EXAMPLE OF TLFN NETWORK TOPOLOGY

Despite their good performance in a variety of applications, ANNs often have been described as a black box. Moreover, the performance of NN is expected to vary with the characteristic of the input data presented. In other words, the less similarity of the new data with the training dataset the more degraded performance we observe. Although a self-learning mechanism is believed to be built during the training stage, the generalization feature of NN cannot be exhaustive.

One method to compensate for the changes in performance of NN is to identify the conditions for which one network or another is performing the best and ultimately to integrate those networks to optimally predict for as many conditions as possible. Another method to optimize the NNs performance, that doesn't necessarily exclude the former, is

to synergistically put them to work with a second type of predictor mechanism. In this research study a memory-based environment such as Case-Based Reasoning (CBR) was investigated. The main reason for this choice is the assumption that a model-based predictor would be more capable of capturing knowledge related to non-recurrent traffic conditions that is believed to be related to most recent past information. On the other hand, the memory-based approach is believed to be more efficient in predicting recurrent traffic conditions due to its memory-like structure. These two assumptions support the decision of integrating the two predictors, ANN and CBR. The combined hybrid system is believed to perform better than employing each technique individually.

## **4.2 CASE-BASED REASONING (CBR)**

Memory-based learning approaches have been cited relatively recently in data mining applications. The concept of case-based reasoning used in several industrial and research applications is a successful example of memory-based approaches. CBR systems are simply defined as a collection of cases that are representations of typical situations with past experience and possible solutions. They are powerful in retaining a memory of the previous examples or cases and their solutions. As a model of reasoning, a CBR system searches its case base for a current case using a set of pre-defined case features. The search can result in finding an identical case or one or more cases that closely match the current query case. If a close match is found, the system will employ an adaptation phase to retrieve the best solution possible. In its simplest form, a CBR system is composed of a case retriever and a case reasoner. While the former is used to find the most appropriate case in the case base, the latter will seek solutions to the retrieved cases.

A simplified illustration of a typical CBR system is shown in FIGURE 9. The figure shows the CBR as a cyclical process involving four actions: retrieve, reuse, revise and retain. When a new problem is presented to a CBR system, a case retriever is activated to find the most similar cases. Case retrieval methods are categorized into two techniques:  $k$ -NN (nearest neighbor) retrieval and inductive retrieval. Inductive retrieval is based on extracting rules or decision trees from past data. This approach is less frequently used in applications, but is often considered more efficient in terms of retrieval speed from large case bases. For relatively small case bases, the  $k$ -NN method is more efficient, and therefore, is proposed for use in this CBR system. The  $k$ -NN method is used to find the  $k$  most similar cases to a target case by matching a set of features  $\{x_1, x_2, \dots, x_n\}$ , where  $n$  is the number of features.

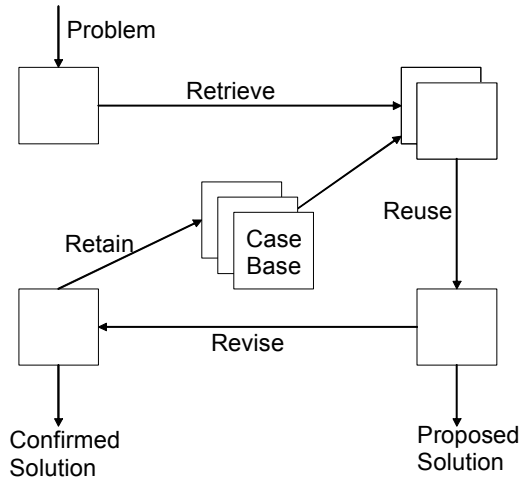


FIGURE 9: ILLUSTRATION OF A TYPICAL CBR SYSTEM

For a given query  $q$  and a case base, the  $k$ -NN method retrieves a set of  $k$  most similar cases using a similarity measure defined by:

$$\text{Similarity}(T, S) = \sum_{i=1}^n w_i f(T_i, S_i) \quad (1)$$

Where

$T$  is the target case

$S$  is the source case

$n$  is the number of attributes in each case

$f$  is a similarity function for attribute  $i$  in cases  $T$  and  $S$

$w_i$  is the weight value assigned to attribute  $i$

When found, one or more cases will be ‘retrieved’. Based on the matching cases, a solution is suggested to be ‘reused’ and tested. If the case match is not close enough, the solution should be ‘revised’ and ‘retained’ as a new case. There are several adaptation procedures to ‘revise’ a new case, but mainly they fall under the category of structural adaptation and derivational adaptation. If the former adaptation principle means to adjust the retrieved solutions, the later means readjusting of the retrieving algorithms and new solutions are generated. Consequently, once a new solution is generated, the ‘reuse’ module simply adds the newly acquired knowledge, the new case and its derived solution, to its case base. More details on CBR systems can be found in (Kolodner, 1993; Watson, 1997).

## 5 METHODOLOGY

The proposed methodology integrated model-based and memory-based approaches into one hybrid system to improve the traffic prediction performance. The model-based approach is derived from a composite family of neural network architectures that belong to function-approximation and temporal-processing classes for their powerful and universal predictive capabilities. The memory-based approach is developed based upon CBR systems. The hybrid approach is motivated by the remarkable analogy with human information processing systems. Model-based approaches are capable of making predictions based on the intrinsic information learned from the training dataset, which is known to be of limited size. In the case of artificial neural networks, the knowledge learned from the training dataset is encoded in the synaptic weights to a certain degree so that the trained network can generalize when presented with cases that are identical or similar to the ones encountered during the training phase. This is believed to be analogous to the human thinking process, except that the latter is often supplemented with memory aid that is explicitly contained in the training dataset before abstraction, as well as the large information source from which the training dataset was sampled.

With the recent emphasis on data mining mechanisms, memory is often realized in the form of large databases that are used to store, query, and retrieve information almost instantaneously to support real-time applications. Currently, databases on traffic conditions are the main information resource in traffic management systems. Unfortunately, the amount of information inherent in such large databases is yet to be fully extracted, even with today's advanced data mining functions. To further illustrate the relevance of memory-based reasoning in the development of traffic prediction systems, we

cite a simplified example from real life travel experiences. In the absence of predictive traffic information systems, travelers, especially commuters, exploit their own travel experiences in the past to support their decision-making processes. With the exceptional memory retention capabilities of the human brain, commuters retrieve their past travel experiences from their memory to predict the traffic conditions that are most likely to prevail at certain locations and time periods during the day. It is not very unusual for some people to further complicate this process by observing day-to-day variations, or perhaps, seasonal and year-to-year variations.

Although predictions that are solely based on historical information are likely to fail under non-recurrent conditions, they still are very useful and intuitive to most people. While model-based approaches often rely on relating forecasted traffic conditions to the most recent past information in the time-space domain, they reportedly fail to incorporate the memory aid that is offered by past travel experiences. Therefore, the focus of this research is on combining both approaches to strengthen the predictive capabilities of on-line traffic prediction systems. The proposed methodology demonstrates a practical framework for a traffic prediction system that is based on a hybrid model-based and memory-based approach.

## **5.1 STUDY AREA AND DATA COLLECTION**

Currently, traffic surveillance systems are key components of the implementation schemes of intelligent transportation systems in urban areas. Equipped with advanced data collection technologies, traffic surveillance systems are continuously streaming real-time information on traffic conditions to traffic management centers to support their operation and management functions. The perpetual data acquisition process has created a

need to establish data warehousing and mining systems for many urban freeway systems. Examples can be found in California, Washington, Florida, Texas, and other states. The massive amount of data, sometimes in the order of a few gigabytes per day, is currently compiled into data archival systems and used to support performance monitoring systems such as Freeway Performance Measurement Project (PEMS) in California (Choe et al. 2002). The data archival system is considered the foundation for both model-based and memory-based learning strategies, as explained next.

This study was conducted using data from a freeway segment of I-4 in Orlando, Florida, as shown in FIGURE 10, the circled area. The traffic surveillance system compiles 30-second data of speed, volume, and lane occupancy from a 40-mile six-lane corridor instrumented with 70 inductive dual loop detector stations spaced at nearly 0.5 miles in both directions. The real time and archived data is accessible via Internet at the I4 Real-Time Traffic Information web-site (<http://www.trafficinfo.org>). The loop detector data is collected in real time via a T1 link between the I-4 Regional Traffic Management Center (RTMC) in Orlando and the intelligent transportation system lab at the University of Central Florida. Speed, volume counts, and lane occupancies are downloaded and compiled into an SQL server that supports multiple publicly accessible web applications such as real time and short-term travel time predictions between user-selected on- and off-ramps. The web-based short-term traffic prediction system was implemented using a nonlinear time series model that was tested extensively in a previous study (Ishak and Al-Deek 2002). For this study speed data was collected from three adjacent stations over a 1-mile section of I-4. The development of the model-based component is based on a total of approximately 40 days that were randomly selected in the years 2001 and 2002.



## 5.2 TRAINING

Training each of the seven network topologies was conducted using NeuroSolutions software package (Lefebvre, Curt - 2001). To achieve optimal performance, different settings were attempted by varying the number and type of inputs to each network. The input to each network was classified into two components: short-term memory (STM) and long-term memory (LTM). The STM component was represented by speed data observed in the past 10 minutes at the three adjacent stations. Input patterns were constructed over time and space to capture temporal and spatial variations of traffic conditions. However, to optimize the performance of the networks, input patterns were constructed from four spatial settings: current station only (y), current and upstream (y, z), current and downstream (y, x), or the three stations combined (x, y, z).

The LTM component was introduced in addition to the STM component to test the network's ability to learn from similar historical traffic conditions observed at the same time on other days. The LTM component was represented by a time index attached to each constructed speed pattern. The time index was referenced to the beginning of the peak period at 6:00 AM and expressed in increments of 1 minute. The essence of using LTM component is to improve the performance at relatively longer prediction horizons by making the network time-cognizant during prediction. Each network was trained to predict the average 5-minute speeds at 5, 10, 15, and 20-minute horizons at current station (y). During the training phase the performance of the networks was monitored via the validation set to avoid overtraining. Training is terminated when the mean square error (MSE) for the cross-validation set does not decrease for 50 consecutive training cycles, a common procedure to prevent overtraining. For the static networks (see FIGURE 2 -

FIGURE 5), except for the hybrid PCA network, the input patterns representing STM were in the form:

Type 1 input:  $\{x(t), x(t-5), y(t), y(t-5)\}$

Type 2 input:  $\{y(t), y(t-5), z(t), z(t-5)\}$

Type 3 input:  $\{x(t), x(t-5), y(t), y(t-5), z(t), z(t-5)\}$

Where

$t$  = refers to the current time

$x(t)$  = average 5-minute speed at downstream station at time  $t$

$x(t-5)$  = average 5-minute speed at downstream station at time  $t-5$  minutes

$y(t)$  = average 5-minute speed at current station at time  $t$

$y(t-5)$  = average 5-minute speed at current station at time  $t-5$  minutes

$z(t)$  = average 5-minute speed at upstream station at time  $t$

$z(t-5)$  = average 5-minute speed at upstream station at time  $t-5$  minutes

The hybrid PCA network has the ability to reduce the dimensionality of the input space by locating the principal components. Therefore, the input patterns were presented in high dimensional vectors of 10 observations taken one minute apart from time  $t$ . In other words, the input patterns were in the form:

Type 1 input:  $\{x(t), x(t-1), \dots, x(t-9), y(t), y(t-1), \dots, y(t-9)\}$

Type 2 input:  $\{y(t), y(t-1), \dots, y(t-9), z(t), z(t-1), \dots, z(t-9)\}$

Type 3 input:  $\{x(t), x(t-1), \dots, x(t-9), y(t), y(t-1), \dots, y(t-9), z(t), z(t-1), \dots, z(t-9)\}$

Due to their ability to build internal short-term memory the dynamic networks didn't need inputs from recent past information. The output vector for each input pattern was constructed in the form  $\{S(t+5), S(t+10), S(t+15), S(t+20)\}$ , where  $S(t+5)$ ,  $S(t+10)$ ,

$S(t+15)$ ,  $S(t+20)$  are the average 5-minute speeds taken at 5, 10, 15, and 20 minute predictions at current station (y), respectively (see FIGURE 6 -FIGURE 8).

### 5.3 TESTING

To test the performance of each network after training, two sets of 10 peak periods collected from 20 different days were presented to the network. For each input pattern in the testing set, multiple predictions were made at 5, 10, 15, and 20 minute horizons. Each predicted value was compared against the actual observed value to calculate two measures of performance: average absolute relative error (*AARE*) and root mean square error (*RMSE*) in speed. Each measure is defined as follows:

$$AARE = \frac{\sum_{i=1}^N \left| \frac{P(i) - O(i)}{O(i)} \right|}{N} \quad (2)$$

And

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P(i) - O(i))^2}{N}} \quad (3)$$

Where

$P_i$  = Predicted speed (mph) for observation  $i$

$O_i$  = Actual observed speed (mph) for observation  $i$

$N$  = Number of observations

The two measures were used to compare the performance of the each of the two NN families under the following settings: desired prediction horizon (5, 10, 15, and 20 minutes), input type (xy, yz, and xyz), and inclusion of the LTM component (Yes/No). In addition to the previous controlled variables, the performance was also checked against various combinations of traffic conditions at each of the three stations. At each station,

traffic conditions were broken down into four levels of congestion: level 1 (speed <20 mph), level 2 (20-40 mph), level 3 (40-60 mph), and level 4 (>60 mph). For all possible combinations of the four levels of congestion at each station, the network performance was evaluated to determine the optimal settings.

This procedure was primarily used to address two questions. First, which network and what settings are optimal for predictions at each of the four horizons tested? Second, what traffic conditions are associated with the largest errors or worst prediction performance so that we can identify the level of confidence in our predictions? Both questions are critical to the successful online implementation of a traffic prediction system. The answer to the first question will identify which network and what settings are best for each traffic condition in terms of minimization of the AARE and RMSE. This allows for the hybrid prediction system that is optimized to produce the best performance under different prediction horizons and traffic conditions. The answer to the second question distinctively identifies the traffic conditions during which prediction accuracy is unacceptable, and therefore, should not be disseminated to the public or at best disseminated with an associate measure of uncertainty. In addition, the two performance measures are used to evaluate the performance of each approach separately, as well as to help integrate the two in a hybrid prediction system.

#### **5.4 HYBRID TRAFFIC PREDICTION SYSTEM**

The developed hybrid model-based and memory-based technique is motivated by the demonstrated synergies of combining both approaches to overcome the shortcomings of applying each individually. As mentioned earlier, the model-based learning is accomplished with artificial neural networks for their powerful and universal predictive

capabilities (Shin and Sang, 2000). Even when knowledge is argued to be stored in the connection weights, it is still difficult to convert to comprehensible symbolic rules (Pal et al. 2001). Such challenge was met recently with a few algorithms that were specifically developed to overcome the lack of knowledge comprehensibility by integrating neural network models with memory-based or instance-based learning approaches (Jang et. al, 1997). The integrated features of both approaches are expected to lead to a more robust prediction system that overcomes the shortcomings of each.

The general outline of the developed hybrid traffic-prediction system framework is illustrated in FIGURE 11. This schematic shows that the model-based and memory-based approaches are executed in parallel. The proposed system features two CBR systems that are constructed from each approach separately. The framework is presented in its full scale on-line implementation mode, although some of its components must clearly be developed off-line at an initialization stage. The hybrid system features mainly two CBR components and a Prediction Query Manager to integrate the outcomes of the two prediction systems. The first CBR system relies upon the model-based approach, which is built upon the static and dynamic ANN. The second CBR is integrating a memory-based approach using the traffic conditions information from a case-based reasoning perspective. Ultimately an integration module, namely Prediction Query Manager (PQM), is invoked to generate an optimize prediction via the outcomes of the two model-based and memory-based approaches. Details on each of three components are provided in the following sections.

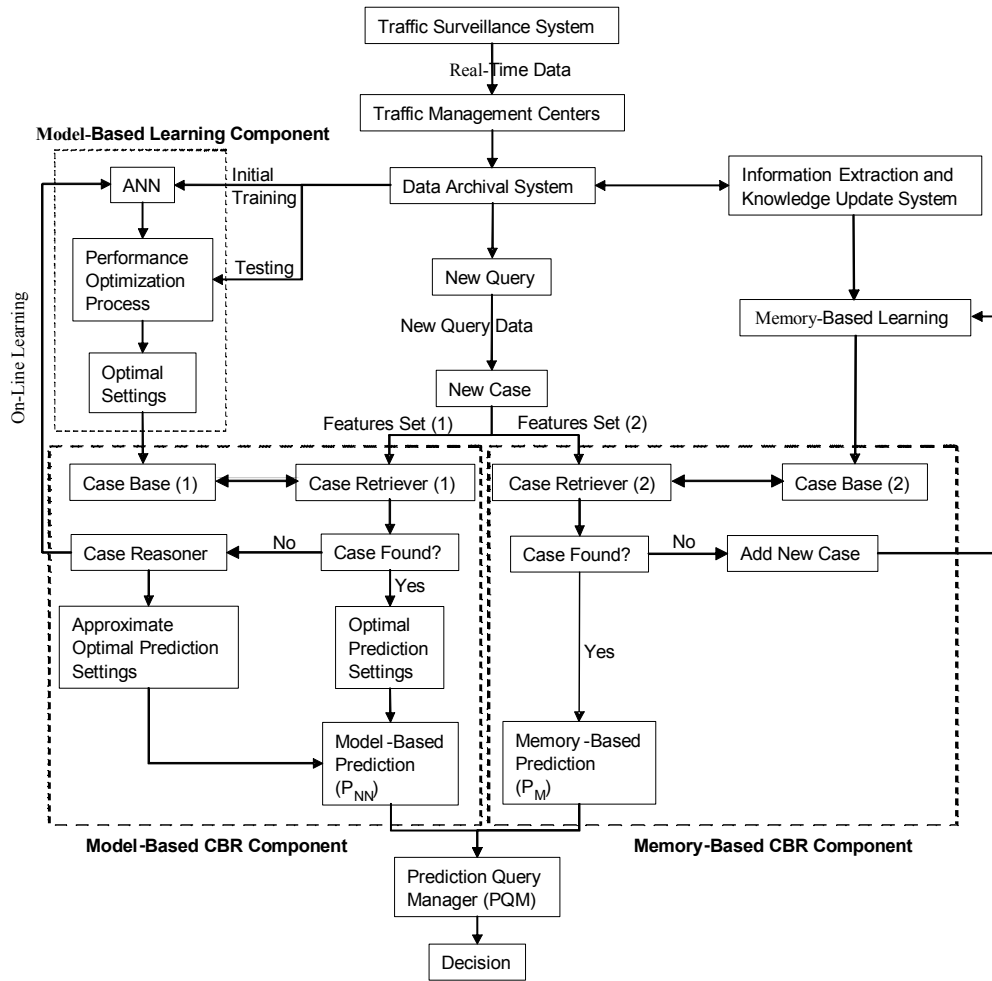


FIGURE 11: HYBRID TRAFFIC PREDICTION SYSTEM OUTLINE

## 5.5 MODEL-BASED APPROACH

Recognizing the variability in performance levels of different neural network topologies, it is believed that if a first optimization stage would exploit multiple network topologies from two classes of architectures: (a) function approximation and (b) dynamic or temporal-processing (Duda et al. 2001, Ham and Kostanic 2001, Haykin 1994, Jang et al. 1997, and Principe et al. 2000) a certain performance improvement will be achieved. In this stage, the performance of the model-based component is optimized under different topological and traffic condition settings. To achieve this goal, data from stations adjacent

to the target station was used. The purpose of using traffic data from adjacent locations was to capture spatial variations in traffic conditions and the effect of shockwave propagation. The model-based approach integrated all NN topologies described in the previous chapter belonging to both static and dynamic NN families. This model-based phase required an initialization of the NNs with an initial training dataset that was extracted randomly from the data archival system. A 15-day training dataset from year 2001 was used to train the seven NNs. In addition, a cross-validation 5-day dataset, from the year 2001, was used to control the training process. The initial learning phase must be initiated off-line due to its time-consuming nature. The full development phase is typically comprised of three stages: training, cross-validation, and testing, so that a third dataset was collected distinctively from the previous data and used in the testing stage.

As stated in previous sections, this research study considered two optimization stages for the developed hybrid traffic prediction system. The optimization of the model-based prediction system relies on identifying the traffic conditions for which a network setup will perform best. The networks setup consists of determining what inputs are needed for a certain network to be employed in the prediction process. The possible inputs are represented as the speed data inputs from the target and the immediate adjacent locations plus another input representing the time stamp of each data sample, such that it is possible to distinguish between different types and number of inputs to be able to build a reduced CBR system. On the other hand, the network outputs, represented by the forecasted speeds at target station for each of the four prediction horizons, compared with the actual values observed after the prediction horizons pass, will produce average network performance errors. Base on minimizing these errors a best prediction network will be chosen as the solution for the model-base CBR approach.

## 5.6 MODEL-BASED CBR APPROACH

The first CBR system in the proposed framework is referred to as “model-based” (see FIGURE 11). The reference to the model-based prediction approach is made due to the fact that the case features are identified with the traffic conditions cases whose prediction performance was optimized in the model-based learning stage. The case base is best described as a collection of cases with different traffic conditions and their associated optimal settings in terms of the optimal topology, input size, and input type for each prediction horizon (PH). For each case, a measure of performance can also be indexed based on the results of the testing phase of the corresponding settings. This performance measure can be considered an information reliability measure of the subsequent information dissemination system.

In order to invoke the model-based CBR system, a set of features was defined for each case. In the developed implementation, the cases were described in terms of the prevailing traffic conditions observed at the target and adjacent upstream and downstream locations of sensors. The traffic conditions were split in four levels, free flow with speeds of 60 mph and up, low congested conditions with speeds between 60-40 mph, mild congested conditions with speeds between 40-20 mph, and heavy congested conditions with speeds of 20 mph or less. Each of the three locations was identified in terms of the four defined congestion levels. When a new case is generated from a new query, the set of features will be passed on to the case retriever, which in turn will search the case base for the most similar case. If a similar case is found, then the optimal prediction settings (topology and input types) are retrieved and applied to obtain the final model-based prediction ( $P_{NN}$ ).

In the event that the case retriever fails to find a closely similar case, then a case reasoner is invoked. The case reasoner will attempt to approximate the optimal prediction settings from cases that are partially similar; i.e. when most (not all) of the features are similar. Using the approximated settings, the model-based prediction can now be estimated. Should the case reasoner fail to relate the target case to any of the cases in the case base, then it will invoke the on-line learning feature of the model-based system by presenting the current case and its predicted response (after prediction time elapses) to the neural network system in order to evaluate the NNs performance for the new case. This will identify the optimal settings, in terms of neural network and its optimal inputs setup that minimize the network error, that correspond to the newly introduced case. The new case and its optimal settings can then be added to the case base for future retrievals.

## **5.7 MEMORY-BASED CBR APPROACH**

The second approach applied in the proposed framework is an explicit memory-based learning system (see FIGURE 11). This system is constructed directly from the data archival system using a set of information extraction and knowledge update procedures. This CBR system was particularly used as a memory representation of the traffic conditions observed in the past. The proposed cases for this CBR system must be capable of encoding the historical information on traffic conditions by location, time of day, day of week, month of year, etc. Such attributes can be effectively used as case features along with the actual observations of flow, speed, or occupancy as the consequent solution for each target case.

Referring to FIGURE 11, as a new case is generated from a new query, it is now routed to this memory-based CBR system with a set of features that identify the new case.

Similar to the previous model-based CBR system, a case retriever will search the case base of historical information to locate the most similar case(s). When at least one similar case is found, the memory-based prediction can now be estimated from the solution(s) of the similar case(s). The solution proposed by the system,  $P_M$ , will be based on the prevailing traffic conditions in the past, which were frequently observed at the same location, weekday, time of day, month of year, and possibly other significant features that could be relevant.

In the event that the target case was not found in the CBR system, the case could be considered new, and therefore, after its solution becomes available should be added to the case base for possible future retrievals. There are two possible actions that can be taken in such event. The most conservative approach is to refrain from making predictions due to the lack of sufficient knowledge. A less conservative approach is to relax one of the features that is thought of as least importance or introduce slight modifications to its attributes and then search for a similar case. An example would be to change the location to another adjacent location, or use another weekday that is known to have similar characteristics. This alternative obviously requires inductive retrieval mechanisms so that a decision tree can be constructed from the case base. For instance, the assumption that traffic conditions at a range of locations are consistently similar at certain times of the day or certain days of the week may support the decision to look for a solution in another location. Such cases can be clustered to reduce the size of the case base and improve the retrieval efficiency of the system.

In the developed implementation memory-based CBR approach several features were investigated. For example in addition to information about traffic conditions at each of the three locations, the time pattern was investigated in the form of probability to find a

certain case in the past within a certain time interval centered around the time of prediction. For example, if a case is described by the following features: the congestion level is free flow (level 4) at upstream, current and downstream locations, then the question becomes: what is the probability of occurrence of a certain case, or traffic conditions setup, at 7:25 am on a Monday, within a 60 minute time-frame window centered around 7:25? In other words, the system needs to estimate the probability of encountering the specified traffic conditions between 6:55 am and 7:55am. It is to the author's knowledge that this concept of probability associated (further referred to herein after as *probability of occurrence*) with the CBR solution has not been previously investigated. For this specific problem the associated probability of occurrence of a certain case is computed as the total number of similar cases within the time specified interval divided to the total number of cases within the specified time interval, see equation (4).

$$p^T(P_M) = \frac{\# \text{ of similar cases at time } T}{\text{total \# of cases at time } T} \quad (4)$$

Where,

$P_M$  - is the average value of the solutions of similar cases found in CBR case base;

$T$  - is the center of the time frame window where the cases were searched for.

In this specific traffic prediction application the time  $T$  is also a feature that helps in describing a case, for another CBR application a different variable may have a specific role and may be used with a similar scope as  $T$ . The associated probability for the solution of a new case presented to the CBR is useful to the extent that it helps in the integration of the two approaches in a final optimized predicted value.

## 5.8 INTEGRATION OF MODEL-BASED AND MEMORY-BASED PREDICTIONS

The prediction values yielded by model-based and memory-based systems are eventually fused using a Prediction Query Manager (PQM) mechanism, which is responsible for making the final prediction decision based on both model-based ( $P_{NN}$ ) and memory-based ( $P_M$ ) predictions. The decision mechanism implemented in this study is based on the associated performance measures of the optimal settings identified during the first optimization phase combined with the associated probability of occurrence defined for the memory-based CBR approach.

The reader is reminded that for each of the traffic conditions identified in the NNs' testing dataset an optimal setting was identified, based on minimizing the defined AARE and RMSE performance measures. On the other hand the CBR's probability of occurrence defined in the previous section is closely related to the performance of the CBR. This assumption is supported by the plot in FIGURE 12 - FIGURE 15, in which the performance envelope of the CBR for a 5-, 10-, 15-, and 20-minute prediction horizon is depicted. The data presented in come from analyzing the CBR prediction over the same testing dataset applied to the seven NNs previously developed. In this figures one may easily identify an envelope that borders to some extent the scatter plots. The envelope approximates an equation of the form  $y = x^{-\alpha}$ , but a precise equation to describe the data was not of relevance. The graphical approximation of the performance envelope is satisfactory for the data analysis performed in this study. The performance envelope is used to relate an expected maximum prediction error with an estimated probability of occurrence for each case presented to the memory-base component of the hybrid prediction system.

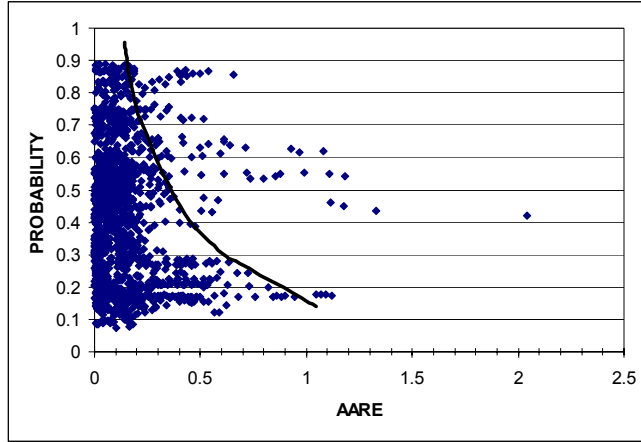


FIGURE 12: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 5-MINUTE PREDICTION HORIZON

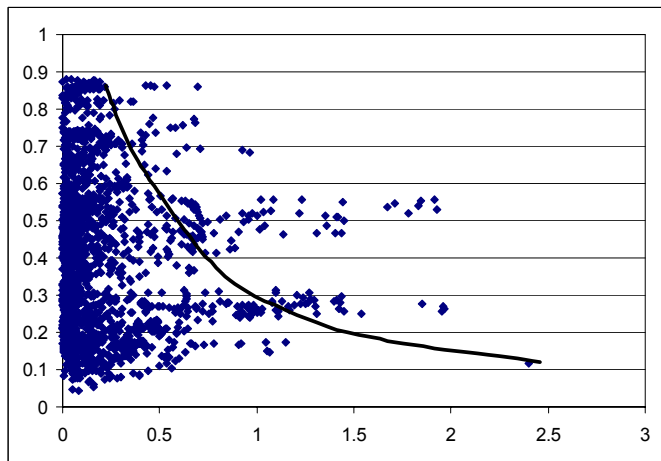


FIGURE 13: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 10-MINUTE PREDICTION HORIZON

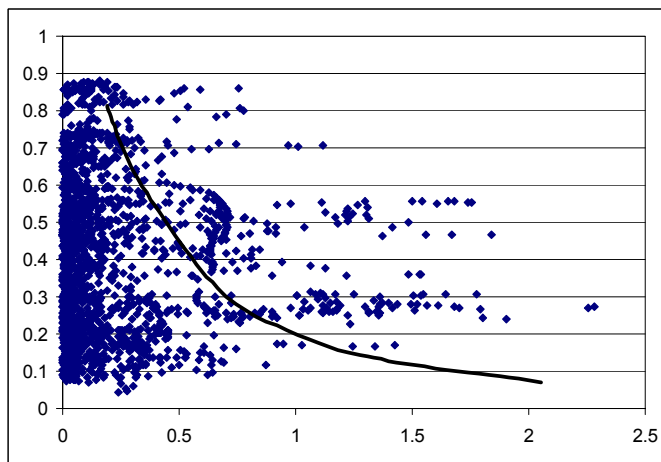


FIGURE 14: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 15-MINUTE PREDICTION HORIZON

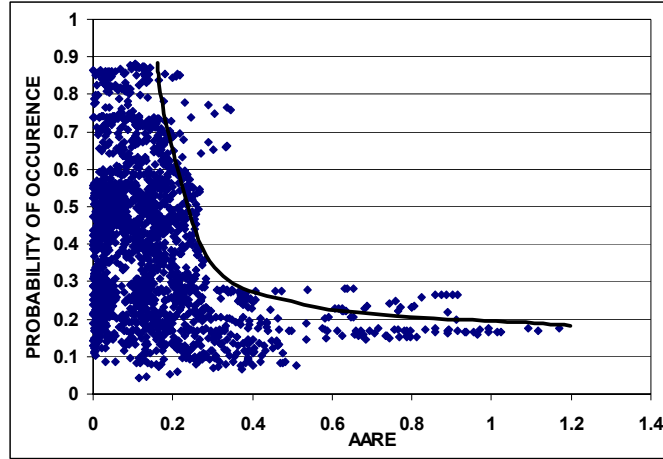


FIGURE 15: PERFORMANCE ENVELOPE OF MEMORY-BASED CBR APPROACH FOR 20-MINUTE PREDICTION HORIZON

To establish a final prediction the PQM processes the following inputs for each treated case:

- traffic conditions (X, Y, Z) information that identifies the optimal settings for the model-based approach;
- pairs of values  $P_{NN}$  - $ERR_{NN}$ , representing the predicted value of the model-based approach and its estimated performance measure value, derived from the optimal settings;
- pairs of values  $P_M$  - $ERR_M$ , representing the predicted value of the memory-base approach and its expected maximum prediction error based on the computed probability of occurrence – using data driven diagrams similar to the ones presented in FIGURE 12 - FIGURE 15. The decision mechanism integrated in the PQM is detailed in FIGURE 16.

The principle that governs the integration mechanism identifies the predictive method that has the smallest estimated error. For this purpose the following inputs are processed:

- the traffic conditions at current, upstream and downstream stations, X, Y, and Z in FIGURE 16; these traffic conditions are also used to identify a case in the model-based CBR component;

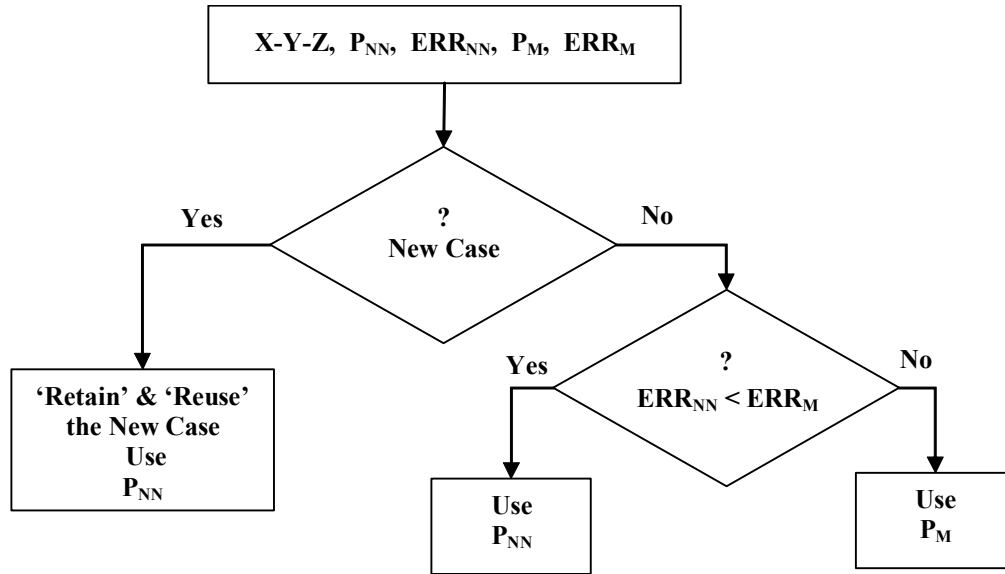


FIGURE 16: PREDICTION QUERY MANAGER DECISION MECHANISM

- the predicted speed based on the optimal NN settings,  $P_{NN}$  along with its estimated error from the testing stage,  $ERR_{NN}$ ;
- the predicted speed from the memory-based CBR approach,  $P_M$  along with its estimated error from the CBR calibration stage,  $ERR_M$ . The estimation of the  $ERR_M$  value should be derived from FIGURE 12-FIGURE 16. For a given probability of occurrence for the each a maximum error can be inferred and used in the PQM decision mechanism.

Should this approach fail and similar cases cannot be found in the case base the model-based approach is trusted to have a better prediction due to its generalization capabilities. In addition, as the predicted information is compared with the available field data the new identified cases are retained for future usage in CBR case base. Eventually the CBR case base will cover exhaustively all possible traffic condition combinations.

## 5.9 APPLYING THE PROPOSED METHODOLOGY

In order to evaluate the performance of the proposed hybrid approach for short-term traffic conditions prediction the following six steps were developed.

**Step 1:** The seven NNs were trained to predict speed at target location (Y). For each NN all possible setups were considered based on speed inputs from the three stations (X, Y and, Z) and the LTM input.

**Step 2:** Then, using a different dataset, each network configuration was tested in order to evaluate its performance in terms of RMSE and AARE.

**Step 3:** Subsequently in the testing dataset, based on the four traffic conditions levels at each of the three stations, the NNs predictions were evaluated for each traffic condition combination (see TABLE 1-TABLE 4). The network setting with the minimum error is selected as the one that should be used further in the prediction for a certain traffic conditions combination. For example, if one wants to estimate the speed in 10-minute prediction horizon, at free-flow prevailing traffic conditions at each of the three stations, and if AARE is considered as performance measure, than in TABLE 2d) case no. 25 identifies the optimal setting as Jordan-Elman NN with inputs from upstream (Z) and current station(Y), downstream (X) and LTM input.

**Step 4:** Next, the CBR prediction is evaluated. For each prediction horizon, the same testing set used for evaluating the NNs optimal settings in Step3 is used for CBR to generate the probability of occurrence plots in FIGURE 12 - FIGURE 15. The performance envelopes are approximated for each of the prediction horizons to be used in the hybrid next steps.

**Step 5:** Next, CBR and NNs performances are evaluated over a second testing dataset. Using the same optimization procedure, based on prevailing traffic conditions at

each of the three stations, optimal settings are derived for NNs. The CBR evaluation is performed by considering the prevailing traffic conditions at the current stations (Y), under a time window of 60 minute, at specific weekdays. The second testing set has ten days, two of each weekday.. The probability of occurrence is computed for each prediction made by CBR and the performance envelopes estimated in Step4 are used to approximate the expected error,  $ERR_M$ .

**Step 6:** In the last step, the prediction values and their associated errors are evaluated by the Prediction Query Manager module. The approach with the smallest estimated error is the chosen to be the winner. If the NN-based prediction fails to find an optimal setting for a new case, then all the networks are tested with the new case once the prediction horizon passes and the actual speeds become available. Next, the optimal settings are derived as in Step 3. Eventually the model-based CBR case base will accumulate enough cases to cover all the possible traffic conditions combinations.

## 6 INTERPRETATION OF THE RESULTS

In the following sections details about the data analysis and its interpretation will be presented. Based on the 4 levels of congestion defined previously the prediction performance of the neural networks was evaluated to determine the optimal settings of the model-based traffic prediction approach. The memory-based approach role improving more the performance of the hybrid traffic prediction system was tested against statistical significance with both Wilcoxon signed rank test and Friedman two-way analysis of variance non-parametric tests.

### 6.1 OPTIMAL SETTINGS OF THE MODEL-BASED APPROACH

The optimal settings of the NN approach were selected for each 5-, 10-, 15-, and 20-minute predictions using *AARE* and *RMSE* independently, with respect to the prevailing traffic conditions at the same station and upstream and downstream stations (see ...). Theoretically, with four levels of congestion at each station, the total number of combinations should be  $4 * 4 * 4 = 64$ . The optimal settings consist in identifying for each combination of traffic conditions the optimal neural network. In addition to the selection of the optimal neural network, also the optimal number of inputs and the type of inputs are identified. This means that the input from all the three stations (current, upstream or downstream) and the long-term memory input are not always needed for all traffic conditions combinations, but specific combinations of these inputs will produce an optimized prediction.

Despite the fact that both performance-measures (*AARE* and *RMSE*), are often used to quantify the performance of prediction models, consistency between the two measures is not, by their mathematical definition, guaranteed. Therefore, optimal settings were selected and presented separately. While some optimal settings were different by measure, others were consistent, indicating that both measures are in agreement. For each prediction horizon, a total of

25 cases were identified as a result of different combinations of traffic conditions at each of the three stations. The 25 cases, however, do not include all possible combinations of traffic conditions, but only the ones covered in the drawn testing data set. However, some combinations were not observed due either to their infrequent occurrence in general or to the limited size of the testing set. Therefore, the optimization was limited to the 25 cases for the considered testing dataset. Should a new case occur in the implementation phase of the hybrid prediction system the on-line learning mechanism should analyze it and integrate it as depicted in FIGURE 11. In other words, the prediction values of all networks under all input combinations will be compared with the actual value when this becomes available. Next, the model-based CBR component will ‘retain’ and ‘reuse’ the identified solution at future occurrences of the newly registered case.

The second optimization stage, that combines predictions of both ANN and CBR methods, was evaluated over a second testing dataset. Considering the same number of level of congestion, in this dataset 38 distinct traffic conditions have been identified, out of which 22 were ‘unseen’ in the first testing set. The rationale of using a second testing dataset was to validate the developed memory-based prediction system in a manner consistent with the approach used for the model-based prediction component.

## **6.2 MEMORY-BASED PREDICTION AND OPTIMAL SETTINGS**

The memory-based CBR case base was deployed in a Microsoft database environment (MSSQL) and a specific Structured Query Language (SQL) interrogation script was performed to search the case base for similar cases with the ones identified in the second testing dataset. The cases of the memory-based approach were characterized by three features:

- the timestamp within a time window of  $\pm 30$  minutes from the current time;
- the weekday of the prediction;

- and the traffic conditions for the target location only.

This set of features was found to be the most effective after several other combinations were evaluated, such as the traffic conditions at the adjacent stations and different time window lengths. A query script developed in Practical Extraction and Report Language (Perl) was employed for interrogating the MSSQL database in the process of retrieving similar cases. Ultimately, through the PQM component, as detailed in the previous chapter, the optimal settings for the hybrid prediction system using the second testing set were identified. The results of this integrated optimization procedure are depicted in TABLE 5, TABLE 6, TABLE 7, and TABLE 8.

The four tables describe similar information. Each record in the table shows the optimal settings for the hybrid traffic prediction system, in terms of: the prediction method (a specific NN type vs. CBR), the type of input (whether to consider the current station, the downstream station and/or the upstream station), the relevance of long-term memory component (whether to consider the LTM input for NN-based predictions only or not), and the comparative *AARE* and *RMSE* for each set of traffic condition combinations. Each table row is distinctly identified by a case number, which corresponds to the traffic conditions combination at the three stations. The tables were sorted by the computed performance measure for the hybrid NN-CBR approach.

Considering the performance measures for the NN approach only, the TABLE 1 through TABLE 4 also show that no particular network topology seemed to have outperformed the others for all cases. The same applies to the type of inputs and the inclusion of LTM component. For 5-minute predictions of speed, the AARE did not exceed 9.5% for 22 out of the 38 cases. Three cases produced errors as high as nearly 13% and the thirteen cases produced errors between 10% and 13%. A possible explanation for this is that the predictions associated with those conditions

TABLE 1 NN OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON *AARE*

(a) 5-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings (5 min)			Average Absolute Relative Error
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
7	2	2	3	MODULAR	XY	No	0.001
8	2	2	4	TLFN	XYZ	Yes	0.001
15	3	2	2	MLP	ZY	Yes	0.001
16	3	2	3	MODULAR	ZY	Yes	0.001
18	3	3	3	TLFN	Y	No	0.001
20	3	4	3	CANFIS	ZY	Yes	0.001
21	3	4	4	Jordan-Elman	Y	Yes	0.001
22	4	3	3	CANFIS	XY	No	0.001
25	4	4	4	TLFN	ZY	Yes	0.001
4	2	1	4	PRN	XY	Yes	0.002
6	2	2	2	PRN	XYZ	No	0.002
19	3	3	4	Jordan-Elman	ZY	Yes	0.002
23	4	3	4	CANFIS	XY	No	0.002
9	2	3	3	TLFN	XY	No	0.003
10	2	3	4	TLFN	ZY	No	0.003
13	3	1	3	Jordan-Elman	ZY	No	0.003
24	4	4	3	TLFN	ZY	No	0.003
17	3	2	4	PCA-hybrid	XY	Yes	0.006
11	3	1	1	CANFIS	ZY	Yes	0.008
14	3	2	1	TLFN	XY	Yes	0.008
3	2	1	3	PCA-hybrid	XYZ	No	0.012
5	2	2	1	PCA-hybrid	XYZ	No	0.014
1	2	1	1	MODULAR	XY	Yes	0.017
2	2	1	2	MODULAR	XY	Yes	0.032
12	3	1	2	PRN	Y	Yes	0.058

(b) 10-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings (10 min)			Average Absolute Relative Error
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
7	2	2	3	MODULAR	XY	Yes	0.001
9	2	3	3	TLFN	ZY	Yes	0.001
15	3	2	2	TLFN	Y	Yes	0.001
20	3	4	3	PCA-hybrid	XYZ	No	0.001
21	3	4	4	Jordan-Elman	XYZ	No	0.001
22	4	3	3	PCA-hybrid	XYZ	No	0.001
23	4	3	4	MLP	XY	No	0.001
25	4	4	4	PRN	XYZ	Yes	0.001
5	2	2	1	PRN	Y	Yes	0.002
6	2	2	2	PRN	XYZ	No	0.002
8	2	2	4	TLFN	Y	Yes	0.002
24	4	4	3	TLFN	ZY	No	0.002
12	3	1	2	MLP	XY	No	0.004
11	3	1	1	CANFIS	ZY	Yes	0.006
18	3	3	3	TLFN	XY	Yes	0.006
14	3	2	1	PRN	ZY	Yes	0.007
4	2	1	4	TLFN	ZY	No	0.012
3	2	1	3	CANFIS	XY	Yes	0.013
17	3	2	4	CANFIS	XY	Yes	0.018
10	2	3	4	TLFN	XYZ	No	0.031
1	2	1	1	CANFIS	XYZ	Yes	0.035
16	3	2	3	CANFIS	XY	Yes	0.043
2	2	1	2	MODULAR	XY	Yes	0.046
19	3	3	4	Jordan-Elman	XYZ	Yes	0.059
13	3	1	3	MODULAR	ZY	No	0.119

TABLE 2 NN OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON AARE

(a) 15-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings (15 min)			Average Absolute Relative Error
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
7	2	2	3	MLP	XY	No	0.001
8	2	2	4	MLP	ZY	Yes	0.001
9	2	3	3	Jordan-Elman	ZY	Yes	0.001
14	3	2	1	Jordan-Elman	ZY	Yes	0.001
15	3	2	2	PCA-hybrid	ZY	Yes	0.001
21	3	4	4	Jordan-Elman	XYZ	Yes	0.001
22	4	3	3	TLFN	XYZ	No	0.001
3	2	1	3	PCA-hybrid	XY	No	0.002
23	4	3	4	Jordan-Elman	Y	Yes	0.002
24	4	4	3	Jordan-Elman	ZY	Yes	0.002
25	4	4	4	CANFIS	ZY	No	0.002
5	2	2	1	TLFN	Y	Yes	0.003
6	2	2	2	TLFN	XYZ	Yes	0.003
12	3	1	2	MLP	XY	Yes	0.008
4	2	1	4	Jordan-Elman	XYZ	Yes	0.01
18	3	3	3	Jordan-Elman	ZY	Yes	0.015
20	3	4	3	Jordan-Elman	XY	Yes	0.02
1	2	1	1	CANFIS	XYZ	Yes	0.037
16	3	2	3	MODULAR	ZY	Yes	0.039
17	3	2	4	PCA-hybrid	ZY	Yes	0.043
2	2	1	2	MODULAR	XY	Yes	0.044
11	3	1	1	MODULAR	XYZ	No	0.05
19	3	3	4	Jordan-Elman	ZY	Yes	0.08
10	2	3	4	PRN	ZY	No	0.113
13	3	1	3	MODULAR	ZY	No	0.181

(b) 20-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings (20 min)			Average Absolute Relative Error
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
18	3	3	3	TLFN	XY	Yes	0.003
23	4	3	4	MLP	XY	Yes	0.003
15	3	2	2	PCA-hybrid	XYZ	Yes	0.004
22	4	3	3	PRN	Y	Yes	0.008
24	4	4	3	PCA-hybrid	XYZ	No	0.008
17	3	2	4	MLP	XY	Yes	0.009
5	2	2	1	MLP	XY	Yes	0.01
8	2	2	4	PRN	ZY	Yes	0.01
9	2	3	3	TLFN	XY	Yes	0.01
14	3	2	1	TLFN	ZY	Yes	0.01
20	3	4	3	TLFN	XY	Yes	0.015
7	2	2	3	CANFIS	XY	Yes	0.016
4	2	1	4	TLFN	ZY	Yes	0.017
3	2	1	3	MLP	XY	No	0.019
25	4	4	4	Jordan-Elman	XYZ	Yes	0.019
12	3	1	2	MLP	XY	Yes	0.026
16	3	2	3	MLP	XY	Yes	0.026
21	3	4	4	TLFN	XY	Yes	0.03
1	2	1	1	MODULAR	XY	Yes	0.04
6	2	2	2	CANFIS	ZY	Yes	0.047
2	2	1	2	MODULAR	XY	Yes	0.049
11	3	1	1	MODULAR	XYZ	No	0.063
13	3	1	3	MODULAR	ZY	No	0.065
19	3	3	4	TLFN	XY	Yes	0.133
10	2	3	4	Jordan-Elman	XYZ	No	0.197

TABLE 3 NN OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON RMSE

(a) 5-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings (5 min)			Root Mean Square Error (mph)
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
4	2	1	4	PRN	XY	Yes	0.0
13	3	1	3	Jordan-Elman	ZY	No	0.2
5	2	2	1	PCA-hybrid	XYZ	No	0.3
22	4	3	3	CANFIS	XY	No	0.5
14	3	2	1	PRN	XYZ	No	0.8
23	4	3	4	MODULAR	ZY	No	0.9
24	4	4	3	PRN	Y	No	0.9
20	3	4	3	MLP	XYZ	No	1.1
25	4	4	4	TLFN	XYZ	Yes	1.7
15	3	2	2	TLFN	XY	No	2.0
11	3	1	1	CANFIS	ZY	Yes	2.1
1	2	1	1	PRN	XYZ	Yes	2.5
2	2	1	2	CANFIS	XYZ	Yes	2.5
12	3	1	2	TLFN	Y	Yes	2.6
21	3	4	4	TLFN	ZY	Yes	2.7
6	2	2	2	TLFN	XY	Yes	2.8
3	2	1	3	MODULAR	XY	Yes	3.0
18	3	3	3	Jordan-Elman	ZY	Yes	5.2
16	3	2	3	PCA-hybrid	XYZ	Yes	5.2
17	3	2	4	PCA-hybrid	XY	Yes	5.6
19	3	3	4	CANFIS	ZY	Yes	5.7
7	2	2	3	MODULAR	XYZ	No	6.5
9	2	3	3	Jordan-Elman	XYZ	Yes	7.7
8	2	2	4	MODULAR	XYZ	No	8.9
10	2	3	4	Jordan-Elman	XYZ	No	9.9

(b) 10-MINUTE PREDICTION HORIZON

Case	Traffic Conditions in terms of Speed (mph)			Optimal Settings (10 min)			Root Mean Square Error (mph)
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
4	2	1	4	Jordan-Elman	ZY	No	0.2
22	4	3	3	PCA-hybrid	XYZ	No	0.5
5	2	2	1	PCA-hybrid	XY	No	0.7
14	3	2	1	Jordan-Elman	XYZ	Yes	0.8
24	4	4	3	CANFIS	XY	No	1.2
20	3	4	3	PCA-hybrid	XY	No	1.4
23	4	3	4	MLP	XY	No	2.0
11	3	1	1	CANFIS	ZY	Yes	2.2
17	3	2	4	PCA-hybrid	XYZ	Yes	2.4
1	2	1	1	TLFN	Y	Yes	2.6
2	2	1	2	MODULAR	XY	Yes	3.0
25	4	4	4	TLFN	ZY	Yes	3.8
15	3	2	2	Jordan-Elman	ZY	Yes	4.1
12	3	1	2	Jordan-Elman	ZY	Yes	4.1
16	3	2	3	MLP	XYZ	Yes	4.5
7	2	2	3	CANFIS	ZY	Yes	5.7
21	3	4	4	TLFN	XYZ	Yes	5.7
6	2	2	2	TLFN	XYZ	Yes	5.7
8	2	2	4	MLP	XYZ	Yes	6.2
3	2	1	3	CANFIS	ZY	Yes	7.4
18	3	3	3	TLFN	XY	Yes	7.8
19	3	3	4	Jordan-Elman	XYZ	Yes	8.7
10	2	3	4	CANFIS	XY	Yes	8.9
13	3	1	3	Jordan-Elman	XY	Yes	9.8
9	2	3	3	Jordan-Elman	XYZ	Yes	11.5

TABLE 4 NN OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON *RMSE*

(a) 15-MINUTE PREDICTION HORIZON

Case	Traffic Conditions in terms of Speed (mph)			Optimal Settings (15 min)			Root Mean Square Error (mph)
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
4	2	1	4	Jordan-Elman	ZY	No	0.2
5	2	2	1	PCA-hybrid	XYZ	No	0.3
22	4	3	3	CANFIS	XY	No	0.5
14	3	2	1	Jordan-Elman	XYZ	Yes	0.8
23	4	3	4	MODULAR	ZY	No	0.9
24	4	4	3	MODULAR	XYZ	No	1.0
20	3	4	3	MLP	XYZ	No	1.1
11	3	1	1	CANFIS	ZY	Yes	2.1
1	2	1	1	MODULAR	ZY	No	2.5
2	2	1	2	CANFIS	XYZ	Yes	2.5
25	4	4	4	CANFIS	ZY	No	2.7
3	2	1	3	MODULAR	XY	Yes	3.0
12	3	1	2	MODULAR	XY	No	3.2
21	3	4	4	CANFIS	XY	No	3.3
6	2	2	2	PCA-hybrid	XY	No	3.5
15	3	2	2	PCA-hybrid	XY	No	4.0
16	3	2	3	PCA-hybrid	XYZ	Yes	5.2
17	3	2	4	PCA-hybrid	XY	Yes	5.6
19	3	3	4	CANFIS	ZY	Yes	5.7
13	3	1	3	MODULAR	ZY	No	6.2
7	2	2	3	MODULAR	XYZ	No	6.5
18	3	3	3	PCA-hybrid	XYZ	Yes	7.7
8	2	2	4	MODULAR	XYZ	No	8.9
10	2	3	4	MODULAR	XYZ	Yes	10.3
9	2	3	3	MLP	XYZ	No	10.7

(b) 20-MINUTE PREDICTION HORIZON

Case	Traffic Conditions in terms of Speed (mph)			Optimal Settings (20 min)			Root Mean Square Error (mph)
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Network	Inputs	LTM/Time Component	
4	2	1	4	TLFN	ZY	Yes	0.3
22	4	3	3	PRN	Y	Yes	0.6
14	3	2	1	TLFN	ZY	Yes	0.9
24	4	4	3	PCA-hybrid	XYZ	No	0.9
5	2	2	1	CANFIS	ZY	Yes	1.3
23	4	3	4	MLP	XY	Yes	1.4
1	2	1	1	TLFN	Y	Yes	3.1
16	3	2	3	MLP	XYZ	Yes	3.3
7	2	2	3	MLP	XYZ	Yes	3.4
17	3	2	4	PCA-hybrid	ZY	Yes	4.2
11	3	1	1	MODULAR	ZY	No	4.6
15	3	2	2	PCA-hybrid	XYZ	Yes	5.1
20	3	4	3	CANFIS	ZY	Yes	5.3
8	2	2	4	MLP	XYZ	Yes	5.3
12	3	1	2	Jordan-Elman	ZY	Yes	5.7
2	2	1	2	Jordan-Elman	XYZ	Yes	6.6
25	4	4	4	Jordan-Elman	XYZ	Yes	7.5
18	3	3	3	TLFN	ZY	Yes	7.5
3	2	1	3	Jordan-Elman	XYZ	Yes	9.1
6	2	2	2	MLP	XY	Yes	9.1
21	3	4	4	Jordan-Elman	XYZ	Yes	9.7
9	2	3	3	MODULAR	XYZ	No	10.1
10	2	3	4	MODULAR	XYZ	No	11.1
19	3	3	4	CANFIS	XYZ	Yes	11.9
13	3	1	3	MODULAR	ZY	No	14.1

TABLE 5 HYBRID OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON AARE

(a) 5-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (5-min PH)			Average Absolute Relative Error	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Prediction Method	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
19	2	4	3	PCA-hybrid	ZY	No	0.002	0.002
29	3	4	3	CANFIS	ZY	Yes	0.011	0.011
30	3	4	4	CBR	Y	n/a	0.029	0.018
38	4	4	4	CBR	Y	n/a	0.048	0.018
9	1	4	4	PCA-hybrid	XYZ	Yes	0.031	0.031
20	2	4	4	MLP	XY	Yes	0.032	0.032
13	2	1	4	PRN	XY	Yes	0.039	0.039
36	4	3	3	CANFIS	XY	No	0.041	0.041
3	1	1	3	MODULAR	XYZ	No	0.046	0.046
18	2	3	4	PRN	ZY	No	0.054	0.054
4	1	2	2	MLP	XY	Yes	0.062	0.062
6	1	2	4	Jordan-Elman	Y	No	0.063	0.063
33	4	1	3	CANFIS	ZY	No	0.069	0.069
8	1	3	4	Jordan-Elman	Y	No	0.071	0.071
31	4	1	1	MODULAR	ZY	No	0.071	0.071
7	1	3	3	MLP	Y	No	0.075	0.075
34	4	2	2	MODULAR	XYZ	No	0.079	0.079
37	4	3	4	CBR	Y	n/a	0.085	0.082
27	3	3	3	CBR	Y	n/a	0.096	0.082
17	2	3	3	CBR	Y	n/a	0.112	0.082
28	3	3	4	CBR	Y	n/a	0.114	0.082
16	2	2	4	TLFN	XYZ	Yes	0.084	0.084
35	4	2	3	CANFIS	XYZ	Yes	0.089	0.089
1	1	1	1	MLP	XY	Yes	0.091	0.091
2	1	1	2	CANFIS	XY	No	0.091	0.091
21	3	1	1	CBR	Y	n/a	0.097	0.093
10	2	1	1	CBR	Y	n/a	0.112	0.093
22	3	1	2	CBR	Y	n/a	0.116	0.093
11	2	1	2	CBR	Y	n/a	0.117	0.093
12	2	1	3	CBR	Y	n/a	0.118	0.093
23	3	1	3	CBR	Y	n/a	0.129	0.093
26	3	2	4	PCA-hybrid	XY	Yes	0.098	0.098
25	3	2	3	CBR	Y	n/a	0.123	0.107
14	2	2	2	CBR	Y	n/a	0.125	0.107
15	2	2	3	CBR	Y	n/a	0.125	0.107
24	3	2	2	CBR	Y	n/a	0.143	0.107
32	4	1	2	PCA-hybrid	ZY	No	0.107	0.107
5	1	2	3	PRN	Y	No	0.116	0.116

(b) 10-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (10-min PH)			Average Absolute Relative Error	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
6	1	2	4	MLP	XY	Yes	0.015	0.015
9	1	4	4	CANFIS	ZY	No	0.020	0.020
29	3	4	3	CANFIS	ZY	Yes	0.028	0.028
38	4	4	4	CBR	Y	n/a	0.042	0.030
30	3	4	4	CBR	Y	n/a	0.089	0.030
13	2	1	4	TLFN	ZY	No	0.056	0.056
36	4	3	3	CANFIS	XY	No	0.056	0.056
33	4	1	3	MLP	XYZ	No	0.060	0.060
3	1	1	3	PCA-hybrid	XYZ	Yes	0.073	0.073
19	2	4	3	Jordan-Elman	Y	Yes	0.084	0.084
1	1	1	1	MLP	XY	Yes	0.092	0.092
8	1	3	4	Jordan-Elman	Y	No	0.095	0.095
7	1	3	3	Jordan-Elman	Y	No	0.099	0.099
26	3	2	4	CBR	Y	n/a	0.109	0.100
24	3	2	2	CBR	Y	n/a	0.142	0.100
15	2	2	3	CBR	Y	n/a	0.145	0.100
14	2	2	2	CBR	Y	n/a	0.169	0.100
25	3	2	3	CBR	Y	n/a	0.173	0.100
16	2	2	4	CBR	Y	n/a	0.217	0.100
37	4	3	4	CBR	Y	n/a	0.110	0.102
27	3	3	3	CBR	Y	n/a	0.113	0.102
18	2	3	4	CBR	Y	n/a	0.114	0.102
17	2	3	3	CBR	Y	n/a	0.117	0.102
28	3	3	4	CBR	Y	n/a	0.134	0.102
5	1	2	3	PCA-hybrid	XYZ	Yes	0.104	0.104
10	2	1	1	MODULAR	XY	Yes	0.107	0.107
2	1	1	2	MODULAR	XYZ	No	0.108	0.108
21	3	1	1	CANFIS	ZY	Yes	0.109	0.109
12	2	1	3	PCA-hybrid	XYZ	No	0.110	0.110
20	2	4	4	CANFIS	ZY	No	0.111	0.111
34	4	2	2	CANFIS	XY	Yes	0.115	0.115
22	3	1	2	MODULAR	XY	No	0.118	0.118
31	4	1	1	CANFIS	ZY	Yes	0.119	0.119
35	4	2	3	PRN	Y	No	0.121	0.121
11	2	1	2	MODULAR	XY	Yes	0.130	0.130
4	1	2	2	TLFN	Y	Yes	0.134	0.134
23	3	1	3	CBR	Y	n/a	0.156	0.144
32	4	1	2	Jordan-Elman	Y	Yes	0.151	0.151

TABLE 6 HYBRID OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON AARE

(a) 15-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (15-min PH)			Average Absolute Relative Error	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
13	2	1	4	Jordan-Elman	XYZ	Yes	0.004	0.004
35	4	2	3	MODULAR	XY	Yes	0.016	0.016
34	4	2	2	MLP	XYZ	Yes	0.021	0.021
16	2	2	4	MLP	ZY	Yes	0.033	0.033
29	3	4	3	MLP	XYZ	Yes	0.038	0.038
30	3	4	4	CBR	Y	n/a	0.055	0.042
38	4	4	4	CBR	Y	n/a	0.109	0.042
9	1	4	4	MLP	XY	Yes	0.044	0.044
36	4	3	3	MODULAR	XYZ	No	0.051	0.051
31	4	1	1	MODULAR	ZY	Yes	0.053	0.053
33	4	1	3	CANFIS	XY	Yes	0.065	0.065
6	1	2	4	MODULAR	XY	Yes	0.086	0.086
1	1	1	1	PCA-hybrid	XY	Yes	0.089	0.089
2	1	1	2	MLP	XY	Yes	0.091	0.091
24	3	2	2	CBR	Y	n/a	0.220	0.097
26	3	2	4	CBR	Y	n/a	0.226	0.097
15	2	2	3	CBR	Y	n/a	0.232	0.097
14	2	2	2	CBR	Y	n/a	0.253	0.097
25	3	2	3	CBR	Y	n/a	0.273	0.097
18	2	3	4	CBR	Y	n/a	0.210	0.111
37	4	3	4	CBR	Y	n/a	0.219	0.111
27	3	3	3	CBR	Y	n/a	0.226	0.111
28	3	3	4	CBR	Y	n/a	0.240	0.111
17	2	3	3	CBR	Y	n/a	0.244	0.111
5	1	2	3	TLFN	XYZ	Yes	0.158	0.158
4	1	2	2	Jordan-Elman	XYZ	No	0.162	0.162
10	2	1	1	CBR	Y	n/a	0.171	0.164
22	3	1	2	CBR	Y	n/a	0.172	0.164
21	3	1	1	CBR	Y	n/a	0.173	0.164
11	2	1	2	CBR	Y	n/a	0.176	0.164
23	3	1	3	CBR	Y	n/a	0.257	0.164
12	2	1	3	CBR	Y	n/a	0.284	0.164
32	4	1	2	Jordan-Elman	ZY	Yes	0.169	0.169
19	2	4	3	MLP	Y	No	0.189	0.189
8	1	3	4	Jordan-Elman	Y	No	0.196	0.196
7	1	3	3	CANFIS	Y	No	0.197	0.197
3	1	1	3	PRN	ZY	Yes	0.199	0.199
20	2	4	4	PCA-hybrid	ZY	No	0.221	0.221

(b) 20-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (20-min PH)			Average Absolute Relative Error	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
33	4	1	3	PCA-hybrid	ZY	Yes	0.009	0.009
34	4	2	2	MLP	XY	Yes	0.010	0.010
35	4	2	3	PCA-hybrid	XYZ	Yes	0.026	0.026
13	2	1	4	TLFN	ZY	Yes	0.032	0.032
16	2	2	4	PRN	ZY	Yes	0.044	0.044
29	3	4	3	TLFN	XY	Yes	0.057	0.057
30	3	4	4	CBR	Y	n/a	0.072	0.063
38	4	4	4	CBR	Y	n/a	0.108	0.063
9	1	4	4	MODULAR	XYZ	Yes	0.063	0.063
32	4	1	2	CANFIS	ZY	Yes	0.065	0.065
28	3	3	4	TLFN	XY	Yes	0.079	0.079
25	3	2	3	MLP	XY	Yes	0.102	0.102
26	3	2	4	CBR	Y	n/a	0.103	0.102
24	3	2	2	CBR	Y	n/a	0.143	0.102
15	2	2	3	CBR	Y	n/a	0.279	0.102
14	2	2	2	CBR	Y	n/a	0.296	0.102
1	1	1	1	MLP	XY	n/a	0.103	0.103
27	3	3	3	CBR	Y	n/a	0.120	0.114
36	4	3	3	CBR	Y	n/a	0.154	0.114
37	4	3	4	CBR	Y	n/a	0.155	0.114
18	2	3	4	CBR	Y	n/a	0.330	0.114
17	2	3	3	CBR	Y	n/a	0.431	0.114
3	1	1	3	Jordan-Elman	Y	No	0.115	0.115
2	1	1	2	TLFN	Y	Yes	0.120	0.120
6	1	2	4	Jordan-Elman	XY	No	0.134	0.134
4	1	2	2	Jordan-Elman	Y	No	0.138	0.138
7	1	3	3	TLFN	XY	Yes	0.145	0.145
5	1	2	3	PRN	Y	Yes	0.150	0.150
10	2	1	1	MODULAR	XY	Yes	0.151	0.151
21	3	1	1	CBR	Y	n/a	0.199	0.176
11	2	1	2	CBR	Y	n/a	0.210	0.176
22	3	1	2	CBR	Y	n/a	0.247	0.176
23	3	1	3	CBR	Y	n/a	0.305	0.176
12	2	1	3	CBR	Y	n/a	0.335	0.176
31	4	1	1	CANFIS	ZY	No	0.180	0.180
20	2	4	4	PCA-hybrid	ZY	No	0.187	0.187
8	1	3	4	PRN	Y	No	0.267	0.267
19	2	4	3	Jordan-Elman	ZY	No	0.335	0.335

TABLE 7 HYBRID OPTIMAL SETTINGS FOR 5- AND 10-MINUTE PREDICTIONS BASED ON RMSE

(a) 5-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (5-min PH)			Root Mean Square Error (mph)	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
19	2	4	3	PCA-hybrid	ZY	No	0.1	0.1
29	3	4	3	CANFIS	XYZ	No	0.6	0.6
3	1	1	3	MODULAR	XYZ	No	0.9	0.9
31	4	1	1	MODULAR	ZY	No	1.1	1.1
4	1	2	2	MLP	XY	Yes	1.2	1.2
1	1	1	1	MLP	XY	Yes	1.3	1.3
33	4	1	3	CANFIS	ZY	No	1.4	1.4
2	1	1	2	CANFIS	XY	No	1.5	1.5
21	3	1	1	MODULAR	ZY	No	1.5	1.5
10	2	1	1	CBR	Y	n/a	1.6	1.5
22	3	1	2	CBR	Y	n/a	2.0	1.5
11	2	1	2	CBR	Y	n/a	2.1	1.5
12	2	1	3	CBR	Y	n/a	3.5	1.5
13	2	1	4	CBR	Y	n/a	3.5	1.5
23	3	1	3	CBR	Y	n/a	4.7	1.5
30	3	4	4	CBR	Y	n/a	1.8	1.7
38	4	4	4	CBR	XY	n/a	2.7	1.7
34	4	2	2	MODULAR	XYZ	No	1.8	1.8
9	1	4	4	PCA-hybrid	XYZ	Yes	1.8	1.8
20	2	4	4	MLP	XY	Yes	1.9	1.9
6	1	2	4	MODULAR	ZY	Yes	1.9	1.9
36	4	3	3	CANFIS	ZY	Yes	2.0	2.0
32	4	1	2	PCA-hybrid	ZY	No	2.3	2.3
14	2	2	2	MLP	XY	No	2.6	2.6
8	1	3	4	Jordan-Elman	Y	No	2.7	2.7
7	1	3	3	Jordan-Elman	Y	No	2.8	2.8
18	2	3	4	CANFIS	ZY	Yes	2.9	2.9
24	3	2	2	CBR	Y	n/a	3.6	3.0
26	3	2	4	CBR	Y	n/a	3.7	3.0
16	2	2	4	CBR	Y	n/a	3.7	3.0
15	2	2	3	CBR	Y	n/a	4.2	3.0
25	3	2	3	CBR	Y	n/a	4.5	3.0
27	3	3	3	CBR	Y	n/a	4.1	3.5
28	3	3	4	CBR	Y	n/a	4.3	3.5
17	2	3	3	CBR	Y	n/a	4.5	3.5
37	4	3	4	CBR	Y	n/a	4.6	3.5
35	4	2	3	PRN	Y	No	3.7	3.7
5	1	2	3	PRN	Y	No	4.4	4.4

(b) 10-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (10-min PH)			Root Mean Square Error (mph)	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
6	1	2	4	MLP	XY	Yes	0.286	0.286
9	1	4	4	CANFIS	ZY	No	1.148	1.148
1	1	1	1	MLP	XY	Yes	1.404	1.404
33	4	1	3	MLP	XYZ	No	1.423	1.423
29	3	4	3	CANFIS	ZY	Yes	1.620	1.620
10	2	1	1	MODULAR	XY	No	1.665	1.665
21	3	1	1	CANFIS	ZY	Yes	1.677	1.677
2	1	1	2	PCA-hybrid	XYZ	Yes	1.715	1.715
38	4	4	4	CBR	XYZ	n/a	2.359	1.773
30	3	4	4	CBR	ZY	n/a	4.358	1.773
31	4	1	1	CANFIS	ZY	Yes	1.832	1.832
22	3	1	2	MODULAR	XY	No	2.007	2.007
3	1	1	3	PCA-hybrid	XYZ	Yes	2.372	2.372
4	1	2	2	MODULAR	XYZ	No	2.498	2.498
11	2	1	2	MODULAR	XY	No	2.614	2.614
24	3	2	2	CBR	Y	n/a	3.767	2.683
14	2	2	2	CBR	Y	n/a	3.788	2.683
26	3	2	4	CBR	Y	n/a	4.340	2.683
15	2	2	3	CBR	Y	n/a	5.464	2.683
25	3	2	3	CBR	Y	n/a	6.852	2.683
16	2	2	4	CBR	Y	n/a	7.295	2.683
12	2	1	3	CBR	Y	n/a	3.891	2.814
13	2	1	4	CBR	Y	n/a	5.236	2.814
23	3	1	3	CBR	Y	n/a	6.132	2.814
36	4	3	3	PCA-hybrid	XY	No	3.231	3.231
20	2	4	4	CANFIS	ZY	No	3.791	3.791
8	1	3	4	Jordan-Elman	Y	No	4.017	4.017
7	1	3	3	Jordan-Elman	Y	No	4.125	4.125
35	4	2	3	PRN	Y	No	4.130	4.130
34	4	2	2	CANFIS	XY	Yes	4.699	4.699
28	3	3	4	CBR	Y	n/a	5.472	4.781
17	2	3	3	PRN	Y	No	4.809	4.809
18	2	3	4	CANFIS	ZY	Yes	5.059	5.059
27	3	3	3	PRN	Y	No	5.069	5.069
37	4	3	4	CANFIS	XY	Yes	5.168	5.168
5	1	2	3	TLFN	Y	Yes	5.728	5.728
32	4	1	2	CANFIS	ZY	Yes	5.972	5.972
19	2	4	3	MODULAR	XYZ	No	7.735	7.735

TABLE 8 HYBRID OPTIMAL SETTINGS FOR 15- AND 20-MINUTE PREDICTIONS BASED ON RMSE

(a) 15-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (15-min PH)			Root Mean Square Error (mph)	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
31	4	1	1	MODULAR	ZY	Yes	0.913	0.913
34	4	2	2	MODULAR	XYZ	Yes	1.129	1.129
1	1	1	1	MLP	XY	Yes	1.365	1.365
6	1	2	4	CANFIS	XY	Yes	1.512	1.512
2	1	1	2	MLP	XY	Yes	1.595	1.595
36	4	3	3	CBR	Y	n/a	3.022	2.230
27	3	3	3	CBR	Y	n/a	6.174	2.230
28	3	3	4	CBR	Y	n/a	6.364	2.230
17	2	3	3	CBR	Y	n/a	6.921	2.230
18	2	3	4	CBR	Y	n/a	7.065	2.230
37	4	3	4	CBR	Y	n/a	9.705	2.230
29	3	4	3	PCA-hybrid	ZY	Yes	2.264	2.264
9	1	4	4	MLP	XY	Yes	2.340	2.340
30	3	4	4	CBR	XY	n/a	3.053	2.400
38	4	4	4	CBR	XYZ	n/a	5.194	2.400
21	3	1	1	MODULAR	XY	No	2.466	2.466
10	2	1	1	CBR	Y	n/a	2.617	2.482
22	3	1	2	CBR	Y	n/a	2.822	2.482
11	2	1	2	CBR	Y	n/a	3.854	2.482
12	2	1	3	CBR	Y	n/a	6.505	2.482
23	3	1	3	CBR	Y	n/a	6.755	2.482
13	2	1	4	CBR	Y	n/a	10.922	2.482
4	1	2	2	MLP	XYZ	No	2.570	2.570
33	4	1	3	MLP	XY	Yes	2.690	2.690
24	3	2	2	CBR	Y	n/a	3.872	3.359
26	3	2	4	CBR	Y	n/a	5.893	3.359
14	2	2	2	CBR	Y	n/a	6.019	3.359
15	2	2	3	CBR	Y	n/a	6.873	3.359
25	3	2	3	CBR	Y	n/a	7.341	3.359
16	2	2	4	CBR	Y	n/a	9.363	3.359
8	1	3	4	Jordan-Elman	Y	No	4.933	4.933
7	1	3	3	Jordan-Elman	Y	No	4.968	4.968
20	2	4	4	PCA-hybrid	ZY	No	5.438	5.438
32	4	1	2	CANFIS	ZY	Yes	5.579	5.579
3	1	1	3	MLP	ZY	Yes	6.371	6.371
5	1	2	3	Jordan-Elman	Y	Yes	7.014	7.014
35	4	2	3	Jordan-Elman	XYZ	Yes	7.196	7.196
19	2	4	3	PCA-hybrid	ZY	No	13.541	13.541

(b) 20-MINUTE PREDICTION HORIZON

Case	Congestion Indicator			Optimal Settings for NN-CBR Approach (20-min PH)			Root Mean Square Error (mph)	
	Downstream Station (X)	Current Station (Y)	Upstream Station (Z)	Model/Memory-based Approach	Inputs	LTM/Time Component	Optimized NN Approach	NN-CBR Approach
33	4	1	3	PCA-hybrid	ZY	Yes	0.527	0.527
34	4	2	2	MLP	XY	Yes	0.547	0.547
28	3	3	4	TLFN	Y	No	1.139	1.139
25	3	2	3	Jordan-Elman	Y	Yes	1.324	1.324
1	1	1	1	MLP	XY	Yes	1.616	1.616
26	3	2	4	Jordan-Elman	Y	Yes	2.022	2.022
6	1	2	4	CANFIS	XY	Yes	2.392	2.392
27	3	3	3	PRN	Y	Yes	2.411	2.411
10	2	1	1	PCA-hybrid	XY	No	2.519	2.519
32	4	1	2	CANFIS	ZY	Yes	2.583	2.583
31	4	1	1	CANFIS	ZY	No	2.816	2.816
2	1	1	2	PCA-hybrid	XY	Yes	2.936	2.936
24	3	2	2	CBR	Y	n/a	3.071	2.949
14	2	2	2	CBR	Y	n/a	7.968	2.949
16	2	2	4	CBR	Y	n/a	16.331	2.949
15	2	2	3	CBR	Y	n/a	16.428	2.949
35	4	2	3	Jordan-Elman	ZY	No	2.970	2.970
9	1	4	4	MODULAR	XYZ	Yes	3.042	3.042
21	3	1	1	CANFIS	ZY	Yes	3.049	3.049
5	1	2	3	PRN	Y	Yes	3.251	3.251
29	3	4	3	PCA-hybrid	XYZ	No	3.385	3.385
22	3	1	2	MODULAR	XY	No	3.434	3.434
30	3	4	4	CBR	Y	n/a	3.568	3.537
38	4	4	4	CBR	Y	n/a	5.125	3.537
11	2	1	2	CBR	Y	n/a	5.469	4.110
12	2	1	3	CBR	Y	n/a	5.558	4.110
13	2	1	4	CBR	Y	n/a	10.019	4.110
23	3	1	3	CBR	Y	n/a	23.307	4.110
7	1	3	3	TLFN	XY	No	5.013	5.013
4	1	2	2	PCA-hybrid	ZY	No	5.090	5.090
36	4	3	3	CBR	Y	n/a	6.731	5.442
18	2	3	4	CBR	Y	n/a	8.499	5.442
37	4	3	4	CBR	Y	n/a	11.077	5.442
17	2	3	3	CBR	Y	n/a	14.150	5.442
3	1	1	3	MLP	ZY	Yes	5.729	5.729
20	2	4	4	MLP	ZY	No	5.999	5.999
8	1	3	4	Jordan-Elman	ZY	Yes	7.569	7.569
19	2	4	3	PCA-hybrid	ZY	No	17.097	17.097

are independent of the information relayed by STM and LTM components, and therefore, the network topologies could not build sufficient internal representations of such cases. For 10-minute predictions, 22 cases exhibited errors less than 12%, 15 less than 17%, and 1 as high as 22%. The same explanation may apply. However, both cases exhibited higher errors for 10-minute predictions than for 5-minute predictions, confirming the intuition that prediction accuracy diminishes with longer horizons. Similar conclusions can be driven from 15- and 20-minute predictions, which show almost consistently higher errors for the worst cases as the prediction horizon increases.

The optimal settings were also selected based on the RMSE for each prediction horizon. Comparisons show some discrepancies between optimal settings for each case based on each measure, but this is the effect of the different meanings of the two performance measures. In addition, by comparing the two performance measures across the prediction horizon for each of the two prediction methods (with and without CBR), one can clearly see the improvement achieved by integrating the memory-based predictor. These results are graphically depicted in FIGURE 17 and FIGURE 18. These figures illustrate the comparative overall RMSE and AARE for each family of NN and for the two optimization procedures developed at each prediction horizon. However, to test if the differences in performance are statistically significant, both Friedman and Wilcoxon tests were conducted for each prediction horizon (see the tests results listed in APPENDIX A – APPENDIX D). The statistical analysis was conducted with the SYSTAT statistical software package.

The Wilcoxon signed rank test is a nonparametric test that compares the median of one column of numbers to a theoretical median. The test uses the difference between paired related measurements (such as before and after, NN and NN-CBR) and evaluates whether the distribution of the paired differences deviates from the zero value. Conversely, the Friedman test computes a

Friedman two-way analysis of variance on selected variables. This test is a nonparametric extension of the paired t test, where, instead of two measures, each subject has n measures ( $n > 2$ ). In this study the four measures ( $n=4$ ) were considered. These measures were the prediction errors associated with each traffic conditions combinations, based on predictions made by static NNs separately, dynamic NN separately, static/dynamic NN optimization, and hybrid NN/CBR optimization. In other terms, the Friedman statistic is used to test the hypothesis that there is no systematic response or pattern across the variables.

The Friedman test showed that for all prediction horizons and all 38 cases, the prediction performance of the four prediction optimization procedures were significantly different from each other. In other words the null hypothesis, which is *the four prediction methods errors come from the same distribution*, was rejected due to the fact that for all predictions horizons and considering both performance measures, *AARE* and *RMSE*, compared with the table value for a 0.05 level of significance ( $\chi^2_{TABLE} = 7.82$ ) the probability was smaller ( $<0.001$  or less) assuming Chi-square distribution with 3 degrees of freedom.

The Wilcoxon signed ranks test also showed that the two-sided probabilities using normal approximation were less than 0.0005 (this is the threshold in SYSTAT) for all the cases when comparing the NN optimal predictions with the NN-CBR optimized approach. In other words the null hypothesis, which is *the two performance measures medians are the same*, was rejected at a 0.05 level of significance. These tests indicate that the reduction in RMSE and AARE of the CBR-NN approach was significantly different from the other approaches.

### **6.3 PERFORMANCE ENVELOPES**

For each of the four levels of prediction horizon the performance envelopes were plotted for each case using the minimum and maximum values of AARE and RMSE. The performance

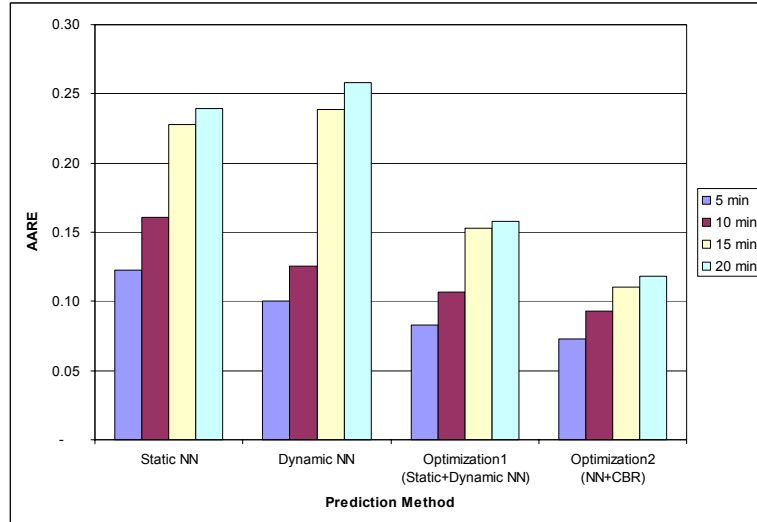


FIGURE 17: *AARE*-BASED PERFORMANCE OPTIMIZATION BY PREDICTION METHOD

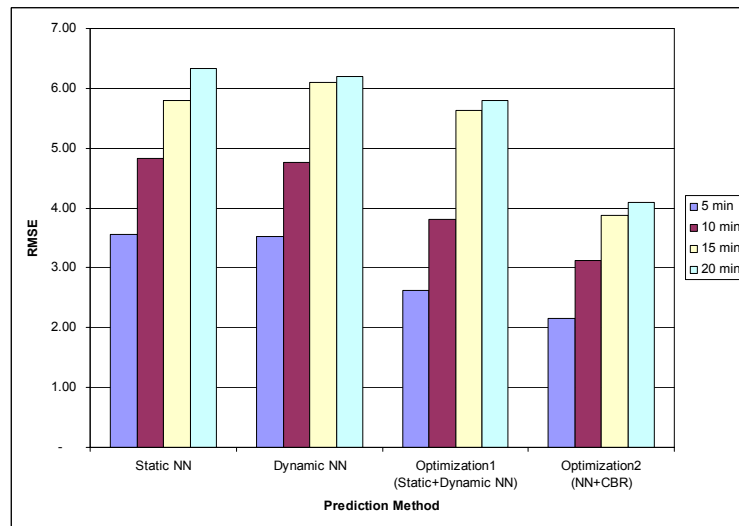


FIGURE 18: *RMSE*-BASED PERFORMANCE OPTIMIZATION BY PREDICTION METHOD

envelopes, shown in FIGURE 19 and FIGURE 20, can be used to identify cases where prediction errors do not meet a maximum acceptable threshold value that is appropriate for online implementation. For instance, if AARE threshold value is set to 10%, then FIGURE 19 can be used to identify all cases where optimal settings do not yield errors less than or equal to 10%. Consequently, when such cases are encountered in real world, the high levels of uncertainty in predictions will then be identified and, perhaps, eliminated entirely from traffic information

dissemination systems. This is an essential requirement for the traveling public to maintain high credibility in the prediction information disseminated by traffic management centers.

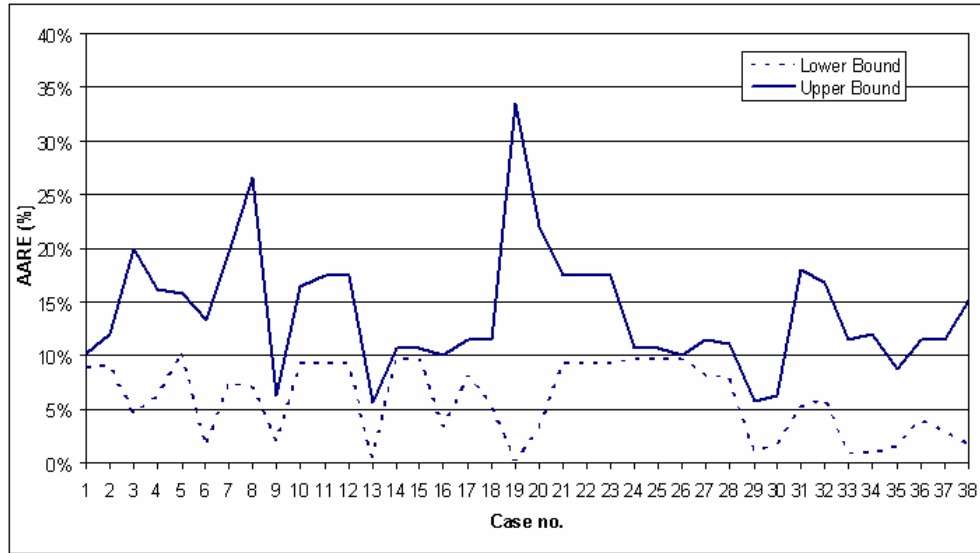


FIGURE 19: PERFORMANCE ENVELOPES BASED ON *AARE*

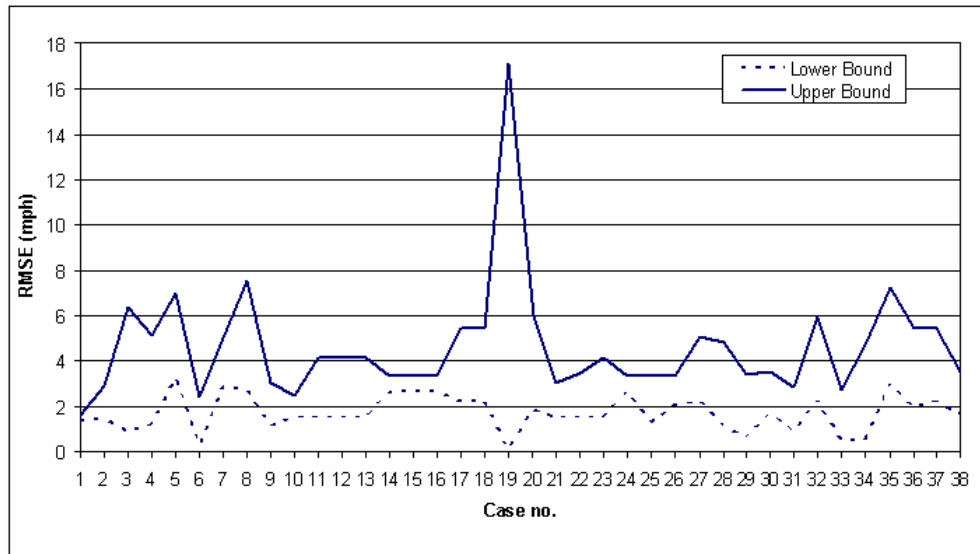


FIGURE 20: PERFORMANCE ENVELOPES BASED ON *RMSE*

It should be emphasized here that the high errors associated with such cases may be attributed to under-representation of those cases in the training data. This often leads to the network's inability to generalize under such conditions. Additional training with data collected from conditions poorly represented could essentially lead to improvement in the overall

prediction performance. Even with the potential improvement of performance as a result of additional training, it is reasonable to expect that there will be certain conditions where prediction accuracy is practically unacceptable. For any traffic prediction system to be successfully implemented, such conditions must be identifiable to recognize the limitations of the prediction model.

#### **6.4 EFFECT OF PREDICTION METHOD AND LTM COMPONENT**

The optimal settings defined in the study included the prediction method, static or dynamic ANNs vs. CBR and the relevance of LTM component as input to the network. Another analysis was carried out to identify the approach that is dominant in the optimal settings for each prediction horizon, as shown in FIGURE 21. The figure shows that no specific prediction approach appears to consistently dominate the optimal settings of all prediction horizons. However, CBR is shown to outperform both ANNs families performance for 15- and 20-minute predictions and to demonstrate comparable performance to the static NNs for 5-minute predictions. The figure also shows that when all cases are combined, CBR outperforms the NN approaches and dominates nearly 42% of the overall optimal settings. Practically, this means that if one would consider only the NN approach, both dynamic and static networks, by employing the CBR in a hybrid solution 42% of the of the cases are improved due to CBR.

Another important factor that was introduced in this study is the relevance of LTM component and its impact on the model-based component prediction performance. As mentioned earlier, the LTM component assists the network in retaining some of the historical information in its weights during the training process. Such memory component is useful in predicting the onset of congestion and in making longer-horizon predictions when predicted conditions are less dependent on information relayed by the STM component. The inclusion of both components can

essentially lead to a model capable of predicting recurrent (LTM component) and non-recurrent conditions (STM component).

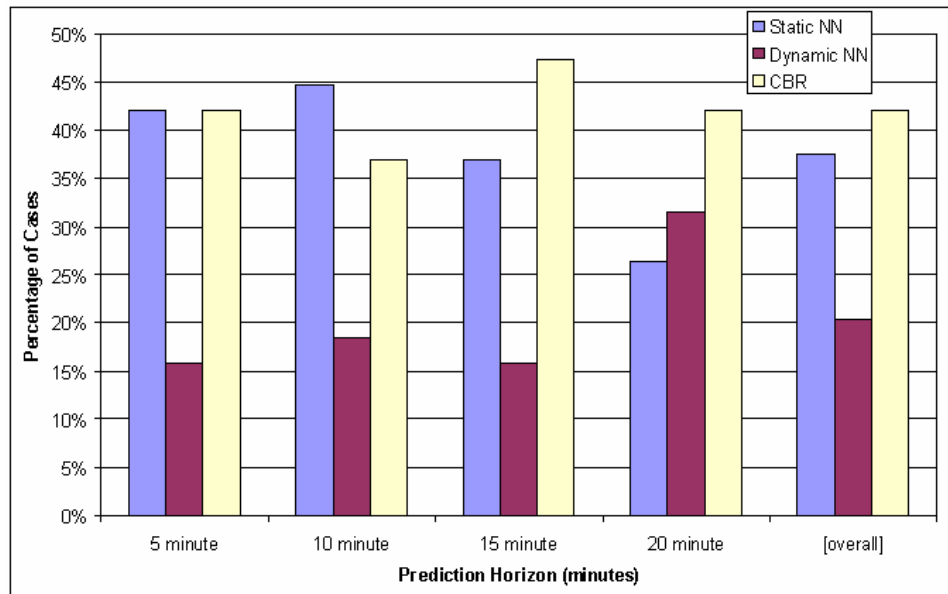


FIGURE 21: FREQUENCY OF OPTIMAL PREDICTION METHODS WITH PREDICTION HORIZONS BASED ON *AARE*

FIGURE 22 illustrates the role of the LTM component in each prediction horizon. The figure shows the percentage of cases whose optimal settings included the LTM component for the optimization approach with and without CBR component. The figure clearly shows that the relevance of LTM is more pronounced in longer-horizon predictions. Another interesting detail depicted in FIGURE 22, is the evidence of the importance of the LTM factor in the optimal settings. As can be easily seen from FIGURE 22 the percentage of cases with LTM in the optimal settings drops when the CBR approach is considered.

The percentage of optimal settings with LTM component increases consistently with the prediction horizon when the model-based approach, the NNs only respectively, is considered. This trend suggests that the critical role of LTM in making predictions more accurate in longer

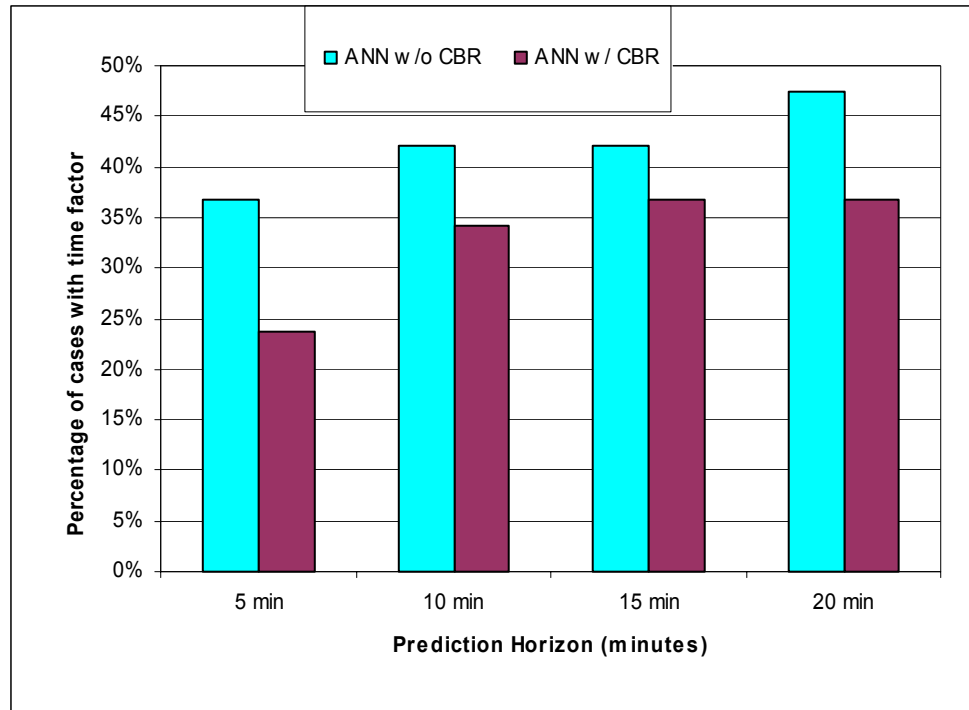


FIGURE 22: EFFECT OF TIME FACTOR ON OPTIMAL SETTINGS

horizons. The reduction of LTM in the hybrid approach could be explained by the fact that the role of LTM in potentially improving the prediction of the NNs in recurrent traffic conditions was overtaken by the CBR predictor. However, the CBR have not taken control for all the cases with LTM input active in the NN optimal settings. This is believed to happen due to insufficient representation of the data in the CBR case base and due to limitation of the CBR to model traffic conditions predictions with simple memory-based process, when traffic is more complex, dynamic and rather non-stationary. Therefore, a hybrid model-based and memory-based traffic prediction system would be more efficient than employing each of the two approaches separately. In the next section more insights about the role of CBR are provided.

## 6.5 EFFECT OF CBR ACROSS TRAFFIC CONDITIONS

More investigation about the role of CBR in the tested implementation of the hybrid system revealed several interesting facts. FIGURE 23 and FIGURE 24 represent the percentage of cases that were using CBR as *AARE*- and *RMSE*-based optimal predictor in the second testing set across the 4 levels of traffic conditions considered. Interestingly enough, from the *AARE* perspective about 60% of the low congested conditions (speeds of 40-60 mph) and more than 40% of the mild congested conditions (speeds of 20-40 mph) were improved by CBR prediction over the existing NN optimal settings. In addition, the same conclusion is supported from the *RMSE* perspective (see FIGURE 24), in which it can be seen that about 48% of the moderately congested conditions were improved by CBR). In other words, the CBR predictor helps predict more efficiently the behavior of traffic during transitions between two more stable states, free-flow conditions and heavily congested conditions.

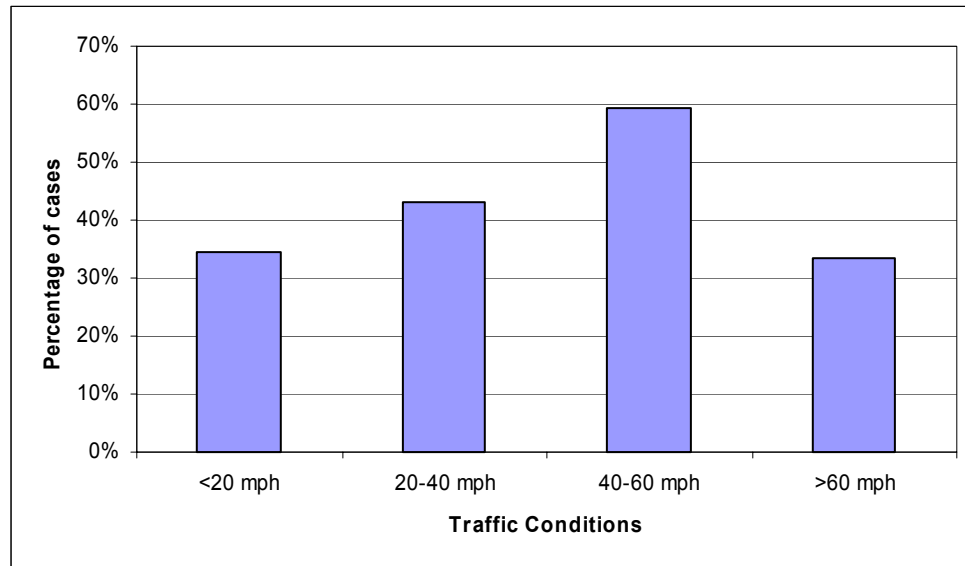


FIGURE 23: *AARE*-BASED OPTIMAL CBR PREDICTION ACROSS TRAFFIC CONDITIONS

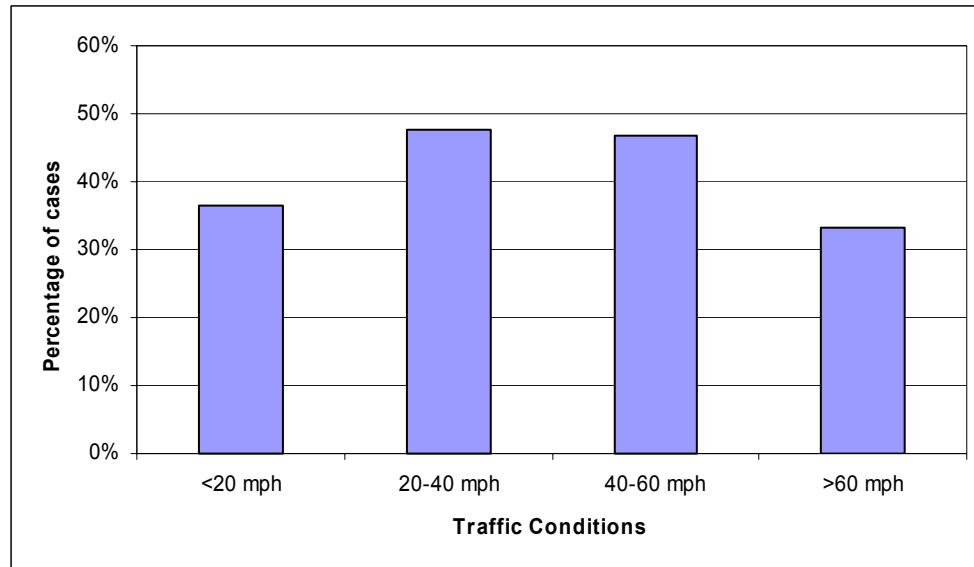


FIGURE 24: *RMSE*-BASED OPTIMAL CBR PREDICTION ACROSS TRAFFIC CONDITIONS

The findings of the analysis prove that a hybrid prediction system that integrates a model-based approach (such as artificial neural networks) and a memory-based approach (such as case-based reasoning) will more effectively address the difficulty associated with short-term traffic prediction during recurrent and non-recurrent conditions.

## 6.6 COMPARISON BETWEEN OPTIMAL AND NON-OPTIMAL PERFORMANCE

Prediction performance was optimized under different network settings and various traffic conditions. In order to quantify the performance improvements achieved by optimization with traffic conditions versus optimization with network settings only, we compare the optimal to non-optimal performance for each case. Non-optimal performance refers to optimization with network settings only, without considering the traffic conditions optimization. This results in selecting the best network topology and input settings for all traffic conditions. Optimal performance in this section, on the other hand, refers to optimization with network settings and traffic conditions. To facilitate the comparison, the reduction in errors of both scenarios was

calculated. FIGURE 25 shows the relative percentage reduction in AARE for each prediction horizon scenario.

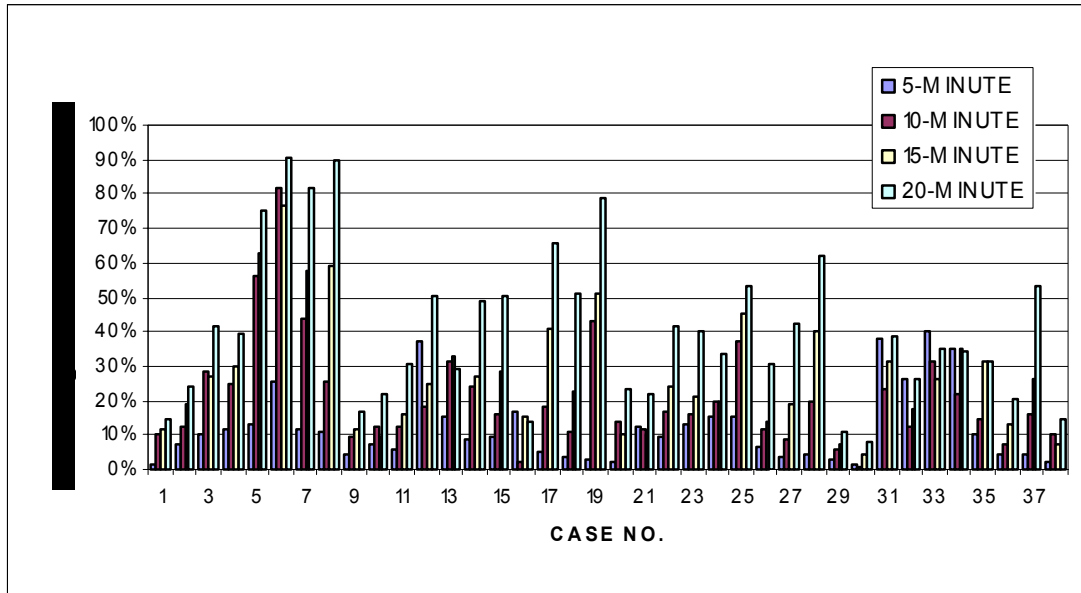


FIGURE 25: PERCENTAGE REDUCTIONS IN *AARE*

The figure shows significant improvements in prediction performance as a result of optimization with traffic conditions (for example, 100% percentage reduction means the performance measure reduced its value to half due to optimization across traffic conditions). The average improvement in terms of percentage reduction of errors was 12.3%, 21.5%, 27.5%, and 40.4% for 5-, 10-, 15-, and 20-minute predictions, respectively. Some improvements were as high as 80% to 90% such as cases 7, 8, 9, and 19. Similar results were obtained in the comparative evaluation of performance based on RMSE. A better illustration of the performance improvements in all cases combined can be seen in FIGURE 26. The figure shows the cumulative percentage of cases with a reduction in AARE that is less than or equal to a specific value. For instance, the figures show that 60% of the cases showed improvements of 90% or less in terms of AARE reduction. The steeper the curve the more cases with small reduction, the less steeper the more cases with higher reduction in error. In FIGURE 26 an average of about 50%

percent of the cases have reduction in AARE between 10% and 60%. Such improvements were exclusively attributed to performance optimization with traffic conditions over optimization with network settings only.

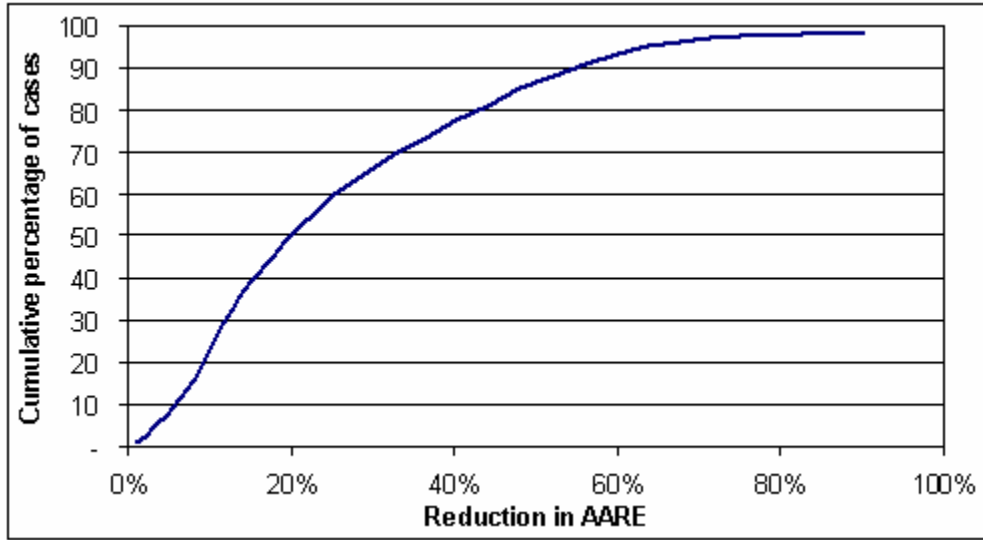


FIGURE 26: REDUCTIONS IN *AARE* BY OPTIMIZATION WITH TRAFFIC CONDITIONS

## 7 SUMMARY AND CONCLUSIONS

This study presented an approach to optimize the performance of freeway traffic prediction systems using a hybrid model-based and memory-based approach. The model-based approach applied four static neural network architectures (MLP, Modular, Hybrid PCA, and CANFIS) and three dynamic neural networks (Jordan/Elman, PRN, and TLFN), trained and tested under various network settings. A case-based reasoning (CBR) prediction system was used as a memory-based approach. The two approaches were then integrated to build a hybrid short-term traffic prediction system. The input to the ANNs was divided into two main components: short-term memory (STM) and long term memory (LTM). The STM component was represented by spatiotemporal information observed in the past 10 minutes and expressed in terms of 5 minute speed averages.

The LTM component was represented by the time stamp associated with each STM component in order to make the trained network time-cognizant. This technique was primarily introduced to allow the networks to learn from the historical information on traffic conditions during similar peak periods in the past. This was necessary to improve the prediction performance during recurring conditions when future predictions are less dependent on information encoded in the STM component. On the other hand, considering the memory-like structure of a CBR prediction, the following features were used to define each case in the case base: the traffic conditions at the target location; the 30-minute time window around the prediction moment; and the weekday for which the prediction is performed. The performance of both approaches was measured in terms of two types of errors: average absolute relative error - *AARE*, and root mean square error -

*RMSE*. Ultimately the two predictors were integrated by a Prediction Query Manager module based on the tested performance of each of the two approaches on a 10-day sample dataset.

The improvement achieved by combining the model-based and memory-based approaches was proved by statistical analysis. Two statistical tests commonly used for testing the differences between two variables, Friedman and Wilcoxon tests, were conducted for each prediction horizon. The Friedman test showed that for all prediction horizons and all 38 cases, the prediction performance of the two optimization procedures were significantly different from each other and with respect to each of the two NN families employed separately in the prediction process. The Wilcoxon signed ranks test also showed that the two-sided probabilities using normal approximation were less than 0.001 for most of the prediction horizons. This indicates that the reduction in *RMSE* and *AARE* of the CBR-NN approach was significantly different from the other approaches.

For the model-based (neural network only) approach the optimal settings were selected to minimize the prediction errors under different network settings, various traffic conditions, and multiple prediction horizons. The optimal settings were based on the testing results obtained from seven network topologies trained with the same data set. The network settings were varied by changing the input type in the STM component and toggling the LTM component for each of the four network topologies considered. Traffic conditions were broken down into four levels at each of the three stations. This resulted in a total of 25 combinations of different traffic conditions. Each of the 25 cases was optimized independently in order to identify the optimal network topology and the optimal network settings.

For the CBR approach, the same testing set used for testing the level of generalization of the NNs was employed, such that besides the previously mentioned performance measures, *AARE* and *RMSE*, a probability of occurrence measure was defined. The probability of occurrence is defined as the number of similar cases found in the case bases divided by the total number of possible cases. From the analysis of the plot of the probability of occurrence versus the prediction errors a direct utility of the probability of occurrence was identified. This means that a probability of occurrence calculated for each new case presented to the CBR prediction system will give an estimated maximum error of the prediction itself. This principle was used in the implementation of the PQM component of the hybrid prediction system.

The study showed that from the NN perspective no particular network topology has consistently outperformed the others for all prediction horizons and all cases. It was also found that the performance optimization under different traffic conditions has the advantage of identifying cases where none of the NN models were able to produce acceptable performance. While additional training with more data may improve the performance for some of those cases, it is still unequivocally critical to identify the major limitations of the prediction model and the cases where its performance falls below the minimum acceptable by traffic management centers. This is a critical issue to the dissemination of reliable information to the public and for the successful implementation of the prediction models.

Another important finding is the effect of the LTM component on the optimal performance. The results showed that the LTM component was more frequently seen in the optimal settings as the prediction horizon increases. An interesting observation is that the importance of LTM drops in the hybrid system, for which the CBR component is

expected to improve the performance from the memory perspective. Moreover, for the model-based approach only nearly 48% of the cases included the LTM component in their optimal settings for 20-minute predictions, as compared to 33% for 5-minute predictions. This trend emphasizes the critical role of LTM in making predictions more accurate in longer horizons. Finally, the study pointed out the comparative evaluation of prediction performance under optimal and non-optimal traffic condition settings. Using the reduction in AARE and RMSE the performance improvement in each case and for each prediction horizon was evaluated. An average improvement in AARE was shown in the range of 10% to 60% and as high as 80% to 90% for a few cases. The reduction in RMSE was also shown to be relatively large.

Based on the analysis presented in this research study the following general conclusions are drawn:

- short-term traffic prediction based on different artificial neural networks architectures have been optimized across prevailing traffic conditions at the same station, and both upstream and downstream stations, with respect to different prediction horizons
- in addition, by using a long-term memory component, more improvement have been achieved, with respect to recurrent and non-recurrent traffic conditions
- further improvements have been obtained, with respect to recurrent traffic conditions, by employing a case-based reasoning approach integrated within a hybrid traffic prediction system

One may consider the prediction approach presented in this thesis as starting point to implement a travel-time estimation algorithm, a useful component of the ATIS implementations at TMCs. The proved improvement of point estimates of the speed in

short-term speed prediction should lead to more reliable travel-time forecasts, assuming a good efficiency of the forecasting algorithms.

## **8 FUTURE WORK**

This study presented an approach for the development of a more efficient traffic prediction system integrating multiple neural network topologies with CBR and traffic condition settings. The conclusions presented in this paper are primarily based on the optimization results derived from the testing data set. Generalization can only be made by assuming that the testing data set is a truly representative and unbiased sample. To verify this assumption, further testing can be made using a larger testing data set. Also, the approach presented in this paper was extensively examined at one location. Based on the results, the approach can be applied to other locations as well. For locations that exhibit similar traffic conditions during the peak periods, the settings obtained in this study may be transferred directly without retraining. However, testing is recommended with data collected from the other locations prior to testing the transferability. If the testing results are not satisfactory, then the current settings may not be applicable without additional performance optimization at the new location by following the steps described in this study. On the other hand, the developed hybrid prediction system is limited to the point traffic conditions prediction. Further research may be conducted in the direction of studying a travel-time link- or path-based prediction system starting from the findings presented in this thesis.

## REFERENCES

- Abdulhai, B., H. Porwal, and W. Recker. Short-Term Freeway Traffic Flow Prediction Using Genetically Optimized Time-Delay-Based Neural Networks. In *Transportation Research Board 78<sup>th</sup> Annual Meeting* (CD-ROM), Washington D.C., 1999.
- Ahmed, M., and A. Cook. Analysis of Freeway Traffic Time Series Data by Using Box-Jenkins Techniques. In *Transportation Research Record 722*, TRB, National Research Council, Washington, D.C., 1982, pp 1-9.
- Ben-Akiva M., A. de Palma, and I. Kaysi. Dynamic Network Models and Driver Information Systems. *Transportation Research A*, Vol. 25, No. 5, September 1991, pp. 251-266.
- Ben-Akiva, M., E. Kroes, J. de Ruiter, S. Smulders, J. Whittaker, and K. Lindveld. "Real-Time Prediction of Traffic Congestion." In *Vehicle navigation & information systems: conference record of papers*. Institute of Electrical and Electronics Engineers. New York, N.Y., 1992, pp. 557-562.
- Capus, L. and N. Tourigny. Road Safety Analysis: A Case Based Reasoning Approach. In *Transportation Research Board 77th Annual Meeting*(CD-ROM), Washington D.C., 1998.
- Chang, J. L. and S.P. Miaou. Real-Time Prediction of Traffic Flows using Dynamic Generalized Linear Models. *Transportation Research Record 1678*, TRB, National Research Council, Washington, D.C., 1999, pp 168-178.
- Chen, M. and S. Chien. Dynamic Freeway Travel Time Prediction Using Probe Vehicle Data: Link-based vs. Path-based. In *Transportation Research Board 80<sup>th</sup> Annual Meeting* (CD-ROM), Washington D.C., 2001.
- Chien S. and C. M. Kuchipudi. Dynamic Travel Time Prediction with Real-Time and Historical Data. In *Transportation Research Board 81<sup>st</sup> Annual Meeting* (CD-ROM), Washington D.C., 2002.
- Choe, T., A. Skabardonis, and P. Varaiya. Freeway Performance Measurement System (PeMS): An Operational Analysis Tool. In *Transportation Research Board 81<sup>st</sup> Annual Meeting* (CD-ROM), Washington D.C., 2002.
- D'Angelo, M., H. Al-Deek, and M. Wang. Travel Time Prediction for Freeway Corridors. *Transportation Research Record 1676*, TRB, National Research Council, Washington, D.C., 1999, p. 184-191.
- Duda, R., P. Hart, and D. Stork. *Pattern Classification. Second Edition*. John Wiley & Sons, Inc., New York, 2001.

- Fu, L. and L.R. Rilett. Estimation of Time-Dependent, Stochastic Route Travel Times using Artificial Neural Networks. *Transportation Planning and Technology*, Vol. 24, No 1, January 2000, pp. 25-36.
- Gazis, D., and C. Knapp. On Line Estimation of Traffic Densities from Time Series of Traffic and Speed Data. *Transportation Science*, Vol. 5, No. 3, March 1971, pp 283-301.
- Ham, F. M. and I. Kostanic. *Principles of Neurocomputing for Science and Engineering*. McGraw Hill, New York, 2001.
- Hamed, M. M., H.R. Al-Masaeid, and Z.M.B. Said. Short-Term Prediction of Traffic Volume in Urban Arterials. *ASCE Journal of Transportation Engineering*, Vol. 121, No. 3, May/June 1995, pp 249-254.
- Haykin, S. *Neural Networks, A Comprehensive Foundation*. Macmillan College Publishing Company, Inc., 1994.
- Ishak, S. and C. Alecsandru. Maximizing Traffic Prediction Performance of Neural Networks under Various Topological, Input, and Traffic Condition Settings. In *Transportation Research Board 81<sup>st</sup> Annual Meeting (CD-ROM)*, Washington D.C., 2003A.
- Ishak, S. and H. Al-Deek. Performance Evaluation of a Short-Term Time-Series Traffic Prediction Model. *ASCE Journal of Transportation Engineering*, Vol. 128, No. 6, November/December 2002, pp 490-498.
- Ishak, S., P. Kotha, and C. Alecsandru. "Optimization of Dynamic Neural Networks Performance in Short-Term Traffic Prediction." Accepted for publication in the *Transportation Research Record Journal*, 2003B. (in press)
- Iwasaki, M. and K. Shirao. A Short Term Prediction of Traffic Fluctuations Using Pseudo-Traffic Patterns. In *ITS world congress conference proceedings*, 1996.
- Jang, J.S.R., C.T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing*. Prentice Hall, N.J., 1997.
- Jiang, Y. Dynamic Prediction of Traffic Flow and Congestion at Freeway Work Zones. In *Transportation Research Board 81<sup>st</sup> Annual Meeting (CD-ROM)*, Washington D.C., 1999.
- Kaysi, I., M. Ben-Akiva, and H. Koutsopoulos. An Integrated Approach to Vehicle Routing and Congestion Prediction for Real-Time Driver Guidance. *Transportation Research Board Annual 72<sup>nd</sup> Meeting (CD-ROM)*, Washington, D.C., 1993.

- Khattak, A. and A. Kanafani. Case-Based Reasoning: A Planning Tool for Intelligent Transportation Systems. In *Transportation Research, Part C – Emerging Technologies*, Vol. 4(5), pp 267-288, 1996.
- Kolodner, J. *Case-Based Reasoning*. Morgan Kaufmann Publishers Inc., San Mateo, California, 1993.
- Kwon, J., B. Coifman, and P. Bickel. Day-to-Day Travel Time Trends and Travel Time Prediction from Loop Detector Data. In *Transportation Research Board 79<sup>th</sup> Annual Meeting* (CD-ROM), Washington D.C., 2000.
- Lee, S. and D. Fambro. Application of the Subset ARIMA Model for Short-Term Freeway Traffic Volume Forecasting. *Transportation Research Record 1678*, TRB, National Research Council, Washington, D.C., 1999, p. 179-188.
- Lefebvre, Curt, *Neuro Solutions, Version 4.10*. NeuroDimension, Inc., Gainesville, FL, 2001.
- Lint, V., S. P. Hoogendoorn, and H.J. Zuylen. Freeway Travel Time Prediction with State-Space Neural Networks. In *Transportation Research Board 81<sup>st</sup> Annual Meeting* (CD-ROM), Washington D.C., 2002
- Lu Jian. Prediction of Traffic Flow by an Adaptive Prediction System. *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990, pp 54-61.
- Mascavan Der Voort, M., M. Dougherty, and S. Watson. Combining Kohonen Maps with ARIMA Time Series Models to Forecast Traffic Flow. In *Transportation Research part C*, 4(5), pp. 307-318, 1996.
- Nihan, N., and G. Davis. Application of Prediction-Error Minimization and Maximum Likelihood to Estimate Intersection O-D Matrices from Traffic Counts. In *Transportation Science*, Vol. 23, No. 2, pp 77-90, 1989.
- Okutani, I. and Y. Stephanedes. Dynamic Prediction of Traffic Volume through Kalman Filtering Theory. In *Transportation Research B*, Vol. 18B, No. 1, pp 1-11, 1984.
- Pal, S.K., T.S. Dillon, and D. S. Yeung. *Soft Computing in Case Based Reasoning*. Springer-Verlag London Limited, Great Britain, 2001.
- Park, D. and L. Rilett. Forecasting Multiple-Period Freeway Link Travel Times Using Modular Neural Networks. In *Transportation Research Record*. No. 1617 p. 163-170, 1998.
- Park, D., L.R. Rilett, and G. Han. Spectral Basis Neural Networks for Real-Time Travel Time Forecasting. *ASCE Journal of Transportation Engineering*, Vol. 125, No. 6, November/December 1999, pp. 515-523.

- Principe, J. C., N. R. Euliano, and W.C. Lefebvre. *Neural and Adaptive Systems*. John Wiley and Sons, Inc., New York, 2000.
- Rilett, L., D. Park. Direct Forecasting of Freeway Corridor Travel Times Using Spectral Basis Neural Networks. In *Transportation Research Record 1752*, TRB, National Research Council, Washington, D.C., 2001, p. 140-147.
- Sadek, A. W. A Hybrid Simulated Annealing and Case-Based Reasoning Approach for Computationally Intensive Transportation Problems: Rationale and Design Issues. In *Transportation Research Record 1774*, TRB, National Research Council, Washington, D.C., 2001.
- Sadek, A. W., B. L. Smith and M. J. Demetsky. A Prototype Case-Based Reasoning Decision Support System for Real-time Freeway Traffic Routing. In *Transportation Research, Part C*, Vol. 9(5), 2001.
- Sadek, A. W., M. J. Demetsky, and B. L. Smith. Case-Based Reasoning for Real-time Traffic Flow Management. In *Journal of Computer-Aided Civil and Infrastructure Engineering*, Vol 14(5), 1999.
- Stamatiadis, C. and W. Taylor. Travel Time Predictions for Dynamic Route Guidance with a Recursive Adaptive Algorithm. Paper presented at the 73rd TRB Annual Meeting, Washington, D.C., 1994.
- Shin, C-K, and Park, S. C. Towards Integration of Memory Based Learning and Neural Networks. In *Soft Computing in Case Based Reasoning*, chapter 5. Pal, S. K., Dillon, T. S., Yeung, D. S. Springer-Verlag, eds. London, England, 2000
- Watson, I. *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufmann Publishers Inc., San Francisco, California, 1997.
- Williams, B. M. Multivariate Vehicular Traffic Flow Prediction: An Evaluation of ARIMAX Modeling. In *Transportation Research Board 81<sup>st</sup> Annual Meeting* (CD-ROM), Washington D.C., 2002
- Yasdi, R. Prediction of Road Traffic using a Neural Network Approach. In *Neural Computing and Applications*, 8, pp. 135-142, 1999.
- Zhang, H. M. Recursive Prediction of Traffic Conditions with Neural Network Models In *ASCE Journal of Transportation Engineering*, Vol. 126, No. 6, pp. 472-481, 2000.
- Zhang, H., S. G. Ritchie, and Z.P. Lo. Macroscopic Modeling of Freeway Traffic using an Artificial Neural Network. In *Transportation Research Record, No. 1588*, pp 110-119, 1997

## APPENDIX A – AARE-BASED WILCOXON SIGNED RANK TEST RESULTS

Wilcoxon Signed Ranks Test Results (5MIN AARE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	17
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.622	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.000	1.000

Wilcoxon Signed Ranks Test Results (10MIN AARE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	14
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.296	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.001	1.000

Wilcoxon Signed Ranks Test Results (15MIN AARE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	18
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.724	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.000	1.000

Wilcoxon Signed Ranks Test Results (20MIN AARE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	16
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.516	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.000	1.000

## APPENDIX B – RMSE–BASED WILCOXON SIGNED RANK TEST RESULTS

Wilcoxon Signed Ranks Test Results (5MIN RMSE)

Counts of differences (row variable greater than column)

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	0	17
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.622	0.000

Two-sided probabilities using normal approximation

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.000	1.000

Wilcoxon Signed Ranks Test Results (10MIN RMSE)

Counts of differences (row variable greater than column)

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	0	12
NN_CBR	0	0

$Z = (\text{Sum of signed ranks}) / \text{square root}(\text{sum of squared ranks})$

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.059	0.000

Two-sided probabilities using normal approximation

	NN_OPTIMIZED	NN_CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.002	1.000

Wilcoxon Signed Ranks Test Results (15MIN RMSE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	20
NN_CBR	0	0

Z = (Sum of signed ranks)/square root(sum of squared ranks)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.920	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.000	1.000

Wilcoxon Signed Ranks Test Results (20MIN RMSE)

Counts of differences (row variable greater than column)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0	14
NN_CBR	0	0

Z = (Sum of signed ranks)/square root(sum of squared ranks)

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	0.000	
NN_CBR	-3.296	0.000

Two-sided probabilities using normal approximation

	NN OPTIMIZED	NN CBR
NN_OPTIMIZED	1.000	
NN_CBR	0.001	1.000

## APPENDIX C – AARE-BASED FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE RESULTS

Friedman Two-Way Analysis of Variance Results for 38 cases (5MIN AARE).

Variable	Rank Sum
STATIC_NN	109.000
DYNAMIC_NN	109.000
NN_OPTIMIZED	91.000
NN_CBR	71.000

Friedman Test Statistic = 15.537  
 Kendall Coefficient of Concordance = 0.136  
 Probability is 0.001 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (10MIN AARE).

Variable	Rank Sum
STATIC_NN	110.500
DYNAMIC_NN	114.500
NN_OPTIMIZED	87.000
NN_CBR	68.000

Friedman Test Statistic = 22.318  
 Kendall Coefficient of Concordance = 0.196  
 Probability is 0.000 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (15MIN AARE).

Variable	Rank Sum
STATIC_NN	98.000
DYNAMIC_NN	136.000
NN_OPTIMIZED	86.000
NN_CBR	60.000

Friedman Test Statistic = 47.305  
 Kendall Coefficient of Concordance = 0.415  
 Probability is 0.000 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (20MIN AARE).

Variable	Rank Sum
STATIC_NN	106.500
DYNAMIC_NN	113.000
NN_OPTIMIZED	91.000
NN_CBR	69.500

Friedman Test Statistic = 17.724  
 Kendall Coefficient of Concordance = 0.155  
 Probability is 0.001 assuming Chi-square distribution with 3 df

## APPENDIX D – RMSE-BASED FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE RESULTS

Friedman Two-Way Analysis of Variance Results for 38 cases (5MIN RMSE).

Variable	Rank Sum
STATIC_NN	100.500
DYNAMIC_NN	122.000
NN_OPTIMIZED	92.000
NN_CBR	65.500

Friedman Test Statistic = 18.663  
Kendall Coefficient of Concordance = 0.164  
Probability is 0.000 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (10MIN RMSE).

Variable	Rank Sum
STATIC_NN	106.000
DYNAMIC_NN	115.000
NN_OPTIMIZED	89.000
NN_CBR	70.000

Friedman Test Statistic = 23.368  
Kendall Coefficient of Concordance = 0.205  
Probability is 0.000 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (15MIN RMSE).

Variable	Rank Sum
STATIC_NN	91.000
DYNAMIC_NN	123.000
NN_OPTIMIZED	97.000
NN_CBR	69.000

Friedman Test Statistic = 25.871  
Kendall Coefficient of Concordance = 0.227  
Probability is 0.000 assuming Chi-square distribution with 3 df

Friedman Two-Way Analysis of Variance Results for 38 cases (20MIN RMSE).

Variable	Rank Sum
STATIC_NN	102.000
DYNAMIC_NN	121.000
NN_OPTIMIZED	95.000
NN_CBR	62.000

Friedman Test Statistic = 28.642  
Kendall Coefficient of Concordance = 0.251  
Probability is 0.000 assuming Chi-square distribution with 3 df

## **VITA**

Ciprian Alecsandru was born in Birlad, Romania, in 1974. He received the degree of Bachelor of Engineering in Electronics for Transportation Engineering from University "Politehnica" of Bucharest, Romania, in 1997, with distinction. He followed another year of specialization in Telematics for Transportation at the same university. Four years before he joined the graduate program in Civil Engineering at Louisiana State University in Fall 2001, he worked as an teaching assistant at the Transportation Faculty at University "Politehnica" of Bucharest, Romania. He expects to receive the degree of Master of Science in Civil Engineering in August 2003 and continues within the doctoral program at the same department.