Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

S.A. Smulders

Filtering of freeway traffic flow

Department of Operations Research and System Theory Report OS-R8806 February

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Filtering of Freeway Traffic Flow

S.A. Smulders
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

In this report an algorithm for the estimation of the state of traffic on a freeway is developed and tested against simulated and real data. The original filter is modified to achieve good traffic state estimates in situations of stationary as well as unstable traffic behaviour. It is shown that use of speed information is necessary to reduce estimator bias and error variance and to achieve filter robustness with respect to modelling errors. A limited use of speed information is shown to be sufficient.

1980 Mathematics Subject Classification : 90B20, 93E11.
Keywords & phrases: freeway traffic, filtering.

Report OS-R8806
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Contents

1. Introduction 3

2. A model of freeway traffic flow 3

3. The optimal filter and approximations 7

4. Analysis of the filter 10
4.1. Detectability and stabilizability 11
4.2. Asymptotical behaviour of the covariance matrix 13
4.3. Speed information and estimator accuracy 16

5. Testing the filter 18
5.1. Performance criteria 19
5.2. Results with simulated data 22
5.3. Results with real data 25
5.3.1 Stationary traffic 26
5.3.2 Unstable traffic 30
5.4. Validation 40

6. Conclusions 45

References 46

Appendices 47
A. Non-optimal filter equilibrium points
B. First order sensitivity analysis
C. Systematic estimation errors

1. Introduction

In an earlier report [15] a traffic control problem was posed and a procedure to solve this problem was presented. As a first step a model was developed and simulated. The model was shown to behave reasonably in various traffic situations. Most noteworthy is the fact that the model showed the instabilities that occur in practice when traffic becomes more and more dense. These instabilities we want to suppress by exercising some kind of control over the flow of traffic, for example by displaying suitable advisory speed signals or by metering the on-ramps.

Although open-loop or time-of-day control is sensible for some traffic control problems, in our case a closed-loop approach seems to be necessary. Such a control is based upon the actual traffic state at any moment. Unfortunately, in the present set up of the Dutch Motorway Control and Signalling System the exact state is not available. Measurements that are available consist of passing times and speeds of vehicles at specific locations along the freeway [13].

The subject of this report is development of an algorithm for the estimation of the state of traffic at any time moment from the available measurements up to that moment. In system and control theory such an algorithm is called a *filter.*

In section 2 the freeway traffic model as developed in [15] will be summarised. The optimal filter will be derived in section 3 and possible approximations to the optimal filter will also be presented. Section 4 is concerned with some theoretical and qualitative aspects of the optimal filter in an ideal traffic situation. In section 5 an approximation of the optimal filter is extensively tested against simulated and real data. Section 6 contains conclusions and suggestions for further research.

2. A model of freeway traffic flow

A detailed derivation and motivation of the traffic model we use is given in an earlier report [15]. In this section we will confine ourselves to summarising the model and listing the parameter values.

The freeway stretch that we will consider is discretised into N sections and for each section $i \in \{1,...,N\}$ we define the state variables

$\rho_i(t)$: density, the number of vehicles in section i per km per lane

$v_i(t)$: the mean speed of the vehicles in section i (km/h)

These state variables are supposed to obey the following stochastic differential equations:

$$d\rho_i(t) = \frac{1}{l_i L_i}[l_{i-1}\bar{\rho}_{i-1}\bar{v}_{i-1} - l_i\bar{\rho}_i\bar{v}_i]dt + \frac{1}{l_i L_i}[dm_{i-1}(t) - dm_i(t)] \tag{2.1}$$

$$\begin{aligned} dv_i(t) = & -\frac{1}{T}[v_i - v^e(\rho_i)]dt - \gamma(L_i l_i)^2[\beta\rho_i + (1-\beta)\rho_{i+1}][\rho_{i+1} - \rho_i]dt \\ & + \frac{l_{i-1}}{l_i L_i} v_{i-1}[v_{i-1} - v_i]dt + dw_i(t) \end{aligned} \tag{2.2}$$

where

$$\bar{\rho}_i = \alpha\rho_i + (1-\alpha)\rho_{i+1}$$
$$\bar{v}_i = \alpha v_i + (1-\alpha)v_{i+1}$$

$$v^e(\rho) = \begin{cases} v_{free} - a\rho \ , & 0 \leq \rho \leq \rho_{crit} \\ b\,[\frac{1}{\rho} - \frac{1}{\rho_{jam}}] \ , & \rho_{crit} < \rho \leq \rho_{jam} \end{cases} \tag{2.3}$$

and

$m_i(t)$: a counting process martingale

$w_i(t)$: a Brownian motion process

N : the number of sections

l_i : the number of lanes of section i

L_i : the length of section i (km)

α : weighting factor $\in [0,1]$

T : relaxation time (h)

v_{free} : free speed, equilibrium mean speed at zero density (km/h)

ρ_{jam} : jam density, at which the equilibrium speed is zero (veh/km/lane)

ρ_{crit} : critical density: the density for which the equilibrium intensity attaines its maximum value

a,b : parameters in $v^e(\rho)$

γ : anticipation factor (km/h^2)

β : weighting factor $\in [0,1]$

For later use we introduce the *intensity*

$$\lambda_i(t) = l_i \bar{\rho}_i \bar{v}_i \,, \qquad i = 0,\ldots,N \tag{2.4}$$

In the model the variables ρ_0 , v_0 , ρ_{N+1} , v_{N+1} are not defined. Suitable boundary conditions at the entrance and at the exit of the freeway stretch have to be chosen to get a closed set of equations. Two possible choices of boundary conditions are given here:

1. prescribed intensity

entrance: $$\rho_0 = [\,\frac{\lambda_0}{v_1} - (1-\alpha)\rho_1\,]/\alpha$$

$$v_0 = v_1$$

exit: $$\rho_{N+1} = \rho_N$$

$$v_{N+1} = [\,\frac{\lambda_N}{\rho_N} - \alpha v_N\,]/(1-\alpha)$$

where λ_0 (t) and λ_N (t) are given as functions of time.

2. stationarity

entrance: $$\rho_0 = \rho_1$$

$$v_0 = v_1$$

exit: $$\rho_{N+1} = \rho_N$$

$$v_{N+1} = v_N$$

In this report we will assume that the freeway stretch under consideration has no on- or off-ramps.

The model as just described is based upon a model derived by H.J. Payne [11]. It was simulated and shown to give reasonable behaviour in various traffic situations in [15]. Most of the parameter values were chosen in an ad hoc way so as to achieve realistic behaviour. A parameter identification procedure was not applied. The equilibrium relation between speed and density however has been estimated from real data for densities below the critical value ρ_{crit}. This estimated part will be used in the filter. For higher densities we will use the relation listed above. The parameters that will be used in the filter are summarised in table 2.1 . A plot of $v^e(\rho)$ is given in figure 2.1.

parameter	value	unit
α	0.85	
T	0.01	h
ρ_{jam}	110.0	veh/km/lane
v_{free}	105.0	km/h
ρ_{crit}	27.0	veh/km/lane
a	0.58	km^2/h
b	3197.0	1/h
γ	6.5	km/h^2
β	0.5	

TABLE 2.1

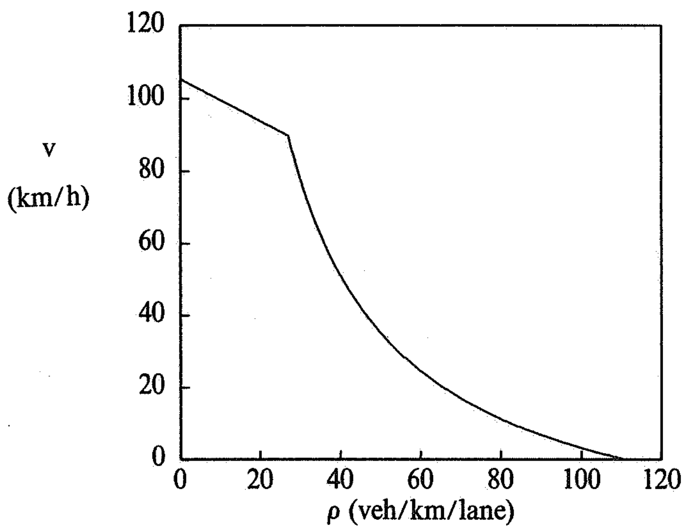


FIGURE 2.1

The measurements that are available consist of passing times and speeds of vehicles over section boundaries. The passing times are represented by the counting processes

$n_i(t)$: the number of vehicles that left section i since t_0

For these processes the following decomposition holds:

$$dn_i(t) = (1+\epsilon_i^f-\epsilon_i^m)\lambda_i(t)dt \; + \; d[\; m_i+r_i^f-r_i^m \;](t) \tag{2.5}$$

Here $\lambda_i(t)$ is given by (2.4) and ϵ_i^f and ϵ_i^m correspond to *counting errors:*

ϵ_i^f : the fraction of false counts at location i

ϵ_i^m : the fraction of missed passages at location i

We will always assume that

$$0 \leqslant \epsilon_i^f, \epsilon_i^m < 1$$

In (2.5) $r_i^f(t)$ and $r_i^m(t)$ are martingales associated with the counting error processes.

For practical reasons the speed domain was divided into M classes $V^1,...,V^M$ where $V^j = [v^{j-1}, v^j)$ for j=1,...,M . The number of classes necessary will be discussed later, in testing the filter. For the moment we restrict ourselves to noting that in case M=1 speed measurements are not modelled, and equation (2.5) suffices. The larger we take M the more specific the speed information will be, at the cost of extra computational effort in the filter. See section 3 for this. The discretisation of the speed domain leads to a distribution of the number of counts n_i over the M classes:

$n_i^j(t)$: the number of vehicles that left section i with speed in class V^j since t_0

Note that $\sum_{j=1}^{M} n_i^j(t) = n_i(t)$.

The measurement equations are now modelled as follows:

$$dn_i^j(t) = (1+\epsilon_i^{f^j} - \epsilon_i^{m^j})\gamma_i^j(\overline{\rho}_i,\overline{v}_i)\lambda_i(t)dt + d[m_i^j + r_i^{f^j} - r_i^{m^j}](t) \qquad (2.6)$$

for i=0,...,N and j=1,...,M , where

$\epsilon_i^{f^j}$: the fraction of false counts in class V^j at location i

$\epsilon_i^{m^j}$: the fraction of vehicles with speed in class V^j missed at location i

$m_i^j(t)$: the martingale associated with $n_i^j(t)$

$r_i^{f^j}(t), r_i^{m^j}(t)$: martingales associated with the counting error processes

γ_i^j : the fraction of vehicles that leave section i with speed in class V^j

The fraction of vehicles passing a location i with speed in class j, γ_i^j , has to be modelled. It is supposed to depend on traffic density and, of course, mean speed at the location. A detailed investigation of real traffic data was made and this has led to the conclusion that for densities below the critical value in case of stationary traffic the passing speed distribution may well be approximated by a normal distribution. The mean of the distribution is equal to the traffic speed $\overline{v}_i$ and the standard deviation was found to decrease with increasing density. Linear regression led to the relation

$$\sigma = 16 - 0.28\,\rho \qquad \text{for } \rho \leqslant \rho_{crit}$$

It is not clear how to estimate the distribution for values above the critical one, as in this case traffic behaves in a nonstationary way. Inspection of a very limited set of data led to the hypothesis that there is a lower bound for σ which is about 6 km/h. Following the relation above this value is reached at ρ = 35.7 veh/km/lane. We therefore chose to model the standard deviation of the probability distribution of passing speeds as follows:

$$\sigma = \begin{cases} 16 - 0.28\rho, & 0 \leqslant \rho \leqslant 35 \\ 6, & \rho > 35 \end{cases} \qquad (2.7)$$

Given the probability distribution of passing speeds F_V as a function of ρ and v the fractions γ_i^j directly follow from

$$\gamma_i^j(\rho,v) = F_V(v^j) - F_V(v^{j-1}) + \frac{1}{M}[\,F_V(v^0) - F_V(v^M)\,]$$

The last term is a correction term to assure that $\sum_{j=1}^{M} \gamma_i^j = 1$.

This concludes our definition of the freeway traffic model on which the filter will be based.

3. THE OPTIMAL FILTER AND APPROXIMATIONS

We will now develop the algorithm for the estimation of the section densities and mean speeds from the measured passing times and speeds.

To simplify notation we introduce the *state vector*

$$X_t = [\rho_1(t), \rho_2(t), \ldots, \nu_{N-1}(t), \nu_N(t)]^T$$

and the *measurement vector*

$$N_t = [n_0^1(t), n_0^2(t), \ldots, n_N^M(t)]^T$$

and the vector of intensities:

$$L(X_t) = [\lambda_0^1(t), \lambda_0^2(t), \ldots, \lambda_N^M(t)]^T$$

Our model and measurement equations may then be summarised as

$$dX_t = F(X_t)dt + dZ_t \tag{3.1}$$

$$dN_t = H(X_t)dt + dM_t \tag{3.2}$$

where F(.) and H(.) and Z and M follow from (2.1), (2.2) and (2.6). For later use we write

$$F(X_t) = \begin{bmatrix} F_1(X_t) \\ F_2(X_t) \end{bmatrix} = \begin{bmatrix} AL(X_t) \\ F_2(X_t) \end{bmatrix}$$

where the constant matrix A follows from equation (2.1) and only contains zeroes and elements of the type $\pm \frac{1}{L_i l_i}$.

For the estimation of the state X_t from the measurements $\{N_s, s \leq t\}$ techniques have been developed in the area of system and control theory. Well-known is the Kalman filter for the case where F and H are linear and Z and M are Brownian motion processes. For the case of counting type measurements theory has been developed also [1]. We will only give a very brief account here.

As can be easily shown, the estimator of X_t that minimizes the estimation error variance is given by

$$\hat{X}_t = E[X_t \mid \mathcal{F}_t^N]$$

It is the mean of X_t conditioned upon the measurements. $\mathcal{F}_t^N$ is the σ-algebra generated by $\{N_s, s \leq t\}$ and represents the information contained in the measurements up to and including time t. For $\hat{X}_t$ the following differential representation can be derived:

$$d\hat{X}_t = E[F(X_t) \mid \mathcal{F}_t^N]dt + \Phi_t(dN_t - E[H(X_t) \mid \mathcal{F}_t^N]dt) \tag{3.3}$$

$$\Phi_t = \left\{ (E[\tilde{X}_t \tilde{H}(X_t)^T \mid \mathcal{F}_t^N] + E[\frac{d}{dt}<Z,M>_t \mid \mathcal{F}_t^N]) E[\frac{d}{dt}<M,M>_t \mid \mathcal{F}_t^N]^{-1} \right\}_{t-} \tag{3.4}$$

Here Φ_t is called the *gain matrix* and

$$\tilde{X}_t = X_t - \hat{X}_t$$

$$\tilde{H}(X_t) = H(X_t) - E[H(X_t) \mid \mathcal{F}_t^N]$$

and $<Z,M>_t$, $<M,M>_t$ are so called predictable (co)variation processes. For a precise definition see [3, Part II, Chapter 7, Theorem 37 & 39]. If we assume that the counting processes in N_t do not have common jumps it may be shown that

$$\frac{d}{dt}<M,M>_t = \text{diag}[H(X_t)]$$

Note that a diagonal matrix which has the components of a vector X as its diagonal elements is denoted by diag[X]. It may be shown that in our case

$$\frac{d}{dt}<Z,M>_t = \begin{bmatrix} AM_t \\ 0 \end{bmatrix}$$

In (3.3) the first term on the righthand-side shows that the filter follows the model, to which the second term is a correction, based on the measurements.

Equation (3.3) gives the exact evolution of $\hat{X}_t$ in time. Unfortunately $\hat{X}_t$ cannot be computed from (3.3) in general. For the evaluation of $E[F(X_t)\,|\,\mathcal{F}_t^N]$ the computation of the entire conditional probability distribution of X_t is needed, which is unfeasible. The same holds for other terms in the representations for $\hat{X}_t$ and Φ_t. We therefore have to resort to approximations.

The usual approach is to develop Taylor series of the nonlinear functions like F and H around the estimated state $\hat{X}_t$ and neglect higher order terms. Depending on whether one only takes first or second order terms into account one speaks of a first or second order filter. We now give the equations of the *truncated second order filter*:

$$d\hat{X}_t = \hat{F}(\hat{X}_t)dt + \Phi_t[\ dN_t - \hat{H}(\hat{X}_t)dt\ ] \qquad (3.5)$$

$$\Phi_t = \left\{\hat{P}_t^x H'(\hat{X}_t)^T \mathrm{diag}^{-1}[\hat{H}(\hat{X}_t)] + \begin{bmatrix} A \\ O \end{bmatrix}\right\}_{t-} \qquad (3.6)$$

$$d\hat{P}_t^x = \left\{\hat{P}_t^x H'(\hat{X}_t)^T + H'(\hat{X}_t)\hat{P}_t^x + \begin{bmatrix} A\,\mathrm{diag}[\hat{L}(\hat{X}_t)]A^T & O \\ O & \Sigma \end{bmatrix} - \Phi_t \mathrm{diag}[\hat{H}(\hat{X}_t)]\Phi_t^T\right\}dt \qquad (3.7)$$

The filter is called truncated because a martingale term in the $\hat{P}_t^x$-equation is neglected. $\hat{P}_t^x$ is an approximation of the matrix of conditional error covariances

$$E[(X_t-\hat{X}_t)(X_t-\hat{X}_t)^T\,|\,\mathcal{F}_t^N]$$

In order to compute first or second order approximations to the optimal estimator of X_t it turns out to be necessary to compute (an approximation of) the conditional error covariance matrix as well. This is why equation (3.7) appears. Σ is the covariance matrix of the Brownian motion processes w_i.

In the equations above $\hat{F}$ is given by the following

$$\hat{F}_i(.) = F_i(.) + \frac{1}{2}\sum_j\sum_k \frac{\partial^2 F_i(.)}{\partial x_j \partial x_k}(\hat{P}_t^x)_{jk}$$

$\hat{H}$ and $\hat{L}$ are given by analogous expressions. To obtain the truncated first order filter one only has to set $\hat{F}_i$ equal to F_i etc. and apply the same equations (3.5) to (3.7) as for the second order filter.

The dimensions of the various vectors and matrices are as follows:

$\hat{X}$: 2N × 1

Φ_t : 2N × (N+1)M

$\hat{P}_t^x$: 2N × 2N

A : N × (N+1)M

Σ : N × N

$\hat{F}$: 2N × 1

$\hat{H}$: (N+1)M × 1

$\hat{L}$: (N+1)M × 1

Computing $\hat{X}$ therefore involves the integration of $2N+(2N+1)N=2N^2+3N$ differential equations ($\hat{P}_t^x$ is symmetric).

The second order filter requires the computation of the Hessian matrices of F_i, H_i and L_i (i=1,...,2N) extra in comparison to the first order filter. This requires the computation of $6N^2(2N+1)$ elements extra. Although in practice many zeros appear in the Hessians and a well-chosen order of computation may save a lot of computational effort it is clear that the second order filter is relatively expensive. It is therefore of interest to investigate whether the extra effort is justified by a considerable gain in estimator precision. This question will be addressed in section 5.

It is well-known that numerical instability may occur in the integration of (3.7) [18]. To reduce the sensitivity to rounding errors so called *square root* algorithms have been developed [7]. Most of these algorithms consider the discrete time version of (3.7). In the square root algorithm of Morf, Lévy and Kailath [9] for the continuous time filter the matrix $\hat{P}_t^x$ is replaced by its Cholesky factor S_t:

$$\hat{P}_t^x = S_t S_t^T$$

where S_t is lower triangular. Using this factorisation (3.7) may be written as (a dot denotes differentiation with respect to time):

$$\dot{S}_t S_t^T + S_t \dot{S}_t^T = S_t S_t^T H'(\hat{X}_t)^T + H'(\hat{X}_t) S_t S_t^T + \begin{bmatrix} A\,\mathrm{diag}[\hat{L}(\hat{X}_t)]A^T & O \\ O & \Sigma \end{bmatrix} - \Phi_t \mathrm{diag}[\hat{H}(\hat{X}_t)]\Phi_t^T \tag{3.8}$$

Note that in computing $\hat{L}$ and $\hat{H}$ and also in computing Φ_t we need $\hat{P}_t^x$ explicitly and therefore have to perform the multiplication of S_t with its transpose. Assuming that $\hat{P}_t^x>0$ for all t so that the inverse of S_t exists, pre- and post-multiplication of (3.8) by S_t^{-1} and S_t^{-T} respectively leads to

$$S_t^{-1}\dot{S}_t + (S_t^{-1}\dot{S}_t)^T = (H'(\hat{X}_t)S_t)^T S_t^{-T} + S_t^{-1}(H'(\hat{X}_t)\,S_t) + S_t^{-1} \begin{bmatrix} A\,\mathrm{diag}[\hat{L}(\hat{X}_t)]A^T & O \\ O & \Sigma \end{bmatrix} S_t^{-T}$$
$$- \; S_t^{-1}\Phi_t \mathrm{diag}[\hat{H}(\hat{X}_t)]\Phi_t^T S_t^{-T} \tag{3.9}$$

Now $S_t^{-1}\dot{S}_t$ is lower triangular and we only need to compute the lower triangle of the righthand-side and divide the diagonal by 2. For this we may use the notation

$$S_t^{-1}\dot{S}_t = \left[(H'(\hat{X}_t)S_t)^T S_t^{-T} + \cdots \right]_{\cdot/2}$$

and so

$$\dot{S}_t = S_t \,[\text{ righthand-side of (3.9) }]_{\cdot/2} \tag{3.10}$$

The advantage in integrating (3.10) over integrating (3.7) is that S_t may be seen as the square root of $\hat{P}_t^x$: the elements are closer together (the difference between two elements, measured in the number of digits, is halved). This will reduce the effect of rounding errors which is especially important when $\hat{P}_t^x$ is almost singular. The gain in precision is at the cost of increased computational effort. Suppose that the computation of H', $\hat{L}$, $\hat{H}$, Φ_t requires Δ operations. Then computing $\hat{P}_t^x$ according to the original equation (3.7) costs

$$\Delta + 10N^3 + 8MN^2(N+1) \text{ operations}$$

and the square root approach requires

$$\Delta + 18N^3 + 12MN^2(N+1) \text{ operations}$$

If the other computations would require relatively little effort (Δ small) the increase due to the square root approach would be over 50 percent. In our case Δ is quite large so the relative effect is smaller, but it is clear that the square root method should only be used if there is a clear indication of possible numerical problems. In our case preliminary investigations showed that $\hat{P}_t^x$ may become illconditioned. This is confirmed by results in section 4.2. We therefore decided to implement the square root algorithm described above.

The square root algorithm proposed by Bar-Itzhack and Oshman [10] is quite different from the one we just described. Here the eigenvalues and eigenvectors of $\hat{P}_t^x$ are propagated. Unfortunately this method is not applicable to our second order filter because $\hat{L}$ and $\hat{H}$ require $\hat{P}_t^x$ explicitly. Furthermore, this eigenfactor solution algorithm requires even more computations:

$$\Delta + 22N^3 + 12MN^2(N+1) \text{ operations}$$

We will therefore use the method of Morf, Lévy and Kailath.

To solve (3.5),(3.6) and (3.10) a FORTRAN computer program was written and implemented on a CDC Cyber 750 mainframe. The integration is simply done by Euler's method, the integration step being dictated by the observed sample path of N_t. Equations (3.5) and (3.10) are integrated in one step from one jump of a component of N_t to the next jump of one of the components of N_t. Next X_t and Φ_t are updated for the jump in N_t and another integration step is carried out. In case there is a small number of counting processes or there are not many jumps (low traffic intensity) the integration step may become too large. To avoid this an upper bound of 0.0001 hour for this step was chosen. This value was tested: the filter turned out to be insensitive to a reduction of the upper bound.

Equation (3.6) requires the inversion of a diagonal matrix. The diagonal consists of the components of $\hat{H}(\hat{X}_t)$:

$$\left[\hat{H}(\hat{X}_t)\right]_{i\times(N-1)+j} = (1+\epsilon_i^{p^j}-\epsilon_i^{m^j})\gamma_i^j(\hat{\rho}_i,\hat{v}_i)l_i\hat{\rho}_i\hat{v}_i$$

$$= (1+\epsilon_i^{p^j}-\epsilon_i^{m^j})\hat{\lambda}_i^j$$

We may expect numerical problems whenever a counting process intensity $\hat{\lambda}_i^j$ becomes small. This may happen when the mean passage speed at a location is far off the center of a speed class which then hardly contains any probability mass. Problems are also to be expected when $\hat{\rho}_i$ or $\hat{v}_i$ becomes small. The problems just mentioned have occurred in practice. As a remedy we chose to take a lowerbound of 10 veh/h for each of the intensities $\hat{\lambda}_i^j$.

The proposed filter had to undergo another small adjustment: whenever $\hat{\rho}_i<0$ or $\hat{v}_i<0$ these variables are reset to 0. Especially the former may occur because of initial uncertainty about the real density. The filter may estimate the density too low and it may become zero while in reality there is still a vehicle in the section. This vehicle may leave the section, necessitating a reduction in the estimated density, which then becomes negative.

As a final remark we note that instead of modelling the passing speed distribution as a normal one as mentioned in section 2, we used a logistics distribution. This distribution is known to approximate the normal distribution well, and has the advantage of having an analytical expression:

$$F_V(v) = \frac{1}{1 + e^{\frac{-\pi(v-\mu)}{\sigma\sqrt{3}}}}$$

This facilitates the computation of the γ_i^j's.

4. Analysis of the filter

Before testing the filter developed in the previous section on data, it may be worthwhile to investigate some of its properties theoretically. Of special interest are the detectability and stabilizability properties: the former is a necessary condition without which part of the state cannot be estimated with finite error variance. Both conditions will be discussed in subsection 4.1. In 4.2 conclusions about the asymptotical behaviour of the covariance matrix $\hat{P}_t^x$ will be drawn. In 4.3 the effect of using speed information will be investigated by computing the asymptotic value of the covariance matrix P_t^x.

4.1 Detectability and stabilizability.
To perform the analysis the filter equations given in the previous section are too complicated. We will therefore start by introducing some simplifying assumptions.

The main difficulty with equation (3.7) is that the righthand-side depends on the estimated state $\hat{X}_t$. This means that this equation is coupled with (3.5) in which the observations N_t appear. Assumptions about these have to be made. We will make the following assumption:

ASSUMPTION *The estimated section density and mean speed are assumed to be nonzero, constant over time and equal in all sections, and the mean speed is assumed to be equal to the equilibrium speed:*

$$\forall\, i \in \{1,..,N\},\ \forall\, t \geqslant t_0 : \hat{\rho}_i(t) = \hat{\rho} \neq 0\ ,\ \hat{v}_i(t) = v^e(\hat{\rho}) \neq 0$$

This eliminates the above formulated difficulty, $\hat{X}_t = (\hat{\rho},\hat{v},\hat{\rho},\hat{v},....)^T$ is now a constant vector. This approach comes down to considering the ideal situation of an homogeneous and stationary traffic stream where the above mentioned assumption holds for ρ_i and v_i. The filter is assumed to produce $\hat{\rho}_i$ and $\hat{v}_i$ that are close enough to ρ_i and v_i to justify application of the assumption in (3.7).

Another problem is that in the second order filter the matrix P_t^x appears in $\hat{L}$ and $\hat{H}$ on the righthand-side of (3.7). We will therefore only consider the first order filter in the following. This leads us to the standard *Riccati equation*:

$$\dot{P}_t = P_t K^T + K P_t - P_t G^T R^{-1} G P_t + Q\ , \qquad P_{t_0} = P_0 \geqslant 0 \tag{4.1}$$

where

$$\begin{aligned} K &= \begin{bmatrix} A\,\mathrm{diag}(\epsilon)\hat{\lambda}' \\ F_2'(\hat{x}) \end{bmatrix} \\ G &= \hat{\lambda}' \\ R &= \mathrm{diag}(1-\epsilon)^{-1}\mathrm{diag}(\hat{\lambda}) \\ Q &= \begin{bmatrix} A\,\mathrm{diag}(\epsilon)\mathrm{diag}(\hat{\lambda})A^T & O \\ O & \Sigma \end{bmatrix} \end{aligned} \tag{4.2}$$

and

$$\begin{aligned} \epsilon &= (\bar{\epsilon},\bar{\epsilon},\cdots)^T \\ \hat{\lambda} &= (l\hat{\rho}\hat{v}, l\hat{\rho}\hat{v},\cdots)^T \end{aligned}$$

Note that $K \in \mathbb{R}^{2N \times 2N}$, $G \in \mathbb{R}^{(N+1)M \times 2N}$, $R \in \mathbb{R}^{(N+1)M \times (N+1)M}$, $Q \in \mathbb{R}^{2N \times 2N}$, $\epsilon \in \mathbb{R}^{(N+1)M}$, $\hat{\lambda} \in \mathbb{R}^{(N+1)M}$.
$F_2' \in \mathbb{R}^{N \times 2N}$ and $\hat{\lambda}' \in \mathbb{R}^{(N+1)M \times 2N}$ are the Jacobian matrices of F_2 and $\hat{\lambda}$ respectively. To simplify computations we have assumed that all sections have an equal number of lanes, l , are equally long, counting error fractions are equal to $\bar{\epsilon}$ at all locations and will further neglect the effect of boundary conditions.

We will now investigate the detectability and stabilizability properties of the pairs of matrices [G,K] and [K,L] respectively. Theorems about the existence of solutions and about the asymptotic behaviour often make use of these notions.

DETECTABILITY

DEFINITION 4.1 *A pair of matrices [G,K] with $G \in \mathbb{R}^{m \times n}$, $K \in \mathbb{R}^{n \times n}$ is called* observable *if*

$$rank([\ G^T \mid K^T G^T \mid ... \mid (K^{n-1})^T G^T\ ]) = n$$

DEFINITION 4.2 *A pair of matrices [G,K] with* $G \in \mathbb{R}^{m \times n}$, $K \in \mathbb{R}^{n \times n}$ *is called* detectable *if*

$$kernel\left(\begin{bmatrix} G \\ GK \\ GK^2 \\ \cdots \end{bmatrix}\right) \subset \chi^-(K) \qquad (4.3)$$

where $\chi^-(K)$ denotes the stable subspace of K. In words: the unobservable subspace corresponding to [G,K] has to be contained in the stable subspace of K. As can easily be seen observability implies detectability. The matrix involved in (4.3) is called the *observability matrix.*

As observability is easier to check than detectability we will first try to establish the former for $G \in \mathbb{R}^{(N+1)M \times 2N}$ and $K \in \mathbb{R}^{2N \times 2N}$ given by (4.2). Three cases will be considered: *i)* M>1 ; *ii)* M=1, N>1 ; *iii)* M=1, N=1.

In case M>1 matrix G may be shown to consist of 2N independent columns with the exception of some rare cases. This means that generally rank(G) equals 2N and so the observability condition is satisfied whatever K is.

In case M=1 and N>1 however $G \in \mathbb{R}^{(N+1) \times 2N}$ so rank(G)<2N and we have to consider $K^T G^T$ and possibly $(K^2)^T G^T$ etc. It turns out that $\text{rank}([G^T \mid K^T G^T]) = 2N$ and observability of [G,K] is guaranteed.

When N=1 and M=1 rank(G) turns out to be smaller than 2N but considering $K^T G^T$ again leads to observability.

We conclude that in all cases the pair [G,K] is observable. Note that the absence of speed information (M=1) does not prohibit the possibility of estimating the traffic state. The counts alone apparently contain enough information.

STABILIZABILITY

DEFINITION 4.3 *A pair of matrices [K,L] with* $K \in \mathbb{R}^{n \times n}$, $L \in \mathbb{R}^{n \times n}$ *is called* controllable *if*

$$rank([\, L \mid KL \mid \ldots \mid K^{n-1}L \,]) = n$$

DEFINITION 4.4 *A pair of matrices [K,L] with* $K \in \mathbb{R}^{n \times n}$, $L \in \mathbb{R}^{n \times n}$ *is called* stabilizable *if*

$$range([\, L \mid KL \mid \ldots \mid K^{n-1}L \,]) \supset \chi^+(K) \qquad (4.4)$$

where $\chi^+(K)$ denotes the unstable subspace of K. Note that controllability implies stabilizability. The matrix involved in (4.4) is called the *controllability matrix.*

Again we will first try to establish the stronger property for the pair [K,L] where $L \in \mathbb{R}^{2N \times 2N}$ such that $LL^T = Q$ and $K \in \mathbb{R}^{2N \times 2N}$ is given by (4.2). Three cases are considered: *i)* $\epsilon>0$, $\Sigma>0$; *ii)* $\epsilon=0$; *iii)* $\epsilon>0$, $\Sigma=0$. (Here $\epsilon>0$ means that all components of ϵ are larger than 0).

If $\epsilon>0$ and $\Sigma>0$ then Q has a unique positive definite root $Q^{\frac{1}{2}}$. Take $L=Q^{\frac{1}{2}}$ then rank(L)=2N, the maximal value. Controllability of [K,L] follows immediately.

In case $\epsilon=0$ it is clear that $L = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma^{\frac{1}{2}} \end{bmatrix}$ and controllability does not follow from the structure of L alone. The matrices KL etc. have to be taken into consideration. Notice that

$$K = \begin{bmatrix} 0 \\ F_2'(\hat{X}_t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ K_1 & K_2 \end{bmatrix}$$

which leads to

$$KL = \begin{bmatrix} 0 & 0 \\ 0 & K_2\Sigma^{\frac{1}{2}} \end{bmatrix}, \qquad K^2L = \begin{bmatrix} 0 & 0 \\ 0 & (K_2)^2\Sigma^{\frac{1}{2}} \end{bmatrix} \qquad \text{etc.}$$

This means that there are at most N independent rows in [L | KL | ...] and so the pair [K,L] is not controllable. Let us now try to establish the weaker property of stabilizability. Instead of investigating the stabilizability of [K,L] however it turns out to be handier to consider the equivalent property of detectability of $[L^T,K^T]$.

Clearly

$$\begin{bmatrix} L^T \\ L^TK^T \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & K_1^T \\ 0 & K_2^T \\ 0 & 0 \\ 0 & \Sigma^{\frac{1}{2}} K_2^T \\ \vdots & \vdots \end{bmatrix} \quad \text{and} \quad K^T = \begin{bmatrix} 0 & K_1^T \\ 0 & K_2^T \end{bmatrix}$$

so K^T has a set of eigenvectors corresponding to the unstable eigenvalue 0 that are in the kernel of the observability matrix. This excludes detectability of [L^T,K^T] and therefore [K,L] is not stabilizable.

In case $\epsilon>0$ and $\Sigma=0$ we take

$$L = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{with} \quad D = D^T > 0 \quad \in \mathbb{R}^{N\times N}$$

The controllability matrix then has the following structure

$$\begin{bmatrix} D & 0 & * & 0 & * & 0 & .. & .. \\ 0 & 0 & K_1D & 0 & K_2K_1D & 0 & .. & .. \end{bmatrix}$$

which means that whenever $\text{rank}(K_1D)=N$ controllability is assured. Now it turns out that whenever $\frac{\dot{v}^e(\hat{\rho})}{T} + \gamma\,(lL)^2\hat{\rho} \neq 0$ matrix K_1 is invertible and so will K_1D implying that $\text{rank}(K_1D)=N$. So in general controllability is assured. There are however realistic parameter values and values for $\hat{\rho}$ for which the condition just mentioned is not satisfied. A more detailed investigation shows that in case $\frac{\dot{v}^e(\hat{\rho})}{T} + \gamma\,(lL)^2\hat{\rho}=0$ controllability is still assured as long as $\gamma\neq0$. The case $\gamma=0$, $\frac{\dot{v}^e(\hat{\rho})}{T}=0$ requires further analysis that will not be carried out here.

To summarize:

- if $\epsilon>0$ and $\Sigma>0$ the pair [K,L] where $LL^T=Q$ is controllable;
- if $\epsilon=0$ the pair [K,L] where $LL^T=Q$ is not stabilizable.

  There are uncontrollable modes corresponding to the eigenvalue 0 of K;
- if $\epsilon>0$ and $\Sigma=0$ the pair [K,L] where $LL^T=Q$ is controllable,

  possibly with the exception of the case where $\gamma=0$ and $\frac{\dot{v}^e(\hat{\rho})}{T}=0$.

4.2 Asymptotical behaviour of the covariance matrix.

Now the results of the previous subsection may be used to draw conclusions about the asymptotical behaviour of the solution P_t of (4.1) which is an approximation of the covariance matrix P_t^x of (3.7). Note that the existence and uniqueness of a solution of (4.1) is guaranteed because it is just a set of ordinary differential equations. In case $\lim_{t\to\infty} P_t$ exists it will have to satisfy the *Algebraic Riccati Equation* (ARE)

$$0 = PK^T + KP - PG^TR^{-1}GP + Q \tag{4.5}$$

It is therefore of interest to investigate whether a solution of the ARE exist, whether it is unique, positive definite etc. We are only interested in real, symmetric, nonnegative definite solutions. As a by-product of a solution $\overline{P}$ of the ARE the eigenvalues of the matrix

$$\overline{K} = K - \overline{P}G^TR^{-1}G \tag{4.6}$$

are of interest, for straightforward manipulations show that

$$\frac{d}{dt}(\hat{X}_t - X_t) = (K - \overline{P}G^TR^{-1}G)(\hat{X}_t - X_t) + \text{a constant term} + \text{noise terms}$$

which means that stability of $\overline{K}$ guarantees exponential reduction of estimation errors. It is clear that only solutions for which $\overline{K}$ is not unstable are valuable.

DEFINITION 4.5 *A solution $\overline{P}$ of the ARE (4.5) is called* strong *if all the eigenvalues of $\overline{K}$ defined by (4.6) lie in the closed left half of the complex plane, and is called* stabilizing *if they all lie in the open left half of the complex plane.*

The classical result about the solution of the ARE is given by the following theorem:

THEOREM 4.6 *If [G,K] is detectable and [K,L] is stabilizable, where $LL^T = Q$, $L \in \mathbb{R}^{2N \times 2N}$, then the ARE (4.5) has a unique nonnegative definite solution $\overline{P}$. Furthermore, $\overline{P}$ is stabilizing and if [K,L] is controllable $\overline{P}$ is positive definite.*

PROOF

See Kuçera [5], Theorem 5 and Wonham [19], Theorem 4.1.

This theorem completely satisfies our needs in case $\epsilon > 0$. When $\epsilon = 0$ however, [K,L] is not stabilizable and a stronger result is necessary. Fortunately, using results which have appeared in literature recently, we are able to state the following theorem:

THEOREM 4.7 *Assume [G,K] to be detectable. Then the following holds:*

i) The ARE (4.5) has a unique strong solution $\overline{P}_{strong}$;

ii) $\overline{P}_{strong}$ is stabilizing if and only if there are no [K,L]-uncontrollable modes corresponding to purely imaginary eigenvalues of K;

iii) $\overline{P}_{strong}$ is positive definite if and only if [−K,L] is stabilizable.

PROOF

i) See Poubelle et al. [12] .

ii) According theorem 3 of Kuçera [5] a stabilizing solution of the ARE exists if and only if [G,K] is detectable and the associated Hamiltonian matrix has no purely imaginary eigenvalues. But this Hamiltonian has such an eigenvalue if and only if there exists a corresponding [G,K] unobservable and/or [K,L] uncontrollable eigenvalue (Lemma 8 in Kuçera [4]). The detectability of [G,K] excludes the first option which means that Kuçera's theorem may be reformulated as "A stabilizing solution exists if and only if [G,K] is detectable and there are no [K,L] uncontrollable modes on the imaginary axis". As there is only one strong solution according to i), a stabilizing solution has to be equal to $\overline{P}_{strong}$.

iii) ←) Theorem 2 of Richardson and Kwong [14] states that given that the Hamiltonian has no purely imaginary eigenvalues a positive definite solution exists if and only if [G,K] detectable and [−K,L] stabilizable. Our statement immediately follows by the fact that [−K,L] stabilizability together with [G,K] detectability implies that the Hamiltonian has no purely imaginary eigenvalues

(Kuçera [4] , lemma 8).

$\rightarrow$) Use theorem 3 of Richardson and Kwong [14]. [G,K] detectability and positive definiteness of a solution of the ARE implies that $\dim(A_2)=0$ in this theorem and stabilizability of $[-K_1,L_1]=[-K,L]$. □

Remark This theorem is the continuous time equivalent of theorem 3.2 of [16].

A unique strong solution thus exists in the case $\epsilon=0$ but the solution neither is stabilizing nor positive definite. In fact the strong solution may easily be shown to be

$$\bar{P}_{strong} = \begin{bmatrix} 0 & 0 \\ 0 & P_3 \end{bmatrix}$$

where P_3 is the unique positive definite solution of

$$0 = P_3 K_2^T + K_2 P_3 - P_3 G_2^T R^{-1} G_2 P_3 + \Sigma$$

in case $\Sigma>0$ and $P_3=0$ if $\Sigma=0$. Here $K = \begin{bmatrix} 0 & 0 \\ K_1 & K_2 \end{bmatrix}$ and $G = [G_1 \ G_2]$. The singularity of $\bar{P}_{strong}$ is caused by the fact that the section densities (and possibly the mean speeds) are estimated exactly asymptotically. In this case

$$\bar{K} = \begin{bmatrix} 0 \\ F_2' \end{bmatrix}$$

which implies that

$$d\hat{\rho}_i = 0 \quad \text{for } i=1,\ldots,N$$

which is logical once $\hat{\rho}_i=\rho_i$. So the non-stabilizability and singularity should not bother us. What should be a concern is the fact that more nonnegative solutions than $\bar{P}_{strong}$ exist:

Theorem 4.8 *A solution of the ARE (4.5) is unique if and only if [K,L] is stabilizable.*

Proof

See Kuçera [5], Theorem 4.

In our problem [K,L] is not stabilizable when $\epsilon=0$. In this case P_t might converge to an unsatisfactory solution. It will have to be investigated under which conditions convergence to the strong solution is guaranteed. Let us now direct attention to the convergence properties of P_t, the solution of (4.1). The classical result is the following:

Theorem 4.9 *If [K,L] is stabilizable and [G,K] is detectable, where $LL^T=Q$, $L \in \mathbb{R}^{2N\times 2N}$ then $\lim_{t\to\infty} P_t=\bar{P}$ for any $P_{t_0}\geqslant 0$. Here $\bar{P}$ is the unique solution of the ARE (4.5).*

Proof

See Kuçera [5], theorem 17.

In case [K,L] is not stabilizable the following result is of interest:

Theorem 4.10 *If [G,K] is detectable and $P_{t_0}\geqslant\bar{P}_{strong}$ then $\lim_{t\to\infty} P_t=\bar{P}_{strong}$.*

Proof

See theorem 3 of Poubelle et al. [12].

A careful selection of P_{t_0} thus assures convergence to the right solution.

To summarize:

- if $\epsilon>0$ the solution of the Riccati equation (4.1) converges to the unique, stabilizing and positive definite solution of the algebraic Riccati equation (4.5).
- if $\epsilon=0$ the solution of the Riccati equation (4.1) converges to the unique strong solution of the algebraic Riccati equation (4.5) if $P_{t_0} \geqslant \overline{P}_{strong}$.
  $\overline{P}_{strong}$ is not stabilizing and not positive definite.

In our filter applications the components of ϵ will generally be larger than 0, but small (in the order of 0.02). Numerical problems in the computation of $P_t^{\hat{x}}$ may therefore be expected. It is for this reason that a square root algorithm as mentioned in section 3 is implemented.
Some idea about the magnitude of the asymptotical estimation errors may be found in the next subsection.

4.3 Speed information and estimator accuracy.
The results of the previous section may now be used to investigate the effect of speed information on the accuracy of the estimates.
In general $\Sigma>0$ and $\epsilon>0$ will hold and the asymptotical error variances will be larger than 0. The magnitude of the errors will depend on the amount of information available in the estimation procedure: the more detailed the information, the more accurate the measurements will be.
The amount of information clearly depends on the number of speed classes M we use. If M=1 no speed information is contained in the measurements and in the limiting case where $M \to \infty$ the entire passing speed distribution is available.
The solution $\overline{P}$ of the algebraic Riccati equation (4.5) will now be computed using the interactive package MATLAB-SC [8]. Parameters are given the values of table 2.1 except that due to filtering results to be presented in the next section α was set to 0.5 and γ to 1.0. Furthermore, we took N=4, l=2, L=0.5 and Σ=diag(10000) and $\overline{\epsilon}=0.015/M$ in accordance with values used in the filter tests of section 5.
M will be chosen 1, 2 and 3 and $\hat{\rho}$ will be taken 20.0, 30.0 and 40.0 veh/km/lane. If M=2 the boundary between the two speed classes is chosen equal to $v^e(\hat{\rho})$ which implies that both classes contain equal probability mass. If M=3 the boundaries are also chosen such that each class contains equal probability mass. The results of the computations are given in table 4.1 and figures 4.1 and 4.2. In figure 4.1 the error standard deviation in $\hat{\rho}$ for several values of M and ρ is plotted. In figure 4.2 the same is done for $\hat{v}$. In table 4.1 the diagonal elements of the covariance matrix are listed, these are approximations of error variances in the state estimates. In figures 4.1 and 4.2 lines are drawn for the eye only, M can only be integer valued. From the results the following conclusions may be drawn:

- increasing the number of speed classes reduces the error variance, especially when density is high;
- the reduction in error variance is considerable when going from 1 to 2 speed classes, and negligible when going from 2 to 3 classes;
- the minimal value of the standard deviation of the error in $\hat{\rho}$ is about 2 veh/km/lane and in $\hat{v}$ about 4 km/h.

In case of an ideal traffic stream, homogeneous and stationary, taking two speed classes apparently is necessary and sufficient to produce accurate estimates. In practice, where the intensity varies considerably, more classes may be necessary. This will be investigated in the next section.

ρ (veh/km/lane)	20.0			30.0			40.0		
M	1	2	3	1	2	3	1	2	3
$\tilde{var}(\rho_1-\hat{\rho}_1)$	6.5	4.5	3.4	16.7	3.1	2.4	25.8	5.2	4.1
$\tilde{var}(\rho_2-\hat{\rho}_2)$	7.7	5.3	4.1	18.6	3.8	3.1	30.7	6.0	4.7
$\tilde{var}(\rho_3-\hat{\rho}_3)$	7.7	5.3	4.1	21.3	3.9	3.9	58.6	5.9	4.7
$\tilde{var}(\rho_4-\hat{\rho}_4)$	6.0	4.2	3.2	19.8	3.4	4.1	83.7	5.0	4.0
$\tilde{var}(v_1-\hat{v}_1)$	17.4	13.7	13.3	44.1	13.8	13.0	38.0	13.2	12.4
$\tilde{var}(v_2-\hat{v}_2)$	21.0	14.2	13.7	51.8	14.1	13.4	47.7	13.6	12.8
$\tilde{var}(v_3-\hat{v}_3)$	22.0	14.2	13.7	67.4	14.1	13.9	92.1	13.5	12.8
$\tilde{var}(v_4-\hat{v}_4)$	22.5	14.5	14.0	82.4	13.7	16.3	137.6	13.0	12.2

TABLE 4.1

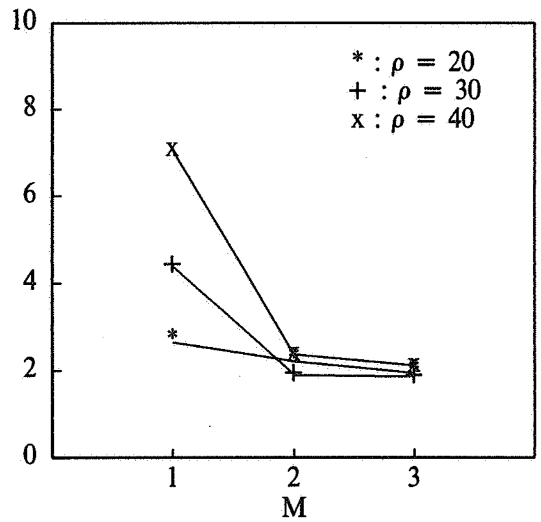


FIGURE 4.1

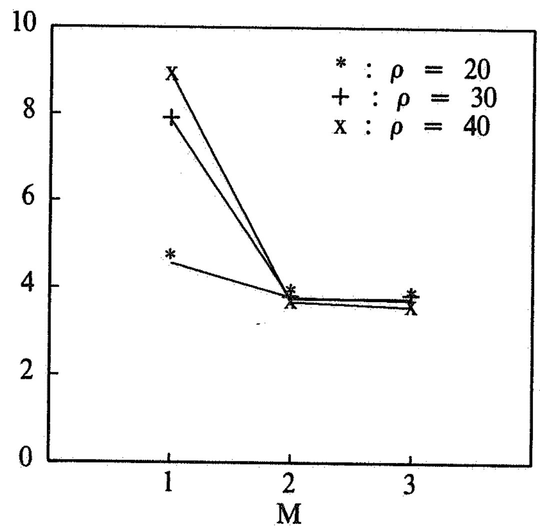


FIGURE 4.2

We may already gain some insight in the number of classes needed by repeating the previous computations for M=2, moving the boundary between the two speed classes, v^1, away from $v^e(\hat{\rho})$. In the limiting cases where either $v^1=0$ or $v^1 \rightarrow \infty$ we are in fact dealing with just one class, expressing itself in a large error variance. So, if we move v^1 away from $v^e(\hat{\rho})$ we expect a gradual deterioration of the estimates.

The results for $\hat{\rho}=30$ veh/km/lane and $v^1 =$ 57.5, 67.5, 77.5 ($= v^e$), 87.5 and 97.5 km/h are shown in table 4.2. Note that in case $v^1=57.5$ and 97.5 km/h only 2 percent of the probability mass of the passing speed distribution is in one class and 98 percent is in the other.

From table 4.2 the following conclusions may be drawn:

- moving away v^1 from $v^e(\hat{\rho})$ generally increases estimator error variance;

- v^1 may be far off the center of the speed distribution (>20 km/h) while the error variance remains small;
- the effect is asymmetric: decreasing v^1 is worse than increasing it. This means that to achieve the same accuracy classes should be narrower at lower speeds or broader at higher speeds.

From the second conclusion it follows that a limited number of speed classes will probably suffice under all circumstances.

v^1 (km/h)	57.5	67.5	$v^e(\hat{\rho})$ 77.5	87.5	97.5
$\tilde{\text{var}}(\rho_1-\hat{\rho}_1)$	8.3	5.0	3.1	2.9	4.2
$\tilde{\text{var}}(\rho_2-\hat{\rho}_2)$	10.8	7.0	3.8	3.7	5.6
$\tilde{\text{var}}(\rho_3-\hat{\rho}_3)$	10.4	7.4	3.9	3.8	5.6
$\tilde{\text{var}}(\rho_4-\hat{\rho}_4)$	9.4	6.0	3.4	3.1	4.5
$\tilde{\text{var}}(v_1-\hat{v}_1)$	28.2	19.0	13.8	16.4	22.2
$\tilde{\text{var}}(v_2-\hat{v}_2)$	29.1	19.6	14.1	16.8	23.6
$\tilde{\text{var}}(v_3-\hat{v}_3)$	28.5	19.6	14.1	16.7	23.6
$\tilde{\text{var}}(v_4-\hat{v}_4)$	30.4	19.2	13.7	16.5	23.5

TABLE 4.2

5. Testing the filter

Now that we have gained some basic knowledge about the filter we want to investigate its behaviour when applied to data. Performance criteria will be proposed in 5.1 before reporting on results of filtering simulated data in 5.2. In 5.3 results with real data from the signalling system will be presented and a validation is carried out in 5.4.

The testing of the filter is carried out in several stages. First the filter is applied to simulated data to obtain some insight in the performance in absence of modelling errors. A separate test is carried out to address the issue of robustness. Next the filter is tested against real data in two steps : first to low density, stationary traffic and next to data showing increasingly severe instabilities. These tests will lead to a structural change in the filter and also in some parameter changes. Next to that some insight is gained in the sensitivity of the filter with respect to the parameters.

An important issue to address is the choice of system noise strength. In our case we can only vary the variance of the acceleration noises $w_i(t)$ in (2.2). Omitting these noises will lead the filter to rely on the model only and discard measurement information on the long run. Large model noises will however lead the filter to rely heavily on the measurements, the model hardly plays a role. In the latter case the filter is insensitive to model parameters. We therefore chose to start with small acceleration noises to gain optimal insight in the effect of model parameters. Later we will increase the noise strength and search for the value that minimizes the estimation errors.

Choosing a small acceleration noise turns out to be problematic in the case of unstable traffic. The filter possesses a line of non-optimal equilibrium points and sometimes ends up in one, leading to a systematic error. In one case the error in the intensity amounted to 60 percent! A detailed explanation of this phenomenon is given in appendix A. Our solution to the problem was to start with a reasonably good estimate of the initial state and set the filter gain to zero. This means that the filter will only adjust the density according to the measured vehicle passages and that the mean speed will exactly follow the model.

Although the topic of speed information was already addressed in the previous section we will seek confirmation of the results found there in tests with real data here. A comparison between the first order and the second order approximation to the optimal filter will also be made. Concluding the tests with real data the performance of the filter on a data set of one hour will be shown and compared to the result of the filter we started with.

5.1 Performance criteria.

We will now discuss several criteria according to which test results in the next subsections will be evaluated.

INTEGRATED SQUARED ERROR

In the case of testing the filter with simulated data the real state is available and direct comparison with the estimates is possible. In this case we will use the *Integrated Squared Error criterion* (ISE) componentwise :

$$\mathrm{ISE}(\hat{x}_i) = \left(\frac{1}{T} \int_{t_0}^{t_0+T} [x_i(t)-\hat{x}_i(t)]^2 dt\right)^{1/2} \qquad (i=1,...,2N) \qquad (5.1)$$

where

T : length of the period considered

t_0 : starting time

x_i : simulated state value (section density or mean speed)

$\hat{x}_i$: estimated state value

N : number of freeway sections considered

LOCAL ESTIMATES

Unforunately, when testing with real data no information about the real state variables is available. We can still use the ISE to compare the results of two filters applied to the same data set, and thereby assess the sensitivity, but we also want to obtain some insight in the actual errors.

We may compute the averages of speed and intensity over measuring locations directly from the data (without using the filter) and compare them with weighted filter estimates. Define

$\overline{v}_i^\delta(t)$: the harmonic mean of the measured passing speeds during the interval $[t-\frac{1}{2}\delta, t+\frac{1}{2}\delta]$ over location i

$\overline{\lambda}_i^\delta(t)$: the number of vehicles that passed location i during the interval $[t-\frac{1}{2}\delta, t+\frac{1}{2}\delta]$, divided by δ (5.2)

$$\hat{\overline{v}}_i(t) := \alpha\hat{v}_i(t) + (1-\alpha)\hat{v}_{i+1}(t)$$

$$\hat{\lambda}_i(t) := (1-\epsilon_i^m+\epsilon_i^f)l_i[\alpha\hat{\rho}_i(t)+(1-\alpha)\hat{\rho}_{i+1}(t)][\alpha\hat{v}_i(t)+(1-\alpha)\hat{v}_{i+1}(t)]$$

for i=0,...,N.

In here δ is the length of the time period over which the detector data is to be averaged, and α a weighting constant.

Now a comparison of $\hat{\overline{v}}_i$ and $\overline{v}_i^\delta$ and of $\hat{\lambda}_i$ and $\overline{\lambda}_i^\delta$ will give some indication of the quality of the estimates. Note that we may compute the ISE for these processes to obtain quantitative results. This procedure is also followed by Cremer [2]. Of course, the results will be sensitive to the choice of the values for δ and α, so these have to be chosen with care.

As for $\overline{v}_i^\delta$ the following argument may suffice. The mean speed $\hat{v}_i$ is defined over the number of vehicles in section i; $\hat{v}_{i+1}$ over the number in section i+1. In all the data sets considered this number will vary between 10 and 40, say. This corresponds to an equilibrium intensity between 1984 and 4096 veh/h respectively. So a time period of 18 to 35 seconds is needed for a group of 10 to 40 vehicles to pass over a measuring location. We conclude that if we take $\delta=25$ seconds the average $\overline{v}_i^\delta$ is taken over about the same group of vehicles as $\hat{\overline{v}}_i$.

For $\overline{\lambda}_i^\delta$ we will take a larger value of δ however (1 minute) to express our interest in the longer term fluctuations of the intensity.

Concerning $\hat{\lambda}_i$ and the choice of α we note the following. Studying the effect of a bunch of vehicles passing a measuring location gives some valuable insight in how to choose a value for α. If α is close to 1.0 the emphasis in $\hat{\lambda}_i$ will be on the upstream section. The density $\hat{\rho}_i$ will have jumps at moments that vehicles enter section i (that is : pass location i−1!) and so $\hat{\lambda}_i$ immediately increases if $\alpha \approx 1.0$. This is somewhat unrealistic because the intensity measured at location i, $\overline{\lambda}_i^\delta$, will only increase after a certain delay, corresponding to the time it takes a vehicle to pass through section i. In short : $\hat{\lambda}_i$ will react too early in comparison to $\overline{\lambda}_i^\delta$ when α is close to 1.0. A complementary effect occurs when vehicles pass location i. The simplest remedy to this problem is to set α equal to 0.5. This will smooth the behaviour of $\hat{\lambda}_i$ and also retard the increase when vehicles pass location i−1 and retard the decrease when the pass location i.

The same value of α will be taken for $\hat{\overline{v}}_i$.

Comparing the estimated speed $\hat{\overline{v}}_i$ and intensity $\hat{\lambda}_i$ with real values in the manner described above is a bit crude. One might argue that what we are doing is to asses the quality of the estimates of a nearly optimal filter, $\hat{\overline{v}}_i$ and $\hat{\lambda}_i$, by comparing them to the values, $\overline{v}_i^\delta$ and $\overline{\lambda}_i^\delta$, of an estimation procedure which we consider to be inferior. We have to be careful in drawing conclusions from these comparisons. One of the situations in which a clear conclusion seems possible will be encountered in experiment 5.9. In this case $\hat{\overline{v}}_i$ is clearly reacting too slow to a speed disturbance. This disturbance has already left section i before $\hat{\overline{v}}_i$ starts to decrease. The local estimate criterion may also be used to detect systematic errors.

INNOVATIONS

A method of assessing the quality of the estimates which is more sound than the previous one is obtained by studying the properties of the innovations process. Recall the filter equation (3.3). The process

$$I_t = N_t - N_{t_0} - \int_{t_0}^{t} E[H(X_s) \mid \mathcal{F}_s^N]ds \tag{5.3}$$

is called the *innovations process* and has some important properties.

THEOREM *For the innovations process defined by (5.3) the following holds :*

i) $E[I_t] = 0$

ii) $E[(I_{t_2}^j - I_{t_1}^j)(I_{t_4}^j - I_{t_3}^j)] = 0$, for $t_1 \leqslant t_2 < t_3 \leqslant t_4$; $1 \leqslant j \leqslant (N+1)M$

iii) $\mathrm{var}[I_t^j - I_s^j] = \int_s^t E[H^j(X_s)]ds = \mathrm{var}[M_t^j - M_s^j]$, for $t_0 \leqslant s \leqslant t$; $1 \leqslant j \leqslant (N+1)M$

PROOF

i) and ii) follow from the fact that $\{I_t,\ t \geq t_0\}$ is a martingale with respect to the history $\{\mathcal{F}_t^N,\ t \geq t_0\}$. Note that $N_t = \int_{t_0}^{t} E[H(X_s) \mid \mathcal{F}_s^N]ds + I_t$ is the Doob-Meyer decomposition of $\{N_t,\ t \geq t_0\}$ with respect to $\{\mathcal{F}_t^N,\ t \geq t_0\}$.

iii) follows by noting that $E[(I_t^j - I_s^j)^2] = E[<I^j,I^j>_t - <I^j,I^j>_s]$ where $<I^j,I^j>$ is the predictable variation of I^j. Now $<I^j,I^j>_t = \int_{t_0}^{t} E[H^j(X_s) \mid \mathcal{F}_s^N]ds$ □

In this theorem i) and ii) are most important. Violation of i) indicates bias and uncorrelated innovation increments indicate non-optimality of the estimates. The sum over j of the variances in iii) is a measure of the distance between the predicted and the real output value on the interval [s,t]. The three properties in the theorem will be checked during the filter tests and are in fact a substitute for checking the maximality of the *likelihood ratio* associated with this estimation problem.

Applying the innovations later on we will compute I_t^j (componentwise) only at jump times, which means that in fact we are studying the process $\{I_{T_n^j}^j,\ n \geq 0\}$ where $\{T_n^j,\ n \geq 0\}$ are the jump times of $\{N_t^j,\ t \geq t_0\}$. For this process the same properties hold as for the original process, thanks to the *optional sampling theorem* [1], assuming that we confine ourselves to considering a finite time interval. This is necessary to satisfy the uniform integrability condition in the theorem. We will not compute the increment of I_t^j over a single jump interval, but over k intervals :

$$I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j, \qquad i = 1,2,...$$

The parameter k will be chosen such that the longer term behaviour of the intensity processes is considered, just as δ was chosen in the local estimate criterion.

We will compute the following statistics :

$$\mu_k^j = \frac{1}{k}\left\{\frac{1}{n}\sum_{i=1}^{n}(I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j)\right\}$$

$$(\sigma^2)_k^j = \frac{1}{k}\left\{\frac{1}{n}\sum_{i=1}^{n}(I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j - k\mu_k^j)^2\right\}$$

$$\rho_{k,l}^j = \frac{1}{nk(\sigma^2)_k^j}\left\{\sum_{i=1}^{n-l}(I_{T_{(i+l)k}^j}^j - I_{T_{(i+l-1)k}^j}^j - k\mu_k^j)(I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j - k\mu_k^j)\right\}$$

The division by k in the above formulas is based upon the following approximate derivations. Assuming that λ^j is constant and $\hat{\lambda}^j \approx \lambda^j$, we have

$$N_{T_{ik}^j}^j - N_{T_{(i-1)k}^j}^j - \int_{T_{(i-1)k}^j}^{T_{ik}^j} \hat{\lambda}^j dt \approx k\left[\frac{\lambda^j - \hat{\lambda}^j}{\lambda^j}\right] \tag{5.4}$$

because $E[T_{ik}^j - T_{(i-1)k}^j] = k/\lambda^j$. So

$$\frac{1}{k}(I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j) \approx \frac{\lambda^j - \hat{\lambda}^j}{\lambda^j}$$

This means that by μ^j we estimate the relative error in $\hat{\lambda}^j$. Furthermore,

$$\mathrm{var}(I_{T_{ik}^j}^j - I_{T_{(i-1)k}^j}^j) \approx k\frac{\hat{\lambda}^j}{\lambda^j} \approx k$$

which means that we expect $(\sigma^2)_k^j$ to be close to one.

Applying the innovations method for each component j of N_t turned out to be problematic in practice : when the real intensity corresponding to N_t^j becomes small and the estimate $\hat{\lambda}^j$ differs from λ^j the resulting innovations increment is far from zero. To see this consider the expression (5.4). If $\hat{\lambda}^j \neq \lambda^j$ and $\lambda^j \rightarrow 0$ then the innovations increment approaches $\pm\infty$. Note that λ^j is the intensity of the counting process corresponding to speed class V^j and may become negligible although traffic is dense. In general we are not interested in the behaviour when λ^j is small, an error in $\hat{\lambda}^j$ is allowed. We have just seen that even a small error in $\hat{\lambda}^j$ may lead to an unbounded error in the innovation increment. It would be preferable to work with the absolute error instead of the relative one as above.

We have experimented with several modifications of the innovations process none of which were very satisfying. The best thing to do turns out to be to select periods a priori for which λ^j is larger than some lower bound and compute the innovations over these periods only. This is quite involved however and the choice of the lower bound is ad hoc. Our final conclusion was to use the innovations only when the intensity cannot become small. This means that we will only compute them per location, summed over all classes j. Only the quality of the intensity estimate is investigated in this way and we will have to use the *local estimate* method to check for the speed estimate. In the following the corresponding statistics will be denoted by μ_k^i, $(\sigma^2)_k^i$ and $\rho_{k,l}^i$ for i=0,...,N.

Finally, we have to note another restriction in the use of the innovations criterion. It has no use trying to optimize the number of speed classes by running different filters and comparing the innovations. The amount of information provided to the filter influences the properties of the innovations process. The more detailed the information is, the more difficult it will be to achieve an uncorrelated process with minimal variance and zero mean. Therefore, a perfect innovations process in case M=1 may be less desirable than a correlated innovations process in case M=2. The two processes just cannot be compared. An illustration of this is in experiment 5.1 in the next subsection, where the mean, variance and correlation in case M=1 are comparable to those in case M=2 (table 5.2), but the ISE's are clearly worse (table 5.1).

5.2 Results with simulated data.

The filter generates its estimates of the state of traffic partly on the basis of measurement information and partly on the basis of the model described in section 2. Using this model we may generate artificial measurements and feed these into the filter.

Details about the simulation procedure and examples are given in [15]. Applying the filter to model generated measurements implies that modelling errors play no role, and we may obtain optimal insight in other errors involved :

- errors due to the 1st or 2nd order approximation to optimal filter (bias);
- errors due to lack of information in the measurements (asymptotical error variance).

A further advantage of using model generated measurements is the availability of the exact state values for comparison with the estimates. We will now present the results of three experiments that illustrate the errors mentioned above.

The second source of error mentioned above was already addressed, in an approximative way, in section 4. There it was found that in case of perfectly stationary and homogeneous traffic the asymptotic estimation error could be considerably reduced by using speed information in the estimation procedure. This result is confirmed by the following experiment.

EXPERIMENT 5.1 (ASYMPTOTICAL ERROR VARIANCE) The traffic model of section 2 is simulated for a stretch of freeway of 4 sections, 0.5 km each, over a period of 15 minutes. All sections have two lanes, there are no on- or off-ramps.

The initial values are $\rho(t_0)=30$ veh/km/lane and $v(t_0)=77.5$ km/h and for the entrance boundary a constant intensity of $\lambda_0=2325$ veh/h/lane was prescribed, whereas for the exit stationarity was assumed. Model parameters are as in table 2.1. The covariance matrix of the acceleration noise was taken to be diag(100) to represent noise occurring in practice. There are no counting errors. The first order filter with the improvements that will be explained in the next subsection is applied to the generated measurements in three different cases : M=1,2 and 3. Recall that M is the number of speed classes. In case M=2 or 3 the classes are chosen like in subsection 4.3 : [0.0,77.5) & [77.5,∞) and [0.0,73.4) , [73.4,81.4) & [81.4,∞) respectively. In stationary traffic an equal amount of probability mass falls in each of the classes in this way and in that sense the choice is optimal. The filter is started in the correct state but with a large initial error variance to represent initial uncertainty about the real state. For all three cases the ISE defined by equation (5.1) is tabulated in table 5.1. Note that the main improvement is in using two speed classes instead of one. Comparing the results here with those of subsection 4.3, table 4.1 shows us that the errors in the speeds are comparable and the errors in the densities are somewhat smaller because of the absence of counting errors. Note that table 4.1 presents variances and table 5.1 presents standard deviations. The statistics presented in table 5.2 illustrate the remark in the last paragraph of the subsection 5.1.

	M		
ISE	1	2	3
$\hat{\rho}_1$	6.4	1.5	1.3
$\hat{\rho}_2$	1.5	0.9	0.9
$\hat{\rho}_3$	2.5	1.5	1.4
$\hat{\rho}_4$	1.9	1.0	1.0
$\hat{v}_1$	16.6	4.1	3.3
$\hat{v}_2$	8.2	3.7	3.8
$\hat{v}_3$	7.7	4.6	4.5
$\hat{v}_4$	6.4	2.9	2.8

TABLE 5.1

	μ^i_{60}		$(\sigma^2)^i_{60}$		$\rho^i_{60,1}$	
	M		M		M	
i	1	2	1	2	1	2
0	0.00	-0.05	0.63	1.11	-0.21	-0.33
1	0.04	-0.03	0.64	0.42	0.22	0.18
2	0.05	-0.00	0.61	0.65	0.08	0.08
3	0.00	-0.02	0.59	0.68	0.09	0.15
4	0.01	-0.01	0.58	0.74	-0.11	0.04

TABLE 5.2

To obtain an idea about the magnitude of bias introduced by the approximations to the optimal filter and about the speed of convergence the following experiment is carried out.

EXPERIMENT 5.2 (BIAS) A series of measurements is randomly generated, not by means of the model but corresponding to a perfectly stationary and homogeneous traffic stream of 15 minutes.
All vehicles drive at constant speed of 77.5 km/h and at a constant distance which is such that the intensity at any location along the freeway equals the equilibrium value (2325 veh/h/lane). Now vehicles pass over section boundaries simultaneously so that the density is constant over time and equal to 30 veh/km/lane. Starting the first order filter that will result from the experiments in subsection 5.3 with a large initial uncertainty we may investigate convergence speed and estimator bias. We compare the cases M=1 and M=2. The final state estimates for one realisation are shown in table 5.3 and the final approximate error variances in table 5.4. Results of another realisation were found to be comparable. We conclude that using two speed classes reduces the bias considerably. The error variances are much smaller also. Convergence of the estimates in table 5.3 was confirmed by

extending the estimation period to 30 minutes. In both case M=1 as well as M=2 the error variances of the densities have to approach zero. The figures in table 5.4 indicate that in case M=1 convergence is much slower.

	M	
	1	2
$\hat{\rho}_1$	28.7	30.3
$\hat{\rho}_2$	26.8	29.4
$\hat{\rho}_3$	27.0	30.0
$\hat{\rho}_4$	25.9	29.3
$\hat{v}_1$	82.4	78.2
$\hat{v}_2$	85.0	78.6
$\hat{v}_3$	86.8	78.1
$\hat{v}_4$	88.0	78.6

TABLE 5.3

	M	
	1	2
$\tilde{\text{var}}(\rho_1 - \hat{\rho}_1)$	4.0	0.3
$\tilde{\text{var}}(\rho_2 - \hat{\rho}_2)$	4.8	0.5
$\tilde{\text{var}}(\rho_3 - \hat{\rho}_3)$	6.6	0.6
$\tilde{\text{var}}(\rho_4 - \hat{\rho}_4)$	2.0	0.3
$\tilde{\text{var}}(v_1 - \hat{v}_1)$	91.1	11.4
$\tilde{\text{var}}(v_2 - \hat{v}_2)$	47.4	12.8
$\tilde{\text{var}}(v_3 - \hat{v}_3)$	40.7	12.7
$\tilde{\text{var}}(v_4 - \hat{v}_4)$	32.6	9.7

TABLE 5.4

To end this section we will now show the improved *robustness* of the filter resulting from using speed information.

EXPERIMENT 5.3 (ROBUSTNESS) A model based series of measurements is generated like in experiment 5.1. The first order filter resulting from subsection 5.3 is applied with an erroneous $v^e(\rho)$-relation : $a = -0.1$ instead of -0.58 and $\rho_{crit} = 105$ instead of 27 veh/km/lane.
From figure 5.1 one notes the difference in using two speed classes instead of one. The dotted line represents the simulated speed values. In case M=2 the modelling error is compensated using the speed measurements. We conclude that using a minimal amount of speed information is necessary to achieve filter robustness with respect to modelling errors.

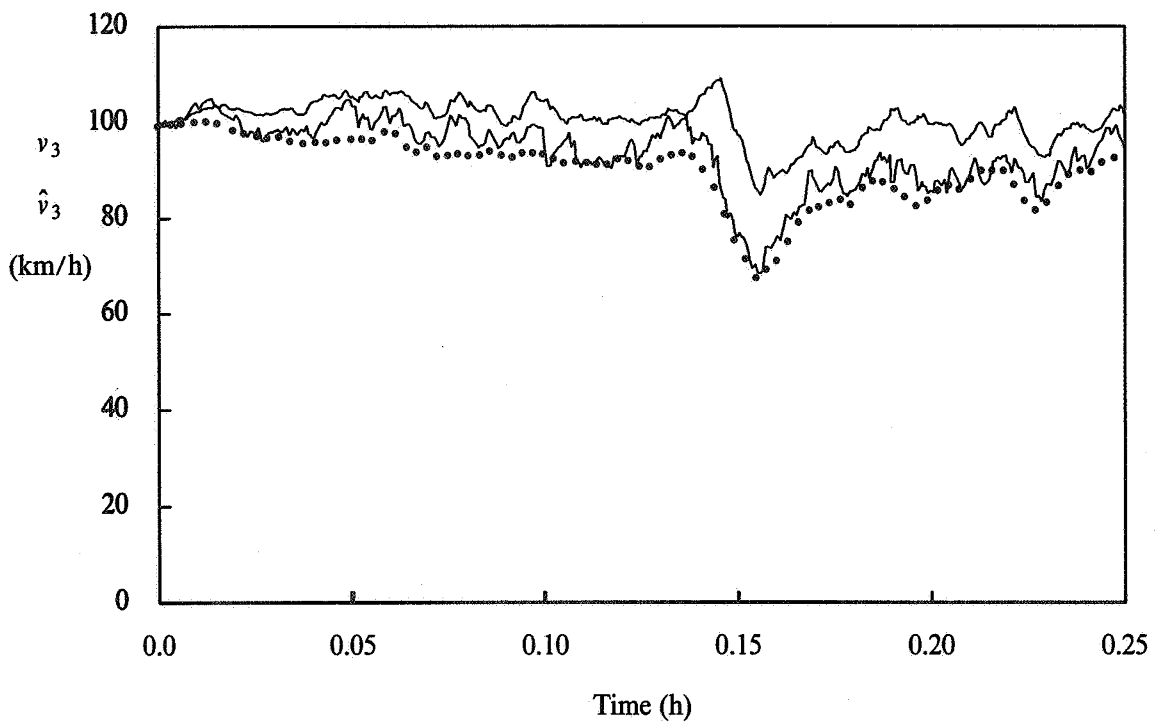


FIGURE 5.1

CONCLUSIONS (SIMULATED DATA) :

- estimator bias is considerable in absence of speed information and almost zero in case speed information is used;
- the asymptotical estimation error is reduced considerably when a limited amount of speed information is used;
- using a minimal amount of speed information is necessary to achieve filter robustness with respect to modelling errors.

5.3 Results with real data.

In the previous subsection filter behaviour is investigated in absence of modelling errors. It is to be expected that modelling errors do play a role in practice: part of the model for freeway traffic is not based on physical laws but on heuristic and limited experimental knowledge about driver behaviour. Therefore experiments with real data are essential to determine the usefulness of the filter in practice. Results of such experiments are presented in this subsection.

Sensitivity with respect to several parameters will be investigated in the following way: first the ISE-criterion will be computed for the estimated state trajectory generated with the nominal parameter value and the trajectory generated with the perturbed parameter. If the ISE is significantly larger than zero for one of the state variables, the innovations will be computed to decide whether the perturbation is beneficial. As the innovations only inform us about the quality of the intensity estimates, the ISE's of $\hat{\bar{v}}_i$ vs. $\bar{v}_i^{\delta}$ and $\hat{\lambda}_i$ vs. $\bar{\lambda}_i^{\delta}$ will be computed as well.

5.3.1 Stationary traffic.

First we will apply the filter to a data set of stationary traffic without serious speed disturbances. The measurements in all the experiments were taken at a two-lane freeway near Utrecht, the northern part of the A12 from km 65.55 to 63.50, on May 9, 1983 from 7:10−7:25 am.

This part of the freeway has no on- or off-ramps. Weather was dry and clear. Cumulating counts over 2 hours of data the relative counting errors among the 5 measuring sites were estimated to be zero: all detectors counted about the same vehicle total. One of the sites was calibrated in an earlier study [17] and shown to miss about 1.5 percent of the vehicles that pass. We will use this value for all locations.

During the period considered traffic is reasonably stationary: speed ranges from about 85 km/h to 110 km/h and the mean intensity is about 1250 veh/h/lane, far below the capacity, which for this part of the freeway was estimated to be about 2215 veh/h/lane [17]. In the experiments we used two speed classes: [0,100) and [100,∞). This amount of speed information proved to be sufficient in a separate test that will not be presented here. The ISE of the estimated state trajectories of the filter with 2 vs. 4 speed classes was negligible.

Model parameters are chosen according to table 2.1. except that ρ_{jam} = 181 veh/km/lane and b = 2835 1/h. This difference is unimportant because the density is not likely to attain a value larger than ρ_{crit} during the period considered. We start with the second order filter in the initial state where all densities are equal to 13.0 veh/km/lane and all section speeds equal the equilibrium value of 97.5 km/h. The initial uncertainty is chosen to be large: $\text{var}[\hat{\rho}_i(t_0)-\rho_i(t_0)]$ = 49 (veh/km/lane)^2 and $\text{var}[\hat{v}_i(t_0)-v_i(t_0)]$ = 49 (km/h)^2. As indicated in the first paragraphs of this section the covariance matrix Σ of the filter is chosen to be small in order to have the filter to rely heavily on the model.

We will now present the results of several tests with the filter.

EXPERIMENT 5.4 (NOMINAL PARAMETER VALUES)

The results in table 5.5 and figure 5.2 show a satisfactory filter behaviour as far as the estimated speed is concerned. The intensity error variance is rather large however and there is a considerable systematic error at the first two locations as well. As it turns out parameter changes will not lead to a significant reduction of this error. It will be possible to remove this error for a large part by changing the structure of the filter, as we will see in the next subsection. The correlations in table 5.6 show a behaviour comparable to that in the case where there are no modelling errors (table 5.2) except for the exit location where heavy correlation exists.

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
	6.4	6.6	5.2	4.2	3.9	693	395	437	371	723

TABLE 5.5

i	μ_{60}^i	$(\sigma^2)_{60}^i$	$\rho_{60,1}^i$
0	-0.10	0.80	0.15
1	-0.10	1.09	-0.15
2	-0.01	0.23	-0.06
3	-0.02	0.26	0.21
4	-0.03	0.26	-0.60

TABLE 5.6

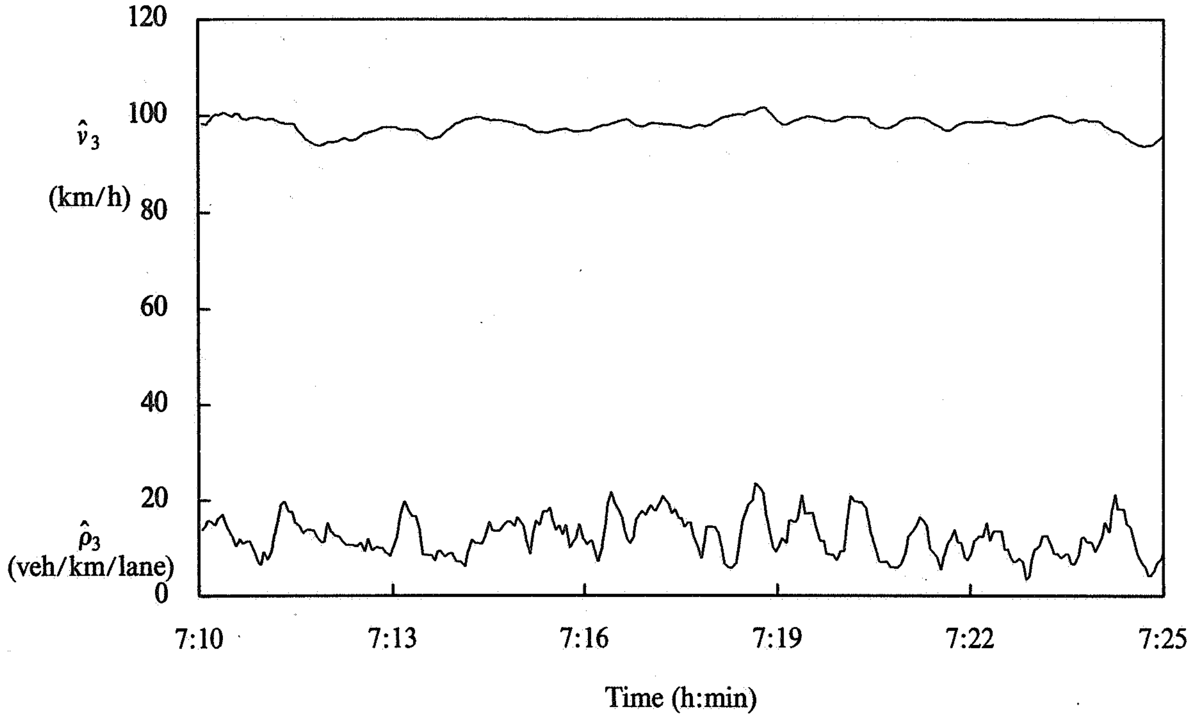


FIGURE 5.2

EXPERIMENT 5.5 (SENSITIVITY W.R.T. T)
The nominal value of T of 0.01 h was changed to 0.0002 h and 0.5 h respectively.
The effect on the estimated state variables is small in case T=0.5h and larger in case T=0.0002h as may be seen from table 5.7. There the ISE's of the state trajectories in case of T=0.0002 and T=0.5 against the nominal trajectory of experiment 5.4 are given. In both cases the relative differences are smaller than 5 percent, so the filter is rather insensitive to the relaxation time.

ISE(.)	$\hat{\rho}_1$ (veh/km/lane)	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$ (km/h)	$\hat{v}_2$	$\hat{v}_3$	$\hat{v}_4$
T = 0.0002h	0.5	0.2	0.1	0.1	4.5	3.3	2.6	2.1
T = 0.5h	0.1	0.2	0.1	0.1	2.8	2.7	2.5	2.3

TABLE 5.7

EXPERIMENT 5.6 (SENSITIVITY W.R.T. α)
The nominal value of α of 0.85 was changed to 0.1.
The effect on both $\hat{\rho}$ and $\hat{v}$ is rather small, see table 5.8. Table 5.9 shows that for $\alpha=0.1$ behaviour is worse than for $\alpha=0.85$.

ISE(.)	$\hat{\rho}_1$ (veh/km/lane)	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$ (km/h)	$\hat{v}_2$	$\hat{v}_3$	$\hat{v}_4$
α = 0.1	1.2	2.1	1.4	0.5	2.2	1.3	0.8	0.5

TABLE 5.8

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
$\alpha = 0.1$	6.2	6.6	5.1	4.2	3.9	847	527	509	434	760

TABLE 5.9

EXPERIMENT 5.7 (SENSITIVITY W.R.T. γ)
The nominal value of γ of 6.5 km/h^2 was changed to 1.0 and 25.0 respectively.
The results are given in table 5.10, table 5.11 and table 5.12. One may note that the filter is more sensitive to this parameter than to the previous ones we investigated. Taking γ = 1.0 km/h^2 gives results that are comparable with γ = 6.5. Taking γ = 25.0 leads to estimates that are worse: the ISE($\hat{\bar{v}}_i$)'s are larger as well as the ISE($\hat{\lambda}_i$)'s. The innovations do not give a clear picture. There is a slight reduction in bias when γ = 25.0 but also an increased variance.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
γ = 1.0 km/h^2	0.5	0.8	0.3	0.2	2.6	1.6	1.3	0.7
γ = 25.0 km/h^2	2.4	1.9	0.7	0.8	8.6	3.9	3.3	1.5

TABLE 5.10

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
γ = 1.0 km/h^2	6.0	5.9	5.1	4.1	4.2	635	381	410	368	706
γ = 25.0 km/h^2	10.6	9.1	5.9	5.3	3.8	756	511	474	409	760

TABLE 5.11

γ = 25.0 km/h^2			
i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
1	0.06	1.14	-0.15
2	0.06	1.50	-0.34
3	0.09	0.42	-0.02
4	0.01	0.30	0.21
5	0.00	0.36	-0.30

TABLE 5.12

EXPERIMENT 5.8 (SENSITIVITY W.R.T. β)
The nominal value of β of 0.5 was changed to 0.1.
The effect is a small, systematic difference in $\hat{v}_i$ for all i, see table 5.13 and figure 5.3. The same effect occurs when β is taken to be 0.9. We conclude that there is no reason to adjust β.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
β = 0.1	0.8	1.6	0.3	0.5	1.7	1.1	0.9	0.5

TABLE 5.13

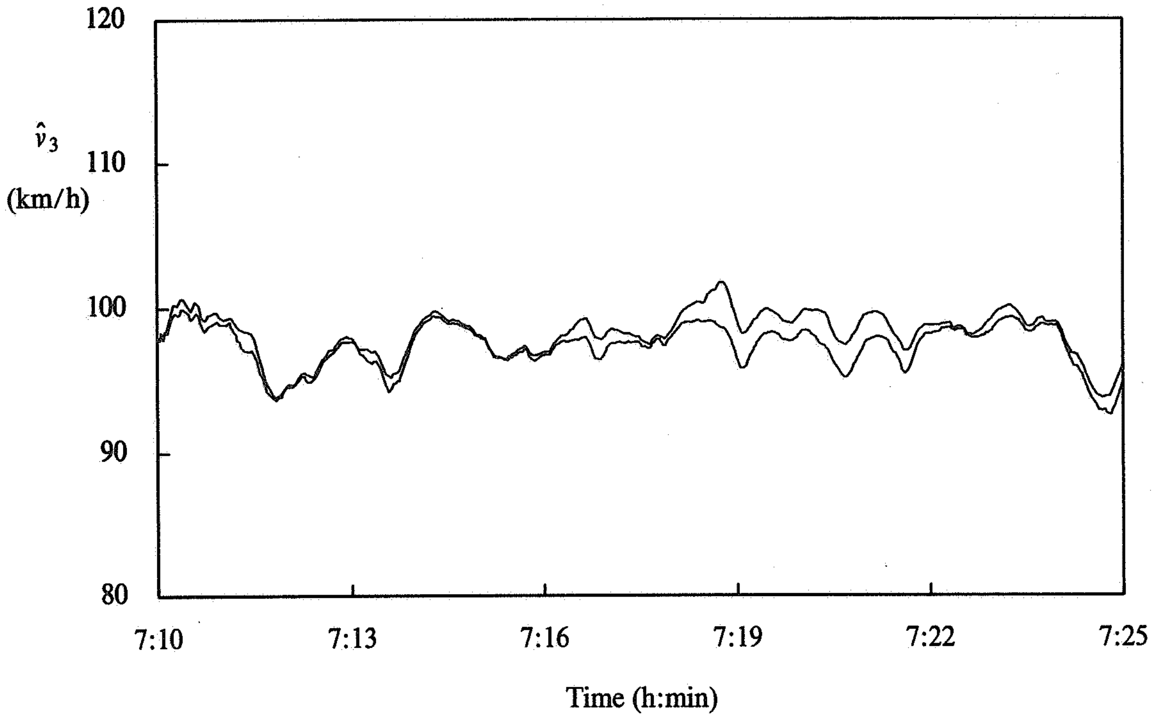


FIGURE 5.3

CONCLUSIONS (STATIONARY TRAFFIC):

- taking more than two speed classes does not lead to a reduction of the estimation error;
- the filter is remarkably insensitive to parameter variations;
- there is no reason to modify the nominal parameter values given in table 2.1;
- the speed estimates are fairly accurate whereas all the intensity estimates have a significant error variance and some show considerable bias.

5.3.2 Unstable traffic.
We will now investigate filter performance in a situation of unstable traffic, in which serious speed disturbances occur. We use measurements from the same freeway stretch as before but now during the period from 8:00–8:15 am.
During this period instabilities occur as may be seen in figure 5.4 where the 25 second averages of passing speeds at several successive measurement locations are shown. The dashed lines correspond to a speed of 90 km/h at the specific location and are drawn at distances of 25 km/h. There are speed drops to 60 km/h, but there is always recovery after some time. The intensity is 1800 veh/h/lane on the average, about 20 percent below the capacity, but reaches the capacity during short periods.

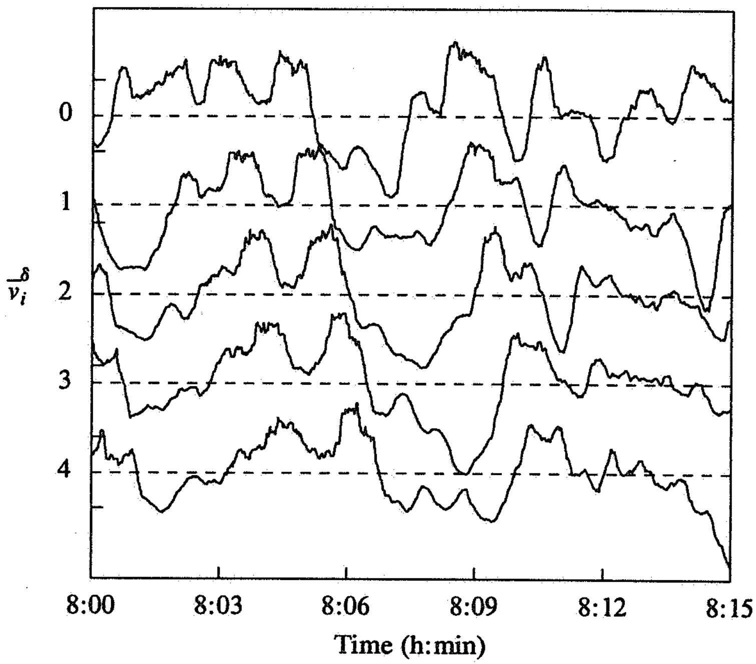


FIGURE 5.4

Before presenting the results we will first discuss a problem that occurred in applying the filter to this set of data and propose a solution.

As mentioned in the beginning of this section the filter possesses non-optimal equilibrium points in case the acceleration noise is small and the filter closely follows the model (Appendix A). To avoid the filter from converging to such a point we decided to start the filter with a small initial uncertainty about the estimates. This prohibits large excursions from these initial values and therefore reduces the probability of ending up in a non-optimal point. Unfortunately this did not work out in practice, probably because of counting and modelling errors which prevent the approximate error covariance matrix $\hat{P}_t^x$ from maintaining a value that is small enough. We therefore resorted to a more crude method: setting the gain matrix Φ_t equal to $\begin{bmatrix} A \\ O \end{bmatrix}$. Then the filter equation (3.5) reduces to

$$d\hat{X}_t = \hat{F}(\hat{X}_t)dt + \begin{bmatrix} A(dN_t - \hat{H}(\hat{X}_t)dt) \\ O \end{bmatrix} \qquad (5.5)$$

or, in more detail,

$$d\hat{\rho}_i(t) = \frac{1}{l_i L_i}\left\{ [dN_{i-1}(t) - dN_i(t)] + (\epsilon^m_{i-1} - \epsilon^f_{i-1})\hat{\lambda}_{i-1}dt - (\epsilon^m_i - \epsilon^f_i)\hat{\lambda}_i dt \right\} \quad (5.6)$$

$$d\hat{v}_i(t) = [\ \cdot\ \cdot\ \cdot\ ]dt \quad (5.7)$$

This means that the densities are adjusted according to the measurements only when vehicles enter or leave a section, apart from a small counting error correction, and that the speed approximately follows the model equation. There are no longer estimate error corrections based on measurement information. This type of approach was also discussed by Van Maarseveen [6]. It should well be noted that this simplification is temporary and only carried out for the benefit of the sensitivity analysis. Later on the full filter equations will be restored and a larger acceleration noise will prevent convergence to a non-optimal equilibrium point.

It is clear from (5.6) and (5.7) that taking more than one speed class in applying this temporarily modified filter is useless, as observations per class are summed by multiplication with the matrix A in (5.5). So we take M=1. We will start with the nominal parameter values of table 2.1. Because of absence of estimate error correction it is necessary to have a reasonable initial estimate of the traffic state. This estimate is generated by applying the filter of subsection 5.3.1 with a large acceleration noise to a 13 minute period of traffic directly preceding the period considered in this subsection. During this 13 minute period traffic is still reasonably stationary and we will get reliable estimates, as indicated by the results of 5.3.1. The resulting estimates are shown in table 5.14 together with the approximate standard deviations obtained from P_t^x.

i	$\hat{\rho}_i$ (veh/km/lane)	$\hat{v}_i$ (km/h)
1	30 ± 1	86 ± 3
2	13 ± 1	94 ± 4
3	19 ± 1	94 ± 4
4	12 ± 1	93 ± 4

TABLE 5.14

EXPERIMENT 5.9 (NOMINAL PARAMETER VALUES)

Results are presented in table 5.15 and 5.16 and figures 5.5, 5.6 and 5.7. The estimates are surprisingly good if we realise that only passing times are available to the filter, which further closely follows the model. No compensation of errors can take place. Speed disturbances are followed by the estimates, but reaction seems to be a bit slow (about 30 seconds late). There is a large intensity error variance. $\hat{\lambda}_3$ shows a systematic error of 7 percent. We conclude that some improvement is necessary, which we will initially try to achieve via parameter changes.

i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
0	0.03	2.75	-0.16
1	0.05	0.83	0.11
2	-0.05	0.62	0.17
3	-0.07	0.76	0.42
4	-0.01	1.00	-0.26

TABLE 5.15

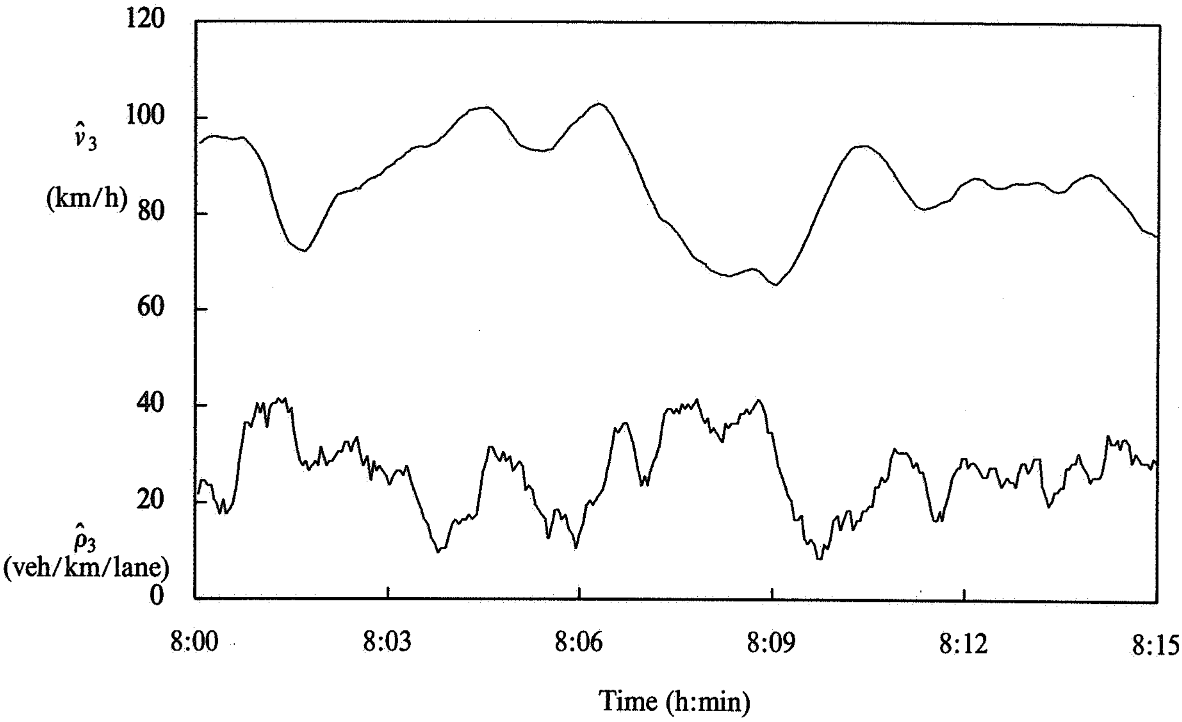


FIGURE 5.5

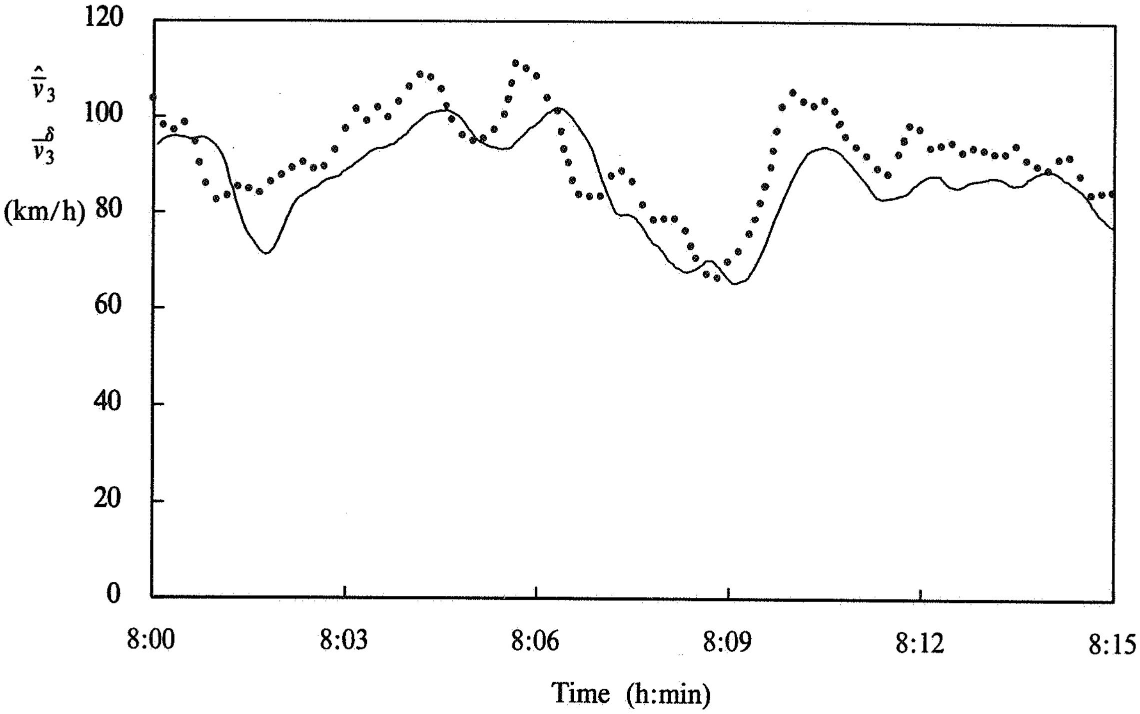


FIGURE 5.6

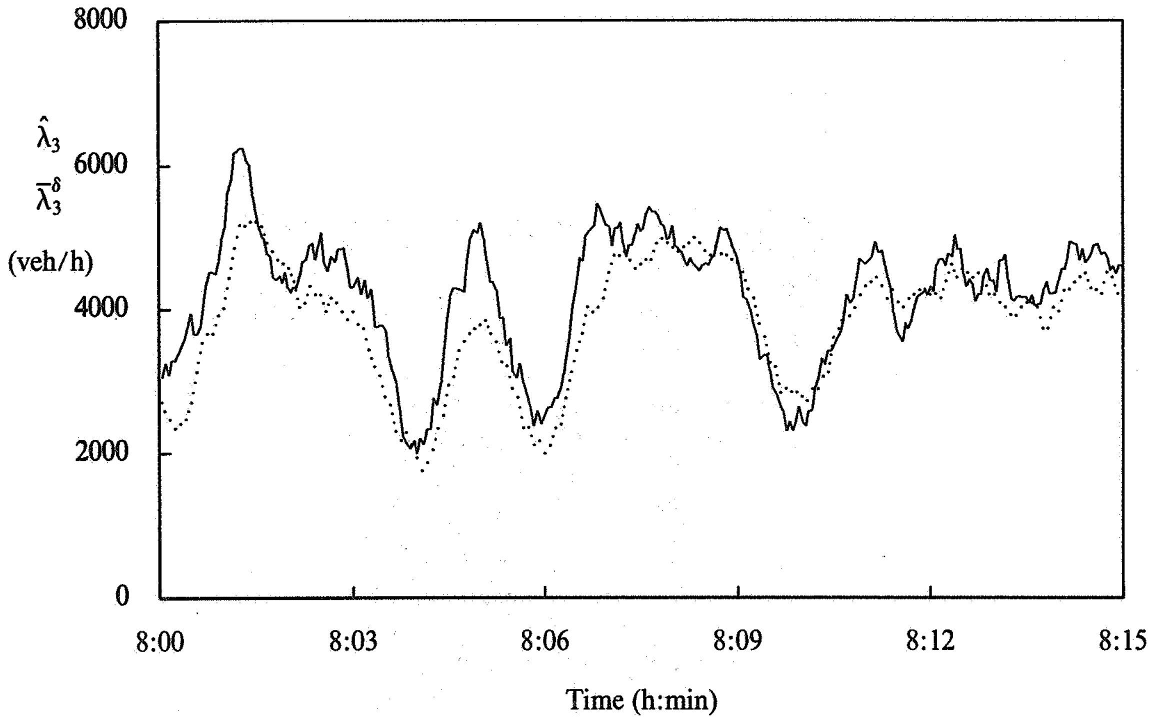


FIGURE 5.7

ISE(.)	$\hat{v}_0$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
	19.8	14.7	10.0	7.1	6.6	1174	709	598	537	774

TABLE 5.16

EXPERIMENT 5.10 (SENSITIVITY W.R.T. ρ_{crit})
The value of ρ_{crit} was changed from 27 to 25 and 30 veh/km/lane respectively. The b-parameter was changed accordingly to maintain continuity of $v^e(\rho)$. Results are shown in table 5.17. Note that it is useless to compute the ISE($\hat{\rho}_i$)'s as these are zero according to (5.6). Our conclusion is that the filter is rather insensitive to ρ_{crit}.

ISE(.)	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
ρ_{crit} = 25 veh/km/lane	2.6	2.8	2.9	2.8
ρ_{crit} = 30 veh/km/lane	2.5	3.0	3.1	3.1

TABLE 5.17

EXPERIMENT 5.11 (SENSITIVITY W.R.T. ρ_{jam})
The value of ρ_{jam} was changed from 110 to 200 veh/km/lane. Table 5.18 shows that the filter is insensitive to this parameter at least for this data set. This is partly caused by the fact that extremely high densities do not occur. To obtain a more definitive idea about

ρ_{jam} it is necessary to apply the filter to congested traffic.

ISE(.)	$\hat{v}_1$ (km/h)	$\hat{v}_2$	$\hat{v}_3$	$\hat{v}_4$
ρ_{jam} = 200 veh/km/lane	1.0	1.2	1.4	1.3

TABLE 5.18

EXPERIMENT 5.12 (SENSITIVITY W.R.T. T)
The value of T was changed from 0.01 to 0.0002 h. The results in table 5.19, 5.20 and 5.21 and figure 5.8 and 5.9 indicate that the filter *is* sensitive to this parameter. The ISE's of $\hat{\bar{v}}_i$ and $\hat{\lambda}_i$ are much better in case T=0.0002h. $\hat{v}_i$ however shows a very wild unrealistic behaviour. The innovations process is better in case T=0.0002 h . The conclusion is that a less drastic decrease of T may be beneficial.

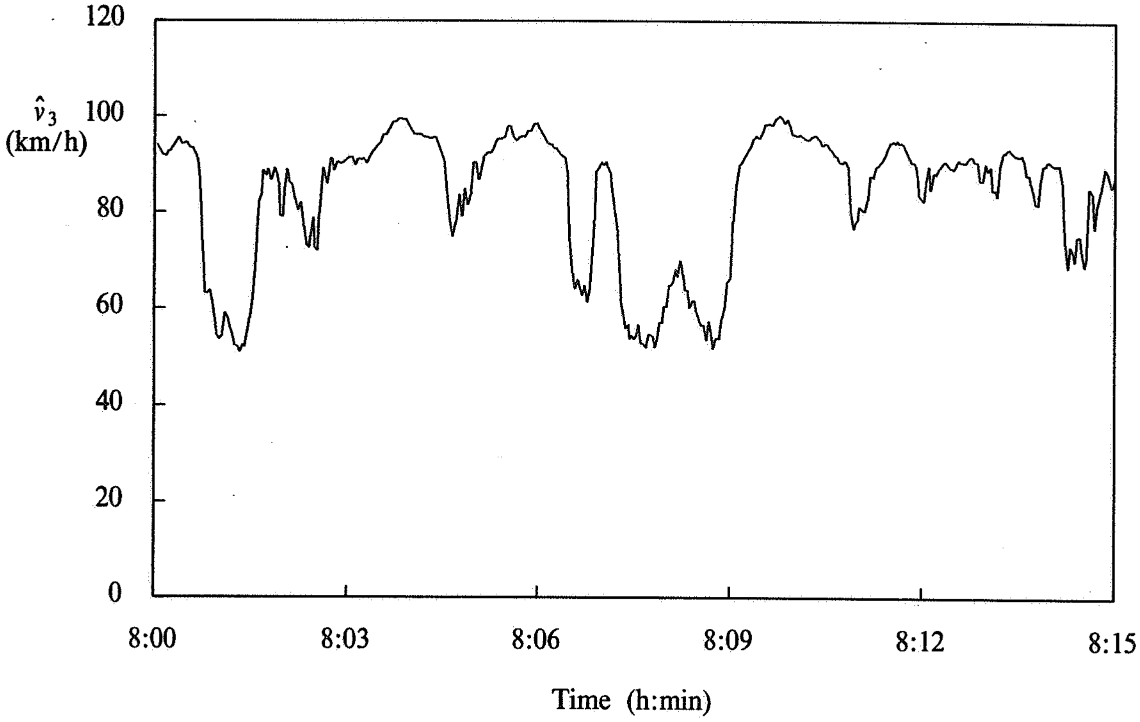


FIGURE 5.8

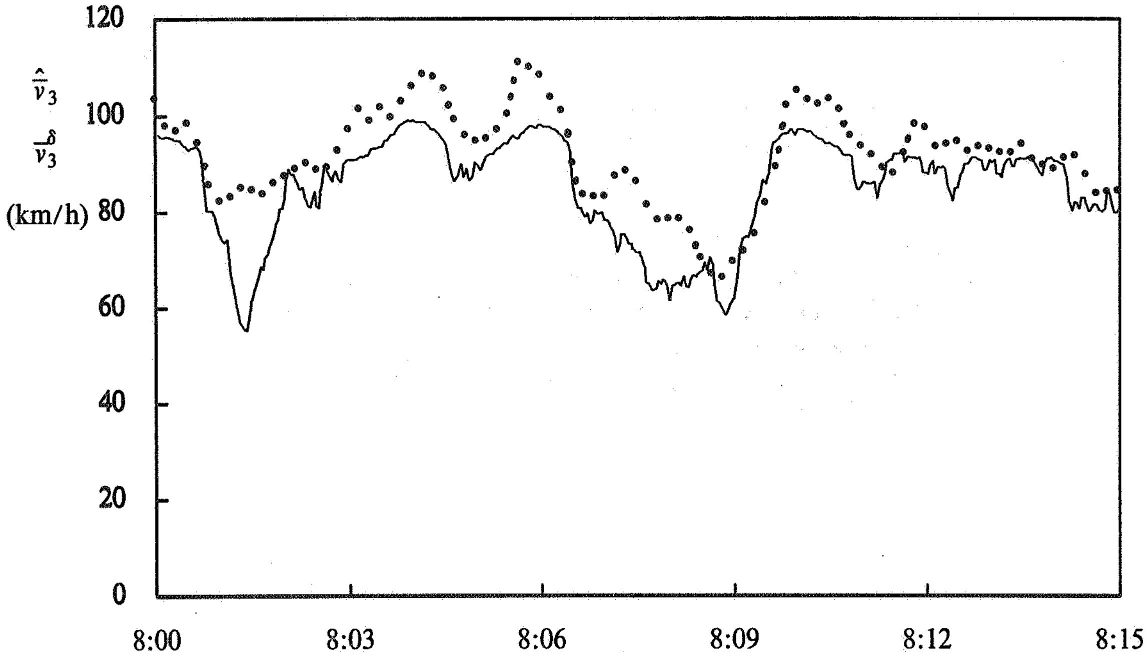


FIGURE 5.9

ISE(.)	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
T = 0.0002 h	19.2	14.1	12.3	7.0

TABLE 5.19

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
T = 0.0002 h	9.8	8.1	7.9	7.5	8.9	712	354	402	432	569

TABLE 5.20

i	μ_{60}^i	$(\sigma^2)_{60}^i$	$\rho_{60,1}^i$
0	0.05	1.25	-0.41
1	0.05	0.22	0.25
2	-0.02	0.65	-0.12
3	-0.04	0.80	0.13
4	-0.00	0.74	-0.54

TABLE 5.21

EXPERIMENT 5.13 (SENSITIVITY W.R.T. γ)
The value of γ was changed from 6.5 to 1.0 km/h^2. The results in table 5.22 and figure 5.10 and 5.11 show that the effect on $\hat{v}_i$ cannot be neglected. A faster reaction to speed

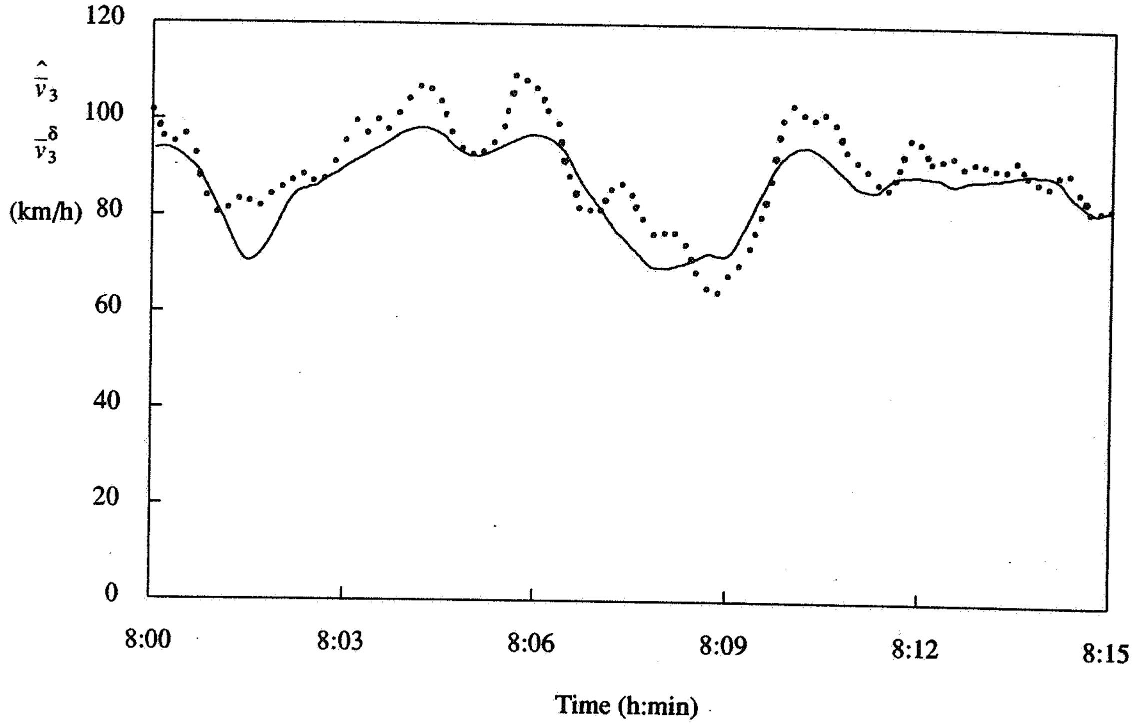


FIGURE 5.10

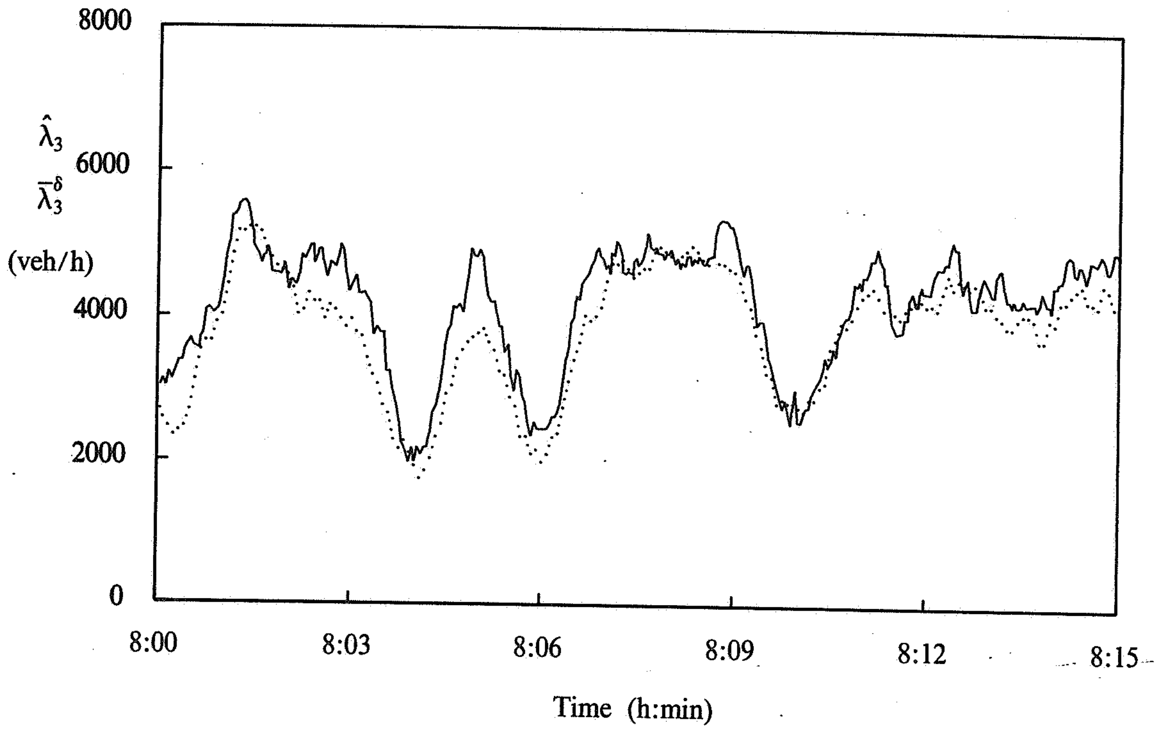


FIGURE 5.11

disturbances is achieved. In contrast to the effect of decreasing T the overall behaviour of $\hat{v}_i$ is not affected and is better than in case $T=0.0002$ h. Compare the ISE($\hat{\bar{v}}_i$)'s. Decreasing γ is preferable to decreasing T and taking $\gamma=1.0$ seems to give satisfactory behaviour. A further reduction of γ was shown to give no further improvement. Note that the systematic error in the $\hat{\lambda}_i$'s is still there. The observed effect of decreasing T or γ is confirmed by a first order analysis presented in Appendix B.

ISE(.)	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
γ = 1.0 km/h^2	11.7	8.1	6.2	3.5

TABLE 5.22

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
γ = 1.0 km/h^2	12.1	8.0	5.5	6.1	6.5	807	426	480	444	690

TABLE 5.23

i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
0	0.03	1.23	-0.50
1	0.03	0.31	0.40
2	-0.06	0.46	-0.07
3	-0.07	0.56	0.22
4	-0.02	0.64	-0.44

TABLE 5.24

Changing β was shown to have the same effect as in experiment 5.8.

In the experiments thus far we have seen that considerable systematic errors in the intensity estimates may occur (10%), which is very unsatisfactory. It is shown that parameter changes do not lead to a significant reduction of these errors. One would expect that increasing the acceleration noise does lead to a reduction, for this would mean that more weight is given to the measurements and correction of $\hat{\rho}$ and $\hat{v}$ should enable a smaller error in the intensity estimates. We investigated this and found that to our surprise the effect was negligible and sometimes the intensity estimates turned out worse. Investigating the filter equations more thoroughly finally led to the conclusion that the errors are caused by a modelling error of the passing speed distribution. This distribution was estimated to depend on the traffic density via its variance, see (2.7). This relation was estimated over a large number of passages ($\pm$10000), during periods of stationary traffic. In practice large deviations from this relation occur during short periods. It turns out that these deviations are serious enough to prevent the filter from estimating the intensity correctly. A more detailed account of this is to be found in Appendix C.

The solution to this problem that we propose is to modify the filter in such a way that the density estimates only react to an error in the total intensity at a location (summed over all speed classes), instead of reacting to errors in intensities per class. This comes down to summing the entries in the upper block of Φ_t, corresponding to the density estimates, over all classes. Define

$$\tilde{\Phi}_{i,kM+j} = \begin{cases} (\sum_{j=1}^{M} \hat{\lambda}_k^j \Phi_{i,kM+j}) / \hat{\lambda}_k & \text{for } 1 \leq i \leq N \\ & \qquad\qquad \text{for } k=0,...,N \\ \Phi_{i,kM+j} & \text{for } N+1 \leq i \leq 2N \end{cases} \tag{5.8}$$

Then our new filter equations are (3.5) to (3.7) where Φ is replaced by $\tilde{\Phi}$. In practice we will compute Φ as before, but before integrating and updating $\hat{X}_t$ we will carry out the summations of (5.8). The following experiment shows that now errors in $\hat{\lambda}$ are compensated for.

EXPERIMENT 5.14 ($\tilde{\Phi}$ INSTEAD OF Φ)
The filter is applied to the same data set as before, but now Φ is no longer set equal to $\begin{bmatrix} A \\ O \end{bmatrix}$. Φ is replaced by $\tilde{\Phi}$ as defined by (5.8). Furthermore , $\gamma=1.0$ km/h^2 and $\alpha=0.5$. The acceleration noise is taken to be small to allow for comparison with the results in the previous experiments. No speed information is used (1 speed class). The initial uncertainty is large, var$[\hat{\rho}_i(t_0) - \rho_i(t_0)]=16$ (veh/km/lane)^2 and var$[\hat{v}_i(t_0) - v_i(t_0)]=16$ (km/h)^2. The innovation results are given in table 5.25 and show that large systematic errors in $\hat{\lambda}_i$ no longer occur.

i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
0	-0.03	1.09	-0.59
1	0.01	0.26	0.19
2	0.02	0.31	-0.21
3	0.01	0.29	-0.01
4	-0.00	0.57	-0.47

TABLE 5.25

As was concluded in section 4, taking speed information into account reduces the estimation error variances and increases robustness with respect to modelling errors. Now, taking M$\geq$1 but Σ small will have hardly any effect: the filter will not put much trust in the presented speed information. We will therefore increase Σ to diag(10000) and take M=4 in the following experiment and investigate whether increasing M further leads to improvement.

EXPERIMENT 5.15 (M=4 VERSUS M=5)
In case M=4 the classes are chosen as [0,75) , [75,85) , [85,95) , [95,∞) and in case M=5 they are [0,70) , [70,80) , [80,90) , [90,100) , [100,∞). The results in table 5.26 show that increasing M to 5 does not lead to significant error variance reduction. The innovations showed no reduction in bias either. Table 5.27 shows the significant improvement with M=4 of the local estimates in comparison to those of experiment 5.9, thanks to the available measurement information.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
M = 5	0.2	0.1	0.1	0.1	1.8	1.3	1.3	1.5

TABLE 5.26

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
M = 4	7.5	5.9	3.9	4.3	3.5	755	389	377	314	680

TABLE 5.27

Now, knowing that enough speed information is provided to the filter, we will investigate whether the second order terms in the approximation to the optimal filter are really necessary, or that a first order filter suffices.

EXPERIMENT 5.16 (1^{rst} VERSUS 2^{nd} ORDER FILTER)
The first order filter is applied with M=4 and compared to the results of the second order filter. Table 5.28 shows that the difference in the estimated state trajectories is negligible. The reduction in computation time however was found to be about 50% . We conclude that using a first order filter is preferable to a second order filter.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
1^{rst} order	0.05	0.01	0.01	0.01	0.4	0.1	0.1	0.1

TABLE 5.28

In section 4 it was suggested that taking two speed classes would probably suffice in most circumstances. This is confirmed by the following experiment.

EXPERIMENT 5.17 (M = 2)
The first order filter is applied with speed classes [0,85) and [85,∞). Table 5.29 shows that the difference with case M=4 is again small.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
M = 2	0.3	0.4	0.2	0.2	2.7	1.9	2.3	2.2

TABLE 5.29

EXPERIMENT 5.18 (M = 1)
It is even true that taking M=1 gives reasonable behaviour, but in this case robustness w.r.t. modelling errors is small and error variances are clearly larger, as follows from table 5.30 compared to table 5.27.

ISE(.)	$\hat{\rho}_1$	$\hat{\rho}_2$ (veh/km/lane)	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$
M = 1	1.6	0.9	0.9	0.6	9.0	5.7	3.9	4.1

TABLE 5.30

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
M = 1	13.7	8.5	5.8	5.8	5.9	769	406	386	306	628

TABLE 5.31

The fact that in case M=1 and even when the correction term of the filter is set to zero ($\Phi = 0$) filter results are reasonable may lead to the conclusion that the density plays a major role and that the speed variable is highly dependent on it. The main task of the filter therefore is to reduce initial uncertainty about these densities and compensate counting errors.

CONCLUSIONS (UNSTABLE TRAFFIC):

- the filter is sensitive to a reduction of T and of γ. Reducing γ to 1.0 leads to better estimates and is to be preferred over a reduction of T;
- the filter is insensitive to ρ_{crit}, ρ_{jam} and β;
- a structural change in the filter equations is necessary to remove a systematic intensity estimate error;
- applying the first order filter is preferable to applying the second order filter;
- using two well-chosen speed classes suffices to produce satisfactory state estimates;
- the evolution of the traffic density largely explains traffic behaviour.

5.4 Validation.
In this subsection the filter that resulted from the tests in preceding subsections will be applied to larger data sets of different dates to check the validity of the results. We will use the same freeway stretch as before (A12 Utrecht, km 65.55 to 63.50) and use data from May 9 and 10 1983. On both days weather is dry and clear.
The experiment on data from May 9 is summarizes results obtained in subsections 5.3.1 and 5.3.2. The experiment on data from May 10 serves as a validation. The filter should work well also on data on which it was not calibrated.

EXPERIMENT 5.19
The period considered is 7:15−8:15 am. The first order filter is applied with $\tilde{\Phi}$, using 3 speed classes: [0,80) , [80,100) and [100,∞). Nominal parameters are used except that $\alpha=0.5$, $\gamma=1.0$ km/h^2 and $\Sigma=\text{diag}(10000)$. The filter is started with a large initial uncertainty. The results are presented in table 5.32, 5.33 and 5.34 and figures 5.12, 5.13 and 5.14. Note the reduced correlation in the innovation increments and the small mean errors in the intensity estimates ($\pm 2\%$). The intensity of the entrance is more difficult to estimate than the others. As one may have noticed the innovation increment correlation criterion has not been of much help until now in testing the filter. Figure 5.15 shows the correlation functions of this experiment and those obtained by running the original filter on the same data set (dashed lines). There is a clear difference in favour of the modified filter.
Running the filter on this data set required almost 6 minutes of CPU-time on the CDC Cyber 750 mainframe.

ISE(.)	$\hat{\bar{v}}_0$	$\hat{\bar{v}}_1$	$\hat{\bar{v}}_2$ (km/h)	$\hat{\bar{v}}_3$	$\hat{\bar{v}}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
	5.5	5.1	3.6	3.6	5.1	742	388	399	358	688

TABLE 5.32

i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
0	-0.07	1.08	-0.01
1	-0.02	0.32	0.07
2	0.02	0.22	0.04
3	-0.02	0.19	0.13
4	-0.02	0.60	-0.25

TABLE 5.33

t = 8:15 h	1	2	3	4
$\tilde{\text{var}}[\hat{\rho}_i - \rho_i]$	1.4	3.6	3.0	1.7
$\tilde{\text{var}}[\hat{v}_i - v_i]$	13.3	15.1	17.6	16.9

TABLE 5.34

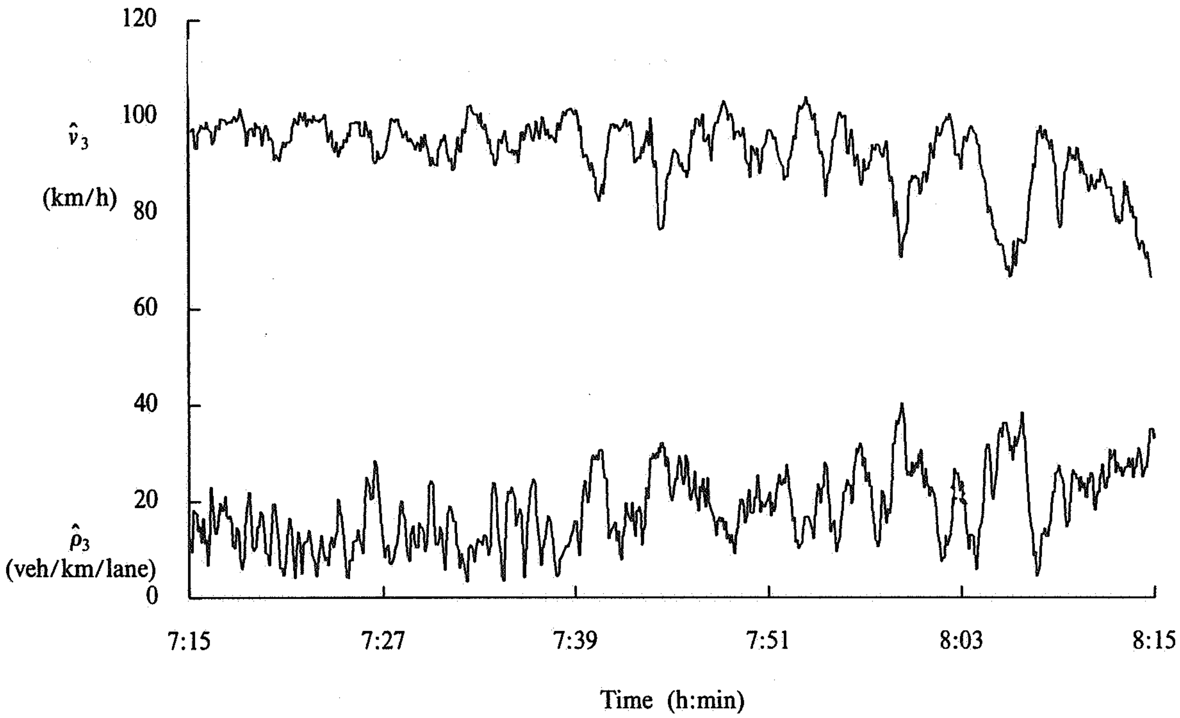


FIGURE 5.12

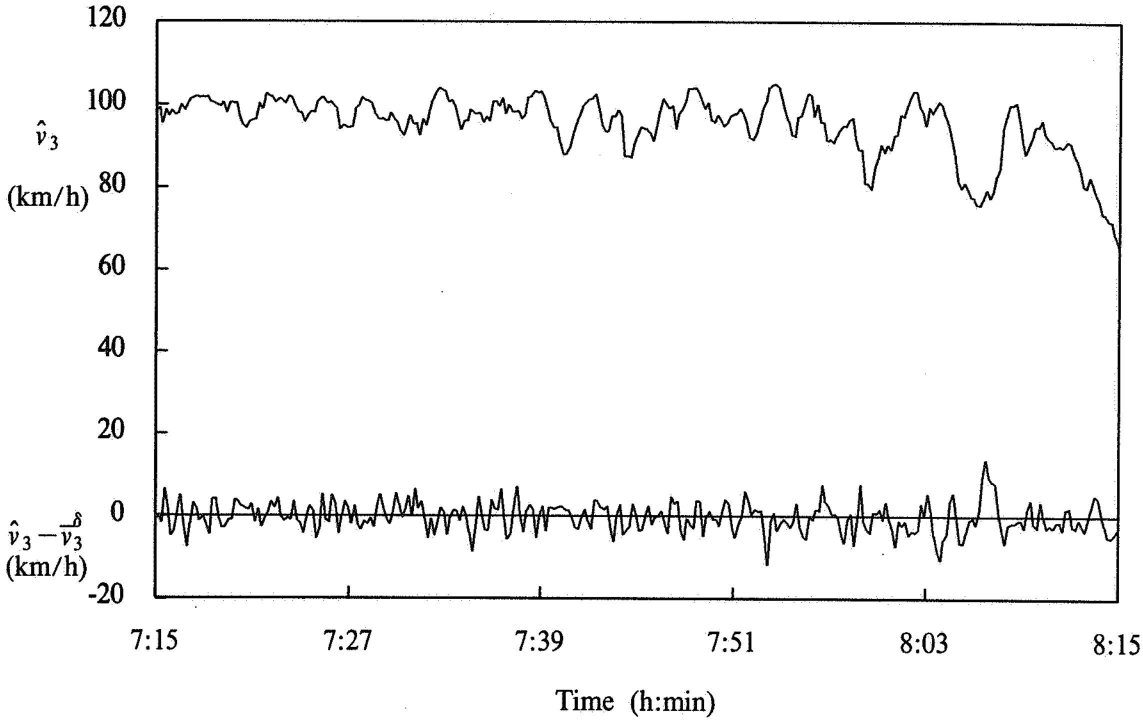


FIGURE 5.13

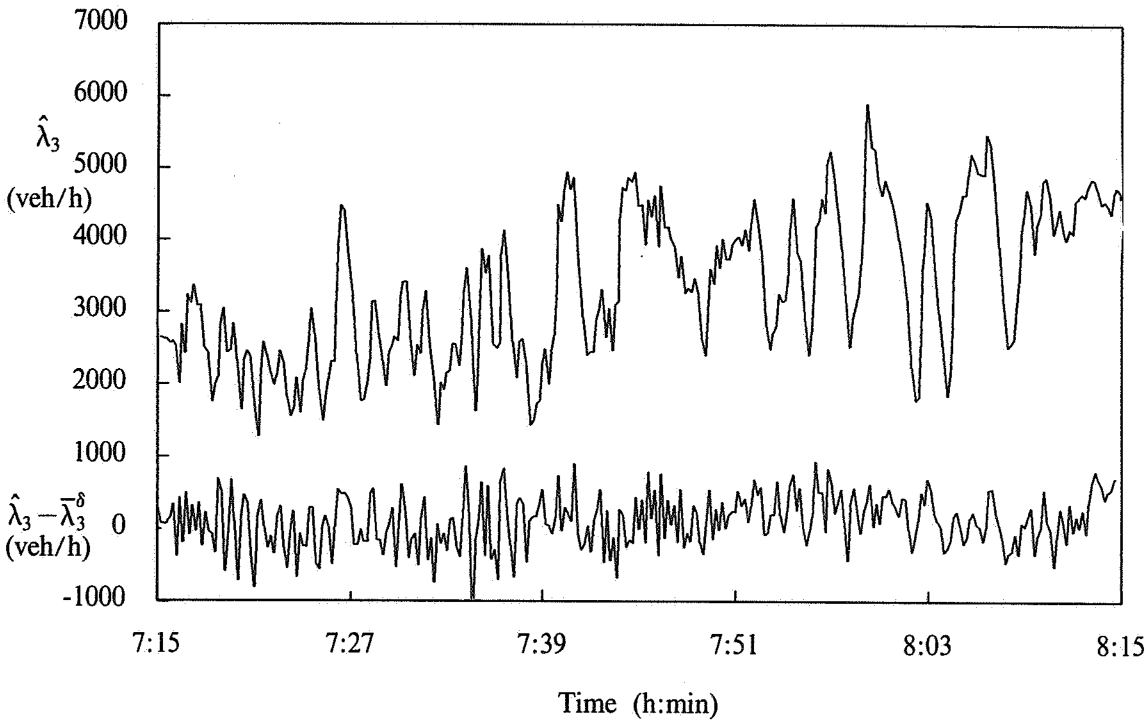
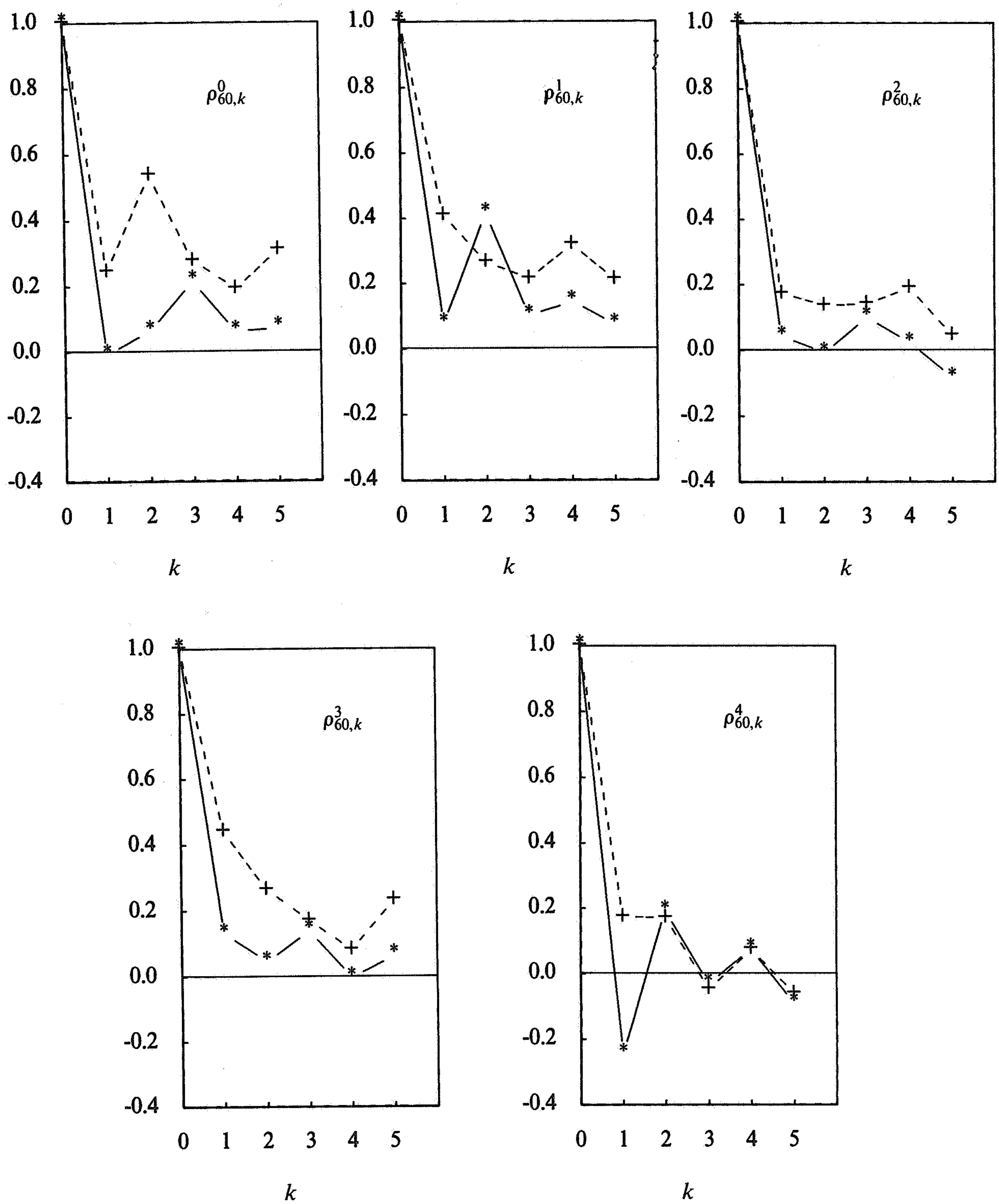


FIGURE 5.14



1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
$\rho^0_{60,k}$
$\rho^1_{60,k}$
$\rho^2_{60,k}$
$\rho^3_{60,k}$
$\rho^4_{60,k}$
0 1 2 3 4 5
k

FIGURE 5.15

EXPERIMENT 5.20

The filter is now applied to data of May 10, 1983 and the period considered is 7:45−8:45 am. This time 2 speed classes are used: [0,80) and [80,∞). The results are in table 5.35 and 5.36 and figure 5.16. Despite the fact that traffic is more dense than in the previous experiment (more instabilities) and only two speed classes are used the results are still satisfactory.

ISE(.)	$\hat{v}_0$	$\hat{v}_1$	$\hat{v}_2$ (km/h)	$\hat{v}_3$	$\hat{v}_4$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$ (veh/h)	$\hat{\lambda}_3$	$\hat{\lambda}_4$
	9.2	5.8	5.1	5.3	4.8	888	427	439	372	784

TABLE 5.35

i	μ^i_{60}	$(\sigma^2)^i_{60}$	$\rho^i_{60,1}$
0	-0.06	1.55	-0.32
1	-0.04	0.37	0.01
2	0.01	0.25	0.08
3	0.01	0.20	0.20
4	-0.02	0.71	-0.04

TABLE 5.36

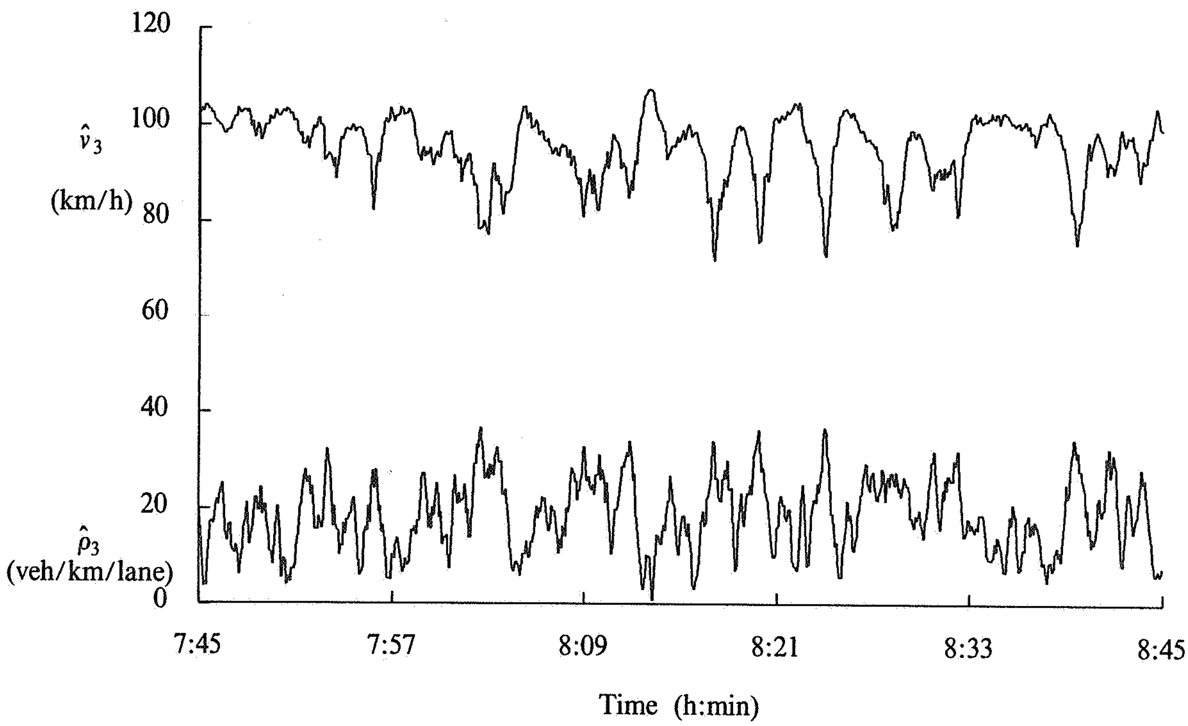


FIGURE 5.16

6. Conclusions

In this report an approximate filter has been derived for the recursive estimation of the state of traffic on a freeway. This filter is based upon a model of freeway traffic flow and uses measurement information from several locations along the freeway. It produces approximations to the optimal conditional expectation estimators.

Detectability and stabilizability of suitable matrix pairs in the filter is investigated and leads to conclusions about the asymptotical behaviour of the error covariance matrix. The effect of using speed information in the filter is investigated by computing the asymptotical error covariance matrix for several different cases. It is concluded that using a limited amount of speed information is necessary to reduce bias and asymptotical error variance and suffices to produce satisfactory estimates.

The proposed filter is tested against simulated and against real traffic data. Performance was evaluated using three criteria : the integrated squared error criterion, the local estimate criterion and the innovation increments criterion. The test against simulated data confirmed the results mentioned in the previous paragraph and also show that using speed information is necessary to obtain robustness with respect to modelling errors. The tests against real data, obtained from the Traffic Engineering Division of the Dutch Ministry of Transport, led to a structural change in the filter and to some parameter changes, resulting in reduced bias and reduced error variance. The filter turned out to be rather insensitive to most parameters. Again a limited amount of speed information proved to be sufficient to produce satisfactory estimates. The evolution of traffic density apparently largely explains traffic behaviour and the speed variable is highly dependent on it. A first order approximation to the optimal filter was shown to be sufficiently accurate. A validation of the results was carried out by applying the modified filter to two sets of one hour of real data.

We will now suggest several subjects for future research.

In order to make the filter more practicable, the possibility of on- and off-ramps has to be taken into consideration. This will be done in the near future. Also, the effect of different weather conditions should be modelled.

To reduce the computational effort required the filter implementation should be simplified. A reduction of approximately 80% in computations would result from implementing a constant or parametrized filter gain matrix. The effect of such a modification of the filter on the accuracy of the estimates has to be investigated. Another considerable reduction of computational effort would result from a simplified implementation of the passage speed distribution.

To improve the quality of the estimates further a different entrance boundary condition could be considered, or a more precise identification of the parameters to which the filter is sensitive might be carried out.

Finally, the behaviour of the filter under congestion conditions may be studied. The effect of omitting measurement information from a location could be considered also.

Acknowledgements

I would like to thank Dr. J.H. van Schuppen, J.A.C. van Toorenburg and the other members of the "Gebruikerscommissie" for their support during the last year and the Traffic Engineering Division of the Dutch Ministry of Transport for providing the necessary and valuable traffic data.

These investigations are supported by the Technology Foundation (STW).

REFERENCES

1. P. BRÉMAUD (1981). *Point processes and queues. Martingale dynamics.*, Springer Verlag, New York.
2. M. CREMER (1979). *Der Verkehrsfluss auf Schnellstrassen*, Springer Verlag, Berlin.
3. C. DELLACHERIE and P.A. MEYER (1975). *Probabilités et Potentiel*, Hermann, Paris.
4. V. KUÇERA (1972). A contribution to matrix quadratic equations, *IEEE Trans. Automatic Control*, AC-17, 344-347.
5. V. KUÇERA (1973). A review of the matrix Riccati equation, *Kybernetika*, 9, 42-61.
6. M.F.A.M. VAN MAARSEVEEN (1982). *Application of martingales in stochastic systems theory - Surveillance and control of freeway traffic flow,* Thesis, Technical University of Twente, Enschede.
7. P.S. MAYBECK (1979). *Stochastic models, estimation and control : Volume 1,2 and 3*, Academic Press, New York.
8. C. MOLER, M. VANBEGIN, and P. VAN DOOREN (1985). *MATLAB-SC Users' Guide,* Technical Note N 166, Philips Research Laboratory, Brussels.
9. M. MORF, B. LÉVY, and T. KAILATH (1978). Square-root algorithms for the continuous time linear least-square estimation problem, *IEEE Trans. Automatic Control*, AC-23, 907-911.
10. Y. OSHMAN and I.Y. BAR-ITZHACK (1985). Eigenfactor solution of the matrix Riccati equation - a continuous square root algorithm, *IEEE Trans. Automatic Control*, AC-30, 971-979.
11. H.J. PAYNE (1971). Models of freeway traffic and control, in *Mathematical Models of Public Systems. Simulation Council Proceedings*, 51-61.
12. M. POUBELLE, I.R. PETERSEN, M.R. GEVERS, and R.R. BITMEAD (1986). A miscellany of results on an equation of Count J.F. Riccati, *IEEE Trans. Automatic Control*, AC-31, 651-654.
13. H. REMEIJN (1982). *The Dutch Motorway Control and Signalling System*, Traffic Engineering Division, Dutch Ministry of Transport, The Hague.
14. T.J. RICHARDSON and R.H. KWONG (1986). On positive definite solutions to the algebraic Riccati equation, *System & Control Letters*, 7, 99-104.
15. S.A. SMULDERS (1986). *Modelling and simulation of freeway traffic flow,* Report OS-R8615, Centre for Mathematics and Computer Science, Amsterdam.
16. C.E. DE SOUZA, M. GEVERS, and R.R. BITMEAD (1986). Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices, *IEEE Trans. Automatic Control*, AC-31, 831-838.
17. J.A.C. VAN TOORENBURG (1983). *Homogeniseren. Effekt van aangepaste adviessnelheid op de verkeersafwikkeling,* Report, Traffic Engineering Division, Dutch Ministry of Transport, The Hague.
18. M. VERHAEGEN and P. VAN DOOREN (1986). Numerical aspects of different Kalman filter implementations, *IEEE Trans. Automatic Control*, AC-31, 907-917.
19. W.M. WONHAM (1968). On a matrix Riccati equation of stochastic control, *SIAM J. Control*, 6, 681-697.

A. Non-optimal filter equilibrium points

In this appendix we will illustrate the problem mentioned in section 5: in case the acceleration noise is small the filter may end up with estimates that contain a large error. Due to the small noise hardly any corrections to the error are made, there is hardly any weight on the measurements. In the first period after the start the filter will however adjust its estimates, provided the initial uncertainty is large. It is during this starting phase that convergence to an incorrect point occurs. The following example illustrates this.

Example

The filter is run on the same set of data as in subsection 5.3.2, but now Φ_t is not set to zero. There is a large initial uncertainty, to allow for corrections during the starting phase. Four speed classes are taken: [0,75) , [75,85) , [85,95) , [95,∞). Parameter values are taken from table 2.1. The filter is run twice, with $\gamma = 6.5$ and $\gamma = 1.0$ km/h^2 respectively. The results in figures A.1 and A.2 show the large systematic difference in $\hat{\rho}_2$ and $\hat{\lambda}_2$ produced by the two filters.
This difference is not to be explained by the different values of γ. It occurs during the first 1.5 minutes and remains constant thereafter. The mean of the innovation increments, μ_{60}^2, was computed and showed that the systematic error in $\hat{\lambda}_2$ amounts to 65% in case $\gamma = 1.0$. The errors in $\hat{\lambda}_1$ and $\hat{\lambda}_3$ are much smaller: +6% and −4% respectively.
The occurrence of such large errors as in the example above may be explained as follows. In case there are hardly any counting errors and the acceleration noise is small, it follows from the conclusions of section 4 that $\lim_{t\to\infty} P_t^x \approx 0$. From (3.6) we then conclude that $\Phi_t \approx \begin{bmatrix} A \\ O \end{bmatrix}$. This again means that the filter equations reduce to

$$d\hat{\rho}_i(t) = \frac{1}{l_i L_i}\left[dN_{i-1}(t) - dN_i(t)\right] \tag{A.1}$$

$$d\hat{v}_i(t) = -\frac{1}{T}\left[\hat{v}_i - v^e(\hat{\rho}_i)\right]dt + \cdots + \text{second order terms} \tag{A.2}$$

From (A.1) we see that $\hat{\rho}_i(t)$ now merely follows the detected vehicle passages and from (A.2) it follows that the $\hat{v}_i(t)$ approximately follows the model. No correction to an error in $\hat{\lambda}_{i-1}$ or $\hat{\lambda}_i$ is possible, which explains why the large errors in the example are not compensated for. The error is generated during the short period when still $P_t^x \neq 0$ and $\Phi_t \neq \begin{bmatrix} A \\ O \end{bmatrix}$. Apparently the covariance matrix P_t^x converges too fast to allow for a correct estimate. It is clear that with the optimal filter the error could not occur, as it produces unbiased estimates. The errors therefore have to be adjusted to the approximations that we made to the optimal filter.

B. First order sensitivity analysis

Consider the equation for $v_i(t)$ where we assume that the density in the neighbouring sections is constant and equals $\rho < \rho_{crit}$ and the speed in those sections is $v^e(\rho)$:

$$\frac{dv_i(t)}{dt} = -\frac{1}{T}\left[v_i(t) - v^e(\rho_i(t))\right] - \gamma\rho(\rho - \rho_i(t)) + 2v(v - v_i(t)) \tag{B.1}$$

Linearisation about the stable equilibrium point $(\rho, v^e(\rho))$ leads to

$$\frac{d}{dt}(\delta v_i) = \left(-\frac{1}{T} - 2v\right)\delta v_i + \left(\gamma\rho - \frac{a}{T}\right)\delta\rho_i \tag{B.2}$$

where $\delta v_i = v_i - v^e(\rho)$ and $\delta\rho_i = \rho_i - \rho$ and we have used the expression (2.3) for v^e. Now (B.2) is a linear system with input process $\delta\rho_i(t)$. The response to $\delta\rho_i(t) = \delta\rho \sin(\omega t + \phi)$

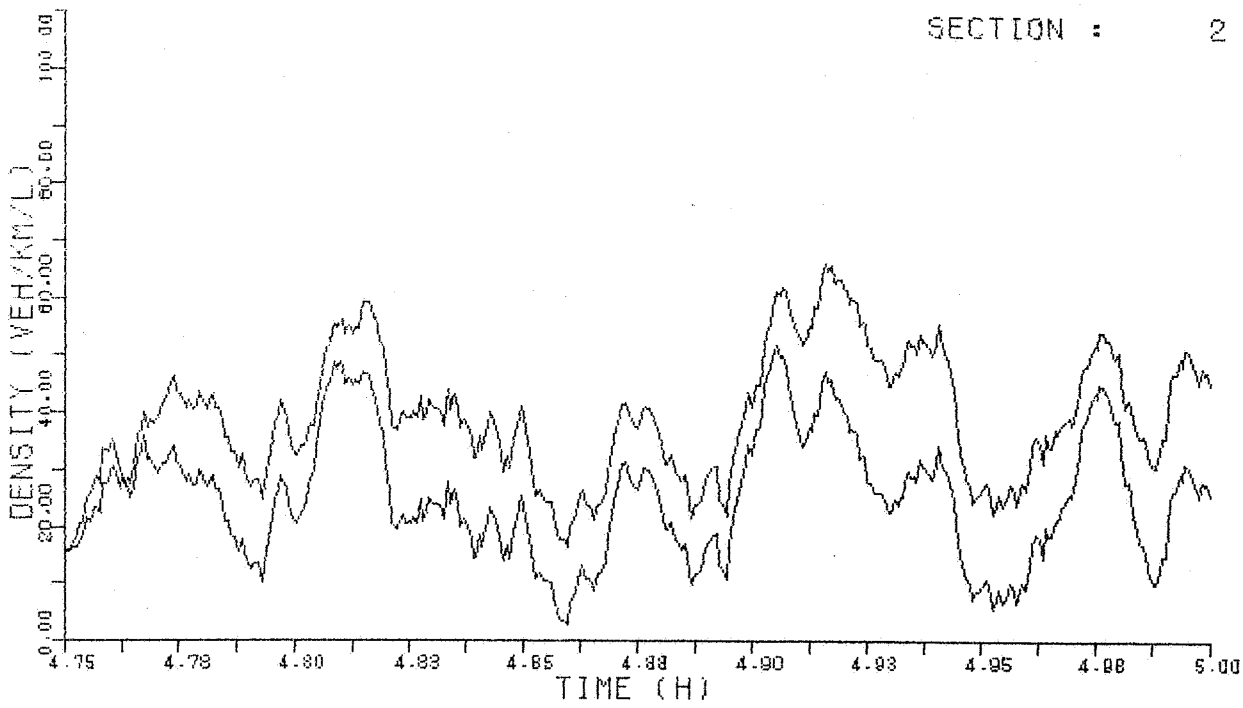


FIGURE A.1

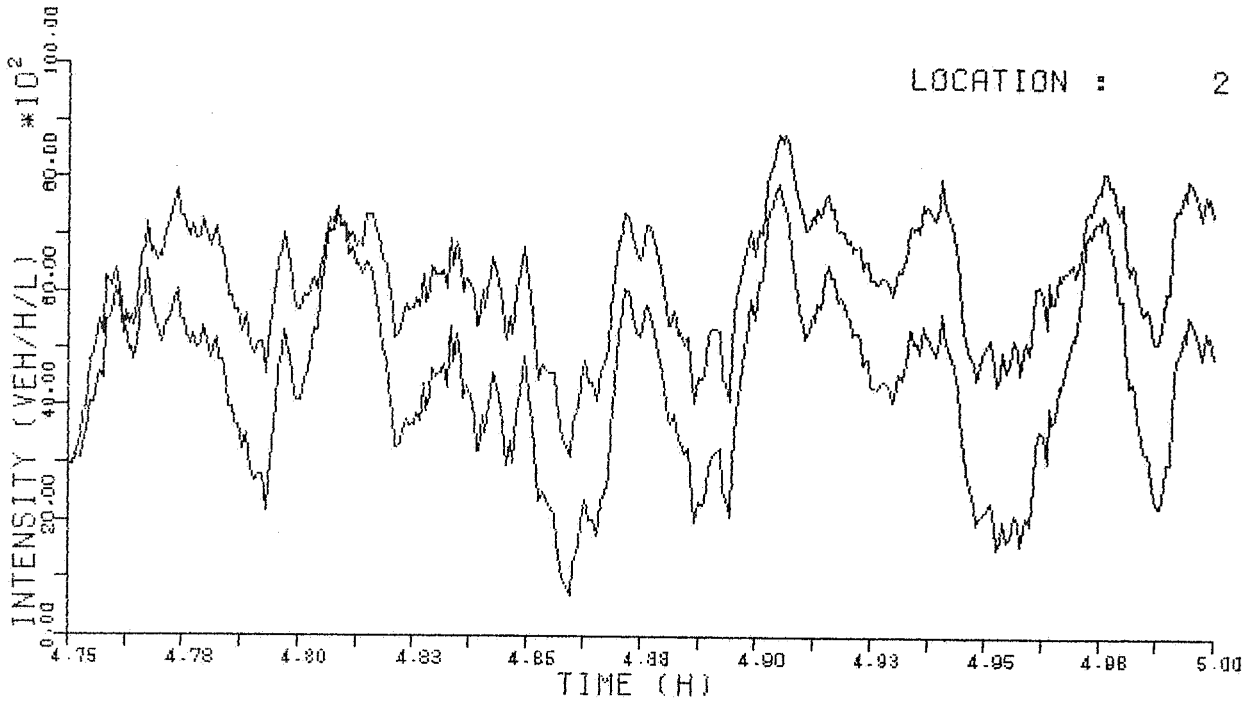


FIGURE A.2

will be

$$\delta v_i = |H(i\omega)| \; \hat{\delta\rho} \; \sin(\omega+\phi+\psi)$$

where $\psi = arg[H(i\omega)]$ and H is the transfer function from $\delta\rho_i$ to δv_i. Now

$$H(i\omega) = \frac{\gamma\rho - \frac{a}{T}}{i\omega + \frac{1}{T} + 2v}$$

so

$$\psi = arg[H(i\omega)] = arctan(\frac{\omega}{-\frac{1}{T}-2v}) + arg(\gamma\rho - \frac{a}{T})$$

We expect the model (B.1) to be such that an increase of ρ_i leads to a decrease of v_i, possibly after a short delay. This means that ψ should be close to π. In section 5 it turned out that for the given parameters the delay was too large. A necessary condition in our model (B.1) for ψ to be able to approach π is that

$$\gamma\rho < \frac{a}{T}$$

For the parameters in section 5 this condition is not satisfied. Decreasing γ or T should lead to better results. According to the condition just mentioned, γ should be less than $58/\rho \approx 2$ km/h^2 if T=0.01 h. And T should be less than $0.09/\rho \approx 0.004$ h if γ=6.5 km/h^2.

C. Systematic estimation errors

During the experiments is subsection 5.3.2 it was noted that considerable systematic errors in the intensity estimates occurred over long periods of time. This was the case even when the acceleration noise was large, enabling the filter to correct the errors using the measurements. Apparently the filter is not always able to perform the right corrections. An explanation may be found in the following analysis.

Consider the case of two speed classes and assume that $l_i=2$, $L_i=0.5$, that there are no counting errors, and $\hat{\lambda}^j_k=\lambda^j_k$ for all k≠i and j=1,...,M. Let us furthermore assume that $\hat{\lambda}_i>\lambda_i$ and that $\overset{\delta}{v}_i \approx \hat{v}_i$. This means that we assume a significant error in $\hat{\lambda}_i$ to exist, but that the estimated local speed is reasonable. We will neglect the stochastic effects and put $dN_k \equiv \lambda_k dt$ for all k. After some manipulations with the filter equations (3.5) and (3.6) one then ends up with

$$d\hat{\rho}_i(t) = \left[\hat{\Phi}_{i,2i+1}+1\right](dN^1_i - \hat{\lambda}^1_i dt) + \left[\hat{\Phi}_{i,2i+2}+1\right](dN^2_i - \hat{\lambda}^2_i dt) \tag{C.1}$$

where

$$\hat{\Phi} = \Phi - \begin{bmatrix} A \\ O \end{bmatrix}$$

Now suppose that $\hat{\lambda}^1_i >> \hat{\lambda}^2_i$ which means that most of the probability mass of the passing speed distribution is in class V^1. Furthermore assume that $\hat{\lambda}^1_i=\lambda^1_i$ and $\hat{\lambda}^2_i>\lambda^2_i$. This means that all of the excess intensity $\hat{\lambda}_i-\lambda_i$ falls in class V^2. By the way, this implies that the modelled passing speed distribution is wrong. See figure C.1.

Now

$$d\hat{\rho}_i = \left[\hat{\Phi}_{i,2i+2}+1\right](\lambda^2_i - \hat{\lambda}^2_i)dt$$

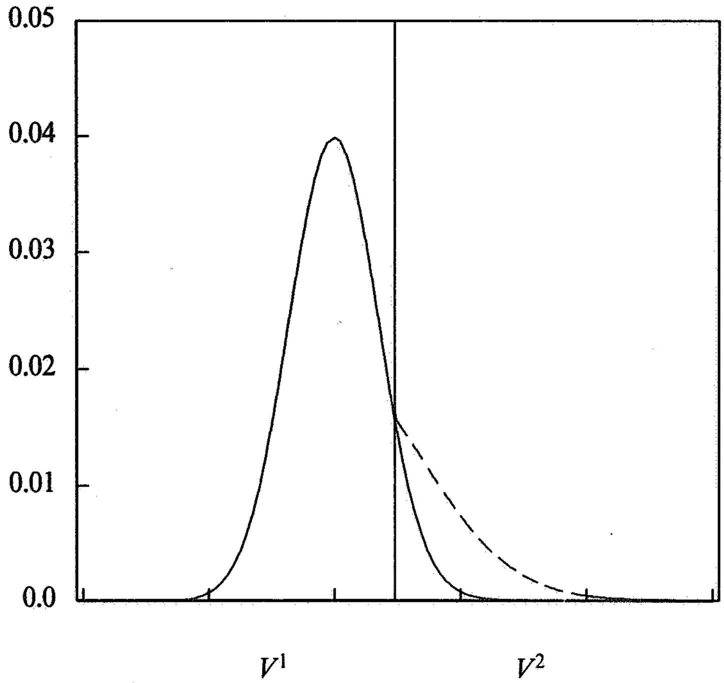


FIGURE C.1

and $d\hat{\rho}_i > 0$ if $\hat{\Phi}_{i,2i+2} + 1 < 0$. The latter condition may well be satisfied: increasing $\hat{\rho}_i$ in this example leads to a reduced speed (the speed distribution in figure C.1 shifts to the left) and to a reduced variance of the speed distribution (according to (2.7)). Both effects lead to a decreased mass in V^2. Increasing $\hat{\rho}_i$ therefore reduces the error in $\hat{\lambda}_i^2$. On the other hand, increasing $\hat{\rho}_i$ will in general increase $\hat{\lambda}_i$, the total intensity, even further. This follows from the definition of intensity (2.4). Therefore there is a tendency to decrease $\hat{\rho}_i$. It will depend on the relative weight of the effects what the resultant effect on $\hat{\rho}_i$ will be. In the numerical results of the filter tests we have found that the described effects in some cases do lead to an increase of the total intensity estimate $\hat{\lambda}_i$ even if it is too large already. The main problem seems to be that the passing distribution does not in all circumstances correspond to the actual one. This also explains why increasing the system noise does not help: the error is in the observation equation, putting more weight on the measurements and thereby emphasizing this relation is likely to increase the error. A similar analysis as the one above shows that a correct distribution would make the occurrence of the described effect unlikely. It is however difficult to model this distribution more precise than we already did by allowing a dependency on the traffic density of the variance: (2.7). It is not clear on what other traffic parameters than speed and density it should depend and how. We therefore decided to change the filter in such a way that the density estimates react as if *no* speed measurements were available (M=1). The justification for this is that the density can be estimated quite accurately from the countdata alone, the speed measurements are not of great importance here. They are of importance in estimating the section speed and will continue to be used in the speed equation of the filter. The elimination of the effect of speed information on the density estimates is achieved by summation of terms in the Φ-matrix as described in section 5.