**Final Technical Report**

Research Project T9233, Task 14
"Improved Inductor Loop"

# IMPROVED ERROR DETECTION FOR INDUCTIVE LOOP SENSORS

by

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

# TECHNICAL REPORT STANDARD TITLE PAGE

| | | |
|---|---|---|
| 1. REPORT NO.<br>WA-RD 300.1 | 2. GOVERNMENT ACCESSION NO. | 3. RECIPIENT'S CATALOG NO. |
| 4. TITLE AND SUBTITLE<br>IMPROVED ERROR DETECTION FOR INDUCTIVE LOOP SENSORS | | 5. REPORT DATE<br>May 1993<br>6. PERFORMING ORGANIZATION CODE |
| 7. AUTHOR(S)<br>Daniel J. Dailey | | 8. PERFORMING ORGANIZATION REPORT NO. |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Washington State Transportation Center (TRAC)<br>University of Washington, JD-10<br>University District Building, Suite 535; 1107 NE 45th St<br>Seattle, Washington 98105 | | 10. WORK UNIT NO.<br>11. CONTRACT OR GRANT NO.<br>T9233, Task 14 |
| 12. SPONSORING AGENCY NAME AND ADDRESS<br>Washington State Department of Transportation<br>Transportation Building, MS 7370<br>Olympia, Washington 98504-7370 | | 13. TYPE OF REPORT AND PERIOD COVERED<br>Final Report<br>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

This report describes the development of an algorithm to detect anomalies in the time series from inductance loop sensors. The anomalies may arise from traffic incidents or loop detector system malfunction. The algorithm uses a statistic produced with inductance loop data to make an optimal prediction of the volume and occupancy values that will occur at the next time step. To guarantee the optimality of this prediction, a Kalman predictor, for use with inductance loop data, is developed. To detect variations from the normal state the optimal prediction is compared with the observed value. Anomaly detection is accomplished by applying thresholds to the difference between the predictions and the observed values. This report demonstrates the use of the anomaly detection algorithm with inductance loop data gathered on Interstate 5 in Seattle, Washington. The report also discusses the scaling and values of thresholds necessary for anomaly detection. This type of dynamic prediction and threshold can be valuable to traffic management systems that rely heavily on inductance loop data.

| | |
|---|---|
| 17. KEY WORDS<br>inductive loops, volume occupancy, Kalman filter, anomaly detection, traffic | 18. DISTRIBUTION STATEMENT<br>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 | 25 | |

# CONTENTS

# List of Figures

## Executive Summary

This report describes the development of an algorithm to detect anomalies in the time series from inductance loop sensors. The anomalies may arise from traffic incidents or loop detector system malfunction. The algorithm uses the statistics associated with inductance loop data to make an optimal prediction of the volume and occupancy values expected at the next time step. To guarantee the optimality of this prediction a Kalman predictor, for use with inductance loop data, is developed. To detect variations from the normal state the optimal prediction is compared with the observed value. Anomaly detection is accomplished by applying thresholds on the difference between the predictions and the observed values. This report demonstrates the use of the anomaly detection algorithm using inductance loop data gathered on Interstate Five in Seattle, Washington. It also discusses the scaling and values of thresholds necessary for anomaly detection. This type of dynamic prediction and threshold can be valuable to traffic management systems that rely heavily on inductance loop data.

# 1. INTRODUCTION

Inductance loops are the principal traffic sensor for a number of metropolitan traffic management systems. The detection of anomalous conditions in the inductance loop data stream has many potential uses, particularly as an indicator of traffic "incidents." This report describes and demonstrates a new predictor/detector methodology which uses an optimal recursive predictor (a Kalman predictor[Boz84]) with real time inductance loop data to identify anomalies in a traffic management system's inductance loop data stream.

A detection system can be used to identify anomalies associated with both traffic incidents and data acquisition system malfunction. If the difference between these types of anomalies can be distinguished, then the immediate detection of anomalies will allow more rapid response to incidents as well as equipment failure. Recursive predictor algorithms have been used in transportation research to examine traffic flow [OS84, Lu90, CS90] and used in other engineering fields to do fault detection. This report presents not only the use of an optimal recursive predictor but also uses information about the correlation between the state variables to construct the system model for the predictor. The specific properties of the time series useful in this effort are the correlation between the data items produced by the inductance loop detectors. These correlations include: (1) the correlation between volume and occupancy (the properties measured by single inductance loops); (2) the correlation between data from adjoining lanes; and (3) the correlation between time series from loops along the path of traffic flow[1]. These properties and a recursive optimal predictor are combined to produce a demonstration of the utility and effectiveness of such a predictor/detector. This demonstration uses inductance loop sensor traffic data taken from Interstate Five in Seattle, Washington.

[1]Recent work has demonstrated the correlation between data from loops separated by relatively long distances of one half to one mile can be used to estimate travel time between the loops[Dai92].

# 2. THEORY

Anomaly detection requires first a prediction of the "normal" state and then a threshold value to be used to identify a measurement that is not in the "normal" range. This section develops a theoretical base for predicting the normal state at a particular time step based on historic data. This prediction uses a vector optimal recursive predictor or vector Kalman predictor.

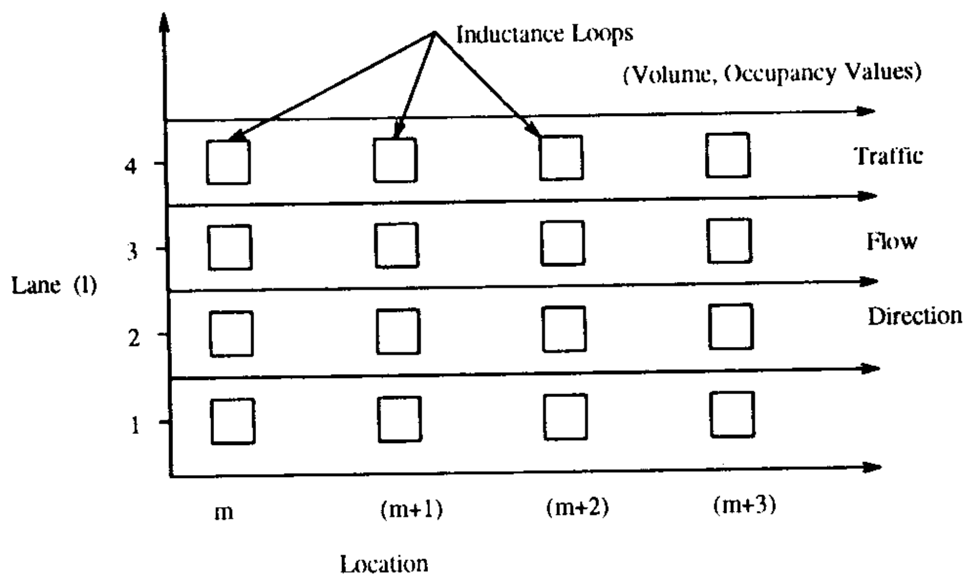


Figure 2.1: Inductance loop layout.

The volume (vehicles per hour) and occupancy (the percent of some total time a segment of highway is occupied by vehicles) time series are modeled as a mean value

$$Y(k) = \begin{bmatrix} V(1,m,k) \\ O(1,m,k) \\ \vdots \\ V(4,m,k) \\ O(4,m,k) \\ V(1,m+1,k) \\ O(1,m+1,k) \\ \vdots \\ V(4,m+1,k) \\ O(4,m+1,k) \\ V(1,m+2,k) \\ O(1,m+2,k) \\ \vdots \\ V(4,m+2,k) \\ O(4,m+2,k) \\ \vdots \\ V(1,m+3,k) \\ O(1,m+3,k) \\ \vdots \\ V(4,m+3,k) \\ O(4,m+3,k) \end{bmatrix} \tag{2.1}$$

Figure 2.2: Inductance loop observation vector.

about which there is a statistical fluctuation. For this report, the observation variable will be a vector combining the volume and occupancy observations. The members of the observation vector at each time step are the volume and occupancy values from loops at several locations. Figure 2.1 represents the loops pictorially and figure 2.2 shows the members of the observation vector in which $V(l,m,k)$ is the volume at loop $l$ at site $m$ at time step $k$; and $O(l,m,k)$ is the occupancy at loop $l$ at site $m$ and time step $k$. The observation vector at time step $k$ is modeled as the combinations of a state vector $X(k)$ with a zero mean random variable term $v(k)$ and is written

$$Y(k) = X(k) + v(k). \tag{2.2}$$

For this detection application we wish to make an optimal prediction for the value of $X$ at the next time step $(k+1)$. The future value is

$$X(k+1) = AX(k) + w(k), \tag{2.3}$$

where $A$ is a system model for $X$ that reflects the dynamics of the system and $w(k)$ is a zero mean random variable or "noise" term. The system model provides a means to estimate the future value based on the present value.

The system model or transition matrix $A$ describes the dynamics of the process. The system model is usually constructed from a differential equation for the process, but in this case there is no simple dynamics equation that relates the observed variables. However, previous work has demonstrated that there is a correlation between the time series from upstream and downstream sites (e.g. a perturbation of the mean volume count at the upstream station will be seen at the downstream station some time later depending on the traffic speed[Dai92]). Additional information about the dynamics of the time series is available from the correlation between the values of volume and occupancy observed in adjacent lanes. In this work the correlated information in the time series arising from different sites and lanes is exploited to construct a system model, or what in this case is better described as a transition matrix since it is based on the past transitions and is used to estimate the future one.

The transition matrix is derived from the prediction equation 2.3 and the correlated information in the time series. Equation 2.3 is post multiplied by $X(k)^T$ (the transpose of $X(k)$), and then taking the expected value results in,

$$E\{X(k+1)X(k)^T\} = E\{AX(k)X(k)^T\} + E\{w(k)X(k)^T\}, \tag{2.4}$$

which, with the assumption that the noise term is uncorrelated, can be rewritten as

$$E\{X(k+1)X(k)^T\} = AE\{X(k)X(k)^T\} + E\{w(k)\}E\{X(k)^T\}. \tag{2.5}$$

Since $w$ is assumed to be zero mean noise $E\{w\} = 0$, and the transition matrix is written,

$$A = E\{X(k+1)X(k)^T\}\left[E\{X(k)X(k)^T\}\right]^{-1}. \tag{2.6}$$

This transition matrix is made up of the covariance matrix at zero lag time,

$$E\{X(k)X(k)^T\} \tag{2.7}$$

and the covariance matrix at one time increment lag,

$$E\{X(k+1)X(k)^T\}. \tag{2.8}$$

Physically this is a measure of the common information in the time series from one time step to the next, normalized by the information common between the members of the state vector. The explicit use of the correlated information (at lag time zero and one) to construct the transition matrix is a new contribution.

The transition matrix just presented provides an estimate of the transition matrix using $(k+1)$ observations. In reality there are only $K$ observations of the time series and therefore only $(K-1)$ transitions. An approximation of the transition matrix based on the $(K-1)$ previous transitions is constructed:

$$A(K) \approx \left[\sum_{k=2}^{k=K} X(k)X(k-1)^T\right]\left[\sum_{k=2}^{k=K} X(k-1)X(k-1)^T\right]^{-1}. \tag{2.9}$$

This transition matrix uses the information in the time series from the last $(K-1)$ transitions to build an estimate to be used in the next transition. With the transition matrix established it is possible to construct an optimal recursive predictor.

In order to make an optimal estimate of the future value of the time series a vector Kalman predictor formulation is used. The vector Kalman predictor for this application consists of three parts, a prediction equation, a gain equation and an error estimation equation which can be written [Boz84]

Predictor:

$$\hat{X}(k+1|k) = A\hat{X}(k|k-1) + G(k)[Y(k) - \hat{X}(k|k-1)] \tag{2.10}$$

Gain:

$$G(k) = AP(k|k-1)[P(k|k-1) + R(k)]^{-1} \tag{2.11}$$

Error:

$$P(k+1|k) = [A - G(k)]P(k|k-1)A^T + Q(k) \tag{2.12}$$

where $\hat{X}(k|k-1)$ is the filtered estimate of $X$ at time step $k$ given the $(k-1)st$ estimate and where,

$A$ is the transition matrix,

$P(k|k-1)$ is the predicted covariance matrix,

$Y(k)$ is the observation vector from the '$k$'th time step,

$Q(k) = E\{ww^T\}$, and

$R(k) = E\{vv^T\}$.

The Kalman predictor formulation in combination with the transition matrix just developed provides the first part needed for an incident detector, namely an optimal prediction of the expected "normal" state to which observations are compared.

The second part needed for an anomaly detector is a set of thresholds for identifying "anomalous" values. These thresholds need to be scaled appropriately for the data. The number of standard deviations from the prediction is the scale on which the size of the deviation from normal is measured. If there are $K$ observations of $Y(k)$ (the observation vector) the standard deviation vector is written

$$\sigma(K) = \frac{1}{K}\left[\sum_{k=1}^{k=K}(Y(k)-\mu(K))^2\right]^{\frac{1}{2}} \tag{2.13}$$

where

$$\mu(K) = \frac{1}{K}\sum_{k=1}^{k=K} Y(k). \tag{2.14}$$

When the $(k+1)$st observation is available it is compared to the predicted value $(\hat{X}(k+1))$ using the metric

$$\mathbf{D} = \frac{Y(k+1)-\hat{X}(k+1)}{\sigma(K)}. \tag{2.15}$$

It is this metric which is calibrated against observed traffic conditions. The actual threshold for anomaly detection is based on the the size of the individual members $d_i$ of $\mathbf{D}$. For example, one criterion might be that the observed volume deviation be for lane one ($d_1$) be less that a criterion $m_1$ and the observed occupancy deviation ($d_2$) be greater that some other criterion $m_2$,

$$d_1 < m_1 \tag{2.16}$$

$$d_2 > m_2. \tag{2.17}$$

This type of criterion might indicate that traffic has stopped or slowed significantly in that lane in the last time step. The selection of these threshold values depends upon actual observation of traffic flows on the freeway in question.

This section has presented a methodology for establishing dynamic thresholds for occupancy and volume observations from individual lanes. It presented the development of a new system function based on the correlation between the observable volume and occupancy values which is then used with an optimal recursive predictor. The deviation of the newest observation from the prediction is scaled by the standard deviation to provide a meaningful metric for measuring the deviation from the optimal prediction.

# 3. Application in Freeway Traffic

This section uses highway occupancy and volume counts from Interstate Five (I-5) in Seattle, Washington to demonstrate the proposed anomaly detector. It demonstrates the use of dynamically predicted values and the dynamic thresholds using loop data. The loop data is from a period of time when the roadway is at an initially low congestion level and then congestion increases until traffic is nearly stalled. This section also demonstrates the use of the proposed method to detect the change in state of the traffic from a steady flow to a "stop and go" condition.

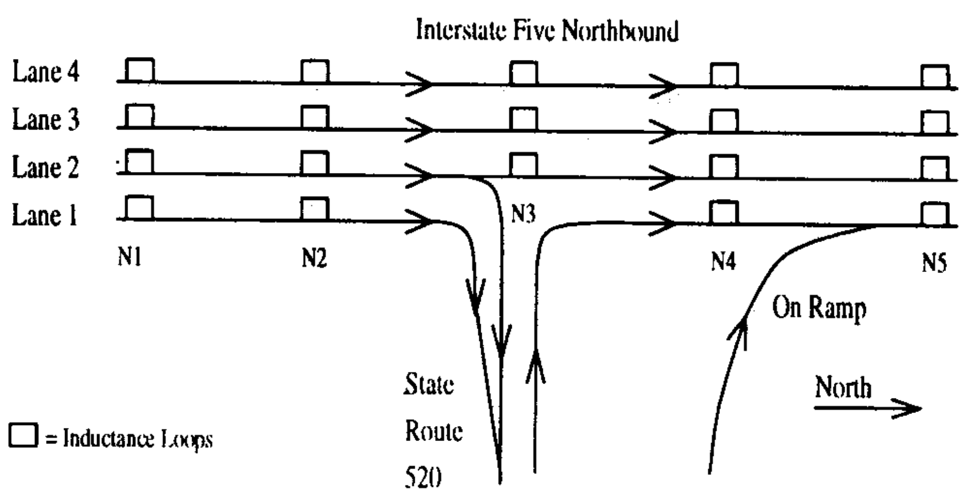


Figure 3.1: Inductance loop locations on northbound Interstate Five in Seattle.

To demonstrate the dynamic nature of the prediction and threshold, volume and occupancy data from five sites on (I-5) were recorded. Sites were selected based on the heavy congestion expected during peak hours at the intersection of State Route 520 (SR-520) and (I-5). The loop layout for the northbound lanes is shown in figure 3.1, the loops are labeled by lane and location. (e.g. Loops locations are labeled N1 through N5 below the lanes, and the lanes are marked 1 through 4.) There are a total of 19 loops in the measurement set, each reporting one minute average volume and occupancy values. Near the intersection between SR-520 and I-5 the widest range of volume and occupancies is expected. During peak periods at least two major observations can be made, (1) SR-520 causes a backup onto I-5 in the right most lanes and (2) the left most lanes are less effected by SR-520.

A good predictor should reproduce these effects. The quality of the predictor depends heavily on the transition matrix. The transition matrix (equation 2.9) is constructed using the time averaged volume & occupancy from the loops. It is created by averaging over a startup period of sixty minutes. After this startup period a prediction of the volume and occupancy values for the loops is made each minute. On I-5 near the SR-520 interchange it is common for the the right most lanes to become heavily congested while the left most lanes are less heavily congested. The global efficacy of the prediction algorithm can be demonstrated by the accurate prediction of the future value of both types of lanes simultaneously. Figure 3.2 shows the observed ($Y(k)$ from equation 2.2) and predicted values ($\hat{X}(k)$ from equation 2.10) of the occupancy in the right most lane verses time. (The southern most site N1 is at the bottom of figure 3.2 and the northern most site N4 is at the top). The prediction and the observation are in agreement for two very different traffic conditions estimated simultaneously. At 2.2 hours into the measurement the occupancy of lane 1 at site N2 shows a heavily congested road way, with traffic nearly stopped. Meanwhile the small occupancy values in lane 1 at site N4 show little congestion. A single transition matrix produces predictions that agree with observations in two very different traffic patterns at the same time but at locations separated by 0.5 miles. The time dependent volume for these sites is shown in figure 3.3. Once again, the significant trends in the volume are well represented by the predictor. The combination of an optimal recursive predictor and the transition matrix presented provides a predictor suitable for use as a baseline for a dynamic threshold anomaly detector.

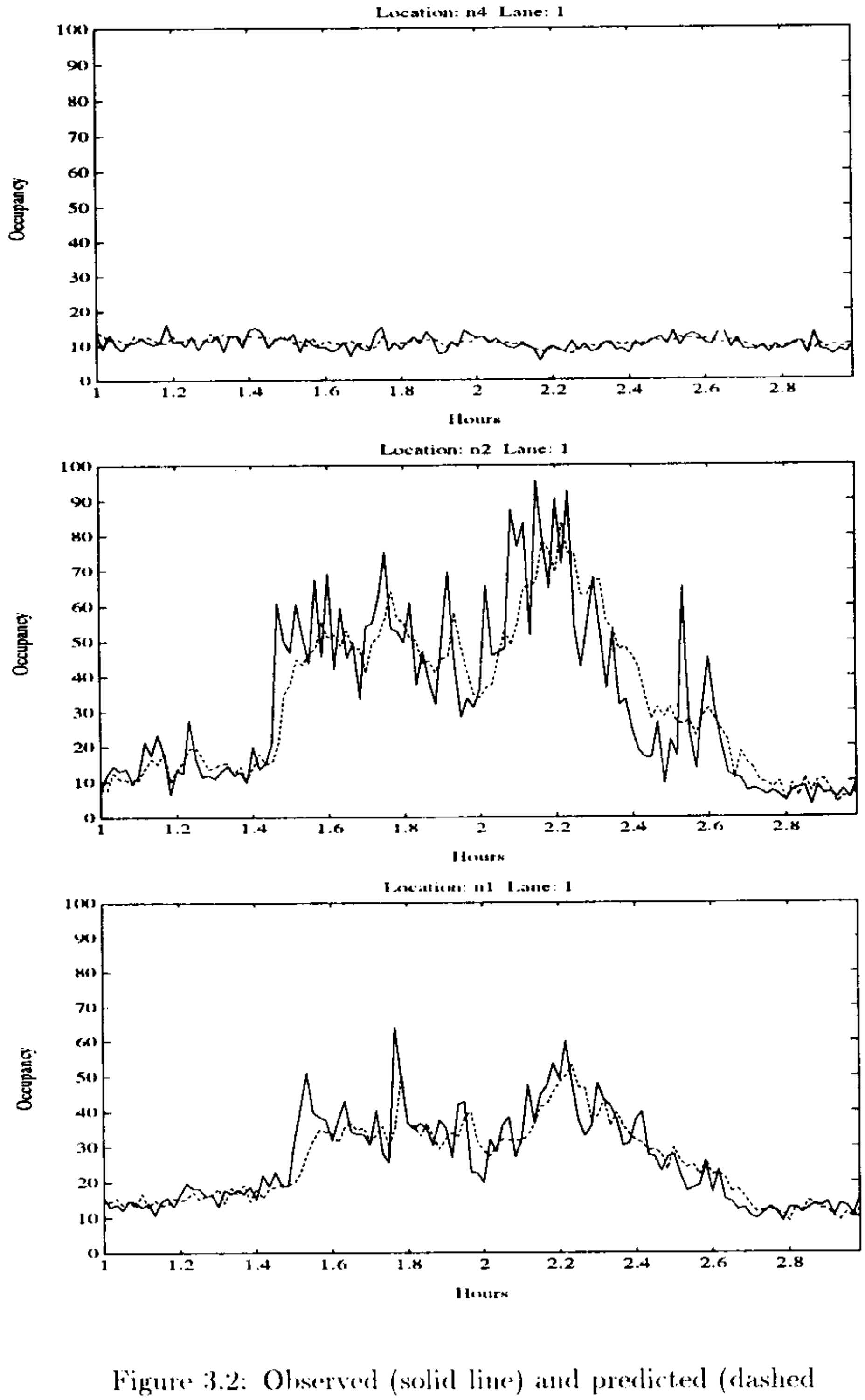


Figure 3.2: Observed (solid line) and predicted (dashed line) values for occupancy at three sites (southern most at the bottom, northern most at the top).

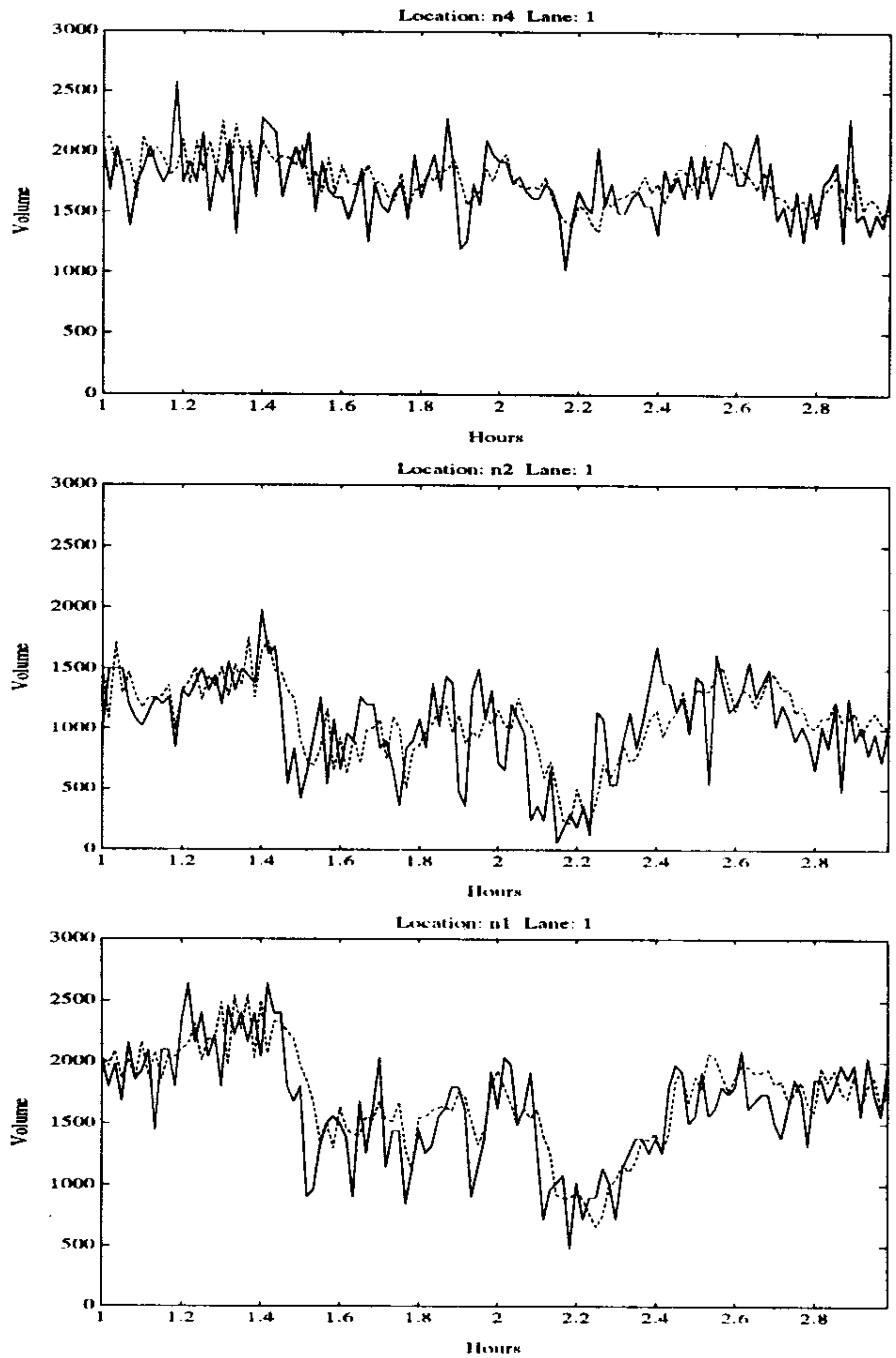


Figure 3.3: Observed (solid line) and predicted (dashed line) values for volume at three sites (southern most at the bottom, northern most at the top).

A dynamic threshold is demonstrated here using a combination of volume and occupancy criteria. This dynamic threshold technique is most effective at identifying a rapid change in the time series being examined. This is especially useful if the volume and occupancy statistics in a free flow condition are very different than those observed during congested periods. Using the same data set, the onset of congestion can be identified using a dynamic threshold. Rapidly increasing occupancy coupled with rapidly decreasing volume is an indication of traffic rapidly slowing and becoming congested. These observations are true for the onset of the congestion in the right most lane of I-5 caused by stopped traffic on SR-520. When there is a stopped traffic on the SR-520 bridge the queue of stopped traffic extends south on I-5. As the length of the queue passes each inductance loop site a relatively rapid and significant change in the volume and occupancy values results.

One possible threshold, to identify a change in state or anomaly, is when the observed occupancy value is three standard deviations larger than the predicted value and the volume value is two standard deviations smaller than the predicted value. For example, at site N1 this is written

$$d_1 = \frac{y_1(k+1) - \hat{x}_1(k+1)}{\sigma_1} < -2 \tag{3.1}$$

$$d_2 = \frac{y_2(k+1) - \hat{x}_2(k+1)}{\sigma_2} > 3 \tag{3.2}$$

where $y_1$ and $y_2$ are the first two member of the observation vector and are the values of the volume and occupancy of lane one at site N1. At site N2 this is,

$$
\begin{aligned}
d_9 &= \frac{y_9(k+1) - \hat{x}_9(k+1)}{\sigma_9} < -2 & (3.3)\\
d_{10} &= \frac{y_{10}(k+1) - \hat{x}_{10}(k+1)}{\sigma_{10}} > 3. & (3.4)
\end{aligned}
$$

More generally the thresholds are,

$$
\left\{ d_i \in \left( \mathbf{D} - \begin{bmatrix} 3 \\ -2 \\ 3 \\ \vdots \\ -2 \end{bmatrix} \right) \right\} \leq 0. \tag{3.5}
$$

When these thresholds are used with the data set presented two anomalous conditions are identified. In figures 3.4 & 3.5 the two standard deviation occupancy threshold is shown as a dashed line and the observed occupancy data as a solid line. In addition the time of the anomalous condition (defined as crossing the specified threshold) is indicated by a vertical line. The first of the two identified anomalies takes place 1.45 hours into the measurement at site N2 indicating the queue is passing the loop location closest to SR-520 on I-5. The second anomaly takes place at 1.5 hours into the measurement indicating that the queue has extended the 0.5 miles separating the loops in approximately 3 minutes.

This type of dynamic threshold is appropriate for identifying rapid changes in the traffic state. It is noteworthy that the anomaly identified is the transition from normal flow to high congestion and that once high congestion is the established state that the detector does not continue to identify the traffic flow as anomalous. In addition those lanes that are unaffected by the backup are not identified as anomalous. This is demonstrated by the data and threshold from site N3 shown in the plot at the top in figures 3.4 & 3.5.

Clearly the use of a threshold based detection system to identify specific traffic events requires knowledge of the expected impact of particular events on the traffic state. The next step in this research is to measure at a set of sites for a sufficient period that an inventory of incidents (as identified by traffic management cameras and state police logs) is available to set threshold conditions for differing types of events.

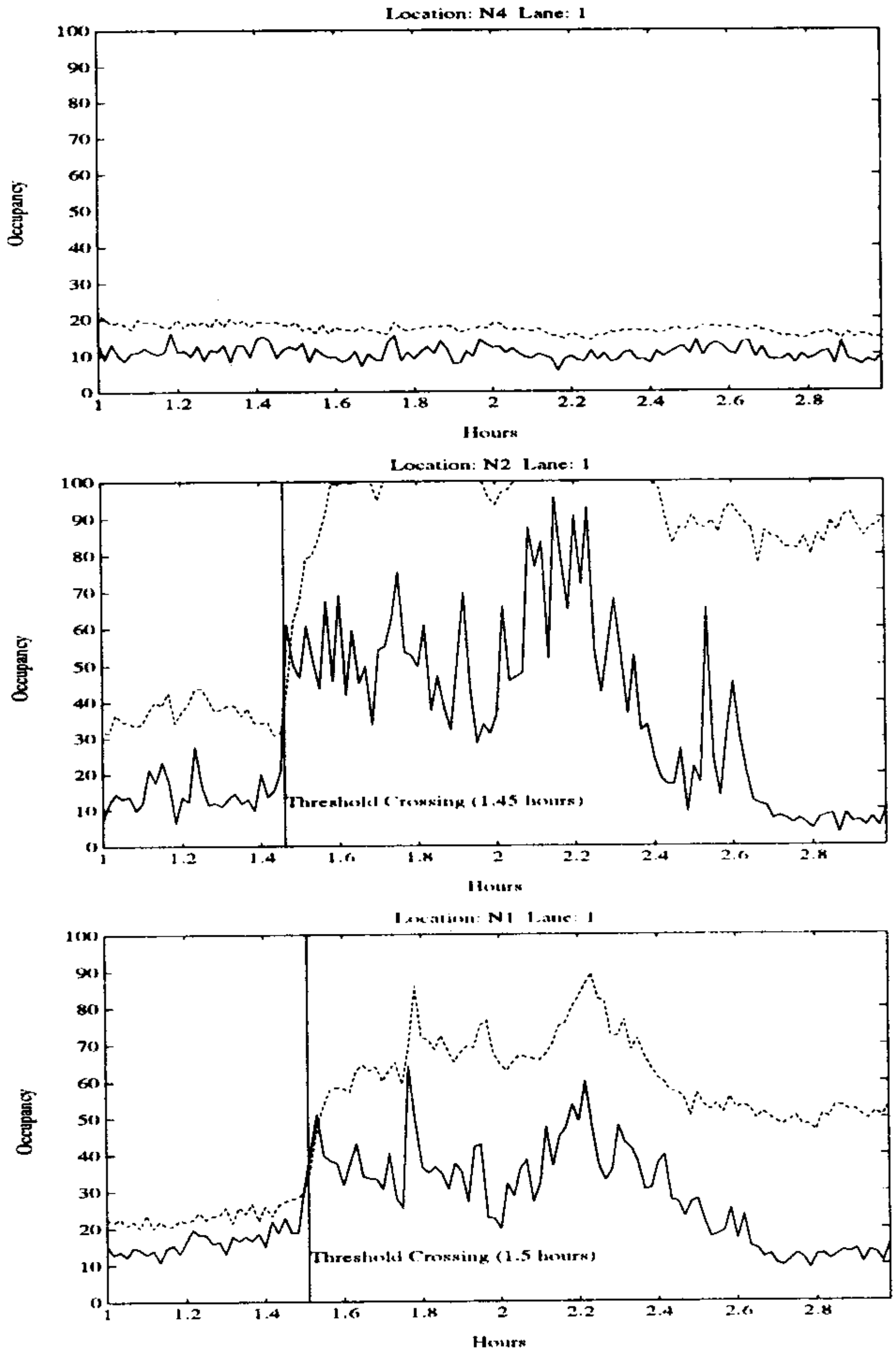


Figure 3.4: Observed occupancies (solid line) with threshold (dashed line) with the single anomaly identified.

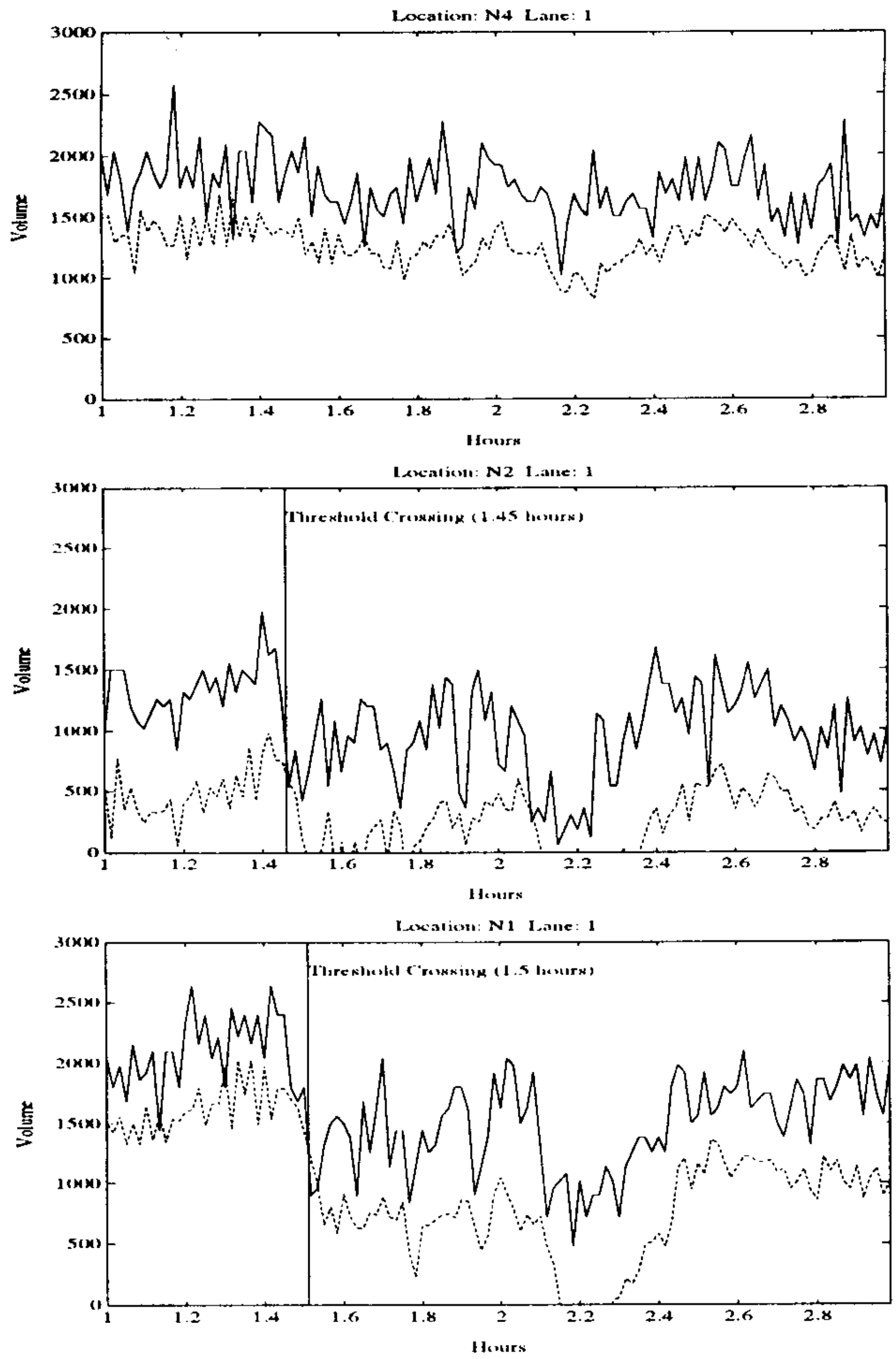


Figure 3.5: Observed volumes (solid line) with threshold (dashed line) with the single anomaly identified.

# 4. Conclusion

This report presented a methodology that can improve the capabilities of a traffic management system based on inductance loop sensors. It described a predictor/detector methodology suitable for identifying changes in the traffic state based on measured volumes and occupancies. The prediction of future values was done using an optimal recursive predictor based on a Kalman predictor formulation. The system model or transition matrix necessary for such a formulation was developed directly from the statistics of the problem. The transition matrix and the predictor formulation were used with inductance loop data from a number of sites on I-5 to demonstrate the validity of the prediction concept. Anomaly detection was presented using the predictor with a set of thresholds. Thresholds for detection of anomalies were developed based on the deviation of the observed values from the predicted values at that time step, and the thresholds were scaled based on the standard deviation of the data. Two simple thresholds (one on volume and one on occupancy) were used with the predictor and data from I-5 to demonstrate the detection of the onset of congestion and the increase in queue length on I-5. The methodology presented in this report combined observations about the properties of inductance loop data with an optimal recursive predictor to produce a new inductance loop based anomaly identification system.

# BIBLIOGRAPHY


[Boz84] S.M. Bozic. *Digital and Kalman Filtering*. Edward Arnold, 1984.

[CS90] M. Cremer and H. Schutt. A comprehensive concept for simultaneous state observation, parameter estimation and incident detection. *11th International Symposium on Transportation and Traffic Theory* , pages 95–111, 1990.

[Dai92] D.J. Dailey. Travel time estimation using cross correlation techniques. *Transpn Res. B*, 1992. Accepted 1/92.

[Lu90] J. Lu. Prediction of traffic flow by an adaptive predictive system. *Transpn Res. Rec. 1287*, 1990.

[OS84] I. Okutani and Y.J. Stephanedes. Dynamic prediction of traffic volume through kalman filtering theory. *Transpn Res. 18B* , pages 1–11, 1984.