Report No. DOT-TSC-RSPA-78-18

TRAFFIC SURVEILLANCE DATA PROCESSING IN URBAN FREEWAY CORRIDORS USING KALMAN FILTER TECHNIQUES

Paul K. Houpt
Michael Athans
Daniel G. Orlhac
William J. Mitchell

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge MA   02139

NOVEMBER 1978
FINAL REPORT

Prepared for

1. Report No. DOT-TSC-RSPA-78-18	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle TRAFFIC SURVEILLANCE DATA PROCESSING IN URBAN FREEWAY CORRIDORS USING KALMAN FILTER TECHNIQUES		5. Report Date November 1978
		6. Performing Organization Code
7. Author's) Paul K. Houpt, Michael Athans, Daniel G. Orlhac, William J. Mitchell		8. Performing Organization Report No. DOT-TSC-RSPA-78-18
9. Performing Organization Name and Address Massachusetts Institute of Technology * Electronic Systems Laboratory Department of Electrical Engineering and Computer Science, Cambridge MA 02139		10. Work Unit No. (TRAIS) RS905/R9516
		11. Contract or Grant No. DOT-TSC-849
12. Sponsoring Agency Name and Address U.S. Department of Transportation Research and Special Programs Administration Office of Transportation Programs Bureau Office of Systems Engineering, Washington DC 20590		13. Type of Report and Period Covered FINAL REPORT April 1975 - April 1978
		14. Sponsoring Agency Code

16. Abstract

Real-time surveillance of traffic conditions on urban freeway corridors using spatially discrete presence detectors is addressed. Using a finite-dimensional (macroscopic) fluid-analog model for freeway vehicular traffic flow, an extended Kalman filter is proposed as a data-processing algorithm to obtain minimum variance estimates of spatial mean speed and density. It is shown that certain model parameters associated with available roadway capacity can be estimated on-line with a variation of the extended Kalman filter, and furthermore, that the time signatures associated with these estimates provide quantitative information concerning the presence of anomalous (i.e., incident) traffic events. Performance of the surveillance algorithm is evaluated using a detailed, multi-lane microscopic vehicle simulation which retains a stochastic mix of driver-vehicle types and passing. Complexity problems associated with the computer implementation of the extended Kalman filter are addressed, and techniques for decentralized realization are proposed.

17. Key Words Estimation, Traffic Variables, Data Processing, Kalman filter, Speed, Headway, Density	18. Distribution Statement DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 198	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

PREFACE

This report is based in part on the S.M. theses by Daniel Orlhac, "Estimation of Traffic Variables via Extended Kalman Filter Methods," and by William J. Mitchell, "Estimation and Simulation of Freeway Traffic Flow Using Car Following and Fluid-Analog Techniques," submitted in August 1975, and in June 1977, respectively.

CONTENTS

CONTENTS (Concl'd)

ILLUSTRATIONS (Contd.)

TABLES

1. INTRODUCTION

1.1 PROBLEM PERSPECTIVE AND WORK SCOPE

The problem addressed in this research is the development of on-line computer algorithms for traffic surveillance data processing in urban freeway corridors. Our focus is on timely estimates of prevailing traffic conditions as measured in terms of aggregate traffic-flow attributes such as volume, speed, and density, over time periods from seconds to minutes. The traffic surveillance activities proposed here are viewed more to serve the needs of real-time traffic control, than the survey type data collection required for long-term traffic management and planning. Results are directed at being a complement to the overall real-time traffic management methodology proposed in this study [FR1] - [FR5]. The methodology has the unifying theme of optimal traffic management in urban freeway corridors. Surveillance needs are thus driven by the information requirements of the real-time control system [FR2], which requires spatially aggregate variables of mean speed* and mean-density* on roadway sections. A fundamental constraint that we have adopted, which reflects practical implementation, is that estimates of these spatially defined variables must be derived from convential "presence" type loop detectors, located at discrete (typically 1/2-mile) intervals along the roadway.

The approach embodies modern estimation theory techniques via the extended Kalman filter. Central to the filter development is a finite-dimensional fluid-analog model in which the freeway is viewed as a coupled set of ordinary non-linear differential equations. Each section or link is modeled by two state variables, the spatial mean speed and mean density. These variables typically describe the aggragate behavior of 1/2 mile roadway sections. Knowledge of parameter values in the link-state differential equations is shown to provide qualitative as well as quantitative information concerning link traffic flow behavior. These parameters vary with time, traffic conditions, and

---

* These quantities are defined precisely in Section 2.

Since the late 1960's, there have been many researchers involved in the development procedures for estimating mean-speed and mean density from loop-detector surveillance data. Typical density estimation schemes are described in Nahi [37], Gazis and Knapp [38], and Gazis and Szeto [20]. Mean speed estimation with detector loops has been proposed in Mikhalkin [13, 14]. All of these methods are based on very simple dynamic or static models of traffic flow. During conditions where flow conditions are very inhomogeneous (such as rush-hour or following incidents), such models do not apply and the corresponding performance of the surveillance algorithm deteriorates: increased error variances and biases typically appear in the estimates.

The Kalman-filtering approach employed here admits the inclusion of more complex dynamic models which govern the underlying variables being estimated. A Kalman filtering approach was used by Gazis and Szeto [20] in the context of tunnel traffic surveillance, but assumes a very simple dynamic model which does not extend to the freeway environment. Their results do, however, demonstrate the effectiveness of the procedure in that traffic surveillance application. Surveillance data are processed recursively as they become available, and estimates of parameters associated with the underlying model become available in addition to estimates of tunnel traffic density.

Essential elements in our approach using the extended Kalman filter are the dynamic model used and the techniques for processing loop detector data to be consistent with that model. Macroscopic traffic models first proposed by Payne [4] were employed. These models have increased complexity to capture the important qualitative traffic attributes during inhomogeneous conditions, but are not so complicated as to preclude algorithm design and implementation.

A closely related Kalman-filtering approach to ours using these models has recently been proposed by Grewal and Payne [28]. The important contribution of the Payne paper is a rigorous demonstration that certain parameters in the model can be identified from observations of mean speed and mean density. Processing of spatially discrete detector data to derive such observations, however, is not addressed. We provide an approach to the latter issue and an alternative parameterization of the model, from which on-line estimates

1.4 ORGANIZATION OF REPORT

The technical content of the remainder of the report is organized as follows: Section 2 introduces the notation, traffic flow, and detector models, and defines the detailed objectives of the surveillance algorithm development. Then in Section 3 the pre-processing procedure for loop-detector data is presented together with details of the extended-Kalman filter data-processing algorithm used in this study. A simple candidate study network and the microscopic vehicle simulation description are provided in Sections 4.1 to 4.3. They are followed by an evaluation of the algorithms from Section 3 in sections 4.4 to 4.6. Finally, implementation considerations for designs in Section 3 are briefly considered in section 4.7. Section 5 presents our conclusions and suggestions for future research.

a compressible fluid analog were made popular by Lighthill and Whitham ([1]), and proliferate in the literature of the later 1950's and early 1960's (see; e.g., Gazis [2]).

Macroscopic fluid-analog models appear to offer attractive advantages of mathematic simplicity while retaining correct qualitative real-time behavior. Although kinetic and continuum models for limited-access roadways have long appeared in the literature [1, 2], only recently have some of the discrepancies between macroscopic models and actual traffic behavior begun to be resolved, stemming from the work of Phillips [3].

2.1.1 Mathematical Model

Macroscopic models simplify the mathematics in the sense that individual vehicle behavior is aggregated into a one-dimensional continuum, analogous to the flow of fluid. However, it is nontrivial to produce a fluid approximation which both retains the correct dynamic qualitative behavior and, from an estimation algorithm design viewpoint, is sufficiently simple.

In a macroscopic traffic description, just as in a fluid of molecular particles, it is common to denote the speed (distance per unit time) of vehicles at position x at time t by $v(x,t)$, and the density (vehicles per unit distance) at position x at time t by $\rho(x,t)$. A third variable of interest is the flow $\phi(x,t)$ (in vehicles per unit time) passing point x at time t.

The model which describes the evolution in time and space of $v(x,t)$ and $\rho(x,t)$ is derived in a sequence of steps:

a. specify a conservation of vehicles (mass-balance) equation, and

b. postulate an explicit form for acceleration response of drivers to local conditions.

Completion of both steps in either the deterministic [4], or statistical [3,5,6], framework results in a pair of simultaneous, non-linear partial differential equations for $v(.,.)$ and $\rho(.,.)$:

$$\frac{\partial}{\partial t} \rho(x,t) + \frac{\partial}{\partial x} \phi(x,t) = 0 \quad , \tag{2.1}$$

$$\frac{dv}{dt}(x,t) \triangleq \frac{\partial v}{\partial t}(x,t) + v\frac{\partial v}{\partial x}(x,t) = f(\rho,v) \quad . \tag{2.2}$$

Equation (2.1) is simply a statement of the principle of conservation of vehicles. It is the continuum equation from the kinetic theory of gases [3,5]. Equation (2.2) postulates that the local acceleration of drivers can be explicitly parameterized as a function of local traffic conditions in the stream.

Over the years, there have been numerous suggestions for the specific choice of $f(\cdot,\cdot)$. The key idea in picking $f(\cdot,\cdot)$ is to posit a mechanism for how drivers physically respond to changing traffic conditions, and then examine the consequences of these behavioral assumptions [7,8].

The most important behavioral attributes which when modeled retain correct qualitative traffic flow behavior appear to be the desire of drivers to maintain:

a. a safe (typically speed-dependent) headway,

and

b. a certain maximum speed provided no other cars are in the way.

The exact choice for the function $f(\cdot,\cdot)$ will establish qualitative properties such as _stability_ and _steady-state_ solutions (if they exist) to (2.1) and (2.2). By formal derivation from classical car following theory (e.g., Gazis [7]), one possible relationship for $f(\cdot,\cdot)$ is

$$\frac{dv(x,t)}{dt} = - \frac{1}{T} \left[v(x,t) - v_e(\rho(x,t)) \right], \tag{2.3}$$

where T represents driver reaction time, and $v_e(\rho)$ is an _equilibrium speed distribution_. The $v_e(\rho)$ curve can be formally derived in terms of parameters of the particular car-following law used. In practice, however, the curve is fitted to steady-state samples from observed traffic data to average _individual_ driver characteristics at a particular location. The $v_e(\rho)$ curve, however obtained, plays a crucial role in traffic behavior analysis discussed in section 2.4.

Intuitively the relationship (2.3) states that on the average (over the ensemble of the driver population), drivers adjust their speeds as a function of _local_ _density_ alone. Empirical evidence has shown that this relationship is typically monotone non-increasing with density.

While the steady-state characteristics are adequately captured by the

τ, ν, and $v_e(\rho)$ jointly affect qualitative dynamic behavior, and are typically obtained from a calibration against historical data or heuristic assumptions. In what follows, we show that the most important parameters affecting the qualitative dynamic behavior are those of the equilibrium speed curve $v_e(\rho)$. Attributes of $v_e(\rho)$ and the closely related the fundamental diagram are discussed further in section 2.4.

Note that the model describes the <u>longitudinal</u> dynamics of traffic flow: there is no explicit representation of the <u>lateral</u> dynamics that result from passing maneuvers and other forms of lane changing. This may seem like a highly unrealistic assumption for describing traffic on freeways. However, the validity of making this approximation will be demonstrated in the performance of the estimation algorithm. We show this by evaluating the surveillance algorithm in a microscopic simulation where passing is present. A significant finding to be demonstrated is that qualitative effects of lateral dynamics on reducing roadway capacity can be modeled by a change in the parameters of the $v_e(\rho)$ curve. Thus, the model can be used for multi-lane roadways.

2.2 FINITE-DIMENSIONAL APPROXIMATION TO PAYNE MODEL

In the form (2.7), (2.8), the Payne continuum model is not very useful for control or surveillance algorithm design since it is a non-linear partial differential equation in $\rho(x,t)$, $v(x,t)$. A pivotal approach to simplification is achieved by spatial discretization of the continuum model. This discretization has two purposes:

a. The model mathematical complexity is greatly reduced, and

b. ariables in the discretized model are more easily observed with available sensors.

The approximation is obtained by partitioning the freeway into N spatially discrete sections, and replacing $\rho(\cdot,\cdot)$ and $v(\cdot,\cdot)$ on each section by a certain spatial average. These definitions are then used in the continuum model using finite difference techniques in a consistent fashion that retains certain boundary conditions.

2.2.1 <u>Spatial Mean Variables</u>

For a section of length Δx beginning at x define

$$M(x,t,\Delta T) \equiv \int_t^{t+\Delta T} \rho(x,\tau)v(x,\tau)d\tau. \tag{2.16}$$

We remark in passing that while (2.15) will always be true, it is generally <u>not</u> the case that $\phi(x; t,\Delta T)$ can be obtained by the product $\bar{v}(x,\Delta x; t)\cdot\bar{\rho}(x,\Delta x; t)$. Conditions when this holds approximately are discussed below.

2.2.2 Finite-dimensional Approximate Model

Payne [4] proposed use of the spatial mean variables defined above together with finite differences to replace spatial derivatives in (2.7) and (2.8), to reduce the partial differential equations in space and time to ordinary differential equations in time alone. The procedure is to:

Step 1

Divide the roadway into spatially homogeneous subsections, $[x_j, x_j+\Delta x_j]$, where x_j partitions need not be uniform.

Step 2

Define: (Figure 2.2)

$$\bar{\rho}_j(t) \equiv \bar{\rho}(x_j, \Delta x_j; t), \tag{2.17}$$

$$\bar{v}_j(t) \equiv \bar{v}(x_j, \Delta x_j; t), \tag{2.18}$$

and approximate spatial derivatives by

$$\frac{\partial \rho}{\partial x}(x,t)\bigg|_{x\epsilon[x_j,x_j+\Delta x_j]} \approx \frac{\bar{\rho}_{j+1}(t)-\bar{\rho}_j(t)}{\Delta x_j}, \tag{2.19}$$

$$\frac{\partial v}{\partial x}(x,t)\bigg|_{x\epsilon[x_j,x_j+\Delta x_j]} \approx \frac{\bar{v}_j(t)-\bar{v}_{j-1}(t)}{\Delta x_j}, \tag{2.20}$$

for $j = 0,1,2,\ldots,N$,

where:

N = number of sections (defined <u>a priori</u>),

Δx_j = length of section j,

$$x_0 = \text{start of freeway section modeled, and}$$

$$x_j = \text{beginning of section } j = x_0 + \sum_{k=1}^{j} \Delta x_k .$$

Remarks

a. Note that (2.19), (2.20) are not the usual finite difference approximations defined in terms of <u>point</u> variables, e.g.,

$$\frac{\partial \rho}{\partial x}(x, t) \approx \frac{\rho(x_{j+1}, t) - \rho(x_j, t)}{\Delta x_j}, \tag{2.21}$$

The spatial mean approximation is used because it is formally convenient to express the model directly in terms of variables which can be observed, as will be demonstrated below.

b. The density gradient is approximated by a <u>forward</u> difference, while the velocity gradient is approximated by a <u>backward</u> difference. Payne [4] motivates the density approximation choice from heuristic physical reasoning (drivers anticipate primarily in a forward direction), and the backward-velocity choice for stability of the discretization.

Step 3

Using the approximations in Step 2, substitute in (2.8) to obtain:

$$\frac{d\bar{v}_j(t)}{dt} = -\bar{v}_j(t) \left[\frac{\bar{v}_j(t) - \bar{v}_{j-1}(t)}{\frac{1}{2}(\Delta x_j + \Delta x_{j-1})} \right] - \frac{1}{\tau} \left[\bar{v}_j(t) - v_e^{\,j}(\bar{\rho}_j(t)) \right]$$

$$- \frac{\nu}{\tau} \frac{1}{\bar{\rho}_j} \left[\frac{\bar{\rho}_{j+1}(t) - \bar{\rho}_j(t)}{\frac{1}{2}(\Delta x_j + \Delta x_{j-1})} \right] . \tag{2.22}$$

The terms on the right hand side of (2.22) are often referred to as "<u>convection</u>," "<u>relaxation</u>,"and "<u>anticipation</u>," respectively, corresponding to the physical interpretation in both gases and car-following theory.

Step 4

Approximate the continuity equation (2.7) by the simple conservation expression
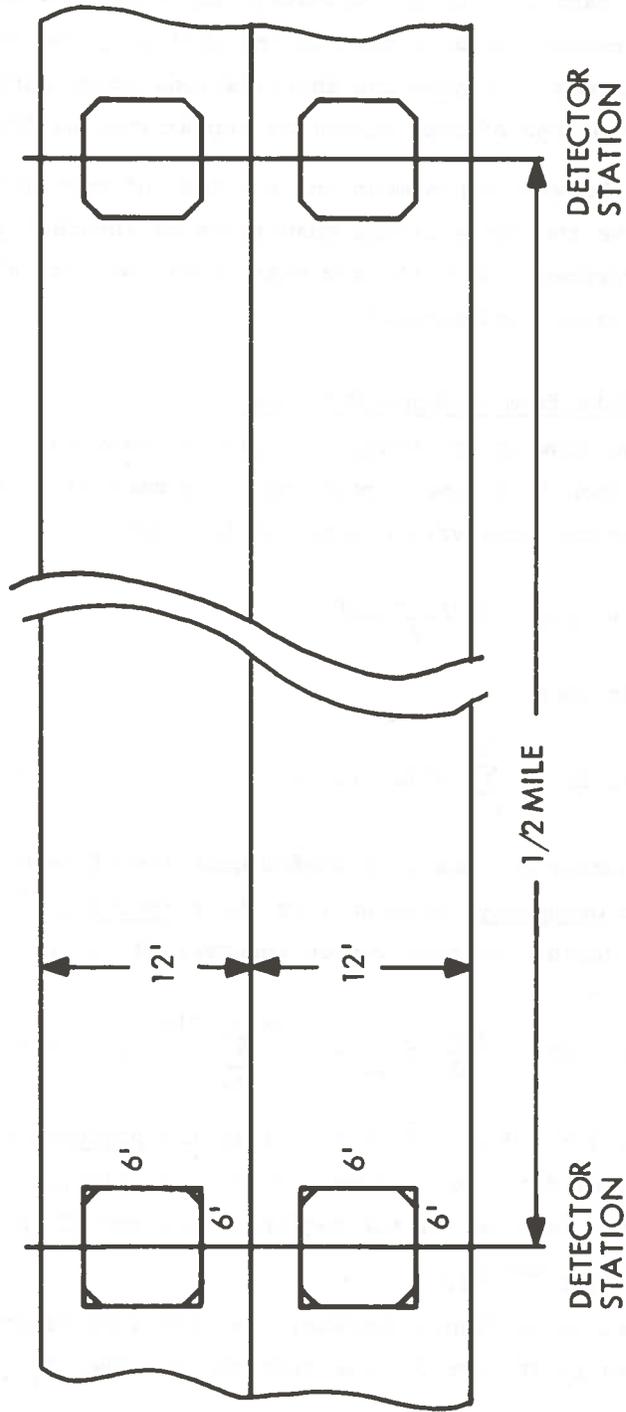
Figure 2.3: Typical Presence Detector Configuration on Freeway
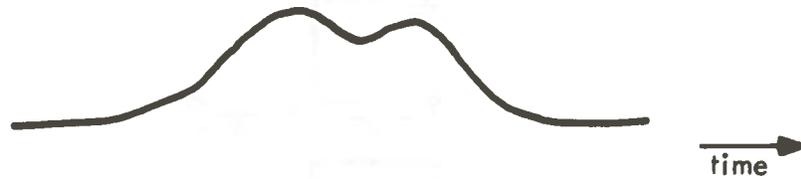
An analog signal resulting from the passage of
a single vehicle

presence
pulse
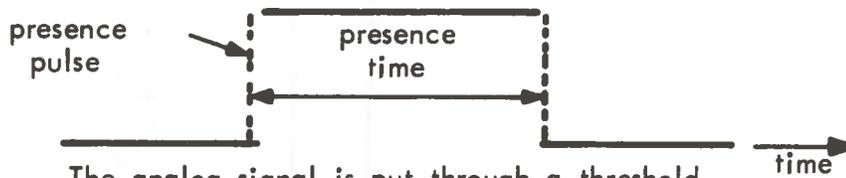
presence
time

The analog signal is put through a threshold
device to yield a binary signal in time

HIGH

LOW

000000000000 11111111111111111 000000000000   time
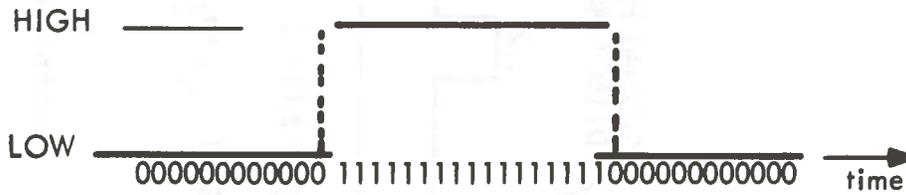
Signal is sampled (15 – 60 time sec.)

1: "vehicle present" bit

0: "vehicle absent" bit

Figure 2.4:   Presence Detector Signal Associated with
Single Vehicle Passage

by sampling errors of $\pm\delta T$, where δT is the sampling period. Errors resulting from $t_{i,I}, t_i$, and F sampling have a more pronounced effect at low density flows, or where short observation intervals, ΔT, result in $M^1(x;t,\Delta T)$ being "small." Analysis and compensation for the effects of such errors will be provided by the Kalman filter algorithm.

Intuitively at least, the quantity occupancy is proportional to density. At higher density, one would expect to see a vehicle over a detector a greater percentage of the time and conversely. However, a precise model for this intuitive relationship must be developed with care, because occupancy is a time-average quantity at a point, whereas density in our model is a spatial-average quantity, at a given time. Conditions under which one can be derived approximately from the other are obtained in Kurkjian [12] and Mikhalkin [13,14], and summarized in section 3.3.4 of this report.

2.4 QUALITATIVE SIGNIFICANCE OF MODEL AND TRAFFIC VARIABLES IN SURVEILLANCE

Knowledge of the macroscopic traffic variables defined in sections (2.2), (2.3), can provide considerable qualitative as well as quantitative insight into prevailing traffic conditions. Since most of this report focuses on quantitative calculation of the relevant quantities, we feel it is important to grasp why certain variables are significant in predicting traffic behavior useful for decisionmaking, and how these quantities relate to the model and surveillance data available. Qualitative aspects are particularly important because for many of the estimated quantities there is no rigorous correspondence between the variable in the model and the real world.

2.4.1 Equilibrium Speed Curve

The equilibrium speed curve, $V_e(\rho)$ in (2.8) and (2.22), plays a particularly simple and yet important role in characterizing qualitative roadway conditions. By defining

$$\phi_e(\rho) \equiv \rho v_e(\rho), \tag{2.27}$$

one obtains a classical parameterization of traffic behavior known as the
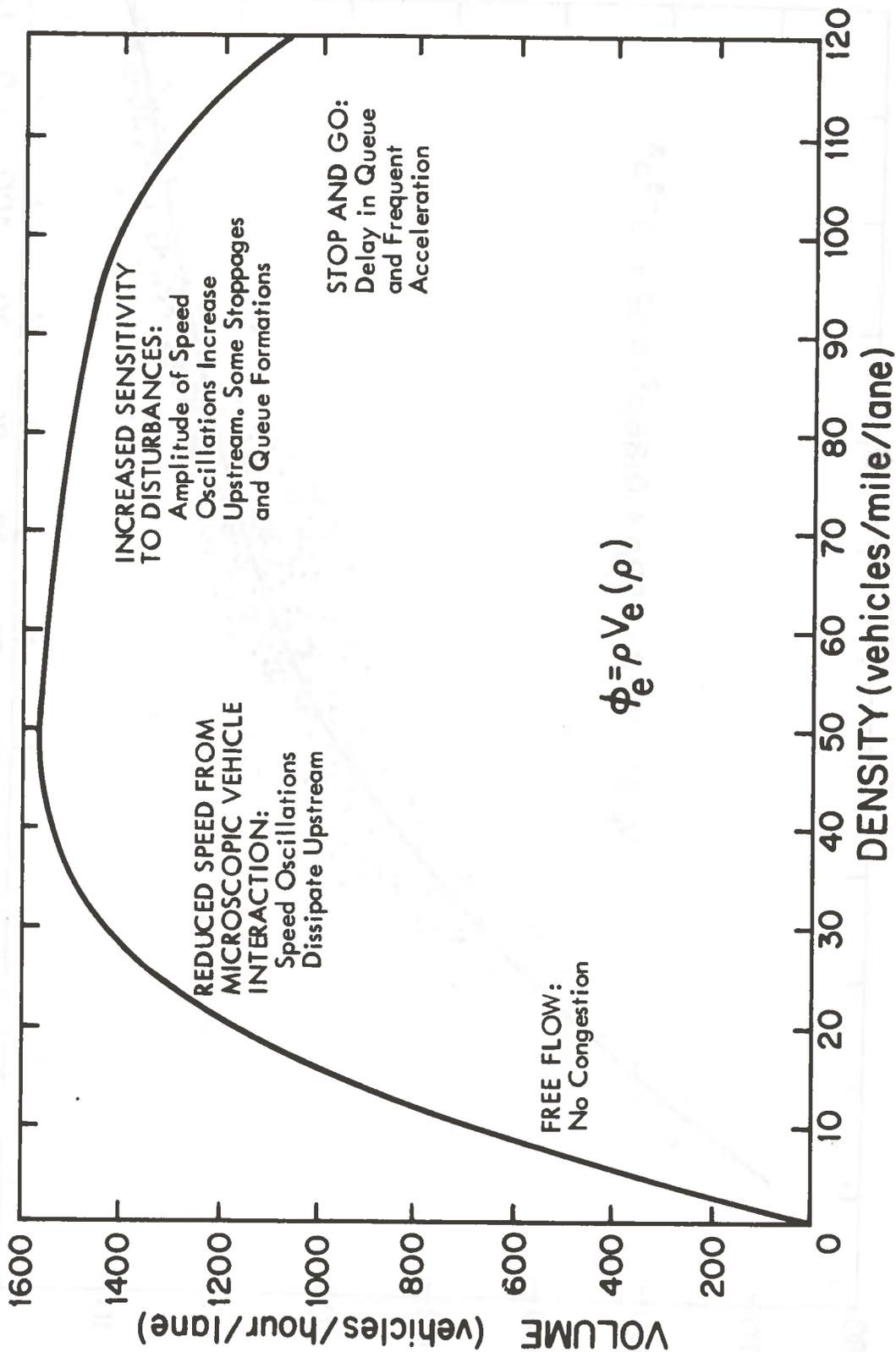
Figure 2.6: Model of Fundamental Diagram

1. Stagnation or <u>jam density</u> (ρ_{jam}):

$$\rho_{jam} \equiv \begin{Bmatrix} \text{density at} \\ \text{which equilibrium} \\ \text{speed} \quad V_e(\rho_{jam})=0 \end{Bmatrix} \quad ,$$

2. <u>Free speed</u> (V_f):

$$V_f \equiv \begin{Bmatrix} \text{speed drivers} \\ \text{assume at} \\ \text{low density} \end{Bmatrix} \quad , \text{ and}$$

3. <u>Available Capacity</u> (ϕ_{max}):

$$\phi_{max} \equiv \begin{Bmatrix} \text{maximum steady} \\ \text{state flow found} \\ \text{from:} \\ \quad \max_{\rho} \quad V_e(\rho) \end{Bmatrix} \quad ,$$

In the model, jam density and free speed have intuitive meaning, and can be relatively easily measured. In real traffic, ρ_{jam} and V_f have meaning only as random variables over the population of drivers. In the model, these are mean values of the observed distribution. Typically, ρ_{jam} is taken to be the bumper-to-bumper concentration of vehicles for the roadway under study, and V_f the speed limit.

The concept of <u>capacity</u> is somewhat more elusive than the other two, and requires some care to establish our parameterization of this quantity. In the traffic engineering community, the capacity of a roadway is defined as a <u>fixed</u> quantity depending on lane widths, surface, grade, curvature, and other geometric properties [17]. One could, therefore, use the value supplied by the appropriate handbook, and define it as the maximum <u>capacity</u>.

Clearly, the steady state ability of a roadway to support a capacity volume of traffic depends on factors <u>other</u> than strictly geometric roadway features. Two important factors in our view are <u>environment</u> associated with weather, and the occurrence of <u>traffic incidents</u>. During inclement weather, the maximum throughput of a roadway is lower than for dry surface conditions, due in part to more conservative driver behavior. Similarly, the occurrence of accidents, presence of stalled vehicles, and spilled loads in traffic lanes reduces the maximum steady-state throughput achievable. We shall refer to this maximum

$$V_e(\bar{\rho}) = v_{free} - \left(\frac{v_{free}}{\rho_{jam}}\right)\bar{\rho}$$

Typical Values

$v_{free} = 55 \text{(mph)}$
$\rho_{jam} = 225 \text{ (veh/mi)}$

v_{free}

50

ρ (veh/mi)

100   200

ρ_{jam}

O
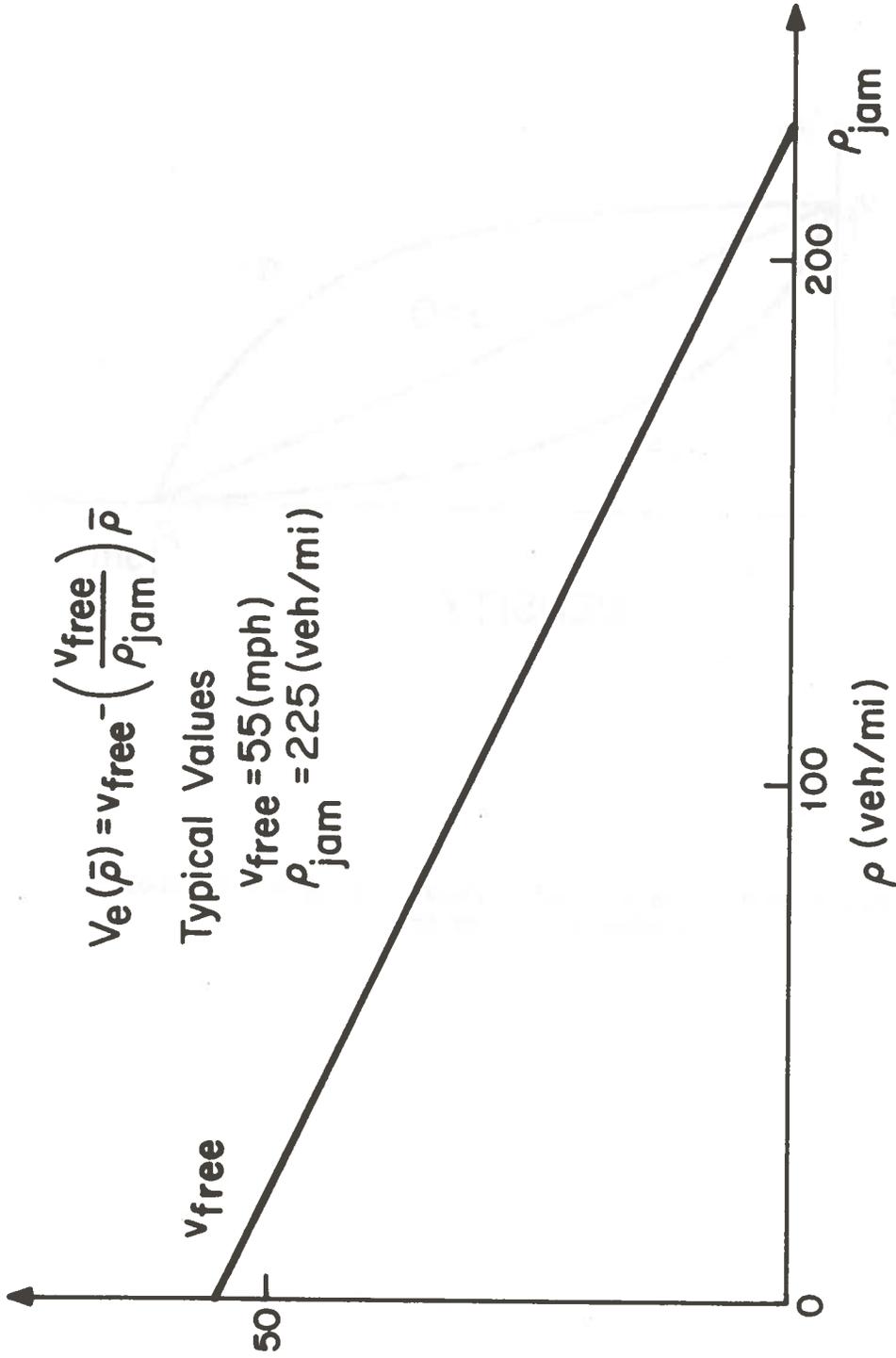
Figure 2.8: Linear Equilibrium Speed-Density Relationship
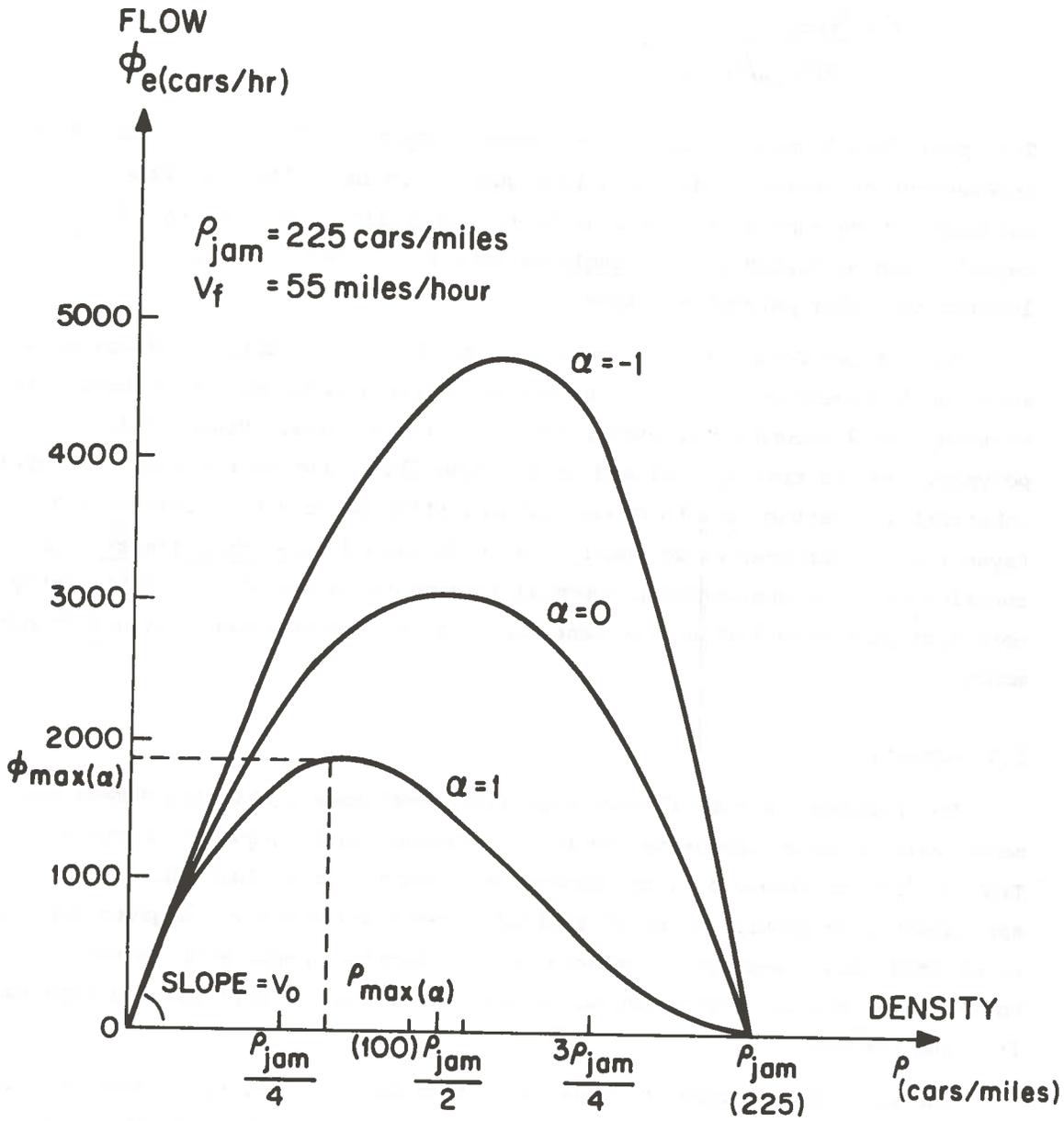
	$\alpha = -1$	$\alpha = 0$	$\alpha = 1$
$\rho_{max}(\alpha)$	$\frac{2}{\sqrt{3}}\left(\frac{\rho_{jam}}{2}\right)$	$\frac{\rho_{jam}}{2}$	$\frac{2}{3}\left(\frac{\rho_{jam}}{2}\right)$
$\phi_{max}(\alpha)$	$\frac{8}{3\sqrt{3}}\left(\frac{V_f \cdot \rho_{jam}}{4}\right)$	$\frac{V_f \cdot \rho_{jam}}{4}$	$\frac{16}{27}\left(\frac{V_f \cdot \rho_{jam}}{4}\right)$

Figure 2.10: Effect of Parameter α in $V_e(\rho)$ on Parobolic Fundamental Diagram on $\phi_e(\overline{\rho})$

The figure shows a plot with v on the vertical axis (miles/hour) ranging from 0 to 60, and ρ on the horizontal axis (cars/mile/lane) ranging from 0 to 250. The curve is flat at v_{free} until ρ_{free}, then decreases logarithmically to zero at ρ_{jam}.

$$v_e(\rho) = \begin{cases} v_{free} & 0 \leq \bar{\rho} \leq \rho_{free} \\ 0 & \rho > \rho_{free} \\ k \ \ln \dfrac{\rho_{jam}}{\bar{\rho}} & \text{otherwise} \end{cases}$$

where

$$k = \frac{v_{free}}{\left[\ln \rho_{jam}/\rho_{free}\right]}$$

typical values:

$$\rho_{free} = 75 \text{ veh/mi}$$

$$\rho_{jam} = 225 \text{ veh/mi/lane}$$

$$v_{free} = 55 \text{ mi/hr}$$
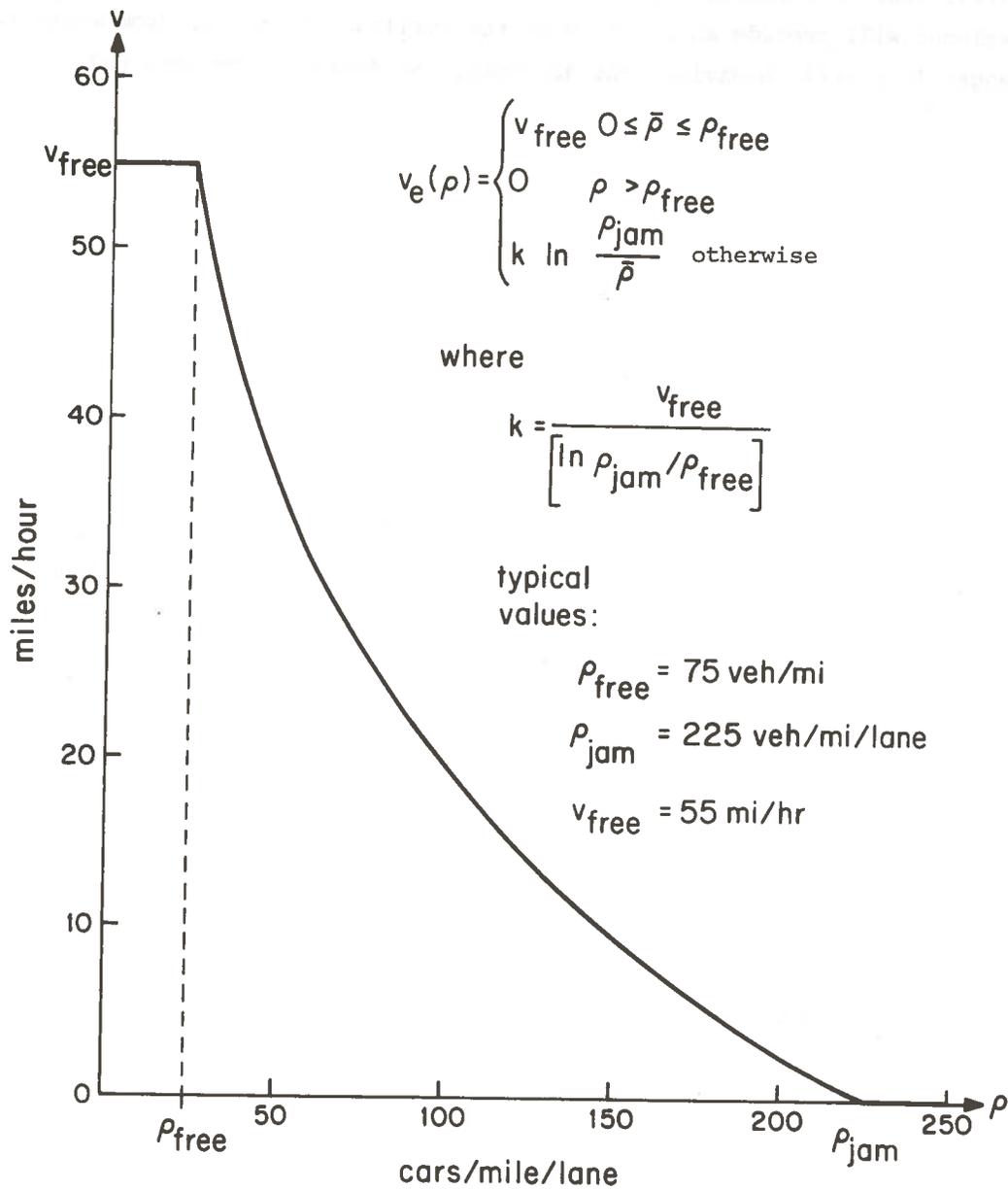
Figure 2.11: Logarithmic Equilibrium -- Speed-Density Relationship

3. SURVEILLANCE ALGORITHM DEVELOPMENT VIA KALMAN FILTER TECHNIQUES

In this section we develop the principal data processing algorithms proposed for application to traffic surveillance. Modern estimation theory and techniques, and in particular, the so called Kalman filtering method form the foundation of our approach. Kalman filtering techniques have enjoyed wide-spread application in the aerospace community since the early 1960's in everything from intertial navigation to control configured aircraft [19]. With the increased use of digital-processing techniques in traffic management, these modern estimation techniques are finding their way into such activities as density estimation on bridges and in tunnels (e.g., and Gazis [20] and Szeto [20]). The "filter" is essentially a recursive algorithm which processes the incoming data as they are received to produce estimates of the underlying system variables which are optimum in the sense of minimum error variance. The underlying state variables in this study are the spatial mean speed and mean density defined in Section 2.2.

The classical Kalman filter algorithm has its roots in linear least-squares optimal estimation; when the system is linear and the uncertainties have prescribed independence properties, the resulting estimates are also optimum in the sense of minimum mean square error. In the context of traffic surveillance, the models are non-linear, and the uncertainties are seldom white Gaussian processes. For the non-linear setting, a widely used filtering technique is to exploit a linearized or "extended" version of the Kalman filter (EKF). Properties of the EKF are summarized in this section as they pertain to the traffic estimation algorithm, including standard procedures by which unknown model parameters can be augmented to the model states and estimated. We will then show in section 4 how the EKF may be used with the model from section 2 to estimate the spatial mean variables and available capacity as parameterized by the fundamental diagram of traffic.

3.1 DISCRETE TIME EXTENDED KALMAN FILTER: STRUCTURE AND UNDERLYING ASSUMPTIONS

3.1.1 System Model

The model described in section 2 for traffic-flow behavior can be written in the form:

$$\underline{y}(t_k) = g\ (\underline{x}(t_k))\ +\ \underline{\lambda}(t_k), \tag{3.4}$$

or in more compact notation*,

$$\underline{y}_k = \underline{g}\left(\underline{x}_k\right) + \underline{\lambda}_k\ , \tag{3.4^1}$$

where:

$\underline{y}_k$ is an m-vector, $m \neq n$ in general, with components $y_k^i = g_i\ (\underline{x}_k) + \lambda_k^i$
i=1, ..., M, which are functions which will be specified in the traffic
setting in section (3.3). The m-vector discrete <u>observation</u> <u>noise</u> sequence,
$\{\underline{\lambda}_k\ \}$ is assumed to have statistics:

$$E\ \{\underline{\lambda}_k\} = 0 \qquad \text{all } k= 1,2,\ldots, \tag{3.5}$$

and positive definite covariance matrix,

$$E\{\underline{\lambda}_k\ \underline{\lambda}_m\} = \Lambda\ \delta_{km}. \tag{3.6}$$

By assumption, $\{\underline{\lambda}_k\ \}$ and $\{\underline{w}\ (t)\}$ are mutually uncorrelated, and uncorrelated
with the <u>initial</u> <u>conditions</u>, $\underline{x}(t_o)$, having

$$E\{\underline{x}\ (t_o)\} \equiv \underline{\bar{x}}_o\ , \tag{3.7}$$

$$E\{(x(t_o)- \underline{\bar{x}}_o)\ (x(t_o) - \underline{\bar{x}}_o)^T\} \equiv \Sigma_o\ . \tag{3.8}$$

---

* For notational convenience, we will equivalently refer to sampled varaibles
via: $\underline{z}_k \equiv \underline{z}(t_k)\ \equiv \underline{z}\ (k\ \Delta\ T)$.

δ_{km} is the Kronecker delta function, $\delta_{km} =1$, k=m; $\delta_{km} =0$ otherwise .

Because the dynamic model is non-linear, the Kalman filter algorithm applied to the <u>linearized</u> system model is strictly suboptimal with respect to the specified performance criterion [19,21]. This <u>extended-</u>Kalman filter, however, has been widely used in non-linear applications with mixed experiences (see [26] for surveys of EKF applications as well as discussion of related non-linear estimation techniques). To summarize the EKF algorithm, we define the following quantities, in addition to $\hat{\underline{x}}_{k|k}$:

$$\hat{\underline{x}}_{k+1|k} \overset{\Delta}{\equiv} \quad \text{\underline{Predicted} estimate of } \underline{x}_{k+1} \text{ before observing} \qquad (3.11)$$
$$\underline{y}_{k+1} \text{ (i.e. given data up to time k),}$$

$$\sum_{k+1|k} \overset{\Delta}{\equiv} \quad \text{error covariance of predicted estimate } \hat{\underline{x}}_{k+1|k}, \qquad (3.12)$$

$$= E \left\{ (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1|k}) (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1|k})^T \right\},$$

$$\hat{\underline{x}}_{k+1|k+1} \overset{\Delta}{=} \quad \text{\underline{updated} or "current" estimate of } \underline{x}_k \text{ given} \qquad (3.13)$$
$$\text{new data } (\underline{y}_{k+1}),$$

$$\sum_{k+1|k+1} \overset{\Delta}{=} \quad \text{\underline{updated} estimate error covariance of } \underline{x}_k \qquad (3.14)$$
$$\text{given new data.}$$

Related quantities of interest are:

$$\hat{\underline{y}}_{k+1|k} \overset{\Delta}{=} \quad \text{\underline{Predicted} \underline{observation} of } \underline{x}_{k+1} \text{ given data up} \qquad (3.15)$$
$$\text{to time k,}$$
$$\overset{\Delta}{=} \quad \underline{g}(\hat{\underline{x}}_{k+1|k}),$$

$$\underline{r}_{k+1} \overset{\Delta}{=} \quad \text{residual process at time k+1,} \qquad (3.16)$$
$$\overset{\Delta}{=} \quad \underline{y}_k - \hat{\underline{y}}_{k+1|k}.$$
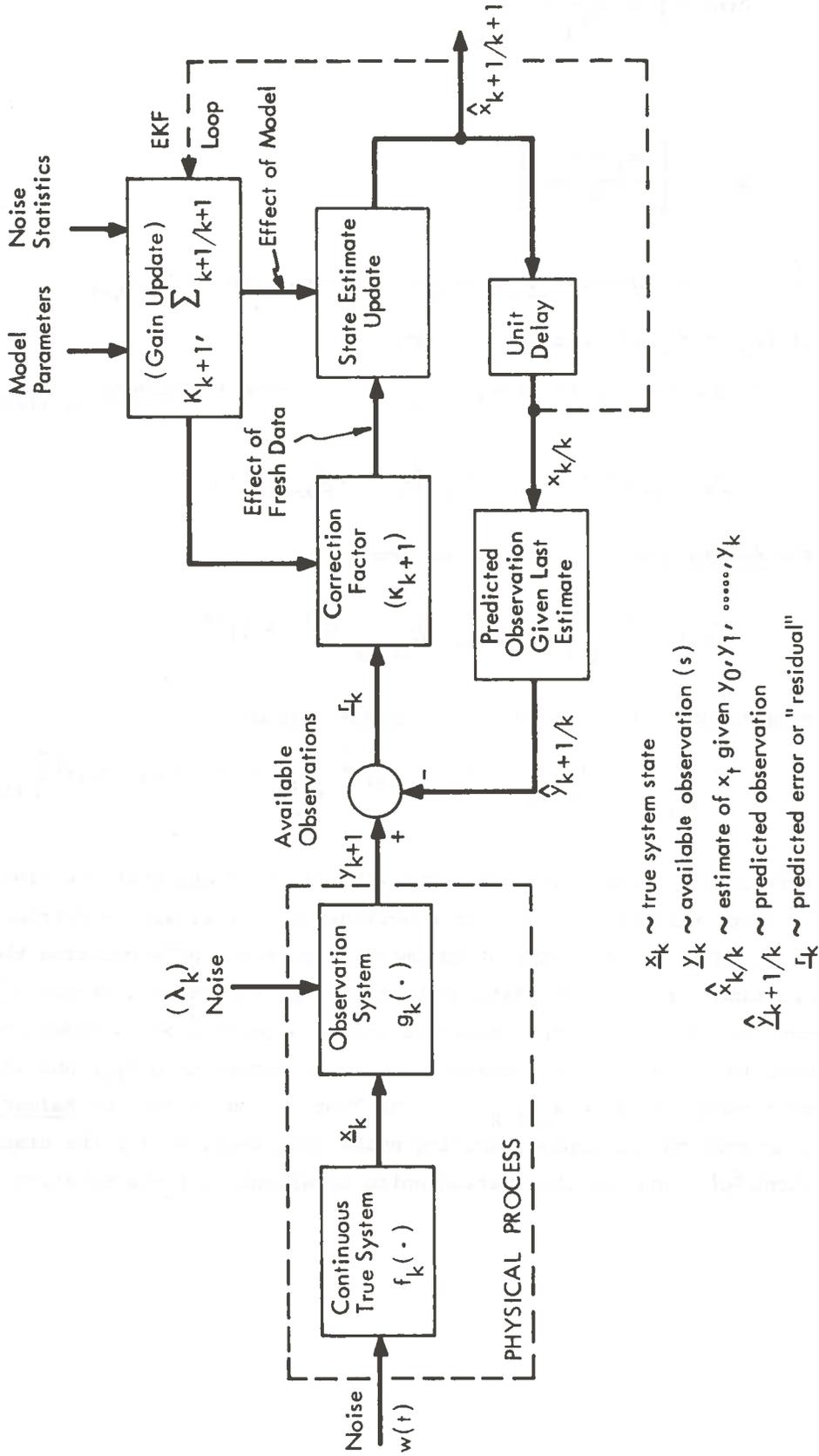
Figure 3.1: Continuous-Discrete Extended Kalman Filter Algorithm

x_k ~ true system state

y_k ~ available observation (s)

$\hat{x}_{k/k}$ ~ estimate of x_t given $y_0, y_1, \ldots, y_k$

$\hat{y}_{k+1/k}$ ~ predicted observation

r_k ~ predicted error or "residual"
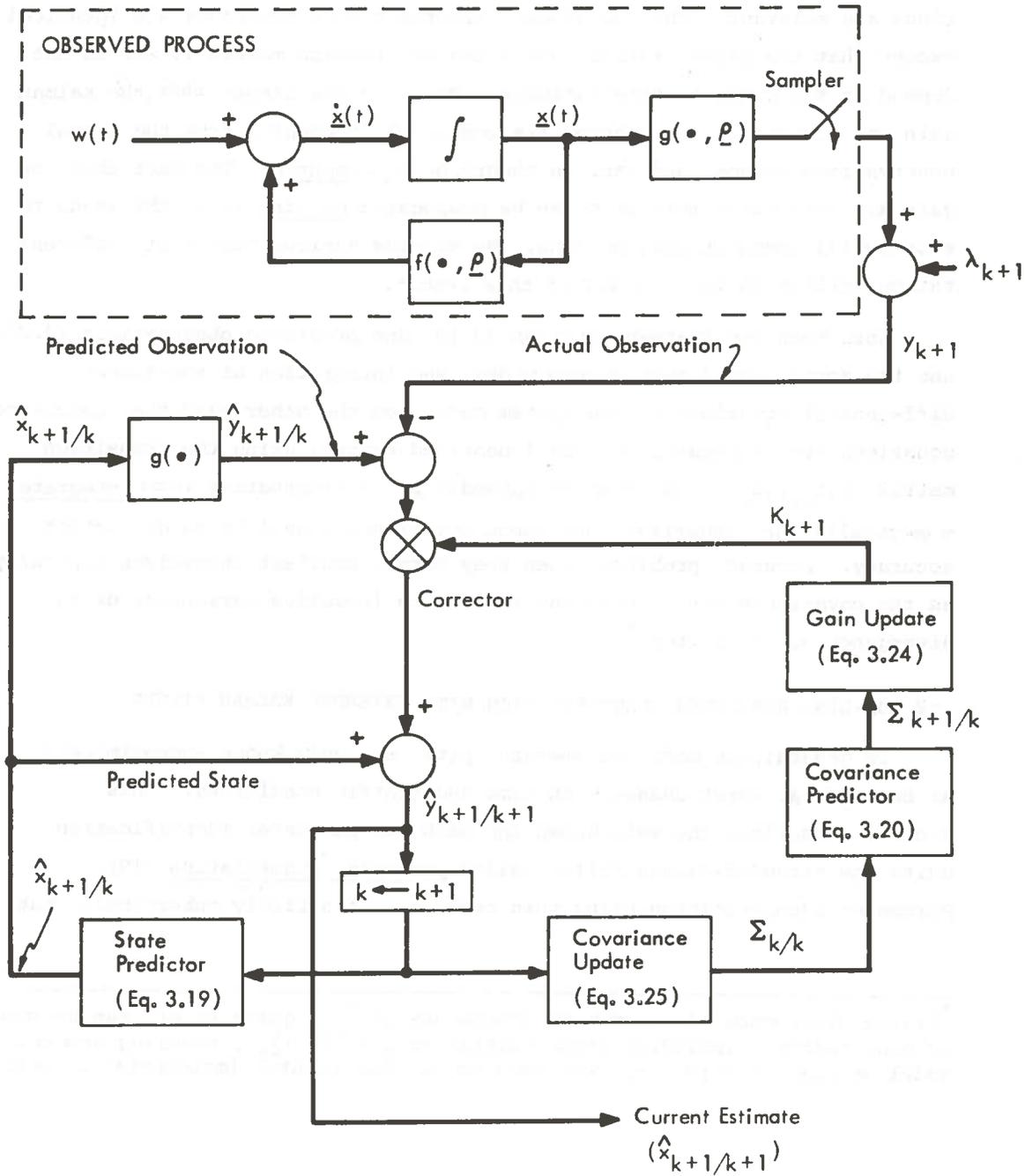
Figure 3.2: Data Flow in Continuous-Discrete Extended Kalman Filter

in a wide spectrum of applications is known to prevent divergence of the filter; even where divergence is not a problem, mean-square error generally improves (see; e.g., (26) for a survey of applications as well as alternative techniques for parameter identification).

3.2.1 Continuous-Discrete Filter: Unknown Parameters

The EKF described in section 3.1 is readily modified to permit application to estimating uncertain parameters in the model. Unknown parameters are treated as state variables and adjoined to the "physical" states in such a way that they appear as explicit estimates in the Kalman filter algorithm. Starting from

$$\dot{\underline{x}}(t) = \underline{f}[\underline{x}(t), \underline{p}(t)] + \underline{w}(t) . \tag{3.1}$$

Consider

$$\delta\underline{x}(t) = \underline{x}(t) - \overline{\underline{x}}(t), \tag{3.26}$$

$$\delta\underline{p}(t) = p(t) - \overline{p} , \tag{3.27}$$

where $\overline{x}(t)$ and $\overline{p}$ are nominal state and parameter values. Then as in Appendix B, the linearized model for small changes in $\delta\underline{x}(t)$, $\delta\underline{p}$ are written

$$\frac{d}{dt} \delta\underline{x}(t) = F(t)\delta\underline{x}(t) + L(t) \delta\underline{p}(t) + \underline{w}(t), \tag{3.28}$$

where:

$$F(t) = \left[\frac{\partial f_i[\overline{x}(t),\overline{p}]}{\partial x_j} \right]_{n \times n} , \tag{3.29}$$

$$L(t) = \left[\frac{\partial f_i(\overline{x}(t),\overline{p})}{\partial p_j} \right]_{n \times np} , \tag{3.30}$$

so that at sampling instants

$$\delta\underline{x}_{k+1} = \Phi(k+1,k)\delta\underline{x}_k + \psi(k+1,k)\delta p + \underline{\xi}_k , \tag{3.31}$$

(i.e, δp_k does not change with time) then (3,31) and (3.34) can be combined into:

$$
\begin{bmatrix} \delta x_{k+1} \\ \hdashline \delta p_{k+1} \end{bmatrix}_{n+np} = \begin{bmatrix} \Phi_{nxn} & \psi_{nxnp} \\ 0 & I_{npxnp} \end{bmatrix} \begin{bmatrix} \delta x_k \\ \delta p_k \end{bmatrix} + \begin{bmatrix} \xi_k \\ \hdashline 0 \end{bmatrix} , \qquad (3,37)
$$

let the _augmented variables_ be denoted:

$$
\delta x_k^a \triangleq (\delta x_{k+1}, \delta p_{k+1}),
$$

$$
\hat{\Phi}^a(k+1,k) \triangleq \begin{bmatrix} \Phi & \psi \\ 0 & I \end{bmatrix} \bar{x} = \hat{x}_{k|k} \quad ,
$$

$$
\xi_k^a \triangleq (\underbrace{\xi_k, \ 0 \ )}_{h_p\text{-entries}} \quad ,
$$

and similarly with:

$$
\delta y_k = [\hat{G}_k \ \hat{N}_k]\delta x_k^a + \lambda_k , \qquad (3.38)
$$

then associating:

$$
\hat{G}_k^a = [\hat{G}_k \ \hat{N}_k] \Big|_{\bar{x} = \hat{x}_{k|k}} , \qquad (3,39)
$$

we have

$$
\delta x_{k+1}^a = \hat{\Phi}^a(k+1,k)\delta x_{k+1}^a + \xi_k^a , \qquad (3.40)
$$

3.2.3 Comments on Parameter Identification

It is obvious that using the extended Kalman filter for adaptive param-
eter identification greatly increases its complexity. Although straight-
forward in theory, the increased computational load required on-line may be
prohibitive. When there is no intrinsic interest in the parameter values
themselves, a number of tricks are possible to eliminate some of the required
equations (see; e.g., Schmidt (27). They are, of course, heuristics (in general)
for which no guarantees of convergence can be provided a priori.

As motivated in the introduction, there is in the traffic context
interest in knowing model parameter values explicitly, since, for example,
the available capacity parameter in the fundamental diagram of traffic pro-
vides qualitative information about the presence of incident conditions.
We have therefore restricted attention to adaptive parameter identification
where the parameters are explicitly estimated. Even with this constraint,
a number of computational simplifications are feasible, and are proposed
further in Section 4.7 of this report.

3.3 MODEL FOR OBSERVATIONS PROCESS

The Kalman-filter design presumes to have available noisy, possibly non-
linear observations of the underlying variables of the mathematical model.
It is not possible within reasonable economic constraints to observe directly
the spatially defined microscopic traffic variables. The problem is thus
to approximate or infer the spatial quantities from appropriate processing of
the microscopic raw detector data obtained from discrete points along the
roadway. In this section, we show how vehicle counts, occupancies and indi-
vidual vehicle speed measurements may be pre-processed and treated as noisy
observations of the spatial mean speed and mean density. In so doing, we obtain
as a byproduct an approximation to the a priori covariances which charac-
terize the random processes in the model.

3.3.1 Background

In section 2.3, we defined microscopic variables associated with presence
type vehicle detectors. There are two types of presence detectors in current
use: magnetic detectors and inductive loop detectors [FR1]. The loop detector

FLOW

detection region

effective loop length

~8'

12' 7' 5' 7' 12'
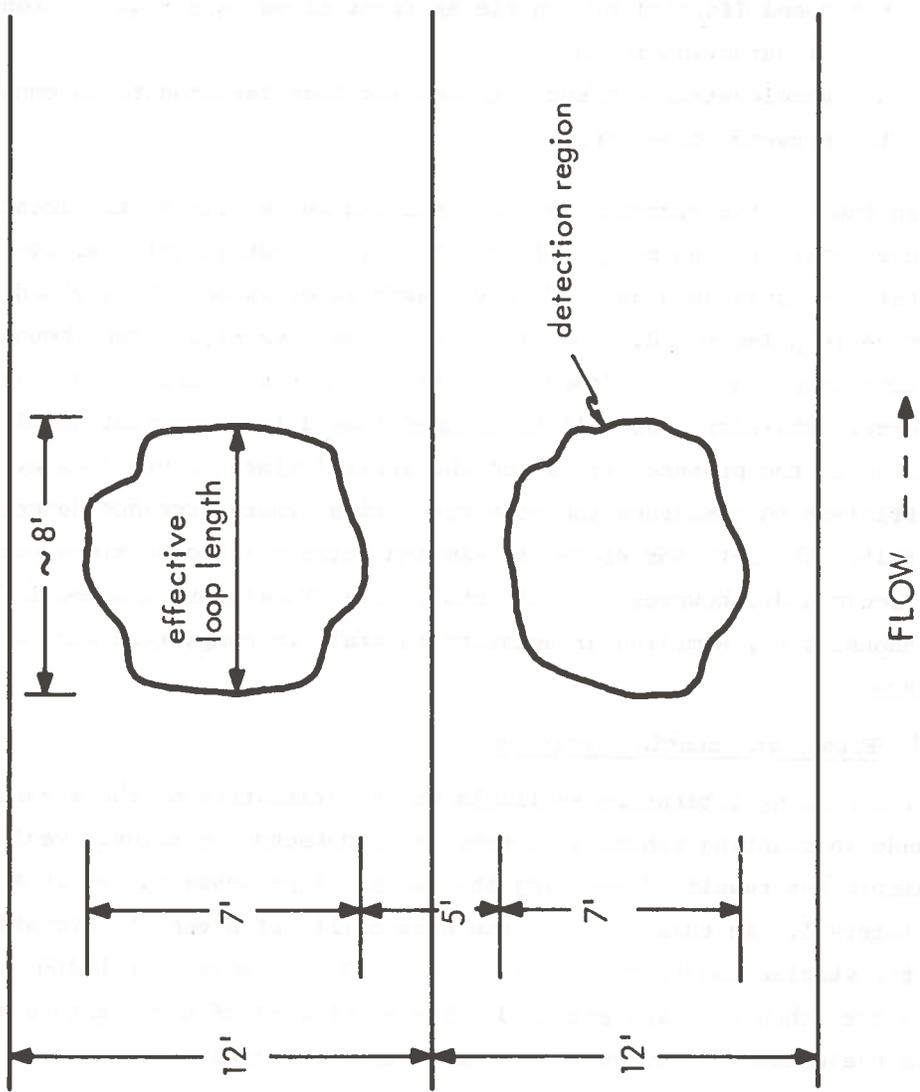
Figure 3.3 Detection Regions at Detector Station

Figure 2.3 shows that the only way for a vehicle to cross a detector station and not produce a presence pulse is for the vehicle to be less than five feet wide and traveling centered over a line separating lanes as it crosses the detector station. It is assumed here that this event never occurs, and, therefore, all vehicles get counted at least once.

Figure 2.3 also shows that it is possible for a vehicle, changing lanes near a detector station, to activate presence detectors in both lanes, and thus, to produce two presence pulses. Figure 3.4 shows a top view of vehicle of length ℓ [ft] and width w [ft] making a lane change. It is moving from center to center of adjacent 12-foot lanes. The lane-changing operation is assumed to take place at a constant speed v [ft/sec] and requires t seconds to complete. Thus, z feet of road are needed for the change where z = vt. Assuming that the detection regions of the loops in adjacent lanes are five feet apart, this vehicle will activate both detectors if, and only if, the detector station is located in the length X of road indicated in Figure 3.4. From the simple geometry of Figure 3.4, the following equation is obtained relating X to ℓ, w, v, and t: (see Kurkjian [12]).

$$X = \frac{vt}{12} \ (w-5) + \ell \ . \tag{3.45}$$

Suppose a vehicle 18 feet long and 6 feet wide makes a lane change at a constant speed of 88 ft/sec, and requires 4 seconds to complete the change. Using Eq. (3.45), this results in X = 47.3 feet. Assuming the lane change is equally likely to occur anywhere along the road, the probability of the lane change resulting in two presence pulses in

$$P = \frac{X}{2640 \ [\text{ft/detector station}]} = \frac{47.3}{2640} = 0.0179. \tag{3.46}$$

Thus, it is rather unlikely that any given lane change will cause an extra count, with 1/2 mile detector spacings. Based on this model for generating errors, the microscopic simulation described in Section 4 and Appendix B was exercised to find empirically error statistics for counting vehicles. The details of this procedure are described in 4.4.

3.3.4 Space-Time Homogeneity and Approximations to Mean-Speed and Density

Occupancy as defined in Section 2.3 is a measurement obtained from data taken over time at a fixed point. Density, on the other hand, is a spatial quantity associated with a fixed time. To relate fixed-time spatial quantities to fixed point temporal quantities in traffic, a model describing the relationship between various instantaneous point variables in traffic is needed. Similarly, point observations of vehicle speeds cannot simply be arithmetically averaged to provide observations of space mean speed.

The approach adopted here is to assume first a certain regularity of traffic flow, under which conditions an explicit and simple relationship between individual vehicle data and the spatial quantities can be shown to exist.

The traffic flow on a section $[x, x+\Delta x]$ over an interval $[t, t+\Delta T]$ is said to be space-time homogenous if the space mean speed and mean density on any subsection of $[x, x+\Delta x]$ at any time within $[t, t+\Delta T]$ is equal to the space-mean speed and density on any other subsection of $[x, x+\Delta x]$ at any other time within $[t, t+\Delta T]$. (For a more rigorous definition, see Breiman [22, 25].) Intuitively, the assumption of space-time homogenous traffic flow means that the traffic conditions do not change either in time or in space. Thus, from observations at a point, spatial quantities can be inferred. Restricting our attention to this condition, the following simplification of notation[*] is possible:

$$\bar{v}(x,\Delta x;t] = \bar{v}(x,\Delta x) ,$$

$$\bar{\rho}(x,\Delta x;t) = \bar{\rho}(x,\Delta x) . \tag{3.47}$$

Breiman [22] showed the following relation to exist between aggregate variables under space time homogenous conditions

$$\phi(x_0,t,\Delta T) = \bar{\rho}(x,\Delta x)\bar{v}(x,\Delta x),$$

$$x_0 \ \varepsilon \ [x, \ x+\Delta x]. \tag{3.48}$$

---

[*] See Section 2.3 of this report.

The omission of the effect of vehicle acceleration results in little loss in accuracy. Only extremely slow speeds (i.e., under 5 miles/hour), or extremely rapid acceleration causes the acceleration term to become significant.

In Eq. (3.52), the presence times, t_j, and the (average) effective loop length, d, are known quantities, but the vehicle lengths, ℓ_j, are unknown. To circumvent this problem, imagine that the ℓ_j are samples of a random variables, ℓ, with a __known__ probability density function, $f_\ell(\ell)$, and replace Eq. (3.52) with its expected value over ℓ. This results in

$$\bar{\rho}(x,\Delta x) = \frac{5280}{TL} E_\ell \frac{1}{\ell+d} \sum_{j=1}^{N(t,T)} t_j \quad \frac{veh}{mile} \quad \text{per lane} , \qquad (3.53)$$

where $E_\ell[\cdot]$ denote expectation over $f_\ell(\ell)$. Note that the 5280 converts the density value from vehicles/foot to vehicles/mile. Comparing Eq. (3.51) with the definition of occupancy, Eq. (2.26) (ignoring the end-effects $t_{i,I}$ and $t_{i,F}$) an approximate relationship between occupancy and density is seen to exist, given by

$$\bar{\rho}(x,\Delta x) = \frac{5280}{100} \ occ(t,\Delta T) E_\ell \left\{ \frac{1}{\ell+d} \right\} \frac{veh}{mile} \text{ per lane} . \qquad (3.54)$$

It is important to have an intuitive understanding of Eqs. (3.50) and (3.54). The density obtained using Eq. (3.54) is actually a time averaged density at a fixed space point, and not the desired spatial average density at a fixed time. __It is the space-time homogeneity assumption which allows time averages to be equated to spatial averages__.

The assumptions and approximations made in deriving Eq. (3.50) and (3.54) should be understood. They are restated and discussed here.

(a) The traffic is assumed to be space-time homogenous. Such an assumption is restrictive.

(b) The harmonic average, Eq. (3.50), is actually an approximation of an expected value (see Breiman [22]). The accuracy of such an approximation increases with $N(t,\Delta T)$. This implies that large time intervals, ΔT, are needed for a given level of accuracy when there are low-flow rates.

$$y_j^\rho(t_k) = \bar{\rho}_j(t_k) + \lambda_j^\rho(t_k), \tag{3.56}$$

for density observations; and

$$y_j^v(t_k) = \bar{v}_j(t_k) + \lambda_j^v(t_k), \tag{3.57}$$

for speed observations; where the noise statistics, $\{\lambda_k^1\}$, $\{\lambda_k^2\}$ are characterized by exercising the microscopic simulation (or eventually from actual data) and comparing true spatial mean variables with those computed using (3.28) and (3.32). In Section 4 of this report, we will show that the error processes $\underline{\lambda}_k^1$, $\underline{\lambda}_k^2$ can be characterized adequately by zero-mean, white processes with specified covariance.

If occupancy data are used both for speed and density observations, it is not hard to anticipate that errors in such observations will be correlated. The raw detector count process, $N(t, \Delta T)$, provides an alternative measurement of the state, which is non-linear:

$$y_j^\phi(t_k) = \frac{N(t, \Delta T)}{\Delta T} , \tag{3.58}$$

and modeled as

$$y_j^\phi(t_k) = \bar{\rho}_j(t_k)\bar{v}_j(t_k) + \lambda_j^\phi(t_k). \tag{3.59}$$

A priori statistics of the processes $\{\lambda_j^\phi(\cdot)\}$, $\{\lambda_j^v(\cdot)\}$ and $\{\lambda_j^\rho(\cdot)\}$ are specified either from the analysis in Section 4 of this report, or, in some applications, from available prior surveillance data. For convenience, the detector pre-processing formulas are summarized in Figure 3.5.

3.4 SUMMARY

In this section we have presented the continuous-discrete extended-Kalman filter algorithm for use in both state and parameter-estimation traffic-surveillance activities. Inputs to the algorithm are discrete-sampled outputs from the continuous-traffic network obtained at (as yet to be specified)

periodic intervals, ΔT. The filter is based on a <u>microscopic</u> model for
vehicular traffic flow in width spatial mean speed and mean density are the state
variables for each section. Outputs of the filter are estimates of these
spatial mean quantities together with estimates of parameters in the model.
The microscopic data available from roadway loop detectors are not directly
compatible with the microscopic traffic variables of interest. We showed
that by harmonically averaging individual vehicle speed measurements, and by
scaling occupancy measurements in a certain way, these quantities could be
approximated to be the spatial-mean variables in the presence of additive noise.

    a.   Dynamic propogation of disturbances between links;

    b.   Facility to study the impact of unmodeled exogenous (ramp-flow) demands;

    c.   Facility to determine sensitivity of surveillance algorithm performance to model parameter variations including a priori noise statistics, model parameter values, sensor data; and

    d.   Sufficiently small network that can be modeled and simulated microscopically (i.e., individual vehicles) to provide benchmark on

       1)   performance of macroscopic model-based filter with microscopic (detector-loop) data;

       2)   answers to sensitivity questions as in (b), (c) above; and

       3)   overall impact of unmodeled lateral dynamics (including lane changing and passing), mixes of driver and vehicle types, and incidents.

The microscopic simulation of the study network identified in paragraph (d) above is detailed in Section 4.3 of this report.

4.2.1 Model Structures

As shown in figure 4.1, the study network consists of three links, each with entry on-ramps and off-ramps. For conceptual purposes, the on-ramps are assumed to be located at the beginning or upstream end of a link section and the off ramps at the downstream end of a section. The ramp locations may typically define link boundaries, or the link boundaries can be defined relative to loop-sensor locations. Where freeways are already instrumented, there is some motivation for taking link boundaries to coincide with detector locations; this issue is discussed further in Section 2 of this report. For purposes of this study, each section has been taken to be 0.5 mile in length, and boundaries are defined by detector-loop locations (Figure 4.1 (b)).

The macroscopic model in Figure 4.1 is described in terms of equations (2.22) and (2.23), which result from the four step approximation procedure outlined in Section 2.2, i.e.*

$$\frac{d\bar{v}_j}{dt} = v_j \frac{\left[v_j - v_{j-1}\right]}{1/2(\Delta x_j + \Delta x_{j-1})} - \frac{1}{\tau}\left[_j - v_e{}^j(\rho_j)\right]$$

$$-\frac{\nu}{\tau}\frac{1}{\rho_j}\left[\frac{\rho_{j+1} - \bar{\rho}_j}{1/2(\Delta x_j + \Delta x_j + 1)}\right] + w_j{}^v , \qquad (4.1)$$

$$\frac{d\bar{\rho}_j}{dt} = \left[\rho_{j-1}\bar{v}_{j-1} - \rho_j v_j + \text{NETFLOW } j\right]\frac{1}{\Delta x_j} + w_j{}^\rho , \qquad (4.2)$$

where

$\bar{v}_j \equiv$ spatial mean speed on section 3,

$\bar{\rho}_j \equiv$ spatial mean density on section j,

NETFLOW j $\equiv$ net flow entering freeway section j ,

$v_e{}^j(\bar{\rho}_j) \equiv$ Equilibrium speed (fundamental diagram) for section j (mi/hr),

$\Delta x_j \equiv$ length of section j (mi) (typically .5 mi here),

$\nu \equiv$ driver-reaction sensitivity coefficient $((\text{mi})^2/\text{hr})$,

$\tau \equiv$ driver-reaction time coefficient (sec), and

$x_j =$ distance coordinate for end of link j.

---

*Time arguments are omitted for clarity of notation.

TABLE 4.1: FUNDAMENTAL DIAGRAMS FOR STUDY NETWORK

TYPE	FUNDAMENTAL DIAGRAM RELATION $(v_e(.))$	FREE PARAMETERS † Number	FREE PARAMETERS † Names	DENSITY $\rho*$ FOR MAX-FLOW	CAPACITY $\phi_{MAX} = \rho*(v_e(\rho*))$
LINEAR	$v_e = v_{free} - s\bar{\rho}$ $s = v_{free}/\rho_{jam}$	2	v_{free} (free speed) ρ_{jam} (jam concentration)	$\rho* = \rho_{jam}/2$	$\rho_{jam}\, v_{free}/4$
LOGARITHMIC	$v_e = \begin{cases} v_{free} & : 0 \le \bar{\rho} < \rho_{free} \\ 0 & : \bar{\rho} > \rho_{free} \\ k\,\ln\dfrac{\rho_{jam}}{\bar{\rho}} & : \text{otherwise} \end{cases}$ $k = v_{free}\left|\ln\left[\rho_{jam}\middle|\rho_{free}\right]\right.$	3	v_{free} ρ_{jam} ρ_{free} (Breakpoint)	$\rho* = \rho_{jam}/\varepsilon$ ††	$\rho_{jam}\,\dfrac{v_{free}}{\varepsilon(\ln \rho_{jam}/\rho_{free})}$
PARABOLIC	$v_e = v_{free}\left(1 - \dfrac{\bar{\rho}}{\rho_{jam}}\right)$ $\times \left((1 - \alpha\dfrac{\bar{\rho}}{\rho_{jam}})\right)$	3	v_{free} ρ_{jam} α : where $-1 < \alpha \le +1$	$\rho* = \rho_{jam}\left\{\dfrac{(1+\alpha)}{3\alpha} + \sqrt{\left(\dfrac{1+\alpha^2}{3\alpha}\right) - \dfrac{1}{3\alpha}}\right\}$ $\alpha \ne 0$ $= \rho_{jam}/2$: $\alpha = 0$	See Figure 2.10 for illustration

† See text discussion.

†† ε denotes base of natural logarithm.

TABLE 4.2   TYPICAL PARAMETER VALUES FOR MACROSCOPIC FREEWAY MODEL

(Logarithmic Fundamental Diagram)

τ	.00139 Hour (5-sec)
ν	15 miles^2/hour
ρ_{jam}	225 cars/miles/lane
ρ_{free}	23.1 cars/miles/lane
v_{free}	55 miles/hour
FLO (INPUT)	1667 cars/hour/lane
NETFLOW$_1$, NETFLOW$_2$, NETFLOW$_3$	0 cars/hour/lane (Nominal)
CAPACITY ϕ MAX.	2000 cars/hour/lane
NUMBER OF LANES	3
w_1^ρ, w_2^ρ, w_3^ρ	3.33 cars/mile/lane (Standard Deviation)
w_1^v, w_2^v, w_3^v	5 miles/hour (Standard Deviation)

Note that $\phi_{in}(t)$ is not strictly Poisson because of the "sampling" of n(t) over (t,t+ΔT), Figure 4.2. However, for simulation purposes, the correlation introduced by sampling n(t) when ΔT is small (1 second) is negligible. As usual, a uniformly distributed pseudo-random number generator will be used for simulation purposes; i.e., let θ_t be a sample of uniformly distributed random variables on (0,1), then n(t) is the smallest integer n such that

$$\sum_{i=0}^{n(t)} e^{-\lambda} \frac{\lambda^i}{i!} \geq \theta_t \tag{4.7}$$

where

$$\lambda \equiv (FLO.\Delta T) .$$

4.2.3 Summary

The equations of motion for the three-section dynamic model which encompasses the above assumptions can be summarized as follows:

$$\frac{d\bar{\rho}_1}{dt} = \left[\phi_{in} - \bar{\rho}_1 v_1 + NETFLOW_1 \right] \frac{1}{\Delta x_1} + w_1^{\rho} , \tag{4.8}$$

$$\frac{d\bar{v}_1}{dt} = \left[\frac{-1}{\tau} \bar{v}_1 - v_e^{-1}(\rho_1) \right] - \frac{\nu}{\tau} \frac{1}{\rho} \left[\frac{\rho_2 - \rho_1}{1/2(\Delta x_1 + \Delta x_2)} \right] + w_1^v , \tag{4.9}$$

$$\frac{d\bar{\rho}_2}{dt} = \left[\bar{\rho}_1 \bar{v}_1 - \bar{\rho}_2 \bar{v}_2 + NETFLOW_2 \right] \frac{1}{\Delta x_2} + w_2^{\rho} \tag{4.10}$$

$$\frac{d\bar{v}_2}{dt} = \frac{-v_2\left[\bar{v}_2 - \bar{v}_1\right]}{1/2\,(\Delta x_2 + \Delta x_1)} - \frac{1}{\tau}\left[v_2 - v_e^{\,2}(\bar{\rho}_2)\right] - \frac{\nu}{\tau}\frac{1}{\bar{\rho}_2}\left[\frac{\bar{\rho}_3 - \bar{\rho}_2}{1/2\,(\Delta x_2 + \Delta x_3)}\right] + w_2^{\,v} \quad , (4.11)$$

$$\frac{d\bar{\rho}_3}{dt} = \left[\bar{\rho}_2\bar{v}_2 - \bar{\rho}_3\bar{v}_3 + \text{NETFLOW}_3\right]\frac{1}{\Delta x_3} + w_3^{\,\rho} \quad , \tag{4.12}$$

$$\frac{d\bar{v}_3}{dt} = \frac{-\bar{v}_3(\bar{v}_3 - \bar{v}_2)}{1/2\,(\Delta x_3 + \Delta x_2)} - \frac{1}{\tau}\left[\bar{v}_3 - v_e^{\,3}(\bar{\rho}_3)\right] + w_3^{\,v} \quad . \tag{4.13}$$

Equations (4.8) to (4.13) together with explicit parameterizations of $v_e^{\,j}(\bar{\rho}_j)$, NETFLOW_j, and initial conditions define the macroscopic freeway model for the study network. Defining the six-vector

$$\underline{x}(t) \equiv (\bar{\rho}_1,\ \bar{v}_1,\ \bar{\rho}_2,\ \bar{v}_2,\ \bar{\rho}_3,\ J_3),$$

we have exhibited the non-linear dynamic model in the form required for filter design,

$$\underline{\dot{x}}(t) = \underline{f}\,(\underline{x}(t)) + w(t), \tag{4.14}$$

where $\underline{f}(.)$ is defined by the right-hand sides of (4.8) to (4.13). In modeling randomness in the dynamic model, we will for the most part consider only $\{w_i^{\,\rho},\ w_i^{\,v}\}$ to be random processes and the NETFLOW_j terms to be zero- or known time functions.*

---

*If NETFLOW_j is random or non-zero mean, it is easy to replace $w_j^{\,\rho}$ by
$$\tilde{w}_j^{\,\rho} = \tilde{\text{NETFLOW}}_j + w_j^{\,\rho}$$
in which $\tilde{\text{NETFLOW}}_j$ is the zero-mean random component of NETFLOW_j.

4.2.4 Free Parameters in Fundamental Diagrams for Study Network

The dynamics of the six-state freeway model (equations (4.8) to (4.13)) implicitly have free parameters in the equilibrium speed curve, $v_e^{\,j}(\bar{\rho}_j)$, which we consider to be in the form:

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t)\ \underline{p}(t)), \tag{4.15}$$

in which we let $p_j^{\,i}(t)$ denote the i-th free parameter on link j. In this study, there are at most two free-parameters in each fundamental diagram so that $\underline{p}(t)$ has a maximum dimension of six:

$$\underline{p}(t) \equiv (p_1^1,\ p_1^2,\ p_2^1,\ p_2^2,\ p_3^1,\ p_3^2)\ .$$

Table 4.3 summarizes the free parameters used in various forms of the fundamental diagram that are used in simulations.

4.3 MICROSCOPIC TRAFFIC SIMULATION

4.3.1 Motivation for Simulation Development

At the commencement of this research project, a need was established to have a viable traffic-data benchmark for evaluating modeling assumptions, and control and surveillance algorithms. A microscopic individual-vehicle simulation was desired. Although a data source for freeway traffic measurements was potentially available for surveillance algorithm evaluation [29], it did not have a benchmark on the spatial mean variables of interest (it is primarily raw occupancy and

(c) Multiple vehicle types classified by weight and acceleration capabilities;

(d) Multiple driver types characterized by aggressiveness in selecting available gaps for passing and merge maneuvers j;

(e) Arbitrary detector/sensor location, and monitoring during simulation run j;

(f) Ability to interface sensors and ramp controls with surveillance and control algorithm logic;

(g) Bookeeping for various simulation statistics, including

(1) arbitrarily defined subsection aggregate variables (spatial mean speed, spatial mean density, and input and output flow);

(2) cumulative individual travel times, aggregate delay, number of stops; and

(3) records of incidents and vehicle maneuver conflicts;

(h) Ability to simulate various traffic incident events including

(1) blocked lane from vehicle failures;

(2) slow-moving vehicles; and

(3) on-ramp demand surges.

The main program logic for exercising vehicle behavior is shown in Figure 4.4. A detailed description of all the modular blocks in Figure 4.4 is provided in Appendix A of this report.
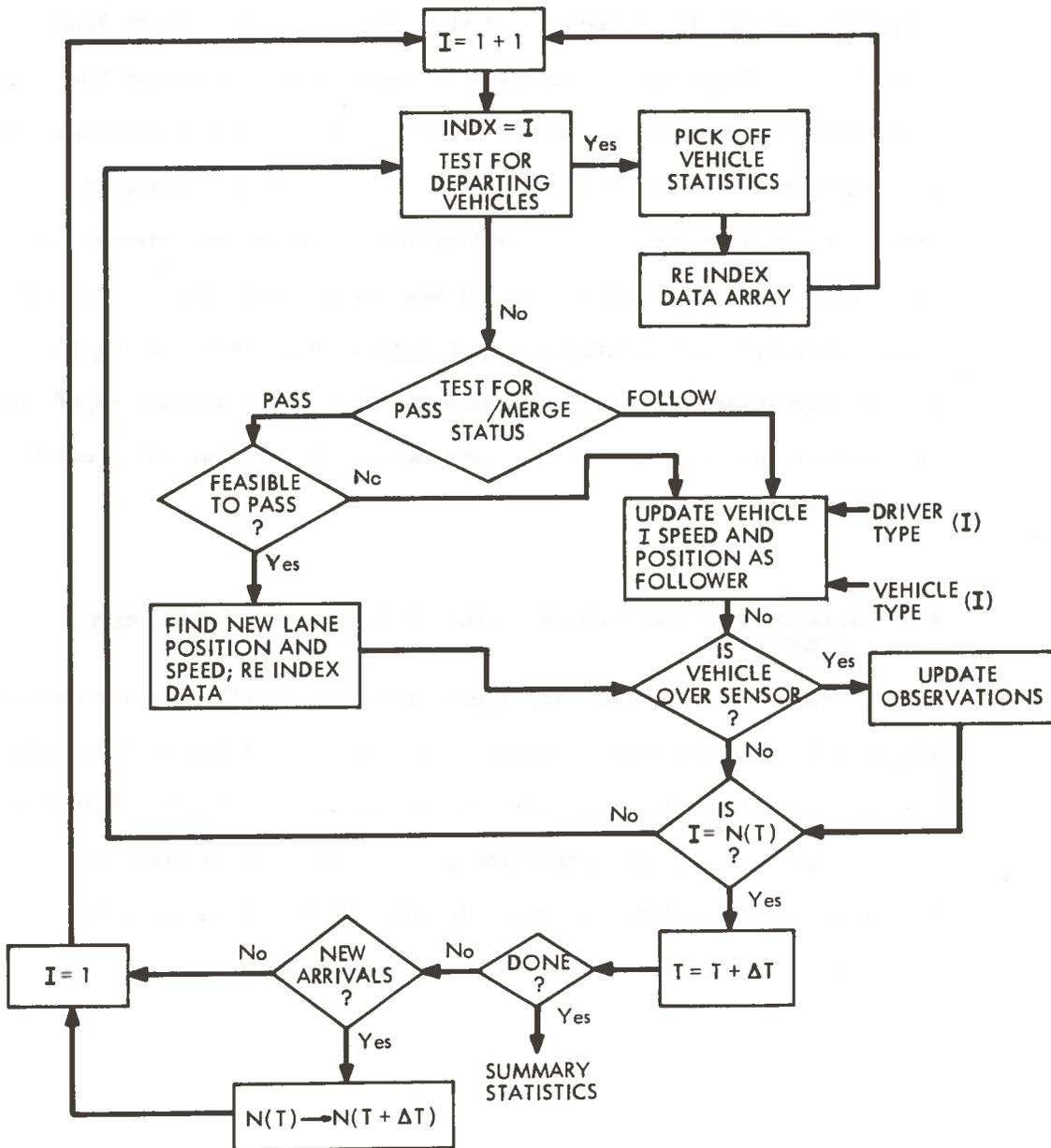
Figure 4.4 Microscopic Vehicle Simulation
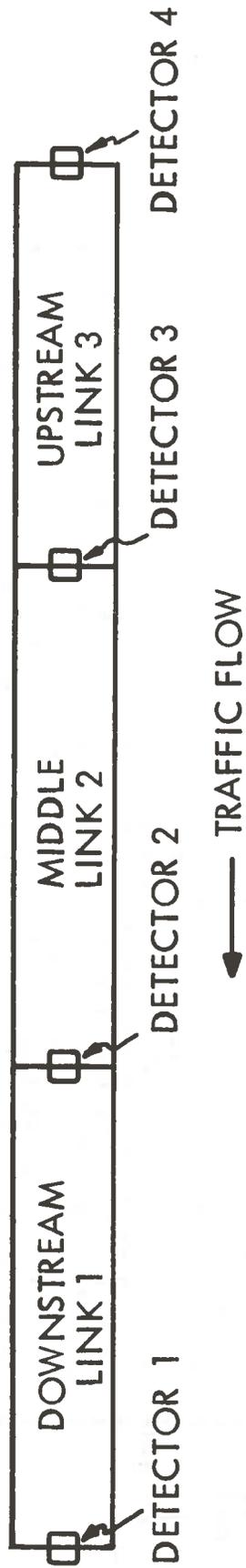
Figure 4.5: Detector Locations for Three-Link Microscopic Vehicle Simulation

X=0

—— VEHICLE MOTION

VELOCITY VS. TIME AND POSITION

X = 5280 (1.0 mi.)

STALLED VEHICLE LOCATION

LANE 1
POSITION
TIME
(sec)

```
      1 6 5    7 6 7 7 7    7 7 7 7      8 7    8 7    5    7 6    9    8 5    6 7
      17 7   6 7 7 7 7    6 7    8       7 5    7      6      6 6    7 7    5 7 7 7 5
      18 7 7 9 9 9     9    9 8    9 8    5 6    8 5    6      8 8    7 7    5 7 7 5
      1      9 9     9    5    5 6 6      6 5    6      6 6    6    6 6 6 7 7 6 6
      1    8     8    9 8 5012 3    8 5    6 6    7    6      5    6 6 6 6    7 7 6 6 7
      1 7   8     8    8 CC0A     5     66    9      5 75    5 5 6 6 6 6 7 7 7 7 7 7
      1 7   8 8    C000E 33     9    5 5 5 5 5    5 6 6 7 7 7 7 7 7 7 7 7 7
      1 8    8 7  CC CC    64      5 5 5 5 5 6 6 6 7 7 8 7 7 8 7 7 7
                OCH2 3  4    5 55 5 5 6 6 6 7 9 8 8 7 7 8 7 7 7
1C0+7   1      4   4C0CC 2333 445 5 5 6 6 6 6 9 9 8 8 8 7 9 7 7 7
      1 7 6 6 5 5 5   6  CCCOACH1 23    5     8      9     8 8    9    7 7 7 7
      1 7 6 6 6 7   8   :CCCCCCCD 3 5     8      9     8 5    7    8 7 7 7
      19 8 8 8   8 9  C0CC:C0C:0     8    8 8    8      7   77    3 9    x
      19      9 8   8  QCIFG112 3  4     9    8 8 7    87    8 9 8 8    9 8    7 7
      1      7   6 7  CCAREGI    5      7      87 8      9 9    8 8 8    9 7 7 7
      1 8 5 6   7  CCC0OA    3     7     9      9     9 9    8 3 8    8 7 7 6
      1 8   9   7  CCFI1123 4 5    6      7      9     E 9    H 6 6    6 7 7 6
      17   5 7 7  CCC0ACH23    3     8     3 8    8    B 3    3 6 6    6 7 7 6
2C0+ 6 7 7   8 7  0CC:C000 5 6    6     7 8      6      6 6 6 6 7    7 7 6
```

FIGURES ARE VELOCITIES IN FEET/SECOND DIVIDED BY 10
A=1 FT/SEC, B=2 FT/SEC, ..., I=9 FT/SEC. X=GT 95 FT/SEC

Figure 4.6 (b)   Sample Simulation Output: 1.0-Mile Freeway, 2/0.5-Mile Links, with Stalled
Vehicle at t=20 (Sec) on <u>Link 1 Curb Lane</u>

Figure 4.7(b): Link 1 Macroscopic Density Variables for Sample Run
                (Cont'd)

Figure 4.7(d): Link 2 Macroscopic Density Variables for Sample Run
(Concl'd)

This conjecture concerning space-time homogeneity has, in fact, been verified using the traffic simulation program. The effects of the other approximations used in deriving Eq. (3.54) (see Section 3.3.4) have also been determined.

The testing of Eq. (3.54) consisted of an examination of the error between the actual spatial density and the density predicted by Eq. (3.54); i.e., with

$$\bar{\rho}(x,\Delta x) = \frac{5280}{100} \; occ(t,\Delta T) E_\ell \left\{ \frac{1}{\ell+d} \right\} \frac{veh}{mile} \quad per \; lane \; . \qquad (3.54)$$

The test used:

    (1)   Values of ΔT ranging from 5 seconds to 1 minute.

    (2)   Values of Δx ranging from 100 feet to 1 mile.

    (3)   Traffic flow conditions ranging from low flow (~750 veh/hr per lane) to high flow (~1600 veh/hr per lane), and included homogeneous and inhomogeneous traffic.

    (4)   A value of $E_\ell \left\{ \frac{1}{\ell+d} \right\}$ equal to 0.034 feet. This was obtained from vehicle-type distribution information (see Appendix A, Table A.3) assuming d = 8 fee.

Before discussing the results, let us recall some notation. Let $y^\rho(k)$ denote the density at time step k obtained using Eq. (3.54) and $\bar{\rho}(k)$ denote the actual density at time step k obtained from the traffic simulation program. The error between the two densities, e(k), can be thought of as a random process $e(k) = y^\rho(k) - \bar{\rho}(k)$. This study is actually a study of the error process e(k), k = 1,2,... .

The results of the test were as follows:

Of course, the problem with increasing ΔT or Δx is that the guarantee of space-time homogeneity is lost. Traffic conditions can and do change drastically at a point over a 1-minute interval and along a 1/2 mile section at a fixed time. Because the density obtained from Eq. (3.54) is actually <u>a time-averaged density localized at a point</u>, this density can be highly dependent upon where in the section the station is located. Furthermore, the section density can change considerably over the time interval. Thus, the selection of a section and a time to which the density of Eq. (3.54) relates is not straightforward, and sometimes results in large unpredictable errors.

4.4.2 Summary

Main findings from this simulation study were:

(1) The temporal density variation on a section or road very local to the detector station can be obtained from occupancy measurements taken over short time intervals. Equation (3.54) is the conversion from occupancy to density, and is valid at all flow levels and inhomogeneous conditions. The noise associated with the conversion (i.e., the difference between the true density and that predicted by Eq. (3.54) is a zero mean white process with a large variance.

(2) Occupancy measurements taken over larger; e.g., one minute, intervals, do not, in general, convert to a section density using (3.54). This is due to the inhomogeneities (i.e., irregularities) that occur in traffic over one-minute intervals.

(3) Section densities on, say, a 1/2-mile section cannot be, in general, obtained from a static conversion of occupancy measurements

4.5 FILTER PERFORMANCE IN MACROSCOPIC SIMULATION

Before proceeding to evaluate the EKF in the microscopic simulation described in Section 4.3, it was decided to generate data with a simulation of the macroscopic Payne model (Figure 4.8). This procedure serves a number of purposes:

a) The filter design can be evaluated with an exact match to the model providing a lower bound on relative error performance,

b) Sensitivity to changes in model and noise parameters can be assessed, including impact of changes in sampling rates, and

c) Structural questions relating to the identifiability and distinguishability of model parameters can be addressed.

In essence, by examining the filter performance when the assumed model agrees exactly in structure with "plant" that generates the data, one can isolate problems which are encountered when actual roadway data is used.

4.5.1 Main Results

The main qualitative findings from this macroscopic filter evaluation were as follows:[†]

4.5.1.1 Known Model Parameters

a) When noise-corrupted observations of mean-speed ($\bar{v}_i$) and density ($\bar{\rho}_i$) were provided to the filter, the estimates converged rapidly to the true values;

b) Estimate error performance was least sensitive to error variance in mean-speed observations, and most sensitive to error variance in density observations,

c) With measurements of speed alone, the filter diverges or develops large biases in density estimates,

[†]All conclusions are based on a simulation of the three-link candidate study network in Section 4.2.

d) With measurements of density alone, the estimates of both mean speed and density converge to the true values at only a slightly slower rate than (a);

e) From a noisy observation of density alone, the filter can provide estimates of densities on neighboring links (without additional observations) but with slower (by a factor of about 3) rate of convergence.

These qualitative conclusions apply over a wide range of both a priori process noise variance and observation noise variance (see experiments below), and

f) With observations of flow ($\bar{\rho}_i \bar{v}_i$) and mean speed ($\bar{v}_i$) on each link instead of density, estimate error convergence was generally achieved, but at a slower rate than any case where density was available.

Using flow as an observation was proposed initially as an alternative to deal with practical difficulty of measuring density with spatially separated point sensors.

4.5.1.2 Unknown Fundamental Diagram Parameters

a) One or two free parameters in each of the candidate fundamental diagrams in Table 3.1 (linear, logarithmic, and parabolic) could be identified per link with the augmented extended Kalman filter (AEKF),

b) With observation scenarios where the estimate converged with known parameters (i.e., except where flow or speed alone are measured), state and parameter estimates converged with rates of convergence that paralleled the cases described in 4.5.1.1 above,

c) Sensitivity of the state and parameter estimates to a priori noise variance paralleled the known parameter case, and

d) With two or more free parameters in the fundamental diagram, convergence was unpredictable. This is related to a structural identifiability and distinguishability problem discussed below.

4.5.2 Sample Run with Known Parameters

A sample simulation run to illustrate the observability of density on

(a) Network
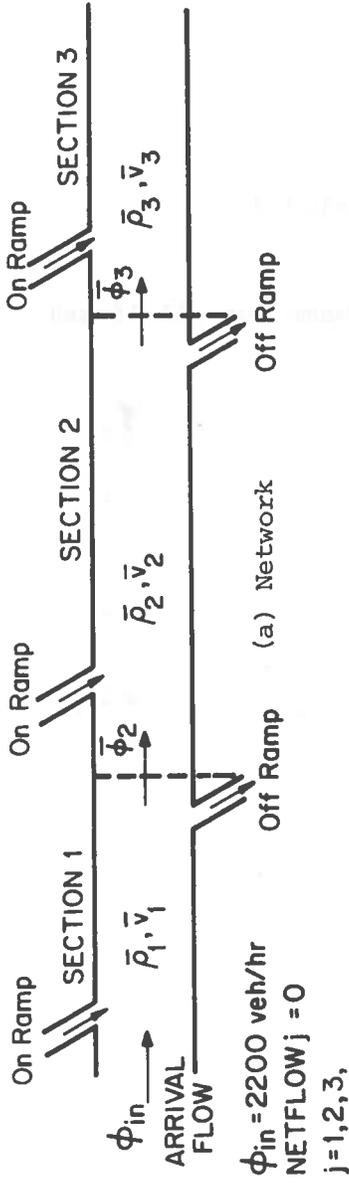
Filter Parameters	Initial time t=0 sec		Final time t=100 sec		Sample time Δt=4 sec.	
	Density $\bar{\rho}$ (veh./mile)	Speed $\bar{v}$ (veh./hr.)	Density $\bar{\rho}_2$ (veh./mile)	Speed v_2 (miles/hr.)	Density 3 (cars/mile)	Speed $\bar{v}_3$ (miles/hr.)
Initial State $\underline{x}(0)$	90	55	40	55	40	55
Initial Estimate $\hat{x}(0)$	110	40	90	30	20	70
Final State $\underline{x}(100)$	43.3	53.45	46.69	53.76	49.66	54.84
Final Estimate $\underline{x}_{100}$	43.5	53.48	46.83	53.93	49.75	54.88
Initial Est. Standard Deviation	4	5	4	5	4	5
Final Est. Standard Deviation	1.4	.55	.97	.51	.5	.18
Meas. Noise (Λ)	80	80	80	80	2	80
Process Noise (w)	.7	.87	.7	.87	.7	.87

(b) Summary Results and Input Data

Figure 4.9: Sample Macroscopic Simulation Run.

$$\underline{y}_k(p) = G(\underline{x}_k, \underline{p}) , \qquad (4.17)$$

then two parameter values, say $\underline{p}$ and $\underline{\alpha}$ (with $\underline{p} \neq \underline{\alpha}$), are said to be <u>indistinguishable</u> if

$$G(\underline{x}_k, \underline{p}) = G(\underline{x}_k, \alpha) , \qquad (4.18)$$

for all k=1,2,... and <u>admissible</u> <u>test</u> <u>sequences</u> $\{\underline{u}_k\}$. Otherwise, $\underline{p}$ and $\underline{\alpha}$ are said to be <u>distinguishable</u>; if for a given parameter value, $\underline{p}$, $\underline{p}$, and $\underline{\alpha}$ are distinguishable for all $\underline{\alpha}$ in a small region (ϵ-ball) near $\underline{p}$, then the system is called <u>locally identifiable</u> [28].

Recently Grewal [28] showed that the ν and τ parameters and one parameter (linear slope) of the fundamental diagram are locally identifiable in this sense in the Payne model. However, the concept of identifiability is a completely <u>deterministic</u> <u>structural</u> property of the particular model parameter-ization used. That a parameter in a system is locally identifiable, however, does <u>not</u> imply anything about the convergence of an estimator of that parameter, even locally.

Essentially, one needs more than identifiability to assure that the parameters can be estimated. A central problem is choosing test signal sequences $\{\underline{u}_k\}$, that are "sufficiently exciting" in the sense that the parameter modulates the output in such a way that its effect is observable. Unfortunately, for non-linear systems, there exists no easy way to characterize sufficiently exciting inputs which render the unknown parameters observable (we refer to [26] for a detailed discussion of these and related issues in system identification).

In traffic flow, the random effects of driver behavior and stochastic demand fluctuation are likely to suffice for the purpose of an identifying test sequence. Based on our preliminary simulation results, some caveats regarding observability of parameters with an extended-Kalman filter can be noted. With two or more free parameters, the mean speed and density need to be changing in time to obtain usable parameter estimates. While no precise
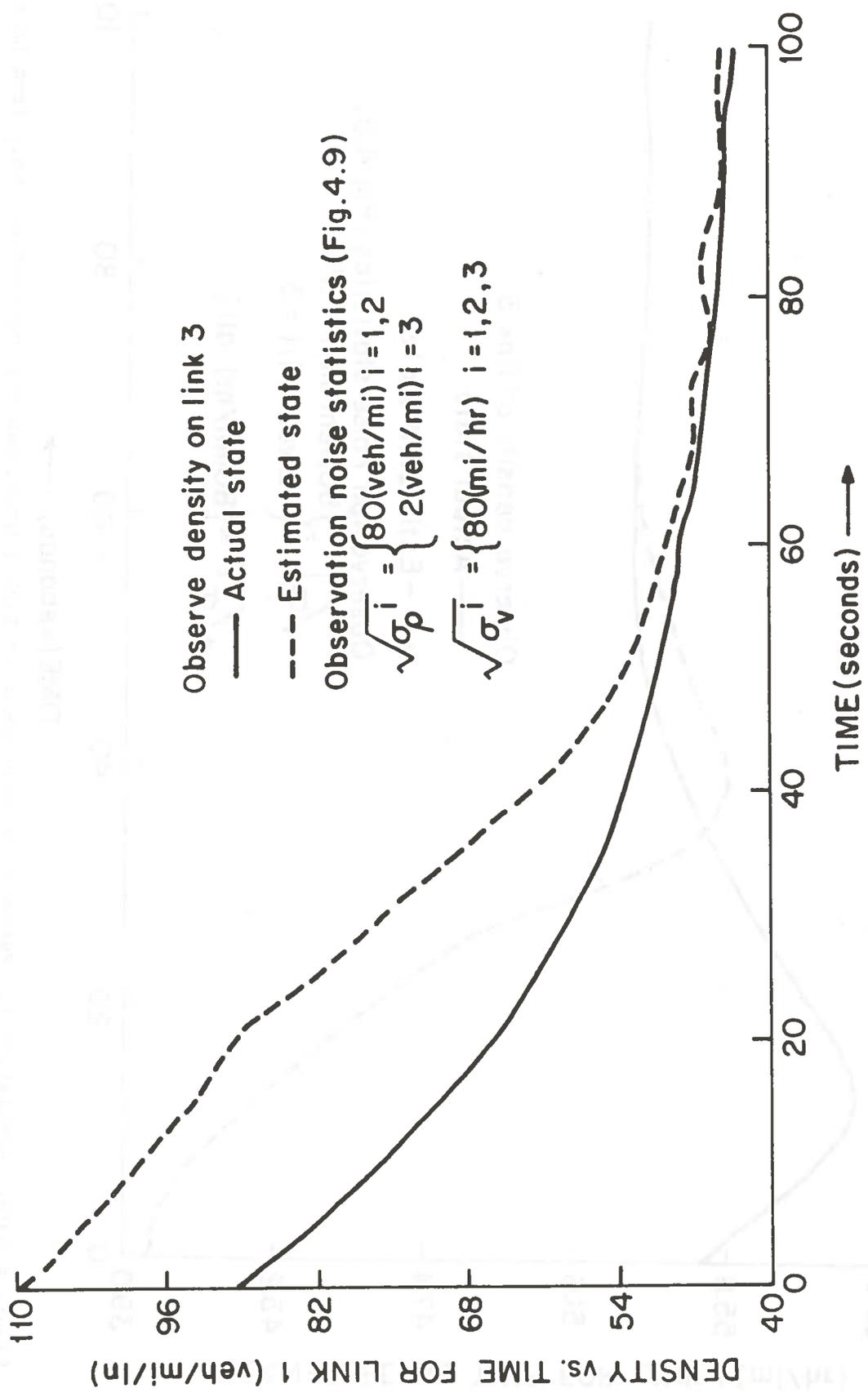
Figure 4.10(a): Simulation 1: Estimate of Density on Link 1 Using Density Observations Only from Link 3

The figure shows a plot of DENSITY vs. TIME FOR LINK 1 (veh/mi/ln) on the vertical axis against TIME (seconds) on the horizontal axis.

Legend:

Observe density on link 3
—— Actual state
---- Estimated state
Observation noise statistics (Fig.4.9)

$$\sqrt{\sigma_p^i} = \begin{cases} 80(\text{veh/mi}) \; i=1,2 \\ 2(\text{veh/mi}) \; i=3 \end{cases}$$

$$\sqrt{\sigma_v^i} = \{80(\text{mi/hr}) \; i=1,2,3$$

Figure 4.11: Simulation l: Estimate of Free-Speed Parameter on Link 2

Simulation 2: Low-Noise Estimation of Slope Parameter (s) in Linear
Fundamental Diagram (Synopsis: Table 4.6)

This run started with identical conditions as in Simulation 1 except
the v_f^i was <u>given</u> and the slope parameter, s^i, estimated on each link. Thus,
the assumed fundamental diagram was:

$$v_e^i(\bar{\rho}^i) = 55 - \hat{s}_i \, \bar{\rho}^i, \quad i=1,2,3. \tag{4.21}$$

Although it required 15 measurements (60 sec) worth of data to converge within
the steady-state error variance, it did so with <u>small</u> true changes in mean speed
and density from initial conditions. Figure 4.12(a) shows the estimated slope
parameter ($\hat{s}_2$) on link 2 over the simulation run. Figures 4.12(b) and 4.12(c)
illustrate the corresponding space mean speed and density on link 2. Note that
the errors in speed and density estimates for Simulation 2 had no statistically
significant difference from simulation 1, despite the fact that in Simulation 2
the flow variable was made pseudo-random as described in Section 4.1.

As a preliminary test of the AEKF's ability to truck parameters which change
with time, Simulation 2 was repeated, except that a step change in the the slope
parameter on link 2 was introduced at t=35 seconds (Figure 4.12(d)). Note
that the change is detected within about 5 measurements (20 sec). This
"detection" result is an optimistic or "best case" since a <u>step change's</u>
effect is easier to observe than when (as in reality) the parameter changes
less discontinuously. We will later show, using microsimulation, that less
abrupt changes resulting in microscopic traffic flow can also be detected, but
the corresponding estimates require longer convergence time.

Simulation 3: Estimation of α Parameter in Parabolic Fundamental Diagram
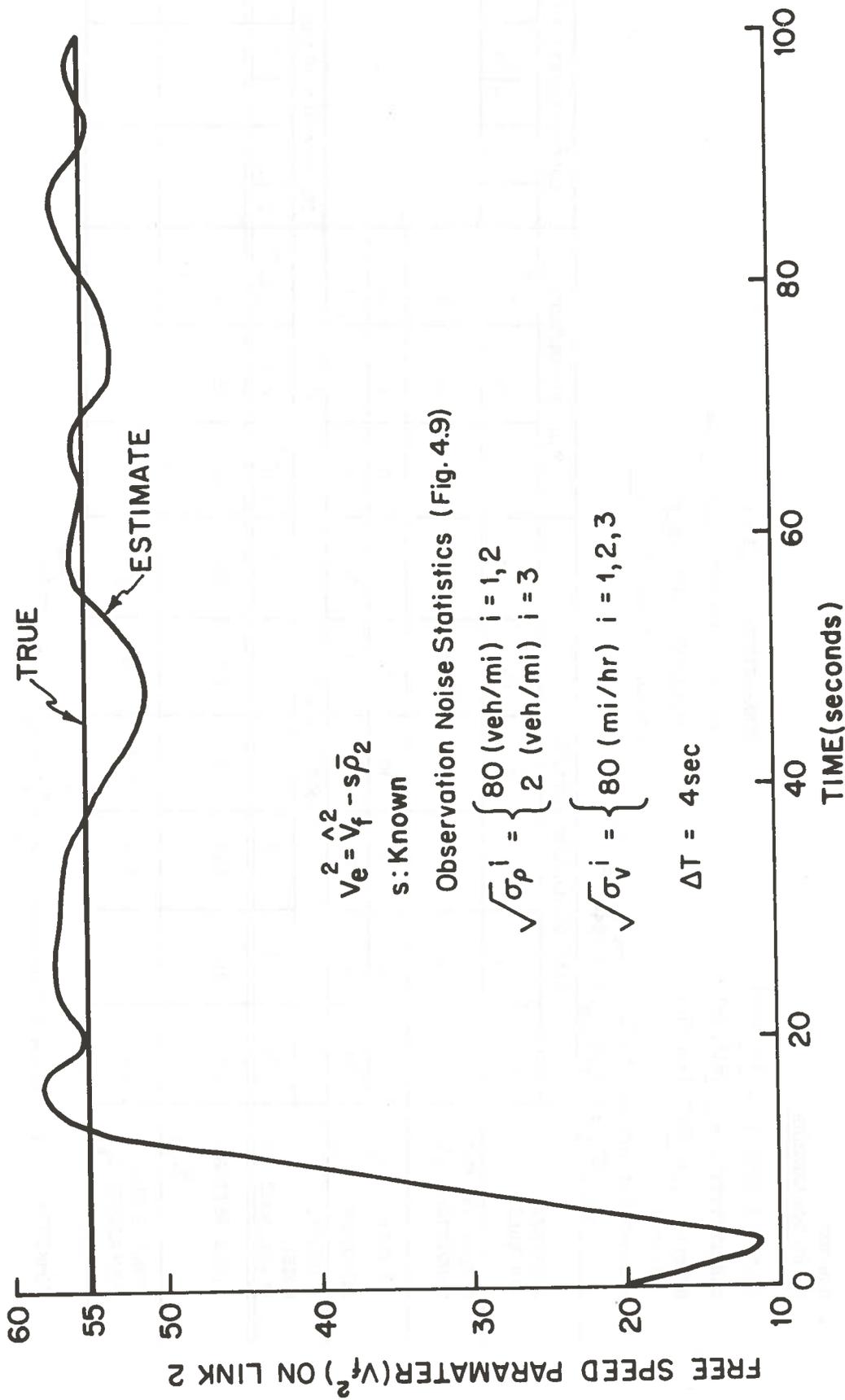(Synopsis: Table 4.7)

The objective of this simulation was to show that with alternative
parameterizations of the fundamental diagram, unknown parameters could be
estimated. In this case, we set the <u>true</u> fundamental diagram on each link to
be <u>linear</u> with
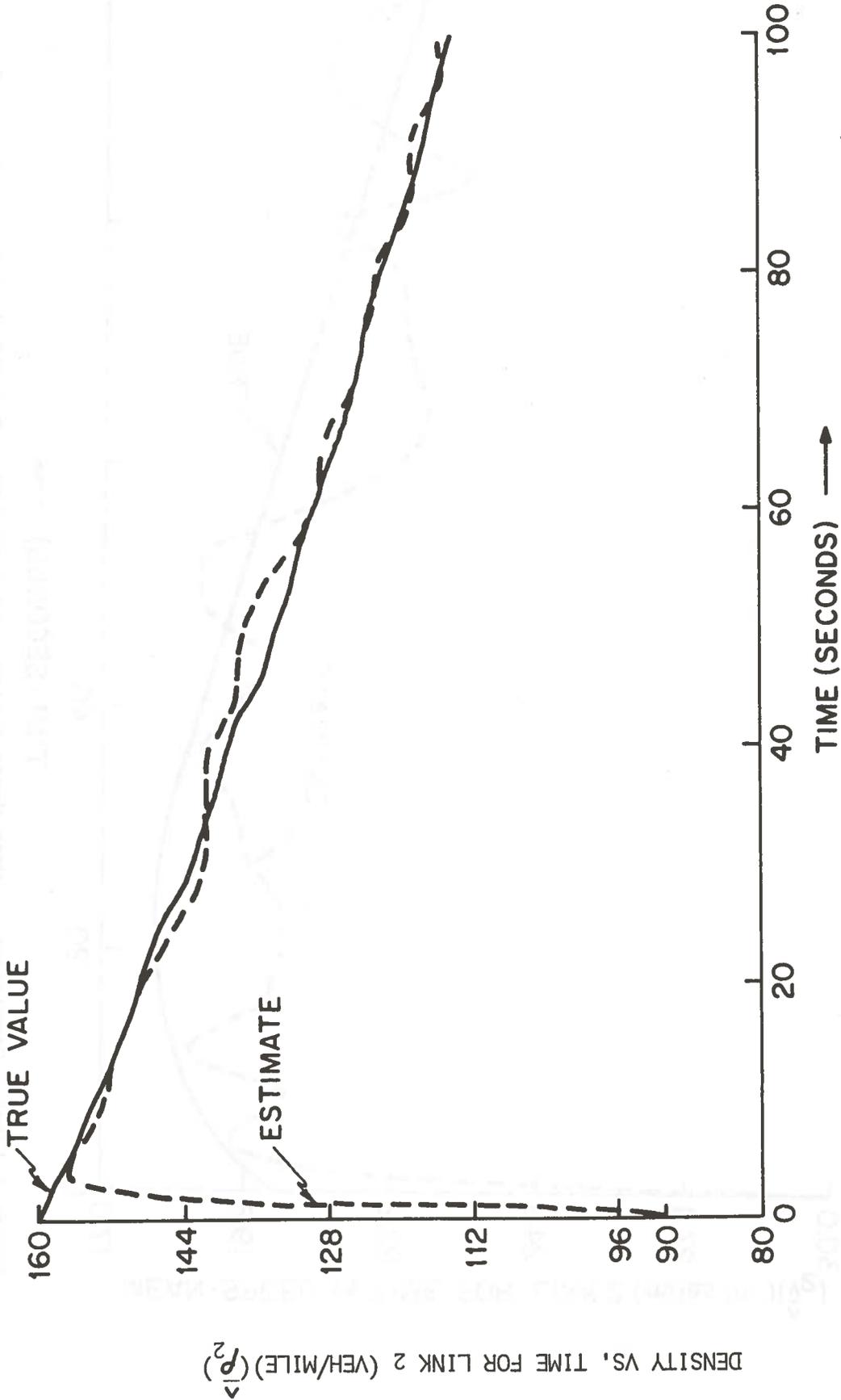
Figure 4.12(b): Simulation 2: Density Estimate on Link 2 with Estimated Slope Parameter $(\hat{S}_2)$
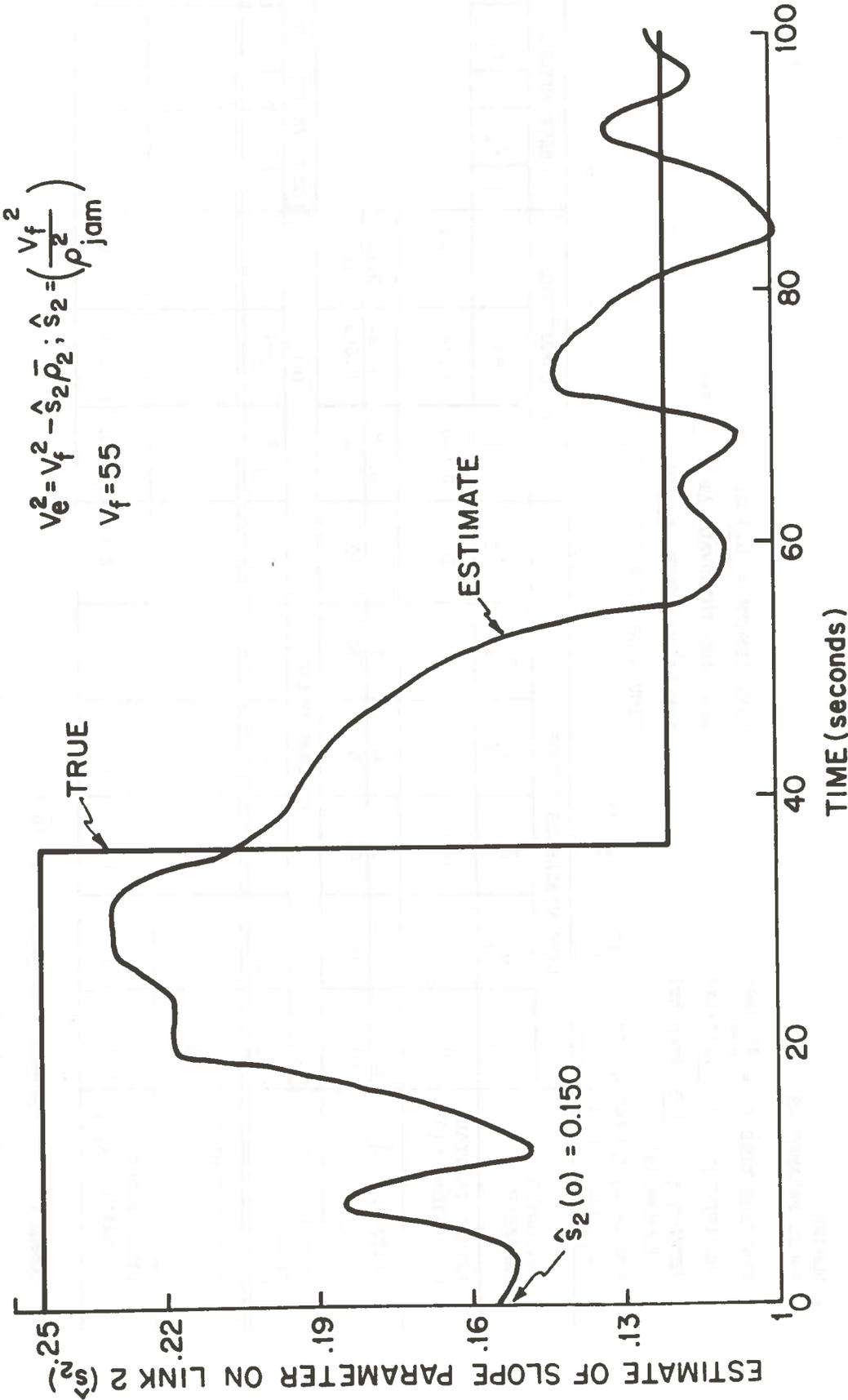
Figure 4.12(d): Simulation 2: Effect of Time-varying Slope Parameter on Its Estimate

$$v_e^i(\bar{\rho}_i) = v_f^i \left(1 - \frac{\bar{\rho}_i}{\rho^i_{jam}}\right), \tag{4.22}$$

with $v_f^i = 55$ and $\rho^i_{jam} = 225$ for $i=1,2,3$. In the filter, however, the fundamental diagram was assumed to be

$$v_e^i(\bar{\rho}_i) = v_f^i \left(1 - \frac{\bar{\rho}_i}{\rho^i_{jam}}\right)(1-\hat{\alpha}^i \frac{\bar{\rho}_i}{\rho^i_{jam}}), \tag{4.23}$$

with $\bar{\rho}_{jam}^i = 225$ and $v_f^i = 55$ (the true known values above), and $\hat{\alpha}_i$ estimated on-line, for $i=1,2,3$. As a result of this mismatch in model structure, one would hope that the filter would estimate $\hat{\alpha}_i$ to be zero ($\hat{\alpha}_i=0$ means the parabola degenerates into a straight line). This is in fact the resulting performance as illustrated in Figure 4.13(a), and Table 4.7. Comparing this run to Simulation 2, a slightly worse parameter r.m.s. error performance can be noted, due to the a priori structural mismatch. The state estimates are illustrated for link 2 in Figure 4.13(b) and (c).

Simulations 4 and 5: Multiple Parameter Estimation with Increased Sampling Intervals and Alternate Measurement Strategies

Two sets of macrosimulation runs were made to determine the effects of increased sampling intervals on estimator performance. In addition, two parameters were allowed to be free in the fundamental diagram. The form of the fundamental diagram was:

$$v_e^i(\bar{\rho}_i) = \hat{v}_f^i(1-\hat{s}_i \bar{\rho}_i), \tag{4.24}$$

for each $i=1,2,3$. For each link the true parameter values were:

---

*Close comparison of Figures 4.12(b), (c) with 4.13(b), (c) shows a very strong correlation in the "random" error. The reason is that the pseudo-random number generator for flows and noises are initialized with the same kernel in both runs, with the result that successive runs appear correlated.

$$v_f^i = 55 \ (\text{mi/hr}) \ , \ \Bigg\}$$

$$s_i = 0.244 \ \text{mi}^2/\text{veh-hr} \ . \ \Bigg\}$$

(4.25)

The unknown parameters and states on each link were measured in two ways. In Simulation 4 (Table 4.8), noisy observations of the true mean-speed and density $(\bar{\rho}_i, \bar{v}_i)$ on each link were available; i.e., from Table 3.3

$$y_\rho^i = \bar{\rho}_i + \lambda^\rho_i \ , \ \Bigg\}$$

$$y_v^i = \bar{v}_i + \lambda^v_i \ , \ \Bigg\}$$

(4.26)

for i=1,2,3, where λ^ρ_i and λ^v_i are the (assumed) white noise terms corrupting the observations. By contrast, in Simulation 5, the observation of true density was not available, but only link flow and spatial mean speed, or, as in Table 3.3,

$$y_v^i = \bar{v}_i + \lambda^v_i \ , \ \Bigg\}$$

$$y_\phi^i = \bar{\rho}_i \ \bar{v}_i + \lambda^\phi_i \ , \ \Bigg\}$$

(4.27)

for i=1,2,3, where λ_i^v is the same as above, and λ_i^ϕ is the white noise term corrupting the flow observation. The purpose of introducing flow-speed observations in Simulation 5 was simply to test the observability of parameters and spatial mean variables with non-linear observations. Intuitively, this was motivated by the practical consideration that flow is a directly observable quantity with detector loops at spatially discrete points.

Results are summarized in Tables 4.8 and 4.9 and Figures 4.14(a) to (d). Note that over the longer (25-minute) simulation run, mean speed and density undergo larger relative changes than for the short runs (100 sec) in Simulations 1 to 3. This "motion" greatly enhances the observability of unknown parameters as demonstrated by Figure 4.14(a). Although the elapsed time for estimating free speed and slope parameter is longer ($\approx$5 min. with full state observations),
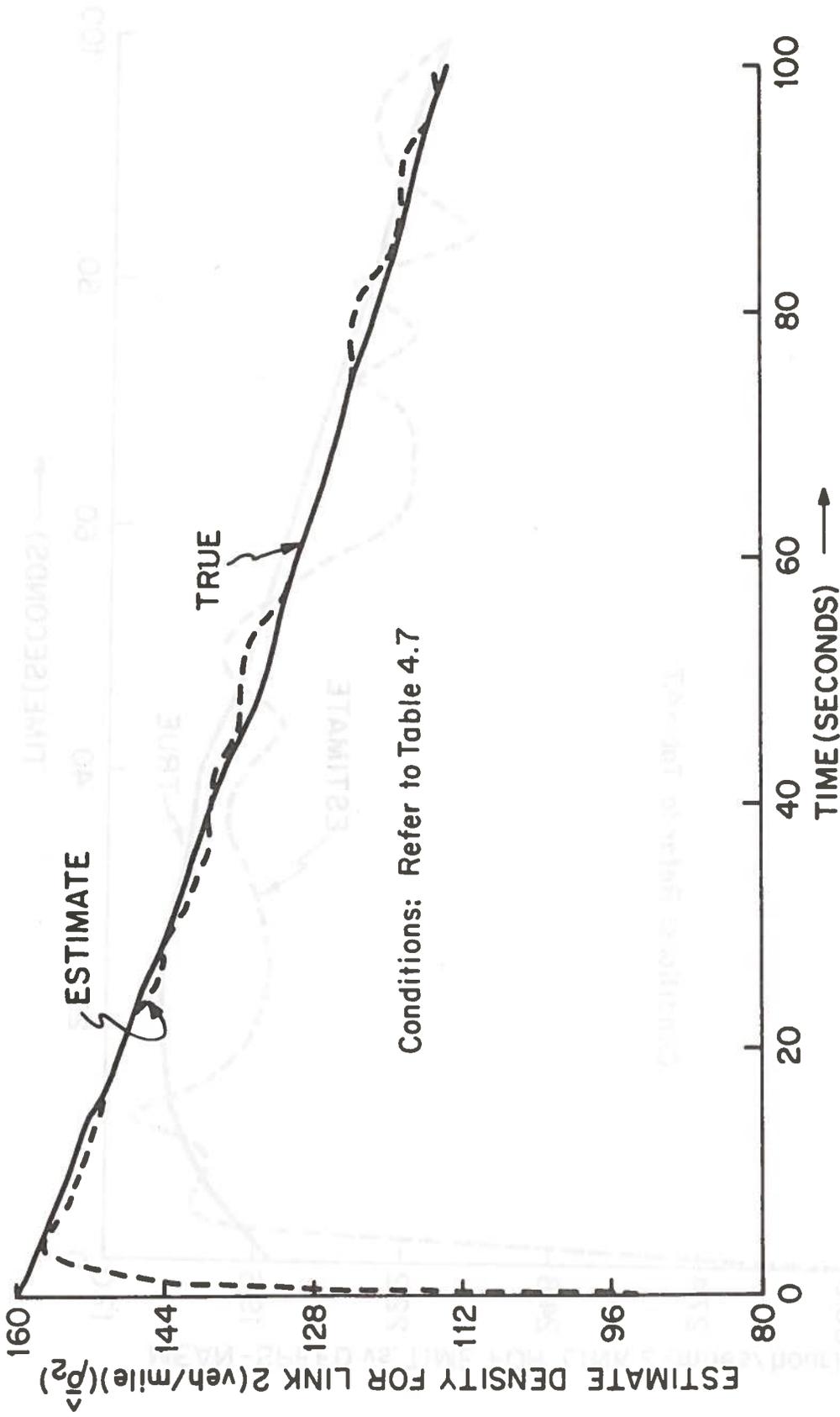
Figure 4.13 (b): Simulation 3: Density Estimate on Link 2 with Estimate α Parameter $(\hat{\alpha}_2)$:

TABLE 4.8   SIMULATION 4 SYNOPSIS: EFFECT OF INCREASED SAMPLE INTERVAL ON TWO-PARAMETER ESTIMATION

NOMINAL
MODEL PARAMETERS

REACTION TIME τ = 15 (sec)   LINK-LENGTH = 0.5 mi

SENSITIVITY ν = 5 (mi^2/hr.)   SAMPLING INTERVAL ΔT = 1 min

NETFLOW ϕ_{in} = 2200 (ven./hr.)   SIMULATION TYPE Macro
(incoming)   SIMULATION RUN TIME 25 min

FUNDAMENTAL DIAGRAM TYPE = Linear, v_f=55, ρ_{jam}=225

ESTIMATED VARIABLE	LINK VARIABLES (STATES)						F.D. PARAMETERS						OBSERVATIONS					
	$\bar{\rho}_1$	$\bar{v}_1$	$\bar{\rho}_2$	$\bar{v}_2$	$\bar{\rho}_3$	$\bar{v}_3$	s_1	v'_f	s_2	v^2_F	s_3	v^3_f	$\bar{\rho}_1$	$\bar{v}_1$	$\bar{\rho}_2$	$\bar{v}_2$	$\bar{\rho}_3$	$\bar{v}_3$
ACTUAL INITIAL CONDITIONS $\underline{x}(0)$	200	10	160	20	130	10	0.244	55	0.244	55	0.244	55						
INITIAL ESTIMATE $\hat{x}_\rho$	120	35	90	30	70	0	0.25	40	0.35	0.35	0.15	20						
Σ_o	90	70	90	70	90	70	0.09	300	0.09	0300	0.09	300						
APRIORI ERROR COVARIANCE Λ (W)	0.5	0.75	0.5	0.75	0.5	0.75	10^{-6}	0.5	10^{-6}	0.5	10^{-6}	0.5	2	2	2	2	2	2
FINAL ESTIMATE $\hat{\underline{x}}_{k\mid k}$	52	42	52	43	51	43	0.22	53	0.25	56	0.23	55						
FINAL ERROR COVARIANCE $\Sigma_{k\mid k}$	1.2	1.3	1.2	1.3	1.2	1.3	10^{-4} 2.8		10^{-3} 4.5		10^{-3} 5.1							

PROCESS NOISE   OBSERVATION NOISE (Λ)

COMMENTS
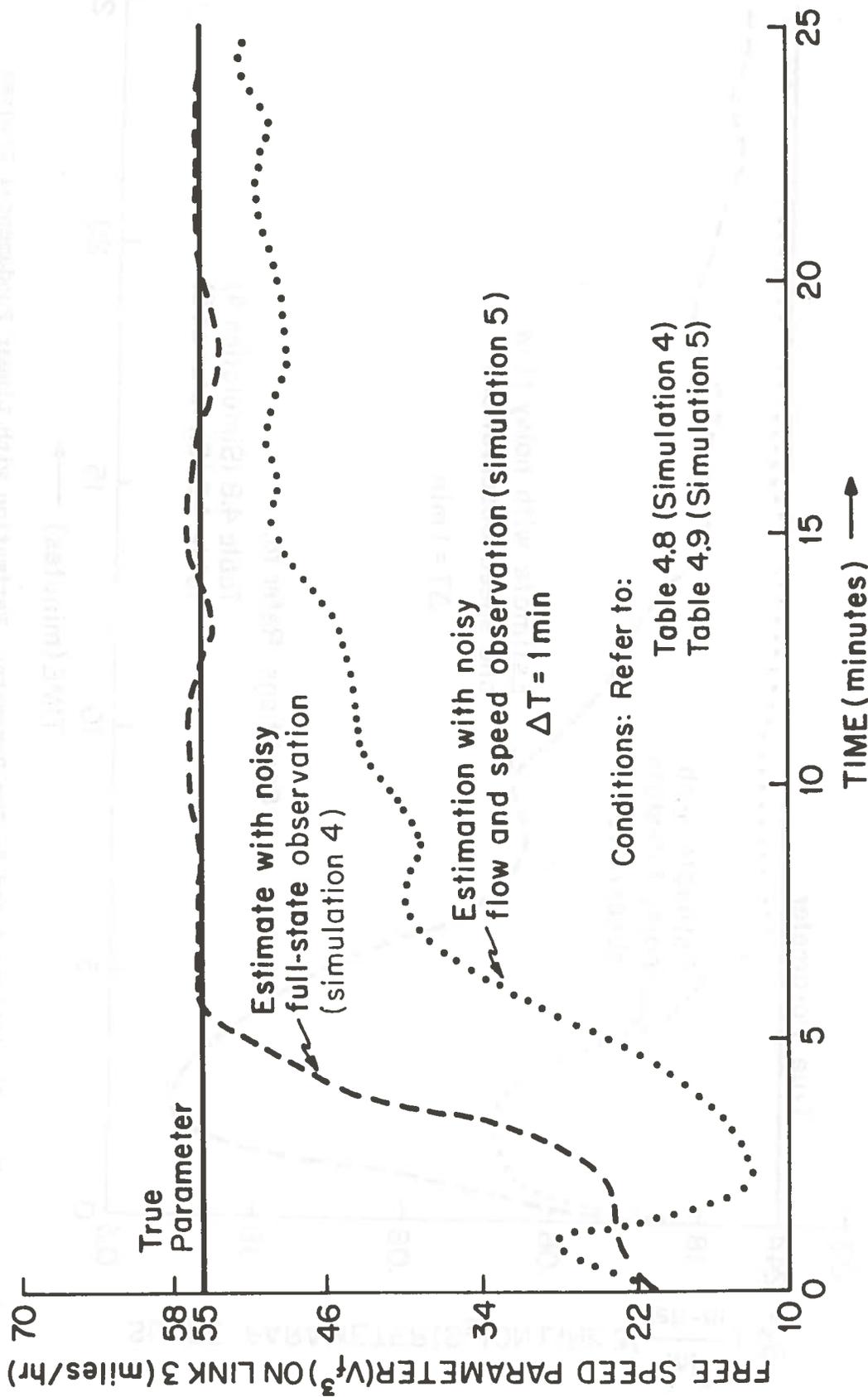
Figure 4.14(a): Simulations 4 and 5: Two-Parameter Estimation with Linear Fundamental Diagram
Free Speed Parameter Estimate on Link 3.
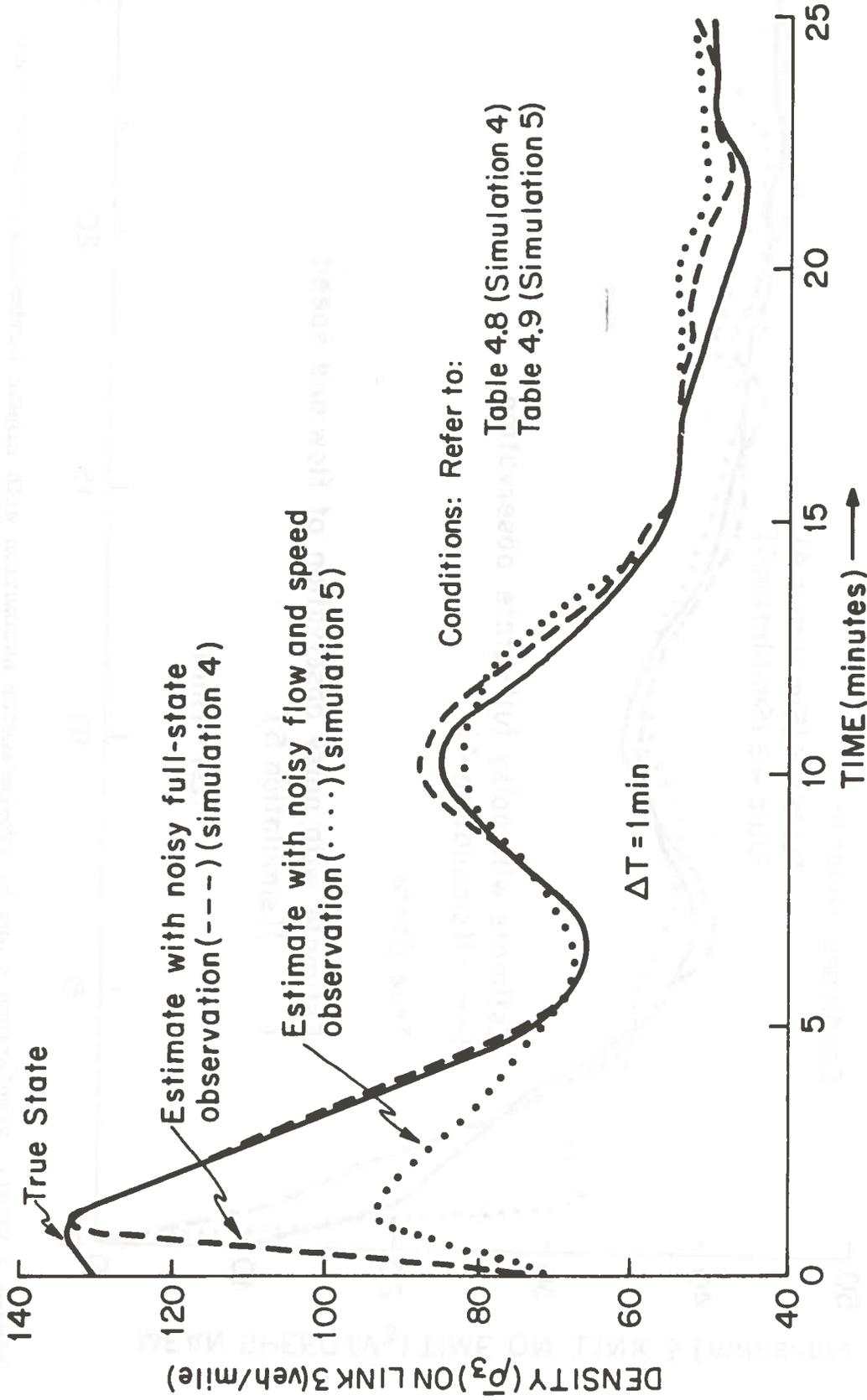
Figure 4.14(c): Simulations 4 and 5: Two-Parameter Estimation with Linear Fundamental Diagram Density Estimate on Link 3

TABLE 4.9   SIMULATION 5 SYNOPSIS: EFFECT OF FLOW/SPEED OBSERVATION CONSTRAINT ON TWO-PARAMETER ESTIMATION

* NOMINAL MODEL PARAMETERS

REACTION TIME τ = 65

SENSITIVITY ν = 5

NETFLOW ϕ_{in} = 2200 (incoming)

FUNDAMENTAL DIAGRAM TYPE = Linear, v_f=55, ρ_{jam}=225, s = 0.244

LINK-LENGTH = 0.5 mi.

SAMPLING INTERVAL ΔT = 1 min.

SIMULATION TYPE   MACRO

SIMULATION RUN TIME   25 min.

ESTIMATED VARIABLE	LINK VARIABLES (STATES)						F.D. PARAMETERS						OBSERVATIONS					
	$\bar\rho_1$	$\bar v_1$	$\bar\rho_2$	$\bar v_2$	$\bar\rho_3$	$\bar v_3$	$-s_1$	v_f^1	$-s_2$	v_F^2	$-s_3$	v_f^3	$\bar\phi_1,\ \bar v_1$		$\bar\phi_2,\ \bar v_2$		$\bar\phi_3,\ \bar v_3$	
ACTUAL INITIAL CONDITIONS $\underline{x}(0)$	200	10	160	20	130	10	.244	55	.244	55	.244	55						
INITIAL ESTIMATE $\hat{\underline{x}}_0$	120	35	90	30	70	0	0.25	40	0.35	35	0.15	20						
Σ_0	90	70	90	70	90	70	0.09	300	0.09	300	0.09	300						
* APRIORI W ERROR COVARIANCE Λ	0.5	0.75	0.5	0.75	0.5	0.75	10^{-6}	0.5	10^{-6}	0.5	10^{-6}	0.5	100	3	100	3	100	3
FINAL ESTIMATE $\hat{\underline{x}}_{k|k}$	53	42	51	44	52	43	0.28	57	0.36	62	0.20	52						
FINAL ERROR COVARIANCE $\Sigma_{k|k}$	3.3	2.4	3.4	2.6	3.2	2.3	1.1 E-2	5.1 E-2	2.4 E-3	8.2 E-3	7.6 E-3	2.5						

(PROCESS NOISE (W) ... OBSERVATION NOISE (Λ))

COMMENTS :

1) $\bar\phi_i \equiv \bar\rho_i\ \bar v_i$ (veh/hr)

2) units: $\bar\rho_i$ (veh/mi.); $\bar v_i$ (mi/hr); variances are squares at these units
   s_i (mi^2/veh.hr);

*Indicates Parameters in Filter Design
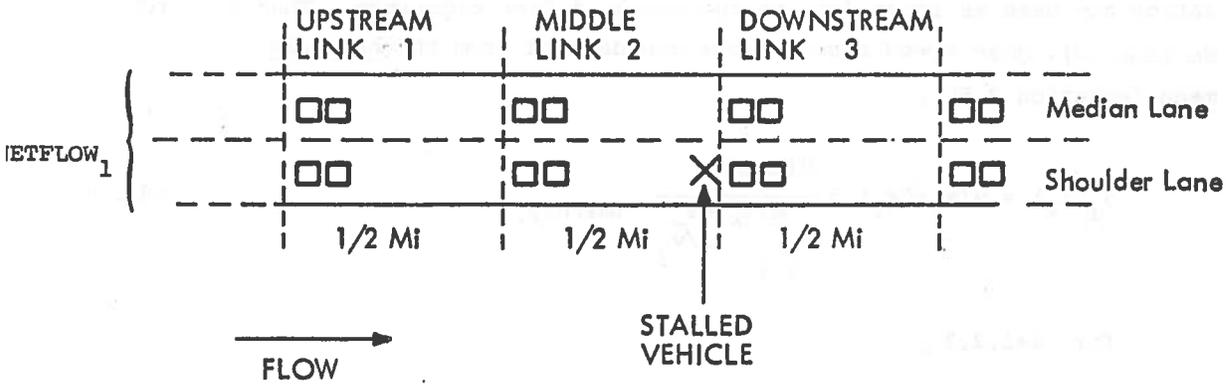
Microsimulation Conditions

1. Exogeneous incoming flow on Link 1, with mean $NETFLOW_1 = 1000$ veh/hr/lane (independent Poisson arrivals on each lane).

2. Standard mix of vehicular traffic (see Appendix A, Table A.3)

3. Run Conditions:

   a. _Non-Incident_: congested middle link 2
      with $\bar{\rho}_2(0) = \rho_{jam}^2$

   b. _Incident_: Initially free flowing traffic, stalled vehicle in right-hand lane at t=300 sec.

Figure 4:15: Study Network for Filter Evaluation in Microscopic Simulation

Incoming (exogeneous) flow was assumed for both runs to be pseudo-Poisson with an intensity of 1000 (veh/hr-lane) or a total of 2000 (veh/hr) for the two-lane freeway. Generation of vehicles, their type and associated driver characteristics, followed exactly the algorithm outlined in Appendix A. For the fixed parameter run (Case 1), we chose initial conditions for vehicle speeds and densities which roughly correspond to the initial conditions used in the macroscopic simulation (Table 4.10); in particular, the middle link was assumed to be initially congested with a density close to the bumper-to bumper concentration.*

4.6.2.2 Filter Design

Using the detector observations described in 4.6.2.1, the EKF algorithm design was developed around the 6-state macroscopic model described in section 4.2, with N=3. The filter state is therefore

$$\underline{x}(t) = (\bar{\rho}_1, \bar{v}_1, \bar{\rho}_2, \bar{v}_2, \bar{\rho}_3, \bar{v}_3) \; .$$

The observations available for each link are modeled in the filter design as (see Table 3.3):

$$\left. \begin{array}{l} y_i^v(t_k) = \bar{v}_i(t_k) + \lambda_v^i(t_k), \\[2mm] y_i(t_k) = \bar{\rho}_i(t_k) + \lambda_\rho^i(t_k), \end{array} \right\} \tag{4.31}$$

for i=1,2,3, where $\{\lambda_v^i, \lambda_\rho^i\}$ are the error processes which model the difference between detector observations and the true spatial mean state. Specification of the appropriate covariances for these processes was initially done based on the detector-loop study, described in section 4.4. Because there is no rigorous correspondence between the error process assumed in the filter design, and the actual error process, some "tuning" of the appropriate covariances will be required. Typical values are given in Table 4.10 Generally change of 4:1 in the assumed observation error covariance matrix entries (from nominal) did not alter the qualitative convergence properties observed below .

---

* ρ_{jam} = 225 (veh/mi) implies a bumper-to-bumper spacing of approximately 24 feet.

or

$$
W = \frac{1}{6}
\begin{bmatrix}
\hat{\bar{\rho}}_1(0) & 0 & \cdots\cdots & 0 & 0 \\
0 & \hat{\bar{y}}_1(0) & \cdots\cdots & 0 & 0 \\
\vdots & & & \hat{\bar{\rho}}_3(0) & 0 \\
0 & 0 & \cdots\cdots & 0 & \hat{\bar{v}}_3(0)
\end{bmatrix}_{6\times 6} .
\tag{4.35}
$$

This amounts to the approximation that the 3-σ (standard-deviation) limits on $\{\rho_i(t), v_i(t)\}$ are within 50 percent of their initial value. Clearly, where actual statistics of these random processes are known or can be measured, the relevant values can be used in lieu of this choice.

The model parameters ν, τ used for the filter design were the same as those used in the macrosimulation runs, and reflect an average of typical values found in the literature ([9-11],[28]). Because of the insensitivity of filter performance to changes in ν and τ, no attempt will be made explicitly to identify them here -- see Grewal [28]. For each link, a <u>logarithmic</u> fundamental diagram was assumed, i.e.:

$$
v_e^i(\rho_i) =
\begin{cases}
v_0^i & : \quad \bar{\rho}_i \le \rho_f^i \\[2ex]
v_0^i \dfrac{\ln[\bar{\rho}_i/\rho_{jam}^i]}{\ln[\rho_f^i/\rho_{jam}^i]} & \rho_f^i < \bar{\rho}_i \le \rho_{jam}^i \\[2ex]
0 & \bar{\rho}_i \ge \rho_{jam}^i
\end{cases}
\tag{4.36}
$$

for i=1,2,3

(see Figure 2. 11). Detailed motivation for this choice of equilibrium speed curve is given in Mitchell [33]. Essentially, this representation appears to provide a better fit at low to medium densities than the simple polynomial fits (linear and parabolic) without increasing the number of free parameters. For the free-flow, fixed-parameter run, the values for v_0^i, ρ_f^i, ρ_{jam}^i obtained by Mitchell were used, and are summarized in Table 4.10.

4.6.3 Results with Fixed-Parameter Design

From the starting conditions in Table 4.10, the microsimulation
was run for a period of approximately three minutes. The time evolution of
traffic behavior with the congested middle link is as shown in Figure 4.16. No
impediments to traffic flow were introduced in this run except for the initial
congestion on both lanes of the middle link at t=0. Performance of the
extended Kalman filter algorithm for a typical simulation run with 5-second
update rate is shown in Figure 4.17 ((a) to (f)). For all links, the error in
density estimation was comparable with the macroscopic simulation runs, while
mean-speed estimates had consistently higher variance (approximately 20 percent
higher with microsimulation data). Mean speed estimation was consistently
poorer with microsimulation data because of the short sampling time (5 sec.)
and inhomogeneous traffic conditions (see Section 4.4). Additional error
sources are related to the detailed operation of the microscopic simulation
(Appendix A). In particular, vehicles entering the freeway from upstream are
given slow initial speeds to avoid overtaking a preceding vehicle the moment
they enter the roadway . This biases the upstream sensor speed measurements
to underestimate the true space mean speed (e.g., on Link 1, Figure 4.17 (b)).

Sensitivity of mean-speed estimation performance to sensor location can
also be seen on the middle link. Initially, the vehicles passing the upstream
detector on link 3 cause slight bias toward over-estimation of mean-speed on
link 2. Note, however, that as traffic speed on the initially congested middle
link increases and conditions tend toward more homogeneous flow, the bias in
mean-speed estimation $(\hat{v}_2)$ tends to vanish. This is entirely consistent with
the analysis in Sections 3 and 4.4.

4.6.4 Results with Adaptive Parameter Estimation

The network in Figure 4.15, and described in Section 4.6.3, was again
employed, with the following modifications:

a. a capacity-reducing incident was introduced on the right-hand
land of an initially free-flowing roadway at t=300 sec;

b. an Augmented extended-Kalman filter was employed to estimate
the available capacity of each link.
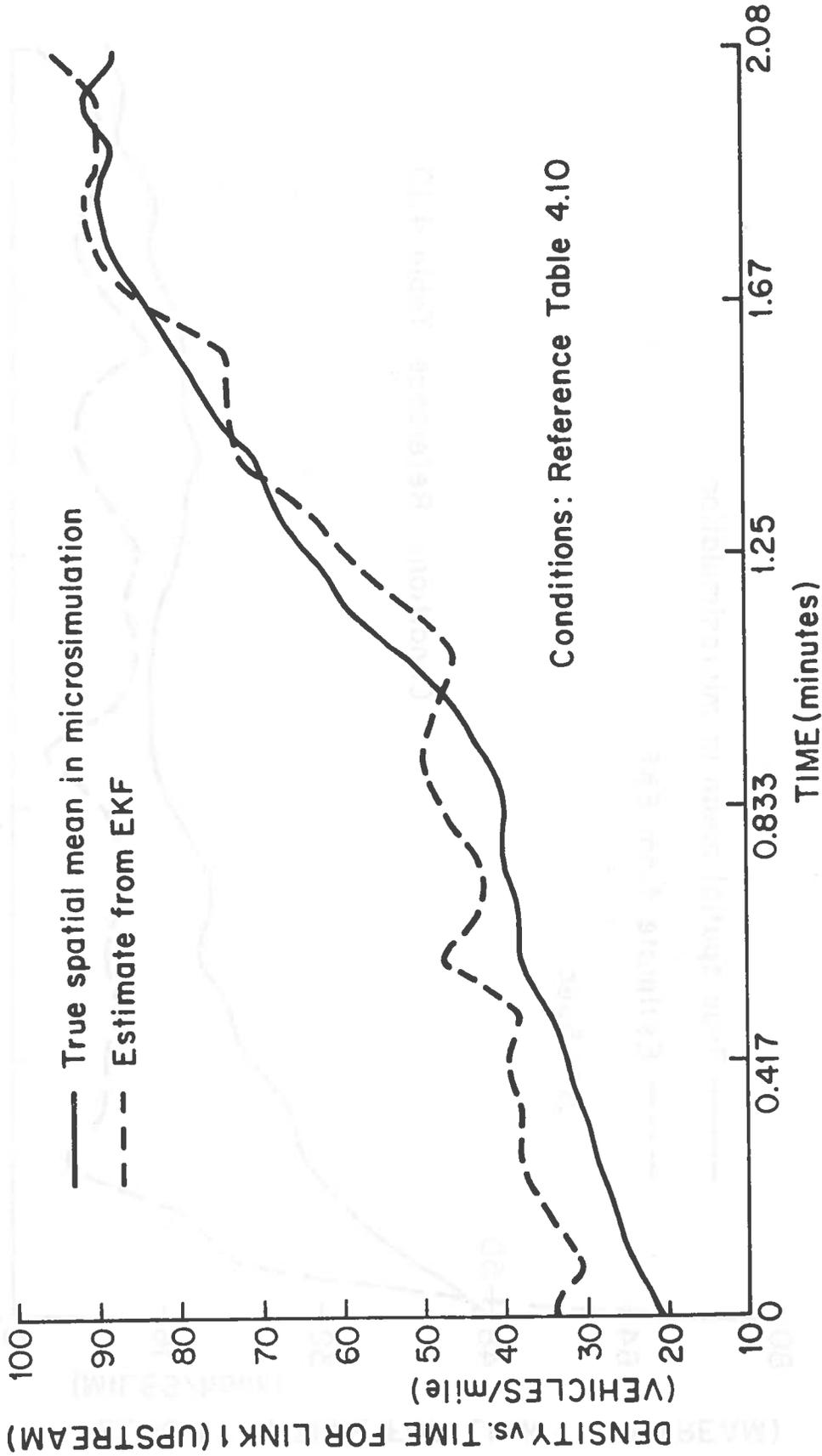
Figure 4.17(a): Fixed-Parameter Macroscopic Filter: Link 1 Density Estimate Versus Time (Congested Middle Link at t=0)
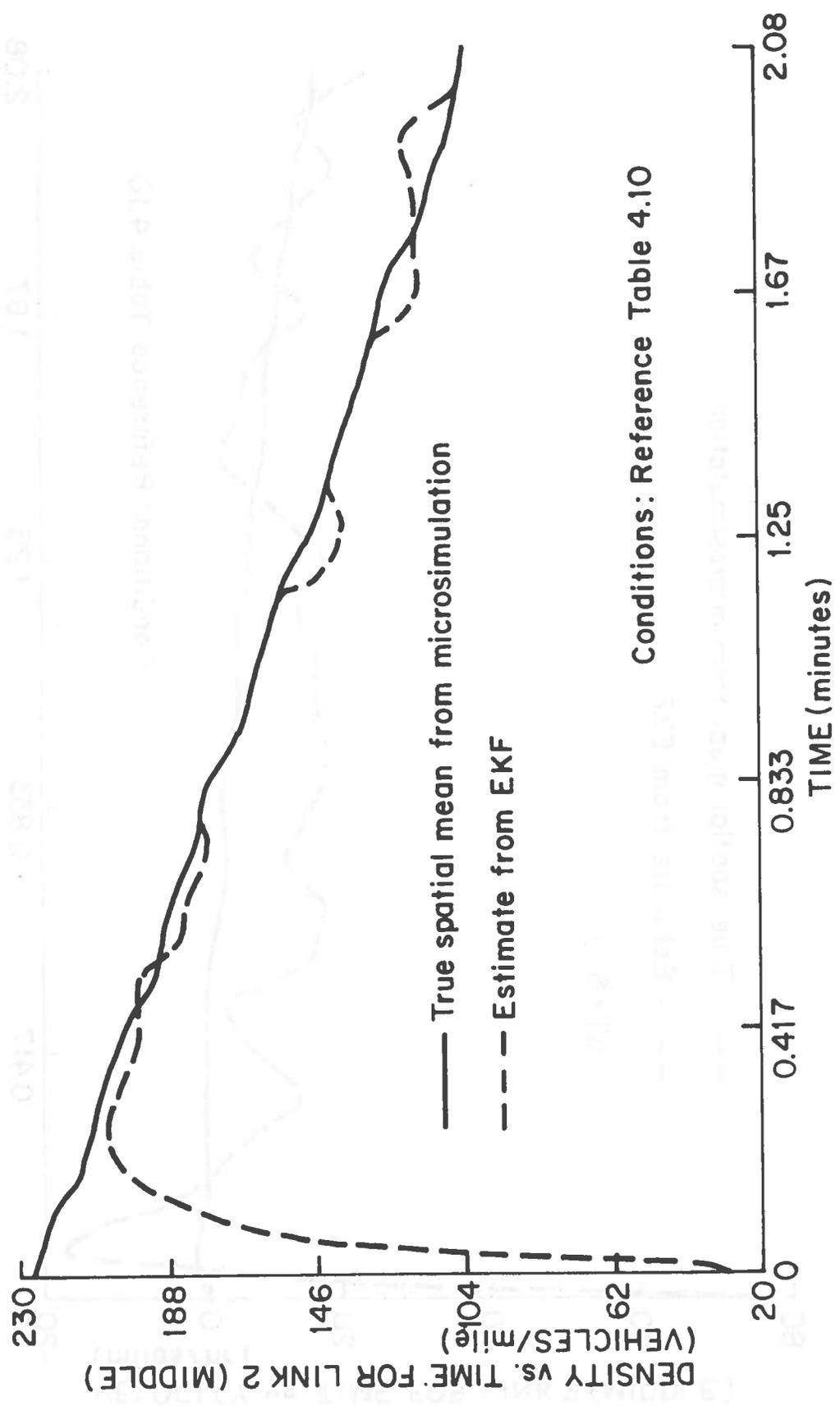
Figure 4.17(c): Fixed-Parameter Macroscopic Filter: Link 2 Density Estimate Versus Time (Congested Middle Link at t = 0)
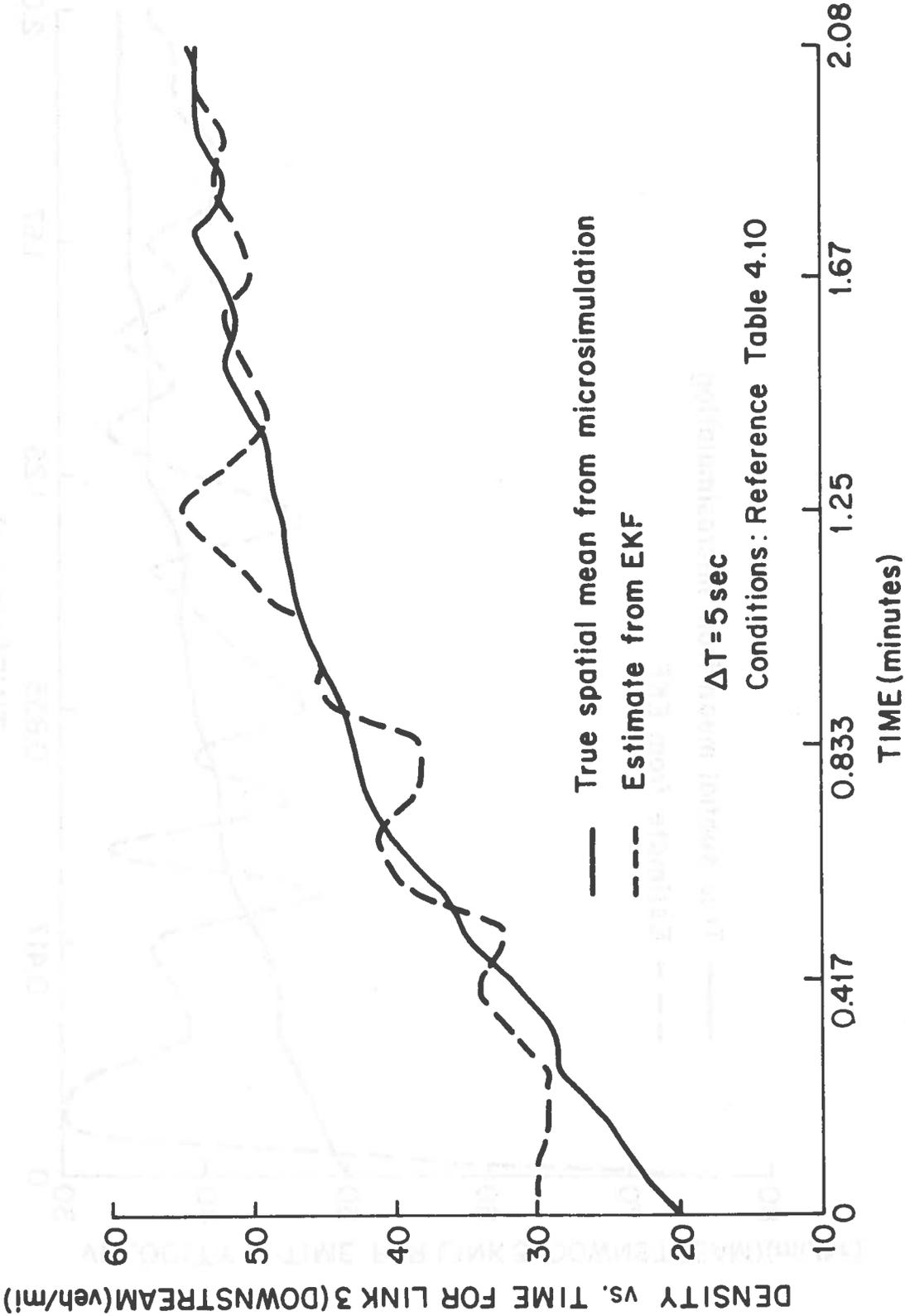
Figure 4.17(e): Fixed-Parameter Macroscopic Filter: Link 3 Density Estimate Versus Time (Congested Middle Link at t = 0)

Recall from Section 2, the available capacity is defined to be

$$\max_{\overline{\rho}_i} \phi_e(\overline{\rho}_i) = \max_{\overline{\rho}_i} \overline{\rho}_i v_e^i(\overline{\rho}_i) , \tag{4.37}$$

for each link i=1,2,3. For the logarithmic fundamental diagram (equation 4.36), each link $v_e^i(\overline{\rho}_i)$ has up to three free parameters, v_0^i, ρ_f^i, and ρ_{jam}^i v_0^i and ρ_{jam}^i were fixed by following the simplifying approach in section 2.4.2 [see also Mitchell [31]]. This leaves one free parameter on each link, ρ_f^i , which parameterizes capacity in a very simple way.

It is easy to show by elementary calculus that with:

$$v_e(\overline{\rho}) = \begin{cases} v_0 & : \quad \overline{\rho} < \rho f \\ v_0 : \dfrac{\ln(\overline{\rho}/\rho_{jam})}{\ln(\rho f/\rho_{jam})} & , \rho_f \leq \rho < \rho_{jam} \\ 0 & : \quad \overline{\rho} > \rho_{jam} \end{cases} , \tag{4.38}$$

$\phi_e(\overline{\rho})$ has a maximum at $\overline{\rho}^*$, where:

$$\overline{\rho}^* = \overline{\rho}^*(\rho_f) = \begin{cases} \dfrac{\rho_{jam}}{e} & : \quad 0 \leq \rho_f < \dfrac{\rho_{jam}}{e} , \\ \rho_f & : \quad \dfrac{\rho_{jam}}{e} \leq \rho_f < \rho_{jam} , \end{cases} \tag{4.39}$$

and

$$\phi_e(\overline{\rho}^*) \equiv C(\rho_f) \equiv \begin{cases} \dfrac{\rho_{jam} v_0}{e(\ln[\rho_{jam}/\rho_f])} & : \quad 0 \leq \rho_f < \dfrac{\rho_{jam}}{e} , \\ \rho_f v_0 : \quad \dfrac{\rho_{jam}}{e} \leq \rho_f < \rho_{jam} . \end{cases} \tag{4.40}$$
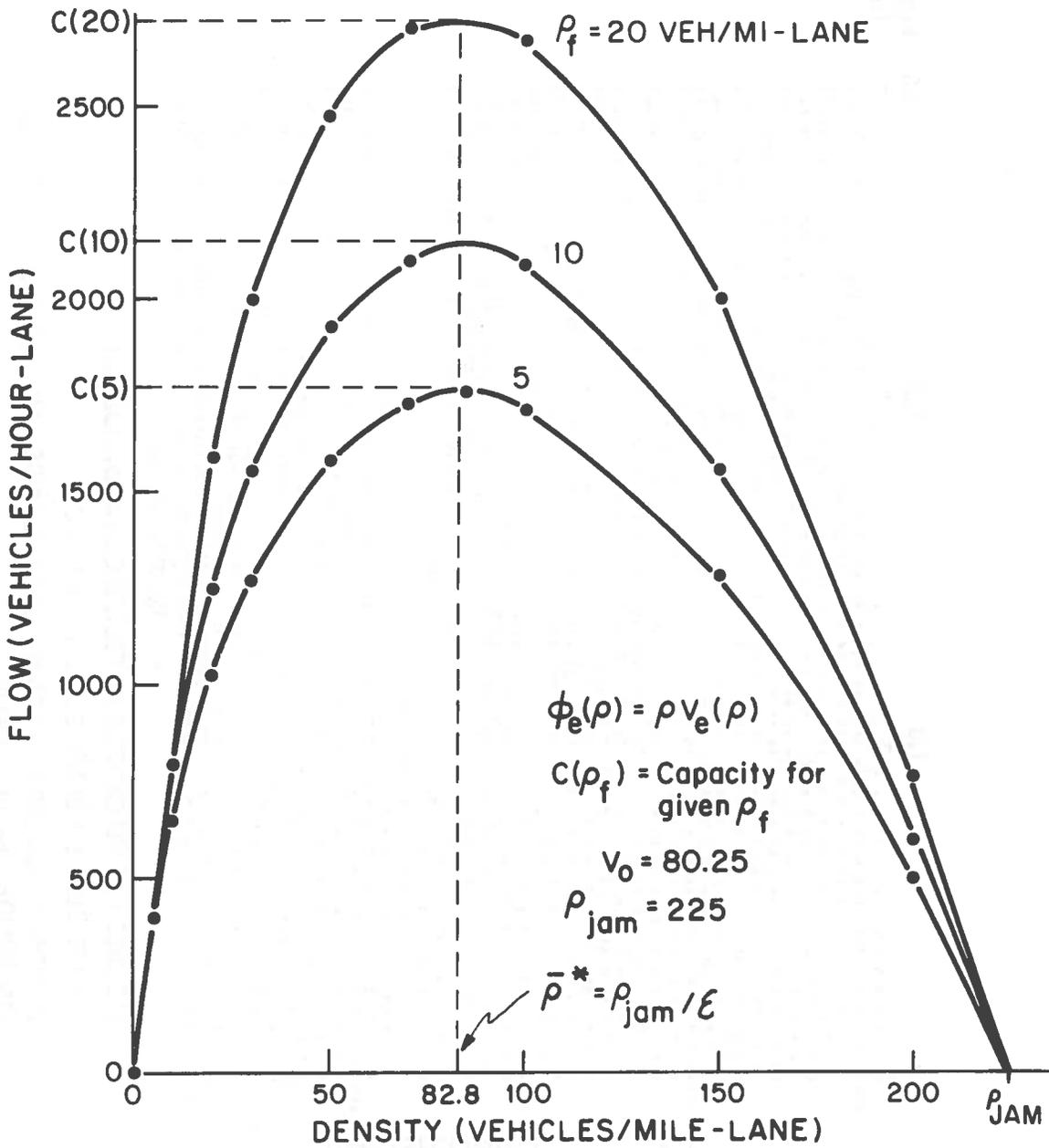
Figure 4.18: Effect of ρ_f Parameter in Logarithmic Fundamental Diagram on Available Capacity, $c(\rho_f)$
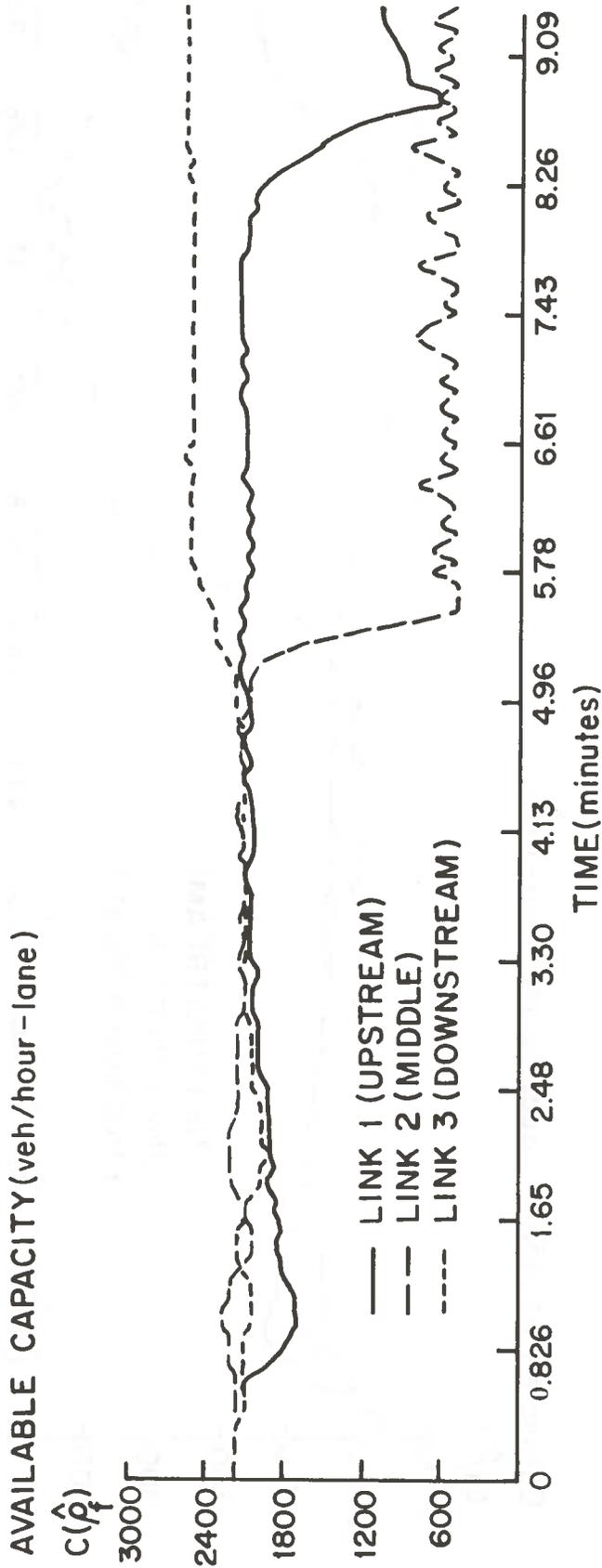
Figure 4.20: Adaptive Estimation of Available Capacity Following Stalled Vehicle Incident: Exact Spatial Mean Variable Observation

stalled vehicle backs into the upstream section (Link 1) at $t \approx 8$ minutes, the corresponding available capacity estimate begins to fall off, reflecting the loss in available capacity downstream.

Although only one lane of the freeway section was physically blocked, note that the estimated available capacity <u>after</u> the incident (≈ 600 veh/hr) is less than one-half of the estimated available capacity before the incident (≈ 2000 veh/hr). This is clearly intuitive since the ability of the remaining lane to carry traffic flow is degraded by queued vehicles behind the stalled one changing lanes to pass.

In figure 4.21, the simulation illustrated in figure 4.20 was re-run, but with detector occupancy and speed measurements employed (instead of the true spatial means) as inputs to the AEKF algorithm. Although the responses differ slightly, the same qualitative "signatures" of the available capacity are exhibited, namely:

a. A marked decrease in the estimated available capacity on the section with the stalled vehicle (link 2).

b. A slight <u>increase</u> in the estimated available capacity on the downstream (uncongested) section (link 3).

c. A finite (approximately 3-minute) response time before the upstream estimated available capacity (link 1) beings to decrease (from queue spill-back).

The main differences between runs with the perfect and detector observations are that with detectors the response to a stalled vehicles is slower (by about 30 seconds), and the estimated available capacity is higher (≈ 900 veh/hr).

It would be erroneous to conclude from these results alone that available capacity can be accurately estimated in absolute terms. Rather, the results show a clear and unambiguous signature that identifies a <u>relative</u> change in capacity after the incident.

The relative change in the estimated available capacity has behavior that is intuitive. Moreover, this behavior of estimated capacity is exhibited with an extremely simple (one-parameter per-link) model. Clearly more free parameters could be added to the available degrees of freedom where circumstances warrant. However, with only this simple one parameter characterization of available capacity, we have subsequently been able to show that systematic approaches to incident detection are feasible, primarily as a consequence of the parameter signature that accompanies a capacity reducing incident [see Willsky et al [34] for details).

With adaptive estimation of available capacity, the error variance of the spatial mean speed and density estimate is also slightly improved (by about 10 percent) over the fixed-parameter case under conditions, the improvement in error Under incident and non-homogeneous flow conditions, the improvement in error variance is approximately 30 percent for runs such as the one illustrated above. Moreover, no bias was observed to develop in the estimates (see Mitchell [33] for details).

A consistent finding in both the fixed parameter and adaptive parameter estimation schemes is the apparent insensitivity to lateral dynamic traffic behavior. In both the non-incident and incident scenarios, passing and lane-changing behavior was intentionally programed into the corresponding microscopic simulations. Even though such lateral behavior is not explicitly modeled in the filter design, filter performance was not degraded.

In some cases, the problems with numerical precision can be handled by use of floating-point and/or multiple precision fixed-point arithmetic. A penalty incurred with this approach is the increased time overhead per operation. When combined with the large number of operations required, this may result in filter update rates $(1/\Delta T)$ which are unacceptably large on a small machine. Alternatively, different _equivalent_ realizations of the Kalman filter, such as the class of _square-root_ algorithms [35] are possible. Square-root algorithms, for example, effectively yield double-precision performance without changing the physical machine word length, albeit at a slight increase in the number of operations that must be performed.

Considerable literature exists on techniques for dealing with finite-word length effects (a highly recommended source is Bierman's Book [35]). Moreover since we did not directly evaluate the impact of finite word-length effects on the traffic surveillance algorithms developed, we now focus attention on the problem of the "dimensionality curse" for which problem specific alternatives can be proposed.

4.7.2 Effects of Algorithm Complexity and "Dimensionality Curse"

In this section we examine implications of the AEKF algorithm complexity on computer-time and storage requirements, and examine possible approaches to deal with the need to operate the algorithm in real time. The essential problem is that the number of equations that must be propagated _on-line_ with the AEKF becomes enormous as the number of states increases, growing essentially as n^2, where n is the total number of filter states, i.e.,

$$n = 2N + n_p \ , \tag{4.42}$$

with, N, the number of freeway sections, and, n_p, the total number of free-parameters being estimated. In Table 4.12, we have summarized the approximate count of state and covariance equations that must be propagated on-line with the AEKF algorithm as a function of the number of freeway sections (N), and free parameters (n_p). Note for example, that with as few as N=8 sections with two estimated parameters per section, there are some 500 differential equations that must be propogated on-line!

The bulk of the computer time and storage required by the continuous-discrete AEKF can be attributed to the state-prediction and covariance propagation, as summarized in Figure B-1, and from Section 3, we have:

Predict Step

$$\hat{\underline{x}}_{k+1|k} = \int_{t_k}^{t_{k+1}} f(\hat{\underline{x}}(\tau|t_k))d\tau \ , \tag{3.19}$$

$$x(t_k|t_k) = \hat{\underline{x}}_{k|k} \ ,$$

$$\textstyle\sum_{k+1|k} = \hat{\Phi}_{k+1|k} \textstyle\sum_{k|k} \hat{\Phi}^T_{k+1|k} + \Xi_k \ , \tag{3.20}$$

Update Step

$$\hat{\underline{x}}_{k+1|k+1} = \hat{\underline{x}}_{k+1|k} + K_{k+1} [\underline{y}_k - \underline{g}[\hat{\underline{x}}_{K=1|k}]) \ , \tag{3.23}$$

$$\textstyle\sum_{k+1|k+1} = (I-K_{k+1} \hat{G}_{k+1}) \textstyle\sum_{k+1|k} \ , \tag{3.25}$$

where the Kalman gain, K_{k+1}, is:

$$K_{k+1} = \textstyle\sum_{k+1|k} \hat{G}^T_{k+1} \left[\hat{G}_{k+1} \textstyle\sum_{k+1|k} \hat{G}^T_{k+1} + \Lambda \right]^{-1} \ , \tag{3.24}$$

the main numerical activities which must be performed on-line can thus be seen to be:

    a. Numerical integration (in (3.19), and solution for transition matrix $\Phi_{k+1|k}$),

    b. Matrix Inversion (in (3.24)), and

    c. Matrix Multiplication (all).

These activities are certainty not unique to Kalman filtering, and we will not attempt to identify "efficient techniques" here.* Rather, we shall attempt to identify properties of the Kalman filter structure, which in the problem specific traffic surveillance context, permit reduction

---

*For an introduction inthe context of Kalman filtering, refer to [21], Chapter 8.

[FR2]. From the structure of the model equations in Section 2, we have that:

$$\frac{d\bar{v}_j}{dt} = -\bar{v}_j \left(\frac{\bar{v}_j - \bar{v}_{j-1}}{c_j} \right) - \frac{1}{\tau} \left(\bar{v}_j - \bar{v}_e^j(\bar{\rho}_j) \right)$$

$$\frac{-\nu}{\tau} \frac{1}{\bar{\rho}_j} \left[\frac{\bar{\rho}_{j+1} - \bar{\rho}_j}{c_j} \right] , \qquad\qquad (2.22)$$

where $c_j \equiv [\Delta xj + \Delta x_{j-1}]/2$ and

$$\frac{d\bar{\rho}_j}{dt} = \frac{\bar{v}_{j-1}\,\bar{\rho}_{j-1} - \bar{v}_j\bar{\rho}_j + r_j - w_j}{\Delta x_j} . \qquad\qquad (2.23)$$

Thus, the $(\bar{v}_j, \bar{\rho}_j)$ dynamics depend <u>only</u> on the dynamics of the <u>nearest neighbor</u> links, $(\bar{\rho}_{j-1}, \bar{v}_{j-1})$ (upstream), and $(\bar{\rho}_{j+1}, \bar{v}_{j+1})$ (downstream). Based on this decoupling from other links ($j \pm k$) for $k > 1$, Looze [FR2] showed that a viable strategy for decoupling of the filter was to:

a. Group the freeway into coupled subsection, each
   with 3 links (6 states) per subsection;

b. Assume that the coupling terms from states associated
   with neighboring subsections are <u>perfectly known</u>; and

c. Design the 3-state EKF for each subsystem assuming the
   known observations are <u>time-varying</u> terms in the
   subsystem dynamics.

The approach is illustrated in Figure 4.22, for a 6-section (N=6 network. In filter "A," assumption (b) enters into the design in (c) by making the local dynamics time-varying by the assumption that $(y_4^v(t), y_4^\rho(t))$ are known. Conversely, $(y_3^v(t), y_3^\rho(t))$ make the local dynamics time-varying in filter "B".

To recapitulate:

1. Decomposition based on the model structure results in N/3 filters, each with only 3(2+m) states;[1]

2. Filters are _decoupled_ except for observations from boundary section (Figure 4.22); and

3. Coupling makes filters time varying, but does not increase their dimension.

The resulting extended-Kalman filter will be suboptimal in a strict mathematical sense. However in evaluating this approach in a _macroscopic_ simulation, Looze showed that even linear time invariant approximations of this type do not degrade estimate performance by more than 5 to 10 percent, when the model parameters are known.[2]

An approximate feeling for the decrease in algorithm complexity can be obtained by examining Table 4.13, in which the total number of state and covariance equations required in a standard and decentralized EKF are compared. Notice that the _relative_ saving in complexity is _greater_ with a larger freeway (or equivalently a larger number of sections for a given freeway). Another advantage of the decentralized EKF is that each filter algorithm can be operated in _parallel_, providing possible opportunities for multi-processing and/or microcomputer architectures.

The above comparison is only approximate since the ultimate program complexity and operation count will depend on the particular machine used. Although we did not evaluate the DEKF performance in the microsimulation, a _macroscopic_ simulation run provides some basis for comparison. EKF and DEKF filters were designed for a 6-link (3.0-mile) model, which was a concatenation of two copies of the one in Section 4.5.[3] Identical speed-density initial conditions were used. For a 25 minute simulation run with 1-minute updates and one unknown parameter per link, the filter computation time was as follows:

---

[1] 6 states + m unknown parameters per subsection, in general.
[2] See Looze et al [FR2].
[3] See Table 4. 8 for conditions and unknown parameter values.

$\begin{cases} \text{Decentralized filter} = 0.684 \text{ (c.p.u-minutes),} \\ \text{central processor time (DEKF) (2 filters, 9 states per filter),} \end{cases}$

$\begin{cases} \text{Centralized filter} = 1.11 \text{ (c.p.u-minutes),} \\ \text{Central processor time (EKF) (1-filter with 18 states),} \end{cases}$

or about a 40 percent decrease with the DEKF implementation. In comparing the filter performance between the EKF and DEKF in the macroscopic simulations, mean speed and density estimates in the DEKF generally took 15 percent longer to converge to within 1-σ of the true state, with no statistically significant difference in accuracy. Error variance in the parameter estimates was about 11 percent higher, however, with the DEKF. These numbers should be taken with the caveat that they do not apply to microscopic freeway data.[*] We would expect similar relative behavior between the EKF and the DEKF, when tested in the microsimulation, but this demonstration remains as a future activity (see section 5).

---

[*] Microsimulation runs were infeasible due to computer budget constraints.

Furthermore, by relating this quantity to a parameter of the fundamental diagram, $v_e(\bar{\rho})$, on each link and estimating this parameter on-line, a different perspective of this quasi-equilibrium relationship emerges. In particular, one does not derive the $v_e(\bar{\rho})$ curve from an independent, off-line (e.g. least-squares) fit of a curve to <u>historical</u> speed-density data, but adapts it on-line to existing conditions. Not only is qualitative incident information provided with this procedure, but, at least in our microscopic simulations, overall estimate accuracy is enhanced, by removing the bias in speed estimates. Clearly, more detailed evaluation of results from <u>both</u> the microscopic simulation and the actual data would be a relevant activity.

From a quantitative performance viewpoint, the extended Kalman filter is a very adequate data-processing algorithm in this application. A substantial hurdle must be surmounted, however, if the algorithm is to be practically implemented in realistically sized corridors. The problem with the straightforward (i.e., standard) EKF implementation is simply the enormous growth in the number of equations which must be solved <u>on-line</u> as the number of freeway sections increases. Based on the sparsely coupled structure of the traffic model, however, there is cause for optimism that some of the decomposition/decentralization ideas presented in Section 4.7 and in Looze [FR2] may provide the necessary dimensional simplification. The DEKF algorithm associated with these procedures possesses the additional structural property that parallel implementation with multi-processor computer architectures is feasible. A high priority should be attached to future developments along these decompostion lines, particularly because of the close correspondence with the decentralized control implementation work [FR2].

Many other performance complexity tradeoffs need to be addressed in the context of practical implementation. Included are the impact on surveillance algorithm performance of changes in loop detector spacing from the 1/2-mile spacing used here and widely deployed in practice. Some of our results for speed and density estimation suggest that density estimation suggest that greater spacing could be used on raodway-segments which have no entry/exit ramps, whereas closer spacing would probably have been advantageous in proximity to ramps, lane drops, and weaving sections if they were present. On the

6. REFERENCES

[1]   Lighthill, J.J., Whitham, G.A., "On Kinematic Waves II. A Theory of
        Traffic Flow on Long Crowdal Roads," Proc. Royal Soc. London, Series
        A, V.229, p. 317-345, 1955.

[2]   Gazis, D. C., (Editor), Traffic Science, John Wiley and Sons, New
        York, 1974, Chapter 1.

[3]   Phillips, W.F., "Kinetic Model for Traffic Flow," Draft Final Report
        DOT-OS-40098, Utah State U., Logan, Utah, Dec. 1976.

[4]   Payne H. J., "Models of Freeway Traffic Control," in Mathematical
        Models of Public Systems, (Simulation Council Proc., Vol. 1., No. 1)
        January 1971.

[5]   Prigogine, I., Resibois, R., Herman, R. and Anderson, R.L. "On a
        Generalized Boltzmann-Like Theory of Traffic Flow," Bull. Acad. R.
        Sci. Lett. Beaux-Arts Belg. Classe des Sciences 48, p. 805-814, 1962.

[6]   Gafarian, A. V., Munjal, P. K. and Pahl, J., "An Experimental Vali-
        dation of Two Boltzmann-Type Statistical Models for Multilane Traffic
        Flow," Transportation Research, 5, p. 211-224, 1971.

[7]   Gazis, D.C.,R. Herman, and R. B. Potts, "Car-Following Theory of
        Steady-State Traffic Flow", Operations Research, Vol. 7, 1959.

[8]   Newell, L., "Nonlinear Effects of the Dyanmics of Car-Following",
        Operations Research, Vol. 9, 1961.

[9]   Isaksen, L., "Supoptimal Control of Large Scale Systems with Appli-
        cation to Freeway Traffic Regulation," Ph D. Thesis E.E. Dept.,
        U. Of Souther California, 1971.

[10]   Isaksen,L.,and H. J. Payne, "Freeway Traffic Surveillance and
        Control", Proceedings of the IEEE, Vol. 61, no. 5, May, 1965.

[11]   L. Isaksen and H.J. Payne, "Regulation of Freeway Traffic",
        Proceedings 1972 Joint Automatic Control Conference, Stanford,
        California.

[12]   Kurkjian, A., "Dynamic Delection and Identification of Incidents on
        Freeway, Vol. II: Approaches to Incident Delection Using Presences
        Detectors," Final Report No. DOT-OS-60137, Also Available as MIT
        Electronic Systems Laboratory, Report ESL-R-765, September 1977.

[13]   Mikhalkin,B., "Estimation of Speed from Presence Detectors,"
        Highway Research Board Record, No. 388, p. 73-83, 1972.

[14]   Mikhalkin, B.,"Estimation of Roadway Behavior Using Occupancy
        Detectors", Ph.D. Dissertation, Dept. of Freeway Operations, Los
        Angeles, California.

[30]   "Network Flow Simulation for Urban Traffic Control System: Phase II,"
         Federal Highway Administration Report No. FHWA-RD-73-87, March, 1974.

[31]   Wicks, D.A., and E.B. Lieberman, "Development and Testing of INTRAS,
         A Microscopic Freeway Simulation Model, Vol. 1, Program Design and
         Parameter Calibration", Inbrim Report to U.S. F.H.W.A. Office of
         Research and Development, prepared by KLD Associates (Report KLD TR-40),
         January 1976.

[32]   A.D. St. John, "Study of Traffic Phenomena Through Digital Simulation,"
         Final Report for U.S. Dept. H.E.W., Public Health Service, Research
         Grant AC-000106, 1966.

[33]   Mitchell, W.J., "The Estimation and Simulation of Freeway Traffic Flow
         Using Car Following and Fluid-Analog Models," S.M. E.E. Thesis, M.I.T.
         Department of Electrical Engineering and Computer Science, June 1977.

[34]   Willsky, A.S., et al, "Dynamic Detection and Identification of
         Incidents on Freeways Volume I: Summary," Final Report for U.S.D.O.T.
         Contract DOT-OS-60137, Available as MIT Electronic Systems Laboratory
         Report ESL-R-764, September 1977.

[35]   Bierman, G.J., _Factorization Methods for Discrete Sequential Estimation_,
         Academic Press, Vol. 128, New York 1977.

[36]   Orlhac, D., "Estimation of Traffic Variables Via Extended Kalman
         Filter Methods," S.M.E.E. Thesis, M.I.T. Department of Electrical
         Engineering and Computer Science, Cambridge, Mass., August 1975.

[37]   Nahi, N.E., "Freeway Data Processing," _Proc. IEEE_, Vol. 61, May 1973
         pp. 537-541.

[38]   Gazis, D.C. and C.K. Knapp, "On-line Estimation of Traffic Densities
         from Time-Series of Flow and Speed Data," _Transportation Science_,
         Vol. 5, No. 3, August 1971, pp. 283-301.

APPENDIX A

MICROSCOPIC VEHICLE SIMULATION PROGRAM

A.1 PROGRAM OVERVIEW

This appendix details the microscopic vehicle simulation used for evaluation of the state estimation and parameter identification schemes documented in this report. It was developed as an interim evaluation device since the original planned use of the INTRAS simulation program [31] was not possible. Some of the microscopic behavioral logic is adapted from St. John et al [32] although modifications to reflect driver types and acceptable risks more typical of urban freeways than rural roads have been made. It was not our attempt to provide a simulation program in this research, and hence, detailed program listings are not provided. FORTRAN listings will be made available upon request.   In Section A.1, we provide a summary of the main program features; Section A.2 motivates the philosophy for program development; and, Section A.3 details block diagrams of the high-level and most-critical low-level routines.

A.1.1 Main Program Features

Two-lane freeway, with lane changing and passing.

Up to 12 entrance ramps in simulated freeway, including microscopic merging logic from freeway-entrance ramps.

Freeway may be divided into, at most, 12 links with different types of grade, and geometry, so that spatial mean values of densities and velocity may be obtained for each link.

Nine different driver personalities (corresponding to various desired speeds and acceptable risk levels) and six different vehicle types (corresponding to performance limits) may be used in the simulation to model the mixture on an actual freeway.

Up to 12 traffic-presence detectors may be included at specified locations of the model freeway, and their presence and velocity outputs as a function of time can be used for real-time estimation or saved for future analysis.

The user can either explicitly specify the initial vehicle statistics for a particular run, or allow the program to generate them for the user.

The input flow to each lane of the freeway and each ramp is an independent Poisson process with user-specified mean.

The user can, if desired, simulate an incident which closes a lane of the freeway by creating a stalled vehicle at a specified time and position of the model freeway.

The microscopic simulation program is a FORTRAN routine which, in its present form, has been implemented on an IBM 370/168 computer. It requires about 340 kilobytes of main storage to compile the FORTRAN source code, using the IBM FORTRAN IV G1 compiler, and to **execute a typical** simulation. Approximately 1.2 minutes of central processor time are required to run a simulation of 1200 steps and a total length of 1.5 miles of high-density traffic on the 168 using the OS/MVT operating system. This time figure is

> linearly dependent upon the total freeway length;
>
> linearly dependent upon the number of time steps;
>
> linearly dependent upon the traffic density (since it is linearly dependent upon the number of vehicles in the system); and
>
> virtually independent of the number of links, ramps, and presence detectors used in the simulation.

This program consists of a main program which keeps track of the current time, and which vehicle is being processed at this time, and subroutines which are called to perform specific parts of the simulation. A qualitative overview of the routines and program organization are given in Table A.1 and Figure A.1.

A.1.2 Input Data Requirements

Data are read into the program from the standard FORTRAN input file.

> Random Number Kernel (used in generating pseudo-random numbers for all stochastic processes used in the program.
>
> Freeway Topology Data
>> Length of Simulated Highway

Number and Position of Link Endpoints, Entrance Ramps, and
Presence Detectors;

Stalled Vehicle Time and Location (if used); and

Probability Distributions for Vehicle and Driver Types.

Initial Conditions

Either

Statistics of all vehicles at t = 0

or

Initial values for spatial mean density and input flow for each
lane of each link of the freeway.

Output flag specifying what data (if any) are to be written to the
output file.

Time Step and Run Duration Data.

A.1.3 Microscopic Simulation Output

The program generates the following plots on the line printer:

1)   Every 10 time steps, vehicle positions and velocities.

2)   At the end of the run (for each lane of each link)

spatial mean density vs. time,

spatial mean velocity vs. time,

average travel time (time required for vehicles leaving this link
during the time interval to reach its end) vs. time, and

spatial mean number of stops vs. time.

3)   At the end of the run (for each lane of the simulation)

velocity vs. time and position.

The program will also, at the user's request, write out the spatial mean
density and velocity data for each link and the presence data for each detector
to output files where they may be punched into cards, printed, or saved in on-
line storage for use by other programs. Output data may be obtained from either
lane of the simulation or as the average of data from both lanes.

x_t POSITION OF FOLLOWING VEHICLE AT TIME t

$\dot{x}_t$ VELOCITY OF FOLLOWING VEHICLE AT TIME t

$\ddot{x}_t$ ACCELERATION OF FOLLOWING VEHICLE AT TIME t

Δ_x DISTANCE TO LEADING VEHICLE

v_L VELOCITY OF LEADING VEHICLE

τ_1, τ_2 TIME CONSTANTS OF LINEARIZED MODEL

Figure A.2:   Definition of Freeway Variables

TABLE A.3

STANDARD MIX OF TRAFFIC

a. Vehicle Type				
Description	Acceleration	Length (ft)	Width (ft)	Percent
car	good	18	6	43.750
car	medium	18	6	39.375
car	poor	18	6	4.375
van	-	20	7	6.250
truck	-	40	8	3.750
truck	-	60	8	2.500

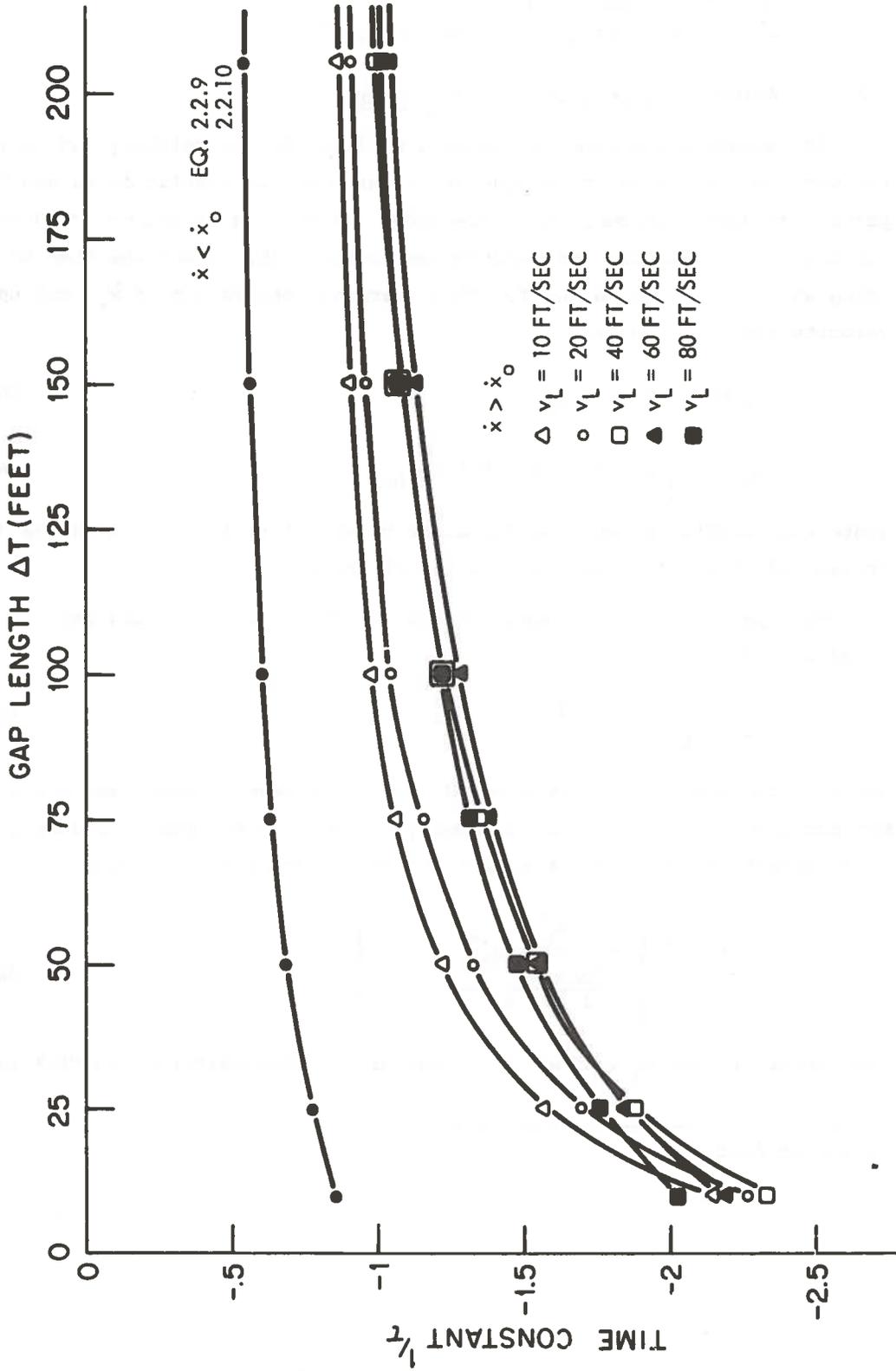| b. Driver Type		
Driver	Risk Level	Percent
slow	low	1
slow	medium	3
slow	high	1
medium	low	9
medium	medium	27
medium	high	9
fast	low	10
fast	medium	30
fast	high	10

Figure A.3: Time Constant $1/\tau$ Versus Gap Length Δx for Medium-Speed, Medium-Risk, Drivers

(which reflect the maximum deceleration they are willing to use in normal maneuvers). Typical values for a medium risk, moderate speed driver are $R_0 = 0.54$, $R_1 = 0.00645$. What we call a <u>standard traffic</u> mix is given in Table A.3.

A.2.3 Ramp-merging Algorithm

A vehicle will merge into the freeway from an entrance ramp if an adequate space gap exists in the right-hand lane of the freeway at the beginning of the time step. An adequate space gap must meet both of the following conditions:

1) the acceleration of the merging vehicle as a function of the new leader must be greater than the risk level of the merging driver; and

2) the acceleration of the new follower as a function of the merging vehicle must be greater than the risk level of the following driver.

A.2.4 Lane-changing Algorithm

For a vehicle to change lanes, all of the following four conditions must be met:

there must be adequate space in the other lane for the lane changing vehicle to enter;

its computed acceleration as a function of its leader in the present lane must be less than the maximum acceleration of which it is capable at its present speed;

if it is traveling in the inside (right-hand) lane, its computed speed as a function of its current leader must be less than its free speed (note that, for this reason, vehicles in the right-hand lane tend to stay there, while vehicles in the left-hand lane tend to move to the right); and

the acceleration as a function of the new leader in the adjacent lane must be greated than its acceleration as a function of its present leader.

A.3 PROGRAM OPERATION

Figure A.1 gives an overview of the logic of the microscopic simulation program. A qualitative description of the subroutines used in the program is given.
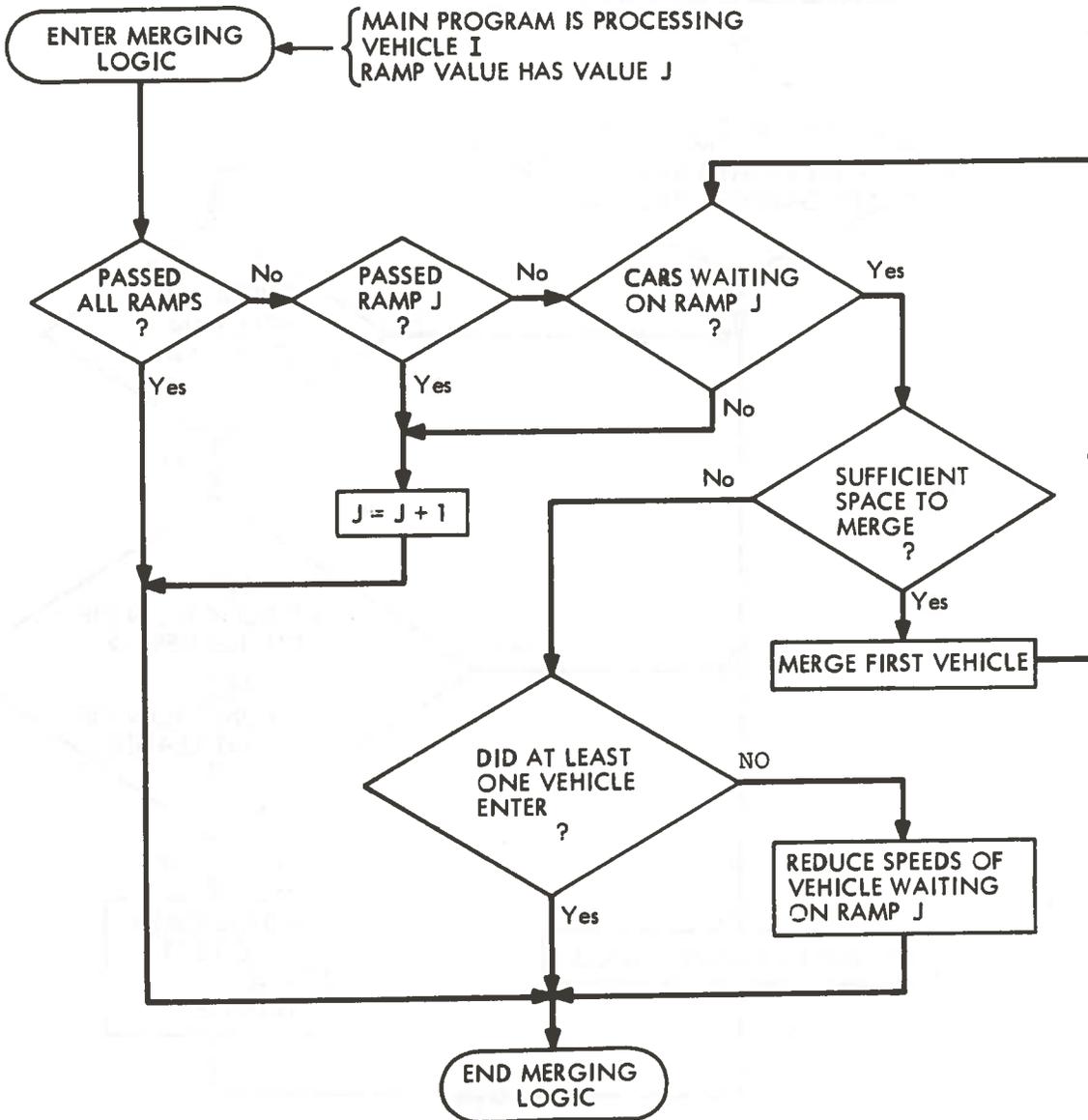
Figure A.4: Merging Logic of Microscopic Simulation

following equations are used to compute the updated speed and position of this vehicle as a function of the leader in its new lane.

A.3.4 Updating of Cumulative Statistics

The program then checks to see if the vehicle leaves the freeway, and if so, it removes its entry from the master file (this entry is now available for assignment to a new arrival). Then the cumulative link statistics

Spatial Mean Density $\quad \overline{\rho}_i = \dfrac{N_1}{x_{i+1} - x_i}$, $\qquad$ (A.11)

Spatial Mean Velocity $\quad \overline{V}_i = \dfrac{\sum\limits_{j\,\epsilon\,\text{link }i} \dot{x}_j}{N_i}$, $\qquad$ (A.12)

Average Travel Time $\quad \overline{\Delta t}_i = \dfrac{\sum\limits_{j\,\epsilon\,\text{link }i} (t - t_{in,j})}{NL_i}$, $\qquad$ (A.13)

Spatial Mean Number of Stops $\quad \overline{NS}_i = \dfrac{\sum\limits_{j\,\epsilon\,\text{link }i} NS_j}{NL_i}$, $\qquad$ (A.14)
of Stops

where N_i = number of vehicles on link i at time t ,

$\quad NL_i$ = number of vehicles leaving link i at time t ,

$\quad \dot{x}_j$ = velocity of vehicle j at time t ,

$\quad t_{in,j}$ = time vehicle j entering freeway,

$\quad NS_j$ = number of stops vehicle j has made during run ,

are finally updated, presence detector data (the number and mean velocity of vehicles crossing this detector during this time step) are recorded, and the processing of the next vehicle begins.

Once all vehicles have been processed, the time step is incremented and new vehicles are placed on the end of the freeway and the ramps (if sufficient unoccupied space exists at the end of the freeway or ramps). Independent Poisson processes are used to generate the vehicle interarrival times for each ramp and lane of the freeway although the program can be adapted to use other types of distributions.

APPENDIX B
SUMMARY OF EQUATIONS FOR CONTINUOUS-TIME DISCRETE
OBSERVATION EXTENDED KALMAN FILTER

This appendix summarizes an alternative form of the discrete-observation continuous-state extended Kalman filter which does not use the transition matrix, $\Phi(t_{k+1}, t_k)$, but relies on numerical integration of the continuous dynamical state and covariance equations directly.

Step (1) Initialization

$$\hat{x}_{0|0} = x_0 \qquad\qquad (B.1)$$

$$\Sigma_{0|0} \quad \Sigma_0 \qquad\qquad (B.2)$$

Step (2) Nonlinear Continuous State Prediction

$$\hat{x}_{k+1|k} \equiv \hat{x}(t_{k+1}|t_k) \, ,$$

$$= \int_{t_k}^{t_{k+1}} f(\hat{x}(\tau|t_k)) \, d\tau \qquad\qquad (B.3)$$

the integration sweep (B.3) generates,

$$\hat{F}(t) \equiv \left[\frac{\partial f_i}{\partial x_j} (\hat{x}(t|t_k)) \right]_{n \times n} \quad t_k \leq t \leq t_{k+1} \, , \qquad (B.4)$$

and

$$\hat{G}_{k+1} \equiv \frac{\partial g_i}{\partial x_j} (\hat{x}(t_{k+1}|t_k))_{n \times n_p} \, . \qquad\qquad (B.5)$$

Step (1)
Initialization

$$\text{INITIALIZE } \hat{x}_{0|0} = \bar{x}_0$$

$$\sum_{0|0} = \sum_0$$

- -

Step (2) Non-Linear
(Continuous)
State-Prediction

(a) Integrate (Predictor) Dynamics

$$\hat{x}_{k+1|k} \equiv \hat{x}(t_{k+1}|t_k) = \int_{t_k}^{t_{k+1}} f(\hat{x}(\tau|t_k))d\tau ; \quad \hat{x}(t_k|t_k) = \hat{x}_{k|k}$$

yields:

$$\hat{F}(\tau) \equiv \left[\frac{\partial f_i}{\partial x_j}(\hat{x}(\tau|t_k)) \right] : \quad t_k \le \tau \le t_{k+1}$$

$$G_{k+1} \equiv \frac{\partial g}{\partial x}[\hat{x}(t_{k+1}|t_k)]$$

$\hat{x}_{k|k} \longrightarrow$

$\sum_{k|k} \longrightarrow$

$\hat{F}(\tau) \qquad \hat{G}_{k+1} \qquad \hat{x}_{k+1|k}$

(b) Linearized transition matrix update

$$\Phi(t_{k+1},t_k) = \int_{t_k}^{t_{k+1}} \hat{F}(\tau)\Phi(\tau,t_k)d\tau)]d\tau$$

$$\Phi(t_k,t_k) = I$$

$\Phi_{k+1|k}$

- -

Step (3),(4)
DISCRETE (3)
Covariane
Prediction

Kalman Gain
update (4)

$$\sum_{k+1|k} = \Phi_{k+1|k} \sum_{k|k} \Phi_{k+1|k}^T + \Xi_k$$

$$K_{k+1} = \sum_{k+1|k} \hat{G}_{k+1}^T [\hat{G}_{k+1} \sum_{k+1|k} \hat{G}_{k+1}^T + \Lambda]^{-1}$$

- -

Step (5),(6)
Estimate and
Covariance
up-date

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(y_k - g[\hat{x}_{k+1|k}])$$

$$\sum_{k+1|k+1} = (I-K_{k+1}\hat{G}_{k+1})\sum_{k+1|k}$$
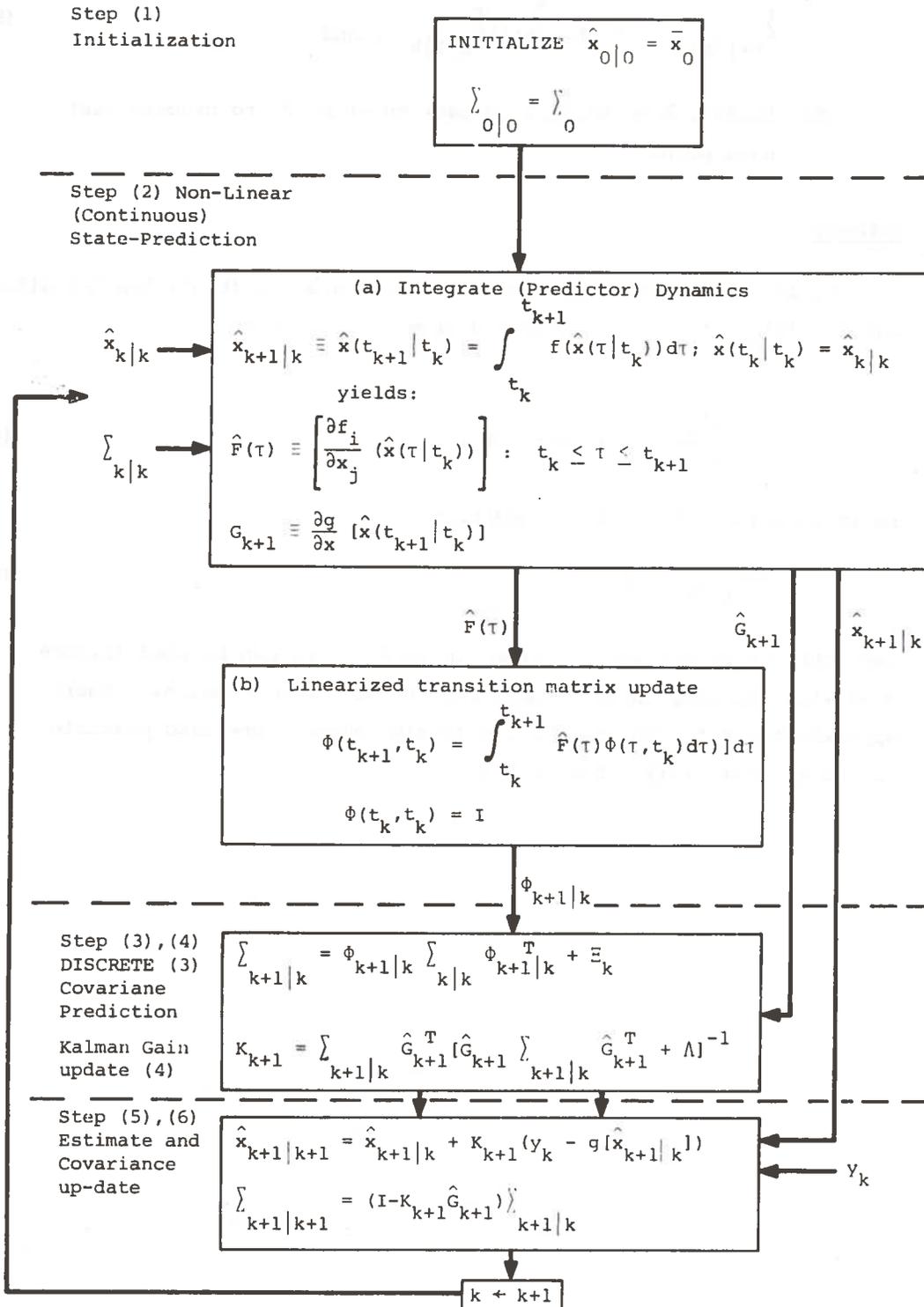
$\longleftarrow y_k$

$k \leftarrow k+1$

Figure B.1: Continuous-Discrete Extended Kalman Filter

APPENDIX C

REPORT OF INVENTIONS

This report contains no inventions or other patentable items; however, it provides advances in techniques for real time freeway traffic surveillance data processing, including estimation of speed, density and available capacity from presence detectors as detailed in Sections 3 and 4. In particular,

Using an intuitive and simple pre-processing procedure for occupancy and speed data from individual vehicles, the processed detector outputs can be modeled as noisy observations of spece mean speed and mean density.

With the pre-processed data from 1/2-mile spaced detectors and free-flowing traffic conditions, accurate estimates of mean speed and mean density can be consistently obtained.

Under non-homogeneous conditions, it is shown that (typically) a single parameter can be used to model the available capacity of a freeway section (via the fundamental diagram of traffic), and that the value of this parameter can be identified with an augmented extended Kalman filter.

Information in the time behavior or signature of the estimated capacity in response to incident conditions is shown to establish a foundation for reliable and rapid incident detection, as well as providing quantitative information for use in the traffic control system [FR3].

110 copies