OFFICIAL DOCUMENT

DO NOT REMOVE FROM
THE RESEARCH OFFICE

IMPROVED ESTIMATES OF TRAVEL TIME FROM REAL TIME INDUCTANCE LOOP SENSORS

WA-RD 292.1

Final Technical Report
May 1993

Washington State
Department of Transportation

Washington State Transportation Commission
Transit, Research, and Intermodal Planning (TRIP) Division
in cooperation with the U.S. Department of Transportation
Federal Highway Administration

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO. WA-RD 292.1	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
4. TITLE AND SUBTITLE IMPROVED ESTIMATES OF TRAVEL FROM REAL TIME INDUCTANCE LOOP SENSORS		5. REPORT DATE April 1993
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Daniel J. Dailey Mark P. Haselkorn Nancy L. Nihan		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Washington State Transportation Center (TRAC) University of Washington, JE-10 The Corbet Building, Suite 204; 4507 University Way N.E. Seattle, Washington 98105		10. WORK UNIT NO.
		11. CONTRACT OR GRANT NO. T9233, Task 5
12. SPONSORING AGENCY NAME AND ADDRESS Washington State Department of Transportation Transportation Building, MS 7370 Olympia, Washington 98504-7370		13. TYPE OF REPORT AND PERIOD COVERED Final Technical Report
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

The Real Time Motorist Information System (RTMIS) under development at the University of Washington uses a ratio of Volume to Occupancy with a correction factor (often denoted as "g") to estimate speed. This correction factor is a function of road topology and occupancy. This project will initially produce a value for the correction factor for every pair of loops being used in the RTMIS. It will further produce a correlation between correction factor and occupancy for each set of loop pairs. The correction factor is produced using speed estimates from cross-correlation time delay measurements of on road data.

17. KEY WORDS Induction Loops, Speed Cross Correlation, Volume Occupancy	18. DISTRIBUTION STATEMENT No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None	50	

Final Technical Report

Research Project T9233, Task 5
"Improved Travel Time Estimates"

IMPROVED ESTIMATES OF TRAVEL TIME FROM REAL TIME INDUCTANCE LOOP SENSORS

by

Daniel J. Dailey
Assistant Professor of Electrical Engineering
(Principal Investigator)

Mark P. Haselkorn
Professor of Technical Communication

Nancy L. Nihan
Professor of Civil Engineering

University of Washington
Seattle, WA 98195

Washington State Transportation Center (TRAC)
University of Washington, JD-10
University District Building
1107 N.E. 45th Street, Suite 535
Seattle, Washington 98105-4631

Washington State Department of Transportation
Technical Monitor
Larry Senn
Urban Systems Engineer, FAME

Prepared for

Washington State Transportation Commission
Department of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

May 1993

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Transportation Commission, Department of Transportation, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Contents

LIST OF FIGURES

List of Tables

Executive Summary

The Real Time Motorist Information System (RTMIS) under development at the University of Washington uses a ratio of Volume to Occupancy with a correction factor (often denoted as "g") to estimate speed. This correction factor is a function of road topology and occupancy. This project will initially produce a value for the correction factor for every pair of loops being used in the RTMIS. It will further produce a correlation between correction factor and occupancy for each set of loop pairs. The correction factor is produced using speed estimates from cross-correlation time delay measurements of on road data.

1. Introduction

The principle source of freeway data available to traffic management and traveler information systems in the Puget Sound region is the DOT inductance loop system. Traffic management systems and traveler information systems currently used in the Puget sound region, such as the "FLOW" map and, the real time motorist information system (RTMIS) named "Traffic Reporter," [MJMB91] use this inductance loop data to determine the status of traffic in the freeway system. While trained traffic management systems personnel can use the volume (the number of vehicles passing over a loop in a given time) and occupancy (the fraction of some total time a loop is occupied) data directly, the needs are different for traveler information systems. It has been asserted that traffic speed and travel time are the properties most useful in providing relevant information to travelers. [MB89] This report presents a methodology to convert the measured volume and occupancy to speed and travel time. It further provides a standard mechanism (with a location specific coefficient) to estimate speed directly using the volume/occupancy ratio. Improved estimated speed from the volume/occupancy ratio allows the traveler information system to provide accurate, real time information to the users. This report is presented in five sections: this introduction, theory, methodology, results, and conclusions.

The second section presents the theory on which travel time estimate are based. The material in this section provides a theory that can be applied to real time traffic data.

The four principle theses of this section are: (1) travel time between inductance loops can be estimated independent of the mean value of volume and occupancy; (2) independent speed estimates can be made from either a cross correlation based travel time estimate or from a volume/occupancy ratio methodology; (3) a relationship between the two independent speed estimates can be used to improve the volume/occupancy ratio estimate; and (4) a total travel time estimate between distant points can be made with detailed knowledge of the speed in each subsegment of the trip.

The third section describes the data format and contents, as well as the data acquisition and data reduction processes. The data format and contents subsection describes the data available from the inductance loop sensor system. The data acquisition and reduction subsection describes the acquisition of loop data from the Seattle metropolitan freeway system, and the details of the numerical processing of the loop data time series, and in addition provides a detailed description of the methodology that converts raw binary loop data into values such as travel time and speed.

The fourth section summarizes the results of the methodology described. First the results of cross correlating inductance loop data to estimate transit time is presented. Second, the mean traffic speed estimates based on the inter-loop distance and the cross correlation travel time are presented. Third the speed estimates from the volume/occupancy ratio techniques are compared to those from the cross correlation technique. Finally, the effect of location on the value of the "speed factor" is presented.

The final section draws conclusions about the results of this research project. Conclusions are drawn in four areas. First, the validity of the "speed factor" concept is examined. Second, the limitations of this concept are enumerated. Third, recommendations about usage of the cross correlation technique are presented.

2. Theory

This section develops the theory on which the improved travel time estimates are made. It discusses the estimation of interloop travel times from the inductance loop times series. Speed estimates are developed based on volume and occupancy counts. A new method to estimate interloop travel delay estimates from the inductance loop time series is summarized. A relationship between the observed speeds and the volume/occupancy ratio is developed. Finally the relationship between speed and overall travel time associated with a driver's total trip is presented.

2.1. Interloop Travel Times

This subsection presents an overview of the methodology for making estimates of the interloop travel time using inductance loop time series data. This subsection reviews theory developed in a previous TransNow report entitled, "Travel Time Estimation using Cross-Correlation Techniques" (see this report for more detail).

The traffic is modeled as propagating down the highway with some mean volume count per unit time. The volume estimates from inductance loops have a mean value around which there is a time dependent statistical fluctuation. If the traffic propagated rigidly and a slight increase in the volume is observed at the upstream station, a similar slight increase would be expected at the downstream station some time later and likewise for slight decreases.[1] The time shift necessary to optimally align the upstream and downstream time series is the mean propagation or travel time of the traffic between the stations. To estimate the mean travel time between stations, the cross correlation coefficient function (CCCF) is used. The CCCF is written,

$$\rho_{12}(\tau, T) = \frac{R_{12}(\tau, T)}{\sigma_1 \sigma_2} \tag{2.1}$$

where

$$R_{12}(\tau, T) = \frac{1}{2T} \int_{-T}^{T} \delta\alpha(x_1, t)\delta\alpha(x_2, t - \tau)dt, \tag{2.2}$$

is the cross correlation function (CCF),

$$\sigma_1^2 = R_{11}(0), \tag{2.3}$$

is the variance, and the continuous time dependent volume at the upstream station is labeled $\delta\alpha(x_1, t)$. The CCF function indicates the correlation between the two time series as a function of the delay time τ. It has the property that the location of the maximum value is the delay time at which the time series from the two inductance loop detector stations are most similar. So, by locating the maximum value in this continuous function, the mean travel time between the inductance loop stations can be estimated.

In practice, the volume data from the on-road inductance loops is a discrete time series. For this analysis, the data is time and spatially averaged and divided into

[1] The manuscript [Dai91] provides more detail on the limits of this methodology.

ensembles of length N. A mean centered volume estimate for the i'th site for the k'th time ensemble is written,

$$V_{ik} = (Vol)_{ik} - \frac{1}{N}\sum_{k=1}^{N}(Vol)_{ik}. \tag{2.4}$$

An estimate of the CCF can be made,

$$\hat{R}_{12}(\tau_k, T) = \sum_{j=1}^{N} V_{ij}V_{2(j+k)} \quad \forall k \quad \in [0, N]. \tag{2.5}$$

The CCF can then be normalized by the standard deviations to produce an estimate of the CCCF,

$$\hat{\rho}_{12}(\tau_k, T) = \frac{R_{12}(\tau_k, T)}{\sigma_1\sigma_2}, \tag{2.6}$$

where,

$$\sigma_1^2 = \frac{1}{N}\sum_{k=1}^{N} V_{1k}V_{1k} \quad \sigma_2^2 = \frac{1}{N}\sum_{k=1}^{N} V_{2k}V_{2k}. \tag{2.7}$$

Since the delay time τ_0 of the maximum value of the CCCF is the mean travel time, a method of estimating this value is needed. To estimate the value of τ_0 from the discrete estimate $(\hat{\rho}(\tau_k, T))$ of $\rho(\tau, T)$, the region near the maximum value in the array $\hat{\rho}(\tau_k, T)$ is least squares fit to an analytical function, and the analytical derivative of this fitted function is used. If the data is approximated as having bandwidth limited Gaussian statistics, the correlation function can be approximated by

$$\rho_{12}(\tau) \approx \frac{sin2\pi B(\tau - \tau_0)}{2\pi B(\tau - \tau_0)}. \tag{2.8}$$

If the trigonometric term in equation (2.8) is expanded in the first two terms of a power series, it can be rewritten,

$$\rho_{12}(\tau) \approx \frac{2\pi B(\tau - \tau_0) \quad - \quad \frac{1}{3!}\left(2\pi B(\tau - \tau_0)\right)^3}{2\pi B(\tau - \tau_0)}, \tag{2.9}$$

which has the general form

$$\rho(\tau) \approx a\tau^2 + b\tau + c \tag{2.10}$$

near the peak value of $\rho_{12}(\tau)$. The analytical derivative of the approximation is

$$\frac{\partial}{\partial \tau}\rho(\tau_0) = 2a\tau_0 + b = 0. \tag{2.11}$$

The parameters a and b can be determined by least squares fitting of the quadratic function to the estimate of $\rho(\tau)$ in the immediate region of the peak value. Then an estimate of τ_0 can be produced from

$$\tau_0 = \frac{-b}{2c}. \tag{2.12}$$

This method provides a mechanism to estimate the mean time for traffic to travel between inductance loops. This method depends on the propagation of small disturbances in the mean volume moving between displaced loops and is independent of the mean volume or occupancy values. An average speed estimate can be made using the distance between the loops and the estimated delay/travel time.

2.2. Speed Estimates

This subsection discusses two methods for estimating vehicle speeds. First, a method that estimates mean traffic speed using the cross correlation time delay developed in Section 2.1 is presented. Second, a method that estimates speed from the volume/occupancy ratio and a single coefficient, along with the framework in which this method was developed, is presented.

2.2.1. Speed Estimates Derived from Cross Correlation Delay Times

Using the cross correlation delay estimate developed in the last section with the measured interloop distance, a mean speed can be calculated,

$$\hat{S} = \frac{(x_2 - x_1)}{\tau_0}. \tag{2.13}$$

This is the constant speed at which a vehicle would travel to traverse the distance between the loops in time τ_0. Estimating speed in this manner implies that the mean traffic speed between the loops is a constant for the duration of the time used to make the time delay estimate.

If this same estimate is used for several contiguous sections of highway, a higher order approximation of the change in the speed can be made. For example, if two sequential areas of the highway are used and two time delays are measured, then the speed can be approximated as changing linearly between two end points. If the delay between x_1 and x_2 is τ_1 and between x_2 and x_3 it is τ_2, an approximation of the speed might be

$$\hat{S} = \left(\frac{\frac{(x_3 - x_2)}{\tau_2} - \frac{(x_2 - x_1)}{\tau_1}}{x_3 - x_1} \right) x + \frac{(x_2 - x_1)}{\tau_1}. \tag{2.14}$$

This approximation implies a linear change in traffic speed between the loops.

This method of estimating speed from the interloop delay time has four principle limitations. First, speed estimates based on delay times depend on the ability to identify a delay peak in the CCCF. This in turn depends on the assumption that there is rigid propagation of the fluctuations in the traffic volume between the loops. This assumption is accurate below about 20% occupancy but is suspect during periods when the occupancy value is high and the volume count is low. Second, the speed is assumed to be either a constant or changing linearly. This assumption is once again more accurate in periods of low occupancy and high volume traffic. Third, the delay is an average delay for all traffic, and is therefore insensitive to individual vehicle speeds. Fourth, this method is computationally expensive. It cross correlates arrays of volume estimates (based on a five-second average) to build the CCCF. This level of computation is not available in the present RTMIS computers, and the five-second time averaged data is not available to the RTMIS. These four constraints limit the range of traffic conditions over which this technique is useful. However, this technique provides the only published estimate of speed from inductance loop data that is independent of the mean value of volume and occupancy. The importance of this independence will become clear in section 2.3.

2.2.2. Speed Estimates Derived from Volume/Occupancy Ratios

Conventional estimates of speed used by traffic management systems rely on the values of volume and occupancy to estimate speed. It has long been asserted that the ratio of flow (q) to concentration (k) is the "space mean speed." [LW55, War52] The quantities flow and concentration are not immediately available from inductance loop sensors. RTMIS presently constructs estimates of these quantities from volume (V), occupancy (O), and a constant factor (g).

The classical relationship that is the basis for the speed estimate is:

$$s = \frac{q}{k} \tag{2.15}$$

where s is the space mean speed in miles per hour, q is the traffic flow in vehicles per hour, and k is the traffic concentration in vehicles per mile. To use this relationship, volume must be converted to flow and occupancy to concentration. Since volume is the number of actuations of a loop per unit time (one minute in RTMIS practice) and an actuation is considered by RTMIS to be a vehicle passing over the loop, only a constant conversion factor is needed to convert volume to flow,

$$q = \frac{60}{T} \times V, \tag{2.16}$$

where T is the number of minutes over which the vehicle count is done.

Occupancy is defined to be the time during which a loop is occupied divided by some total time (or fraction of the total time that a car is present). The conversion from occupancy to flow is not straightforward. The time a single vehicle occupies the inductance loop is the sum of two components: the time for the front of a vehicle to pass over the detector's region of sensitivity (l_d) and the time for a vehicle of length

(l_v) to exit the region. Time of occupation of a loop is:

$$T_O = \int_{l_d} \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx \tag{2.17}$$

where $s(x)$ is the speed as a function of position. The "total time" in a region of length D is:

$$T = \int_{D} \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx. \tag{2.18}$$

The value of occupancy for one vehicle normalized to a distance of D is:

$$O_1 = \frac{\int_{l_d} \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx}{\int_{D} \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx.} \tag{2.19}$$

The speed is assumed constant over the sensing region resulting in:

$$O_1 = \frac{\frac{1}{s}(l_d + l_v)}{\frac{1}{s}(D + l_d).} \tag{2.20}$$

This is the non-dimensional value of occupancy for one vehicle normalized to a distance D. The relationship between traffic concentration and occupancy is assumed to be:

$$k = g \times O, \tag{2.21}$$

indicating that g (the "speed factor") scales between the occupancy value scale $[0, 1/O_1]$ and the concentration scale $[0, 1]$. Using,

$$\begin{aligned} s &= constant \\ l_d &= 6(feet) \\ l_v &= 18(feet/vehicle) \\ D &= 5280(feet/mile), \end{aligned}$$

produces $g = 220$ (vehicles per mile). This is the value that RTMIS uses to convert the fractional occupancy to concentration.

At the present time the RTMIS is using a single constant with data averaged over one minute to define a linear relationship between speed(S) and the ratio of volume (V) to occupancy (O)

$$S = \frac{60 \times V}{g \times \frac{O}{100}} \tag{2.22}$$

where $g = 220$. While this estimate is computationally simple, it depends on the accuracy of the estimate of vehicle length and sensitivity length of the loop detectors both of which are statistical properties. Since any traffic flow contains a mixture of vehicles of various lengths, the g factor is really an empirical constant that is to be determined for a particular set of loops. The empirical nature as well as the site and occupancy dependency of the value for g have been topics of several publications [PH88, HP88]. It is this empirical value issue that limits the utility of this method. To calibrate g, a means of measuring the time and spatially averaged mean speed is needed (typical single vehicle measurements using closely spaced paired loops or travel time estimates from license plate tracking produce single vehicle speeds). The correlation method developed here addresses this issue. The delay and approximate mean speed are estimated independent of the mean values for volume and occupancy. This independent speed estimate can then be used to calibrate g for a particular site over a range of occupancies. Once this calibration is in place, the mean speed using volume and occupancy measurements can be produced with greater accuracy.

This section has described two different methods for determining mean traffic speed from the data available from the inductance loop system. The first method uses a cross correlation technique that provides an estimate of interloop travel time from which speed is estimated. This estimate is independent of the mean volume and occupancy values. The second method used the ratio of mean volume to mean occupancy, with an empirical constant to provide a speed estimate. The computationally expensive independent speed estimate from the correlation technique can be used to calibrate the empirical constant, and provide a more accurate estimate of speed from the volume/occupancy ratio.

2.3. Speed Factor Calculation (g)

RTMIS and the Traffic Systems Management Center (TSMC) computers use the volume/occupancy ratio technique to estimate speed, and therefore an accurate value for the speed factor is very important. To improve the estimate of speed from volume and occupancy, the speed factor g must be estimated for each loop station [HP88]. To do this, time averaged volumes (V) and occupancies (O) as well as estimates of mean speed must be made simultaneously at each of the sites. The speed factor at the "ith" site can then be estimated,

$$g_i = (\frac{V_i}{O_i}) \times \frac{1}{S_i}. \tag{2.23}$$

However, the speed factor is also a function of occupancy as well as location. So the estimate of speed from volume and occupancy should be written,

$$S_i = (\frac{V_i}{O_i}) \times \frac{1}{g_i(O_i)}. \tag{2.24}$$

The site and occupancy dependency of g is examined in more detail in the following section.

If an accurate empirical relationship between the volume/occupancy ratio and speed can be developed using g, then the speed estimates from RTMIS will be more accurate and reliable. Further, other quantities that depend on the estimation of speed (such as total travel time) will be more reliable with an accurate estimate of speed.

2.4. Total Travel Times

One of the important pieces of information travelers need to make decisions is the estimated travel time between their starting point and their destination. The RTMIS estimates this total travel time using speed estimates along the route of travel. A speed estimate is available at each set of loops; and if the speed of traffic is assumed to change linearly between stations, the speed as a function of position is written,

$$s(x) = s_i + \frac{s_{i+1} - s_i}{x_{i+1} - x_i}. \tag{2.25}$$

The travel time between any two stations located at x_i and x_{i+1} is:

$$\tau_{i+1} = \int_{x_i}^{x_{i+1}} \frac{1}{s(x)} dx. \tag{2.26}$$

Now define,

$$\Delta s = s_{i+1} - s_i \tag{2.27}$$

$$\Delta x = x_{i+1} - x_i \tag{2.28}$$

and the travel time can be written:

$$\tau_{i+i} = \int_{x_i}^{x i+1} \frac{1}{s_i + x\frac{\Delta s}{\Delta x}} dx. \tag{2.29}$$

Integration results in:

$$\tau_{i+1} = \frac{\Delta x}{\Delta s} ln \left(x\frac{\Delta s}{\Delta x} + s_i \right) \Bigg|_{x_i}^{x_{i+1}}. \tag{2.30}$$

Applying the limits,

$$\tau_{i+1} = \frac{\Delta x}{\Delta s} \left\{ ln \left(x_{i+1} \left(\frac{\Delta s}{\Delta x} \right) + s_i \right) - ln \left(x_i \left(\frac{\Delta s}{\Delta x} \right) + s_i \right) \right\}. \tag{2.31}$$

Now, if a linear transformation is used, each τ_{i+1} can be used for total trip time estimates.

$$x_i \quad \rightarrow 0 \tag{2.32}$$

$$x_{i+1} \quad \rightarrow (x_{i+1} - x_i) = \Delta x \tag{2.33}$$

results in:

$$\tau_{i+1} = \frac{\Delta x}{\Delta v} ln \left(\frac{v_{i+1}}{v_i} \right). \tag{2.34}$$

This result, while analytically correct, is difficult to implement numerically in the area where $v_i \approx v_{i+1}$ and $\Delta v \approx 0$. However, much of the operating envelope for the RTMIS is in this region. To overcome this difficulty, expand the result as a (Taylor) power series:

$$\tau_{i+1} = 2\Delta x \left\{ \frac{1}{v_{i+1} + v_i} + \frac{(\Delta v)^2}{3} \left(\frac{1}{v_{i+1} + v_i} \right)^3 + \frac{(\Delta v)^4}{5} \left(\frac{1}{v_{i+1} + v_i} \right)^5 \cdots \right\} \tag{2.35}$$

It is noteworthy that this series expansion has none of the numeric difficulties inherent in the analytical result, particularly in the most likely region of operation for the RTMIS project. In addition, these polynomials can be evaluated quickly, providing a computationally cheap estimate of travel time.

The number of terms used in this series will have an effect on the accuracy of the resulting travel time estimate. The percentage error in the estimate is:

$$\epsilon\,(n, v_i, v_{i+1}) = \left(1 - \frac{\displaystyle\sum_{n=0}^{N} \frac{(\Delta v)^{2n}}{2n+1} \left(\frac{1}{v_{i+1} + v_i} \right)^{2n+1}}{\displaystyle\frac{\Delta x}{\Delta v} ln \left(\frac{v_{i+1}}{v_i} \right)} \right) \times 100. \tag{2.36}$$

Numerical experiments indicate that the use of three terms in the series provides a maximum error range between 3% and 4%. An additional five terms (total eight) reduces this only by 1%. These ranges are approximate since the error depends on several variables as well as the region over which the travel time is being estimated.

Travel time between stations can be estimated using the three term polynomial, and the total travel time (T) between two sites will be the sum of all the travel time estimates (τ_{i+1}) for the stations between the sites.

$$T = \sum_i \tau_{i+1} \tag{2.37}$$

where

$$\tau_{i+1} \approx 2\Delta x \left\{ \frac{1}{v_{i+1} + v_i} + \frac{(\Delta v)^2}{3} \left(\frac{1}{v_{i+1} + v_i} \right)^3 + \frac{(\Delta v)^4}{5} \left(\frac{1}{v_{i+1} + v_i} \right)^5 \right\}. \tag{2.38}$$

Using this technique, the total travel time between any two sites, beginning and terminating at stations, can be calculated. This technique provides a means to produce the trip time estimates needed for traveler information support.

2.5. Summary

This section on theory has developed the foundation for improving travel time estimates from inductance loop data. It presents four principal themes. First, it reviewed a correlation technique to estimate inter loop travel times independent of the mean values of volume and occupancy. Second, it discussed the estimation of speed using two methods: (1) the direct ratio of cross correlation estimated time delay to distance; and (2) the use of the ratio of volume/occupancy using an empirical constant (speed factor g). Third, a method that uses the correlation technique with measured volume and occupancy, to improve the accuracy of speed estimates by producing a site specific value of g, was presented. Once an accurate mean traffic speed can be predicted, this speed estimate can be used to estimate total trip time between distant point within the highway system. And finally, this section presented a detailed calculation scheme to provide an accurate travel time estimate for a set of speed estimates.

3. Methodology

This section presents the methodology used in both data acquisition and data reduction. Data from many locations on Interstate 5 (I-5) north were recorded during this effort. Once recorded, the data was transferred to the University of Washington (UW). At the UW, the loop data is reformatted and analyzed to produce estimates for quantities such as time averaged volume, occupancy, speed, and speed factor. In this section, the mechanics and format of the data acquisition are first detailed. Then the methodology used for the estimation of the the speed factor function is presented.

3.1. Data Formats and Contents

This section gives an overview of the data formats and contents. At many sites on I-5 where there are stations consisting of inductance loops and "cabinets" containing microprocessors. Metropolitan Seattle uses two types of loop sensor stations: (1) mainline-type stations and (2) ramp-type stations. The first type is a set of loops that estimates volume and occupancy on the mainline of the freeway. The second type of loop station collects mainline data, ramp data and controls the ramp meter signal. This distinction is important to this project because each type of station is polled by a different computer. In either case the loops (one per lane) are sampled 60 times per second to determine if the inductance of the loop indicates the presence or absence of vehicles.

The first of two possible measurements is the number of initial activations per second. The value is between 0 and 3 and is the number of transitions from the "off" state to the "on" state, indicative of the number of cars that passed over a loop during this second. The second measurement is the total number of samples that were in the "on" state for that second. The vehicle count is be used to derive volume (vehicles/hour) and the total "on" period count to derive occupancy (the fraction of time a loop is occupied).

These two numbers are encoded and sent via modem to TSMC. The encoding is: the two most significant bits (msb) are the count (0-3); and the six least significant bits (lsb) are the number of activations (0-60). Each 8-bit byte is followed by a parity bit for error detection. This results in an effective 9-bit data communication scheme (see Figure 3.1). Nine-bit data communications is a very unusual scheme and requires special data acquisition programming. A series of 8 bytes (9 bits each) is transmitted,

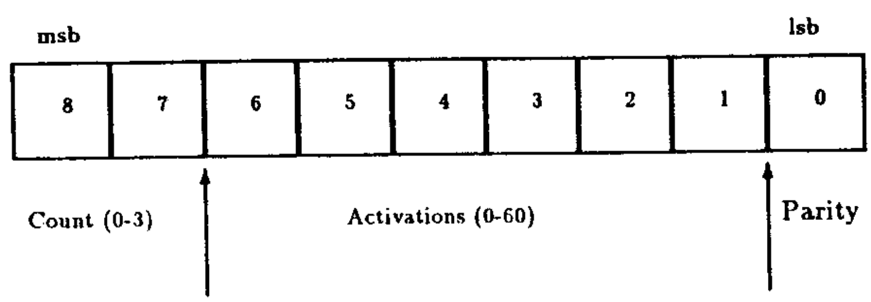

Figure 3.1: TSMC data byte bit format.

and then a byte consisting of the "exclusive or" of all 8 bytes is transmitted. A sample from a station is therefore a "frame" of nine bytes (9 bits in each byte, see Figure 3.2). Northbound and southbound samples are transmitted at one second intervals.

1 byte	2 byte	3 byte	4 byte	5 byte	6 byte	7 byte	8 byte	XOR byte

Figure 3.2: TSMC data frame byte format.

Each sample is transmitted on a one half second boundary. The result is a two 9 byte frames per second. This is the "raw" data from the inductance loop system that will be analyzed.

3.2. Data Acquisition, Decoding, and Reduction

This section describes the methodology used to obtain the inductance loop data from the I-5 corridor. It also describes the decoding of the information as sent back from the on-road cabinets. It then examines the data reduction logic used to estimate travel time between loop stations.

3.2.1. Data Acquisition and Decoding

The on-road data is normally received by a Perkin-Elmer (PE) computer located at TSMC. This project needed access to data directly from the field stations, and so a separate computer was used for data acquisition. The station data, identical to the data received by the PE, is provided to this alternative computer by a parallel modem connection. Unfortunately, because the PE is in control of polling the stations, the data acquisition computer used can only receive the resulting data from the parallel modem asynchronously. The lack of control or synchronization signal means that the data is initially recorded on disk but must later be synchronized and decoded off line.

To get a representative sample that includes both morning and evening commute periods, loop data is recorded for 24-hour periods at up to five sites simultaneously. This information (approximately 1.5 megabytes per station per 24-hour period, or about 10 megabytes per measurement) is transmitted via phone line from TSMC to the UW where it is analyzed. The first step in the analysis is to decode the binary data into volume and occupancy values. To decode the data, the data files must first be synchronized for all the stations being used. The logic behind aligning the data in the 9 byte frame is:

> The "exclusive or" (xor) operation is applied to the first 8 bytes of the data stream and is compared to the 9th byte. If the results are the same the operation is done again to verify the alignment. If the xor result and the 9th byte differ, the entire process slides one byte forward in the data stream (e.g. the 2nd through 9th bytes are examined and compared to the 10th byte.) This process is continued until the two sequential frames align. The alignment is continuously checked during the decoding process, and an error is flagged if the xor result and the 9th relative byte are not identical (this can happen several times in a 24-hour period due to communication

or other errors in the cabinets).

The data acquisition and decoding portion of this project required a software development effort, in particular "C" language routines were developed as follows:

- Custom operating system kernel modifications to deal with the 9-bit data communication.
- Custom terminal line handlers for data transmission to the UW.
- Routines for simultaneous data acquisition on multiple I/O ports.
- Routines to decode the binary data and verify the checksum byte.

The data acquisition requires recording data for 24-hours at one second intervals from up to five loops per site at five sites. This implies each measurement consists of 7.5 megabytes of data and 1.5 megabytes of timing signals that are required to guarantee the synchronization of the data channels. Once recorded at TSMC this data is sent back to the UW over 2400 baud modem connection. The initial physical connection to the sites, data verification, recording and transfer takes a minimum of 2 weeks per site, assuming the loops, amplifiers and TSMC modems are operating properly. In many cases the requirement for one second data from five sites simultaneously and for the entire 24-hour period requires repair of the TSMC infrastructure (e.g. replacing field and base amplifiers and modems). When hardware failures in the TSMC infrastructure occur it can take as much as three weeks to get into the repair schedule and have the repairs accomplished. Once the repairs have been affected data recording can proceed.

Once the data has been transferred to the University of Washington and decoded into one-second volume and occupancy values, it is analyzed using $Matlab^{TM}$ procedures.

3.2.2. Data Reduction

The goals of the data reduction process for this project are:

1. Determine average transit times between loops.
2. Estimate the speed of traffic.
3. Estimate the speed factor g.

This section presents the steps used to accomplish the goals. These functions were implemented using *MATLAB* procedures. Figure 3.3 shows a block diagram of the data reduction process.

The individual lane loop data at one-second intervals is spatially averaged across the highway. The value of spatial averaging is primarily in the low occupancy periods. If a lane-by-lane value is used, the lane volume over extended periods may be very small, while the total traffic flow on the highway (which is what the spatial average measures) may be significant. In the work presented here, a spatial average across the principle lanes (excluding HOV lanes) is used. The one-second volume data from TSMC is then averaged for five seconds. This is done so that there is sufficient information to examine the flow of traffic (e.g. in one second only a few cars can ($V \leq 3$) pass over the loop). The fluctuation about some mean value is the information that is of

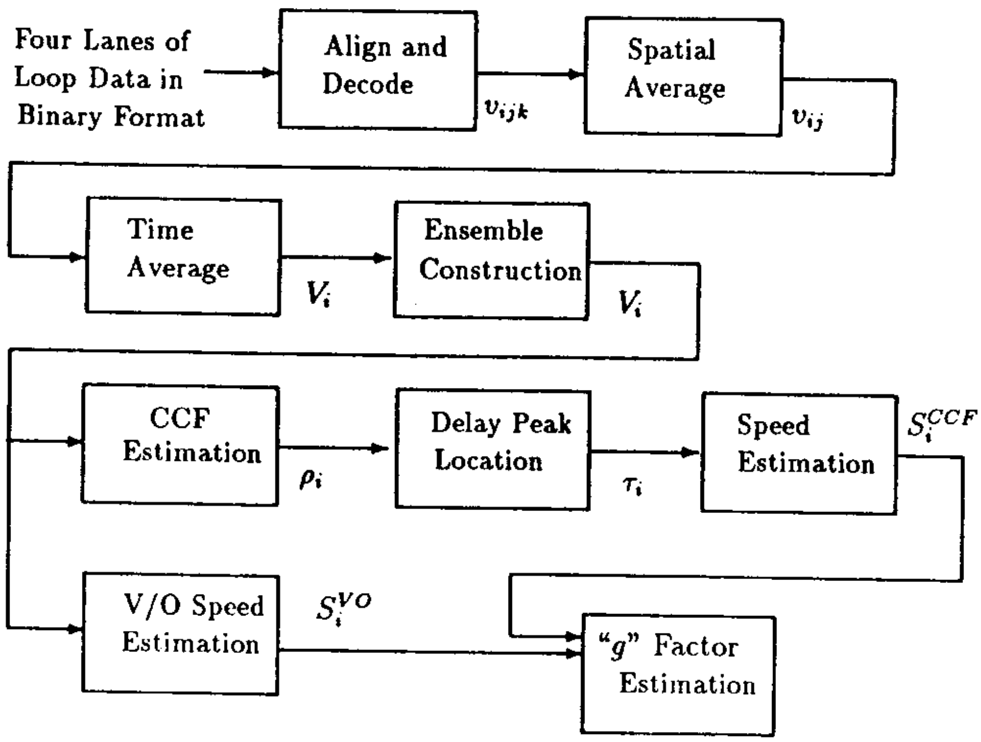

Figure 3.3: Data reduction logic

interest in this study. The time-average provides a realistic mean value for volume. The time-averaged data points are then divided into ensembles, typically 64 points in length. This provides a five minute record of the volume fluctuation at each loop set. The ensemble is then mean centered by forming the mean value over the ensemble and subtracting that value from each point. This results in a zero mean time ensemble at each loop set (see equation 2.4). This time ensemble is used to estimate the CCF (equation 2.5) from which the CCCF (equation 2.6) is derived. The delay peak in the CCCF is located using the technique described in Section 2.1. This delay time is used with the known loop separation (Δx) to make a speed estimate at the ith location and the jth time,

$$S_{ij}^{CCF} = \frac{\Delta x_i}{\tau_{ij}}. \tag{3.1}$$

The speed estimate (S_i^{VO}), based upon the ratio of volume/occupancy, is constructed using time averaged volume and occupancy and RTMIS's value of 2.2 for the speed factor. The volume and occupancy values are produced by averaging over the same five minute period that is used for the CCF technique measurement. Using the two estimates of speed, the speed factor for each period at a site can be estimated,

$$g_{ij} = 2.2 \times \frac{S_{ij}^{VO}}{S_{ij}^{CCF}}. \tag{3.2}$$

This is the improved value of g for the ith site and the jth time. Implicit in the time dependence is the occupancy dependence (since an estimate of occupancy can be made for each time point). This dependence is examined in more detail in the Section 4 Results. A mean value for a site can be made by averaging over all N of the i time samples,

$$\hat{g}_i = \frac{1}{N}\sum_{j=1}^{N} g_{ij} \tag{3.3}$$

in the 24-hour period (N=288). This is the improved estimate of g that can be used by the RTMIS for real time speed estimation.

3.2.3. Summary

This section has summarized the methodology used in this project. The data, as received from the traffic management center inductance loop equipment, contents, and formats were described. The data acquisition and decoding algorithm were presented. The functional blocks necessary to estimate delay time, speed, and speed factor (g) were outlined. The results of this data reduction are presented in the next section.

4. Results

This section presents the results of the data reduction effort. The goal of this effort is to improve speed and travel time estimates made using inductance loop data. These estimates depend on the use of a cross correlation technique; therefore the results of cross correlating inductance loop data to estimate inter-loop travel time is presented first. Second, a mean traffic speed estimate using the cross correlation derived transit time is presented. This speed estimate, which is independent of the mean volume and occupancy values, is then compared to that based on the conventional volume/occupancy ratio technique. This comparison allows for the improvement of the speed factor constant used by the conventional technique. Finally the effect of geographic location on the speed factor is presented.

4.1. Overview

The cross correlation technique used to estimate inter-loop travel times implicitly assumes that the time series being cross correlated have the same time base. As mentioned in subsection 3.1 the system has two type of stations mainline-type and ramp-type each polled by different computers. In examining delay estimates, the mainline-to-mainline and the ramp-to-ramp cross correlations are deemed most reliable, however ramp-to-mainline cross correlations are used when they are the only delay estimates available. Table 4.1[1] shows the numeric labels for the stations, their location (based on the mile marker), direction of traffic (north, south or both) and the type of controller (ramp or mainline). Data from a series of contiguous stations was recorded and processed to estimate the delay times. The set of measurements made is summarized in Table 4.2.

4.2. Delay Estimates

The time series from each loop location in a measurement set can potentially be cross correlated with the other members in that measurement set. This means that if four sites are recorded, there are six possible delay estimates. Figure 4.1 shows this relationship pictorially.

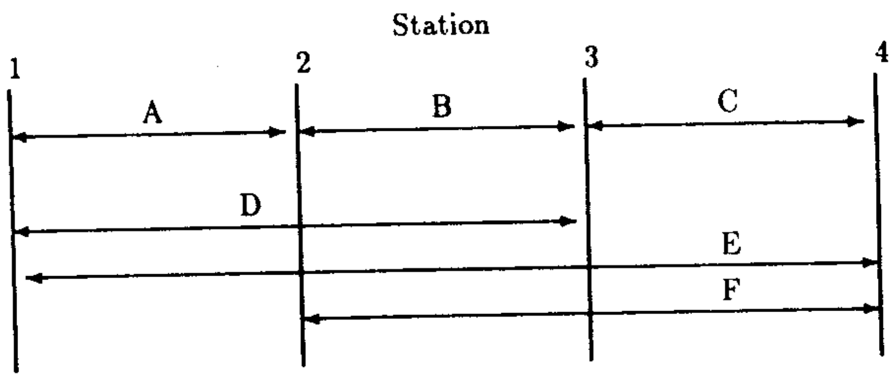

Figure 4.1: Possible delay estimates.

The data presented in this work is based on a five-second average of the time

[1] In the case of stations marked with "*" results are not presented due to equipment failures.

Table 4.1: Inductance loop station description.

Station ID	Mile Marker	Directions	Type
6.5*	165.49	N,S	Mainline
7.0	165.80	N,S	Mainline
8.0*	166.34	N,S	Mainline
9.0	167.00	N,S	Mainline
9.6	167.30	N,S	Mainline
10.0	167.70	N,S	Mainline
10.1	168.00	N,S	Mainline
10.2	168.30	N,S	Mainline
10.3	168.80	N,S	Mainline
10.4	169.20	N,S	Mainline
10.6	169.20	S	Ramp
11.0	169.40	S	Ramp
11.4*	169.20	N	Ramp
11.6	170.00	S	Ramp
12.0	170.25	N,S	Mainline
12.4	170.80	S	Ramp
13.0	170.80	N	Ramp
15.0*	172.15	S	Mainline
15.4	172.15	N,S	Mainline
16.0	172.65	N,S	Ramp
16.2	172.90	S	Ramp
16.4	172.90	N	Ramp
17.0*	173.30	N,S	Mainline
18.0	173.75	N,S	Ramp
18.4*	174.15	N,S	Mainline
18.8	174.60	S	Ramp
19.0	174.60	N	Mainline
20.0	175.10	N,S	Mainline
21.0	175.50	N,S	Mainline
21.6*	176.10	S	Ramp
23.0*	176.70	N,S	Mainline
23.2*	177.20	N,S	Mainline

Table 4.2: Measurement set.

Station ID's	Date
6.5-7.0-8.0-9.0	02/03/92
9.6-10.0-9.0	08/02/91
10.0-10.1-10.2-10.3-10.4	01/06/92
11.0-10.6-10.3-10.2	10/21/91
11.4-11.6-12.0-12.4-13.0	02/18/92
15.0-15.4-16.0-16.2-16.4	04/22/92
16.4-17.0-18.0-18.4-18.8	05/26/92
18.8-19.0-20.0-21.0-21.6	06/11/92, 06/24/92
23.0-23.2	11/28/91

dependent volume counts. It requires approximately five minutes of data to construct the ensembles for use in equation (2.5). However, the next time delay estimate can be made as soon as the next five-second average volume count is available. This means that the resolution of the time delay estimate is equal to the averaging time of five seconds. The fine time resolution possible with the cross correlation technique provides a mechanism to estimate mean vehicle speed on a near real time basis. For illustrative purposes, this report presents the delay estimates that do not overlap in time and are on a time scale more appropriate to examine large scale traffic phenomena.

For the purposes of demonstration in this report, the cross correlation coefficient was constructed and used to produce delay estimates every five minutes for 24-hour periods. An example of these time dependent delay estimates for one site is shown in Figure 4.2. The actual analysis produces such a delay verses time history for every possible pairing of stations (e.g. there are 62 possible such time histories for the measurement set outlined in Table 4.2). However, delay estimates from widely separated stations (e.g. separations grater than 1.5 miles) are unlikely to produce valid delay estimates.

The validity of the delay estimate is related to the maximum value in the CCF. In the work presented here, the delay estimate is deemed invalid if the CCF peak value is less than (0.4). The CCF maximum values for the same example time period are shown in Figure 4.3. A significant number of the CCF estimates do not meet the criteria mentioned above and are not used in any subsequent calculations. These delay estimates are used with the interloop distances to provide speed estimates.

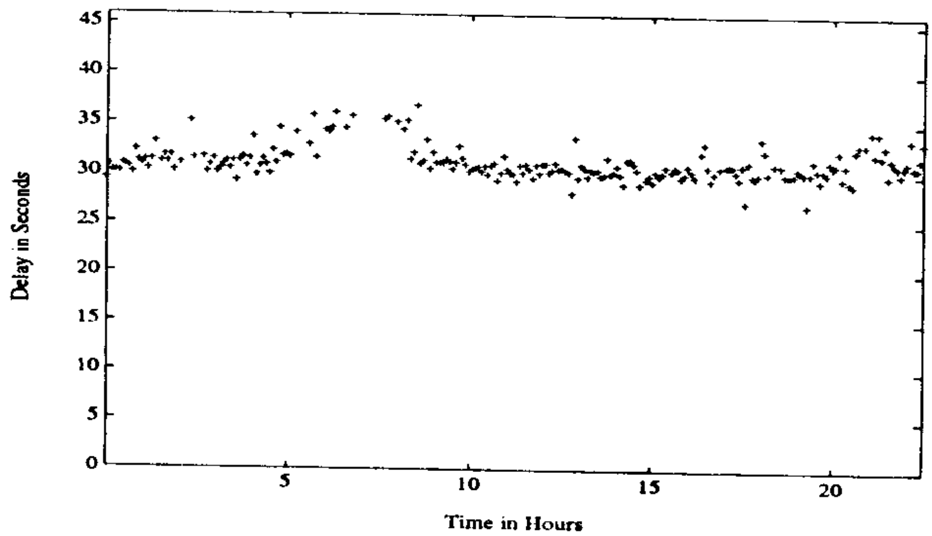

Figure 4.2: Delay verses time.

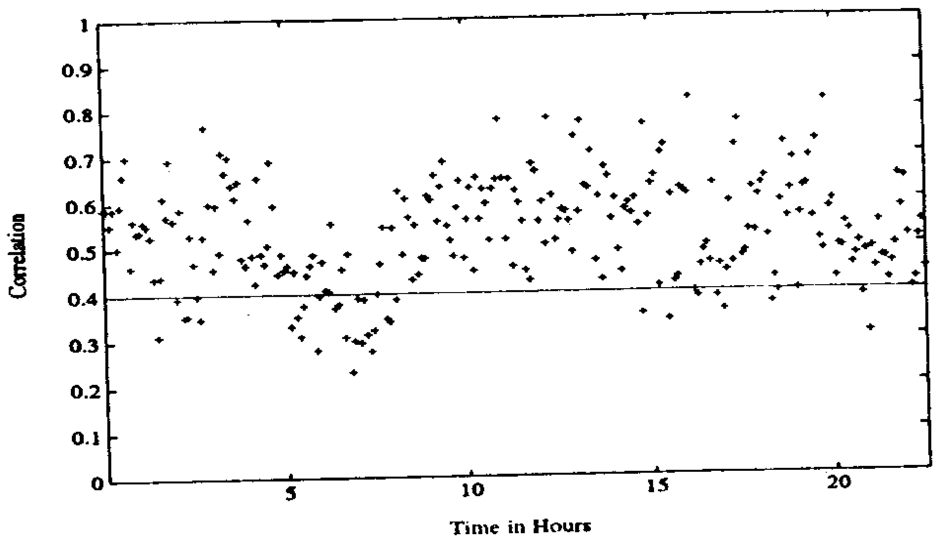

Figure 4.3: Maximum value in the cross correlation coefficient function.

4.3. Speed Estimates

The mean traffic speed is estimated using the interloop distance and the cross correlation travel time measurement in equation (2.13). The speed estimates for the same 24-hour period presented above are shown as "+" symbols in Figure 4.4. The speed estimates for each site using the "g" value accepted by the local traffic management system are also shown as solid lines in figure 4.4. From the comparison of the volume/occupancy based speed estimate to the measured speed in Figure 4.4 it can be concluded that the volume/occupancy ratio with the existing constant (based on an assumed vehicle length) does not accurately predict the mean speed. The speed results suggest that the use of the cross correlation technique to estimate a mean speed factor could provide a mechanism to improve speed estimates available from single inductance loops. The utility of the cross correlation technique does depend on certain assumptions, the most important being the "rigidity" assumption. The rigidity of the traffic flow is directly related to occupancy. As occupancy increases, the traffic flow becomes compressible; and the fluctuation about the mean volume count will not propagate rigidly between the sites. The time dependent occupancy value for each of the sites is shown in Figure 4.5. A comparison of Figures 4.3 and 4.5 reinforces the idea of the occupancy dependence of the rigidity assumption. Figure 4.6 explicitly demonstrates this relationship and is based on road data that is a composite of a number of days of measurements that estimate $\hat{\rho}(\tau_0)$ over a range of occupancies. It is clear that at high occupancies, the 0.4 criteria placed on the CCF maximum value is unlikely to be met. It can be concluded that at high occupancies the CCF maximum value is reduced, and eventually reaches a point where the delay peak can no longer be readily identified. Since the maximum value of the CCF depends on the rigidity assumption, the higher the occupancy of the highway, the less valid the assumption becomes. The cross correlation technique provides a speed estimate independent of the mean volume and occupancy values but is constrained to operate in regions of modest occupancy.

An improved speed factor estimate for use by traffic management and information systems can be made by combining the speed estimate from the cross correlation technique with equations (3.2) and (3.3). Using equation (3.2), the speed factor as a function of time, for the demonstration data, is presented in Figure 4.7. The apparent time dependence of the speed factor is actually related to the changes in occupancy over time. Figure 4.8 shows the speed factor as a function of occupancy. The mean values of $g_1 = 2.0$ and $g_2 = 3.0$ from equation (3.3) are 10% lower and 36% higher than the 2.2 value presently used by the traffic management and information systems. It is clear that a constant value for the occupancy to density conversion coefficient (or speed factor) is an approximation that is invalid at high and very low occupancies [HP88]; however, it is an approximation used by traffic management and information systems. As such, having the best estimate for this coefficient for each inductance loop site can improve the speed estimates available to travelers and traffic management personnel.

With quantitative knowledge of the occupancy to density conversion (or speed factor), speed estimates that use the mean values of the volume and occupancy are possible. Since the inductance loop measurements are time averaged volume and occupancy, the ability to use these quantities to accurately estimate speed is of great value to traffic management systems that rely heavily on inductance loop sensors.

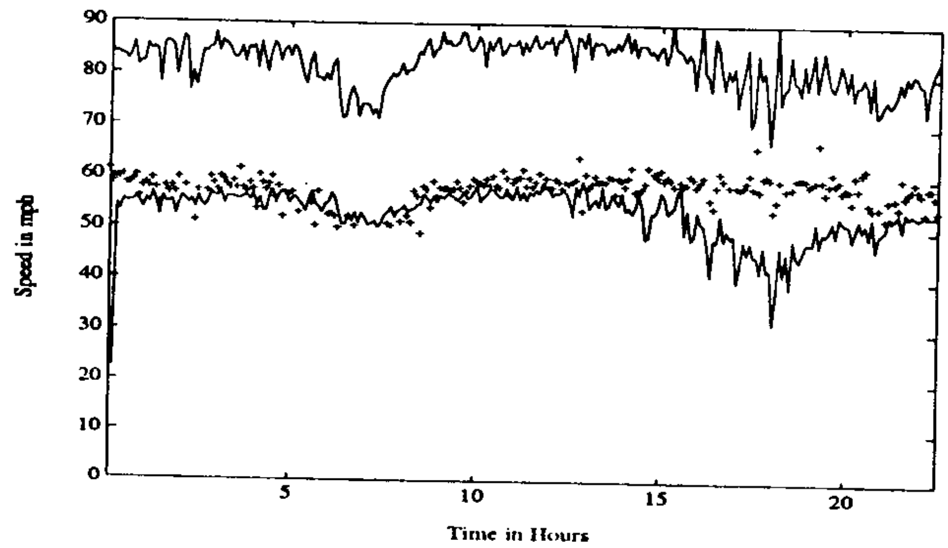

Figure 4.4: Speed as a function of time.

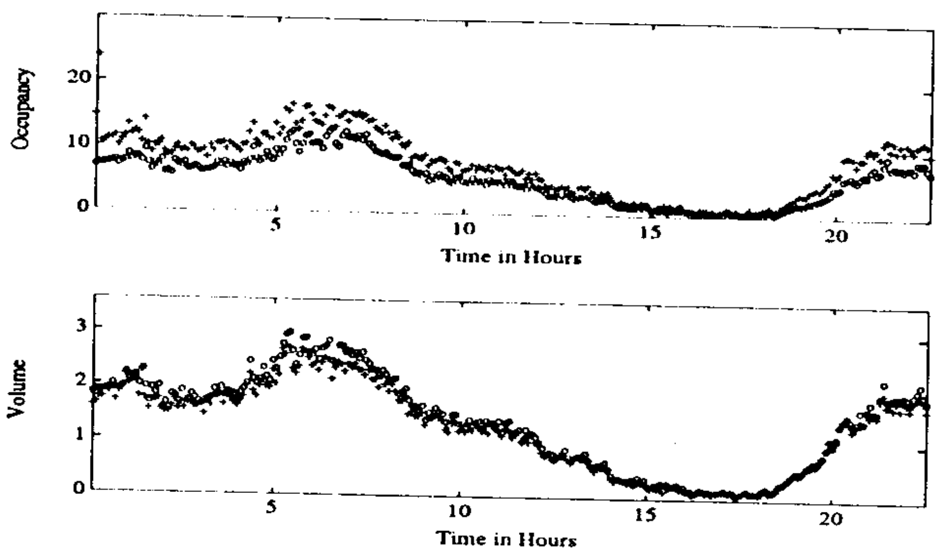

Figure 4.5: Occupancy and volume as a function of time.

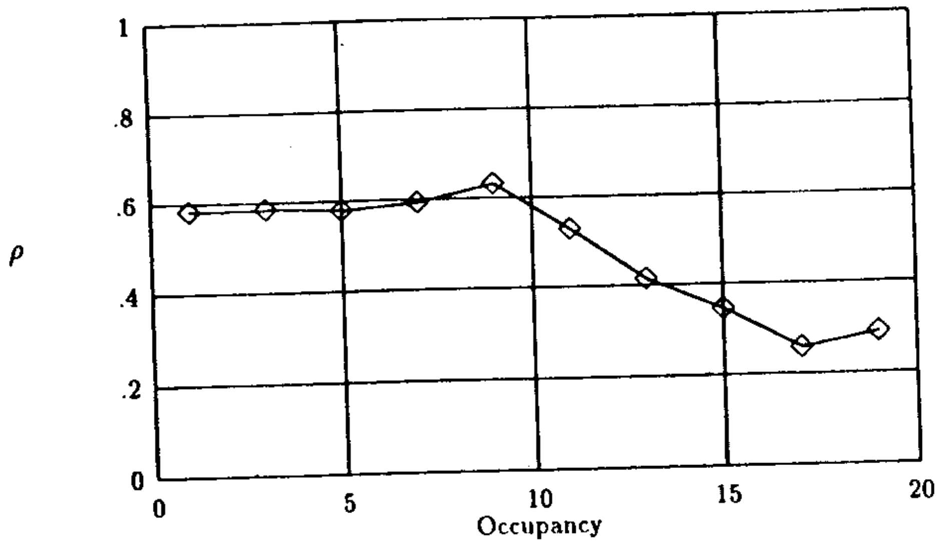

Figure 4.6: Correlation coefficient as a function of percent occupancy.

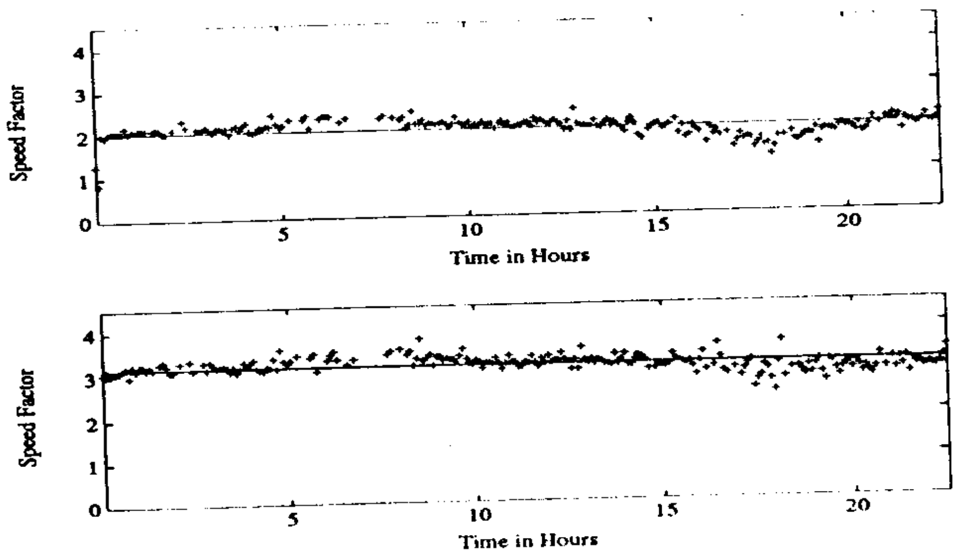

Figure 4.7: Cross correlation coefficient as a function of time.

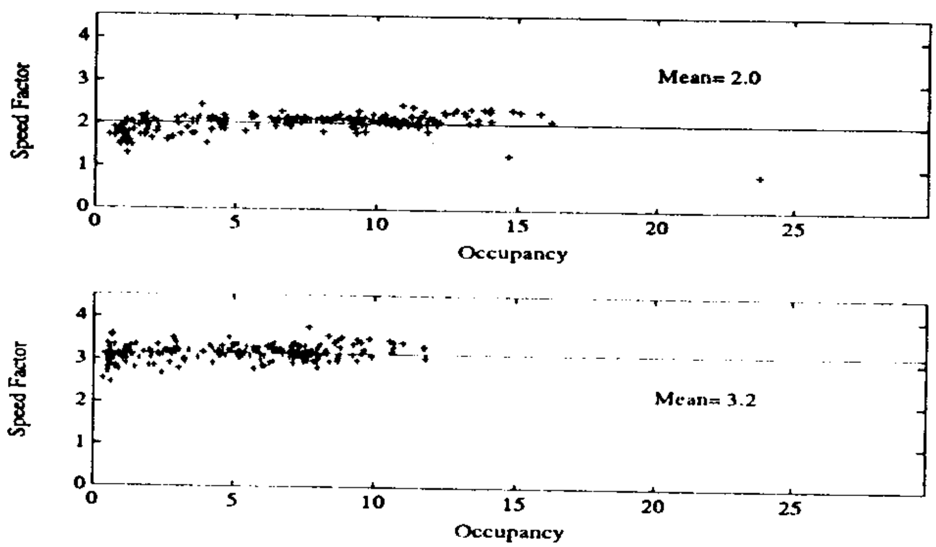

Figure 4.8: Cross correlation coefficient as a function of occupancy.

4.4. Speed Factor as a Function of Location

The work just presented provides a method to estimate the occupancy to density conversion factor, or speed factor. This factor has different values at different geographic locations on the highway. To use the volume and occupancy data at a particular set of loops, the value for the speed factor for that site must be determined. In this work, 23 sites (north and south bound stations where available) are examined and the value of g is estimated. Figure 4.9 displays the estimated mean g values for the sites at which measurements were made. The station ID number on the ordinate axis matches the values from the first column of Table 4.1. Table 4.3 lists the northbound and southbound mean g values for the stations considered in the report. In examining Table 4.3, two items must be considered: (1) some stations have loops only in the northbound or southbound lanes; (2) the value for g is present only if the number of observations in which $\hat{\rho}(\tau_0) > 0.4$ is greater than 25. The second criterion removes estimates that come from station pairs that do not produce a valid delay peak for a large fraction of the observation time.

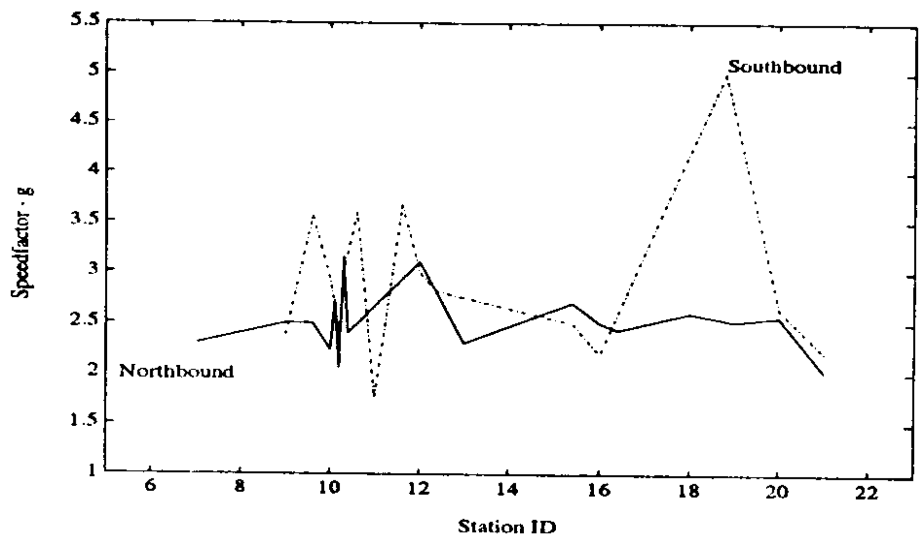

Figure 4.9: Estimated mean speed factor (g) as a function of location.

Table 4.3: Speed factor as a function of location.

Station ID	Northbound "g"	Southbound "g"
7.0	2.3	-
9.0	2.5	2.4
9.6	2.5	3.6
10.0	2.2	2.9
10.1	2.7	2.7
10.2	2.1	2.3
10.3	3.2	3.0
10.4	2.4	-
10.6	-	3.6
11.0	-	1.8
11.6	-	3.7
12.0	3.1	3.0
12.4	-	2.8
13.0	2.3	-
15.4	2.7	2.5
16.0	2.5	2.2
16.2	-	2.4
16.4	2.4	-
18.0	2.6	-
18.8	-	5.0
19.0	2.5	-
20.0	2.6	2.6
21.0	2.0	2.2

5. Conclusions and Recommendations

This section presents conclusions concerning (1) the validity of the estimation of inter-loop travel time and speed using cross correlation techniques, and (2) the conclusions about the occupancy to density conversion. Finally, recommendations on methodologies for the use of speed factors and loop data with traffic information systems are made, and future possible work is suggested.

5.1. Conclusions

Conclusion #1. The cross correlation technique produced valid CCF peak values only in the low (0-15%) range of occupancy values. This is true because the notion of using the CCF estimate depends on the assumption that disturbances to the mean value of the volume count propagate rigidly between the measuring stations. During high occupancy periods the perturbations to the traffic volume count do not propagate rigidly, and therefore the cross correlation technique has a limited range of occupancies over which it is applicable. However, within the lower occupancy range, the cross correlation technique provides a valuable tool for estimating mean speed independent of the mean volume and occupancy values.

Conclusion #2. The volume/occupancy speed estimate used by traffic management and information systems depend on the conversion of occupancy to density. Using the classic flow/density/speed relationship, the conversion between occupancy and density can be made if the speed can be independently estimated. The cross correlation speed estimate is such an independent measurement in that it does not depend on the mean value of volume or occupancy. This report presented the value of the occupancy to density conversion ("g") determined using the observed volume, occupancy, and speed. The value of speed factor was shown to vary with time as a result of the change of occupancy with time. The value of speed factor as a function of occupancy was presented. It was concluded that in the 5-10% occupancy range, a mean value is representative of the "g" to occupancy relationship. It was also concluded that outside this range a single mean value was an inadequate representation of the "g" occupancy relationship. The existing traffic management and information systems in Seattle use a single value for the occupancy density conversion, and the mean value of "g" will provide an improved estimate of speed using volume and occupancies.

Conclusion #3. The value for the occupancy to density conversion constant (speed factor) depends on the topography of the highway. This in turn means that the value for the speed factor is site dependent. This site dependency was considered by calculating the speed factor at a number of sites. Mean values for the speed factor were calculated for 17 northbound and 17 southbound locations. These values should provide improved speed estimates for management and information systems using single loop volumes and occupancies to estimate traffic speeds. While these mean values are actually only valid over the 5-15% range of occupancy, they do provide a geographically specific value for the mean value of the speed factor.

5.2. Recommendations

Recommendation #1. The cross correlation technique speed estimate is reliable only in low to modest occupancy ranges. Some other technique is needed to provide the occupancy to density conversion at high occupancies if the volume/occupancy ratio speed estimate is to be the basis for traveler information systems. Preliminary research shows that a predictor technique such as a Kalman predictor with speed as one of the state variables might be the next logical step. This predictor would use the occupancy/density conversion applicable to the modest occupancy range to estimate the observed speed but would used the filtered estimate of the speed in updating the state variables. This would provide a mechanism that produces a speed and occupancy-to-density conversion that is time dependent and is driven by the recent history of traffic conditions.

Recommendation #2. Only if a new technique that can provide a reliable estimate of speed over a wider range is developed and this technique is dependent upon the calibration possible using the cross correlation technique should further calibration efforts be made.

BIBLIOGRAPHY

[Dai91] D.J. Dailey. Travel time estimation using cross correlation techniques-Final Report. Technical Report TNW91-02, **TransNow** - University of Washington, 1991.

[HP88] F.L. Hall and B.N. Persaud. Estimating speeds from freeway traffic management systems. *ITE 1988 Compendium of Technical Papers*, pages 166–171, 1988.

[LW55] M.J. Lighthill and G.B. Whitham. On kinematic waves, a theory of traffic flow on long crowded roads. *Proc. of the Royal Society of London*, 229:317–345, 1955.

[MB89] M.Haselkorn and W. Barfield. Understanding commuter behavior for the design of motorist information systems. In *69th Annual Meeting of the Transpn Res. Board, Paper N. 890800*, 1989.

[MJMB91] M.Haselkorn, J.Spyrdakis, B.Gobel M.Garner, and W. Barfield. Implementing a pc based graphical, interactive, real-time travler information system that meets commuter needs. In *Proceedings of Vehicle Navigation & Information Systems (VNIS'91), Deerborn:IEEE*, pages 1045–1048, 1991.

[PH88] B.N. Persaud and V.F. Hurdle. Some new data that challenge some old ideas about speed-flow relationships. *Transpn Res. Rec.*, 1194:191–198, 1988.

[War52] J.G. Wardrop. Some theoretical aspects of road traffic res. *Proc. of the Institution of Civil Engineers*, 3:158, 1952.