Final Report
TNW91-02

TRAVEL TIME ESTIMATION USING CROSS CORRELATION TECHNIQUES

by

Daniel J. Dailey
Asst. Professor
Electrical Engineering
(Principal Investigator)

Mark P. Haselkorn
Professor
Technical Communication

Nancy L. Nihan
Professor
Civil Engineering

University of Washington
Seattle, Washington

Transportation Northwest (TransNow)
Department of Civil Engineering
135 More Hall, FX-10
University of Washington
Seattle, WA 98195

October 1991

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. This document is disseminated through Transportation Northwest (TransNow) Regional Center under the sponsorship of the Department of Transportation UTC Grant Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. Sponsorship for the local match portion of this research project was provided by the Washington State Department of Transportation and the University of Washington. The contents do not necessarily reflect the views or policies of the U.S. Department of Transportation or any of the local sponsors. This report does not constitute a standard, specification, or regulation.

Abstract

This project demonstrates the viability of using cross-correlation techniques with inductance loop data to measure the propagation time of traffic. The propagation time between loops separated by 0.5 miles is measured using time averaged activation counts ("Volume") from inductance loops. The resulting time delay is used to estimate the mean speed. This independent speed estimate is used to improve the Volume/Occupancy ratio estimate of speed. A functional form relating speed to the Volume/Occupancy ratio is developed and evaluated using real traffic data. This result can be used to improve the total travel time estimates for commuters in the I5 corridor in Seattle, WA.

CONTENTS

List of Figures

1 . Introduction

Inductance loops are the primary data source for real-time monitoring of traffic conditions on Puget Sound freeways. However, the use of loop data for speed and travel time estimates has been a subject of controversy. This work validates the use of loop data for time delay estimates. This effort uses time series techniques to measure the transit time between loops in real traffic conditions. From these transit time estimates speed estimates are developed. These speed estimates are compared to those arrived at from a commonly used Volume/Occupancy ratio technique. [1] [HP88] Finally, travel times estimates are used to correct the method incorporated into a motorist information system. [MB89]

The objectives of this report are to:

1. Demonstrate that transit times can be estimated from inductance loops which have a large relative separation (approximately 0.5 mile) for using as a speed trap.
2. Present an improved method for speed estimation from inductance loop volume and occupancy data.
3. Present improved total travel time estimates from inductance loop data.
4. Describe how the improved travel time estimates can be used by a motorist information system.

The specific tasks accomplished by this project were:

1. Obtain loop data from Traffic Systems Management Center (TSMC) at varying sample rates for the same time period. This to be done in an area with as little dispersion as possible (e.g. a length of highway with no on/off ramps).
2. Make travel time estimates of individual vehicles using license plate number collection over the interloop region(To be done over the same time period that loop data is recorded).

[1] Volume is the number of vehicles passing a point in a a period of time (see 2.2.1 for more detail) and Occupancy is the fraction of some total time that a specific area of highway has a vehicle present (see 2.2.2 for more detail)

3. Produce travel time estimates using cross-correlation techniques on the loop data (at all sampling rates).

4. Establish the relationship between the "particulate" or single vehicle travel time estimates and the time averaged cross-correlation estimates.

5. Develop the Volume/Occupancy ratio to speed relation.

6. Identify minimum sample rate for reliable results.

7. Test minimum sample rate at other locations with favorable geography.

8. Develop the Volume/Occupancy ratio relation for all areas tested.

9. Produce initial final report.

10. Implementation of results, and revision of final report.

Each of these specific tasks was accomplished to fulfill the goals of this project.

This project had several phases involving loop data from TSMC. The loop data is available at varying frequencies (e.g. 1 second, 20 seconds, 1 minute, 5 minute). Since the goal is to estimate time delays of order 30 seconds the one second data is the most desirable sampling rate for the proof of principle aspect of this project. Operations and control is most often done using data of larger granularity (e.g. 1 to 5 minute averages). Based on the operations and control aspects the applicability of this cross-correlation technique will also be tested on the time averaged data at 20 and 60 second scales.

This report presents the results of this project in four sections. First the underlying theory is presented. The theory for the use of cross-correlation delay estimates is presented along with a method to quantify errors in these estimates. The conventional method for speed estimates using inductance loop data is reviewed, and a functional form for a "speedfactor function" is developed. Second, the methodology for obtaining inductance loop data is described and the data reduction methodology is discussed. Third results are presented. The results include a demonstration of principle, an initial evaluation of the operational envelope, and the development of a volume/occupancy to speed correlation. Finally conclusions are drawn about the effectiveness of the method in comparison to conventional methods. The work described here demonstrates that travel time and speed estimates based on loop data can be made reliable and useful.

2. Theory

This chapter has three sections that present the background theory used to estimate speed and travel time based on inductance loop sensors. The first section develops a new continuous model for traffic flow. This model is based upon the rigid propagation of fluctuations from the mean vehicle/time count. The model is used to develop the form of time series descriptors that can be used to measure time delay between widely displaced inductance loops. The second section describes the conventional single loop method for estimating speed which is based upon a ratio of volume and occupancy values and mentions the short comings of such a technique. The third section develops a new function that when calibrated using onroad data, can be used with the conventional single loop method to provide improved speed estimates.

2.1. Cross-correlation Transit Time Estimation

This section introduces a new model for traffic flow and develops statistical function for such a model. These statistical function are then used to estimate travel time. The key to estimating inter-loop time delay for loops separated by a relatively large distance, is the use of a time averaged estimation process. Traffic propagates down the highway with an average concentration about which there is some statistical fluctuation. The value of the concentration at some point down stream (x_2) is a function of several contributions:

- the concentration upstream (x_1) which has propagated to the new point
- an attenuation factor for the dispersal of that upstream concentration (b)
- the change in concentration from addition or removal of cars ($\Delta\alpha(\Delta x, x_1, t)$)
- noise or uncorrelated fluctuations ($n(t)$)

and is written

$$\alpha(x_2, t) = b\alpha(x_1, t + \tau) + \Delta\alpha(x_1, x_2, t) + n(t). \tag{2.1}$$

Representing the concentration as the sum of a mean value α_0 and a fluctuating component

$$\alpha(x, t) = \alpha_0(x) + \delta\alpha(x, t). \tag{2.2}$$

This work is concerned only with the fluctuation component, the mean value is removed, and the fluctuating component of the down stream time series is written:

$$\delta\alpha(x_2, t) = b\delta\alpha(x_1, t + \tau) + \Delta\alpha(x_a, x_2, t) + n(t). \tag{2.3}$$

The auto-correlation function (ACF) of this quantity is written:

$$R_{11}(\tau, T) = \frac{1}{2T}\int_{-T}^{T} \delta\alpha(x_1, t)\delta\alpha(x_1, t - \tau)dt \tag{2.4}$$

This function has the property that the value of the auto-correlation at zero lag time is the variance squared,

$$\sigma_1^2 = R_{11}(0). \tag{2.5}$$

The correlation coefficient function

$$\rho_{11}(\tau) = \frac{R_{11}(\tau)}{\sigma_1^2} \tag{2.6}$$

is an indication of the information in the time series that is correlated at varying lag times (e.g. at zero lag the correlation coefficient is unity indicating that the time series is completely correlated with itself). [BP86]

The cross correlation function (CCF) between the two time series is written:

$$R_{12}(\tau, T) = \frac{1}{2T}\int_{-T}^{T} \delta\alpha(x_1, t)\delta\alpha(x_2, t - \tau)dt. \tag{2.7}$$

Substituting equation (2.3) into equation (2.7),

$$\begin{aligned} R_{12}(\tau, T) &= \frac{b}{2T}\int_{-T}^{T} \alpha(x_1, t)\delta\alpha(x_1, t - \tau_0 - \tau)dt \\ &+ \frac{1}{2T}\int_{-T}^{T} \alpha(x_1, t)\Delta\alpha(x_1, x_2, t - \tau)dt \\ &+ \frac{1}{2T}\int_{-T}^{T} \delta\alpha(x_1, t)n(t - \tau)dt. \end{aligned} \tag{2.8}$$

Note that the first term is the auto-correlation function shifted in time by τ_0 and multiplied by a scale factor. So

$$\begin{aligned} R_{12}(\tau, T) &= bR_{11}(\tau - \tau_0, T) \\ &+ \frac{1}{2T}\int_{-T}^{T} \alpha(x_1, t)\Delta\alpha(x_1, x_2, t - \tau)dt \\ &+ \frac{1}{2T}\int_{-T}^{T} \delta\alpha(x_1, t)n(t)dt. \end{aligned} \tag{2.9}$$

If there are no "on" or "off" ramps and no noise or loop failure the ideal cross-correlation is the auto-correlation displaced by a time lag of τ_0 and scaled by the dispersion factor b:

$$R_{12}(\tau, T) = bR_{11}(\tau - \tau_0, T) \tag{2.10}$$

This states that the cross-correlation function maximum takes place at a time delay τ_0. Further the fraction of the informative which is common to the two time series is the correlation coefficient function

$$\rho_{12}(\tau) = \frac{R_{12}(\tau, T)}{\sigma_1 \sigma_2} \tag{2.11}$$

$$= \frac{bR_{11}(\tau - \tau_0, T)}{\sigma_1 \sigma_2}. \tag{2.12}$$

In the ideal case the value of the correlation coefficient function is a maximum at τ_0 and the value of this maximum represents the dispersion or lack of rigidity in the propagating traffic.

By constructing the correlation coefficient function it is possible to use the correlation coefficient value ($\rho(\tau_0)$) as an indication of the commonality of information in the time series. Further the location of the maximum (τ_0) is the mean time for traffic to propagate between the stations. This mechanism provides a time delay estimate and a check on the validity of such an estimate.

The estimates of the covariance function are being made using discrete information. There are errors in the estimates of the magnitude of the covariance and in the location of the lag time. The normalized mean square error for the cross-correlation function is [1].

$$\epsilon\left[\hat{R}xy(\tau)\right] \approx \frac{1}{\sqrt{2BT}}\left[1 + \rho_{xy}^{-2}(\tau)\right]^{1/2}. \tag{2.13}$$

This provides a method of evaluating the error in the estimate of the magnitude of the covariance function that depends upon the bandwidth of the signal (B) and the total time used in the cross-correlation estimate (T). The value of T is a parameter selected during data analysis, the bandwidth B is a property of the time series and must be determined.

To estimate the error in the estimate of the time delay an assumption about the functional form of the cross-correlation function near the delay time is necessary. Specifically the timeseries is assumed to be approximately band limited white noise over some bandwidth. This suggests a functional form for the cross-correlation function,

$$R_{xy}(\tau - \tau_0) = R_{xx}(0)\left(\frac{sin 2\pi B(\tau - \tau_0)}{2\pi B(\tau - \tau_0)}\right). \tag{2.14}$$

[1] Equation 8.103 on page 273 [BP86]

This form provides a mechanism to estimate the error in the time lag and to estimate B based upon fitting this functional form to the cross-correlation function near the maximum value. The 95 % confidence interval for determining the location where the peak value occurs is [2]

$$\left(\frac{-2}{\pi B}\right)\left(\frac{3}{4}\right)^{1/4}\left\{\epsilon\left[\hat{R}_{xx}(-\tau_0)\right]\right\}^{1/2} \leq \tau \leq \left(\frac{2}{\pi B}\right)\left(\frac{3}{4}\right)^{1/4}\left\{\epsilon\left[\hat{R}_{xx}(-\tau_0)\right]\right\}^{1/2}. \quad (2.15)$$

This quantifies the errors in estimating the time delay using cross-correlation techniques.

The speed can be estimated using the time delay and the assumption that the speed is approximately constant over the segment between the inductance loops,

$$S = \frac{D}{\tau_0}. \quad (2.16)$$

This speed estimate can be used for evaluating the accuracy of the conventional volume/occupancy ratio method described in the next section.

This section has presented a new method for estimating time delay and speed that is independent of the actual values of occupancy or volume. The next section reviews the conventional method which depends on the mean value of the occupancy and volume.

[2] Equations 8.124 and 8.125 page 277 of [BP86]

2.2. Conventional Speed Estimates

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. TSMC presently constructs estimates of these quantities from volume, occupancy and a constant factor.

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

$$s = \frac{q}{k} \tag{2.17}$$

where s is the vehicle speed in miles per hour, q is the traffic flow in vehicles per hour, and k is the traffic concentration in vehicles per mile.

2.2.1. Volume to flow conversion:

The relationship between the first quantity measured by TSMC, labeled volume (V), and traffic flow is straightforward. Volume is the number of actuations of a loop per unit time (five minutes in TSMC practice) and an actuation is considered by TSMC to be a vehicle passing over the loop. Therefor a constant conversion factor is all that is needed,

$$q = \frac{60}{T} \times V, \tag{2.18}$$

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

2.2.2. Occupancy to density conversion:

The second value available from TSMC is called Occupancy (O) and 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) times 100. The conversion from Occupancy to flow is less straightforward.

The time in the region of the detector per vehicle 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:

$$T_O = \int_{l_d} \frac{1}{s(x)}dx + \int_{l_s} \frac{1}{s(x)}dx. \tag{2.19}$$

The "total time" in a region of length D is:

$$T = \int_{D} \frac{1}{s(x)}dx + \int_{l_s} \frac{1}{s(x)}dx. \tag{2.20}$$

Occupancy per vehicle, as defined by TSMC is:

$$O = \frac{\int_{l_d} \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx}{\int_D \frac{1}{s(x)} dx + \int_{l_v} \frac{1}{s(x)} dx} \tag{2.21}$$

where the "total distance " in the denominator is normalized to one mile. The speed is assumed constant over the sensing region resulting in:

$$O = \frac{\frac{1}{s}(l_d + l_v)}{\frac{1}{s}(D + l_d)} \tag{2.22}$$

The relationship between traffic flow and occupancy is assumed to be:

$$k = g \times O \tag{2.23}$$

A value for g is determined assuming:

$$\begin{aligned} s &= constant \\ l_d &= 8(feet) \\ l_v &= 16(feet/vehicle) \\ D &= 5280(feet/mile). \end{aligned}$$

The value $g = 220$ (vehicles per mile) converts the fractional occupancy to flow. Using these two constants the relationship between the volume/occupancy ratio and speed can be developed.

2.2.3. Implementation

At the present time WA Traffic Systems Management Center (TSMC) is using a single constant to define a linear relationship between speed and the ratio of volume (V) to occupancy (O).

$$s = \frac{V}{gO} \tag{2.24}$$

From the literature [HP88] and observations of TSMC data, the relationship between speed and the Vol./Occ. ratio is only linear over a limited speed range. To address this concern the functional relationship between speed and Vol./Occ. must be more accurately calculated. This effort will empirically determine the relationship between speed and the Vol./Occ. ratio and define a function $g(Occ.)$ which will properly convert the ratio to speed.

$$s = \frac{V}{g(O)O} \tag{2.25}$$

Much of the operating envelope of the traffic management system is outside the region of where the speed vs Vol./Occ. is linear. Outside of the zone of linearity, estimates of speed are incorrect. To address this shortcoming of previous efforts this report suggests a functional form for the relationship between the actual speed and the volume to occupancy ratio. This new "speedfactor" is presented in the next section.

2.3. "Speedfactor" Functional Form

Previous researchers have noted a relationship between flow and concentration. This relationship has the general form of an inverted parabola. [LW55] This same relationship applies to volume and occupancy. An example of the volume occupancy plane is shown in figure(2.1). The data for this figure comes from I5 northbound. It consists of

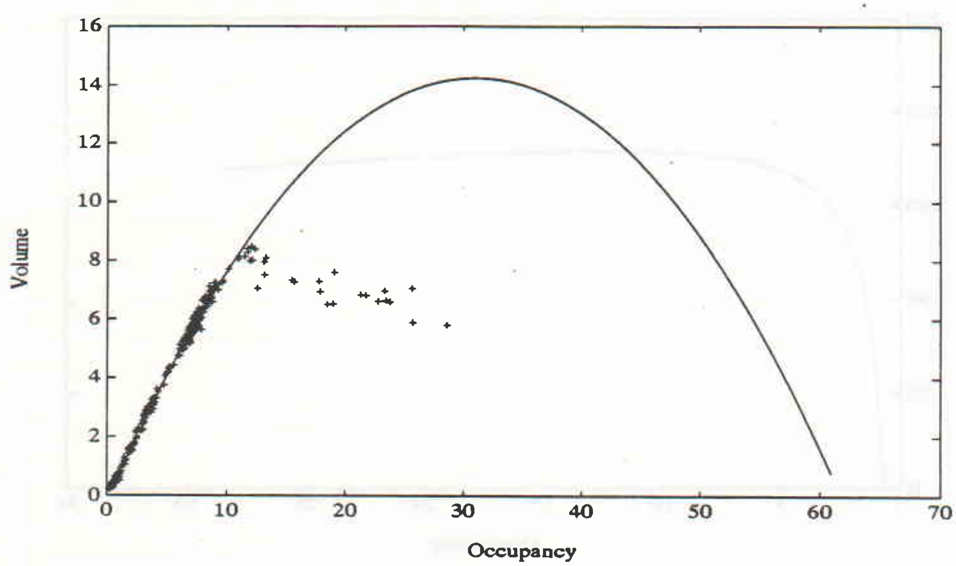

Figure 2.1: Volume-Occupancy plane with best fitting parabola based upon five second averages made over a 24 hour period.

five second average estimates of volume and occupancy made at five minute intervals over a 24 hour period. The parabola shown is a minimum least squares fit to the observed values. The maximum value of occupancy is a concentration of 200 vehicles per mile which corresponds to other observations for traffic jams on long straight roads. [Gre35] This suggests that the functional form of the volume (V) vs occupancy (O)

equation is,

$$V \approx aO^2 + bO + c. \tag{2.26}$$

Since the speed is estimated from a ratio (see equation 2.25) a functional form for the speedfactor can be constructed by substituting equation 2.26 into equation 2.25 to get,

$$g(O) \approx \frac{1}{s(O)}\left(aO + b + \frac{c}{O}\right). \tag{2.27}$$

This form is still dependent upon the speed which at larger occupancies must be a function of occupancy. Figure(2.2) shows the function $g(O)s(O)$ for the parameters from the flow vs concentration parabola derived from figure(2.1). This provides an

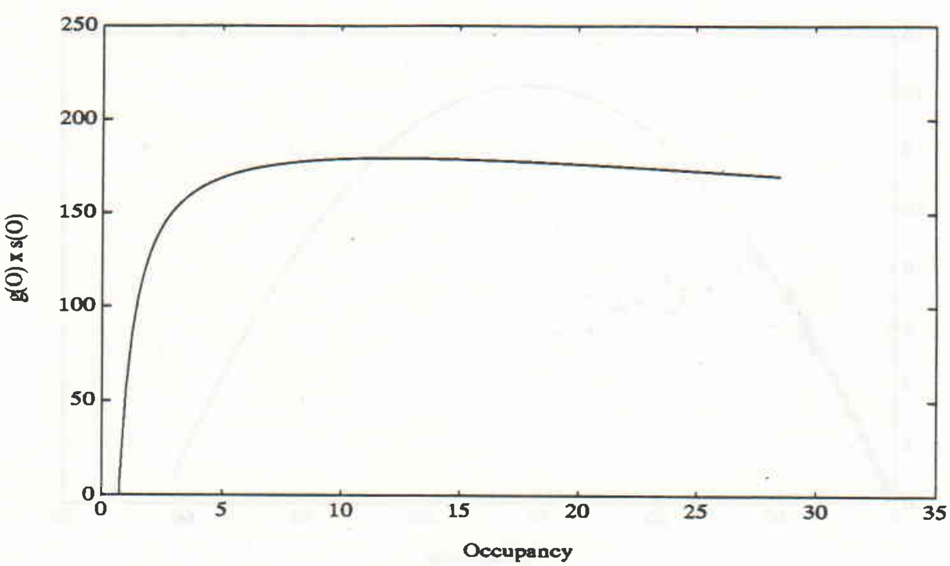

Figure 2.2: Speedfactor function $g(O)$ times speed ($s(O)$) for the parameters from a parabolic fit of flow vs concentration.

indication of the functional form for the $g(O)$ function but is biased by the speed. Since at low occupancies the speed is nearly independent of occupancy the shape of this function below perhaps 10% occupancy is predictive of the results to be expected experimentally. The general form of this function will be used with experimental data to derive the empirical $g(O)$ in chapter 4.

2.3.1. Summary

This chapter presented several ideas. First the underlying theory, on which the new method of speed estimation developed for this project is based, was presented. This theory is based on a continuous model of traffic flow. Second, the conventional method used for speed estimation based upon volume and occupancy was presented. The shortcomings of the conventional method is discussed. Third, a function that when calibrated using the new speed estimation method will provide improved volume to occupancy ratio speed estimates is proposed. These sections have presented a theoretical framework. To apply these ideas to real traffic situations several additional steps must be taken. Real traffic data must be obtained and the discretized version of the functions developed in section (2.1) must be constructed. The following chapter describes the data acquisition methodology and presents the discrete version of the ideas from this chapter that can be used with real traffic data.

2.4.1 Summary

This chapter presented several ideas. First, the underlying theory on which the new method of speed estimation developed for this project is based, was presented. This theory is based on a continuum model of traffic flow. Second, the current single loop method used for speed estimation based upon variance and occupancy was presented. The shortcomings of this conventional method is discussed. Third, a function that when calculated using the new speed estimation method will provide improved [illegible] to accurately make speed estimates is proposed. These sections have presented a theoretical framework. To apply these ideas to real traffic conditions several additional steps must be taken. Real traffic data must be acquired and the discretized version of the function developed in section [illegible] must be constructed. The following chapter describes the data collection methodology and presents the discrete version of the ideas from this chapter that can be used with real traffic data.

3. Methodology

This project required the acquisition of onroad data, as well as the reduction and analysis of this data. To demonstrate the viability of this cross-correlation technique real data from the onroad cabinets at 185th and 195th on I5 North, in the southbound direction, was recorded. This chapter describes the mechanics of the data collection and reduction to volume and occupancy as a function of time.

3.1. Data Formats and Contents

This section gives an overview to the data collection aspect of the project. At 33 sites on I5 north there are inductance loops and "cabinets" containing microprocessors. 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 micro processors sum the results of two types of events. The first datum is the number of initial activations per second. The value is between zero and three and is the number of transitions from the "off" state to the "on" state, indicative of the number of cars that passed over during this second. The second datum is the total number of samples that were in the "on" state for that second. The vehicle count is intended to 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 by the microprocessors 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 eight bit byte is followed by a parity bit for error detection. This results in an effective nine bit data communication scheme (see figure (3.1)). Nine bit data communications is a very unusual scheme.

A series of eight bytes (9 bits each) is transmitted and then a byte consisting of the "exclusive or" of all eight bytes is transmitted. A sample from a station is therefore a "frame" of nine bytes (9bits in each byte, see figure (3.2)).

Northbound and southbound samples are transmitted at one second intervals. Each sample is transmitted on a one half second boundary. The eighth channel southbound (unused for traffic data) was set so that the loop appeared to be in the "on" state continuously and generated a value of 60 in the low six bits of that byte. This provides

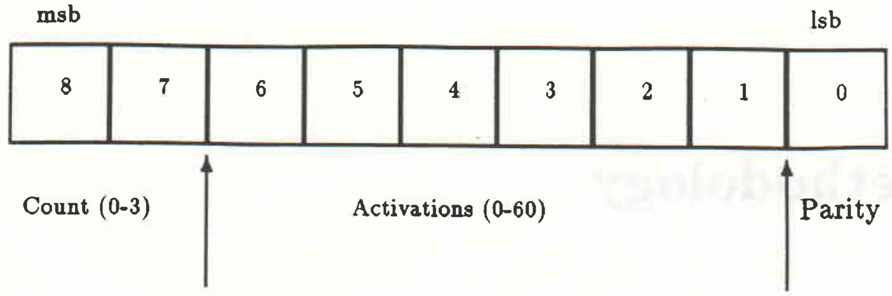

Figure 3.1: TSMC data byte bit format.

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.

an assurance of the correct directionality in the data acquisition. The result is a two nine byte frames per second with the southbound frame marked by the value of 60 in the six lsb of the eighth byte.

3.2. Data Acquisition, Decoding and Reduction

This section describes the methodology used to obtain the inductance loop data from the I5 corridor. It also describes the decoding of the information as sent back from the onroad cabinets. It then examines the data reduction logic used to estimate travel time between loop stations.

3.2.1. Data Acquisition and Decoding

The onroad data is normally received by a Perkins-Elmer (PE) computer located at TSMC. Since this project needed access to the minimum possible sample time data, directly from the field, a separate computer was used for data acquisition. The data, identical to the data received by the PE, is provided to this alternative computer by a parallel modem connection. Unfortunately the PE is in control of the polling of 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 latter be synchronized and decoded off line.

This information (approximately 1.5 megabytes per 24 hour period) is transmitted via phone line from TSMC to the University of Washington (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 nine byte frame is:

> The "exclusive or" (xor) operation is applied to the first eight bytes of the data stream and is compared to the ninth byte. If the results are the same the operation is done again to verify the alignment. If the xor result and the ninth byte differ the entire process slides one byte forward in the data stream (e.g. the second through ninth byte are examined and compared to the tenth 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 ninth relative byte are not identical (this can happen several times in a 24 hour period due to communication or other errors in the cabinets.

Once the data has been received at the University of Washington and decoded into one second volume and occupancy values it is analyzed using *Matlab*^{TM} procedures.

The data acquisition and decoding portion of this project required some software development effort, in particular "C" language routines where developed:

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

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 "speedfactor" g.

To accomplish these goals, several *Matlab*^{TM} procedures were developed. This section presents the functional blocks used to accomplish the goals. appendix (??). 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.

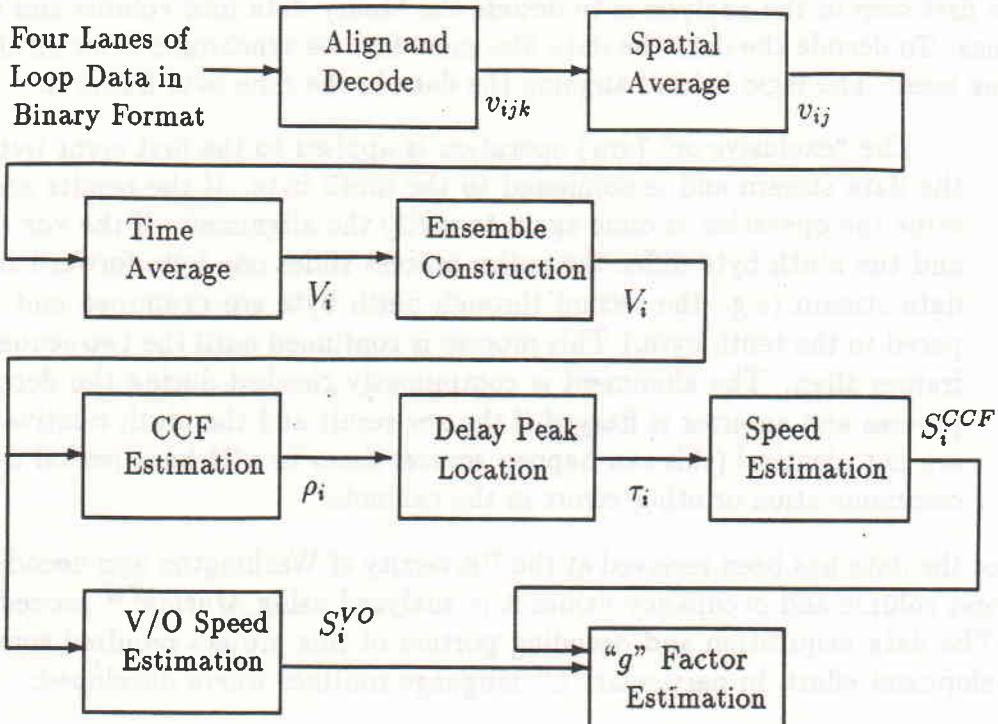

Figure 3.3: Data Reduction Logic

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) is used.

The one second volume from TSMC is then averaged together for some period of time, five seconds per data point is typical. 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 which is of interest in this study. The time average provides a method of establishing a realistic mean value for volume.

The time averages 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:

$$V_{ik} = (Vol)_{ik} - \frac{1}{N}\sum_{k=1}^{N}(Vol)_{ik} \quad (3.1)$$

The ensembles from two sites are cross-correlated to produce a cross-correlation function,

$$R_{12}(\tau_k, T) = \sum_{j=1}^{N} V_{ij}V_{2(j+k)} \quad \forall k \quad \epsilon[0, N]. \quad (3.2)$$

The function is then normalized to produce a covariance function.

$$\rho_{12}(\tau_k, T) = \frac{R_{12}(\tau_k, T)}{\sigma_1\sigma_2} \quad (3.3)$$

where,

$$\sigma_1^2 = \frac{1}{N}\sum_{k=1}^{N} V_{1k}V_{1k} \quad \sigma_2^2 = \frac{1}{N}\sum_{k=1}^{N} V_{2k}V_{2k} \quad (3.4)$$

It is this covariance function which contains the time delay information. The average delay time is determined by the location of the maximum value in the covariance function, or the value of τ_0 such that

$$max\left[\frac{\partial\rho_{12}(\tau_0, T)}{\partial\tau} = 0\right]. \quad (3.5)$$

To actually estimate the proper value for τ_0 the discrete covariance function is approximated by

$$\rho(\tau) \approx a\tau^2 + b\tau + c \quad (3.6)$$

about the maximum value in the $\rho_{12}(\tau, T)$ array. This function is analytically differentiable so that

$$\frac{\partial}{\partial\tau}\rho(\tau_0) = 2a\tau_0 + b = 0 \quad (3.7)$$

and

$$\tau_0 = \frac{-b}{2c}. \quad (3.8)$$

Since the functional form is least squares fit about the maximum value of $\rho(\tau_i)$ this meets all the constraints of equation (3.5). In this way the mean time delay between the loops is estimated. This process can be done at the resolution of the time averaging. For example, if five second averages are used and an ensemble of 64 points is cross-correlated, it takes approximately five minutes to accumulate the data for the first time delay estimates. Thereafter an interloop delay estimate can be made every five seconds when a new ensemble member is available. Thus the time resolution of the method is equal to the time average applied to the on road data. In practice, the data is analyzed over a 24 hour period. The time delay estimate is made every five minutes as a default but can be estimated at any interval greater than the time average period. This time delay estimate is used to construct the speed estimate along with the assumption of constant mean speed between the loops

$$S_i^{CCF}(T) = \frac{\Delta x}{\tau_i(T)}. \tag{3.9}$$

The speed estimate based upon the ratio of volume to occupancy (and a speed factor) is constructed using time averaged volume and occupancy values (S_i^{VO}). This is constructed using TSMC's value of 2.2 for the speed factor. The volume and occupancy values are averaged over the same period (T) that is used for the correlation coefficient function (CCF) technique. As a result there are five arrays produced by the analysis procedures: delay estimates, CCF speed estimate, volume-occupancy ratio speed estimates, time averaged volume and time averaged occupancy.

Using the two estimates of speed, the speedfactor for each period can be estimated,

$$2.2 \times \frac{S_i^{VO}}{S_i^{CCF}} = g_i. \tag{3.10}$$

This then provides an array of values of speedfactor with an average occupancy associated with the time interval used in the estimation.

Using the data derived above the following plots are made:

- Time delay vs Time (Figure (3.4))
- S^{VO} speed vs Time (Dotted lines in Figure(3.5))
- S^{CCF} speed vs Time (Solid lines in Figure (3.5))
- g vs Occupancy (Figure (3.6))

Figures 3.4 through 3.6 are examples of the output for the 24 hour period beginning 8:31 on 28th of November 1990. In addition, the functional form of the speedfactor developed in section (2.3), equation (2.27), is fit to the speedfactor/occupancy pairs. The speedfactor function and data points vs occupancy is shown in figure(3.7). The data reduction effort produced a set of procedures to allow the data from TSMC inductance loops to be used to estimate:

1. Average transit times,
2. Average traffic speed,
3. "g" factor as a function of occupancy.

The validity of the results of the data reduction is examined in the next section.

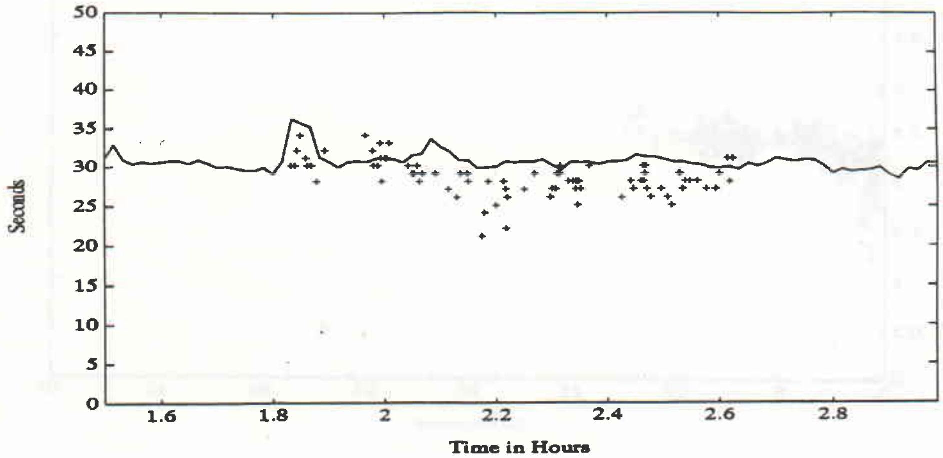

Figure 3.4: Southbound time delay, cross-correlation and onroad data, as a function of time.

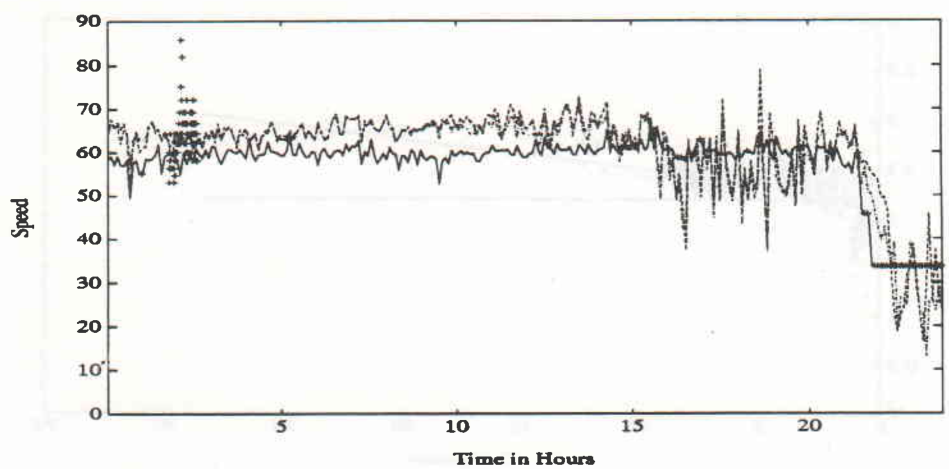

Figure 3.5: Southbound speed, V/O, cross-correlation and onroad data, as a function of time.

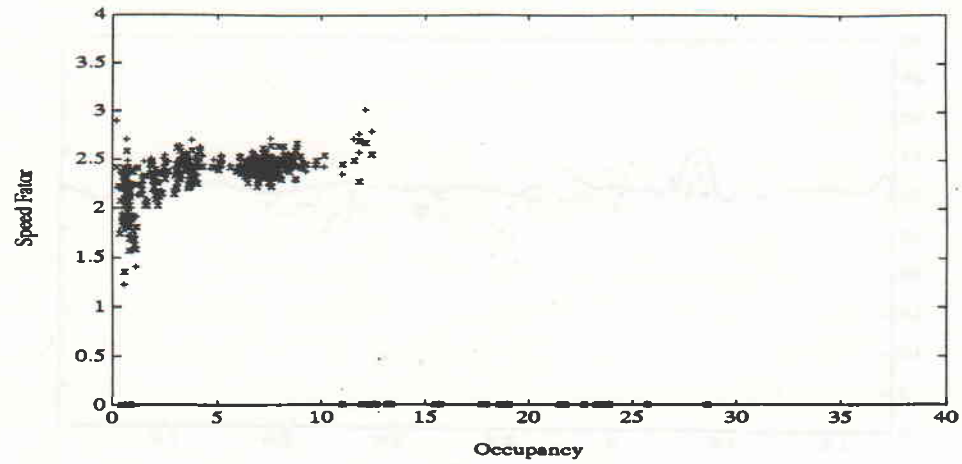

Figure 3.6: Speedfactor (shown as a "+") as a function of occupancy, southbound I5 between 185th and 195th. (A zero value indicates that the correlation coefficient was of an unacceptable value.)

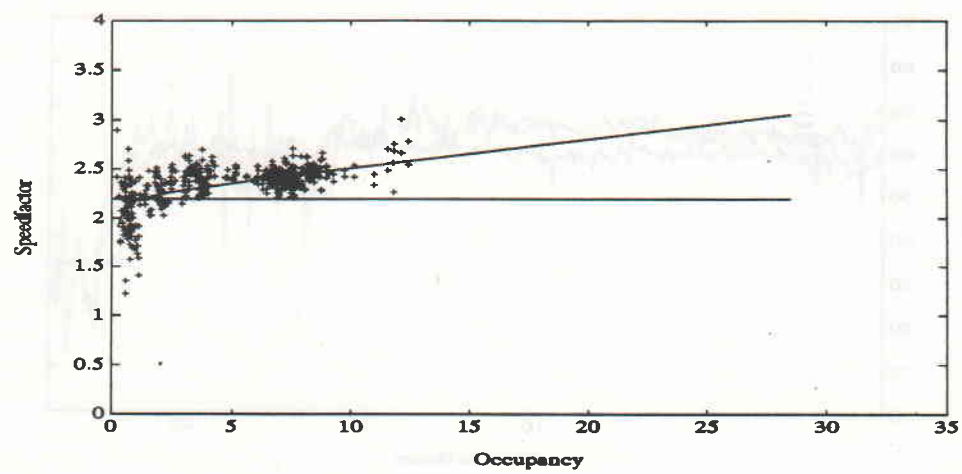

Figure 3.7: Speedfactor vs occupancy with fitted function, southbound I5 between 185th and 195th. The horizontal line is at 2.2 the value presently used by TSMC.

4 . Results

In this chapter the theory developed in chapter 2 is used with real traffic data to demonstrate the viability of using correlation techniques to measure the propagation time of traffic between inductance loops one half mile apart. This delay time is converted to a mean speed estimate that can be compared to onroad estimates by laptop computers as well as the conventional volume to occupancy ratio. The last section of this chapter presents a new functional relationship between observed speed and the volume to occupancy ratio. This new function is then used to improve the speed estimates from the conventional ratio technique.

4.1. Demonstration of Principle

The data used in this study is the count of activations of an inductance loop. Each activation is indicative of a vehicle passing over the loop. The number of activations is summed over some period. Figure(4.1) is the number of vehicles at each of the chosen stations summed over four lanes and five seconds. This timeseries is used to create auto- and cross-correlation functions. There is no clear relationship between the timeseries obvious from the plots. This is the reason more sophisticated time series techniques involving the auto and cross-correlation coefficient functions are needed.

The validity of the time delay estimates depend upon the value of the correlation coefficient between the volume at the two stations. Figure(4.2) top shows and example of the autocorrelation coefficient function for one of the stations. Note that the correlation function has a maximum at zero time delay, and this maximum is normalized to unity using the variance (see equation 2.5). Figure(4.2) bottom is the cross correlation coefficient function between two stations 0.5 miles apart. Equation 2.12 predicts that under ideal conditions this function should be the variance normalized ACF shifted but a time lag of τ_0 and scaled by the dispersion factor. Comparing the upper and lower plots in figure(4.2) the noteworthy factors are; first the CCF has a lower maximum peak value; and second the peak is shifted away from zero time lag. Both of these features are predicted by equation 2.12. The asymmetry about the peak in the CCF is indicative of the effect of the noise component from equation (2.10).

The CCF for loops 0.5 miles apart consistently have the predicted shape. The value of the cross-correlation coefficient is generally above 0.4. The cross-correlation

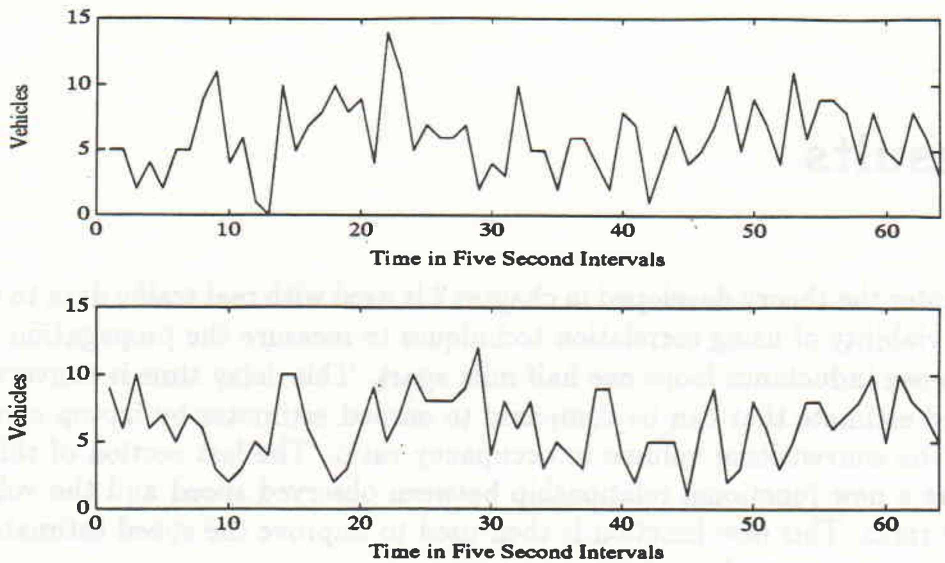

Figure 4.1: Number of vehicles in five seconds as a function of time.

coefficient values, every five minutes for a 24 hour period, are shown in Figure(4.3). The occupancy for this same period is shown in Figure (4.4). Note that the values of ρ_{12} are above 0.5 in nearly all cases where the occupancy is below 12%. The correlation coefficient is indicative of the common information in the two signals. For example if $\rho = 0.6$ then 77.5% of the information present in the signals at the first loop is observed at the second loop. This indicates that during a significant portion of a day a reliable and accurate estimate of the average delay time between inductance loops can be made using this cross-correlation technique. The data presented is from southbound I5 and begins at 8:22 am. The high occupancy period beginning at the 20 hours point is the morning rush hour for 28th of November 1990.

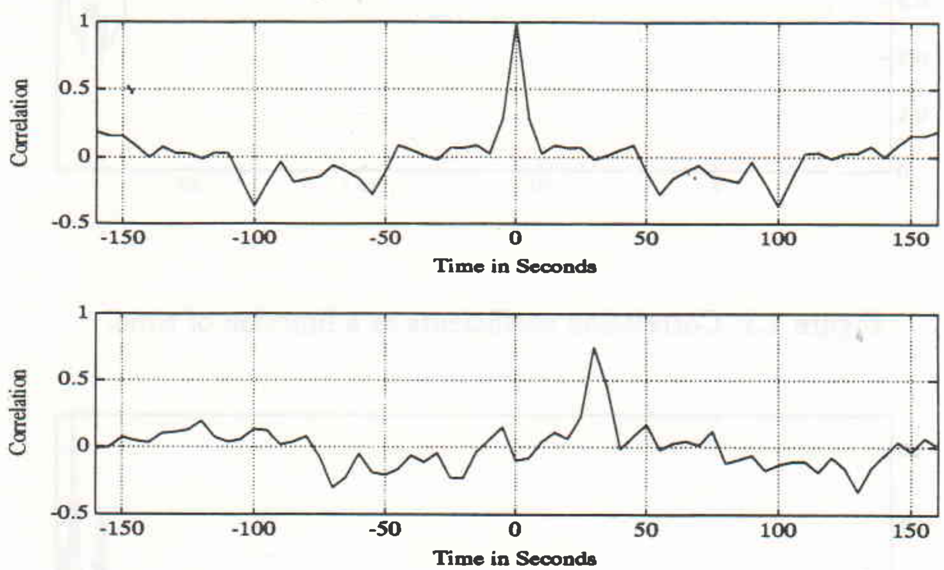

Figure 4.2: Autocorrelation (top) and Cross-correlation function (bottom) for traffic time series.

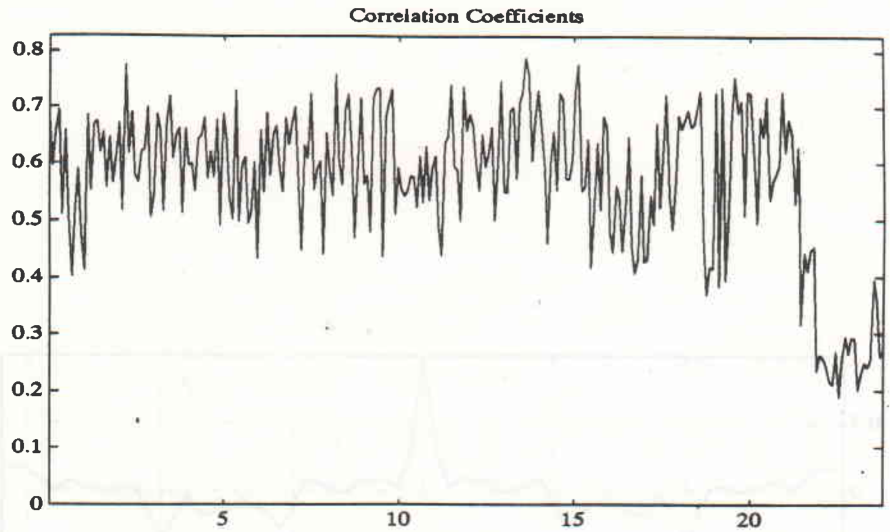

Figure 4.3: Correlation coefficients as a function of time.

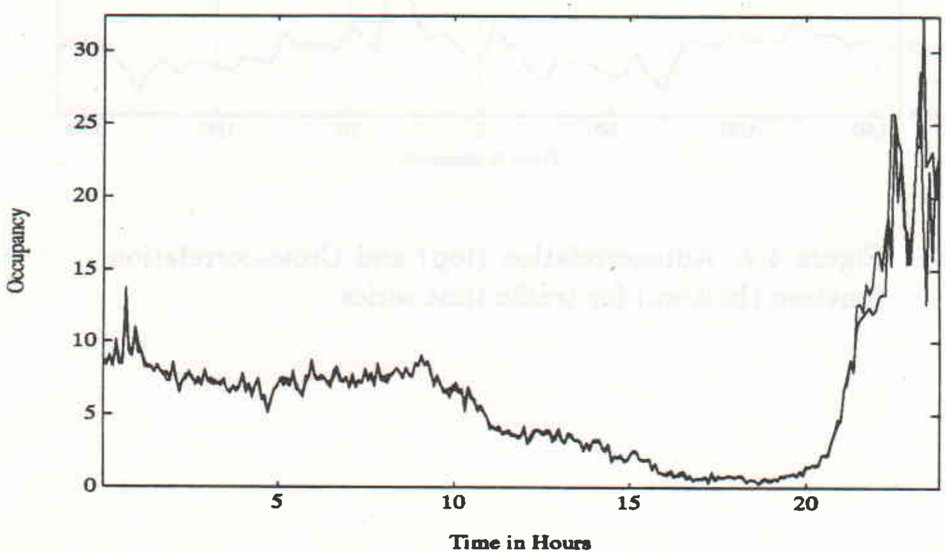

Figure 4.4: Southbound occupancy as a function of time.

4.2. Comparison to "Onroad" time delay estimates

The members of TRAC and TSMC have used portable computers "onroad" to estimate travel time. They place personnel with laptop computers at two locations and record license plate numbers and a time stamp for each license. These computers have their clocks synchronized so that by matching license numbers and determining the difference in the time stamps a delay estimate is possible. This technique is often used over long stretches of freeway and has errors associated with the visual perception and typing ability of the personnel operating the laptops that are difficult to quantify. At the request of, and in cooperation with TRAC a set of delay times was measured on 28th of November 1990. The data from TSMC was recorded in parallel to compare the results of the "laptop" and CCF methods. Figure (4.5) show the results of the CCF

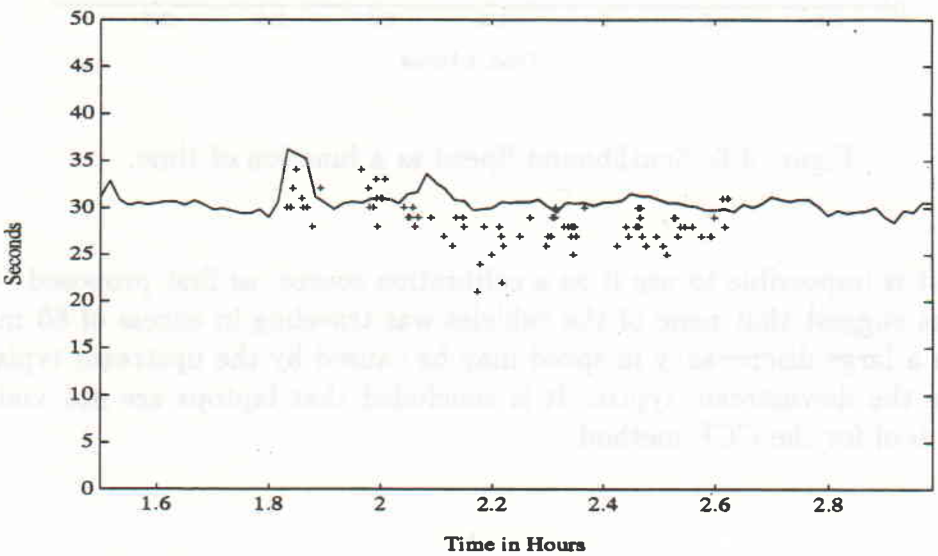

Figure 4.5: Southbound time delay, cross-correlation and onroad data, as a function of time.

method (the solid line) and the "laptop" method (indicated by '+'). This data is from the 185th and 195th street overpasses (separated by 0.5 miles) and is for the period of 9am until 11:30 am. This data indicates that the time delay between sites using the two methods is similar. The "laptop" method has a much larger variance. Figure (4.6) shows the speed estimates produced using the three methods. The dashed line is the volume/occupancy ratio with $g = 2.2$; the solid line is the CCF method, and the "+" are the laptop estimates. The laptop method has such a large unquantifiable

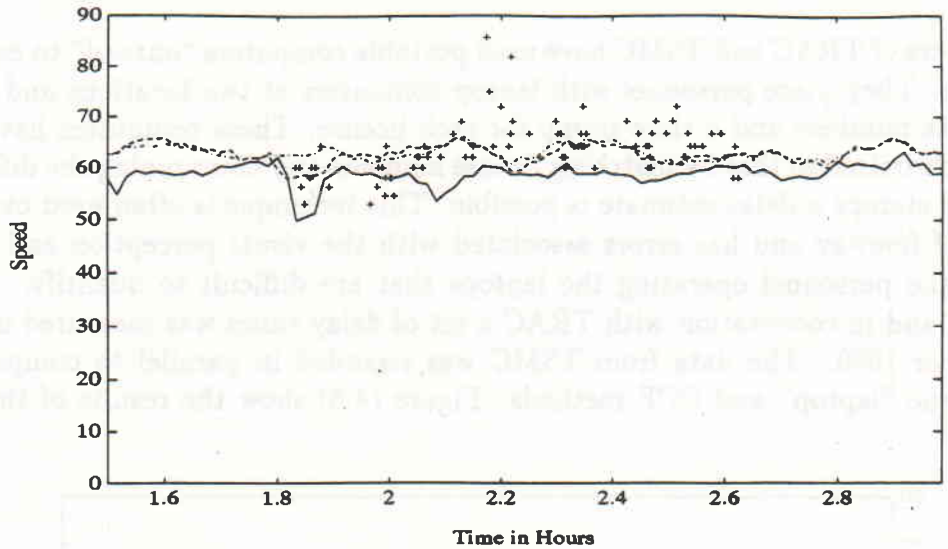

Figure 4.6: Southbound Speed as a function of time.

error that it is impossible to use it as a calibration source, as first proposed. Onroad observations suggest that none of the vehicles was traveling in excess of 80 miles per hour. Such a large discrepancy in speed may be caused by the upstream typist being slower than the downstream typist. It is concluded that laptops are not viable as a calibration tool for the CCF method.

4.3. Occupancy Limits of the Methodology

This method has a range of occupancy values over which it is valid. The limit on the CCF method is actually associated with the size of $\rho_{xy}(\tau_0)$. If this value is too low, the confidence of identifying the time delay is reduced or eliminated. Figure

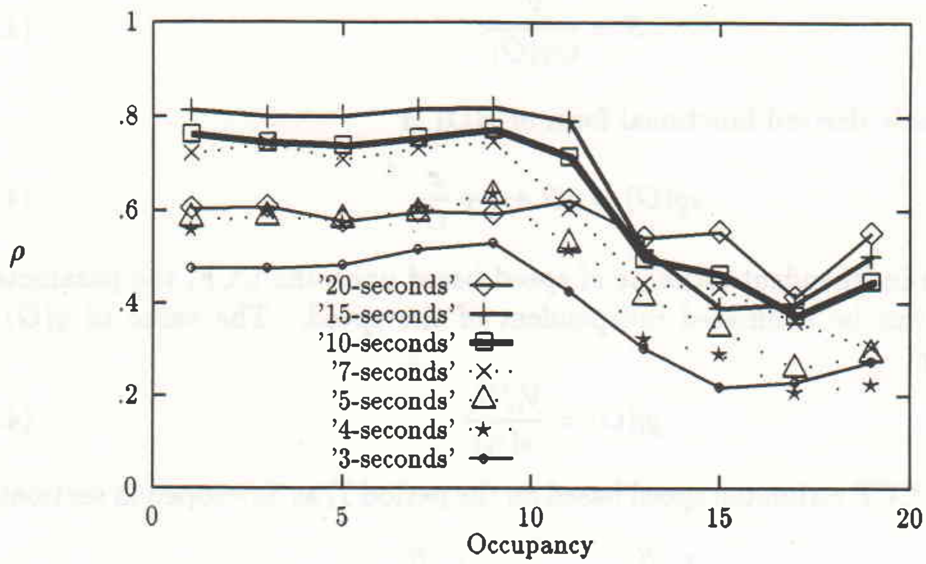

Figure 4.7: Correlation coefficient as a function of occupancy and averaging time.

(4.7) shows the correlation coefficient as a function of Occupancy and sample average time (see Section 3.2.2). This shows that the correlation coefficient decreases with increasing occupancy. In general the usability of the CCF speed estimate requires that the correlation coefficient be greater than about (0.4). . By limiting the value of acceptable (ρ) we limit the Occupancy range over which the CCF method may be used. Based upon the constraints on the data presented in this study,

- spatial averaging over several lanes,
- 5 second time average per point,
- $\rho \geq 0.4$,

the maximum occupancy at which this CCF method is reliable is approximately 15%. This maximum occupancy is less than heavy rush hour occupancy values or incident occupancy values. However, a large portion of the day has occupancies below 15% as was clear from figure(4.4). Additional work is in progress to increase the occupancy range over which the CCF speed estimate is reliable, but is incomplete at this writing.

4.4. Correlation of the Volume to Occupancy Ratio with Speed

The correlation between the volume-occupancy ratio leads to the development of the "g" factor and the "g factor functional form". As developed earlier, the speed estimated based upon volume and occupancy is

$$S = \frac{V}{Og(O)}. \tag{4.1}$$

And the empirically derived functional form of g(O) is

$$sg(O) = aO + b + \frac{c}{O} \tag{4.2}$$

Since we have an independent estimate of speed based upon the CCF, the parameters of the function can be estimated independent of the speed. The value of $g(O)$ is approximated by

$$g_i(O) = \frac{\hat{V}_i/\hat{O}_i}{s(\tau_i)} \tag{4.3}$$

where $s(\tau_i)$ is the CCF estimated speed based on the period T_i as developed in section(3.2.2), and

$$\hat{V}_i = \frac{1}{N}\sum_{j=1}^{N} V_{ij} \quad \hat{O}_i = \frac{1}{N}\sum_{j=1}^{N} O_{ij} \tag{4.4}$$

when there are N samples in the period T. Figure (4.8) shows the estimates of $g_i(0)$ for a 24 hour period. The general distribution of g_i vs occupancy is similar in shape to the functional form developed in Chapter (2.3). The dependency on Occupancy is more dramatic at very low occupancy as predicted in the functional form. The mean value for g at each occupancy is shown in figure(4.9) The shape of $g(O)$ is parameterized using equation (4.2). Fitting the data set to equation (4.2) results in the correlation

$$g(O) = 0.0242 \times O + 1.8566 - \frac{0.0672}{O}. \tag{4.5}$$

The function and data are shown in figure (4.10). This function

$$Speed = \frac{V}{0.0242 \times O^2 + 1.8566 \times O - 0.0672} \tag{4.6}$$

makes an improved speed estimate, based on volume and occupancy estimates.

This result is topologically localized as is suggested by Hall and Persaud [HP88]. This implies that the parameter set needs to be estimated for each location on the highway that will use this improve speed estimate. Figure (4.11) is the original speed based upon V/O and CCF using the TSMC simple constant 2.2 for g. Figure (4.12) is the improved speed estimate using the new $g(O)$ function with the V/O ratio. It is clear that the speed estimate is closer to the independent estimate of speed. This improved speed estimate can be used by any application program that has the parameters for the correlation at each site. For example, a RTMIS which used the coefficient for each set of loops on a highway will provide a more accurate speed estimate.

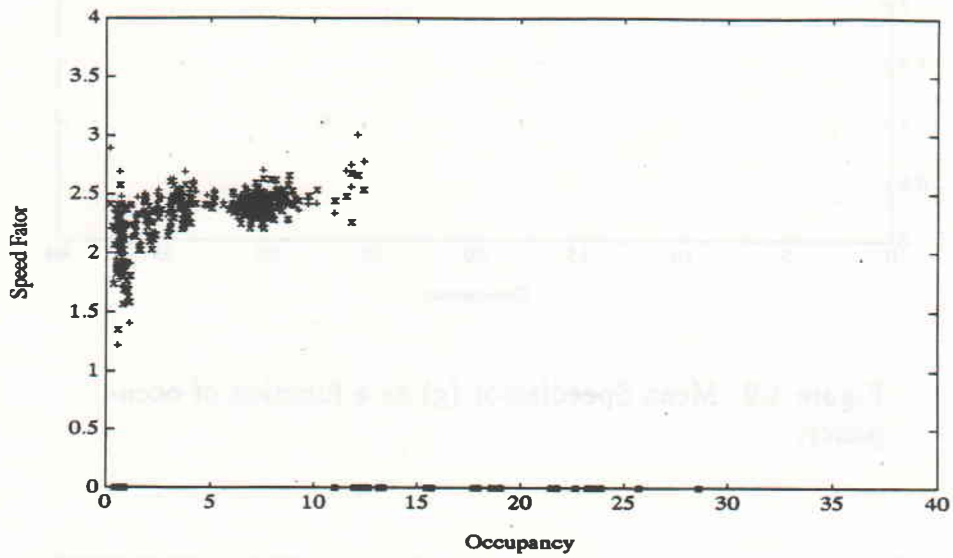

Figure 4.8: Speedfactor (g) as a function of occupancy.

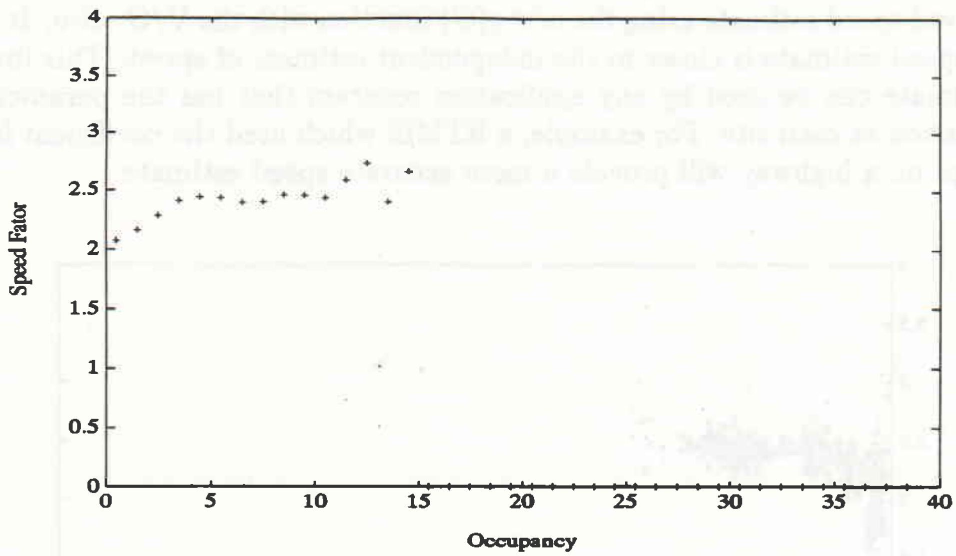

Figure 4.9: Mean Speedfactor (g) as a function of occupancy.

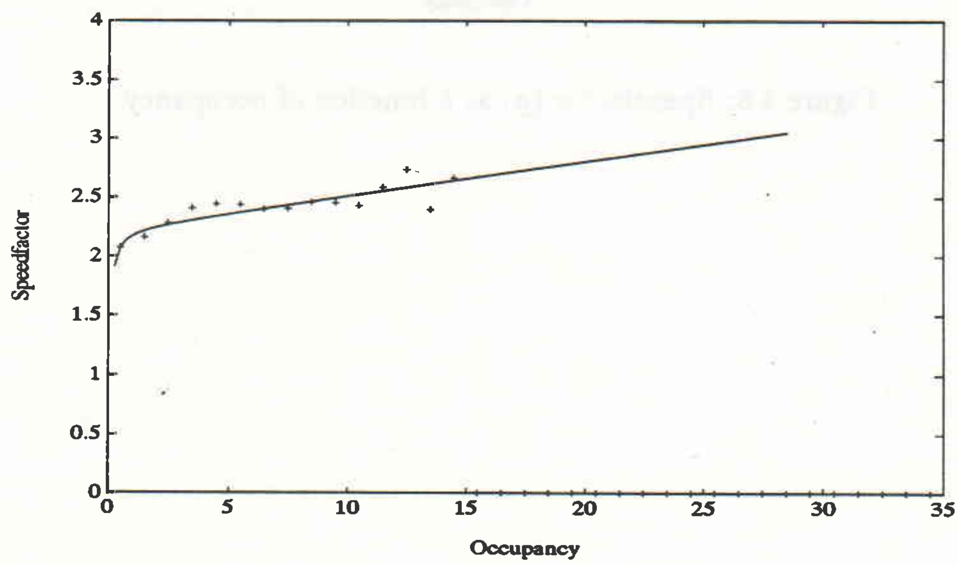

Figure 4.10: Mean speedfactor data and function (g) as a function of occupancy.

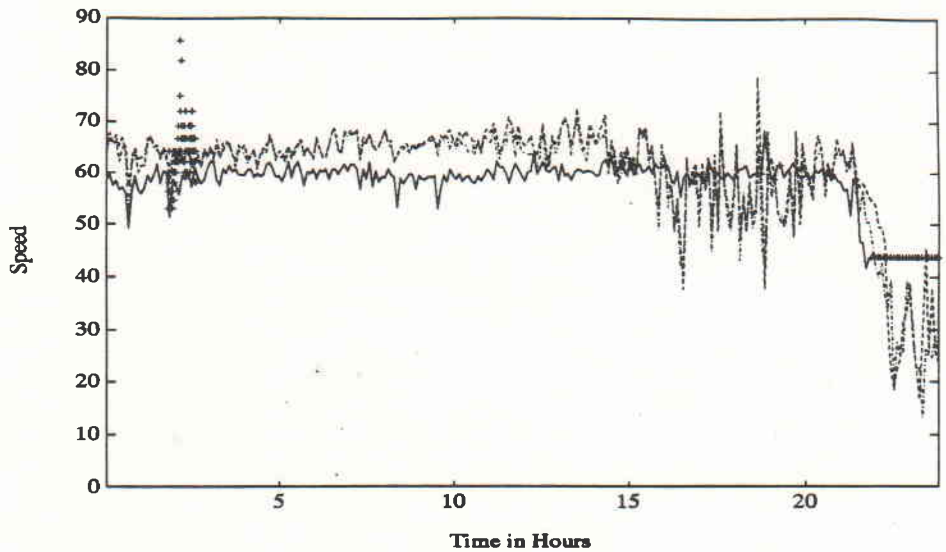

Figure 4.11: Original speed estimates, CCF method - solid line; V/O method - dashed line; "laptop" estimates - "+".

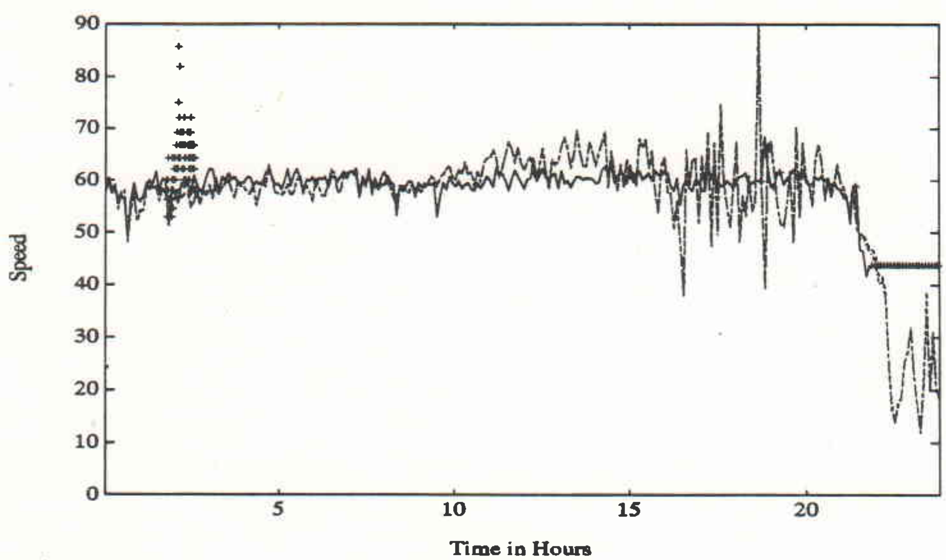

Figure 4.12: Improved speed estimates.

5 . Conclusions

This research effort has demonstrated the viability of using inductance loops to measure speed. The cross-correlation coefficient function created from the time dependent volume estimates at locations 0.5 miles apart provides a reliable estimate of the mean time delay between the loops. Together with the measured distance this can be used to estimate the mean speed of traffic. This method provides reliable results for occupancies below 15%. The minimum averaging time for reliable results is of order five seconds. A delay estimate can be made every five seconds based on the averaging time. This estimate is independent of the volume/occupancy ratio which has been the basis of previous speed estimates by WSDOT, USDOT and UW projects. Because of the independent estimate of speed from the same data stream a corrected volume/occupancy based speed estimate can be made which accounts for the occupancy dependence inherent in the V/O relationship.

This project has proposed a functional form for the "speedfactor" relating the volume/occupancy ratio to the speed. This functional form has three parameters and is site specific. Using the independent speed estimate available from the cross-correlation coefficient function the parameters of the speedfactor function can be estimated. Once this is done the calibrated speedfactor function can be used with a single loop set's volume and occupancy values to provide an improved speed estimate. This resolves the occupancy dependence problem, in using loop data for speed, that has been raised by Hall & Persaud [HP88].

Future work includes extending occupancy range for reliable CCF results, and the use of the correlation coefficient function in identifying failed loops.

Bibliography

[BP86] J.S. Bendat and A.G. Piersol. *Random Data: Analysis and Measurement Procedures.* John Wiley & Sons, 2nd edition, 1986.

[Gre35] B.D. Greenshields. A Study of Traffic Cappacity . *Proceedings of the Highway Research Board, Washintgton,* 14:488, 1935.

[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.

[Leu88] W. Leutzbach. *Introduction to the Theory of Traffic Flow.* Springer-Verlag, 1988.

[LW55] M.J. Lighthill and G.B. Whitham. On Kinematic Waves, A Theory of Traffic Flow on Long Crowded Roads. *Proceedings 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 Syetems. In *69th Annual Meeting of the Transportation Research Board, Paper N. 890800,* 1989.

[War52] J.G. Wardrop. Some Theoretical Aspects of Road Traffic Research . *Proceedings of the Institution of Civil Engineers,* 3:158, 1952.