

**TRAFFIC CONTROL SYSTEMS ANALYSIS BY
MEANS OF DYNAMIC STATE AND INPUT-OUTPUT MODELS**

H. Strobel

**RR-77-12
June 1977**

Research Reports provide the formal record of research conducted by the International Institute for Applied Systems Analysis. They are carefully reviewed before publication and represent, in the Institute's best judgment, competent scientific work. Views or opinions expressed therein, however, do not necessarily reflect those of the National Member Organizations supporting the Institute or of the Institute itself.

**International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria**

Preface

The issues addressed by the Human Settlements and Services Area at IIASA range from short-term planning problems, such as the real-time management of urban traffic and emergency services, to long-term problems, such as the formulation and implementation of national policies of urban growth. This paper focuses on the control of urban traffic systems. It reports on research conducted by the Automatic Traffic Control Study at IIASA. In the first part, Professor Strobel presents a state-of-the-art review of urban traffic control models. The potential practical applications of the models and the estimation of their parameters are then illustrated by means of a case study: the North-South-Connection in Dresden, German Democratic Republic.

Frans Willekens
April 1977



Summary

This report deals with the following questions: which dynamic models and which advanced methods of identification theory are used or could be used in urban traffic control systems analysis; and which problems are still unsolved. A survey of basic approaches is presented, and particular attention is devoted to the state and input-output models. Their significance for the analysis of traffic control systems is then discussed. Finally, the paper reports on real-time identification methods for the determination of input-output model parameters. The implementation of the methodology is illustrated by a case study.

Contents

MOTIVATION	1
PART I: DYNAMIC TRAFFIC FLOW MODELS: A SURVEY	
STATE MODELS	2
Microscopic Traffic State Models	4
Macroscopic Traffic State Models	8
Aggregated Macroscopic Traffic State Models	11
INPUT-OUTPUT MODELS	16
Microscopic Input-Output Models	16
Macroscopic Input-Output Models	17
THE ROLE OF STATE AND INPUT-OUTPUT MODELS IN THE ANALYSIS OF TRAFFIC CONTROL SYSTEMS	23
Traffic Control Concepts	23
Models Needed for Route Control Analysis	23
Models Needed for Flow Control Analysis	25
Models Needed for Vehicle Movement Control Analysis	28
Conclusions	29
PART II: IDENTIFICATION OF DYNAMIC INPUT-OUTPUT MODEL PARAMETERS: A CASE STUDY	
THE MODELS AND IDENTIFICATION METHODS STUDIED	31
The Aim of the Case Study	31
Introduction of Stationary Input and Output Signals and of Modified Model Structures	31
The Explicit Methods Used	35
The Implicit (Recursive, Adaptive) Methods Used	37

SIMULATION STUDIES	39
Stationary Traffic Considerations	39
Nonstationary Traffic Conditions	49
ANALYSIS OF A REAL TRAFFIC PROCESS	52
The Process Studied	52
The Results Obtained	53
REFERENCES	55

Traffic Control Systems Analysis by Means of
Dynamic State and Input-Output Models

MOTIVATION*

A general systems analysis approach to transportation involves three levels:

- transportation systems planning (from a socio-economic point of view);
- operational planning (scheduling, routing); and
- control and guidance.

For an analysis of the problems occurring at these levels, a set of specific mathematical models is needed to enable us to describe the dynamics of the relevant processes. This paper deals with the dynamic models needed for an analysis of the lowest level of the hierarchy, i.e. large-scale computerized traffic control and guidance systems restricted to urban street and freeway traffic. The main consideration is which dynamic models and which advanced methods of identification theory are already used or could be used in traffic control systems analysis, and which problems are still unsolved? For this purpose, a survey of basic approaches of dynamic traffic flow modelling is given first, and the significance of the different state and input-output models then discussed, with respect to the role these models play in the analysis of traffic control systems. The conclusion is that the real-time identification of input-output model parameters, though important for the implementation of route guidance and other advanced traffic control systems, so

*Parts of this report have been presented in an invited paper entitled "Application of Parameter and State Identification Methods in Traffic Control Systems" at the 4th IFAC Symposium "Identification and Systems Parameter Estimation" held in Tbilisi, USSR, September 21-27, 1976, as well as in a survey paper presented at the IFAC-IIASA Workshop "Optimization Applied to Transportation Systems" held in Vienna, Austria, February 17-19, 1976.

far did not set much interest in fundamental or applied research. The second part of the paper is therefore a case study of the application of explicit and implicit identification methods for determining input-output models of road sections, long streets with signals at intersections, and street networks.

PART I: DYNAMIC TRAFFIC FLOW MODELS: A SURVEY

STATE MODELS

The dynamic behavior of a system can be described by two different types of models [70]:

- State models, which describe the relations between a set of input signals or control variables, a set of output signals or measurable reactions, and a set of state variables characterizing the state of a system in a rather general sense. One obtains these models by means of laws valid for specific systems, e.g. the Maxwell or Kirchheff laws for electrical systems, the Newton laws for mechanical systems, the laws of traffic flow for traffic systems, etc. The so-called state equations, in the form of a system of linear or nonlinear differential or difference equations of first order, result.

- Input-output models, which describe only the relations between the input or control variables and the output variables, i.e. they do not consider the state variables in an explicit form. Hence, input-output models can describe the dynamics of a system completely, only

- if all state variables can be changed by the input (control) variables in a prescribed manner, i.e. if the system is controllable, and
- if all state variables can be reconstructed by evaluating the (measurable) output variables, i.e. if the system is observable.

Therefore an input-output model can describe only a system which is controllable and observable, or the observable and controllable

part of a general system, respectively*. On the other hand, input-output models have the advantage that they can be designed in many cases without a detailed knowledge of the mathematical laws valid for the specific system. This is of special interest in those systems (e.g. social, environmental) for which such mathematical laws are not available, so that the model can be constructed only on the basis of sets of input-output data. As a result of the application of a parameter identification procedure, one obtains

- nonparametric models, in the form of impulse responses or frequency responses (in the linear case), or Volterra expansions (in the nonlinear case), or
- parametric models, in the form of linear or nonlinear differential or difference equations of the nth order, transfer functions, Hammerstein models, Wiener models, etc. (see [70] for more details).

In this section the description of traffic by state models is discussed; input-output models are the subject of the next section. The following two basic approaches may be used for obtaining dynamic traffic flow models [19]:

- Description of the traffic flow starting from a model of the movement of the individual vehicles (*microscopic traffic flow models*);
- Consideration of the traffic as a fluid continuum (*macroscopic traffic flow models*).

By these two approaches, a fairly well developed and documented traffic flow theory evolved during the fifties and sixties (cf. [10, 17-19, 22-24, 27, 33, 34, 36, 45, 60-62, 77]). The following consideration uses those elements of the traffic flow theory important to dynamic modelling of freeway and street traffic.

*Mathematical conditions for observability and controllability were presented by Kalman at the beginning of the sixties[70].

Microscopic Traffic State Models

In the microscopic traffic modelling approach, it is assumed that every driver who finds himself in a single-lane traffic situation reacts according to the relation:

reaction of driver i at time $t = \lambda_i \{\text{stimulus at time } t - \tau_i\}$

to a stimulus from his immediate environment, especially from the car, $i - 1$, in front of his own car, i (Figure 1). The reaction of the driver may be expressed by the acceleration, $\dot{s}_i(t)$, of his car. λ_i describes the sensitivity of the driver's reaction to a given stimulus, and τ_i is a reaction time lag. It has been shown that the main stimulus is caused by the speed difference, $v_{i-1}(t) - v_i(t) = \dot{s}_{i-1}(t) - \dot{s}_i(t)$, resulting in the nonlinear state {car-following} model

$$\begin{aligned}\dot{s}_i(t) &= v_i(t) \\ \dot{v}_i(t) &= \lambda_i \{v_{i-1}(t - \tau_i) - v_i(t - \tau_i)\},\end{aligned}\tag{1}$$

with

$$\lambda_i = \lambda_{i0} \frac{\{v_i(t - \tau_i)\}^m}{\{s_{i-1}(t - \tau_i) - s_i(t - \tau_i)\}^l}\tag{2}$$

containing the position, s_i , and the speed, v_i , of car i as state variables, and the speed, v_{i-1} , of the leading car, $i - 1$, as the control (input) variable. Equation (2) describes the observations that the sensitivity of the reaction of a driver depends on the speed, $\dot{s}_i = v_i$, of his own car, and the distance, $s_{i-1} - s_i$, between his own car and that in front of him. This is illustrated by the signal flow diagram shown in Figure 2a for a system of two cars only. In the case of a string of N vehicles, one has to couple N of these driver-car models, resulting in a highly nonlinear model for the whole system that is very difficult to handle

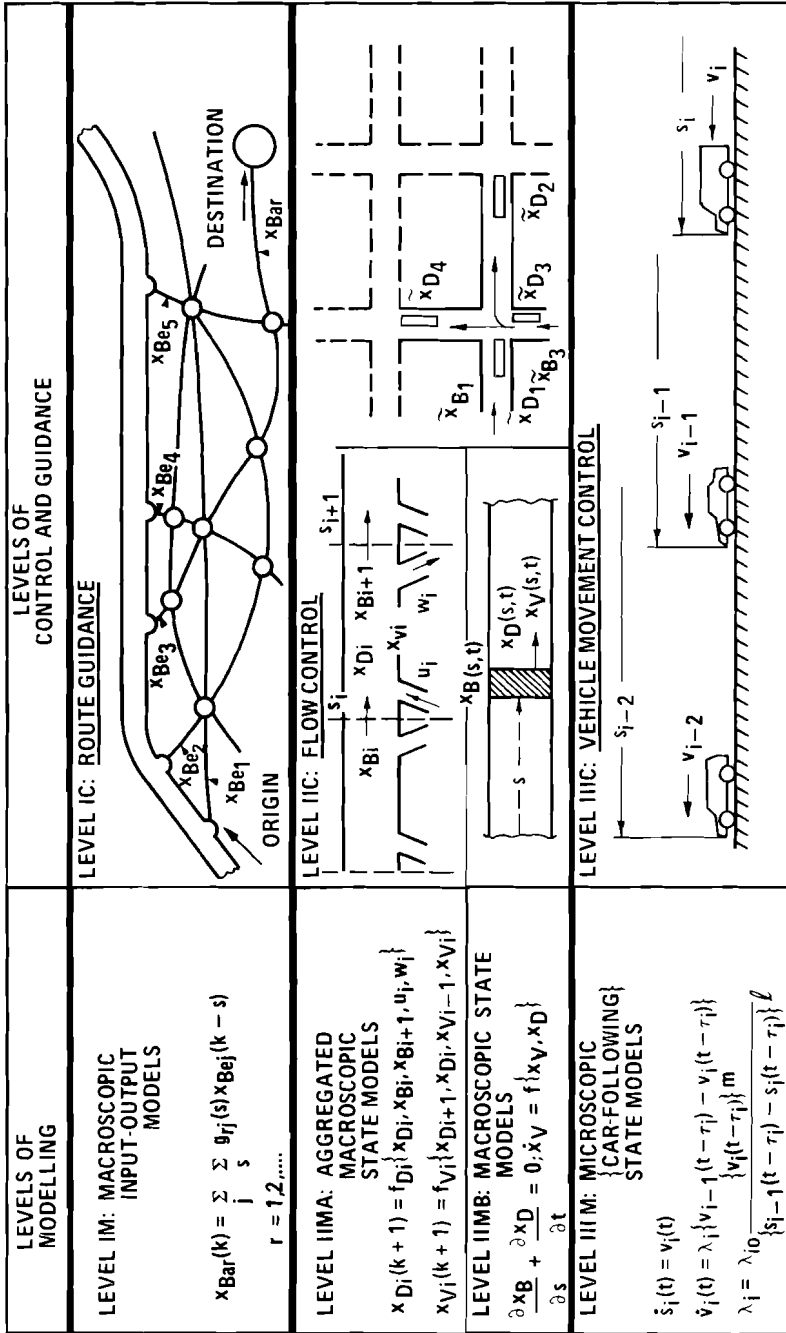


Figure 1. Levels of traffic modelling and control.

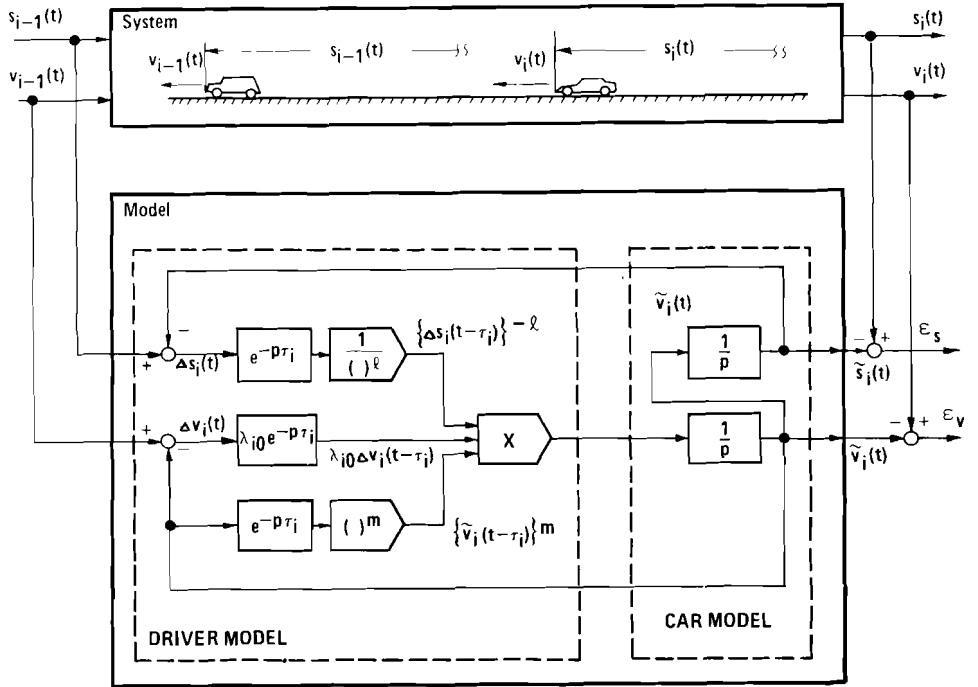


Figure 2a. Nonlinear car-following model for a system of two cars (cf. equations (1) and (2)).

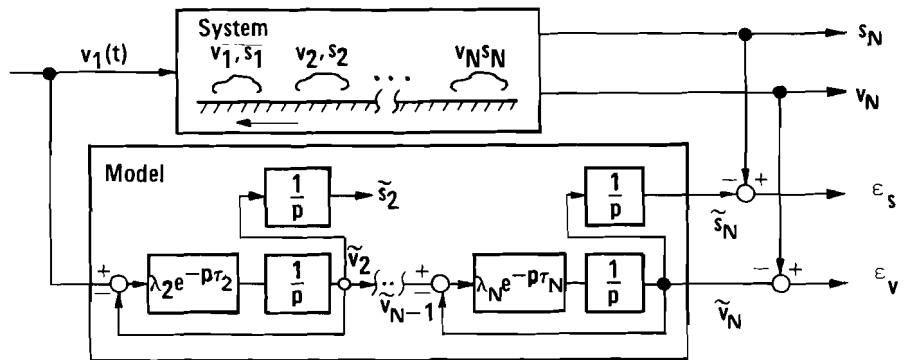


Figure 2b. Linear car-following model for a string of N cars (cf. equations (3) and (4)).

in studying traffic flow dynamics. Therefore, a special case of λ_i from Equation (2) with $\ell = m = 0$ is very often preferred. The so-called linear car-following model,

$$\dot{s}_i(t) = v_i(t) \tag{3}$$

$$\dot{v}_i(t) = \lambda_{i0} \{v_{i-1}(t - \tau_i) - v_i(t - \tau_i)\} ,$$

is obtained, which, for a string of N cars, results in the general state model

$$\dot{\underline{x}}(t) = (A)\underline{x}(t - \underline{\tau}) + (B)u(t - \tau_2) , \tag{4a}$$

with the state vector

$$\underline{x}^T(t) = (s_2(t), v_2(t), \dots, s_N(t), v_N(t)) , \tag{4b}$$

the systems matrix

$$(A) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & -\lambda_{20} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & \lambda_{30} & 0 & -\lambda_{30} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & \lambda_{40} & 0 & -\lambda_{40} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} , \tag{4c}$$

and the control variable $u(t)$. This equals the speed of the leading car,

$$u(t) = v_1(t) , \tag{4d}$$

and the input matrix (B), which in this special case is a vector

of the simple form

$$(B)^T = (0 \quad \lambda_{20} \quad 0 \quad 0 \quad \dots) \quad (4e)$$

(cf. Figure 2b).

As shown in Figures 2a and 2b, this model of traffic flow may be considered as a series of interconnected control loops. Each control loop can become unstable for certain values of the time delays, τ_i , and the sensitivity coefficients, λ_{i0} . Instability means in this context that collisions will occur even if one considers a system of two cars only. On the other hand, stability of the individual control loops shown in Figure 2b will not give any guarantee that a collision will not happen in a long string of vehicles. To check the conditions for the occurrence of collisions in a queue of N cars, one has to study the time responses of the distances, $\Delta s_i = s_{i-1} - s_i$, between the cars resulting from changes of the control variable (4d), i.e. of the speed of the leading car. The model ((3) to (4e)) permits such studies only if estimates of the model parameters τ_i and λ_{i0} are available. By means of special experimental studies on the driver-car performance, Herman and his co-workers (cf. [19, p. 89 ff.]) have obtained for 8 different drivers the following estimates for λ , τ , and $\lambda\tau$ and the mean values $\bar{\lambda}$, $\bar{\tau}$, and $\bar{\lambda\tau}$:

$$\begin{aligned} 0.17 \text{ s}^{-1} &\leq \lambda_{i0} \leq 0.74 \text{ s}^{-1} && \text{with } \bar{\lambda}_0 = 0.37 \text{ s}^{-1} \\ 1.0 \text{ s} &\leq \tau_i \leq 2.2 \text{ s} && \text{with } \bar{\tau} = 1.55 \text{ s} \\ 0.18 &\leq \lambda_{i0}\tau_i \leq 1.04 && \text{with } \bar{\lambda}_0\bar{\tau} = 0.56 \end{aligned} \quad (5)$$

For more details on microscopic traffic models see [10, 19, 22-24, 33, 34].

Macroscopic Traffic State Models

Macroscopic traffic flow theory was founded by Lighthill and Whitham [45] and by Richards [62] during the fifties. They

considered a traffic stream as a fluid continuum described by the 3 aggregated traffic variables

- volume $x_B(s,t)$ (cars/h)
- density $x_D(s,t)$ (cars/km)

and

- speed $x_V(s,t)$ (km/h)

(cf. Figure 1). Since these three variables are related to each other by

$$x_B(s,t) = x_D(s,t)x_V(s,t) \quad , \quad (6)$$

it is sufficient to introduce only two of them, e.g. x_D and x_V , as state variables. Hence only two differential equations of the first order are needed for a traffic state model of a single long lane for which overtaking may not be permitted. Using the principle of conservation of cars, one gets the first state equation in the form of the well-known partial differential equation

$$\frac{\partial x_D(s,t)}{\partial t} + \frac{\partial x_B(s,t)}{\partial s} = 0 \quad . \quad (7)$$

For the second state equation, which has to describe the acceleration of the traffic stream, Payne [55] and Isaksen and Payne [39] have proposed the relation

$$\frac{dx_V(s,t)}{dt} = - \frac{1}{T} \{x_V(s,t) - \bar{x}_V(x_D)\} - \frac{v}{T} \frac{1}{x_D(s,t)} \frac{\partial x_D(s,t)}{\partial s} \quad , \quad (8)$$

where the term containing v represents the average reaction of drivers to a change in density ahead. The parameters v and T may be considered as the sensitivity coefficient and the reaction time constant, respectively. The first term on the right-hand side of equation (8) takes into account the average behavior of drivers to keep the speeds of their cars close to the speed

$\bar{x}_V(\bar{x}_D)$ that occurs under constant traffic conditions. For this speed, the so-called Greenshields model

$$\bar{x}_V = \bar{x}_f (1 - \bar{x}_D / \bar{x}_{Dmax}), \text{ with } \bar{x}_f = \lambda / \bar{x}_D, \quad (9)$$

can be determined by means of the linear car-following model (3), where \bar{x}_f is the free traffic speed and \bar{x}_{Dmax} , the jam concentration [19, 77] (cf. Figure 3). The nonlinear car-following model (2) delivers, for $\ell = 1$ and $m = 0$, the so-called Greenberg model [27, 22, 19]

$$\bar{x}_V = \lambda \ln \{ \bar{x}_{Dmax} / \bar{x}_D \} . \quad (10)$$

There are, of course, further possibilities for describing the interrelation between the stationary values, \bar{x}_V and \bar{x}_D , of speed and density--for example, the equation

$$\bar{x}_V = \bar{x}_f - \sum_{i=1}^n a_i \bar{x}_D^i, \quad (11)$$

used successfully by Isaksen and Payne [38] with $n = 3$ for the Los Angeles Freeway, the relation

$$\bar{x}_V = \bar{x}_f (1 - \bar{x}_D / \bar{x}_{Dmax}) (1 - \alpha \bar{x}_D / \bar{x}_{Dmax}), \quad (12)$$

with $-1 \leq \alpha \leq 1$ (cf. [54]), or the expression

$$\bar{x}_V = b \exp \left\{ - \frac{1}{2} \left[\frac{\bar{x}_D}{a} \right]^2 \right\} \quad (13)$$

successfully applied for studies in New York's Lincoln Tunnel by Szeto and Gazis [76]. These static models represent, together with the equation

$$\bar{x}_B = \bar{x}_D \bar{x}_V \quad (14)$$

for the traffic volume (see equation (6)), the "fundamental diagram of traffic" (see Figure 3) which explains some essential traffic flow phenomena as shock waves (see [10, 18, 19, 33, 43] for more details).

Equations (6)-(14) describe a nonlinear distributed parameter model that can be applied for the analysis of traffic control systems in a simplified, i.e. aggregated, form only.

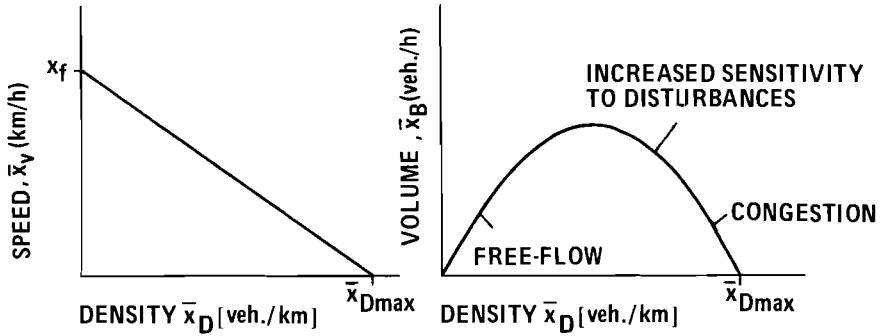


Figure 3. Fundamental diagram of traffic.

Aggregated Macroscopic Traffic State Models

The necessary simplification is effected by dividing the freeway, the tunnel, or the bridge into sections, Δs_i , of the length, and by introducing the aggregated state variables (cf. Figure 1)

$$\text{section density, } x_{Di}(k) = \frac{1}{\Delta s_i} \int_{s_i}^{s_i + \Delta s_i} x_D(s, t_k) ds \quad , \quad (15)$$

and

$$\text{section speed, } x_{Vi}(k) = \frac{1}{\Delta s_i} \int_{s_i}^{s_i + \Delta s_i} x_V(s, t_k) ds \quad , \quad (16)$$

in the form of spatial means of $x_D(s, t)$ and $x_V(s, t)$ for discrete time intervals $t_k = k\Delta t$. For the traffic volume at the section

boundaries, the temporal means,

$$x_{Bi}(k) = \frac{1}{\Delta t} \int_{t_k - \Delta t}^{t_k} x_B(s, t) dt \quad , \quad (17)$$

are used. With these definitions, it is possible to approximate the partial differential equation (7) by the simple difference equation

$$x_{Di}(k + 1) = x_{Di}(k) + \frac{\Delta t}{\Delta s_i} \{x_{Bi}(k) - x_{Bi+1}(k)\} \quad , \quad (18)$$

if $\partial x_D / \partial t$ is replaced by $\{x_{Di}(k + 1) - x_{Di}(k)\} / \Delta t$ and $\partial x_B / \partial s$ by $\{x_{Bi} - x_{Bi+1}\} / \Delta s_i$. This equation illustrates the principle of conservation of cars. It is quite obvious that the number of cars, $\Delta s_i x_{Di}(k + 1)$, at time $(k + 1)\Delta t$ in section i results from the number of cars, $\Delta s_i x_{Di}(k)$, stored at $k\Delta t$, plus the difference of the numbers of cars entering, $\Delta t x_{Bi}(k)$, and leaving, $\Delta t x_{Bi+1}(k)$, the section during the time interval $(k - 1)\Delta t < t \leq k\Delta t$. For a freeway section connected with on- and off-ramps carrying the traffic volumes $u_i(k)$ and $w_i(k)$ (cf. Figure 1 for level IIC), equation (18) has to be enlarged to the more general relation

$$x_{Di}(k + 1) = x_{Di}(k) + \frac{\Delta t}{\Delta s_i} \{x_{Bi}(k) - x_{Bi+1}(k) + u_i(k) - w_i(k)\} \quad . \quad (19)$$

If the same simplification method as used in equation (7) is applied to (8), then the nonlinear difference equation

$$x_{Vi}(k + 1) = x_{Vi}(k) - \frac{2\Delta t x_{Vi}(k)}{\Delta s_{i+1, i-1}} \{x_{Vi}(k) - x_{Vi-1}(k)\} + \frac{\Delta t}{T} \left\{ x_{Vi}(k) - \bar{x}_{Vi} - \frac{2v}{\Delta s_{i+2, i}} [x_{Di+1}(k) - x_{Di}(k)] / x_{Di}(k) \right\} \quad , \quad (20)$$

with $\Delta s_{rj} = s_r - s_j$, is obtained (see [54, 55] for more details).

The second term introduced on the right-hand side of this equation allows for the section speed, $x_{Vi}(k)$, possibly changing at the section boundary, i.e. for $x_{Vi}(k) \neq x_{Vi-1}(k)$ (cf. Figure 1). For the first section ($i = 1$), it is obviously justified to choose $x_{V0} = x_{V1}$ and to neglect that term, while for the last section ($i = N$), it is reasonable to assume that x_{DN} is equal to x_{DN+1} and to exclude the term involving the sensitivity coefficient, v . In this way, with

$$x_{Bi}(k) = x_{Di-1}(k) x_{Vi-1}(k) \quad (21)$$

(cf. equation (6)) for a three section freeway as an example (cf. Figure 1), one gets the state equations

$$\begin{aligned} x_{D1}(k+1) &= x_{D1}(k) + \frac{\Delta t}{\Delta s_{21}} [-x_{D1}(k)x_{V1}(k) + x_{B1}(k) + u_1(k) - w_1(k)] \quad , \\ x_{V1}(k+1) &= x_{V1}(k) + \frac{2v\Delta t}{T\Delta s_{31}} \left[1 - \frac{x_{D2}(k)}{x_{D1}(k)} \right] - \frac{\Delta t}{T} [x_{V1}(k) - \bar{x}_{V1}] \quad , \\ x_{D2}(k+1) &= x_{D2}(k) + \frac{\Delta t}{\Delta s_{32}} [-x_{D2}(k)x_{V2}(k) + x_{D1}(k)x_{V1}(k) + u_2(k) - w_2(k)] \quad , \\ x_{V2}(k+1) &= x_{V2}(k) + \frac{2v\Delta t}{T\Delta s_{42}} \left[1 - \frac{x_{D3}(k)}{x_{D2}(k)} \right] \quad (22) \\ &\quad - \frac{\Delta t}{T} [x_{V2}(k) - \bar{x}_{V2}] - \frac{2\Delta t}{\Delta s_{31}} x_{V2}(k) [x_{V2}(k) - x_{V1}(k)] \quad , \\ x_{D3}(k+1) &= x_{D3}(k) + \frac{\Delta t}{\Delta s_{43}} [-x_{D3}(k)x_{V2}(k) + x_{D2}(k)x_{V2}(k) + u_3(k) - w_3(k)] \quad , \end{aligned}$$

and

$$x_{V3}(k+1) = x_{V3}(k) - \frac{\Delta t}{T} [x_{V3}(k) - \bar{x}_{V3}] - \frac{2\Delta t}{\Delta s_{42}} x_{V3}(k) [x_{V3}(k) - x_{V2}(k)]$$

representing a nonlinear traffic state model of the form

$$\underline{x}(k + 1) = \underline{f}(\underline{x}(k), \underline{u}(k)) \quad , \quad (23)$$

with the six-dimensional state vector

$$\underline{x}^T(k) = (x_{D1}(k) x_{V1}(k) x_{D2}(k) x_{V2}(k) x_{D3}(k) x_{V3}(k)) \quad , \quad (24)$$

and the three-dimensional control vector

$$\underline{u}^T(k) = (u_1(k) u_2(k) u_3(k)) \quad . \quad (25)$$

This freeway traffic model can be changed to a tunnel or bridge traffic model by putting $u_i = w_i = 0$ and introducing the input traffic volume $x_{B1}(k)$ at the first section as a control variable, $u_1(k)$.

The aggregated traffic model (19)-(22) is still complicated, especially because of the nonlinear equation (20). In specific applications, it may be necessary and possible to substitute a further simplified expression for this relation. Nahi and Trivedi [52], in connection with a density estimation problem in one freeway section, have used the difference equation

$$x_V(k + 1) = \alpha x_V(k) + \eta(k) \quad , \quad (26)$$

here $\eta(k)$ is a stochastic disturbance causing speed changes. For the solution of the same problem in tunnel traffic, Szeto and Gazis [76] successfully applied static model (13) as a dynamic model after introducing a noise term $\xi(k)$:

$$x_V(k) = b \exp \left\{ -\frac{1}{2} \left[\frac{x_D(k)}{a} \right]^2 \right\} + \xi(k) \quad . \quad (27)$$

The application of equations (20), (21), (26), and (27), given here as models of the speed behavior of a traffic stream, is not possible in the case of urban street networks containing signals at intersections. The dynamics of traffic flow in urban

streets changes rapidly with traffic volume, and so it is not possible to present one dynamic model valid for all possible traffic conditions.

A relatively simple model can be designed for the complicated case of a network of oversaturated intersections characterized by the traffic supply being smaller than the traffic demand so that queues of cars are always waiting. In this special case, the travel time between intersections is much smaller than the waiting time at the intersections, and so the speed dynamics in equation (20) can be neglected; only the principle of conservation of cars expressed by equation (19) need be used. This applied to the simple one-way network shown in Figure 1, for example, results in the state model

$$\begin{bmatrix} \tilde{x}_{D1}(k+1) \\ \cdot \\ \cdot \\ \tilde{x}_{D4}(k+1) \end{bmatrix} = \begin{bmatrix} \tilde{x}_{D1}(k) \\ \cdot \\ \cdot \\ \tilde{x}_{D4}(k) \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 & 0 \\ s_1 & -1 & r_3 & 0 \\ 0 & 0 & -1 & 0 \\ r_1 & 0 & s_3 & -1 \end{bmatrix} \cdot \begin{bmatrix} \tilde{u}_1(k) \\ \cdot \\ \cdot \\ \tilde{u}_4(k) \end{bmatrix} + \begin{bmatrix} \tilde{x}_{B1}(k) \\ 0 \\ \tilde{x}_{B3}(k) \\ 0 \end{bmatrix} \quad (28)$$

$$\hat{=} \tilde{\underline{x}}_D(k+1) = \tilde{\underline{x}}_D(k) + (B) \cdot \tilde{\underline{u}}(k) + \underline{z}(k) ,$$

where $\tilde{x}_{Di}(k)$ is the number of cars waiting in link i at the corresponding intersection, $\tilde{u}_i(k)$, the control variable (i.e. the number of cars leaving the link when the green traffic light is flashing), and $\tilde{x}_{Bi}(k)$, the number of cars arriving at link i during the time interval $(k-1)\Delta t < t \leq k\Delta t$. The parameters s_i and r_i describe the percentage of cars $\tilde{u}_i(k)$ going straight ahead, s_i , or turning to the right or left, r_i (see [3, 16, 20, 46, 73] for more details).

If it is not reasonable to neglect the travel times τ_i , then equation (28) has to be changed to the more complicated

$$\tilde{x}_D(k+1) = \tilde{x}_D(k) + (B)\tilde{u}(k) + (B)_1\tilde{u}(k-\tau_1) + (B)_2\tilde{u}(k-\tau_2) + \dots + \underline{z}(k) \quad (29)$$

as shown by Singh, et al. [68].

For light traffic conditions, without permanent queues at the intersections, the dynamics of the group of cars (platoons) formed at intersections with signals have to be considered. Platoon models, however, are discussed in the following paragraph on input-output models.

INPUT-OUTPUT MODELS

An overview of models is now presented where the state variables are not considered explicitly, and where the main interest is in the dynamic relations between certain control or input variables of the system and the reactions of the system measured by output variables.

Microscopic Input-Output Models

If one introduces as the input signal the speed of the leading car, $x_e(t) = v_1(t)$, and as the output signals the speed, $x_{a1}(t) = v_N(t)$, and the position, $x_{a2}(t) = s_N(t)$, of the Nth car in a string of N cars, then the linear car-following model illustrated by Figure 2b results in the two transfer functions:

$$G_1(p) = \frac{Z\{v_N(t)\}}{Z\{v_1(t)\}} = \prod_{i=2}^N \frac{\lambda_{i0} e^{-p\tau_i}}{p + \lambda_{i0} e^{-p\tau_i}} = \frac{x_{a1}(p)}{x_e(p)}$$

and

$$G_2(p) = \frac{Z\{s_N(t)\}}{Z\{v_1(t)\}} = \frac{1}{p} G_1(p) = \frac{x_{a2}(p)}{x_e(p)} .$$

These input-output models have not yet attracted much attention in traffic flow theory literature (cf. [10, 19, 33]), though they provide an excellent basis for the application

of the well-developed classical control theory on stability and collision analysis.

Macroscopic Input-Output Models

This approach to dynamic traffic flow modelling has so far been studied only by a rather small number of authors (cf. [11, 74]). The input-output model describes the interrelations between macroscopic traffic variables at those points of a freeway, tunnel, rural road, or even urban street network that are of special interest and may be defined as inputs and outputs of the traffic system. Traffic volume will be introduced here as the input and output variables.

First, a single traffic link and a long street including intersections, or a freeway with on- and off-ramps will be considered.

The Traffic Route Model

For a single driving route, it is reasonable to introduce as input and output variables the number of cars entering, $x_{Be}(k)$, or leaving, $x_{Ba}(k)$, the route during the time interval $(k-1)\Delta t < t \leq k\Delta t$ (cf. Figure 5). According to

$$x_{Ba}(k) = \overset{O}{x}_{Ba}(k) + x_{Bz}(k) \quad (30)$$

the output traffic volume $x_{Ba}(k)$ consists of the number of cars coming from the input, $\overset{O}{x}_{Ba}(k)$, and the number of cars entering the route through other access points, $x_{Bz}(k)$. $x_{Bz}(k)$ may be considered as a disturbance, while $\overset{O}{x}_{Ba}(k)$ is assumed to depend on those values $x_{Be}(k-m), \dots, x_{Be}(k-n)$ of the input traffic volume $x_{Be}(k)$ that are delayed by the travel time interval

$$T_{Rmin} = m\Delta t \leq T_R \leq n\Delta t = T_{Rmax}$$

in comparison with the time $t_k = k\Delta t$:

$$\dot{x}_{Ba}^0(k) = f\{x_{Be}(k-m), \dots, x_{Be}(k-n), k\} \quad . \quad (30a)$$

By developing $f\{\dots\}$ into a Taylor series and neglecting the nonlinear terms, one obtains

$$\dot{x}_{Ba}^0(k) = \tilde{g}(m, k)x_{Be}(k-m) + \dots + \tilde{g}(n, k)x_{Be}(k-n) \quad ,$$

where the parameters $\tilde{g}(s, k)$ describe the percentage of cars reaching the output within the travel time interval $(s-1)\Delta t < T_R \leq s\Delta t$. If we assume that the expectations, $E\{\tilde{g}(s, k)\} = g(s)$, of the time varying parameters are constant, then it is reasonable to use the model

$$\begin{aligned} x_{Ba}(k) &= x_{Ba}^0(k) + x_{Bz}(k) \\ &= g(m)x_{Be}(k-m) + \dots + g(n)x_{Be}(k-n) + x_{Bz}(k) \quad , \end{aligned} \quad (31)$$

where the sum

$$h(s) = \sum_{i=m}^s g(i) \quad (32)$$

characterizes the proportion of cars with a travel time $T_R \leq s\Delta t$, and $h(n)$ describes the number entering the route via the input and leaving it via the output. This parameter $h(n)$ (henceforth called the "split coefficient") can take values within the interval

$$0 \leq h(n) \leq 1 \quad . \quad (33)$$

For $h(n) = 1$, no car out of $x_{Be}(k)$ leaves the route before reaching the output; for $h(n) = 0$, all cars $x_{Be}(k)$ leave the route before reaching the output.

From $h(n)$ and $g(s)$, it is possible to obtain the relation

$$F(s) = h(s)/h(n) = P\{T_R \leq s\Delta t\} \quad (34)$$

for the probability distribution of the travel time $T_R = s\Delta t$, and the corresponding formula for the density

$$f(s) = g(s)/h(n) = P\{(s - 1)\Delta t \leq T_R \leq s\Delta t\} , \quad (35)$$

resulting in

$$T_{RM} = \sum_{s=m}^n (s\Delta t) f(s) \quad (36)$$

for the mean travel time, T_{RM} , that the traffic stream takes to pass between input and output (cf. Figure 5).

The coefficients, $g(s)$, in equation (31) may be considered as discrete values of the impulse response, $g(t)$, of the street. Thus, equation (31) results, with $g(s) = 0$ for $s = 0, \dots, m-1, n+1, \dots$, in the convolution sum

$$x_{BMA}(k) = \sum_{i=0}^{\infty} g(i)x_{Be}(k - i) , \quad (37)$$

or in Duhamel's convolution integral

$$x_{BMA}(t) = \int_0^{\infty} g(\tau)x_{Be}(t - \tau)d\tau = g(t)*x_{Be}(t)$$

if very small sampling time intervals, Δt , and continuous time functions are used. Two interesting statements can thus be made:

- The impulse response of the route is proportional to the travel time probability density function (cf. equation (35))
- The step response of the road describes the probability distribution function according to equation (34), with the final value $h(n)$ being the split coefficient that characterizes the percentage of cars passing the whole route between the input and the output without leaving it anywhere (cf. Figure 5).

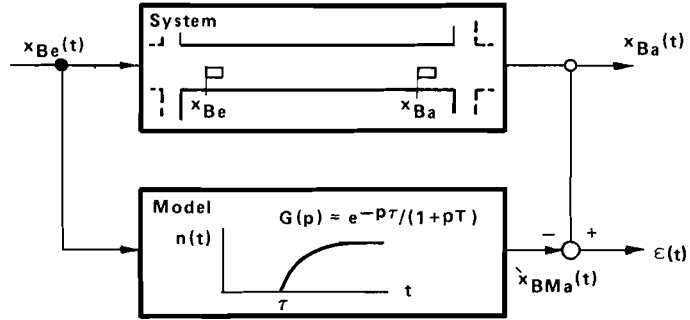


Figure 4. Parametric input-output model obtained from Robertson's [63] platoon model.

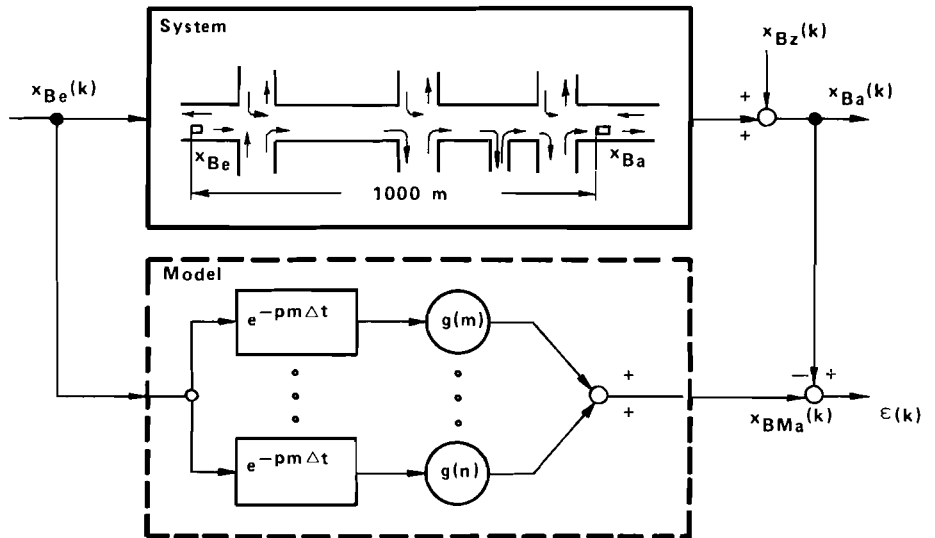


Figure 5. Macroscopic input-output model illustrated for the North-South-Connection in Dresden ($\Delta t = 10, \rho =$ electronic traffic detectors).

The model (38) obtained here is a nonparametric one [70], giving rise to questions about the applicability of parametric models in the form of difference or differential equations. The so-called platoon dispersion model

$$x_{BMa}(k + 1) = (1 - F)x_{BMa}(k) + Fx_{Be}(k - N + 1) \quad , \quad (38)$$

with $\tau = N\Delta t = 0.8T_{RM}$ and $F = 1/\{1 + 0.5N\}$, proposed by Robertson [63] (see Figure 4), may be considered as such a parametric model that has been successfully used for the simulation of platoon dynamics, i.e. for traffic links connecting neighboring inter-sections. For small sampling intervals, Δt , the difference equation (38) can be approximated by the differential equation

$$\dot{x}_{BMa}(t) + \tilde{F}x_{BMa}(t) = \tilde{F}x_{Be}(t - \tau) \quad , \quad (39)$$

with $\tilde{F} = F/\Delta t$, resulting in the well-known transfer function

$$G(p) = \frac{e^{-p\tau}}{1 + p/\tilde{F}} \quad . \quad (40)$$

The applicability of other parametric models, e.g. of the Åström-model (see [70]), has been studied by Doormann [11].

The Traffic Network Model

The traffic route model (31) can be enlarged to a network model by using

$$\begin{bmatrix} x_{Ba1}(k) \\ \vdots \\ x_{Bal}(k) \end{bmatrix} = \begin{bmatrix} g_{11}(k) \dots g_{1r}(k) \\ \vdots \\ g_{l1}(k) \dots g_{lr}(k) \end{bmatrix} * \begin{bmatrix} x_{Be1}(k) \\ \vdots \\ x_{Ber}(k) \end{bmatrix} + \begin{bmatrix} x_{Bz1}(k) \\ \vdots \\ x_{Bzl}(k) \end{bmatrix} \quad , \quad (41)$$

where * is the convolution symbol. For the network shown in Figure 6, for example, one obtains for output 1

$$x_{Ba1}(k) = \sum_{j=2}^4 \sum_{s=m_{1j}}^{n_{1j}} g_{1j}(s)x_{Bej}(k-s) + x_{Bz1}(k) \quad . \quad (42)$$

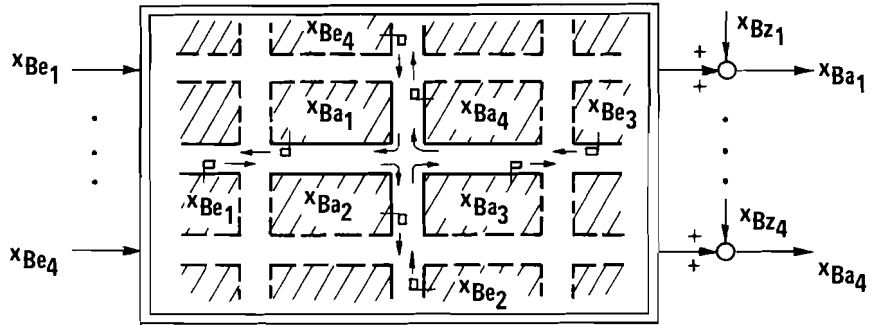


Figure 6. Macroscopic input-output model of a traffic network.

The percentage of cars traveling from the input x_{Bej} to the output x_{Bai} is given by the split coefficient

$$h_{ij}(n_{ij}) = \sum_{s=m_{ij}}^{n_{ij}} g_{ij}(s) \quad , \quad (43)$$

and for the corresponding travel time distribution functions one obtains the general expression

$$f_{ij}(s) = g_{ij}(s) / \sum_{s=m_{ij}}^{n_{ij}} g_{ij}(s) \quad , \quad \text{with } g_{ij}(s) \geq 0 \quad , \quad (44)$$

resulting in

$$T_{RM_{ij}} = \sum_{s=m_{ij}}^{n_{ij}} (s\Delta t) f_{ij}(s) \quad (45)$$

for the mean travel time $T_{RM_{ij}}$ that the individual traffic streams need for going from the input j to the output i .

THE ROLE OF STATE AND INPUT-OUTPUT MODELS IN THE ANALYSIS OF TRAFFIC CONTROL SYSTEMS

We now describe what role the models summarized in the two preceding sections already play or are expected to play in the analysis of computerized traffic control systems. A brief summary of the existing fundamental traffic control concepts is given first.

Traffic Control Concepts

The basic concepts proposed for freeway and area traffic control may be considered as hierarchically structured control systems containing three levels (cf. [71, 72] and Figure 1):

- Optimal guidance of main traffic streams through a network of freeways and service streets (route control)
- Optimal traffic flow control on freeways, in tunnels, on bridges, and in urban street networks
- Vehicle movement control, e.g. merging control and distance regulation in a string of moving vehicles.

A survey of the state of implementation of these concepts and the experiences gained from real applications is given by Strobel in [71, 72]. Here only those methodological aspects of interest from the identification viewpoint are described.

Models Needed for Route Control Analysis*

A route control system assists drivers in finding the (in some sense) best route from a certain origin to a desired destination. Due account is taken of changing traffic conditions in different parts of the network caused, for example, by accidents, weather, and maintenance operations (cf. Figure 1, level IC).

*These are described in [2, 4, 7, 14, 53, 64, 71, 72, 81].

A computing system is provided and, from information given by traffic detectors, it evaluates the traffic situation and determines the optimal routes in real-time operation. These routes are shown to drivers by changeable computer-controlled road signs located at freeway off-ramps and essential intersections of the arterial street network [1, 8, 12]. The use of displays within cars [15, 31, 50] is sometimes included.

For the determination of an optimal route, two criteria should be taken into account:

- Minimization of mean travel times between origin and destination points (Wardrop's first principle)
- Minimization of traffic density in different parts of the network, i.e. of the weighted sum of all densities (Wardrop's second principle).

In route control systems so far installed, the travel time criterion is preferred in general. Thus, a route control algorithm contains the following two parts:

- An identification part dealing with the estimation of the mean travel times, and the split coefficients describing the distribution of the traffic streams within the network. This task has to be solved in an on-line real-time operation mode with the use of traffic detector data only.
- A real-time optimization part dealing with the selection of optimal routes.

It is quite obvious that the solution of the identification task of the first part needs a traffic model similar or equal to that presented in the section on macroscopic input-output models. But it is interesting that none of the route control systems implemented in the past have used such a macroscopic input-output model. One may conclude that the application of input-output models in route control systems presents an unsolved problem which is a subject of fundamental research (cf. [11, 74]).

Models Needed for Flow Control Analysis

Freeway and Tunnel Traffic*

The capacity of a traffic lane decreases if the traffic density gets larger than an optimal value. This well-known phenomenon, illustrated by the fundamental diagram in Figure 3, explains the occurrence of natural congestion on freeways and in tunnels when too many cars enter traffic links. The aim of a traffic flow control system is therefore to maintain traffic demand along all parts of the freeway below the critical level by restricting freeway access by means of traffic lights at the entrance ramps. The necessary traffic light control algorithm requires the solution of the following two optimization problems.

- Static optimization and open-loop control [26, 38-40, 58]: With the use of demand patterns obtained from historical data, nominal values for the input flow rates, u_i , have to be determined in such a way that the overall traffic throughput is maximized. These control variables, u_i , are, of course, no longer the optimal ones if disturbances, e.g. an accident, occur. For such situations, one uses:
- Dynamic optimization and feedback control [21, 35, 37-40, 47, 59, 77]: The task of this control system is to minimize deviations between the nominal, precomputed state variables density, x_{Di} , and speed, x_{Vi} , and their actual values, by real-time computation of corrections to the nominal values that take account of control variables, u_i (cf. Figure 1, level IIC).

Fundamental contributions to the solution of these problems were presented by Isaksen, Payne and their associates [29, 30, 37-40, 49, 55-59]. They used for the first time the aggregated macroscopic state model (19)-(25). The application of this model requires the solution of the following state and parameter identification problems which is still a subject of fundamental

*See [5, 12, 17, 19, 21, 25, 26, 28-30, 35-42, 47, 49, 51, 52, 54-59, 66, 71, 72, 80, 81].

research [5, 9, 25, 28, 30, 35, 42, 49, 51, 52, 54-57]:

- The parameters of the dynamic equations (20) and (22), and of the static models (9)-(13), change with weather conditions, traffic incidents, etc. They have to be determined by means of an on-line real-time method.
- It is not possible to take direct measurements of the state variables section density, $x_{Di}(k)$, and section speed, $x_{Vi}(k)$. Traffic detectors measure traffic volumes, $x_{Bi}(k)$, and mean speeds only at fixed points, i.e. at the section boundaries (cf. Figure 1).

One has to deal therefore with a combined state and parameter estimation problem, leading to the application of the extended Kalman filter [13]. This problem has been studied by Orhac, et al. [54] for the three section freeway model (22). The complicated form of this model has not yet permitted algorithms reliable and robust enough for practical application to be developed. Thus this problem, too, is still a subject of fundamental research, especially with respect to the use of decentralized principles for control and identification [54, 57, 67]. The situation is different if one considers only one freeway section and tries to solve the state identification problems for the individual sections independently of each other. First successful applications of modern identification methodology to this simplified problem have been reported by Gazis and his co-workers [5, 25, 76] and later by Nahi [51, 52]. Szeto and Gazis [76] used equations (18) and (28) for a model of a tunnel section and introduced the time varying model parameters, a and b , as additional state variables with the simple state equations

$$a_{k+1} = a_k \quad \text{and} \quad b_{k+1} = b_k \quad .$$

Experimental tests carried out for New York's Lincoln Tunnel delivered reliable and sufficiently accurate estimates for the section density and speed by the application of the extended Kalman filter (see [25, 76] for more details). Similar promising

results have been obtained by Nahi and Trivedi [51, 52] for a freeway section with the use of a recursive minimum square estimator, and model equations from relations (18) and (27).

Urban Street Networks*

The most widely used traffic control concept is traffic light control and coordination. The methods in use can be classified under:

- Precomputation of optimal signal programs for time-of-day, open-loop control by heuristic methods, mathematical programming methods, or simulation techniques;
- traffic-responsive signal program selection, i.e. adaptive open-loop control;
- traffic-responsive signal program modification and generation, i.e. feedback control (cf. [10, 19, 36, 48, 65, 78]).

Problems of parameter and state estimation do not play an important role for the first two methods. A particular exception is the application of simulation programs for the determination of optimal signal programs. These simulation programs use a simple model describing the principle of the conservation of cars at intersections (cf. equation (18)), and a platoon dispersion model simulating the traffic flow between intersections. This is true, for example, for the well-known and widely used TRANSYT simulation program of Robertson [63] which contains the input-output model (38) as the platoon model.

Situations where modern identification methods are going to be an important and useful tool occur under complicated traffic conditions that require implementation of feedback control algorithms. In such situations, the traffic control problem must be handled by a multicriterion approach with consideration of the following hierarchy of criteria [36]:

- stoppage mode for light traffic,

*See [3, 16-20, 36, 46, 48, 63, 65, 68, 73, 77-79].

- delay mode for medium traffic,
- capacity mode for saturated intersections,
- queue mode for very dense traffic, and
- jam mode for congested conditions.

The first two criteria are generally used for the pre-computation of signal programs [10, 17-19, 36, 48, 65, 78]. The last three, needed for heavy traffic conditions, can be implemented only as a feedback control algorithm. For the last criterion, for example, an optimal control strategy for time optimal congestion removal can be obtained by means of the state model (28) with the use of Pontragin's maximum principle [16, 46, 73]. Here the identification problem consists in the determination of parameters s_i and r_i , i.e. of the percentage of cars driving straight ahead, and turning to the right and left. As shown earlier, such a task is the same as the identification of the split coefficient, $h_{ij}(n_{ij})$, according to equation (43) and the macroscopic input-output model (42).

The same is true for the model (29) that has been used for the design of control strategies fulfilling the third and fourth optimizing criteria (cf. [68]). For both models (28) and (29), a state estimation problem occurs if one has to determine the queue lengths, i.e. the numbers, $\tilde{x}_{Di}(k)$, of waiting cars at different intersections, by means of noisy detector measurements of traffic volume at selected points along the traffic links [73, 76].

Models Needed for Vehicle Movement Control Analysis*

The lowest level of the control hierarchy shown in Figure 1 mainly concerns the problem of distance regulation in a string of moving highway vehicles, with the aim of reducing the danger of collisions and increasing the freeway capacity. It is obvious that such problems require the application of microscopic models and the microscopic traffic flow models discussed earlier can

*These are described in [6, 32, 69, 71, 72, 75, 78].

be used to illustrate the significance of the problem. The linear car-following model ((1), (3), and (4)), for example, permits the phenomenon of traffic queue instability to be explained. It can be shown by equations (3) and (4) that a system of two cars is unstable for $\lambda_{i0}\tau_i > \pi/2$, and that oscillations with damped amplitude for $1/e < \lambda_{i0}\tau_i < \pi/2$ result. Instability and oscillations occur if a driver reacts too slowly (large τ_i) or too much (large λ_{i0}) to speed changes of the leading car. Small speed changes of that car are amplified resulting, in long strings of vehicles, in collisions of the cars at the end of the queue. If one assumes the same model for all drivers, i.e. $\lambda_{i0} = \lambda, \tau_i = \tau$, then this result occurs as soon as $\lambda\tau > 0.5$ --a value of the same order of magnitude as the experimentally determined values given in equation (5). This agrees with the feeling of many drivers that, when driving in long strings of vehicles at high speeds, they are often close to the limit of stability. To reduce this danger of collision, one should provide the drivers with certain driving aids that would assist them to have stable control parameters, λ_{i0} and τ_i . Radar distance measuring devices [69] and special head-up driver displays might fulfill this task [75], but since these problems are the subject of fundamental research, certain identification problems may occur concerning the simulation of driver behavior. On the other hand distance regulation systems already play a significant role today in the development of so-called "automated guideway transit systems" (cf. [71, 72]).

Conclusions

The following conclusions may be drawn from the state-of-the-art survey presented above:

- The essential foundations for dynamic traffic flow modelling were created by the development of the microscopic and macroscopic traffic flow theory during the fifties and sixties [10, 19, 33]. The application of modern identification methods, however, has been the subject of theoretical and experimental studies carried

out during the last five years or so, and is still a subject of fundamental research work.

- This is especially true for the application of the extended Kalman filter and related methodology for parameter and state estimation problems occurring with the computer control of freeway and tunnel traffic [5, 19, 25, 28-30, 41, 42, 49, 54-57]. Research activities have resulted in an enlargement of macroscopic traffic flow theory by introducing the Isaksen-Payne model ((8) and (2)). The papers published during the last 2 years give the impression that, in several parts of the world, control scientists are dealing with the application of modern identification methodology to traffic state and parameter estimation problems with respect to traffic flow control tasks (cf. Figure 1, Level II), and that they are focusing on the development of decentralized algorithms that can be implemented by spatially distributed control systems with microprocesses [35, 54].
- On the other hand, it is interesting to note that identification problems at the first level (cf. Figure 1) have not yet had much attention from control scientists (cf. [11, 76]), though, as shown here, the identification of macroscopic input-output models could play an important role for certain high level traffic control problems, e.g. for the creation of route guidance systems.

Therefore, it seems to be useful to complete the general survey presented above by a special case study on the identification of macroscopic input-output models.

PART II: IDENTIFICATION OF DYNAMIC INPUT-OUTPUT
MODEL PARAMETERS: A CASE STUDY

THE MODELS AND IDENTIFICATION METHODS STUDIED

The Aim of the Case Study

The knowledge of models (31) and (41) described earlier presents very valuable information on the distribution of traffic streams within a traffic network, and on the corresponding mean travel time. If it were possible to determine the model parameters, $g(s)$ or $g_{ij}(s)$, automatically, by a computer coupled via traffic volume detectors with the street network, then a very valuable tool for the solution of several traffic control and guidance problems would have been obtained (cf. the last section of Part I).

The purpose of this Part is to investigate if this identification task can be solved, and how accurate it would be for conditions near to real traffic situations.

Introduction of Stationary Input and Output Signals and of Modified Model Structures

Whether parameter estimation algorithms may be considered as optimal depends mainly on the statistical properties of the noise signals and the form of the mathematical model (cf. $x_{Bz}(k)$ in equations (31) and (41)). Figure 7 shows a stochastic disturbance, $x_{Bz}(k)$, obtained at the so-called North-South-Connection in Dresden (Figure 5). This disturbance is caused by traffic entering the route via intersections located between the two traffic detectors shown in Figure 5. The mean value, \bar{x}_{Bz} , of the disturbance is, of course, larger than zero. The same is true for the mean values, \bar{x}_{Be} and \bar{x}_{Ba} , of the input and output traffic volumes (cf. Figure 7). Moreover, these mean values change in the course of the day. There are two possible ways of avoiding the application of nonstationary stochastic signals in an identification algorithm:

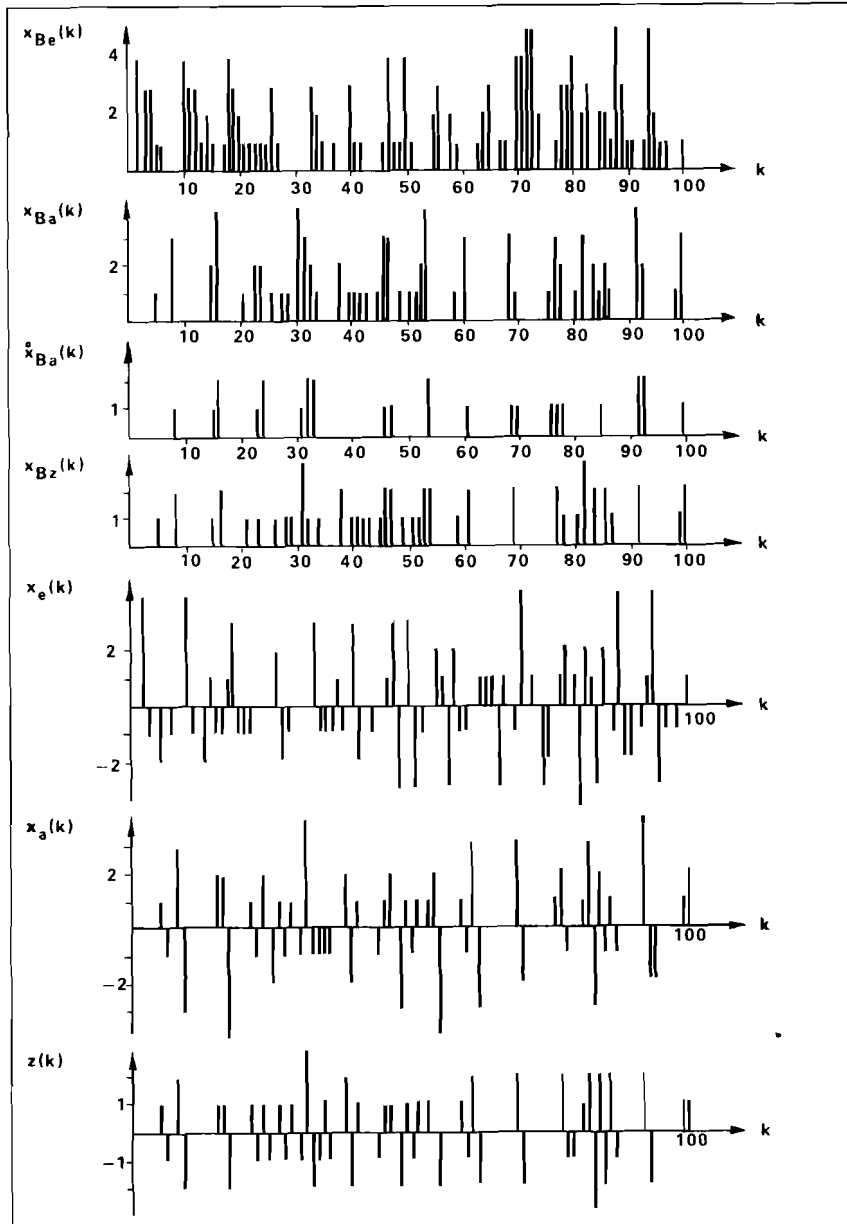


Figure 7. Traffic volumes, x_{Be} , x_{Ba} , $\overset{\circ}{x}_{Ba}$, and x_{Bz} , and input, x_e , output, x_a , and noise, z , signals obtained for the Dresden N-S-Connection (cf. Figure 5).

- to use the differences between the volumes and their mean values according to

$$\begin{aligned}\tilde{x}_e(k) &= x_{Be}(k) - \bar{x}_{Be} \\ \tilde{x}_a(k) &= x_{Ba}(k) - \bar{x}_{Ba} \\ \tilde{z}(k) &= x_{Bz}(k) - \bar{x}_{Bz} \quad ,\end{aligned}\tag{46}$$

or--what is more convenient for real time computations--

- to use the differences between the volumes at times $k\Delta t$ and $(k - 1)\Delta t$, i.e.

$$\begin{aligned}x_e(k) &= x_{Be}(k) - x_{Be}(k - 1) \\ x_a(k) &= x_{Ba}(k) - x_{Ba}(k - 1) \\ z(k) &= x_{Bz}(k) - x_{Bz}(k - 1) \quad .\end{aligned}\tag{47}$$

For the output signals $\tilde{x}_a(k)$ and $x_a(k)$, one now obtains instead of equation (31) the new relation

$$\begin{aligned}\tilde{x}_a(k) &= g(m)\tilde{x}_e(k - m) + \dots + g(n)\tilde{x}_e(k - n) + \tilde{z}(k) \\ &= \tilde{x}_{aM}(k) + z(k) \quad ,\end{aligned}\tag{48}$$

and (with certain approximations)

$$\begin{aligned}x_a(k) &= g(m)x_e(k - m) + \dots + g(n)x_e(k - n) + z(k) \\ &= x_{aM}(k) + z(k) \quad .\end{aligned}\tag{49}$$

In the same manner, one obtains from equation (41) the new multivariable model of a street network*:

*For convenience, here and in the following equations, a restriction is made to the variables (47).

$$x_{ai}(k) = x_{aMi}(k) + z_i(k) = \sum_{j=1}^r \sum_{s=m_{ij}}^{n_{ij}} g_{ij}(s)x_{ej}(k-s) + z_i(k) \quad (50)$$

for $i = 1(1)\ell$.

The noise terms, $\tilde{z}(k)$ and $z(k)$, may now be considered as stationary stochastic discrete signals that are, moreover, approximately uncorrelated for $k\Delta t$ and $k\Delta t + s\Delta t$. This is illustrated by the correlation functions $\psi_{zz}(\tau)$ and $\psi_{x_{Bz}x_{Bz}}(\tau) = \psi_{zz}(\tau) + \bar{x}_{Bz}^2$ shown in Figure 8. The relatively large negative value of

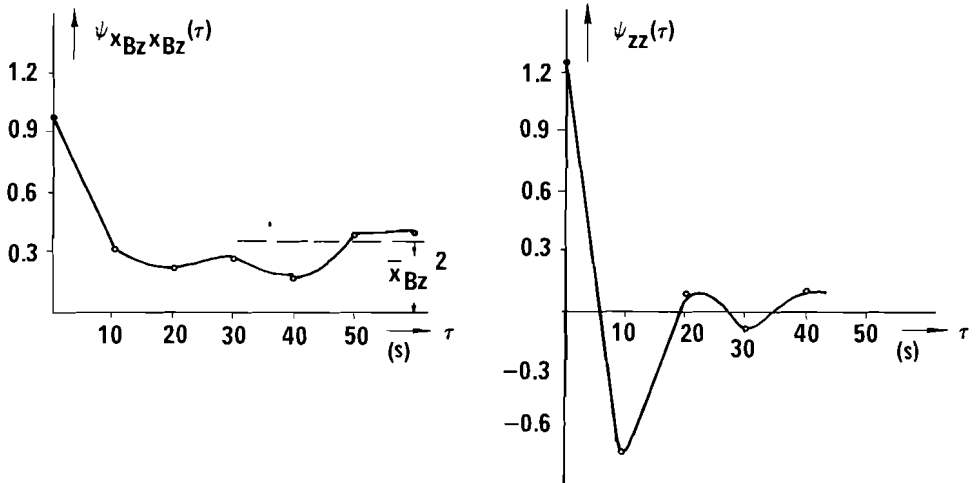


Figure 8. Estimates of the correlation functions of the noise signals, $x_{Bz}(k)$ and $z(k)$ (cf. Figure 7).

$\psi_{zz}(\tau)$ at $\tau = \pm\Delta t$ is caused by computing $z(k)$ with the neighboring values $x_{Bz}(k)$ and $x_{Bz}(k-1)$ taken from equation (47). In spite of this, the correlation between $z(k)$ and $z(k+s)$ may be considered as low, at least for $|s| > 1$, and so the application of a minimum squares estimation technique seems to be justified if the noise amplitudes $z(k)$ are normally distributed. This is, of course, not the case, as can be seen from the frequency distribution $f(z)$ (Figure 9) computed by means of the disturbance $z(k)$ (Figure 7), obtained from real traffic measurements (cf.

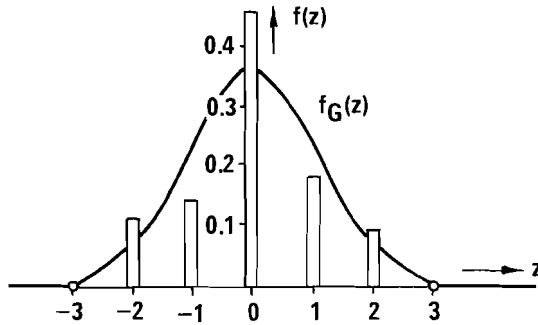


Figure 9. Empirically determined frequency distribution of the noise magnitudes, $z(k)$, compared with the Gaussian distribution, $f_G(z)$.

Part I, Conclusions). Nevertheless the principal form of this distribution $f(z)$ and the Gaussian distribution are similar (Figure 9). Therefore, it has been decided not to use the maximum-likelihood method, but to prefer least squares estimation techniques. This decision also seems reasonable from the viewpoint that the chosen estimation algorithm should be applicable with acceptable storage and computing time requirements in a real-time process computing system.

The following algorithm has been taken into account:

The Explicit Methods Used

For the one-dimensional model (49), one obtains with the error equation system

$$\begin{bmatrix} \varepsilon(k) \\ \cdot \\ \cdot \\ \varepsilon(k-N) \end{bmatrix} = \begin{bmatrix} x_a(k) \\ \cdot \\ \cdot \\ x_a(k-N) \end{bmatrix} - \begin{bmatrix} x_e(k-m) & \dots & x_e(k-n) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_e(k-m-N) & \dots & x_e(k-n-N) \end{bmatrix} \cdot \begin{bmatrix} g(m) \\ \cdot \\ \cdot \\ g(n) \end{bmatrix} \quad (51)$$

$$\hat{\underline{\varepsilon}} = \underline{\underline{x}}_a - (U) \cdot \underline{\underline{b}} ,$$

and the minimization criterion

$$Q = \underline{\underline{\varepsilon}}^T \underline{\underline{\varepsilon}} = \min , \quad (51a)$$

the well-known normal equation system

$$(U)^T(U)\hat{\underline{b}} = (U)^T\underline{x}_a = (A)\hat{\underline{b}} = \underline{a} \quad (52)$$

for the estimate $\hat{\underline{b}}$, where the elements, a_{ij} and a_i , of the matrix (A) and the vector \underline{a} , respectively, are given by

$$a_{ij} = \sum_{v=0}^N x_e(k - i - v)x_e(k - j - v) \quad (53)$$

$$a_i = \sum_{v=0}^N x_e(k - i - v)x_e(k - j - v) .$$

These elements are for stationary signals approximately proportional to the correlation functions

$$\begin{aligned} \psi_{x_e x_e}(s) &\approx \frac{1}{N-s} \sum_{k=1}^{N-s} x_e(k)x_e(k+s) , \\ \psi_{x_e x_a}(s) &\approx \frac{1}{N-s} \sum_{k=1}^{N-s} x_e(k)x_a(k+s) . \end{aligned} \quad (54)$$

This leads to the equation

$$\begin{bmatrix} \psi_{x_e x_e}(0) & \dots & \psi_{x_e x_e}(m-n) \\ \vdots & & \vdots \\ \psi_{x_e x_e}(n-m) & \dots & \psi_{x_e x_e}(0) \end{bmatrix} \cdot \begin{bmatrix} \hat{g}(m) \\ \vdots \\ \hat{g}(n) \end{bmatrix} = \begin{bmatrix} \psi_{x_e x_a}(m) \\ \vdots \\ \psi_{x_e x_a}(n) \end{bmatrix} \hat{=} (\tilde{A})\hat{\underline{b}} = \tilde{\underline{a}} , \quad (55)$$

which is well-known for its application to the estimation of discrete impulse response values by correlation analysis [70], i.e. the convolution integral (38), and the resulting Wiener-Hopf's integral equation (cf. (67)).

It is to be expected, and will be shown in the following, that both estimation algorithms, i.e. equations (52) and (55), deliver similar estimation results both for the use of relationship (47), and for stationary signals generated by equation (46).

The Implicit (Recursive, Adaptive) Methods Used

The intended application of a process computer requires recursive computing techniques. In general, recursive estimation algorithms use the relation

$$\hat{\underline{b}}(k + 1) = \hat{\underline{b}}(k) + \underline{\Gamma}(k) \varepsilon(k + 1) \quad , \quad (56)$$

with

$$\varepsilon(k + 1) = \underline{x}_a(k + 1) - \underline{u}^T(k + 1) \hat{\underline{b}}(k) \quad , \quad (56a)$$

and

$$\underline{u}^T(k) = (\underline{x}_e(k - m) \dots \underline{x}_e(k - n)) \quad , \quad (56b)$$

i.e. the $k + 1$ estimate $\hat{\underline{b}}(k + 1)$ is determined by the k th estimate $\hat{\underline{b}}(k)$, plus a certain correction term containing the model error $\varepsilon(k + 1)$, and a weighting vector $\underline{\Gamma}(k)$. The latter can be calculated in the case of the recursive least squares method by the formula

$$\underline{\Gamma}(k) = (C[k + 1]) \underline{u}(k + 1) \quad , \quad (57)$$

with

$$C[k + 1] = \frac{1}{w^2} \left\{ C[k] - \frac{(C[k]) \underline{u}(k + 1) \underline{u}^T(k + 1) (C[k])}{w^2 + \underline{u}^T(k + 1) (C[k]) \underline{u}(k + 1)} \right\} \quad , \quad (57a)$$

and $w = 1$. For $w < 1$, equation (57a) corresponds to the recursive least squares method with "exponential forgetting". This uses the minimization criterion

$$Q(k + 1) = \sum_{i=1}^{k+1} \varepsilon^2(i) w_i^2 = \varepsilon^2(k + 1) w_{k+1}^2 + \dots + \varepsilon^2(1) w_1^2 = \min \quad , \quad (58)$$

with $w_i^2 = w^{2[k+1-i]}$, and $w < 1$, (58a)

i.e. $w_{k+1}^2 = w^0 = 1, w_k^2 = w^2, \dots, w_1^2 = w^{2k}$. (58b)

Criterion (58) makes it possible that the latest measured values, $x_a(k+1), x_a(k), \dots$ etc., will influence the estimated result more than the older ones, i.e. $x_a(1), x_a(2), \dots$. Figure 10 shows this "forgetting" factor w_i^2 for $0.950 \leq w \leq 0.999$. One can see that for $w = 0.980$, for example, only the last 20 measured values influence criterion (58) with a weight greater than 0.4, while this is true for the last 110 measured values for $w = 0.996$.

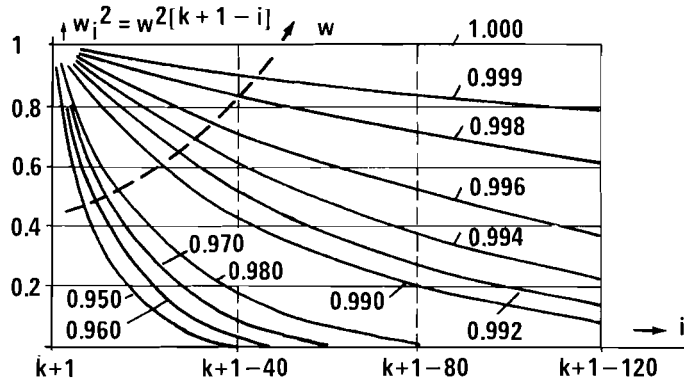


Figure 10. The "forgetting" factor, w_i^2 .

This property of "forgetting" old measured values is obviously very important for nonstationary traffic conditions with changing travel times, T_{RM} , and parameters, $g(s)$.

Besides the recursive least squares method, the following simplified algorithms need to be taken into consideration [70]:

- the stochastic approximation, with

$$\underline{\Gamma}(k) = \frac{c_1}{k+1} \underline{u}(k+1), \quad (59)$$

- the relaxation method, with

$$\underline{\Gamma}(k) = c_2 \frac{\underline{u}(k+1)}{\underline{u}^T(k+1) \underline{u}(k+1)}, \quad (60)$$

and

- the so-called quick and dirty regression, with

$$\underline{\Gamma}(k) = c_3 \underline{u}(k+1) \left\{ \frac{1}{\sum_{r=1}^{k+1} \underline{u}^T(r) \underline{u}(r)} \right\}. \quad (61)$$

It has been reported that these methods require remarkably less computing time than the least squares method: 8 times less for the first algorithm, 5 times for the second, and 3 times for the third (cf. [70, 91]).

The methods summarized here will be studied in the following paragraphs for:

- Stationary traffic conditions, i.e. the model parameters $g(s)$ and $g_{ij}(s)$ of equations (49) and (50) are considered as time invariant.
- Nonstationary traffic conditions with parameters $g(s)$ and $g_{ij}(s)$ varying, i.e. the mean travel times are supposed to be changed by accidents, changing weather conditions, etc.

First, the studies are carried out with simulated traffic processes. The results obtained are then compared with those from an experiment carried out under real traffic conditions in the North-South-Connection in Dresden (cf. Figure 5).

SIMULATION STUDIES

Stationary Traffic Considerations

The measured real traffic volumes $x_{Be}(k)$, shown in Figure 7, are used in the following studies also as input volumes, in order to have simulated traffic conditions similar to real ones. With the use of $x_{Be}(k)$ and an uncorrelated discrete noise signal, $x_{Bz}(k)$, input, $x_e(k), \tilde{x}_e(k)$, and output, $x_a(k), \tilde{x}_a(k)$, signals

have been determined by equations (46)-(49). To describe the noise/signal ratio, the coefficient

$$vz = \frac{\bar{x}_{Bz}}{\bar{x}_{Ba}} \cong \frac{\text{sum of cars entering the route between input and output}}{\text{sum of cars flowing from input to output}} \quad (62)$$

is used. For the travel times, a probability distribution as shown in Figure 11 is considered valid.

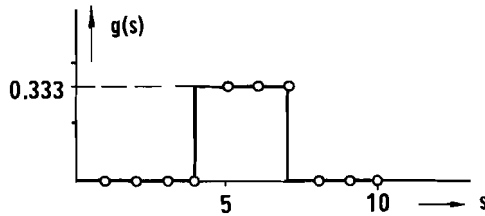


Figure 11. Travel time distribution used in simulation studies.

Explicit Methods

Table 1 summarizes the results obtained by explicit methods. The following conclusions may be drawn:

- There is no significant difference between using equation (46) and (47) for the creation of stationary signals (cf. rows Z and D in Table 1), so equation (47) is used in the following.
- As expected, there are no remarkable differences in the estimation of the mean travel time, T_{RM} , and the split coefficient, $h(n)$, if one uses correlation analysis (cf. equation (55)) instead of the least squares method (cf. equation (52), and examples 1.1 and 2.1, with 1.2 and 2.2 in Table 1).
- For the undisturbed system, with $vz = 0$ representing a street section without intersections between input

detector and output detector, it is possible to obtain very accurate estimates for the impulse response values, $g(s)$, as well as for the mean travel time and the split coefficient (cf. Table 1).

Table 1. Application of explicit identification methods (LS, least squares estimation; CA, correlation analysis; Z, equation (46) used; D, equation (47) used).

			s = 4	s = 5	s = 6	s = 7	s = 8	$T_{RM/s}$	$h(n) = \frac{n}{\sum_{s=m}^n g_s}$	
True values			$g(s)$ 0	0.333	0.333	0.333	0	60.0	1.00	
			$f(s)$ 0	0.333	0.333	0.333	0			
Street section $v_z = 0$	1.1	Z	$\hat{g}(s)$ 0.00	0.32	0.33	0.30	0.02	60.2	0.97	
		D	$\hat{f}(s)$ 0.00	0.33	0.34	0.31	0.02	60.3	0.99	
	1.2	Z	$\hat{g}(s)$ -0.01	0.33	0.35	0.29	0.02	60.0	0.98	
		D	$\hat{f}(s)$ 0	0.33	0.35	0.29	0.02	60.4	1.03	
	Route with inter-sections $v_z = 1$	2.1	Z	$\hat{g}(s)$ 0.10	0.31	0.28	0.35	-0.02	58.3	1.02
			D	$\hat{f}(s)$ 0.10	0.30	0.27	0.33	0	58.3	0.94
2.2		Z	$\hat{g}(s)$ 0.09	0.30	0.26	0.32	-0.03	59.1	0.98	
		D	$\hat{f}(s)$ 0.09	0.31	0.27	0.33	0	59.1	1.00	
		CA	Z	$\hat{g}(s)$ 0.06	0.32	0.28	0.35	-0.03	59.1	0.98
			D	$\hat{f}(s)$ 0.06	0.32	0.28	0.35	0	59.1	1.00
			$\hat{g}(s)$ 0.07	0.32	0.27	0.37	-0.03	59.1	1.00	
			$\hat{f}(s)$ 0.07	0.31	0.27	0.36	0			

- For the large noise/signal ratio of $v_z = 1$, i.e. the number of cars entering the route via intersections or on-ramps between input and output is equal to the number of cars flowing from the input to the output, large estimation errors occur for $g(s)$ values. If one normalizes them with respect to the mean value

$$\bar{\hat{g}} = \frac{1}{n - m + 1} \sum_{s=m}^n \hat{g}(s) \quad , \quad (63)$$

then the relative error takes values within the interval $-35\% \leq \Delta\hat{g}(s)/\bar{\hat{g}} \leq 50\%$, where negative estimates, $\hat{g}(s)$, occur for $g(s) = 0$ at $s = 4$ and $s = 8$. Since a negative frequency or probability does not have any physical meaning, these negative estimates of $\hat{g}(s)$ cannot be used for the calculation of $f(s)$, and the mean travel time, T_{RM} , has to be determined by means of the non-negative estimates of $\hat{g}(s)$ only:

$$\hat{f}(s) = \begin{cases} 0 & \text{for } \hat{g}(s) < 0 \\ \hat{g}(s) / \sum_{s=m}^n \hat{g}(s) & \text{for } \hat{g}(s) \geq 0 \end{cases} \quad (64)$$

$$\hat{T}_{RM} = \sum_{s=m}^n (s\Delta t) \hat{f}(s) \quad . \quad (64a)$$

On the other hand, there is no reason to neglect the negative estimates of $\hat{g}(s)$ in the estimate

$$\hat{h}(n) = \sum_{s=m}^n \hat{g}(s) = [n - m + 1] \bar{\hat{g}} \quad (65)$$

for the split coefficient $\hat{h}(n)$. If one assumes that the estimations, $\hat{g}(s)$, are unbiased, i.e. $E\{\Delta\hat{g}(s)\} = 0$, and that the estimation errors, $\Delta\hat{g}(s)$, are nearly uncorrelated, then the determination of the mean value, $\bar{\hat{g}}$, by equation (63) results in a certain smoothing effect of the statistical errors, $\Delta\hat{g}(s)$. Therefore one may expect that the estimate of the split coefficient obtained by equations (62) and (65) is more accurate than the estimates of the impulse response values, $g(s)$. This statement holds true for examples 2.1 and 2.2 where, in spite of large estimation errors, $\Delta\hat{g}(s)$, the relative errors of the split coefficient lie

within the relatively small interval $-6\% \leq \Delta \hat{h}(n) \leq 3\%$. Even more accurate estimates have been obtained by equations (64) and (64a) for the mean travel time with $-2.8\% \leq \Delta \hat{T}_{RM}/T_{RM} \leq -1.5\%$ (cf. Table 1).

This interesting result is characterized by small travel time errors, in spite of large errors of the impulse response values, and the following explanation can be given: If one assumes that all cars need the same time, \tilde{T}_R , for traveling between input and output, then one would get a response for the route in the form of a Dirac-impulse with an area equal to the split coefficient $h(n)$:

$$g(t) = h(n) \delta(t - \tilde{T}_R) \quad . \quad (66)$$

Substituting $g(t)$ in the Wiener-Hopf integral equation

$$\psi_{x_e x_a}(\tau) = \int_0^{\infty} g(\theta) \psi_{x_e x_e}(\tau - \theta) d\theta \quad , \quad (67)$$

one finds the very simple relation

$$\psi_{x_e x_a}(\tau) = h(n) \psi_{x_e x_e}(\tau - \tilde{T}_R) \quad (68)$$

between auto- and cross-correlation functions illustrated by Figure 12. To estimate the travel time, \tilde{T}_R , it is obviously sufficient to know only rough estimates of the cross-correlation function $\psi_{x_e x_a}(\tau)$. It is completely sufficient to know the position of the maximum of $\psi_{x_e x_a}(\tau)$, which in general can be determined reliably with just a small number, N , of measured values of $x_e(k)$ and $x_a(k)$, e.g. $N \approx 100$ as in the examples of Table 1.

The explanation given here for an idealized traffic situation holds true also for a real one, as was proved by estimation of the travel time in the way shown in Figure 12 for several sets of data obtained at real traffic processes.

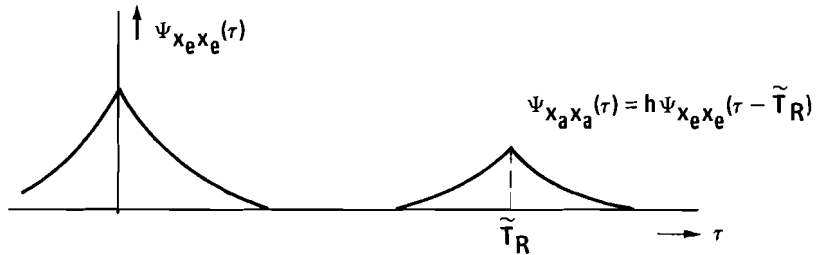


Figure 12. Determination of a rough estimate of the travel time by means of the cross-correlation function.

Implicit Methods

The implicit methods characterized by equations (56)-(61) are studied under the same conditions chosen for example 2.1 Z of Table 1 for the investigation of explicit methods. The results obtained after 100 iterations for the initial estimates ($\hat{g}_0(s) = 0.2$ for $s = 5, 6,$ and $7,$ and $\hat{g}_0(s) = 0$ for $s = 4$ and 8) are summarized in Table 2.

Table 2. Estimates of $\hat{g}(s), \hat{f}(s), \hat{T}_{RM},$ and $\hat{h}(n)$ obtained by means of implicit methods (cf. equations (56)-(61)).

		s = 4	s = 5	s = 6	s = 7	s = 8	$\hat{T}_{RM/s}$	$\hat{h}(n)$
True Values	$\hat{g}(s)$	0.000	0.333	0.333	0.333	0.000	60.0	1.00
	$\hat{f}(s)$	0.000	0.333	0.333	0.333	0.000		
Stochastic Approx.	$\hat{g}(s)$	0.00	0.30	0.31	0.40	-0.09	61.6	0.92
	$\hat{f}(s)$	0.00	0.30	0.31	0.40	0		
Relaxation	$\hat{g}(s)$	0.08	0.31	0.26	0.29	-0.08	58.6	0.86
	$\hat{f}(s)$	0.09	0.33	0.28	0.31	0		
Quick and Dirty Reg.	$\hat{g}(s)$	-0.03	0.40	0.38	0.11	-0.01	56.7	0.85
	$\hat{f}(s)$	0	0.45	0.43	0.12	0		
Recursive Regression	$\hat{g}(s)$	0.10	0.31	0.28	0.35	-0.02	58.3	1.02
	$\hat{f}(s)$	0.10	0.30	0.27	0.33	0		

The following conclusions may be drawn:

- The recursive minimum squares estimation method gives, of course, the same results as the explicit version of this method (cf. equation (52) and Table 1).
- The stochastic approximation provided the sufficiently accurate estimates shown in Table 2 for a special value of c_1 in equation (59) only. By several trials, the optimal value, $c_1 = 0.005$ depending on the statistical properties of the noise signal $z(k)$, was found. It appeared that even small deviations from this value result in serious convergence problems.
- The same is true for the coefficients c_2 and c_3 of the relaxation method and the quick and dirty regression (cf. equations (60) and (61)). Small changes of the empirically determined optimal values $c_2 = 0.1$ and $c_3 = 0.01$, for which the results are given in Table 2, led to quite different estimates of $\hat{g}(s)$, \hat{T}_{RM} , and $\hat{h}(n)$.

Since it seems not possible to precompute reliable values of c_1 , c_2 , and c_3 by means of a priori information available from real traffic measurements, the practical applicability of equations (59)-(61) is very uncertain. The following studies are therefore restricted to the application of the recursive regression where the identification of the multivariable model (50) is first investigated.

Identification of Multivariable Models

For the street network shown in Figure 6, for example, one obtains for the output volume, x_{Ba1} , and the input volumes, x_{Be2} , x_{Be3} , and x_{Be4} , the model equation

$$\begin{aligned}
 x_{a1}(k) = x_a(k) = & \sum_{s=m_{12}}^{n_{12}} g_{12}(s)x_{e2}(k-s) \\
 & + \sum_{s=m_{13}}^{n_{13}} g_{13}(s)x_{e3}(k-s) + \sum_{s=m_{14}}^{n_{14}} g_{14}(s)x_{e4}(k-s) + z_1(k) \quad , \quad (69)
 \end{aligned}$$

if one uses the differences $x_{ei}(k) = x_{Bei}(k) - x_{Bei}(k - 1)$ and $x_{a1}(k) = x_{Ba1}(k) - x_{Ba1}(k - 1)$ as the necessary stationary signals (cf. equation (47)).

For this model, the vectors \underline{u} and \underline{b} from equations (56a,b) take the forms

$$\underline{u}^T(k) = (x_{e2}(k-m_{12}) \dots x_{e2}(k-n_{12}) \ x_{e3}(k-m_{13}) \dots x_{e3}(k-n_{13}) \ x_{e4}(k-m_{14}) \dots x_{e4}(k-n_{14})) \quad (70)$$

and

$$\underline{b}^T = (g_{12}(m_{12}) \dots g_{12}(n_{12}) \ g_{13}(m_{13}) \dots g_{13}(n_{13}) \ g_{14}(m_{14}) \dots g_{14}(n_{14})) \quad , \quad (71)$$

where the true values of $g_{1j}(s)$ shown in Figure 13 have been chosen corresponding to the true travel times $T_{RM12} = 60.0$ seconds, $T_{RM13} = 45.0$ s, $T_{RM14} = 30.0$ s and the true split coefficients $h_{12}(n_{12}) = h_{13}(n_{13}) = h_{14}(n_{14}) = 1.00$.

The following examples have been studied (cf. Table 3):

Example 1: The noise/signal ratio, vz , of equation (62) can be reduced to zero if one measures the whole traffic entering a route on a network by volume detectors. For the network studied here (cf. Figure 6), such a situation occurs if there are no traffic input volumes other than the 4 variables x_{Be1} , x_{Be2} , x_{Be3} , and x_{Be4} . If one assumes that the identification of the mean travel time, T_{RM12} , and the split coefficient, $h_{n12}(n_{12})$, is of primary interest, then the use of additional measurements for x_{Be3} and x_{Be4} should result in a more accurate estimate. As shown in column 1 of Table 3 and the estimates, $\hat{g}_{1j}(s)$, marked by crosses in Figure 13, it is actually possible in such a case to obtain unknown parameters with very small errors.

Example 2: The experiment of example 1 is repeated for an undisturbed two-dimensional model, i.e. with $g_{14}(s) = 0$, resulting again in very accurate estimates (cf. Table 3 and example 2a of Figure 13). Next it is assumed that no measurements are taken for x_{Be3} , i.e. the number, $x_{Be3}(k)$, of cars entering the network

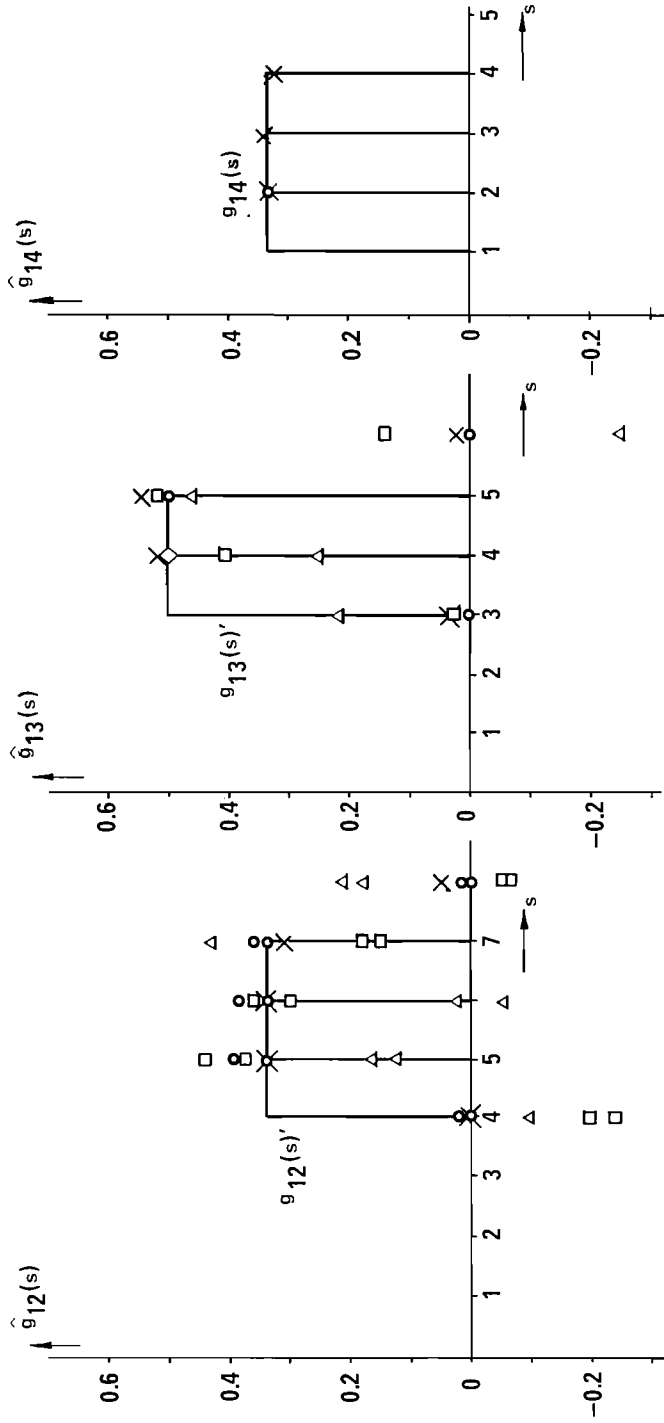


Figure 13. Estimates of $\hat{g}_{\gamma_j}(s)$ for $v_z = 0$ (examples 1[x], 2a[o], 2b[o]), $v_z = 1, 2$ (examples 3a[□], 3b[□]), and $v_z = 2, 3$ (examples 4a[△], 4b[△]) (cf. Table 3).

at input 3 is now playing the role of a disturbance (example 2b of Figure 13). A comparison of the estimates obtained for these two examples from Figure 13 illustrates (cf. Table 3) that the additional disturbance caused by the non-measured input volume, x_{Be3} , leads to much larger estimation errors of the impulse response values, $\hat{g}_{12}(s)$, and the split coefficients, $\hat{h}_{12}(n_{12})$. On the other hand, the estimation error for the mean travel time \hat{T}_{RM12} remains small (only -0.98%). Thus one may conclude that the algorithm used is capable of withstanding disturbances in travel time estimation. This statement holds true if additional disturbances occur (see examples 3 and 4 below).

Table 3. Estimated travel times, T_{RMij} , and split coefficients, $h_{ij}(n_{ij})$, for the simulated traffic network shown in Figure 6.

Example	1	2a	2b	3a	3b	4a	4b
vz	0	0	0	1.2	1.2	2.3	2.3
\hat{T}_{RM12}	60.55	60.01	59.51	57.55	57.10	69.55	68.08
\hat{T}_{RM13}	44.95	45.00	-	47.02	-	42.40	-
\hat{T}_{RM14}	30.03	-	-	-	-	-	-
δT_{12} (%)	+0.92	+0.01	-0.98	-4.3	-4.8	+15.9	+13.5
δT_{13} (%)	-0.11	0.00	-	+4.5	-	-5.8	-
δT_{14} (%)	+0.10	-	-	-	-	-	-
$h_{12}(n_{12})$	1.014	0.997	1.158	0.521	0.705	0.646	0.728
$h_{13}(n_{13})$	1.103	0.997	-	1.094	-	0.733	-
$h_{14}(n_{14})$	0.994	-	-	-	-	-	-
δh_{12} (%)	+1.4	-0.3	15.8	-48.0	-29.5	-35.4	-27.2
δh_{13} (%)	+10.3	-0.3	-	+9.4	-	-26.7	-
δh_{14} (%)	-0.6	-	-	-	-	-	-
w	0.996	0.995	0.995	0.995	0.995	0.997	0.997

Examples 3 and 4: Here the same two-dimensional model is used as for the preceding example, but, instead of $v_z = 0$, a noise/signal ratio of $v_z = 1.2$ (example 3a, 3b) and $v_z = 2.3$ (example 4a, 4b) is assumed, i.e. the number of cars entering the network unobserved and passing output 1, is 1.2 times (example 3a, b), or 2.3 times (example 4a, b), the number of cars going from input 2 and input 3 to output 1 (cf. Figure 6). Also, under these very complicated circumstances, it was possible to obtain sufficiently accurate travel time estimates even when large errors in the estimates of the split coefficients, and even larger ones in the impulse response values, were occurring (cf. Figure 13 and Table 3).

In summary, one may conclude that the method studied will lead to robust and reliable travel time estimates, even under complicated traffic conditions, i.e. if many cars are entering the analysed route or network unobserved. But larger estimation errors for the split coefficients are to be expected under these conditions.

Nonstationary Traffic Conditions

It will be assumed in the following that changes of travel time may occur. The question is: How fast and accurately can these changes be identified? As is well-known [70], any adaptive identification method implies a compromise between the "quickness" and the "accuracy" of the identification process. For the recursive regression method of equations (56)-(57a), this compromise has to be made by a suitable choice of the forgetting factor, w (cf. Figure 10). A fast reaction to parameter changes requires a small value of w ; a high accuracy requires a large one that deviates only slightly from the maximum value, $w = 1$. To study the problems of choosing a proper forgetting factor is, therefore, the purpose of the investigations presented here. These are carried out for the extreme case of a sudden increase in travel time from $T_{RM} + 35$ s for $t < 0$, to $T_{RM} = 75$ s for $t \geq 0$, i.e. for a sudden change of the impulse response (or travel time distribution) values, $g(s)$, from

$$g(s) = \begin{cases} 0.5 & \text{for } s = 3, 4 \\ 0 & \text{for } s < 3, s > 4 \end{cases} \quad \text{to} \quad g(s) = \begin{cases} 0 & \text{for } s = 7, 5, 8 \\ 0.5 & \text{for } s = 7, 8 \end{cases}$$

(cf. Figures 14 and 15). The following examples are considered.

- Road sections without intersections between input and output; i.e. $v_z = 0$. The curves obtained for the estimates $\hat{g}_k(s)$ in Figure 14 and $\hat{T}_{RM}(0, k)$ in Figure 15 illustrate that for $0.986 \leq w \leq 0.994$ the beginning of a travel time change can be discovered after 20 iterations. After 80 iterations the remaining relative travel time errors are 4% for the smaller value $w = 0.986$, and 11% for the larger $w = 0.994$ (cf. Figure 16, $v_z = 0$). For this important special traffic process a forgetting factor of $w = 0.986$ or even a little smaller could be considered as a suitable choice.

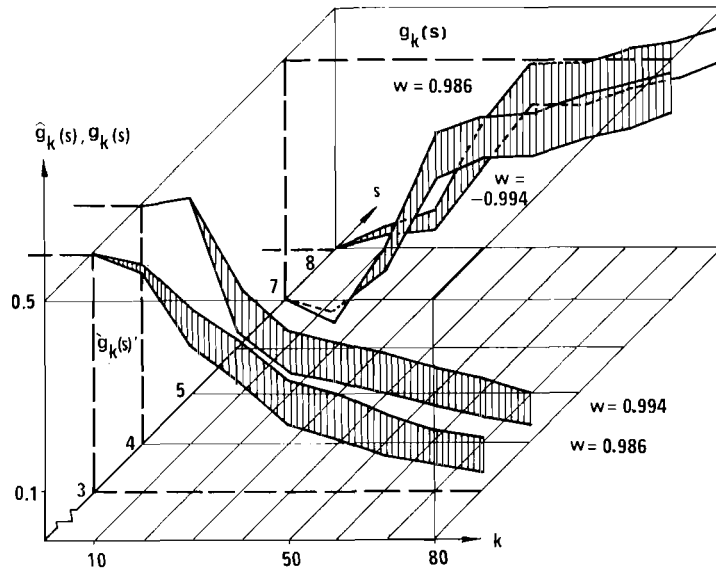


Figure 14. Estimates, $\hat{g}_k(s)$, of the changing parameters $g_k(s)$ for $v_z = 0$.

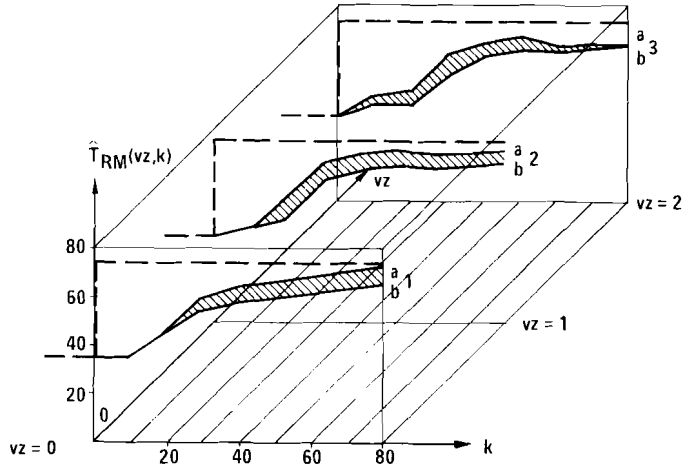


Figure 15. Estimates, $\hat{T}_{RM}(k, vz)$, of the changing travel time, T_{RM} , for $vz = 0; 1; 2$ and $w = 0.986$ {a} and $w = 0.984$ {b} .

- Traffic routes with intersections or on-ramps between input and output, i.e. $vz = 1$ and $vz = 2$. For these, the "step responses" of the travel time identification algorithm are shown in Figure 15 (curves 2 and 3), while Figure 16 illustrates the dependence of the travel-time error, $\delta T_{RM} = \Delta \hat{T}_{RM} / T_{RM}$ (values \square and Δ), and the mean impulse response error,

$$\delta g = \Delta \bar{g} / \bar{g} = (1/\bar{g}) \sqrt{(1/(n - m)) \sum_{s=m}^n [\hat{g}(s) - g(s)]^2}$$

(values \square and Δ), of the forgetting factor w . One observes that for large noise/signal ratios ($vz = 2$), small values of w may no longer be considered as preferable. Nevertheless, a choice of a value of w within $0.986 \leq w \leq 0.994$, will very likely result in acceptable estimates for a wide variety of noise/signal ratios, i.e. for very different traffic conditions.

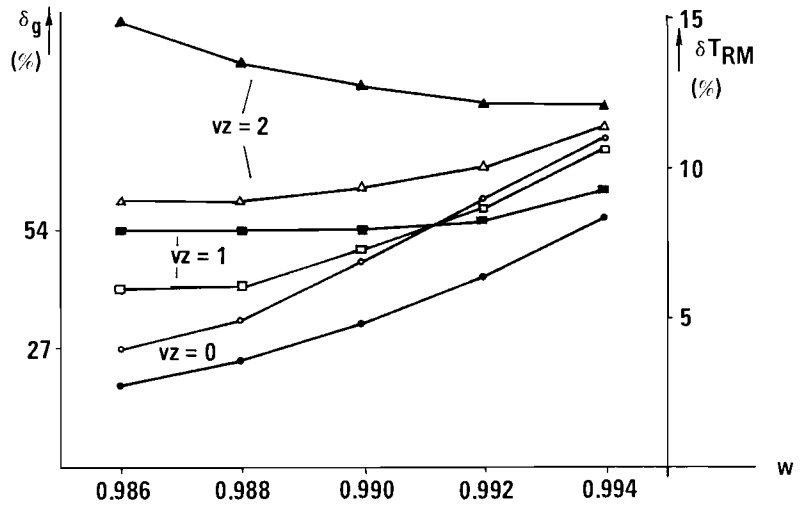


Figure 16. Relative estimation errors, δT_{RM} and δg , resulting from 80 iterations.

ANALYSIS OF A REAL TRAFFIC PROCESS

The aim of this part of the case study is to check if the presumptions made in the simulation studies may be considered as realistic.

The Process Studied

As a study subject, a 1000 m long part of the so-called North-South-Connection in Dresden (cf. Figure 5) has been chosen, and this has been coupled with a process computer located in the Dresden Hochschule für Verkehrswesen via traffic detectors and telephone lines. Figure 7 shows one of the sets of measured traffic volumes, $x_{Be}(k)$ and $x_{Ba}(k)$, obtained by this experimental installation at sampling intervals of 10 s, and used in the following identification experiments. True values of the mean travel time, T_{RM} , and the split coefficient, $h(n)$, are needed for these experiments, in order to have a basis for judging the accuracy of the estimates. These can be obtained

very accurately by means of a manual off-line method--the so-called licence plate method (cf. [44, 81]). This uses the following simple basic idea: The licence numbers of the cars passing the input and output of the route are visually identified by special measuring personnel, and immediately stored in a suitable form, e.g. on a magnetic tape by means of a telex typewriter. Using stored information for input and output simultaneously, a digital computer determines the travel time distribution, $f(s)$, the mean travel time, T_{RM} , and the split coefficient, $h(n)$, with high accuracy. Moreover, it can compute the number, $\bar{x}_{Ba}^Q(k)$, of cars going from input to output, as well as the number, $x_{Bz}(k)$, of cars entering the route between input and output and leaving via the output. Both volumes are shown in Figure 7, which illustrates that the mean value, $\bar{x}_{Bz}(k)$, of the disturbance, $x_{Bz}(k)$, is about twice as large as the mean value, \bar{x}_{Ba}^Q , of $x_{Ba}(k)$, i.e. the noise/signal ratio of equation (62) is about 2, which is of the same order of magnitude as used in the simulation studies described earlier.

Three intersections with signals are located between the input and output detectors (cf. Figure 5) where, because of certain specialities, the traffic stream in general has to stop once at the middle intersection. One could argue that this stop will cause problems concerning the applicability of the input-output model (49). But the results obtained by the least squares method and presented in Figure 17 and Table 4, illustrate that those doubts are not justified in the traffic system analyzed.

The Results Obtained

One can see that the estimated impulse response values, $\hat{g}(s)$, describe only very roughly the shape of the true values, $g(s)$ (cf. Table 4), while the frequency distribution, $\hat{f}(s)$, determined by equation (64) fits the true values, $f(s)$, much better (cf. Figure 17). This is the reason why the estimate, \hat{T}_{RM} , of the mean travel time can be determined with the extraordinarily small error of 1.6%.

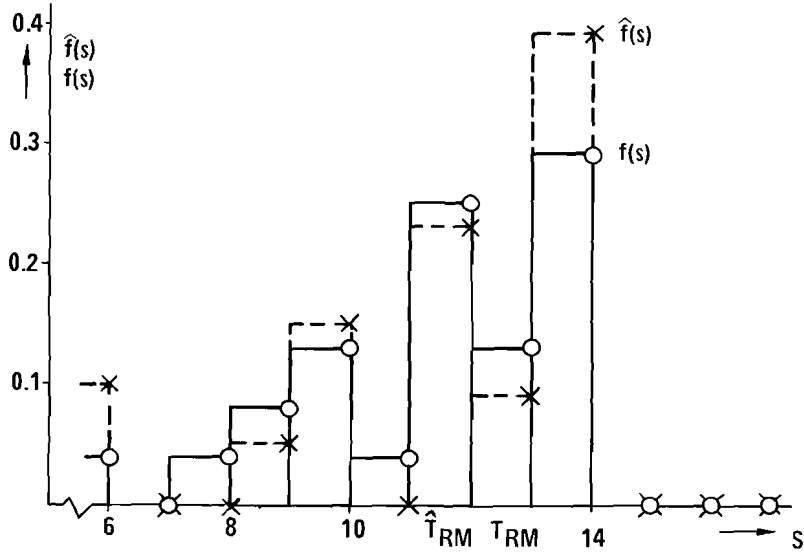


Figure 17. Estimated, $\hat{f}(s)$, and "true", $f(s)$, values of the travel time frequency distribution obtained from traffic volumes measured at the North-South-Connection in Dresden (cf. Figure 5 and Figure 7).

Table 4. True values, $g(s)$ and $f(s)$, and estimates, $\hat{g}(s)$ and $\hat{f}(s)$, obtained from x_{Be} and x_{Ba} shown in Figure 7 for the Dresden North-South-Connection.

		s = 6	s = 7	s = 8	s = 9	s = 10	s = 11	s = 12	s = 13	s = 14	s = 15	T_{RM}	$h(n)$
Exact values	$g(s)$	0.01	0	0.01	0.02	0.03	0.01	0.06	0.03	0.07	0	121s	0.24
	$f(s)$	0.04	0	0.04	0.08	0.12	0.04	0.25	0.12	0.29	0		
Estimates	$\hat{g}(s)$	0.11	-0.07	-0.13	0.05	0.16	-0.29	0.25	0.10	0.43	-0.28	119s	0.33
	$\hat{f}(s)$	0.10	0	0	0.05	0.15	0	0.23	0.09	0.39	0		

From the simulation studies, an error of the order of 5 to 10% can be expected (cf. Table 3, example 4), and will very likely occur if one repeats the identification with different sets of data. For the split coefficient $h(n)$, an estimation error of about 38% was obtained, which is of a similar order of magnitude as in the simulation studies (cf. Table 4 with Table 3 for example 4). To reduce this error, an additional traffic detector should be installed between the input and output of the route as shown in Figure 5. Considering the complex structure of that route, such an investment is obviously reasonable.

Nevertheless, the conclusion that the methodology presented here for the development of macroscopic traffic input-output models and the identification of model parameters, is successfully applicable under real traffic conditions, and delivers valuable information for different classes of traffic control problems (cf. Part I, last section) is justified. Further studies of real traffic processes, e.g. under nonstationary conditions, are, of course, necessary and it is intended to carry these out.

REFERENCES

- [1] Berger, C.R., and L. Shaw, *Diversion Control of Freeway Traffic*, paper presented at the Sixth IFAC World Congress, Boston, Mass., August 1975.
- [2] Bolle, G., System ALI: Verkehrslenkung durch Computer Prognose, *Bild der Wissenschaft*, 4 (1976), 76-82.
- [3] Burhardt, K.K., and R. Kulikowski, Optimum Control of Oversaturated Traffic Systems, *Bulletin de l'Académie Polonaise des Sciences*, XVIII, 7 (1970) 59-75.
- [4] Butrimenko, A., and H. Stobel, On an Algorithm for the Adaptive Automatic Control of the Main Vehicular Traffic Streams in Urban Areas, internal paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1974.
- [5] Chang, M.-F., and D.C. Gazis, *Traffic Density Estimation with Consideration of Lane-Changing*, IBM Watson Research Center, Yorktown Heights, N.Y., 1975.

- [6] Chu, K., Decentralized Control of High-Speed Vehicular Strings, *Transportation Science* (November 1974), 361-384.
- [7] Chu, K., and D.C. Gazis, *Dynamic Allocation of Parallel Congested Traffic Channels*, RC 4545, IBM Watson Research Center, Yorktown Heights, N.Y., 1973.
- [8] Coste, I.-F., M. Liger, and P. Vidieu, L'Operation Ponte du Languedoc, *Revue Generale des Routes et des Aero-dromes*, 475 (1972), 50-62.
- [9] Cremer, M., Modellbildung und Messwertaufbereitung für den Verkehrsablauf auf Schnellstrassen, *Strassenverkehrstechnik*, 20 (1976), 86-92.
- [10] Drew, D.R., *Traffic Flow Theory and Control*, McGraw Hill, N.Y., 1968.
- [11] Doormann, I., Dynamische Modelle in der Strassenverkehrstechnik, *Veröfftl. des Institutes für Stadtbauwesen TH Braunschweig*, 11 (1973).
- [12] Everall, P.F., *Urban Freeway Surveillance and Control: The State-of-the-Art*, Research Report FCP 22E 1613, Department of Transportation, Federal Highway Administration, Washington, D.C., 1972.
- [13] Eykhoff, P., *System Identification*, Wiley, London, 1974.
- [14] Filomeni, P., F. Guavi, and G.M. Secco-Suardo, A New Strategic Approach to Real-Time Traffic Routing, *IEEE Trans. on Systems, Man and Cybernetics*, SMC-3, 1973.
- [15] Friedel, H., Autofahrer-Lenkungs- und Informationssystem, Lenkung der Kfz-Ströme - ALI. *Radio Elektronik Schau*, 5 (1975), 266-269.
- [16] Gazis, D.C., Optimum Control of a System of Over-Saturated Intersections, *Operations Research*, 12 (1964), 815-831.
- [17] Gazis, D.C., Traffic Control: From Hand Signals to Computers, *Proceedings of the IEEE*, 59, 7 (1971), 1090-1099.
- [18] Gazis, D.C., Traffic Flow and Control: Theory and Applications, *American Scientist*, 60, 4 (1972), 414-424.
- [19] Gazis, D.C., ed., *Traffic Science*, Wiley, New York, 1974.
- [20] Gazis, D.C., Modeling and Optimal Control of Congested Transportation Systems, *Networks*, 4 (1974), 113-124.

- [21] Gazis, D.C., and R.F. Foote, Surveillance and Control of Tunnel Traffic by an On-Line Digital Computer, *Transportation Science*, 3 (1969), 255-275.
- [22] Gazis, D.C., R. Herman, and R.B. Potts, Car-Following Theory of Steady-State Traffic Flow, *Operations Res.*, 7 (1959), 499-505.
- [23] Gazis, D.C., R. Herman, and R.W. Rothery, Analytical Methods in Transportation: Mathematical Car-Following Theory of Traffic Flow, *J. Eng. Mechanics Div.*, 89 (1963), 29-46.
- [24] Gazis, D.C., R. Herman, and R.W. Rothery, Nonlinear Follow the Leader Models in Traffic Flow, *Operations Research*, 9 (1961), 545-567.
- [25] Gazis, D.C., and C.H. Knapp, On-Line Estimation of Traffic Densities from Time-Series of Flow and Speed Data, *Transportation Science*, 5, 3 (1971), 283-301.
- [26] Gershwin, S.B., *Dynamic Stochastic Control of Freeway Corridor Systems; Volume II: Steady State Optimal Traffic Assignment Using the Accelerated Gradient Projection Method*, Electronic System Laboratory Report, ESL-R-609, MIT, Cambridge, Mass., 1975.
- [27] Greenberg, H., An Analysis of Traffic Flow, *Operations Res.*, 7, 1 (1959), 79-85.
- [28] Grewal, M.S., *Modeling and Identification of Freeway Traffic Systems*, Ph.D. dissertation, Dept. Elec. Eng., University of Southern California, Los Angeles, 1974.
- [29] Grewal, M.S., G.A. Bekey, and H.J. Payne, Parameter Identification of Dynamical Systems, *Proc. of the 1974 IEEE Conference on Decision and Control*, CH0900-1CS, Phoenix, 1974.
- [30] Grewal, M.S., and H.J. Payne, Identification of Parameters in a Freeway Traffic Model, *IEEE Trans. on Systems, Man and Cybernetics*, SMC-6 (1976), 176-185.
- [31] Groth, G., and O. Pilsack, An Electronic System for Individual and Active Route Guidance on Motorways, [78].
- [32] Hartwich, E., Entwurf eines automatischen Abstandsreglers für ein Strassenfahrzeug, *Regelungstechnik und Prozessdatenverarbeitung*, 22, 7 (1974), 193-224.
- [33] Herman, R., ed., *Theory of Traffic Flow*, Elsevier, Amsterdam, 1961.

- [34] Herman, R., E.W. Montroll, R.B. Potts, and R. Rothery, Traffic Dynamics: Analysis of Stability in Car-Following, *Operations Research*, 7 (1959), 86-106.
- [35] Houpt, P.K., and M. Athans, *Dynamic Stochastic Control of Freeway Corridor Systems; Volume 1: Summary*, Electronic Systems Laboratory Report, ESL-R-608, MIT, Cambridge, Mass., 1975.
- [36] Inose, H., and T. Hamada, *Road Traffic Control* (Translated from Japanese by E.C. Posner) Division of Engineering and Applied Science, California Institute of Technology, 1974.
- [37] Isaksen, L., *Suboptimal Control of Large-Scale Systems with Application to Freeway Regulation*, Ph.D. dissertation, Dept. Elec. Eng., University of Southern California, Los Angeles, 1971.
- [38] Isaksen, L., and H.J. Payne, Freeway Traffic Surveillance and Control, *Proceedings of the IEEE*, 61, 5 (1973), 526-536.
- [39] Isaksen, L., and H.J. Payne, Regulation of Freeway Traffic, Joint Automatic Control Conference, Proceedings (1972), paper 4-5, 63-72.
- [40] Isaksen, L., and H.J. Payne, Suboptimal Control of Linear Systems by Augmentation with Application to Freeway Traffic Regulation, *IEEE Trans. on Automatic Control*, 18, 3 (1973), 210-219.
- [41] Kahn, R., *Incident Detection Logics for the Los Angeles Freeway Surveillance and Control Project*, revised research report, State of California Business and Transportation Agency, Los Angeles, 1974.
- [42] Knapp, H., Traffic Estimation and Control at Bottlenecks, in *1972 IEEE System, Man and Cybernetics Conference*, New York, 1972.
- [43] Leutzbach, W., *Einführung in die Theorie des Verkehrsflusses*, Springer Verlag, Berlin, Heidelberg, 1972.
- [44] Lätzsch, L., *Vorbereitende Untersuchungen zur verkehrsabhängigen Steuerung eines Einzelknotens*, forschungsbericht, Hochschule für Verkehrswesen, Dresden, 1973.
- [45] Lighthill, M.J., and G.B. Whitham, On Kinematic Waves, II.A. Theory of Traffic Flow on Long Crowded Roads, *Proc. Roy. Soc. London*, 229A (1955), 317-345.
- [46] Lohse, E., *Beitrag zur optimalen Regelung des Strassenverkehrs*, Dissertationschrift, Hochschule für Verkehrswesen, Dresden, 1976.

- [47] Looze, D.P., P.K. Houpt, and M. Athans, *Dynamic Stochastic Control of Freeway Corridor Systems; Volume III: Dynamic Centralized and Decentralized Control Strategies*, Electronic Systems Laboratory Report, ESL-608, MIT, Cambridge, Mass., 1975.
- [48] May, A.D., Some Fixed-Time Signal Computer Programs, in [78].
- [49] Mikhalkim, B., H.J. Payne, and L. Isaksen, Estimation of Speed from Presence Detectors, *Highway Research Board Record*, 388 (1972), 73-83.
- [50] M.I.T.I., *Comprehensive Automobile Control System*, Preliminary Report, Japanese Ministry of International Trade and Industry, Tokyo, 1975.
- [51] Nahi, N.E., Freeway-Traffic. Data Processing, *Proceedings of the IEEE*, 61, 5 (1973), 537-541.
- [52] Nahi, N.E., and A.N. Trivedi, Recursive Estimation of Traffic Variables: Section Density and Average Speed, *Transportation Science*, 7, 3 (1973), 269-286.
- [53] Nguyen, S., An Algorithm for the Traffic Assignment Problem, *Transportation Science* (1974), 203-216.
- [54] Orlhac, D.G., M. Athans, J. Speyer, and P.K. Houpt, *Stochastic Control of Freeway Corridor Systems; Volume IV: Estimation of Traffic Variables Via Extended Kalman Filter Methods*, Electronic System Laboratory Report, R-611, MIT, Cambridge, Mass., 1975.
- [55] Payne, H.J., Models of Freeway Traffic and Control, in *Simulation Council Proceedings, Mathematical Models of Public Systems*, 1, 1 (1971).
- [56] Payne, H.J., *Evaluation of Estimates of Traffic Parameters from Presence Detectors - Part I: Generation of Freeway Traffic Data and Related Software*, USCEE Report 462, Dept. Elec. Eng., University of Southern California, Los Angeles, 1974.
- [57] Payne, H.J., D.N. Goodwein, and M.D. Teener, *Evaluation of Existing Incident Detection Algorithms*, Research Report FHWA-RD-75-39, Technology Service Cooperation, Santa Monica, California, 1975.
- [58] Payne, H.J., and W.A. Thompson, Allocation of Freeway Ramps Metering Volumes to Optimize Corridor Performance, *IEEE Transactions on Automatic Control*, AC-19,3 (1974), 177-186.
- [59] Payne, H.J., W.A. Thompson, and L. Isaksen, Design of a Traffic-Responsive Control System for a Los Angeles Freeway, *IEEE Trans. on Systems, Man and Cybernetics*, 3, 3 (1973), 213-224.

- [60] Pipes, L.A., An Operational Analysis of Traffic Dynamics, *J. Appl. Phys.*, 24 (1953), 274-281.
- [61] Reuschel, A., Fahrzeugbewegungen in der Kolonne bei gleichförmig beschleunigtem oder verzögertem Leitfahrzeug, *Zeit. Osterreich. Ing. und Archit. Ver.*, 95 (1950), 59-62, 73-77.
- [62] Richards, P.J., Shock Waves on the Highway, *Operations Research*, 4 (1956), 42-51.
- [63] Robertson, D.I., *Transyt: A Traffic Network Study Tool*, RRL Report LR 253, Road Research Laboratory, Crowthorne, Berkshire, 1969.
- [64] Robillard, P., Estimating the O-D Matrix from Observed Link Volumes, *Transp. Res.*, 9 (1975), 123-128.
- [65] Schnabel, W., *Lichtsignalgesteuerter Strassenverkehr*, Transpress, VEB Verlag für Verkehrswesen, Berlin, 1975.
- [66] Sendaula, H.M., and C.H. Knapp, Optimal Control of Hyperbolic Distributed Parameter Systems with Application to Traffic Bottlenecks, *Proc. of the 1972 International Conference on Cybernetics and Society*, IEEE, N.Y., 1972.
- [67] Shah, M.M., Supplemented Partitioning Approach to Linear Filtering Theory, paper presented at the IFAC Symposium on Large Scale Systems Theory and Application, Udine, Italy, June 16-20, 1976.
- [68] Singh, M.G., S.A.W. Drew, and J.F. Coales, Comparison of Practical Hierarchical Control Methods for Interconnected Dynamical Systems, *Automatica*, 11 (1975), 331-350.
- [69] Shefer, J., R.J. Klensch, H.C. Johnson, and G.S. Kaplan, A New Kind of Radar for Collision Avoidance, paper presented at the Automatic Engineering Congress, Detroit, Mich., February 25-March 1, 1974.
- [70] Strobel, H., *Experimentelle Systemanalyse*, Akademie Verlag, Berlin, 1975.
- [71] Strobel, H., *Transportation, Automation and the Quality of Urban Living*, IIASA Research Report RR-75-34, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.
- [72] Strobel, H., Computerized Urban Traffic Control Systems, paper presented at the IFAC Symposium on Large Scale Systems Theory and Applications, Udine, Italy, June 16-20, 1976.

- [73] Strobel, H., and E. Lohse, Über die Anwendung des diskreten Maximumprinzips zur Ermittlung einer optimalen Regelungsstrategie für ein System Übersättigter Strassenkreuzungen, Bericht über das XVIII, *Int. Wiss. Kolloquium der TH Ilmenau*, Vortragsreihe B 1 (1973), 111-116.
- [74] Strobel, H., and W. Naumann, Über ein Korrelationsverfahren zur Abschätzung von Reisezeiten im Strassenverkehr, *Messen, steuern, regeln*, 19, 5 (1976), 167-171.
- [75] Strobel, H., and W. Ullmann, *Die optimale Abstandsregelung im Kolonnenverkehr auf Autobahnen*, paper presented at the Wissenschaftliche Konferenz der Technischen Universität, Prague, September 1975.
- [76] Szeto, M.W., and D.C. Gazis, Application of Kalman Filtering to the Surveillance and Control of Traffic Systems, *Transportation Science*, 6 (1972), 419-439.
- [77] Tabak, D., Application of Modern Control and Optimization Techniques to Transportation Systems, *Advances in Control Theory*, 10 (1973), 345-435.
- [78] *Traffic Control and Transportation Systems*, Proceedings of the Second IFAC/IFIP/IFORS World Symposium in Monte Carlo, September 16-21, 1974, North-Holland, Amsterdam, 1974.
- [79] Todtli, J., Contribution to the Problem of Urban Road Traffic Control During Rush Hours, in [78], 31-42.
- [80] Yagoda, N., The Dynamic Control of Automotive Traffic at a Freeway Entrance Ramp, *Automatica*, 6 (1970), 385-393.
- [81] Ziegler, M., Wegewahl als Regelkreis-Untersuchungen über den Zusammenhang zwischen Wegewahl, Reisezeiten und Verkehrsmengen, *Strassenbau und Strassenverkehrstechnik*, 9 (1970).