Estimation of Annual Average Daily Traffic Volumes using Neural Networks

Mario Adamo

A Thesis Submitted to the Department of Mathematics and Computer Science in partial fulfillment of the requirements for the Degree of Specialized Bachelor of Science in Computer Science.

Laurentian University

Supervised By: Professor Pawan J. Lingras PhD

Department of Mathematics and Computer Science
Faculty of Science
Laurentian University
April 1994

Abstract

This study compared the estimations of annual average daily traffic (AADT) volumes using the conventional method(Factors), multiple regression analysis, and the neural network approach. All three approaches were compared using three different classification schemes as well as different duration of traffic counts. The neural network and multiple regression approaches consistently performed better than the conventional approach, and the neural network approach in many cases slightly outperformed the multiple regression approach. Apart from providing a good modeling tool for estimating AADT, the results also provide useful insight into the duration of the short term traffic counts and the classification schemes for the highway sites.

Acknowledgments

To my wonderful family, my wife Johanna and daughter Amanda, who have endured the pain and sacrifices throughout this project and beyond. To my mother-in-law and father-in-law, Tuire and Angelo Gaccione, for their love and support. Also I would like to thank my mother, Michelina Adamo for her unconditional love.

I would like to thank my advisor, mentor, and most importantly, friend, Dr. Pawan Lingras. He has been an inspiration to me as a computer scientist. I would like to thank him for his time and efforts which go beyond the scope of generosity and this project. Financial support from Natural Sciences and Engineering Research Council(NSERC) of Canada is much appreciated. The traffic data obtained from Alberta transportation was essential for the experiments in this project. I want to thank Alberta Transportation and Dr. S. C. Sharma, Professor, Faculty of Engg., University of Regina for supplying the data used for the experiments.

Lastly I would like to thank Algoma Steel Inc. for supplying the use of their mainframe. I couldn't of done the multiple regression analysis for the 168 hour counts without it.

Table of Contents

List of Figures

List of Tables

List of Graphs

Chapter 1

Introduction

Highway agencies collect traffic volume data from various seasonal and permanent traffic counters over a number of years. Since the installation of a permanent traffic counter (PTC) on every road section is not economically feasible, highway agencies routinely use sample traffic counts. The sample traffic counts are obtained using seasonal traffic counters (STCs). The data obtained from seasonal traffic counts is routinely used to estimate important traffic parameters for the overall highway network.

The present study deals with the estimation of an important traffic parameter called *annual average daily traffic* (AADT). The AADT provides a measure of overall utilization of the highway facility. The results obtained from the conventional methods are compared with those obtained from the neural networks and multiple regression analysis. The estimation of AADT in conventional methods is typically done as follows. PTC sites are grouped together into similar patterns of temporal volume variations and road classes according to driver population such as commuter, long distance, and recreational (DiRenzo, et al. 1985; Sharma et al. 1986). Average traffic factors are determined for each PTC or road group. These factors are then used in estimation of required parameters from sample counts (Sharma and Allipuram, 1993).

Classification and estimation of parameters using inductive learning techniques is one of the major functions of neural networks. Recent developments in neurocomputing (Hecht-Nielsen, 1990) are making it possible for relative novices to employ neural networks in their analysis as a substitute for more elaborate statistical procedures. This study uses one of the frequently used neural networks called *multi-layer feedforward backpropagation network.* Supervised learning is achieved by a well established learning technique called the *generalized delta rule*.

Today the estimation of AADT volumes is done using a technique called factors. This approach uses average factors for estimating AADT. An average factor is computed for each of annual classification groups. STC data is then classified into a one of the groups, and the factor for that group is multiplied with the average daily traffic for that sample period to obtain an estimated AADT.

Estimation of AADT volumes can also be done by using multiple regression analysis. The AADT is predicted by a linear combination of independent variables. In this case the independent variables are represented by hourly counts. For 24, 48, and 168 hour counts, 24, 48, and 168 independent variables respectively will be used to represent the independent variables(regressor variables). For our multiple regression models, the principle of least squares will be used to produce estimates of AADT volumes that are the best linear unbiased estimates under classical statistical assumptions.

For the purpose of training and testing the three models, the PTC sites are grouped together based on similar traffic patterns. The short term classification is done using the Kohonen neural networks to establish five different road classes. A few PTC sites from each group are used for testing and short term counts similar to the STCs are extracted. The PTC groups excluding the test PTC sites are used in the development of factors, multiple regression constants, and neural networks to estimate AADT. Since the classification of highway sites plays an important role in the AADT estimation, actual development of factors, multiple regression constants, and neural networks are carried out using different classification schemes. On the one end of these classification schemes, there is the true classification established using the complete annual traffic data. However, the seasonal traffic counters do not provide the complete annual data necessary for the true classification. Hence, on the other end, all the PTCs are grouped in one class and the three models are developed for the single class. In between these two extreme classification schemes, there are classification schemes based on the traffic patterns collected during the short term counts. The models are developed and tested for 24 hour, 48 hour and week long counts.

This document will begin with a review of literature. This chapter will provide a review of the three estimation models, as well as the various classification schemes. Next, an explanation of how the experimental data was classified, and the results of that classification will be presented in the study data and classification schemes chapter. Following that chapter, the details of the three estimation models used in the study will be explained in the description of models chapter. Next, the results of the experiments will be explicitly shown and explained in the results and analysis chapter. Lastly, the summary and conclusions will be provided in chapter 6.

Chapter 2

Review of Literature

This section reviews the conventional procedure for estimation of AADT. A brief review of various classification schemes, neural networks and multiple regression analysis is also provided. Annual Average Daily Traffic (AADT) volumes are calculated by dividing annual traffic volume by the number of days in the year. The estimation of AADT is done in two stages:

1. Classification of PTC and STC sites into different road classes.
2. Development of factors for estimation of AADT for each road class.

2.1 Road Classification

Different highway sections in a given highway system have different traffic stream characteristics. Highway sections with similar traffic characteristics can be grouped together to simplify the analysis. In Canada, there are usually 30 to 60 PTC sites in a province, which are located throughout the provincial highway system so that continuous data on the traffic patterns and characteristics of all classes of highways are collected (Garber and Hoel, 1988). Grouping of PTC sites into similar seasonal traffic patterns is required to establish various types of road classes. These road classes are then used in the development of average expansion factors to estimate parameters such as AADT from sample counts.

In a very commonly used classification system, roads are classified on the basis of trip purpose and trip length characteristics (Sharma *et al.*, 1986); examples of resulting classes are *commuter, business, long distance*, and *recreational*. Such a classification simplifies the analysis, because instead of analyzing individual highway sections it is possible to consider a fewer number of classes. Trip purpose provides information about the road users. It is one of the important criteria in a variety of traffic engineering analyses. Trip purpose can be determined by obtaining the information directly from the road users. Unfortunately, it is not possible to obtain such information from all the road users using all the highway facilities. Therefore, in order to classify roads, traffic engineers study various traffic patterns obtained from PTC sites and sample surveys of a few road users. Sharma and Werner (1981) used monthly factors which are defined as the ratios of average daily traffic (ADT) for a month to AADT, to classify PTC sites based on hierarchical grouping (a clustering technique) and Scheffe's *S*-method of multiple group comparison. The classification technique such as hierarchical grouping is applied to the PTC data to establish the road classes based on complete traffic patterns. The STC sites only provide short term counts. The STC sites are assigned to one of the road classes using the incomplete traffic patterns obtained from short term counts using measures such as *least mean-squared error*.

Recently, Lingras (1995) used the Kohonen Neural network for classification of traffic patterns. The results of classification using the Kohonen neural network are shown to be similar to the hierarchical grouping method. The Kohonen networks are computationally more desirable for a large number of patterns and can also be used to classify incomplete patterns obtained from STCs. The following subsections were adapted from (Lingras, 1995).

2.1.1 Classification of Highway Sections Using Statistical Methods

The seasonal and permanent traffic counters scattered across the highway network are the major sources of traffic data. These traffic counters provide the *traffic volume* -- the number of vehicles that have passed through a particular section of a lane or highway in a given time period. Traffic volumes can be expressed in terms of hourly or daily traffic. More sophisticated traffic counters provide additional information such as the speed, length and weight of the vehicle. Highway agencies generally have records from traffic counters over a number of years. In addition to the data obtained from traffic counters, traffic engineers also conduct occasional surveys of road users to get more information. In Canada, there are usually 30 to 60 PTC sites in a province, which are located throughout the provincial highway system, so that continuous data on the traffic patterns and characteristics of all classes of highways are collected (Garber and Hoel, 1988, Sharma and Allipuram, 1993).

The PTC sites are grouped together to establish various types of road classes. In a commonly-used classification system, roads are classified on the basis of trip purpose and trip length characteristics (Sharma and Werner, 1981); examples of resulting classes are commuter, business, long distance, and recreational. Trip purpose provides information about the road users. It is an important criterion in a variety of traffic engineering analyses. Trip purpose information can be obtained directly from the road users, but since all users cannot be surveyed, traffic engineers study various traffic patterns obtained from seasonal and permanent traffic counters and sample surveys of a few road users. Some of the important traffic patterns are as follows:

- Hourly traffic pattern: Variation of hourly traffic volume in a given day.
- Daily traffic pattern: Variation of daily traffic volume in a given week.
- Monthly traffic pattern: Variation of monthly average daily traffic volume in a given year.

The classification is done using different statistical procedures for grouping similar objects. Sharma and Werner (1981) used monthly traffic patterns to classify PTC sites based on hierarchical grouping (a clustering technique) and Scheffe's S-method of multiple group comparison.

2.1.2 Hierarchical Grouping

The hierarchical grouping technique is used to group a set of k-dimensional vectors. Each vector represents a pattern. For example, a monthly traffic pattern is a twelve-dimensional vector (Sharma and Werner, 1981). The ratio of monthly average daily traffic (MADT) over the annual average daily traffic (AADT) is used as the component for each month in the monthly traffic pattern. The process of grouping n patterns starts with n groups -- one group for each pattern. By selecting two groups which produce the least amount of *within-group* error, the number of groups is reduced by one to n-1. The within-group error is calculated by summing up the difference $E(\mathbf{x},\mathbf{y})$ between all pairs of vectors $(\mathbf{x},\mathbf{y})$ in that group given by:

$$E(\mathbf{x},\mathbf{y}) = \frac{\sum_{i=1}^{k}(x_i - y_i)^2}{k}, \qquad (1)$$

where x_i and y_i are i^{th} components of the vectors $\mathbf{x}$ and $\mathbf{y}$, respectively. The remaining n-1 groups are further reduced by one by combining two groups, such that the within-group error is minimum. This process of grouping continues until the number of groups is reduced to one.

The errors associated with the successive stages of the grouping process represent the marginal cost of reducing the number of groups by one. The error associated at a particular stage is greater than or equal to the error associated with the previous stage of grouping. Hierarchical grouping does not specify the optimum number of groups. However, the graph of error associated with successive stages of grouping usually reveals a *knee-of-curve* -- that part of the graph where the error starts increasing at rapid rate.

The PTC groups are used to develop guidelines for construction, maintenance and upgrading of the highway sections. In order to apply these guidelines to all the highway sections, it is necessary to assign these highway sections to one of the PTC groups. Highway agencies attempt to use STC programs to classify all the highway sections in their jurisdiction under one of the PTC groups. Such a classification usually involves assigning incomplete patterns to the existing groups of patterns. The next section describes a statistical approach for such an assignment.

2.1.3 Classification of Incomplete Patterns

The schedule of STC programs suggested in the literature and practised by highway agencies is diverse (Albright, 1991). Moreover, there is a lack of systematic method to determine the PTC group that best fits the traffic pattern at each STC location. Recently, Sharma and Allipuram (1993) used the least mean square error to assign incomplete traffic patterns to the existing groups. Each group of patterns is represented by a vector **g**. Each component g_i of **g** is given by:

$$g_i = \frac{\sum_{\text{for all } \mathbf{x} \text{ in the group}} x_i}{\text{number of vectors in the group}}. \tag{2}$$

The difference $E'(\mathbf{g},\mathbf{z})$ between a group vector **g** and an incomplete pattern **z** is given as:

$$E'(\mathbf{g},\mathbf{z}) = \frac{\sum_{1 \le i \le k,\ z_i \text{ is known}} (g_i - z_i)^2}{\text{number of available components of } \mathbf{z}} \tag{3}$$

If the pattern **z** is complete, $E'(\mathbf{g},\mathbf{z}) = E(\mathbf{g},\mathbf{z})$. Hence, the measure E' is a generalization of the measure E. An STC site with traffic pattern **z** is assigned to that PTC group for which the difference $E'(\mathbf{g},\mathbf{z})$ is minimum, where **g** is the traffic pattern for the group.

The existing approaches (Sharma and Allipuram, 1993; Sharma and Werner, 1981) are generally used with the monthly traffic variation. Day-to-day and hour-to-hour variations can also play an important role in assigning STC sites to one of the PTC groups. Proper representation of daily and hourly variations in the existing statistical procedure is difficult due to a large number of daily and hourly patterns.

The next section explains how neural networks may be used to group patterns. A combination of the hierarchical grouping technique and neural networks may make it possible to group a large number of daily and hourly patterns.

2.1.4 The Kohonen Neural Network (for Classification)

Neural networks are used successfully in a wide variety of applications including investment, medicine, science, engineering, marketing, manufacturing, and management (Lawrence, 1993). Neural networks learn from experience (using inductive learning) and not from programming. Neural networks are good at recognizing patterns, generalizing, and predicting trends. They are fast and tolerant of imperfect data, and do not need formulae or rules from the experts in the application domain.

Researchers have proposed different types of neural networks for solving a variety of problems (Hecht-Nielsen, 1990; Lawrence, 1993; Zahedi, 1990). In its most general form, a neural network consists of several neurons. Each neuron receives inputs from other neurons and (optionally) from the external environment and produces an output.

There are two different stages in the development of a neural network model: training and testing. During the training stage, the network uses inductive learning principle to learn from a set of examples called the training set. The learning process for the Kohonen Neural Network is *unsupervised.*. In the unsupervised learning, the desired output from the neurons is not known. The network attempts to classify patterns from the training set into different groups. The Kohonen rule (Kohonen, 1988) is used for unsupervised learning which is an example of a learning equation

Since the actual classification is not known, an unsupervised learning model may be more suitable for the traffic pattern classifications. Researchers have proposed two different types of neural networks for unsupervised learning (Haykin, 1994). The unsupervised learning proposed in Linsker's model (Linsker, 1986) can be used for networks with several output neurons with non-zero outputs. The unsupervised learning using the Kohonen rule (Kohonen, 1988) uses competitive learning approach. In competitive learning, the output neurons compete with each other. The winner output neuron has the output of 1, the rest of the output neurons have outputs of 0. The competitive learning is suitable for classifying a given pattern into exactly one of the mutually exclusive classes. The behaviour of such a network is similar to the hierarchical grouping discussed earlier. Hence, this study uses the Kohonen network which is based on the competitive learning approach. Fig. 2.1 shows an example of the Kohonen network.

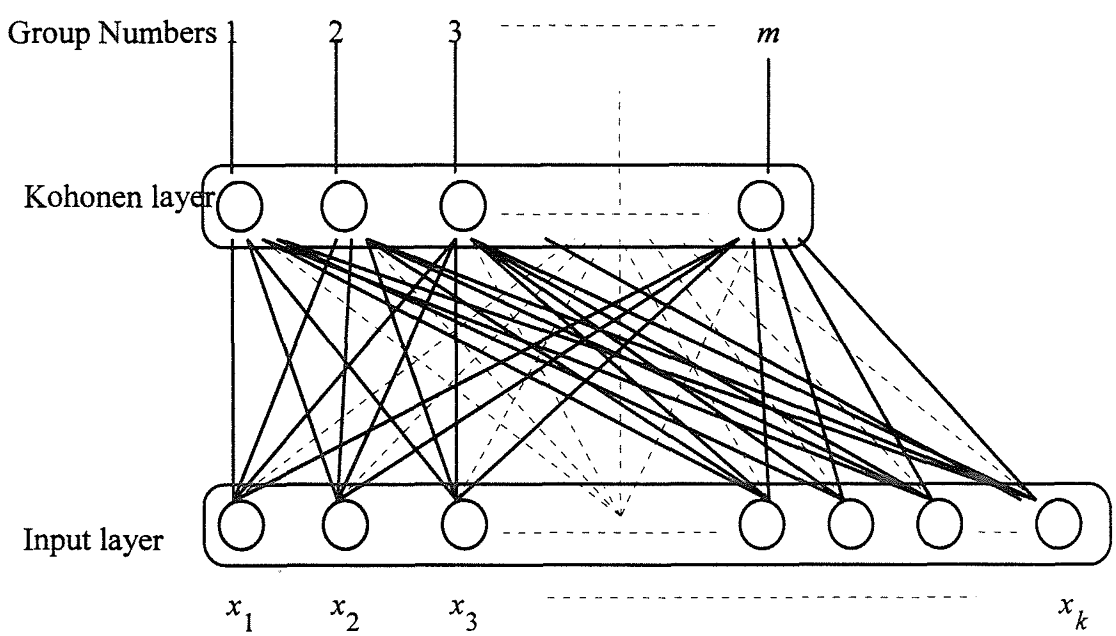

FIG. 2.1 The Kohonen Neural Network

The network is used to group patterns represented by k-dimensional vectors into m groups. The network consists of two layers. The first layer is called the input layer and the second layer is called the Kohonen layer -- named after the scientist who developed the network. The network receives the input vector for a given pattern. If the pattern belongs to the i^{th} group, then i^{th} neuron in the Kohonen layer has a output value of one and other Kohonen layer neurons have output values of zero. Each neuron in the Kohonen layer is connected to the input layer as shown in Fig. 2.2.

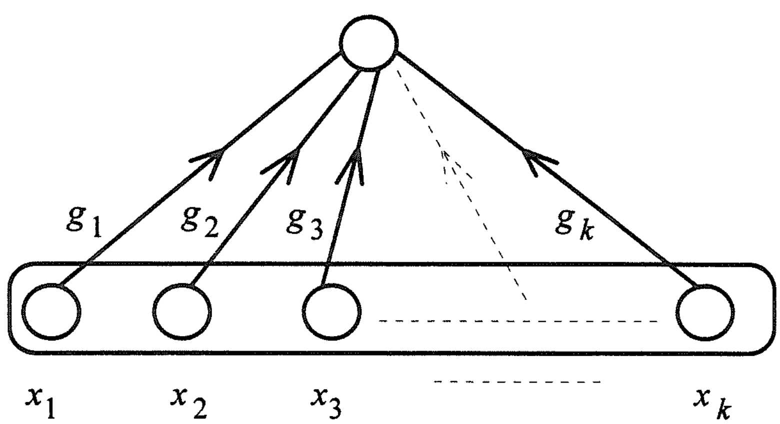

FIG. 2.2 Connections between a Kohonen Layer Neuron and Input Layer

Each connection is assigned a weight g_i. Weights of all the connections to a Kohonen layer neuron make up a k-dimensional weight vector **g**. The weight vector **g** for a Kohonen layer neuron is the vector representation of the group corresponding to that neuron. For any input vector **z**, the network compares the input with the weight vector for a group using the measure such as $E'(\mathbf{g}, \mathbf{z})$. The pattern **z** belongs to the group with minimum value for $E'(\mathbf{g}, \mathbf{z})$. The grouping of patterns in the Kohonen rule is similar to the classification of patterns (possibly incomplete) used by Sharma and Allipuram (1993). However, Sharma and Allipuram's approach uses the existing groups created using hierarchical grouping. The Kohonen neural network, on the other hand, generates the groups through a learning process as follows: Initially, the network connections are assigned somewhat arbitrary weights. The training set of input vectors is presented to the network several times. For each iteration the weight vector **g** for a group that is closest to the pattern **z** is modified using the equation:

$$\mathbf{g}_{new} = \mathbf{g}_{old} + \alpha(t) \times \mathbf{z}, \qquad (4)$$

where $\alpha(t)$ is a learning factor which starts with a high value at the beginning of the training process and is gradually reduced as a function of time.

Neural networks can be used to substitute the statistical techniques for grouping of traffic patterns. Comparing grouping of traffic patterns using the hierarchical grouping method and the Kohonen neural network, it has been shown(Lingras 95) that the Kohonen neural network can be used to approximate the hierarchical grouping technique. The Kohonen neural network also provides an ability to classify incomplete patterns which is similar to the existing least mean square approach. Hence, one can say that the Kohonen neural network integrates the hierarchical grouping of complete patterns and the least mean square approach for classifying incomplete pattern. It may be advantageous to use hierarchical grouping on a small subset of typical traffic patterns to determine the appropriate number of groups and the initial weights for the neural network. The neural network can then continuously group new and larger sets of patterns and change its parameters to reflect the changing traffic patterns. Such an approach may be useful in using hour-to-hour and day-to-day traffic variations in addition to the monthly traffic volume variations in classifying highway sections.

2.2 Neural Networks (for AADT estimation)

Neural networks are good at recognizing patterns, generalizing, and predicting trends. They are fast and tolerant of imperfect data, and do not need formulae or rules. Researchers have proposed different types of neural networks for solving a variety of problems (Hecht-Nielsen, 1990; Lawrence, 1993; Zahedi, 1990). In its most general form, a neural network consists of several layers of neurons. Each neuron receives input from other neurons and external environment and produces output. The output from a neuron can be sent to the external environment or to other neurons in the network.

This study used a slightly restricted type of neural networks based on

multi-layered, feed-forward, and backpropagation design for unsupervised and supervised learning. As shown in Fig. 2.3, the networks used in the study consist of one input layer, one output layer and one hidden layer of neurons. The input layer neurons accept input from the external environment. The output from input layer neurons is fed to the hidden layer neurons. The hidden layer neurons feed their output to the output layer neurons which send their output to the external environment. Neurons from each layer feed the output only to the next layer and hence the network is called *feed forward.*

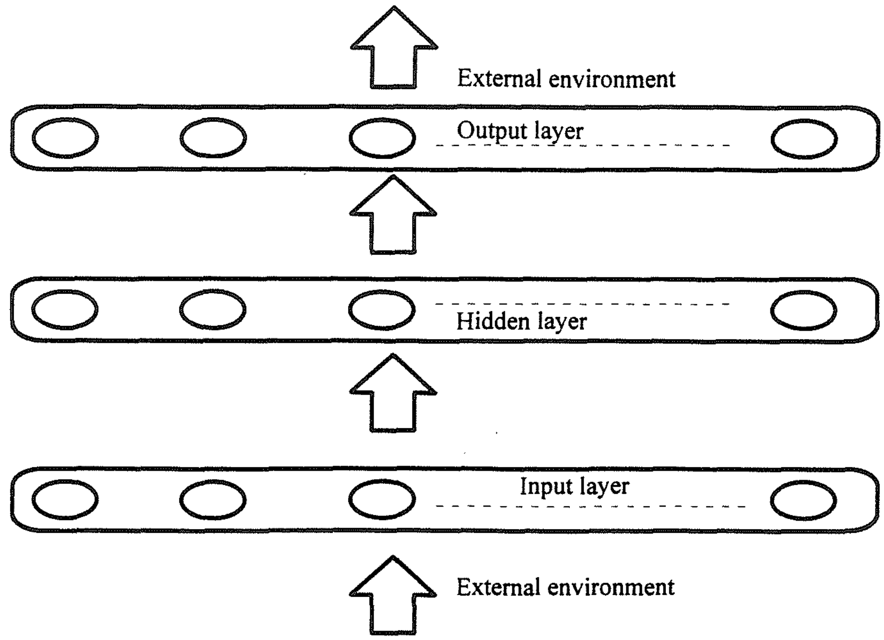

Fig. 2.3 Multi-layered, Feed-forward Neural Network

The input and output of a neuron are governed by certain mathematical equations. Fig. 2.4 is used to illustrate the input and output equations used in this study. It is assumed that $neuron_i$ in a given layer is connected to all the neurons $(neuron_1, neuron_2, \ldots, neuron_n)$ in the previous layer. The connection from $neuron_j$ to $neuron_i$ has the weight w_{ji}. The weights of the connections are initially assigned an arbitrary value between 0 and 1. The appropriate weights are determined during the training phase. Input to the $neuron_i$ is obtained using the following equation:

$$input_i = \sum_{j=1}^{n} w_{ji} \times output_j \qquad (5)$$

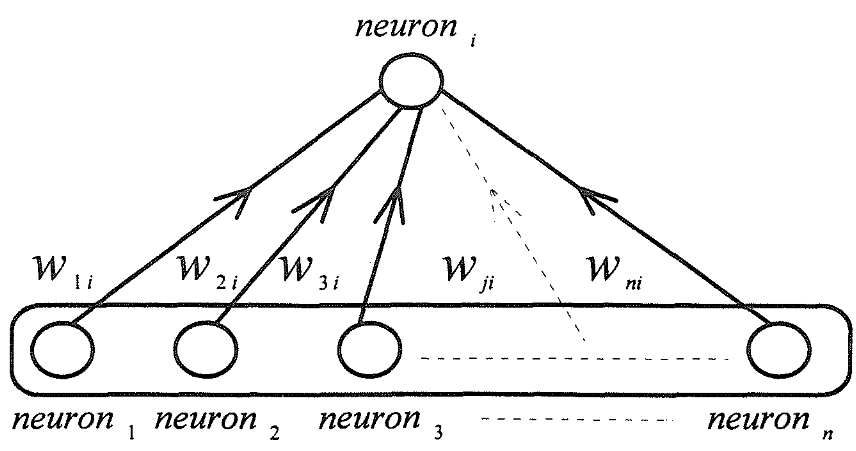

Figure 2.4 Illustration of feed forward and backward propagation

Output from the *neuron*$_i$ is calculated using the sigmoid transfer function as:

$$output_i = f(input_i) = \frac{1}{1 + e^{gain \times input_i}}, \tag{6}$$

where *gain* is a system parameter determined by the system designer to specify the slope of the sigmoid function around input value of zero. In the proposed network, a value of 2 is used as the *gain*. There are several other functions for determining the output from a neuron. The sigmoid transfer function is chosen because it produces a continuous value in the 0 to 1 range.

There are two different stages in development of neural network model, training and testing. During the training stage the network uses inductive learning principle to learn from a set of examples called the training set. The learning process can be *unsupervised* or *supervised*. In unsupervised learning, the network attempts to classify examples from the training set into different groups based on input patterns. In supervised learning, the desired output from output layer neurons for the examples in the training set is known. The network attempts to adjust weights of connections between neurons to produce the desired output. During this process, the error in the output is propagated back from one layer to the previous layer for adjusting weights of the connections, i.e. the network uses *backpropagation* method for propagating the error.

The weights of the connections are modified iteratively. The network is presented with the training set repeatedly and is allowed to change weights after one (or more) iteration(s). The weights are modified using a learning equation. This study uses two of the most popular learning equations, the generalized delta rule for supervised learning, and the Kohonen rule for unsupervised learning.

The generalized delta rule is a variation of the delta rule. The delta rule was developed for engineering applications. The objective of the delta rule is to minimize the sum of the squared errors. The $error_i$ for $neuron_i$ is given by

$$error_i = output_i - desired\_output_i, \qquad (7)$$

where $output_i$ is the actual output and $desired\_output_i$ is the desired output for $neuron_i$. The generalized delta rule is used for non-linear transfer functions such as the sigmoid transfer function used in this study. The generalized delta rule calculates the weights using the following equation:

$$w_{ji}^{new} = w_{ji}^{old} + \alpha \times output_j \times error_i \times f'(input_i), \qquad (8)$$

where $f'(input_i)$ is the derivative of the transfer function evaluated at $input_i$ and α is the learning parameter which represents the speed of learning. For the sigmoid transfer function used in this study,

$$f'(input_i) = input_i \times (1 - input_i) \qquad (9)$$

In the testing stage, the network is tested for another set of examples for which the output from the output layer neurons is known. After the neural net model is tested successfully, it is used for predictions.

2.3 Multiple Regression Analysis

Multiple regression analysis is a technique that employs several independent variables to predict the value of a dependent variable; hence, each of these predictor variables explains part of the total variation of the dependent variable. The independent variables are represented by the hourly counts, where as the dependent variable is represented by the actual AADT volumes.

Multiple regression analysis allows us to exercise statistical control and to determine the influence of any X(independent variable) on Y(dependent variable). The techniques of multiple regression are straightforward extensions of those of simple regression. This study uses three multiple regression models which vary only in the amount of independent variables. 24, 48, and 168 independent variables are used for the 24, 48, and 168 hourly counts respectively. It is very difficult to illustrate hyperplanes of these degrees. Therefore for the purpose of simplicity, multiple regression will be explained in the following section using two independent variables. The following subsections were adapted from Kohler 88.

2.3.1 Multiple Regression (2 independent variables)

Let us consider the case in which one dependent variable, Y, is related in linear fashion, to two independent variables, X_1 and X_2. (Y may be actual AADT, X_1 and X_2

may be two hourly traffic volumes). The first goal of the analysis of such a case is the establishment of an estimated multiple-regression equation, such as

$$Y' = a + b_1X_1 + b_2X_2. \tag{10}$$

This equation gives us the estimated value, Y', of the dependent variable for any specified pair of values of the independent variables. There exists three estimated regression coefficients, a, b_1 and b_2. Their meaning is easy to comprehend: a is the estimated value of Y when $X_1 = X_2 = 0$; b_1 gives us the change in Y(also referred to as the partial change or net change in Y) associated with a unit change in X_1 when X_2 is held constant; while b_2, similarly, equals the change in Y associated with a unit change in X_2 when X_1 is held constant. The values of b_1 and b_2 are called the estimated partial-regression coefficients; they are, in fact, the partial derivatives of Y with respect to either X_1 or X_2. This multiple regression equation corresponds to a plane in three-dimensional space. The multiple regression equations used in this study correspond to hyper-planes of degree 24, 48, and 168.

2.3.2 The Regression Plane

In our three-variable case, three observations are made for each sample unit: one for the value of Y, one for X_1, and one for X_2. These observations are depicted in Figure 2.5.

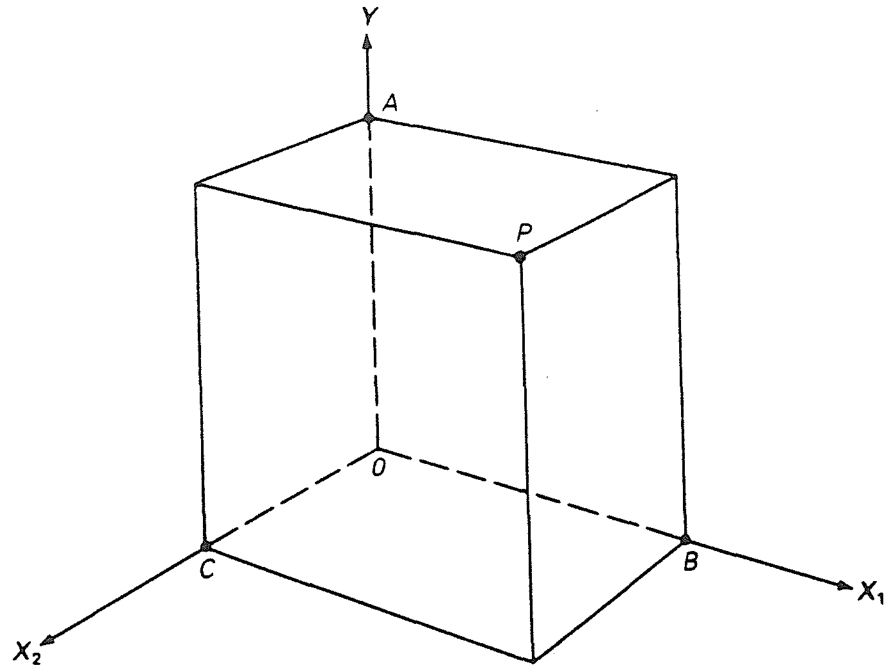

Figure 2.5 Sample Point in three-dimensional Scatter Diagram

If there are many sample observations, such as those depicted in figure 2.6, a plane will be represented. The three-variable multiple regression technique estimates the multiple regression equation (10) in such a way that all the estimates derived from it fall on a surface, such as the shaded area ABCD in our graph(figure 2.6), that is called the regression plane and that is positioned among the sample points in such a way as to minimize the sum of the squared vertical deviations between these sample points and their associated estimates.

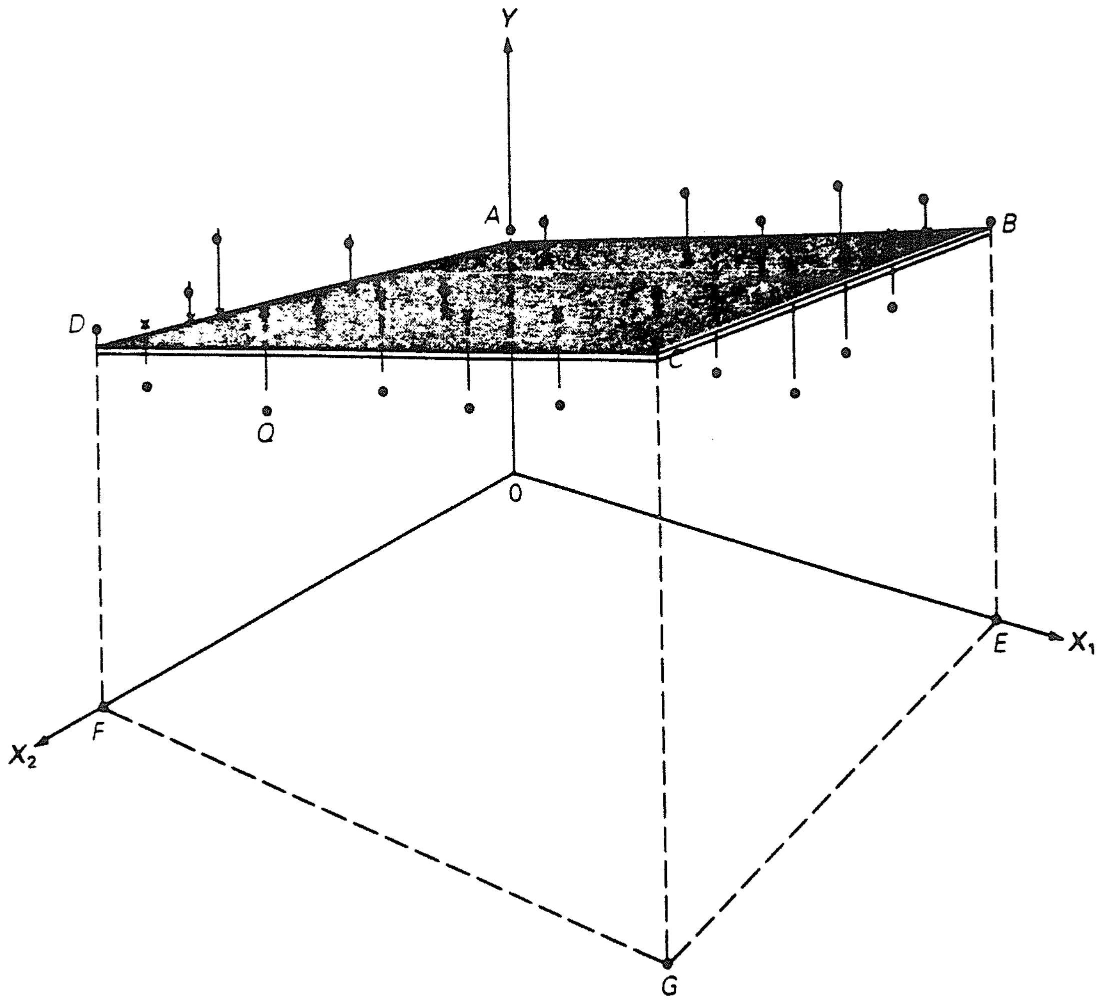

Figure 2.6 The Regression Plane

In figure 2.6, the observed sample points(or actual values of Y for any given combination of X_1 and X_2) are represented by dots which are paired with x's by a line. The associated estimates, Y', are represented by the x's lying on the shaded plane. These crosses, of course, are positioned immediately below or above the paired dots, depending on whether the latter are suspended above or below the regression plane.

We can also note the values of the regression coefficients in figure 2.6. The value of a estimates Y for $X_1=X_2=0$; hence, it is the Y intercept, equal to distance 0A in the graph. The value of b_1 is the slope of the regression plane when holding X_2 constant(at any desired level). Imagine cutting the regression plane parallel to the X_1 axis at $X_2=0$. The cut would trace line AB in the YX_1 plane and show the value of Y rising from 0A at $X_1=0$ to EB at $X_1=0E$. The slope of line AB, relative to 0E, equals b_1. Imagine instead cutting the regression plane parallel to the X_1 axis at $X_2=0F$. The cut would trace line DC and show the value of Y rising form FD at $X_1=0$ to GC at $X_1=0E$. The slope of line DC, relative to FG, also equals b_1. Thus b_1 always indicates how Y changes with X_1, while not changing X_2.

Finally b_2, is the slope of the regression plane when holding X_1 constant(again at any desired level). Imagine cutting the regression plane parallel to the X_2 axis at $X_1=0$. That cut would trace line AD in the YX_2 plane and show the value of Y rising(even though slightly) from 0A at $X_2=0$ to FD at $X_2=0F$. The slope of line AD, relative to 0F, equals b_2. Imagine instead cutting the regression plane parallel to the X_2 axis at $X_1=0E$. The cut would trace line BC and show the value of Y rising from EB at $X_2=0$ to GC at $X_2=0F$. The slope of line BC, relative to EG, also equals b_2. Thus, b_2 always indicates how Y changes with X_2, while not changing X_1.

2.3.3 Multiple Regression with three or more independent variables.

When simple regression analysis turns into multiple regression analysis, hand calculations become quit burdensome. The principles involved remain the same as the number of variable rises, but the calculations, luckily can nowadays be performed by computer.

When three independent variables are involved, the computer finds an estimated least-squares regression of the form,

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3, \tag{11}$$

and this procedure can be extended without end when additional predictor variables are included in the analysis(by adding b_4X_4, b_5X_5, and so on to the equation).

The moment the analysis involves four variables(one dependent and three independent ones), the data can no longer be pictured in a scatter diagram. The human mind can envision up to three dimensions, but it balks at four dimensions and more.

Chapter 3

Study Data and Classification Schemes

PTC data collected in the years 1985 to 1989 and 1991 was used in the experiments. The traffic pattern for each year at each PTC was treated as a separate annual traffic pattern. This resulted in 264 annual traffic patterns. In the first classification scheme, the true classification was established using the complete annual traffic data. All the available annual traffic patterns were grouped using the Kohonen Neural networks based on twelve monthly factors (Sharma and Werner, 1981) into five groups. A monthly factor is given by:

$$\text{monthly factor} = \frac{\text{MADT}}{\text{AADT}}, \tag{12}$$

where

MADT = monthly average daily traffic volume.

Second column in table 3.1 shows the number of annual patterns in each of the five groups. The annual patterns were then separated into train and test sets by randomly choosing a few annual patterns as test patterns. This division resulted in 211 training patterns and 53 test patterns. Columns 3 and 4 in table 3.1 show the number of training and test patterns for each of the groups.

Group	Total	Train	Test
0	9	7	2
1	117	94	23
2	7	5	2
3	18	13	5
4	88	71	17
Total	264	211	53

Table 3.1 Number of Annual Patterns

The short term Seasonal Traffic Counts (STCs) are carried out at different times and they also vary in terms of the duration of the count. The schedule of STC programs suggested in the literature and practised by highway agencies is diverse (Albright, 1991). In order to compare results for diverse STC programs, this study used different time periods and duration of short term seasonal counts. The preliminary experiments indicated that the seasonal counts in July generally provide better estimations than any other month. Moreover, many of the short term counts generally take place during summer months. Three time durations, namely one day (24 hours), two days (48 hours) and week long (168 hours) starting on a Monday in July were chosen to limit the number of possible combinations. Tables 3.2 shows the number of short term traffic patterns used for training and testing for the five groups obtained from classification based on annual patterns as well as for the combined class.

	24 Hour Patterns		48 Hour Patterns		168 Hour Patterns	
Group	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
0	36	9	34	9	28	7
1	524	56	501	52	426	45
2	28	5	27	4	23	4
3	83	15	81	14	67	12
4	408	44	390	42	330	37
Total	1079	129	1033	121	874	105

Table 3.2 Number of Patterns
Grouping Based on Annual Patterns

The seasonal traffic counters do not provide the complete annual data necessary for the true classification. Hence, on the other end, all the PTCs are grouped in one class and the factors and neural network models are developed for the single class. The numbers corresponding to the combined class are shown in table 3.3.

	24 Hour Patterns		48 Hour Patterns		168 Hour Patterns	
	Train	Test	Train	Test	Train	Test
Total	973	235	926	228	784	195

Table 3.3 Number of Patterns
Non-Grouping

In practice, it is not possible to classify the STCs based on the annual data. However, combining them in one single group does not provide good estimations. Traffic engineers use their subjective judgment to classify the STCs into one of the traffic classes established by classification of PTCs. Such a subjective judgment may or may not be based on the available short term traffic count data. This study used an objective classification of the short term traffic patterns obtained from the training set of annual patterns using the Kohonen neural networks. The Kohonen network thus obtained was used to classify the short term traffic patterns obtained from the test set of annual patterns. Table 3.4 shows the number of short term traffic patterns in each of the five groups.

	24 Hour Patterns		48 Hour Patterns		168 Hour Patterns	
Group	Train	Test	Train	Test	Train	Test
0	117	20	45	7	164	41
1	130	28	242	72	105	27
2	210	36	203	36	168	14
3	221	60	96	15	290	81
4	295	91	340	98	157	32
Total	973	235	926	228	784	195

Table 3.4. Number of Patterns
Grouping Based on Short Term Patterns

Chapter 4

Description of Models

This section describes the details of neural network models, multiple regression models and the conventional factor-based model used in the study.

4.1 Factor Model

The conventional approach uses average factors for estimating AADT. The average factor is defined as:

$$\text{factor} = \frac{\sum_{i=1}^{n} \frac{\text{AADT}}{\text{SADT}}}{n}, \qquad (13)$$

where:

SADT = Average Daily traffic volume recorded during the short term count,
AADT = actual AADT,
n = number of short term patterns.

The factors thus calculated are used in the estimation of AADT for an STC site as follows. First, the STC site is classified into one of the five groups. The SADT, computed from the STC data is used along with the factor of the group to estimate AADT as:

$$\text{estimated AADT} = \text{SADT} \times \text{factor}. \qquad (14)$$

4.2 Neural Network Model

The neural network approach uses three network designs for the 24-hour, 48-hour and 168-hour short term patterns. Each network design consists of three layers of neurons, an input layer, a hidden layer and an output layer. The number of neurons in each layer for the three designs are shown in table 4.1.

Duration	Input	Hidden	Output
24 hours	24	13	1
48 hours	48	24	1
168 hours	168	85	1

Table 4.1 Number of Neurons in a Layer

The input pattern consists of the ratios of

$$\frac{\text{hourly traffic volume}}{\text{SADT}} \qquad (15)$$

for each hour in the short term count. The output of the neural network consists of the ratio

$$\frac{\text{AADT}}{\text{SADT}}. \qquad (16)$$

For each training set described in section 3, a neural network is trained using the appropriate neural network design.

4.3 Multiple Regression Model

The Multiple Regression approach also incorporated three designs for the 24-hour, 48-hour and 168-hour short term patterns. The number of independent variables were directly proportional to the degree of the hourly count(i.e. 24, 48, 168 independent variables were used for the 24-hour, 48-hour and 168-hour short term patterns respectively) . As in the neural network model the input patterns consisted of the ratios shown in equation 15, for each hour in the short term count. The dependent variable of analysis consisted of the ratio shown in equation 16.

For each training set described in section 3, a multiple regression model was constructed using the appropriate design.

4.4 Testing the Models

The conventional factors, multiple regression models and the trained neural networks were tested using the test set. Errors in estimation for the test set may originate from two sources. One of the sources of errors is the sampling process. The number of patterns in training and test sets might be very small, or the samples may not provide a good representation of the universe. The other source of error is the estimation method itself. In order to get an indication of the errors from these two different sources, the conventional factors and the trained neural networks were tested for training set as well as the test set. Testing the models using the training set indicates how well the training method works by itself.

The values of estimated and actual values of AADT are compared using the following percent difference measure.

$$\Delta = \frac{\text{estimated} - \text{actual}}{\text{actual}} \times 100 \qquad (17)$$

where

Δ = percent error
actual = actual AADT
estimated = estimated AADT

The maximum and average errors for each set are used to compare the results of estimation. The average error provides a measure of the overall accuracy, while the maximum error describes the worst case.

Chapter 5

Results and Analysis

Tables 5.1(a) and 5.1(b), show the results obtained from neural networks, multiple regression, and the conventional method for 24 hour counts using the true classification with the train and test data respectively. For the training set(Table 5.1(a)), the average errors for neural networks range from 5.7% to 13.1%, while the maximum errors range from 27.7% to 86.2%. The average errors for the conventional method range from 6% to 20.3%, and the maximum errors range from 68.8% to 160.6%. The average errors for the multiple regression model range from 3.9% to 7.5%, and the maximum errors range from 20% to 70.1%. The errors for the test data(Table 5.1(b)) are somewhat higher due to sampling errors. The average errors for neural networks range from 5.1% to 19.7%, while the maximum errors range from 20.6% to 46.1%. The average errors for the conventional method range from 5.4% to 24.3%, and the maximum errors range from 24.4% to 102.4%. The average errors in the regression model range from 5.0% to 22.6%, and the maximum errors range from 19.7% to 54.4%. The errors for the neural network and multiple regression models are consistently lower than the conventional method with the exception of group 0 test data. The neural network model produces slightly lower errors than multiple regression in half the cases.

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg	Max	Avg
0	86.2%	13.1%	160.6%	20.3%	23%	6.2%
1	58.6%	5.7%	96.8%	6.0%	70.1%	5.7%
2	45.4%	8.0%	114.8%	12.9%	20.0%	3.9%
3	36.7%	8.3%	138.0%	14.9%	37.5%	7.5%
4	27.7%	6.5%	68.8%	10.4%	29.7%	6.5%

Table 5.1(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	44.7%	14.5%	92.0%	16.3%	42.4%	22.6%
1	20.6%	5.1%	24.4%	5.4%	20.4%	5.0%
2	46.1%	19.7%	102.4%	24.3%	36%	20.8%
3	45.6%	10.9%	81.5%	17.5%	54.4%	11.5%
4	20.8%	6.0%	42.4%	8.5%	19.7%	6.2%

Table 5.1(b) Test Set

Table 5.1 Errors for 24 hour count for Grouping Based on Annual Patterns

The errors for 48 hour counts (Tables 5.2(a) and 5.2(b)) are significantly lower than those for 24 hour counts. For the training set(Table 5.2(a)), the average errors for neural networks range from 3.7% to 6.3%, while the maximum errors range from 9% to 22.9%. The average errors for the conventional method range from 5.2% to 16.1%, and the maximum errors range from 39.1% to 83.9%. The average errors for the multiple regression model range from 4.1% to 5.0%, while the maximum errors range from 16.6% to 23.6%. The multiple regression model contains table entries containing N/A, this indicates there were not enough observations to complete the analysis. Multiple regression analysis requires at least as many observations as independent variables. Table 3.2 shows the number of observations used in each group for the annual grouping classification method. For groups 0 and 2 there were less than 48 observations, thus insufficient to carry out the multiple regression analysis. This becomes even more of a problem in the case of 168 hourly counts as we will see. The errors for the test data(Table 5.2(b)) are somewhat higher due to sampling errors. The average errors for neural networks range from 4.5% to 15.3%, while the maximum errors range from 13.3% to 42.9%. The average errors for the conventional method range from 4.9% to 16.9%, and the maximum errors range from 16.1% to 55.8%. The average errors for the multiple regression model range from 4.2% to 22.4%, while the maximum errors range from 12.8% to 61.2%. The error for the neural network model and multiple regression model are consistently lower than the conventional method. The multiple regression and neural network model are very similar in their results.

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	12.3%	5.8%	83.9%	16.1%	N/A	N/A
1	22.9%	4.3%	48.8%	5.2%	23.6%	4.1%
2	9.0%	3.7%	63.0%	9.5%	N/A	N/A
3	17.0%	6.3%	71.1%	12.4%	16.6%	4.7%
4	20.6%	5.4%	39.1%	8.2%	17.8%	5.0%

Table 5.2(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	19.0%	6.2%	54.0%	13.7%	N/A	N/A
1	14.3%	4.5%	16.1%	4.9%	12.8%	4.2%
2	13.3%	6.3%	48.9%	16.9%	N/A	N/A
3	42.9%	15.3%	55.8%	15.6%	61.2%	22.4%
4	22.2%	6.4%	26.8%	6.6%	22.0%	6.1%

Table 5.2(b) Test Set

Table 5.2 Errors for 48 hour count for Grouping Based on Annual Patterns

The 168 hour counts show further reduction in errors. For the training set(Table 5.3(a)), the average errors for neural networks range from 2.3% to 4.1%, while the maximum errors range from 5.5% to 13.7%. The average errors for the conventional method range from 4.4% to 11.6%, and the maximum errors range from 24.6% to 36.3%. The average errors for the multiple regression model range from 1.9% to 2.1%, while the maximum errors range from 8.4% to 10.0%. Groups 0,2, and 3 had a insufficient number of observations(i.e. # of observations < 168, table 3.2) to carry out the multiple regression analysis. The errors for the test data(Table 5.3(b)) are somewhat higher due to sampling errors. The average errors for neural networks range from 3% to 7.7%, while the maximum errors range from 15.2% to 25%. The average errors for the conventional method range from 3.1% to 22.7%, and the maximum errors range from 9.3% to 37.5%. The average errors for the multiple regression model range from 4.3% to 4.5%, while the maximum errors range from 16.1% to 17.3%. The reduction in errors from 48 hour counts to 168 hour counts is not as significant as the reduction in errors from 24 hour counts to 48 hour counts. This observation indicates that 48 hour counts may be cost effective solution for the estimation of AADT volumes.

Group	**Neural Nets**		**Factors**		**Regression**	
	Max	**Avg.**	**Max**	**Avg**	**Max**	**Avg**
0	13.7%	4.1%	34.9%	11.4%	N/A	N/A
1	7.1%	3.1%	24.6%	4.4%	10.0%	2.1%
2	5.5%	2.3%	35.4%	11.6%	N/A	N/A
3	10.1%	3.1%	35.0%	7.8%	N/A	N/A
4	8.0%	2.7%	36.3%	4.9	8.4%	1.9%

Table 5.3(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	17.6%	5.5%	15.1%	4.2%	N/A	N/A
1	16.6%	3.0%	9.3%	3.1%	16.1%	4.3%
2	18.1%	7.4%	37.5%	22.7%	N/A	N/A
3	25.0%	7.7%	20.0%	8.5%	N/A	N/A
4	15.2%	3.5%	13.3%	4.5%	17.3%	4.5%

Table 5.3(b) Test Set

Table 5.3 Errors for 168 hour count for Grouping Based on Annual Patterns

When a single group is used, the maximum errors for the three time intervals range from 34.8% for 168 hour counts to 312.8% for 24 hour counts (Tables 5.4(a and b)). The average errors range from 4.7% to 13.8%. Very high errors for 24 hour counts further reinforces the conclusion that at least 48 hour counts should be used for AADT estimation.

It is interesting to note that the errors for neural networks without classification (tables 5.4(a and b)) compare favourably with errors for the conventional method with the true classification (tables 5.1(a and b)). This observation seems to suggest that the neural networks implicitly classify patterns during the estimation of AADT. This observation is important because in practice the true classification of STC sites is unknown. But the results from this study indicate that the neural networks with no classification of STC sites can perform as good an estimation as conventional method which assumes that true classification of STC sites is known.

The neural network model consistently outperformed the multiple regression model(with exception to the train set 24 hour counts) for the non-grouping classification scheme. However the differences in error levels between the two models were marginal.

Duration	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg	Max	Avg
24 hours	174.5%	8.4%	312.8%	13.8%	213.0%	9.3%
48 hours	51.7%	6.0%	186.3%	11.2%	83.8%	6.9%
168 hours	34.8%	4.9%	123.1%	11.3%	37.6%	4.7%

Table 5.4(a) Train Set

Duration	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
24 hours	77.6%	10.0%	181.5%	15.0%	102.0%	11.0%
48 hours	79.3%	8.9%	143.7%	12.7%	84.9%	9.3%
168 hours	60.6%	7.5%	130.0%	13.6%	57.7%	7.8%

Table 5.4(b) Test Set

Table 5.4 Errors for different duration counts with no Grouping

In practice, the true classification of STC sites is not known. However, the extreme classification scheme of combining all the highway sites in one group may not necessarily be the correct approach to AADT estimation. Tables 5.5 to 5.7 show the results obtained by grouping the highways sites based on short term patterns. Barring a few anomalies, neural networks and multiple regression analysis provide better results than the conventional method for 24 hour, 48 hour, and 168 hour counts. The neural network model marginally outperformed the multiple regression model. As expected the classification based on short term patterns provides better results than the classification scheme which uses a single group (tables 5.4(a) and 5.4(b)). However, these results are not as good as those obtained using the true classification (tables 5.1 to 5.3).

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg	Max	Avg
0	171.0%	15.7%	270.2%	24.7%	164.8%	16.2%
1	56.5%	10.0%	149.5%	17.0%	59.4%	10.2%
2	12.1%	3.8%	21.1%	4.1%	14.4%	3.5%
3	51.5%	6.9%	54.8%	8.6%	51.8%	6.8%
4	52.5%	5.7%	64.2%	6.2%	53.7%	5.5%

Table 5.5(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	49.0%	16.4%	51.7%	20.3%	77.7%	21.3%
1	57.2%	17.6%	130.9%	31.4%	64.0%	19.1%
2	12.4%	4.8%	14.2%	4.1%	13.0%	4.8%
3	66.5%	8.9%	58.2%	8.45%	65.5%	9.2%
4	71.8%	7.8%	84.5%	8.0%	78.3%	7.6%

Table 5.5(b) Test Set

Table 5.5. Errors for 24 hour count for Grouping Based on Short Term Patterns

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg	Max	Avg
0	16.1%	6.7%	111.0%	17.3%	N/A	N/A
1	60.7%	8.0%	92.3%	11.7%	58.7%	7.7%
2	12.9%	3.7%	20.9%	4.2%	13.3%	3.2%
3	9.4%	4.2%	38.1%	11.0%	9.0%	2.8%
4	36.9%	5.7%	46.4%	5.9%	32.7%	5.2%

Table 5.6(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	80.8%	21.1%	79.6%	23.7%	N/A	N/A
1	72.8%	10.7%	102.5%	13.0%	76.1%	11.3%
2	14.3%	4.2%	9.8%	3.8%	12.8%	4.5%
3	15.4%	5.7%	21.9%	10.3%	21.8%	7.9%
4	70.7%	7.9%	76.6%	7.2%	68.0%	7.5%

Table 5.6(b) Test Set

Table 5.6. Errors for 48 hour count for Grouping Based on Short Term Patterns

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg	Max	Avg
0	13.7%	5.0%	65.9%	13.9%	N/A	N/A
1	6.2%	2.0%	23.2%	4.6%	N/A	N/A
2	8.8%	2.4%	98.2%	11.6%	N/A	N/A
3	16.9%	4.6%	84.5%	7.6%	13.5%	2.9%
4	13.9%	2.6%	36.0%	6.3%	N/A	N/A

Table 5.7(a) Train Set

Group	Neural Nets		Factors		Regression	
	Max	Avg.	Max	Avg.	Max	Avg.
0	60.0%	14.6	101.2%	21.1%	N/A	N/A
1	28.6%	7.2%	9.2%	4.3	N/A	N/A
2	85.1%	10.4	98.6%	12.7	N/A	N/A
3	57.5%	7.9%	46.4%	8.1%	48.4%	7.6%
4	11.6%	3.8%	15.6%	6.1%	N/A	N/A

Table 5.7(b) Test Set

Table 5.7. Errors for 168 hour count for Grouping Based on Short Term Patterns

Due to the tabular nature of the resultant data, it is difficult to visualize the results. Therefore I have provided three graphs which represent the three classification schemes. Within each graph, the three estimation models are compared to each other. The comparison measure is the percentage error in estimation of the AADT volume versus the hourly count period. For the annual grouping and short term classification schemes I have only shown the results of one group in the graphs.

For the annual grouping classification scheme(Graph 5.1), we see that the factor model produces the worst performance(greatest error). It is represented by the line with the triangular nodes. The factor model is consistently worst than the neural network model(line with the square nodes), and almost consistently worst than the multiple regression model(line with circular nodes) with the exception of the 168 hourly count. This exception is most likely due to sampling errors. The neural network model is marginally worst than the multiple regression model for the 24 and 48 hourly counts. It should be noted that when viewing all the data from all five groups for the annual grouping scheme, the neural network model more often performs marginally better than the multiple regression model. Another important point that the graph illustrates, is the error levels decrease as the duration of the hourly count increase. From 24 to 48 hours, all three models increase in performance. This is to expected due to the extra data. The same is true for the transition to the 168 hourly count from the 48 hourly count, with exception to the multiple regression model, which is most likely due to sampling errors.

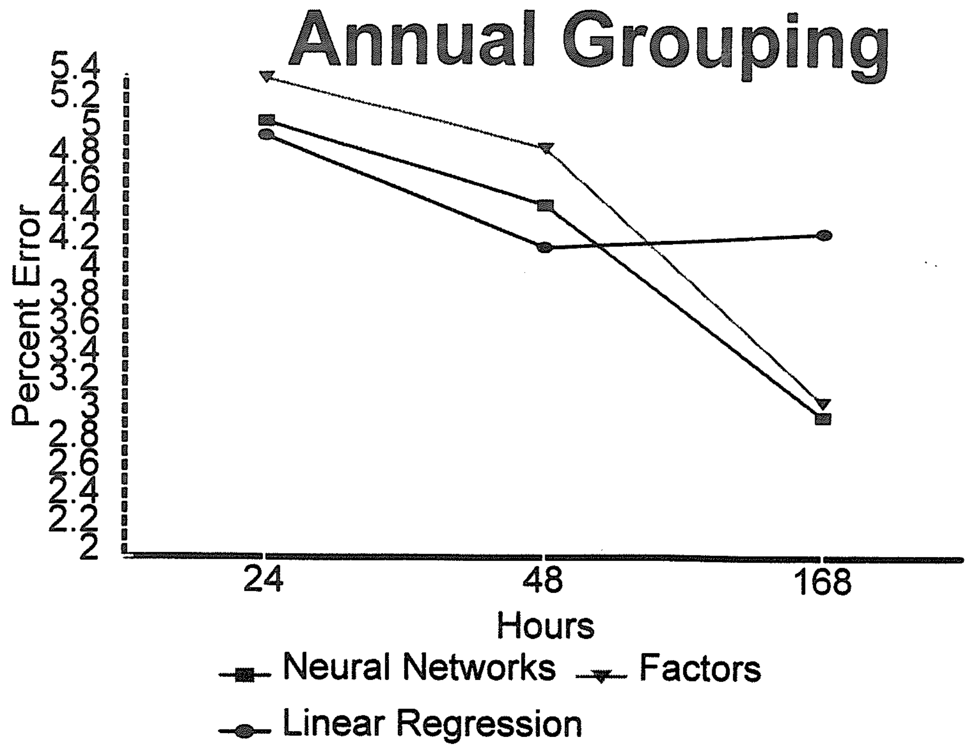

Table 5.1 Annual Grouping Classification scheme for Group 1 - Test Data

The next graph(Graph 5.2) depicts the results for the non-grouping scheme. This graph reflects the majority of the results. The factor model produces the worst results, the linear regression model is marginally higher in error levels than the neural network model, and the neural network model produces the best results compared to the other two models. This graph also shows the direct correlation between error level and hourly count(as sample duration increases, the error level decreases). This is in exception to the factor model for the 48 to the 168 hourly count, which is likely due to sampling errors. The amount of performance benefits between 24 and 48, and, 48 and 168, is approximately the same. This demonstrates that the performance gain is not as drastic from the 2 day count to the 7 day, as from the 1 day count to the 2 day count, considering the need for the STC on site for the extra five days.

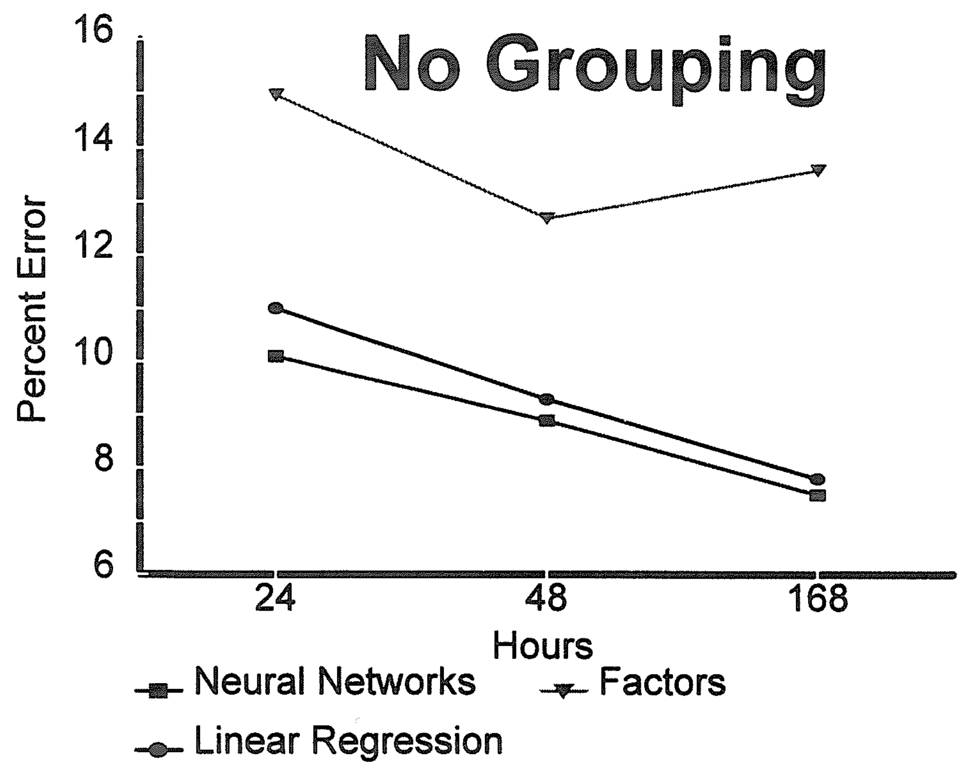

Table 5.2 Non-grouping Classification Scheme - Test Data

The last of the classification schemes is the one based on short term patterns and is shown in graph 5.3. Sampling errors are more evident in this graph, and are demonstrated by the factor model decreasing in performance between the 24 and 48 hour count, and the neural network model decreasing in performance between the 48 and 168 hourly count. For this graph we will focus our attention on the 48 hourly count. The neural network model performs the best followed by the multiple regression model and then the factor model.

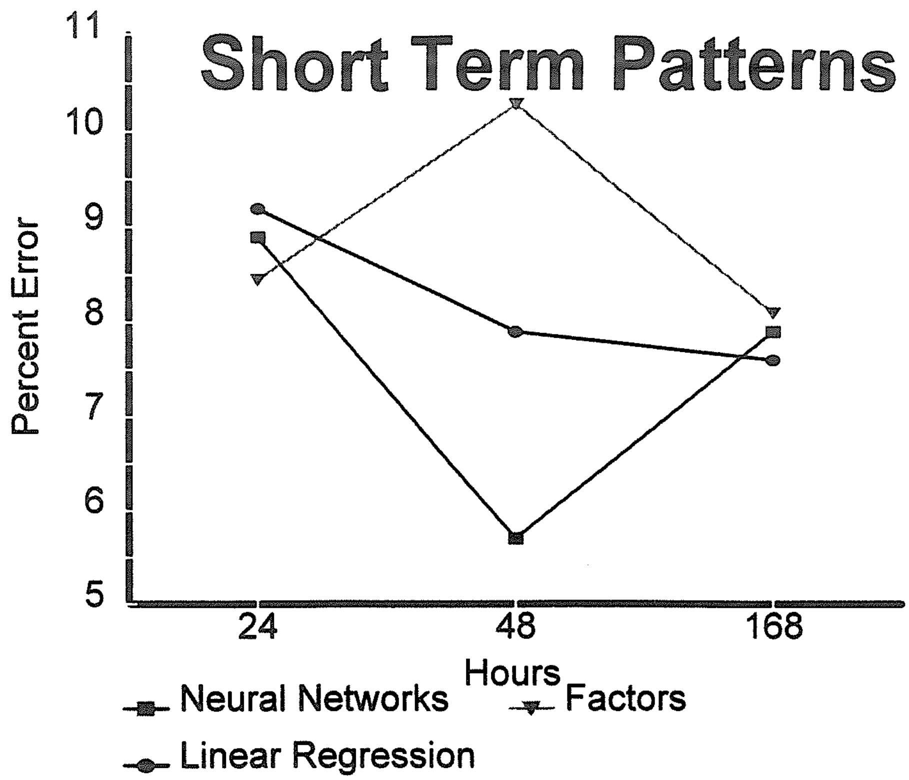

Table 5.3 Classification based on short term patterns for Group 3 - Test Data

By viewing and comparing all three graphs, it is evident the annual grouping classification scheme produces the best results. The non-grouping scheme provides the worst results and the classification based on short term patterns is in the middle.

Chapter 6

Summary and Conclusions

This study compared the estimation of AADT using the conventional factor based approach, multiple regression analysis, and the neural network approach. The estimation was based on 24 hour, 48 hour and 168 hour short term counts. Classification of highway sites is an important factor in the estimation procedures. Hence, three different classification schemes were used. The first classification scheme used the true classification based on annual patterns. Since the annual patterns are unavailable for the seasonal traffic counter sites, the second classification scheme combined all the highway sites in a single group. Between these two extreme classification schemes there was a third classification scheme which classified highway sites according to the short term traffic patterns. These short term traffic patterns are available for all the highway sites under investigation.

For all the classification schemes used in the study, the neural networks and multiple regression consistently outperformed the conventional method. The neural network model in many cases slightly outperformed the multiple regression model. The true classification scheme provided the best results for all three approaches. The errors for the second classification scheme, which combined all the highway sites in a single group, were the highest. The third classification scheme, based on the short term traffic patterns, provided results which were better than the second scheme but worse than the first scheme. Even though an accurate classification scheme yields the best results, such a classification is not possible for STCs. The classification of STCs is generally based on the judgment of traffic engineers who are the domain experts. The third classification scheme based on the short term traffic patterns, provides an objective classification method which can be used in practice. The third classification scheme also provides reasonably accurate estimates of AADT.

Another interesting conclusion that follows from this study is related to the duration of the short term traffic counts. The results obtained from 48 hour traffic counts are substantially better than the 24 hour traffic counts. However, the improvement with week-long counts over the 48 hour counts is not significant enough to justify an additional five days of counting. That is, 48 hour traffic counts are cost-effective data collection schedules for the estimation of AADT.

References

DeGarmo, E.P., Sullivan, W.G. and Canada, J.R. 1984. *Engineering Economy*, Macmillan Publishing Co., New York, N.Y., pp. 264-266.

Garber, N.J. and Hoel, L.A. 1988. *Traffic and Highway Engineering,* West Publishing Co., New York, N.Y., pp. 97-118.

Fwa, T.F. and Chan W.T. 1993. Priority Rating of Highway Maintenance Needs by Neural Networks, Journal of Transportation Engineering, American Society of Civil Engineers, 116, 3, pp. 419-432.

Lingras(1995) Classifying Highways: Hierarchical Grouping Vs Kohonen Neural Networks, *to appear in the Journal of Transportation Engineering*, American Society of Civil Engineers.

Sharma S.C. and Allipuram, R.R. 1993. Duration and Frequency of Seasonal Traffic Counts, Journal of Transportation Engineering, American Society of Civil Engineers, 116, 3, pp. 344-359.

Sharma S.C. and Werner, A. 1991. Improved Method of Grouping Province wide Permanent Traffic Counters, Transportation Research Record 815, Transportation Research Board, Washington D.C., pp. 13-18.

Hecht-Nielsen, R. *Neurocomputing*, Addison-Wesley Publishing, Don Mills, Ontario.

Lawrence, J. 1993. *Introduction to Neural Networks: Design Theory and Applications*, California Scientific Software Press, Nevada City, California.

White, H. 1989. Neural Network Learning and Statistics, *AI Expert*, 4, 12, 48-52 .

Zahedi, F. 1990. *Intelligent Systems for Business: Expert System with Neural Networks*, Wadsworth Publishing Company, Belmont, California.

Kohler, Heinz, 1988. *Essentials of statistics*, Scott, Foresman and Company, Glenview, Illinois.

Glossary

AADT = Annual average daily traffic.

Classification = A process of grouping data patterns into classes of similarity.

$E(\mathbf{x},\mathbf{y})$ = Mean square error between two complete patterns **x** and **y**.
$E'(\mathbf{g},\mathbf{z})$ = Mean square error between a complete pattern **g** and an incomplete pattern **z**.

Generalized Delta Rule = A learning technique used in a supervised learning neural network configuration.

g, x, z, y = Vectors representing traffic patterns.

Hierarchical grouping = technique used to group a set of k-dimensional vectors
Kohonen Neural Network = A neural network configuration for unsupervised learning.

MADT = Monthly average daily traffic.
Multiple Regression = A type of statistical analysis procedure which requires the use of more than one independent variable(i.e. 24, 48, or 168).

Pattern = a k-dimensional vector of degree 24, 48 or 168, where each dimension is represented by a hourly count.

PTC = Permanent Traffic Counter

SADT = Seasonal Annual Average Daily Traffic

STC = Seasonal Traffic Counter

t = Time.
$\alpha(t)$ = Learning parameter in the Kohonen rule expressed as a function of time.