Report No. DOT-TSC-RSPA-78-12

DYNAMIC STOCHASTIC CONTROL OF FREEWAY CORRIDOR SYSTEMS
Summary and Project Overview

Michael Athans
Stanley B. Gershwin
Paul K. Houpt

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Electronic Systems Laboratory
Operations Research Center
Department of Electrical Engineering and Computer Science
Cambridge MA 02139

DECEMBER 1978

FINAL REPORT

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
RESEARCH AND SPECIAL PROGRAMS ADMINISTRATION
Office of Transportation Programs Bureau
Office of Systems Engineering
Washington DC 20590

Technical Report Documentation Page

1. Report No. DOT-TSC-RSPA-78-12	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DYNAMIC STOCHASTIC CONTROL OF FREEWAY CORRIDOR SYSTEMS Summary and Project Overview		5. Report Date December 1978
		6. Performing Organization Code
7. Author(s) Michael Athans, Stanley B. Gershwin, Paul K. Houpt		8. Performing Organization Report No. DOT-TSC-RSPA-78-12
9. Performing Organization Name and Address Massachusetts Institute of Technology* Electronic Systems Laboratory Operations Research Center Department of Electrical Engineering & Computer Science Cambridge MA 02139		10. Work Unit No. (TRAIS) OS850/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 Apr. 1975 - Sep. 1977
		14. Sponsoring Agency Code

15. Supplementary Notes

*Under contract to: U.S. Department of Transportation, Reaearch and Special Programs Administration, Transportation Systems Center, Cambridge MA 02142

16. Abstract

This report provides a summary and overview of the results obtained in the completion of this Contract, "Application of Modern Control Theory to Urban Freeway Corridor Control." The focus of this activity has been to develop systematic methodological approaches to overall traffic management from both short-term (real-time) and long-term (planning) perspectives. Our approach embodies formulation and solution of interrelated mathematical problems from operations research and optimal control theory, including:

1) selection of traffic models, both static and dynamic;
2) optimal (steady-state) traffic assignment;
3) dynamic ramp-metering control algorithm design; and
4) traffic surveillance data-processing algorithm design.

Theoretical results and designs are evaluated using a variety of computer simulation tools. Practical implementation of these procedures for traffic management are addressed in all phases of the research through decomposition and decentralization techniques.

17. Key Words Freeway Corridor Control Optimal Traffic Assignment Dynamic Ramp Metering Traffic Surveillance	18. Distribution Statement DOCUMENT IS AVAILABLE TO THE PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED	74	

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

PREFACE

This summary report presents the joint efforts of faculty, staff, and students associated with the M.I.T. Electronic Systems Laboratory. The authors gratefully acknowledge contributions from Nathan Gartner and John D.C. Little of the M.I.T. Operations Research Center. Students in the Department of Electrical Engineering and Computer Science who made direct contributions to results reported herein include Pierre Dersin, Douglas P. Looze, Rafael Lopez=Lopez, William J. Mitchell, and Daniel G. Orlhac. Additional inputs were provided by Han-Ngee Tan, Samuel Chiu and Michael McIlrath. For contributions to the initial phases of this research, the authors are appreciative to Jason L. Speyer, University of Texas (formerly of the C.S. Draper Laboratory). Finally, we acknowledge with thanks the continued support, encouragement, and guidance provided by our TSC technical monitor, Diarmuid O'Mathuna, and by the administrator of Transportation Advanced Research Programs, Robert Crosby.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1. INTRODUCTION AND MOTIVATION

This report presents an overview of a research program directed at the optimal management of traffic flow in urban-freeway corridors. Detailed results of our findings are provided in the six companion reports, (FR1 to FR6), which are listed separately in the references. In this study, a freeway corridor is assumed to consist of one or more multi-lane limited-access roadways connected through entry and exit ramps to surface streets. The surface-street network may consist of parallel service roads, both signalized and unsignalized, as well as sections of a traffic grid associated with a downtown central business district (CBD).

With existing corridor networks, the increased demand resulting from vehicle-population growth has resulted in frequent occurence of congestion. Delay and energy consumption which accompany this congestion pose significant economic and social costs. In the past, the approach to relieving congestion was to expand the roadway network, at significant capital cost, and usually little impact on traffic-flow congestion. In recent years, however, the trend has been toward making more efficient use of available facilities through a combination of good traffic-engineering planning in conjunction with real-time surveillance and control, made possible in part by digital-computer technology.

The broad research objective in this study then has, as its focus, development of systematic methodologies for the management of freeway corridors in which tradeoffs in design performance can be quantified and optimized. We show in particular how the highly complex overall traffic-management problem can be simplified by a coordinated application of techniques from operations-research and modern-system theory. Mathematical programming, and optimal control and estimation theory, play central roles in our analysis and design.

There are two kinds of results presented in the six other reports covered by this summary report: (1) explicit computer algorithms which perform the required optimizations, surveillance data-processing, and decision making associated with our approach, and (2) identification of generic methods from which the algorithms in (1) result which can be applied to similar corridors. We emphasize the generality of our approach which is not tied to details of the simple corridors used for illustration purposes, in this context the methods are as important as the specific results.

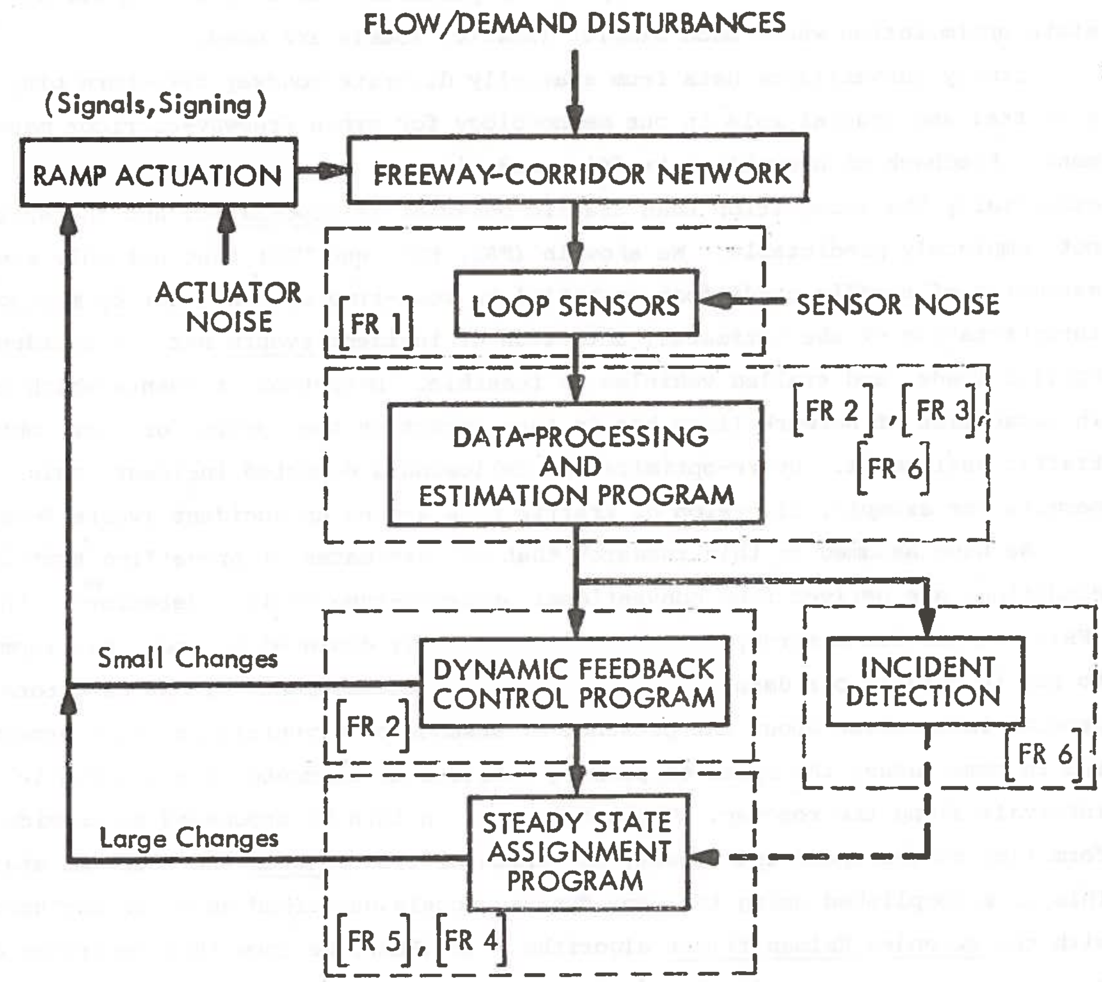

Figure 1.1 Approach to Traffic Management and Real-Time Control in Urban-Freeway Corridors

Note: Entries in brackets [ ] refer to the relevant Final Report listed in the References.

Considerable attention has therefore been directed at techniques for further decomposition and/or physical decentralization of these data-processing activities. For example in (FR4), we provide a sensitivity analysis of the optimal assignment procedure to show the effect of changes in exogeneous demand or incidents on a previously computed optimal assignment. The results provide a systematic technique to decide if re-optimization (on-line) of the network is warranted, and if so, how much of the corridor network to include (see (FR4) for details). By exploiting the natural decoupling inherent in the dynamic freeway-corridor model, strategies for decentralization of the dynamic control and surveillance algorithms are proposed in (FR2) and (FR6). Simulated performance of the suboptimal decentralized estimation and control strategies shows relatively little performance deterioration, but appears to promise substantial decreases in algorithm complexity, including the possibility of multi-parallel-processor implementation.

1.2 Report Structure

The remainder of this summary is structured as follows. In section 2 we review the model for traffic behavior used in the various design problems, and define the principal traffic-management research objectives in terms of attributes of these mathematical models. Then in section 3, we turn to the formulation and solution of the optimal static traffic-assignment problem reported in (FR5), including the closely related work on sensitivity analysis in (Fr4). Section 4 presents the main findings in traffic-surveillance algorithm development proposed in (FR3) and (FR6). Real-time dynamic ramp-metering control and its decentralized implementation (FR4) conclude the summary in section 5. Some conclusions and suggestions for further research are provided in section 6.

congestion conditions (see section 3).

Keeping track of individual vehicles in a dynamic model is an extremely complex task, although it would appear necessary in view of the above performance measures. Modeling vehicular flow on a microscopic basis requires an essentially infinite-dimensional model structure from which little if any insight into the control problem for a network can be gained. The philosophical problem from a pragmatic point of view is: How can one aggregate the dynamics of network flows in a way that preserves the essential behavior, and from what, performance measures of interest can be predicted?

A fruitful approach has been a network characterization in terms of freeway-section flows (mean-speed and density) and arterial-link queues. Most performance measures in usual engineering practice can be directly related to arterial link and freeway-flow behavior, in both steady-state and dynamic settings.

2.2 Selection of Freeway-Corridor Structure and Its Models

As discussed in section 1, models to describe both long-(or steady-state) and short-term (or dynamic) behavior are required. For both phenomena, it is appropriate to partition the corridor into discrete components. As indicated in Figure 2.1, the freeway corridor is viewed as interconnected subsystems consisting of freeway sections (including ramps) and arterial links.

The partition of the freeway into sections is dictated by the physical topography of the system, based on ramp locations, and geometric properties such as curves, grades, and other natural characteristics. Experienced observation of key weak points such as where bottlenecks occur (narrowing roads or other predictable disturbances) may also suggest a subdivision of the freeway. The fundamental assumption is that for each section, the model is spatially homogeneous in an appropriate sense. The sense of homogeneity is that model parameters are approximately constant with respect to position in each section or link. The actual parameterization used depends on whether steady-state or dynamic behavior is to be modeled.

The approach to aggregation of (microscopic) individual vehicle behavior has been to model behavior and performance in terms of vehicular flow (vehicles per unit time.) This macroscopic description is used in both steady-state and dynamic models.

2.2.1. Dynamic Fluid-Flow Models for Freeway and Long Arterial Sections

The macroscopic dynamic behavior of traffic has been extensively studied with a compressible fluid-dynamic analog. Lighthill and Whitham [8] were first to develop this analog. Mathematically, one obtains a partial differential equation in fluid molecule speed and density from a continuity of flow assumption coupled with an equation that describes the acceleration of individual vehicles as a function of local speed and density (see FR3 for details).

A relationship which parameterizes the solution of the fluid analog equations has come to be known as the fundamental diagram of traffic (Figure 2.2). This fundamental diagram has provided a convenient parameterization of the volume-speed density interaction. In addition to the density that maximizes flow volume (when it exists), qualitative stability properties and wave behavior have also been deduced using this relationship. The early applications were to single-lane flows such as on bridges or tunnels [Greenberg [3], Crowley and Greenberg [5], Herman [6], Gazis [7]]. In addition, the simpler fluid models have been verified empirically. The stability properties deduced from the fundamental diagram can be verified by analysis of the microscopic interaction of vehicles in the "car-following" literature [Chandler [9], Gazis [7], Duckstein and Unwin [10]].

Despite considerable analysis of fluid dynamic models (e.g., Newell [11], Prigogine and Herman [12]), none are very useful for control of a freeway since the equations have both spatial and temporal dependence, and hence are infinite dimensional.

A finite-dimensional approach, which was adopted in the present research, is to use an ordinary differential equation approximation to the fluid model. The technique was suggested by Payne [13]. The fundamental idea is to discretize spatially the freeway corridor sections, replacing speed and density by their spatial means (see (FR6), Section 2). A heuristic interpretation of this approach is shown in Figure 2.3. What results is a collection of coupled, ordinary non-linear differential equations in the section (spatial) mean-speed and (spatial) mean-density.

The dynamic equations for each of the N sections are:

(Space mean-speed)

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

where

x_0 = start of freeway modeled

x_j = {beginning of section j} = $x_0 + \sum_{k=1}^{j} \Delta x_k$, $1 \leq j \leq N$, and,

Δx_j = length of section j,

and (Space mean-density)

$$\frac{d\bar{\rho}_j(t)}{dt} = \frac{\bar{v}_{j-1}(t)\bar{\rho}_{j-1}(t) - \bar{v}_j(t)\bar{\rho}_j(t) + r_j(t) - w_j(t)}{\Delta x_j} \tag{2-2}$$

where $r_j(t)$ and $w_j(t)$ are on-ramp and off-ramp flows (respectively), and ν and τ are parameters which model driver reactions.

The model is a blend of macroscopic (fluid) and microscopic (car-following) theories. It exhibits most of the qualitative dynamic behavior one would expect from more complex models. Moreover, it provides an analytically tractable model for dynamic control system design; each section has two state variables (mean speed and mean density) and controls corresponding to the number of on and off ramps. The actual number of state variables required in any corridor depends on how fine a partition into sections is required to mimic observed behavior. Sections approximately 1/2 to 1-1/2 miles in length appear typical.

2.2.2 Steady-State Flow Models for Freeway and Arterial Links

The problem of modeling the steady-state distribution of macroscopic corridor traffic flow is complicated by the autonomy of individual drivers. A classical assignment problem starts by assuming a flow dependent utility function for drivers, which may be different for each link. By assuming various driver behavioral rules apply (e.g., via Wardrop's [14] principles), one can iteratively make systematic incremental changes in the flow distribution until net utility (e.g., delay) is consistent with driver preferences and link capacities. The classical assignment-problem formulation and algorithms for solution are discussed in Yagar [15], Nguyen [16], and Payne and Thompson [17]; other references are provided in section 3 of this summary report.

Flow behavior on long or unsignalized arterials could be modeled with second order non-linear differential equations similar to those in section 3.2.1. On shorter, signalized arterials in an urban CBD or on merging/exit ramps, off/on, this approach does not apply.

The complication is that most delay accrues from time spent in queue at a signal light, or waiting to merge onto the freeway. The dynamic behavior of stop-line queues on signalized ramps and arterials depends jointly on:

3. STATIC ALLOCATION OF CORRIDOR NOMINAL TRAFFIC FLOW AND SIGNAL SETTINGS

Because of the importance of predicting driver responses to decision variables selected for corridor control, we consider both user-optimized and system-optimized assignments in the steady-state analysis. Methods of traffic assignment, traffic signal setting, and modal-split analysis are combined in a set of computer-assisted optimization programs. Here, we describe travel-time functions for freeways, freeway entrance ramps, and signalized streets. Energy is parameterized by fuel consumption for cars, carpools, and buses; modal split analysis is described and integrated with assignment, so that the effect of favoring buses or carpools with exclusive use lanes (e.g., diamond lanes) can be assessed. We begin by considering some of the related system-optimized traffic-assignment procedures.

3.1 Steady-State System-Optimized Assignment

The signal synchronization problem for arterial networks was the earliest area where optimal traffic assignment has been studied. Systematic control techniques were applied. The assumption of undersaturated link flows led to the queuing theory results of Webster [21] for the isolated signalized intersection. The undersaturated flow assumption permits an average delay to be obtained as a function of signal parameters. The average delay in traveling an arterial of signalized intersections can then be related to the relative signal offsets. Offset-synchronizing strategies were obtained for the arterial by Little [22] using mixed integer programming. The extension of arterial signal synchronization to a network of more general structure was obtained by Hillier and Rothery [19]. Hillier's results were applied to an experimental control scheme in Glasgow, Scotland (Hillier [24]).

Programs have been written to tackle some of the difficult static, functional minimizations required in the synchronization approaches, including SIGOP (Traffic Research [25]) and TRANSYT (Robertson [26]). The TRANSYT algorithm used a microscopic dynamic-simulation procedure to calculate delay by running the system model and changing signal parameters in a hill-climbing procedure until delay was minimized. More recently, systematic and compuoccupancy, and (c) a greater energy consumption than cars or car pools. These concepts are quantified below.

The values of traffic signal parameters, such as cycle time and green split (the fraction of a cycle that the signal is green), have an important effect on delays on signalized links, and thus, on the assignment. The problem of <u>optimal signal setting</u> is that of calculating the parameters that minimize delay or other cost function. This requires models of traffic signal delays and an optimization algorithm. In this section, models of signal delays are presented. In the following section, methods of calculating signal parameters are discussed.

Different classes of vehicles may experience different levels of service. For example, if cars are prohibited from traveling on some links, they will probably encounter longer travel times than for buses. Travelers can be expected to respond to such discrepancies by choosing more desirable transportation modes. The calculation and prediction of this behavior are the problem of <u>modal split</u>.

3.2.1 <u>Flows and Constraints</u>

Let $\phi_{ij}^{(n)}$ be the flow, in vehicles per hour, of traffic on link i whose destination is node j, and which is in class n. In this report, n=1 refers to single-passenger cars; n=2 refers to carpools; and n=3 refers to buses.

We have <u>positivity constraints</u>; i.e.,

$$\phi_{ij}^{(n)} \geq 0; \quad \text{all } i,j,n. \tag{3.1}$$

We also have <u>conservation of flow</u> constraints, which require that the total flow into a node (including that which arrives from outside, called the origin-destination requirements) equals the total flow out of the node (including that which leaves the system).

The <u>total passenger flow rate</u> P_i on link i is tion that expresses this is

$$\tau_i = \frac{1}{E_i - \phi_i} + t_i(\phi_i), \tag{3.5}$$

where E_i is the effective capacity of link i, and is a function of ϕ_ℓ.

Signalized arterials. Two signalized arterial travel time functions have been considered in this study. One is the Webster formula [34], and is most appropriate for isolated intersections; i.e., where signalized intersections are relatively far apart. The travel time is given by $\tau_i = t_i(\phi_i) + Q_i$, where $t_i(\phi_i)$ is the traversal time, and Q_i is the queueing time which is a function of the cycle time, the saturation flow (the capacity of link i if the signal were always green), and the green split time (the fraction of time the signal is effectively green to link i). The other is the MITROP cost [35]. This travel time function has been rewritten to be compatible with the assignment algorithms, and is discussed in [33].

3.2.2.2 Energy Consumption

In this study, we also investigate the total energy consumed per hour by the vehicles on the network. Again, there are different functional forms for different kinds of links.

Freeway links. Claffey [36] found data for $G^{(n)}(v)$ the gasoline consumption per class n vehicle per mile as a function of velocity. We fitted polynomials to his tabulated results. The total fuel consumption on link i is given by

$$F_i = \sum_n \ell_i \phi_i^{(n)} G^{(n)}(v_i), \tag{3.6}$$

where $\phi_i^{(n)} = \sum_j \phi_{ij}^{(n)}$ is the total flow of class n vehicles on link i, and ℓ_i is the length of link i. When n=1 or 2 (cars or carpools), we use Claffey's results for automobiles. Claffey has no results for buses, but for n=3, we use his results for two axle, six-tire trucks.

as a whole. Let cost C be given by

$$C = W_1 \sum_i F_i + W_2 \sum_i \tau_i P_i , \tag{3.9}$$

where W_1, W_2 are constants specified by the programmer. If $W_1 = 0$, the program minimizes total travel time (person-hours per hour) in the network. If $W_2 = 0$, the program minimizes total energy consumption. If both are positive, they can be thought of as the dollar cost of fuel and time, and C is a total dollar cost per hour.

3.2.4 Assignment Principles

The assignment principle is the rule by which network attributes, which depend on flow values, influence the distribution of flow in a network traffic assignment.

Two assignment principles were formulated by Wardrop [38]. The system-optimization assumption is that traffic is distributed in a way that minimizes some criterion function. (Wardrop suggested average travel time. We consider C in equation (3.9).) Under the user optimization assumption, drivers are assumed to choose the shortest paths available. As a consequence, trip times on different paths traveled between the same pair of points tend to equalize.

3.2.5 Modal Split

In investigating mode split, we use the logit model [39]. Define the passenger demand $R_{ij}^{(n)}$ to be the rate of arrivals of passengers at node i who wish to go to node j on mode n. Note that

$$R_{ij}^{(n)} = r_{ij}^{(n)} w^{(n)} , \tag{3.10}$$

where $r_{ij}^{(n)}$ is the rate of arrival of vehicles of class n at node i whose drivers wish to go to node j, and $w^{(n)}$ is the average passenger occupancy of vehicles of class n.

Define the total passenger demand, R_{ij} from i to j as equalities, the Kuhn-Tucker conditions [42]. Suppose instead, we satisfy a different set of nonlinear equations and inequalities; we find flows to equate travel times on paths defined by the extremals already generated. If this procedure converges, it creates flows that equate times on paths utilized between origins and destinations - the user optimization solution.

3.3.2 Traffic-Signal Setting

Single intersection. In case of a single (isolated) intersection, the only variables to be determined are the cycle time and the green splits. The most common method for setting signals at isolated intersections is due to Webster [34]. We treat cycle time as a fixed constant to simplfy computations. Cycle time is much less sensitive to changes in flows than green splits.

Interconnected intersections. When two or more signal-controlled intersections are in close proximity, the simple Webster formulas are insufficient. The signals then form a network which has to be synchronized (i.e., a common cycle time has to be found) and coordinated (relative phases, or offsets, must be determined). The method used is MITROP (Mixed-Integer Traffic Optimization Program) which optimizes simultaneously all the network control variables: offsets, splits, and cycle time. MITROP is formulated as a mathematical program and provides a globally optimal solution [35].

3.3.3 Overall Procedures

Assignment and signal settings. Figure 3.1 illustrates the overall program flow when an assignment and a set of signal settings are calculated together. Signal settings are guessed (1) and the corresponding assignments calculated (2). Given the resulting flows, a new set of signal settings is found (3). A stopping criterion is tested (4), and if the procedure does not terminate, a new assignment is calculated (2) and the procedure repeats. We have no proof of convergence, but this has converged in the examples considered.

Exclusive ramps and lanes. Networks can have potentially exclusive lanes. For example, in a network in [33] the leftmost lane in both directions of a freeway is treated separately from the two right lanes. The system optimized assignment may forbid a class of vehicles from the exclusive lane. It is the programmer's option to make the lane potentially exclusive by separating it in this way. The program decides if it actually should be.

If the assignment is user optimized, exclusive lanes can be separated as above. Here, however, the programmer must specify which lanes are to be exclusive, and which vehicles are excluded. This is because reserving a lane is for the good of the system, but the assignment is chosen by the users.

Modal split. Figure 3.2 contains the flow chart for integrating mode split with assignments and signal setting calculations. The mode split - that is, the entire origin destination demand matrix broken down by mode - is guessed (1). The corresponding assignment and green splits are then calculated (2) as in Figure 3.1. A new requirement matrix is calculated (3) from (3.10 to 3.13). A stopping criterion is tested, and if the procedure is not terminated, it proceeds to step (2). This is heuristic, as is the assignment-signal setting procedure, but it has performed satisfactorily.

3.4 Applications Example

Several optimization studies were performed on an artificial candidate network to compute assignments with and without signal optimization and modal split options being exercised. Detailed results are described in (FR5). For example, a small freeway corridor network with 56 links and 28 nodes was studied. The left lanes in both directions were isolated in the manner described above. Results of some of the assignment and signal setting cases are displayed in Table 3.1. The solution to the energy optimization problem included diamond lanes in both directions, whereas the travel-time optimizations did not. Nonetheless, very little difference in energy consumption resulted from the diamond lane in user optimized assignment. Other numerical experiments were performed in which the logit model was added to the assignment and signal setting techniques to explore modal split. User optimized assignments were compared with and without diamond lanes. Results from one typical run are shown in Table 3.2. It is clear that since buses and carpools are favored in the diamond lane case, some drivers switch to these modes.

TABLE 3.1 SAMPLE NUMERICAL COMPARISONS

Run	Energy $\sum_i F_i$	Travel Time $\sum_i \tau_i P_i$
System Optimization -- Energy	762.0	914.5
System Optimization -- Travel Time	787.6	877.9
User Optimization - - Without Diamond Lanes	792.3	908.9
User Optimization - - With Diamond Lanes	791.5	891.5

TABLE 3.2 MODAL-SPLIT RESULTS

Characteristic	Without Exclusive Lanes	With Exclusive Lanes
Energy Cost (gallons/hour)	646.6	598.9
Travel Time Cost (passenger-hours/hour)	540.2	496.1
Total Vehicle Flow (car equivalents)	11,246.5	10,036.1

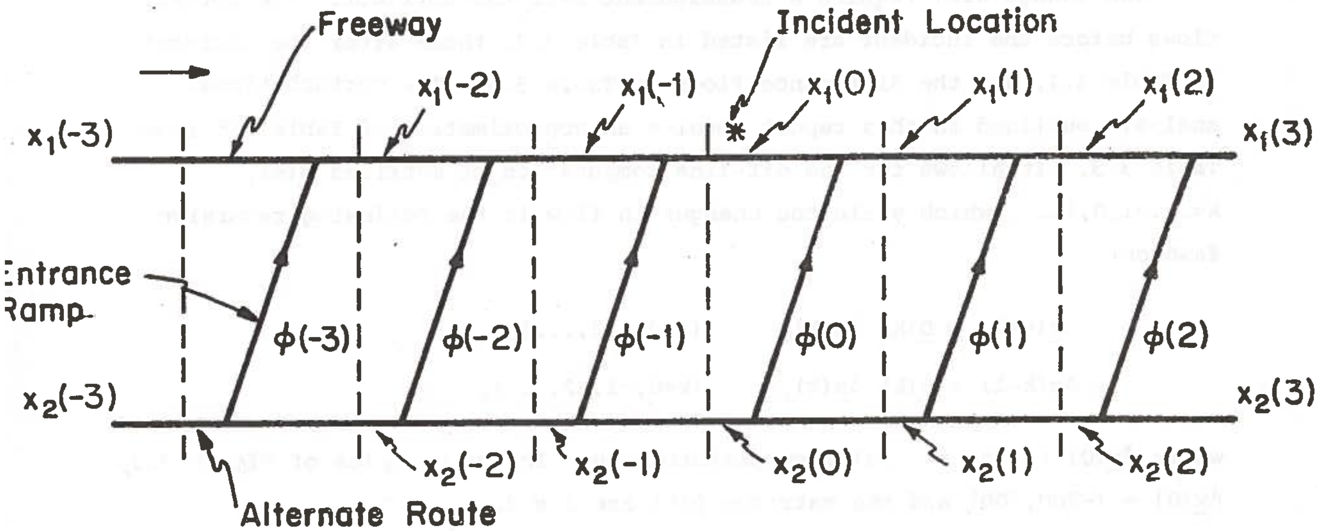

Figure 3.3. Application of Perturbation Analysis to Rerouting Traffic Around an Incident

TABLE 3.3 FLOWS BEFORE INCIDENT

k	-4	-3	-2	-1	0	1	2	3
$x_1(k)$	3500	4000	4500	5000	5200	5400	5600	5700
$x_2(k)$	3000	2500	2000	1500	1300	1100	900	800
$\phi(k)$		500	500	500	200	200	200	100

TABLE 3.4 FLOWS AFTER INCIDENT

k	-4	-3	-2	-1	0	1	2	3
$x_1(k)$	3320	3740	4140	4500	4500	4900	5240	5440
$x_2(k)$	3180	2760	2360	2000	2000	1600	1260	1060
$\phi(k)$		420	400	360	0	400	340	200

TABLE 3.5 CHANGES OF FLOWS

k	-4	-3	-2	-1	0	1	2	3
$\delta x_1(k)$	-180	-260	-360	-500	-700	-500	-360	-260
$\delta x_2(k)$	+180	+260	+360	+500	+700	+500	+360	+260
$\delta\phi(k)$		-80	-100	-140	-200	+200	+140	+100

c. Capacity constraints are not dealt with explicitly, but implicitly through penalty terms in the cost function.

d. It is assumed that the network as a whole has the required capacity to accommodate traffic. In real traffic systems, perturbations often are not confined to a small region. Before reaching steady state, they may grow to cover most of the corridor. In the context of this report, we would then conclude that the network does not have the capacity to carry the traffic, either for lack of total capacity of freeways or arterials or on account of unsatisfactory lateral access. Thus, in that case (d), the steady-state assumptions are violated, and static-optimization techniques, including the present perturbation analysis, are not applicable.

The above assumptions have been made to formulate the problem of traffic perturbation analysis, and have enabled us to make stronger quantitative statements than otherwise. It is important to observe that the resulting mathematical formulation remains physically meaningful however.

parameters, a fairly standard approach of augmenting the unknown parameters to the estimated states is proposed and evaluated with the extended Kalman filter algorithm.

Surveillance algorithms designed are evaluated in a detailed individual vehicle, microscopic vehicle simulation which has been adapted from literature sources. Features not modeled in the macroscopic differential equation model are included in this microsimulation which include: (i) stochastic variation among driver types and vehicle performance capabilities; and (ii) explicit multi-lane passing and lane-changing.

4.2 Background

Observation of prevailing traffic conditions is a pivotal element in responsive (real-time) traffic management. A broad spectrum of results is available for obtaining on-line estimates of macroscopic traffic variables such as mean speed and density. Application of the surveillance information to control systems design is provided in the context of real-time control.

A practical constraint imposed on algorithm development has been the use of conventionally configured loop detectors [FR1] as the source of surveillance data. Such sensors are typically located at spacing of 1/2 mile on limited access roadways. A detailed discussion of typical realizations of such detectors with their properties is available in Nahi [43], and our own study in [FR1].

The data provided by loop detectors consist of a series of ones and zeroes corresponding to the presence (1) or absence (0) of a vehicle in its proximity when interrogated by the surveillance computer. These data are very often averaged for 20 to 30 seconds, and converted to a dimensionless number between 0 and 100 called occupancy (percent) which is proportional to density (veh/mi) under homogeneous-flow conditions [43]. By appropriately interpreting the detector outputs over the averaging interval, the corresponding vehicular volume (veh/hr) can also be observed.

parameters in the model can be identified from observations of mean speed and 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 of available roadway capacity can be derived. Moreover, we show how qualitative information such as the presence of incident conditions can be inferred from such estimates. We believe such information to be central to the overall approach to efficient and systematic real-time traffic management.

4.3 Approach

We have developed a version of the so-called extended Kalman filter algorithms to obtain estimates of the spatial mean speed and density variables defined in Section 2. The state variable for this filter thus becomes the 2N-vector

$$\underline{x}(t) = (\underline{\rho}_1(t), \underline{v}_i(t); \ldots; \underline{\rho}_N(t), \underline{v}_N(t)),$$

where N is the number of freeway sections. Observations of the "state" $\underline{x}(t)$ are made available at discrete time instants:

$$t = t_k = k\Delta T, \qquad k=1,2,\ldots.$$

Average occupancy during (t_k, t_{k+1}) is processed as a measurement of density. Using formulas derived in (FR6), we model the observation on each link (i) as:

$$y_i^\rho(t_k) \equiv \bar{\rho}_i(t_k) + \lambda_i^\rho(t_k), \tag{4-1}$$

where $\lambda_i^\rho(t_k)$ models the error in the occupancy-to-density transformation, obtained from

$$y_i^\rho(t_k) = c \quad OCC(t_k, \Delta T) \quad (\text{veh/mi}), \tag{4-2}$$

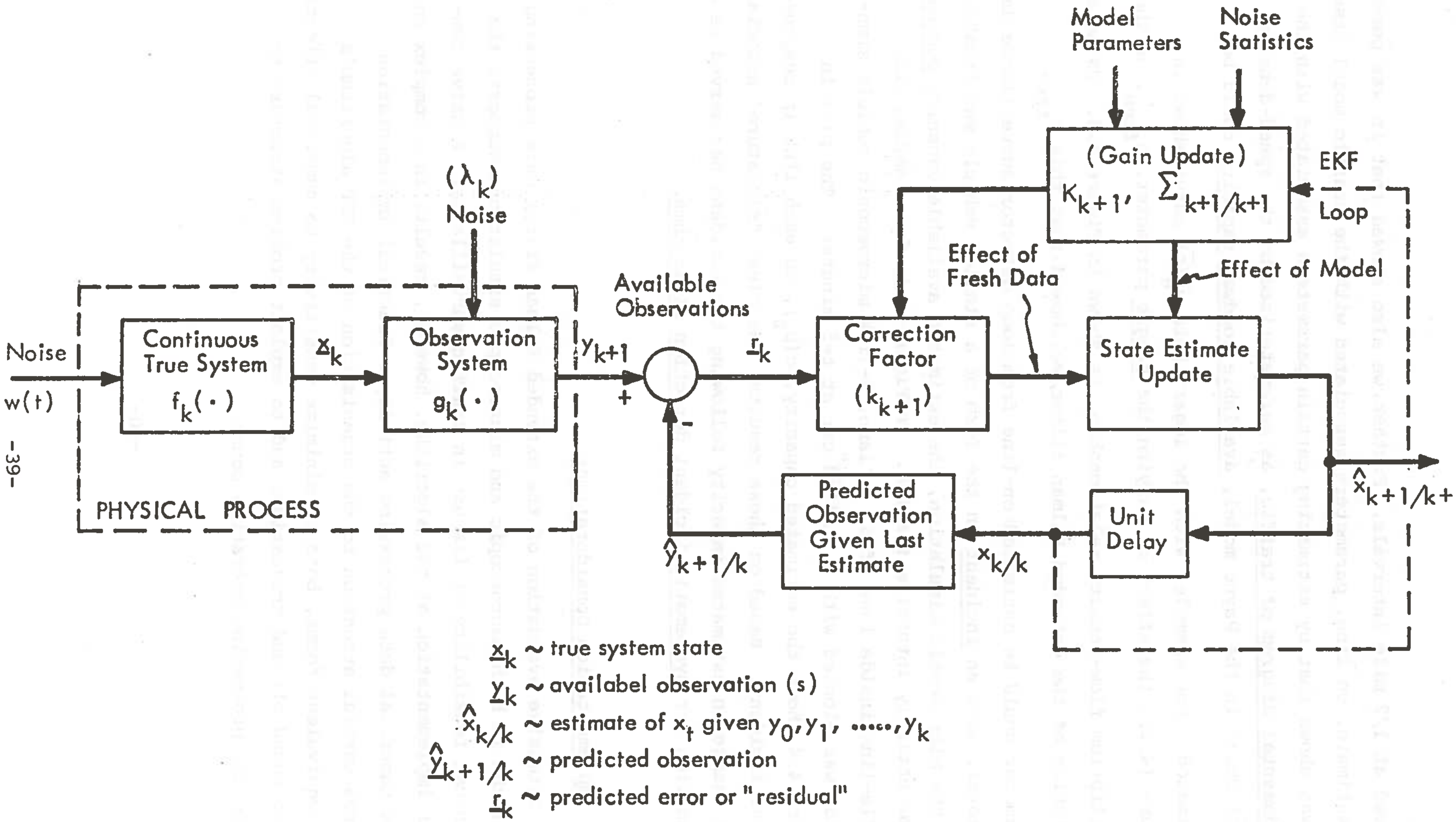

Figure 4.1 Continuous-Discrete Extended Kalman Filter Algorithm

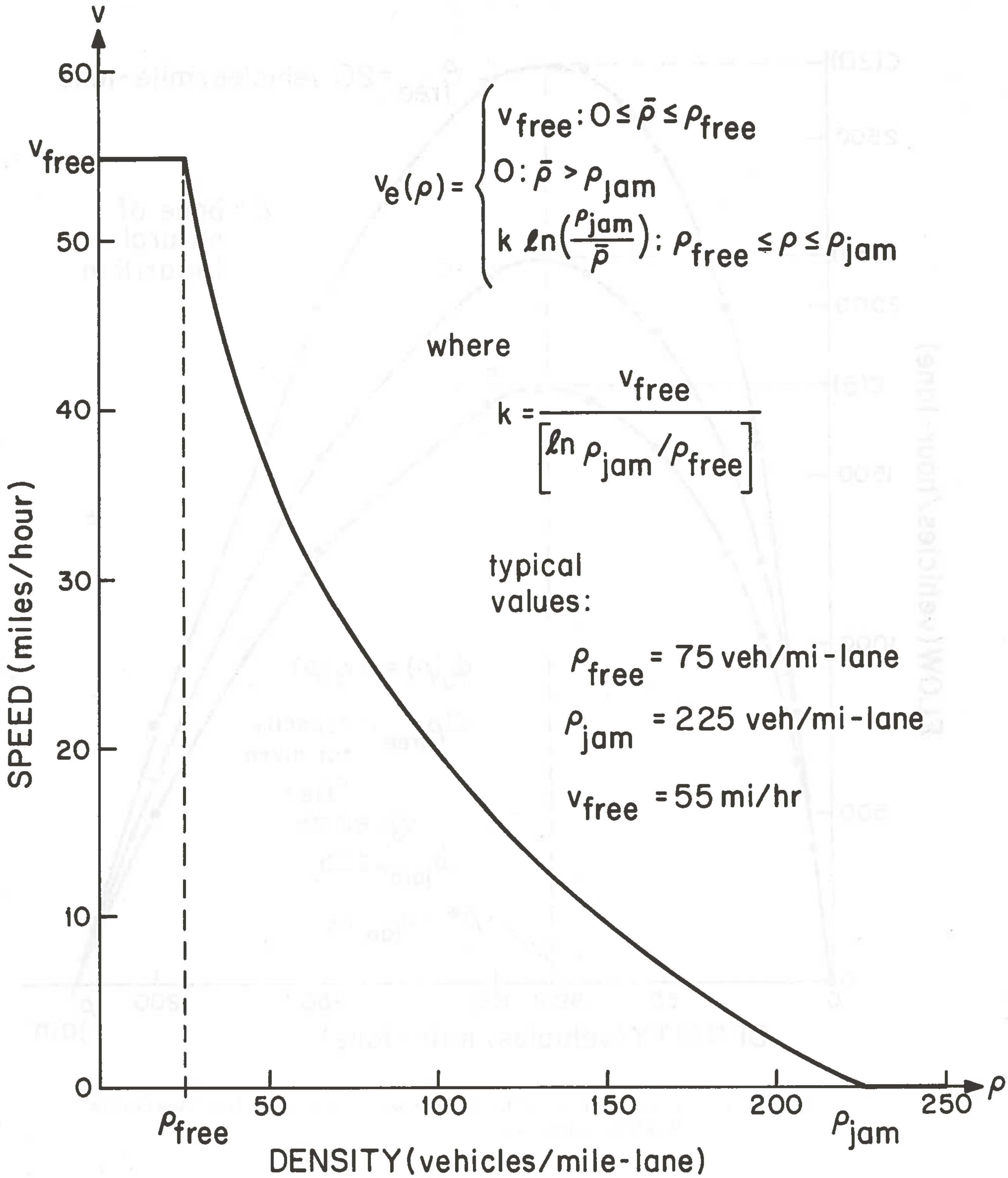

Figure 4.2 Logarithmic Equilibrium -- Speed-Density Relationship

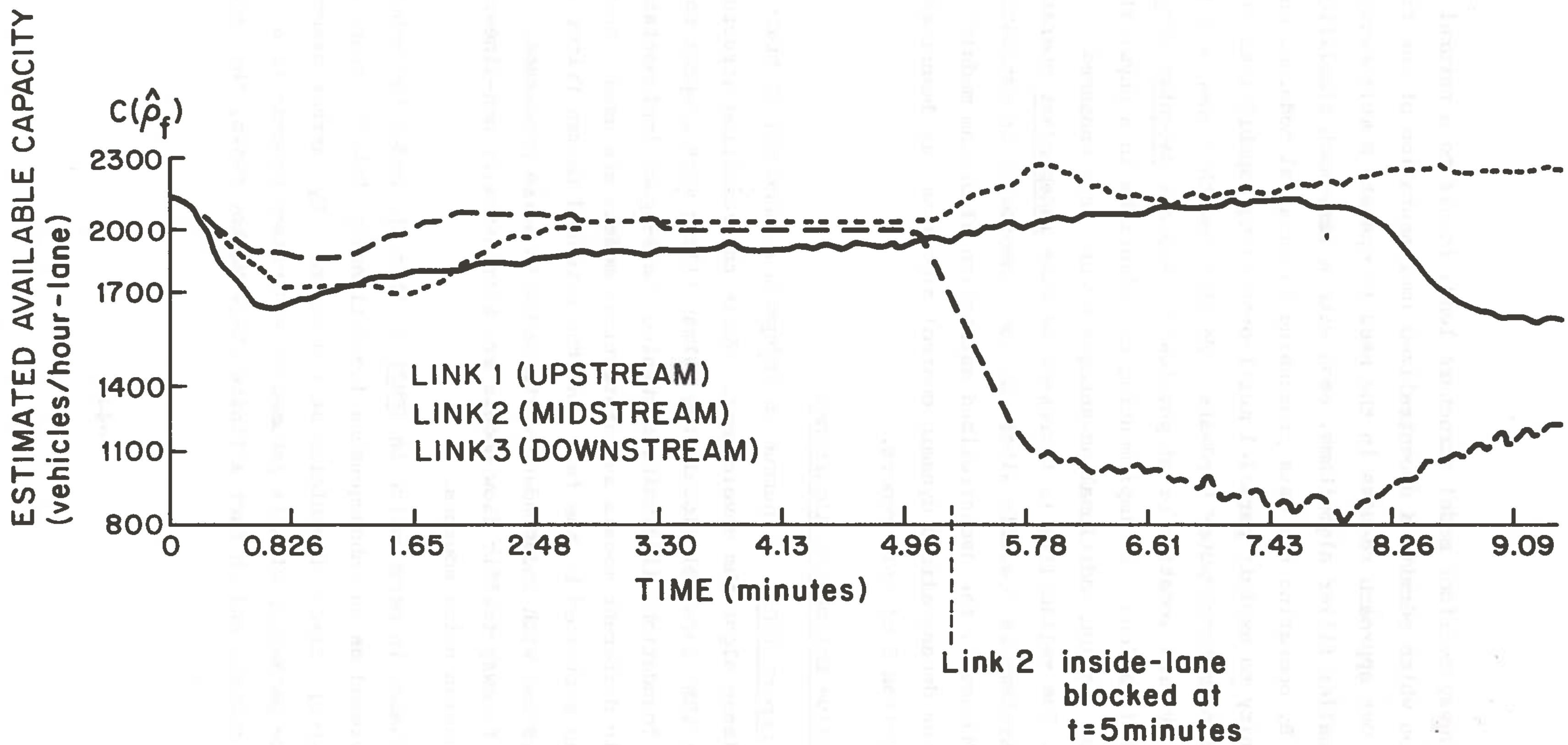

Figure 4.4 Extended Kalman Filter Capacity Estimate

traffic model is viewed as what are called "double-stochastic" processes (see e.g., Snyder [49].) The advantage of modeling vehicular traffic flow in this way has its payoff in the resulting filter algorithm (Segall [53]). First of all, the input to the filter is the presence detector output essentially without pre-processing since vehicle interarrival times are the characteristic of the traffic process observed. Averages of occupancy or harmonic means of vehicle speeds are not computed. Second, and more significant, is the fact that the model describing the rate parameter stochastic process can be non-linear without compromising optimality of the resulting filter.

The disadvantage of this approach is that the underlying rate parameter model must be finite state. And obtaining a finite-state approximation which on the one hand is sufficiently accurate and on the other does not have too many states is a non-trivial task, the issues and approach for which are explored in (FR3). With the modeling question aside, the resulting filter equations turn out to be very simple to implement for on-line processing (see (FR3) for details). We strongly recommend further model verification and development with the finite-state structure. In particular, the finite-state modeling and associated filter development appear to hold considerable promise for application to signalized arterial streets.

4.7 Summary of Main Contributions in Surveillance Algorithm Development

The principal contributions of this research are to the design and evaluation of an extended Kalman filter algorithm which uses conventional occupancy and speed data from spatially discrete presence type loop detectors. Specifically,

a. 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 space mean speed and density.

5. DYNAMIC FLOW REGULATION

The dynamic flow control algorithm receives as input the desired operating conditions derived from the steady-state assignment algorithm (Section 3 of this report and [FR3]). Feedback control is based on the linearized Payne dynamic model [13], and assuming quadratic integral performance index. The dynamic control uses estimates derived from traffic surveillance data to compute changes in nominal freeway ramp volumes or diversion fractions so as to track the desired operating speeds and densities.

5.1 Previous Results on Flow Regulation

The dynamic behavior of freeway flow has received considerable recent attention in particular with respect to responding to disturbances from the nominal flow conditions desired. Adjustment of metering rates to balance the queues on entrance ramps has received attention from Yuan and Kreer [50] and Payne, Meisel, and Teener [51]. An approach to the use of local ramp controls in the vicinity of a bottleneck or accident where vehicular flow has begun to queue has been suggested by Shaw [52,54] and McNeil [55].

A more global dynamic theory of perturbational regulation using modern control theory techniques has been proposed by Isaksen and Payne [56], and is discussed further in [FR 2]. Some sub-optimal deterministic simplifications to the regulator problem are considered by Isaksen [57] and more recently by Isaksen and Payne [4].

5.2 Method of Dynamic Control-System Design

The non-linear ordinary differential equation model for freeway mean-speed and density of Payne and Isaksen is linearized about the allocated operating point. Then a standard Linear Quadratic Gaussian (LQG) optimal-control problem is formulated (See [FR2], Section 2; and Athans [18]).

To facilitate the LQG regulator design, relative state and control weights in the integral cost functional must be specified. Following the procedure outlined in [FR 2], Section 4, the weights are chosen such that during periods of maximum expected deviations from nominal assigned flows, the fraction of flow detoured to or from freeway sections just reaches a pre-specified thresh-

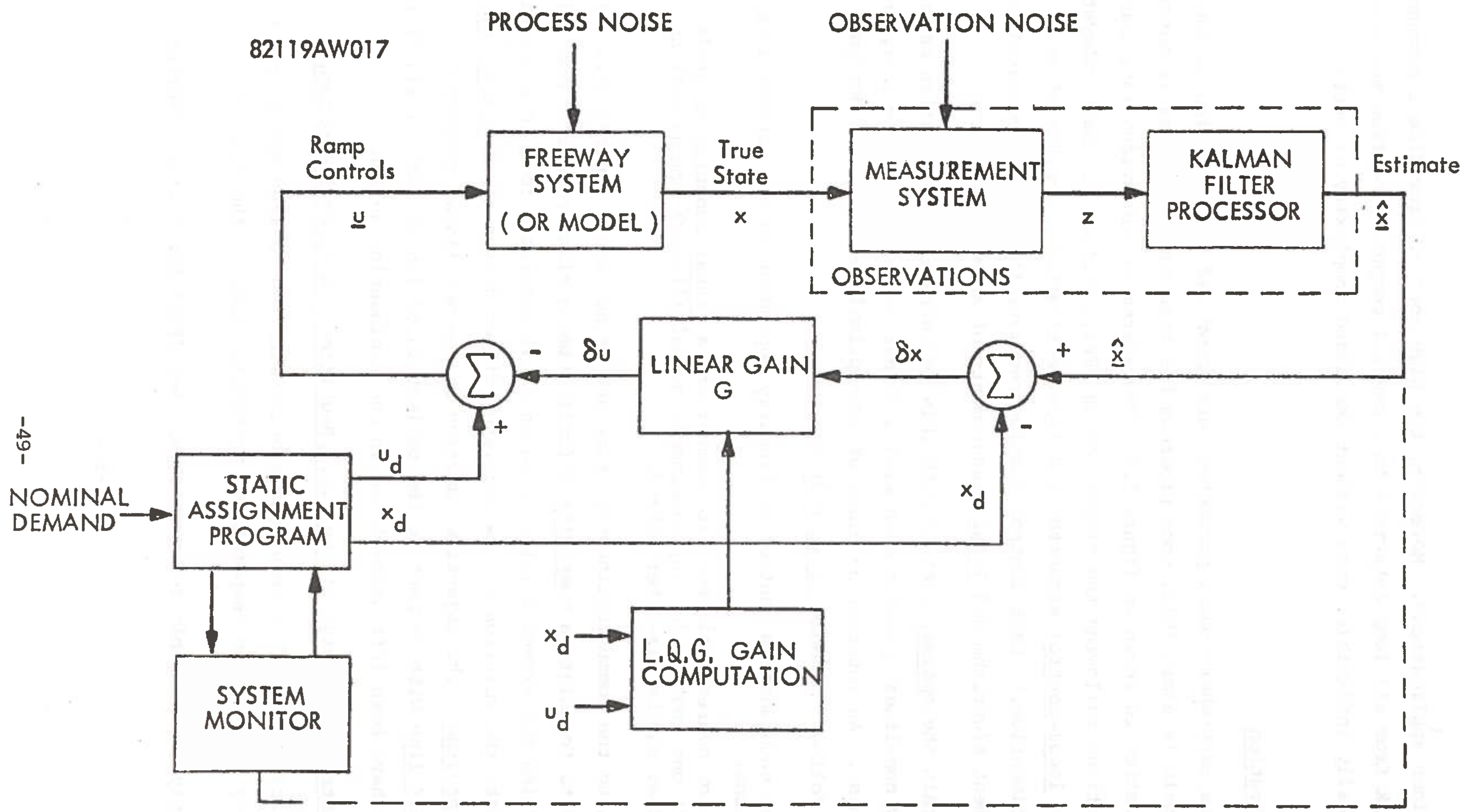

Figure 5.1 Linear Feedback Control in Overall System

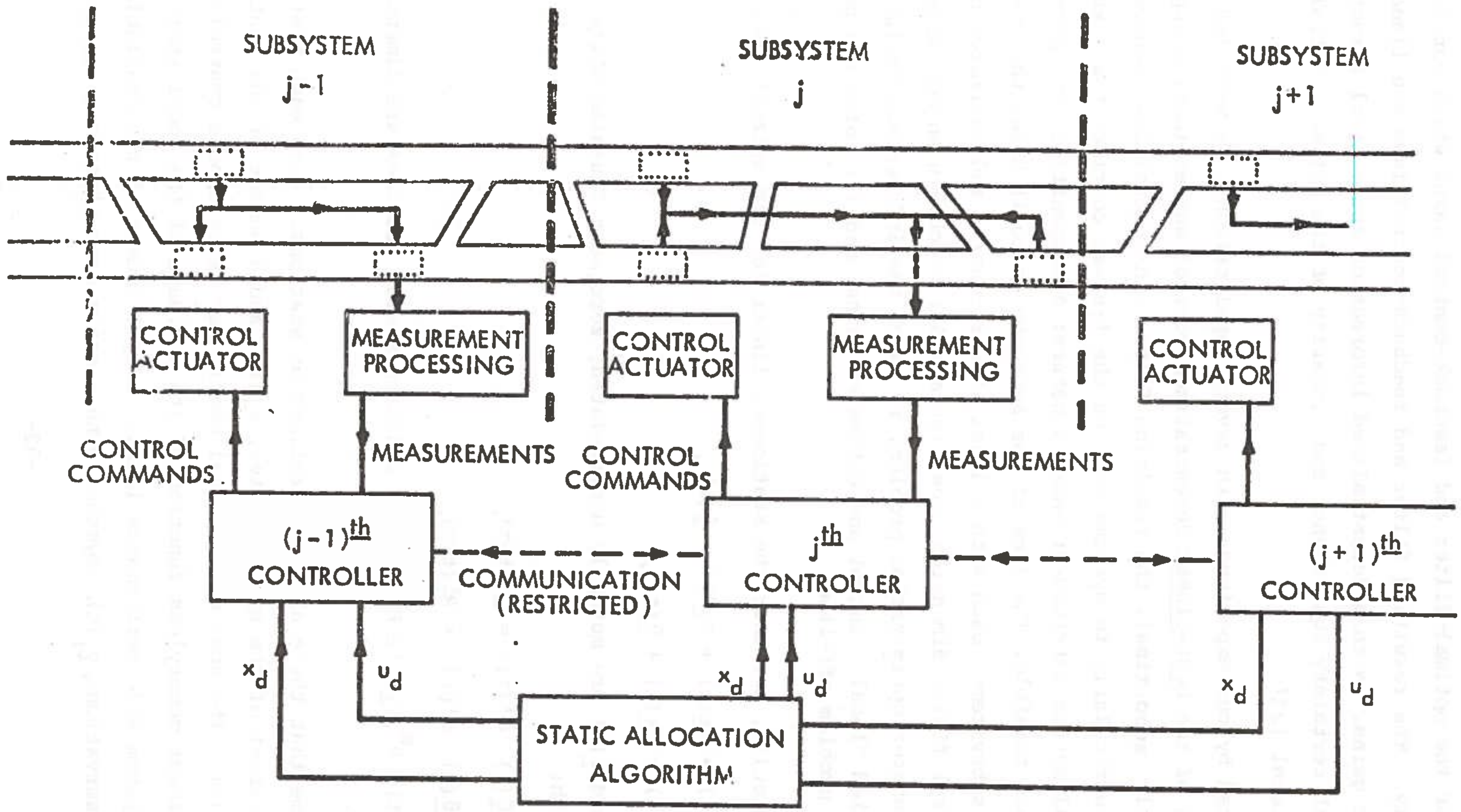

Figure 5.2 Structure of Decentralized Controller

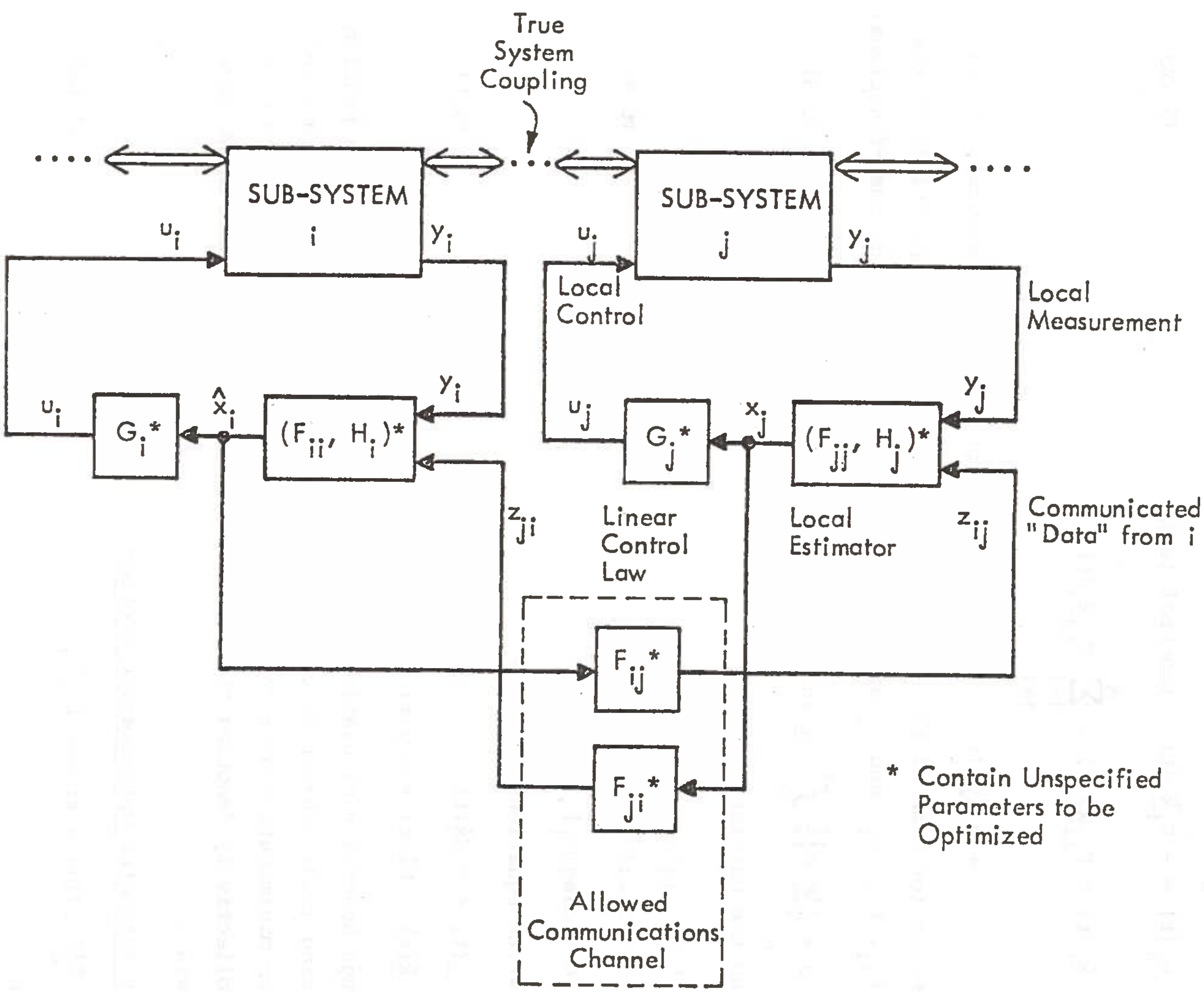

Figure 5.3 Linear (Suboptimal) Decentralized Estimation and Control Scheme

$$S(\alpha) = \begin{bmatrix} S_1 & 0 \\ 0 & G(\alpha)S_2G(\alpha) \end{bmatrix},$$

$$A(\alpha) = \begin{bmatrix} A & -BG(\alpha) \\ H(\alpha)C & F(\alpha)-BG(\alpha)-H(\alpha)C \end{bmatrix},$$

$$L(\alpha) = \begin{bmatrix} I & 0 \\ 0 & H(\alpha) \end{bmatrix},$$

$$\Xi = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

Proof of Equivalence: See [FR2,23]

The design of the suboptimal decentralized L.Q.G. regulator is now reduced to a nontrivial mathematical programming problem whose numerical solution is considered in FR2 and [23]. Some important observations are appropriate.

1) The really difficult part of the problem is obtaining the decomposition leading to equations (5.2) as in figure 5.3. While it will be desirable to have general quidelines for this decomposition, no such theory is presently available. Insight from particular applications such as the freeway problem together with engineering judgment and heuristics will be necessary for a successful decentralization strategy.

2) Since control and estimator gains are selected simultaneously subject to the constraints of the information pattern, there is no separation principle applicable in the decentralization case. On the other hand, the parameter-optimization technique can be applied to the centralized (classical) case, and the parameter, α, will contain the usual filter and control gains. The solution to (5.6a) subject to (5.6b) will then be identical to the (unique) centralized solution obtained via the usual Riccati equations.

5.4.3 Summary of Main Results

The problem of designing a decentralized, suboptimal L.Q.G. regulator was approached using an equivalent static mathematical programming formulation. Independent variables of optimization in the non-linear program include the collection of all feedback control and Kalman filter gains admissible under a prespecified information pattern. Numerical solution for the undetermined gains was obtained using a modified Davidon-Fletcher-Powell descent algorithm in which a

6. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Historically, traffic engineers, managers, and planners have had as their goals the delivery of safe, convenient, and reliable transportation service to travelers. These goals have been accomplished through the installation of traffic-control devices and by the construction of new roadways.

More recently, the economic and political climate has changed, and this has required a re-evaluation of priorities, and a re-examination of the means by which goals are accomplished. The drastic increases in energy costs have made the minimization of energy consumption one of society's more important current objectives. Also, there is now great public resistance to the construction of new roadways. Consequently, it is now crucial for transportation professionals to achieve the best possible service, while reducing the consumption of fuel and by making the best possible use of existing facilities.

The options available to traffic engineers, managers, and planners include encouraging the use of mass transit and installing sophisticated, computer-oriented automatic controls which are responsive to the dynamic and random nature of traffic. Among available options, there is always an optimization problem: how to make the best possible use of limited, often expensive, resources. In vehicular transportation, the resources include the roadway network, fuel, and control devices.

In quantifying this problem so that options can be explored systematically, the analyst is required to provide tools in the form of mathematical models and computer algorithms which can perform the tradeoffs or on-line policy decisions quickly. Our reported and proposed work encompasses the modeling of traffic behavior in urban areas over both long-and short-time horizons, together with the development of tools for mathematical optimization of traffic behavior with respect to average delay and energy criteria. Modern control, estimation, and optimization theory play central and crucial roles in performing the required analysis in a systematic fashion.

Future efforts must be directed at a continuation of the modeling and optimization of vehicular-traffic flow in urban grids. Our approach considers two broad areas which are characterized by the time horizons examined, and the level of detail in which traffic behavior is modeled.

control is expected as a consequence. To reduce some of the complexity in the dynamic control analysis we recommend attention to sub-system geometries such as merges and weaving sections which are typically critical network sections.

The results we have obtained to date, we believe, form the foundation for systematic methodologies for urban-freeway, and indeed, general urban roadway networks. Considerable attention has been directed at implementation and theoretical questions associated with optimal use of urban roadway resources. However, there remain significant non-trivial problems, both theoretical and practical, which have been identified during this research that limit the short-term applicability to the practicing engineer. In view of the potential long term payoffs in energy and traffic-congestion improvements, we believe that future investment in such fundamental-research studies will ultimately yield practical tools and data processing methods which are cost effective and attractive to planners and engineers.

[16] S. Nguyen, "An Algorithm for the Traffic Assignment Problem," Transp. Science, Vol. 8, No. 3, pp. 203-216, August 1974.

[17] N.J. Payne and W.A. Thompson, "Allocation of RAmp Metering Volumes to Optimize Corridor Performance," IEEE Trans. Auto. Control, Vol. AC-19, No. 3, pp. 177-186, June 1974.

[18] M. Athans, "The Role and Use of the Stochastic LQG Problem in Control System Design," IEEE Trans. Auto Control, Vol. AC-16, No. 6, pp. 529-552 , December 1971.

[19] J.A. Hillier and R. Rothery, "The Synchronization of Traffic Signals for Minimal Delay," Transportation Sciences, Vol. 1, 81-94 , 1972.

[20] N. Gartner, J.D.C. Little, and H. Gabbay, "Optimization of Traffic Signal Settings in Networks by Mixed-Integer Linear Programming," MIT Operations Research Center Technical Report No. 91, March 1974.

[21] F.V. Webster, "Traffic Signal Settings," Road Research Laboratory, Tech. Report No. 39, London, Her Majesty's Stationery Office, 1958.

[22] J.D.C. Little, "The synchronization of Traffic Signals by Mixed-Integer Linear Programming." Operations Research 14(4) pp. 568-594, 1977.

[23] D.P. Looze, P.K. Houpt, N.R. Sandell, and M. Athans, "On Decentralized Estimation and Control with Application to Freeway Ramp Metering," IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, pp. 268-275, April 1978.

[24] J.A. Hillier, "Appendix to Glasgow's Experiment in Area Traffic Control." Traffic Engineering and Control, Vol. 7, No. 9, 569-571, 1971.

[25] Traffic Research Corporation, "SIGOP: Traffic Signal Optimization Program," Report prepared for U.S. Bureau of Public Roads under contract CPR-11-2862, (PB 173 738) September 1966.

[26] D.I. Robertson, " 'TRANSYT' Method for Area Traffic Control," Traffic Engineering and Control, 11 pp. 276-281, October 1969.

[27] N. Gartner, "Optimal Synchronization of Traffic Signal Networks by Dynamic Programming." Traffic Flow and Transportation (G.F. Newell, Editor) American Elsevier Publishing Compl. New York, 281-295 , 1970.

[28] J.A. Wattleworth and D.S. Berry, "Peak Period Control of a Freeway System -- Some Theoretical Investigations," Highway Research Record 89, Highway Research Board, Washington, pp. 1-25, 1965.

[29] C.F. Wang, "On a Ramp-Flow Assignment Problem," Transporation Science 6, pp. 114-130, May 1972.

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

[45] D.C. Gazis and M.W. Szeto, "Design of Density Measuring Systems for Roadways," P. 44-52,.Highway Research Board Record, No. 388, 1972.

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

[47] B. Mikhalkin, "Estimation of Roadway Behavior Using Occupancy Detectors", Ph.D. Dissertation, Dept. of Freeway Operations, Los Angeles, CA., 1970.

[48] M.S. Grewal and H.J. Payne, "Identification of Paramenters in a Freeway Traffic Model," IEEE Trans. on Systems, Man Cybernetics, Vol. SMC-6, No. 3, pp. 176-185, March 1976.

[49] D.L. Snyder, Random Point Processes, Academic Press, New York, 1976.

[50] L.S. Yuan and J.B. Kreer, "Adjustment of Freeway Ramp Metering Rates to Balance Entrance RAmp Queues," Transportation Research, pp. 127-133, June 1971.

[51] H.J. Payne, W.S. Meisel, and M.D. Teener, "Ramp Control to Relieve Freeway Congestion Due to Traffic Disturbances." Highway Research 469, Highway Research Board, Washington, pp. 52-64, 1973.

[52] L. Shaw, "On Traffic Jam Queues," Transporation Research, Vol. 4, pp. 281-292, October 1970.

[53] A. Segall and T. Kailath, "The Modeling of Randomly Modulated Jump Processes," IEEE Trans. Information Theory, Vol. IT-21, pp. 64-71, March, 1975.

[54] L. Shaw, "On Optimal Ramp Control of Traffic Jam Queues," IEEE Transactions on Automatic Control, AC-17, No. 5, October 1972.

[55] D.R. McNeil, "Growth and Dissipation of a Traffic Jam," Transp. Res., Vol. 3, pp. 115-121, 1969.

[56] L. Isaksen, and H.J. Payne, "Regulation of Freeway Traffic", 1972 JACC Proceedings, Paper No. 4-5, p. 63.

[57] L. Isaksen, "Suboptimal Control of Large Scale Systems with Application to Freeway Regulation," Ph.D. Thesis, Elect. Engineering Dept., University of Southern California, 1971.

APPENDIX REPORT OF INVENTIONS AND NEW TECHNOLOGY

While there have been no inventions as a result of this work, certain theoretical advances and technological innovations have been made and are detailed in the other reports of this series (see p. 64). These achievements include respectively:

1. An optimization procedure for traffic control in a freeway corridor network involving assignment, signal setting and effects of modal split (DOT-TSC-RSPA-78-10).

2. A detailed analysis of the propagation of traffic-flow perturbatious including shock waves (DOT-TSC-OST-77-62).

3. The theoretical basis for systematic decentralized dynamic control and estimation (DOT-TSC-OST-77-10).

4. A comprehensive evaluation of the various types of available vehicle detectors (DOT-TSC-OST-77- 9).

5. A systematic scheme for identification and estimation of traffic variables using techniques based on the extended Kalman filter (DOT-TSC-RSPA-78-18).

6. An alternative procedure for estimation of traffic variables based on Markov point-process theory (DOT-TSC-RSPA-78-9).

220 copies