Title

Integration Of Probe Vehicle And Induction Loop Data: Estimation Of Travel Times And Automatic Incident Detection

Authors

Westerman, Marcel
Litjens, Remco
Linnartz, Jean-paul

Publication Date

1996

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Integration of Probe Vehicle and Induction Loop Data – Estimation of Travel Times and Automatic Incident Detection

Marcel Westerman
Remco Litjens
Jean-Paul Linnartz

California PATH Research Report
UCB-ITS-PRR-96-13

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department of Transportation, Federal Highway Administration.

June 1996

ISSN 1055-1425

Integration of Probe Vehicle and Induction Loop Data

-Estimation of Travel Times and Automatic Incident Detection

Marcel Westerman
Delft University of Technology, Dept. of Civil Engineering
Currently working at: Netherlands Organization for Applied Scientific Research (TNO)
P.O. Box 6041, 2600 JA Delft, Tbe Netherlands, e-mail: mwe@inro.tno.nl

Remco Litjens
University of California at Berkeley, Dept. of Electrical Engineering and Computer Sciences
e-mail: remco@crescent.eecs.berkeley.edu

Jean-Paul Linnartz
University of California at Berkeley, Dept. of Electrical Engineering and Computer Sciences
Currently working at: Philips Research Labaratories
Prof.Holstlaan 4, 5656 AA Eindhoven, e-mail: linnartz@natlab.research.philips.com

Integration of Probe Vehicle and Induction Loop Data
- Estimation of Travel Times and Automatic Incident Detection -

Marcel Westerman, Remco Litjens, Jean-Paul Linnartz

1996

Abstract - Integration of road traffic data collected by means of non-infrastructure based probe vehicles and infrastructure based induction loop detectors allows enhancement of the quality of real-time traffic information that can be obtained for ATMIS. Appropriate fusion of both data sources requires knowledge about deficiencies in present data collection and processing techniques based on infrastructure based traffic detectors and fundamental insight in techniques for processing probe vehicle data. In this research report methods for estimating real-time travel times and performing automatic incident detection for ATMIS based on induction loop or probe vehicle data alone are developed. By properly incorporating additional traffic data from the other source the performance of the developed methods is shown to improve.

The research described in this PATH report has yielded theoretical and practical,insight into the issue of how data delivered from non-infrastructure probe vehicles concerning spatial fluctuations in the traffic process can be included to enhance the reliability and the accuracy of travel time estimates and AID performance based on data measured by traditional infrastructure based (induction loop) detectors concerning temporal fluctuations in the traffic process.

From the results obtained in this PATH research report we learn that fusing probe vehicle data and induction loop data can considerably enhance the quality of derived traffic information. By implementing the data processing techniques described in this PATH research report adequate real-time road traffic information for ATMIS can be attained. Hence, constructing a combined induction loop and probe vehicle monitoring system for ATMIS is indisputably recommended.

Keywords: Traffic Surveillance, Advanced Traffic Management Systems, Advanced Traveler Information Systems, Data Collection, Data Processing, Probe Vehicles

Contents

1. Introduction

The objective of the research presented in this report is to investigate methodologies for fusing and further processing real-time road traffic data collected by probe vehicles and induction loop detectors in order to obtain relevant traffic information that is essential for effective deployment of ATMIS.

Integration of data from probe vehicles and induction loops is useful provided that their collective performance exceeds the performance of both data collection techniques on their own. However, few sound theories exist for processing elementairy traffic data collected by probe vehicles or induction loops alone into useful ATMIS information. Hence, before adressing the isssue of fusing probe vehicle and induction loop data, we will study methods for processing induction loop and probe vehicle data alone.

In the research that is described in this report we have limited ourselves to obtaining information concerning real-time travel times and to performing automatic incident detection. Both issues are considered to be of major importance for effective deployment of ATMIS.

The structure of this report is as follows.

In chapter 2 we will describe a novel algorithm for estimating real-time travel times on freeways using data measured by induction loop detectors. Since ATMIS is concerned, which relates to an extensive road network, the mutual distance between consecutive induction loop detectors is large, typically in the order of 5 to 10 kilometers. The developed algorithm for estimating real-time travel times from data measured by widely spaced induction loop detectors has been applied to factual traffic data, and has been calibrated using factual travel times measured by means of the moving observer method. Although, in general, the obtained results have been found to be acceptable for ATMIS, the developed algorithm remains susceptible to significant stochastic measuring errors that are characteristic of induction loops.

In chapter 3 processing of probe vehicle data is studied. We have concisely addressed the issue of processing probe vehicle data into information about density and flow, and have found that this type of data is principally unsuitable for this purpose. Subsequently, we have developed a comprehensive algorithm for estimating real-time travel times based on probe vehicle data. For this purpose, a multi-layered speed model is introduced. This speed model also allows for detecting significantly deviating traffic conditions; this function has been defined as automatic congestion detection. By assuming realistic parameter values for the various parameters in this algorithm, we have computed the accuracy and the reliability of the algorithm and the estimate.

Chapter 4 is dedicated to the issue of fusing probe vehicle and induction loop data. For this, we have developed distinct methodologies and algorithms.
In section 4.1 probe vehicle data is integrated into the algorithm for estimating real-time travel times from induction loop data described in chapter 2. Simulation results have demonstrated that

integrating probe vehicle data considerably enhances the accuracy and reliability of the travel time estimations. Above all, the estimates are resistant to serious measuring errors.
In section 4.2 induction loop data is integrated into the algorithm for estimating real-time travel times from probe vehicle data described in chapter 3. Computations with realistic basic assumptions have demonstrated that the travel times produced by the combined algorithm are more realistic than those based on probe vehicle data alone.
In the final analysis presented in section 4.3 of this chapter, the procedure for performing automatic congestion detection (i.e. detecting significantly deviating traffic conditions) based on probe vehicle data is enhanced to include induction loop data as well. This has resulted in an extensive algorithm for performing automatic incident detection. All modules of this combined probe vehicle and induction loop AID system are described in detail in this chapter.

Chapter 5 contains conclusions and formulates recommendations. Chapter 6 closes with references used throughout the report.

2. Infrastructure Based Traffic Detectors

This chapter deals with processing data measured by infrastructure based induction loop detectors. As the costs of installing induction loop detectors are high they should be utilized economically on a network-wide scale, implying large mutual spacings (typically 5 to 10 kilometers) of freeways. Furthermore, when processing the data collected by these induction loop detectors a significant (systematic as well as stochastic) error has to be taken into consideration (Bolte, 199 1).

Four macroscopic traffic flow parameters can be distinguished that are of particular interest for ATMIS, viz. flow q, density k, speed v and the prevailing traffic regime (i.e. free flow or congested).

Traffic volume q and (time-mean) spot speed are directly measured by induction loop detectors, as well as occupancy (i.e. the total time an induction loop has been occupied by vehicles in relation to the time of the total measurement). From measured occupancy the density k on a road link can be computed (estimated), although such computations can only accurately be performed in the case of homogenous traffic and hence not under congestion conditions (the only proper method for determining road link density under congestion conditions would be by aerial photography). The latter demonstrates the major shortcoming of induction loops (and, as a matter of fact, of all infrastructure based traffic detectors, possibly except for video camera's when they are not employed as substitution of regular fixed traffic detectors) for ATMIS purposes.

Infrastructure based traffic detectors measure traffic data at one cross-cut, while ATMIS require information about the actual traffic data over an entire road link, or at least over a lengthy , longitudinal section of a road link. In other words, infrastructure based traffic detectors measure temporal fluctuations in the traffic process at one specific location, while ATMIS require information about spatial as well as temporal fluctuations in the traffic process on a complete road network. Regrettably, extraction of longitudinal section data from cross-section data requires complex computations.

Detection of disturbances in the traffic flow on a road link (selection of the prevailing traffic regime or Automatic Incident Detection, AID) from measured infrastructure based detector data has been the subject of much research and several algorithms have been developed. None of the existing algorithms, however, is capable of performing this task with both sufficient accuracy (they either suffer from a high false alarm rate...) and reliability (...or they suffer from a low detection rate) from traffic data collected by infrastructure based traffic detectors with large mutual distances (of circa 5 to 10 kilometers), as is required for network-wide ATMIS traffic monitoring. Seeing that detection of incidents using infrastructure based traffic detectors for ATMIS will not be adequately possible, this subject is deferred to chapter 4, where infrastructure based traffic detectors are combined with non-infrastructure based detectors.

For the above-mentioned reason, in this chapter a new method for determining real-time travel

times on freeways from induction loop data is developed. The method is based on simple counting in combination with correlation techniques. Section 2.1 discusses the traffic flow model that constitutes the foundation for the algorithm developed in section 2.2. This algorithm is implemented into a computer model called COMETT (Computer Model for Estimation of Travel Times using induction loop detectors), which is applied to factual (measured) road traffic data. The results are verified in section 2.3 and calibrated in section 2.4. Section 2.5 closes with a discussion concerning the issue of collecting real-time traffic data by means of infrastructure based traffic detectors.

In this section we will develop and test a new algorithm for determining real-time travel times on freeways using data measured by induction loop detectors with large mutual spacings, typically of 5 to 10 kilometers.

2.1 Traffic Flow Model

Consider an arbitrary stretch of a freeway[1] as sketched in figure 1. This stretch or road link is bounded by two locations, labelled x_1 and x_2, with a mutual distance ranging from typically 5 to 10 kilometers. Each of these two locations represents a measuring site equipped with a pair of induction loop detectors installed on each lane, such that both the traffic flow and the speed of each vehicle passing one of these two locations is measured on a per lane basis. Furthermore, we deploy a discretization interval of δt minutes, implying that the installed equipment provides the number of vehicles that have passed the concerned location during a particular time period of δt minutes, while the speed of these vehicles is provided as an average over this same interval δt. A common discretization interval for ATMIS is 1 to 5 minutes. Finally, the detection equipment at both measuring sites is assumed to employ synchronized time stamps.

We define the travel time at a point in time t, $\tau(t)$, on this stretch of road link as the average time that all vehicles that are detected passing the location x_2 at point in time t, have experienced while traveling from x, to x_2.

The traffic flow on this road link can be modeled by a discrete flow (i.e. a discretized flow) which value fluctuates around a certain mean value. By analogy with (Dailey, 1993) we assume this flow to be a function of several contributions:

---

[1] Analogous to (Papageorgiou, 1991) throughout this dissertation we will use the term freeway to characterize a road traffic system with the following characteristics:

(a) a long road,
(b) unilateral flow,
(c) more than one lane,
(d) no traffic lights along the mainstream, and
(e) interactions with adjacent roads limited to particular on-ramps and off-ramps.

This means that highways (USA), motorways (UK), Autobahnen (Germany), autostrada (Italy), autoroutes (France) and autosnelwegen (NL) belong to the class of freeways.

the concentration, $\alpha(x,t)$, from an upstream point (the position of an induction loop detector located at x_1) that has propagated to this location (the position of an induction loop detector located at x_2) at time t. The concentration $\alpha(x,t)$ is defined by the number of vehicles that have passed location x during time period t.
change in concentration from addition of removal of vehicles due to on and off ramps, $S(x,t)$
a dispersion factor, $\beta(t)$, that represents fluctuations in the traffic flow relative to the mean concentration
noise due to loop failure, $\mu(t)$

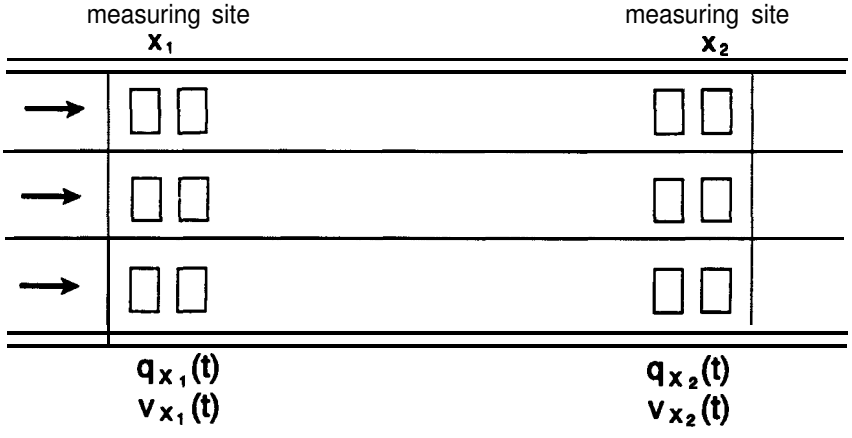

Figure 2.1 Road link bounded by two measuring sites x_1 and x_2

Accordingly, the traffic flow at the downstream measuring site x_2 of this road link can be written

$$\alpha(x_2, t) = \beta(t)\,\alpha(x_1, t-\tau(t)) + \sum_{x_1}^{x_2} S(x,t)\,dx + \mu(t)$$

(Adapted from (Dailey, 1993))

where $\tau(t)$ is the mean time for the traffic to propagate from x_1 to x_2, i.e. the travel time at point in time t. This travel time is computed every discretization interval, i.e. every δt minutes.

2.2 Algorithm

As stated in the previous section, the algorithm for determining real-time travel times from data measured by induction loop detectors is based on simple counting by constituting and interrelating two cumulative curves that represent the measured vehicle passings. This part is

called the basic principle. This basic principle will be shown to be rather susceptible to measuring errors. Moreover, it has been argued that induction loop detectors often fail to detect passing vehicles while on and off ramps between the two successive measuring sites will cause a deviation in the volumes measured at x, and x_2. Hence, a control mechanism is required to repress the effect of these deviations on the estimated travel times. This control mechanism produces so-called re-calibrations to compensate for the latter effect. Both the basic principle and the re-calibrations will be explained below.

Basic principle $\alpha(x_1,t)$ and $\alpha(x_2,t)$

We take into consideration a road link as sketched before. The induction loop detectors located at x_1 and x_2 measure a vehicle passing a detector and record the point in time. Disregarding the disturbances due to on and off ramps ($S(x,t)$) and loop failures ($\mu(t)$) and not going into the fluctuations ($\beta(t)$) as yet, leaves the relation

$$\alpha(x_2,t) = \alpha(x_1,t-\tau(t))$$

which directly relates the concentrations at both measuring sites by means of the (mean) travel time $\tau(t)$.

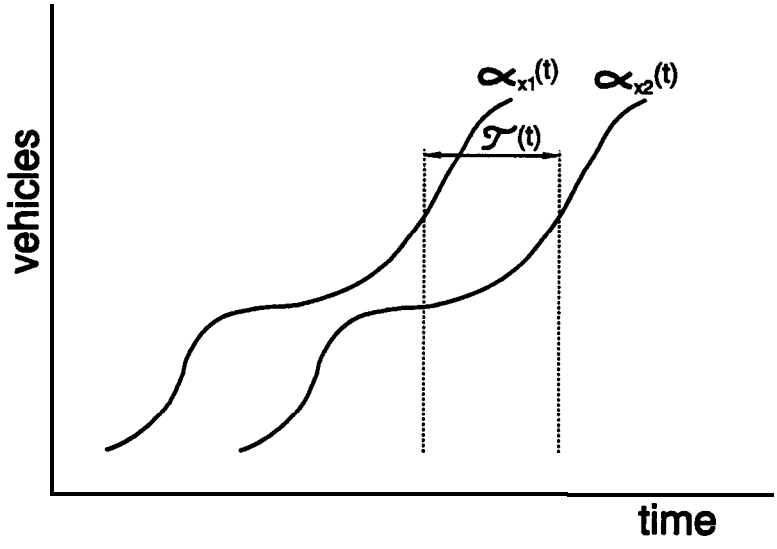

Figure 2.2 The time-dependent cumulative distributions $\alpha_{x1}(t)$ and a,(t)

By using this relation for determining the travel time $\tau(t)$ at the point in time t the spreading in the travel time will be much smaller than the spreading created by means of traditional methods for estimating the travel time, using delay-functions to relate the speed, flow or density and the capacity of a part of a freeway (Wardrop, 1952; Ford; 1956; Fumess; 1965) and consequently introduce significant deviations (Brandston, 1976, Gunter and Hall, 1986; Hall and Gunter, 1986). Therefore, it should be possible to get more accurate results when using a method based on the described direct relation than is possible with the traditional methods.

The vehicle passings measured at the locations x, and x_2 are added up cumulatively for each measuring site. This results in two cumulative time-dependent distributions $\alpha_{x1}(t)$ and $\alpha_{x2}(t)$ as indicated in figure 2. The determination of the travel times takes place by correlating these distributions, where the time shift needed to correlate both distributions gives the travel time between the measuring sites.

Let us assume that t_{x2} is a point in time at which the travel time of vehicles passing x_2 at t_{x2} is known (see figure 3);

$$\tau\,(t_{x2}) = \text{known.}$$

Seeing that

$$t_{x2} - t_{x1} = \tau(t_{x2})$$

the former point in time t_{x1} at which the *same* vehicles passed in x_1 is known;

$$t_{x1} = \textbf{known.}$$

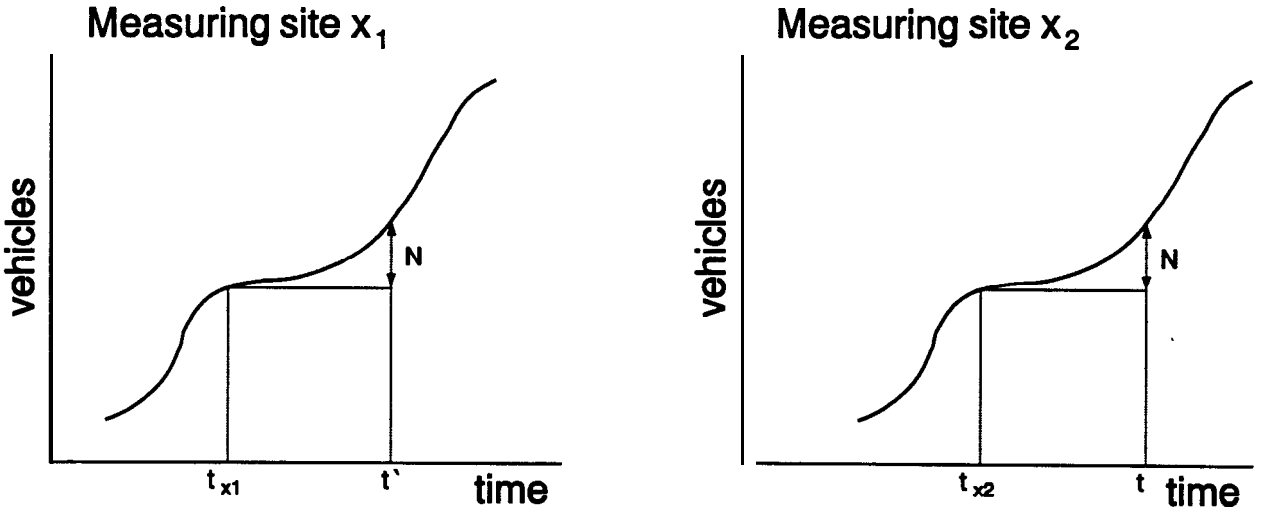

Figure 2.3 The basic principle with cumulative distributions $\alpha_{x1}(t)$ and a,(t)

The travel time $\tau(t)$ at a later point in time *t* can now easily be determined as follows. In the cumulative distribution for counting point x_2, *a,(t), the* number of vehicles *N* that has passed between t_{x2} and *t* in x_2 is given by

$$\alpha_{x2}(t) - \alpha_{x2}(t_{x2}) = \text{N}$$

Then in the curve of measuring site curve x,, $\alpha_{x1}(t)$, the point in time *t'* at which since t_{x1} also N

vehicles have passed is given by

$$\alpha_{x1}(t') = \alpha_{x1}(t_{x1}) + N$$

The travel time for vehicles passing x_2 at t then can be calculated by means of

$$z(t) = t - t'.$$

When δt is the aggregation period (e.g. 1 minute) t can essentially walk through. the whole registration-interval with steps of δt. For each aggregation period δt a corresponding characteristic travel time $\tau(\delta t)$ will be found.

Noise due to loop failure $\mu(t)$ and on and off ramps $S(x,t)$

The basic principle described above is capable of determining travel times faultlessly. However the assumption of disregarding disturbances due to on and off ramps ($S(x,t)$) and loop failures ($\mu(t)$) is solely theoretical. In practice, every induction loop detector has a certain fault rate and in between successive measuring sites several on and off ramps may be located. In these cases after some time the direct and accurate relation between the concentrations $\alpha(x,t)$ at the successive sites becomes increasingly inaccurate and slowly disappears. This means that the equation used for the basic principle does not hold interminably and should read

$$\alpha(x_2, t) = \alpha(x_1, t\text{-}z(t)) + \sum_{x_1}^{x_2} S(x, t)\, dx + \mu(t)$$

The fault rate of the type of induction loop detectors that are currently used in the Netherlands (ES06 detectors; see for instance (Westerman, 1995B)) was studied by (Donk and Versluis, 1985). In this empirical study they concluded that the deviation of the this type of detectors does not exceed 2% and generally amounts about 1% ("missings"). The consequences of the induction loops located at x_1 and at x_2 frequently missing vehicles can be analyzed by assuming that the chance of missing a vehicle in the case of individual detections is equal to p. We also assume that these missings are mutually independent and that p is independent of the flow, road type, weather, and suchlike. Further, we restrict ourselves to total volume, so no division into vehicle categories, which would essentially be possible (Westerman, Immers and Hamerslag, 1992; Western-ran, 1993). After N vehicle passings the mean error is pN (missings). With a flow $q(t)$ it also takes

$$\varepsilon_t = pN / q(t)$$

of time longer before the correct number (= N) of vehicles has passed the induction loop located at measuring site x,, assuming the worst case that the induction loop located at x_1 counts correctly and the one located at x_2 does not. The error in the travel time also increases with N.

Seeing that $N = T\ q(t)$

the error in the travel time is

$$p(t) = \varepsilon_t = pT,$$

where T is the period from the point in time at which the starting points in time t_{x2} and t_{x1} were determined to the point in time at which the current travel time determination takes place (t).

When counting further using p = 2% and assuming the flow $q(t)$ being constant (also if $q(t)$ varies this valuation is in this same order), then at a travel time τ of for example 6 minutes (speed 100 km/hour and distance between the measuring sites 10 kilometers), the relative inaccuracy due to loop failure in τ after T = 30 minutes has increased up to 0,6 / 6 = 10%. This means that in the worst case after circa half an hour the travel times determined in this way become considerably inaccurate. This depends on the distance between the measuring sites, the mean travel speed and of course highly on p. When the measuring error is systematic or when more detectors are used the seriousness of this adverse effect curtails.

The divergences in the measured concentrations due to traffic leaving or entering the freeway through on and off ramps, can be many times larger than the inaccuracies due to loop failures and so can lead to larger disturbances in the direct relation between the concentrations at these both measuring sites. Assume a two lane freeway with an one lane on ramp in between two successive induction loop detectors located at position x_1 upstream and location x_2 downstream. In the case of severe congestion the ratio between traffic on the main road and traffic on the on ramp will amount approximately three to one, in other cases this ratio will be higher (Schuurman and Vermijs, 1993). Hence, in case of congestion (worst case) the error in the concentration at position x_2 could increase to

$$\alpha(x_2, t)^{-1} \sum_{x_1}^{x_2} S(x, t)\, dx \approx 25\%.$$

Calculating with this percentage as done for the noise due to loop failures leads to a considerably short time interval in which the determined travel times will be accurate and reliable. However, in the case of congestion the travel time $\tau(t)$ increases, so the relative error $(= p * T / \tau(t))$ decreases. The method developed so far is capable of determining travel times in the case of free flow and congested traffic. Due to loop failures and, in exceptional situations also due to on and off ramps, the accuracy and the reliability of the determined travel times slowly decreases. To guarantee a continuing reliability and accuracy of the determined travel times it is necessary to determine the starting points in time t_{x1} and t_{x2} in the cumulative distribution curves $\alpha_{x1}(t)$ and $a_{,,}(t)$ over and over in time.

Fluctuations $\beta(t)$

We showed that, especially when the method developed so far has been utilized for some time,

the error due to loop failure is too large to produce a sufficiently accurate travel time. To correct this error, as well as to increase the reliability of the determination of the travel times, the starting points in time t_{x1} and t_{x2} have to be fixed regularly so that $t_{x2} - t_{x1}$ equals the travel time at the point in time t_{x2}. These repeated *'re-calibrations'* take place by correlating a dispersion factor that represents fluctuations in the traffic flow relative to the mean concentration ($\beta(t)$). Fluctuations in the traffic flow can occur on several levels of scale. We make a distinction between:

Meta-fluctuations
Fluctuations in the traffic flow per day of the week or per month of the year.
Macro-fluctuations
Slow trends per day, like the run up to peak hour.
Meso-fluctuations
Characteristic fluctuations over periods from half a minute to various minutes which are preserved over several kilometers (Weits, 1990).
Micro-fluctuations
Momentary fluctuations e.g. due to 'strange' manoeuvres of individual car drivers, which are uncorrelated and not characteristic for the traffic flow.

It has been shown in (Weits, 1990) that when fluctuations at meso-level occur, they are preserved during several kilometers. Accordingly, these can be used to relate the point in time (t'), at which a certain characteristic fluctuation was measured by the induction loop detector located at position x_1, and the point in time (t), at which that same fluctuation arrived at measuring site x_2. The theoretical foundation of this assumption has first been formulated by (Sakasita and May, 1975) in which it was assumed that "***...the arrival patterns of vehicles on a freeway lane remains essentially unchanged under non-incident conditions.".***

The correlation of these characteristic fluctuations gives, very accurately, the time, for these fluctuations to propagate from x_1 to x_2. As this time equals the travel time, the error in the determination of the travel time introduced by the basic principle due to loop failures can be corrected. Correlating the points in time t' and t is done by minimizing the surface between parts of the curves of the cumulative distributions $\alpha_{x1}(t)$ and $\alpha_{x2}(t)$, located at a certain time interval around t' and t, by means of the least squares method. The horizontal shift (time difference) needed for this correlation gives the mean travel time $\tau(t)$ over the stretch of freeway bounded by positions x_1 and x_2 of the vehicles that passed the induction loop detector located at measuring site x_2 at the point in time t. A vertical shift of the part of the curve of measuring site x_2, compared to the part of the curve of measuring site x_1, is performed to correct for loop failures.

This process is illustrated in figure 4 and can be described by the following formulas. The travel time at the point in time t found by correlating fluctuations in the cumulative curves $\alpha_{x1}(t)$ and $\alpha_{x2}(t)$ is given by

$$\tau_{correlation}(t) = t - t''$$

in which t'' is the calculated point in time in the cumulative curve $\alpha_{x1}(t)$ and is a correction of the point in time t' which was calculated by means of the basic principle

$$t'' = t' + \delta t$$

The correction factor δt is given by minimizing the surface between parts of both curves

$$\min \quad Q_{\delta t}, \qquad -d_{x1} + d_{x2} \le \delta t \le d_{x1} - d_{x2}$$

This surface is calculated by application of the least squares method

$$Q_{\delta t} = \sum_{i=-d_{x2}}^{d_{x2}} \left(\alpha_{x1}(t' + \delta t + i) - \alpha_{x2}(t + i) \right)^2$$

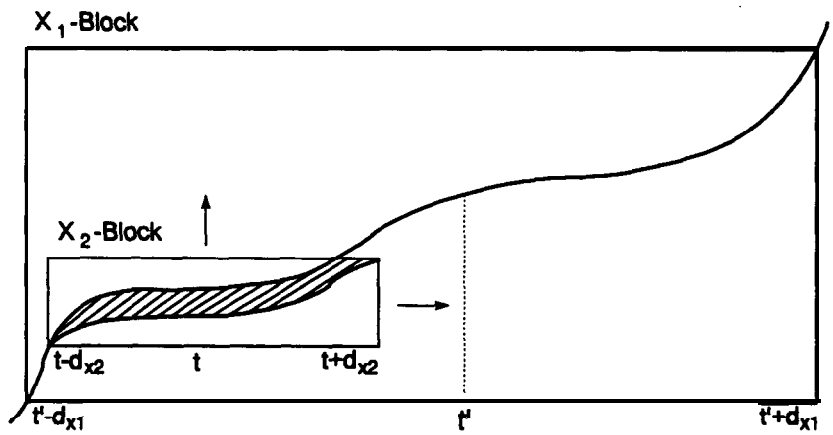

Figure 2.4 Least squares method applied to correlate parts of the cumulative curves $\alpha_{x1}(t)$ and a,(t)

When this surface is sufficiently small, compared to the size of the applied x,-block and x_2-block

$$Q_{\delta t} \le T d_{x2}$$

with T the correlation threshold which has to be determined empirically.
It may occur that coincidental resemblances in the fluctuations in the traffic flow at both measuring sites arise, which will cause the surface between the parts of the two cumulative

curves to be sufficiently small (i.e. smaller than $T\,d_{x2}$). To prevent the re-calibration process from being unjustly triggered by such coincidental resemblances, an additional check-criterion is included which has to be met in order to allow a re-calibration to be performed. This check-criterion examines whether the travel time found by to correlating characteristic fluctuations is realistic and corresponds to the speeds measured at both measuring sites. In other words, the check-criterion computes the difference between $\tau_{speed}(t) = \frac{1}{2} L_{x1,x2} / v_{x1}(t) + \frac{1}{2} L_{x1,x2} / v_{x2}(t)$, with $L_{x1,x2}$ the length of the concerned stretch of road, and $\tau_{cor}(t)$. In the case that this difference is smaller than a certain threshold, and an appropriate correlation was found, a re-calibration is allowed to be performed.

Hence, as last step in this re-calibration process, the starting points in time are redefined

$$t_{x1} = t$$
$$t_{x2} = t''$$

and the total number of measured vehicle passings in both measuring sites are equalized. The whole procedure (basic principle and re-calibration process) can be applied to the traffic as a whole as well as to separate vehicle categories. The latter case will result in several (different) travel times for cars, trucks, heavy trucks, etcetera. Furthermore, this will have the advantage of enhanced utilization of the re-calibration techniques, as even during congestion the composition of the traffic flow with respect to the arrangement of e.g. heavy trucks will practically remain equal. Under conditions of 'smooth' congestion this might enable some additional re-calibrations (Westerman and Immers, 1992).

2.3 Verification

In the previous section a novel algorithm has been developed and presented for determining real time travel times on freeways in the case of free flow and congested traffic. This algorithm has been implemented into a computer model called COMETT (Computer Model for Estimation of Travel Times using induction loop detectors). This section describes the application of COMETT to factual freeway data in order to verify its results. The traffic data used was measured by installing research units in the ES06 induction loop detectors of the Motorway Control and Signalling System (MCSS; the present Motorway Traffic Management system, MTM) (Klijnhout and Langelaar, 1987). For an extensive description of this MTM system we refer to (Westerman, 1995B)).

A sketch of the test-site, a part of the freeway from Utrecht to Amsterdam (A20), is shown in figure 5. The COMETT computer model has been applied to data measured at various days at this measuring site, with various combinations of measuring sites, with large and small mutual distances, with no, one or various on/off ramps in between, etcetera. An extended presentation of these results can be found in (Westerman, 1990). The basic data collected by these research units in the ES06 induction loop detectors concerned individual vehicles passings which. have been aggregated into half minute flows (Van Der Linden and Valk, 1990). The verification concerns

the measurement that started at 16:34 on 16 November 1989 and ended at 06:07, because for this day and time qualitative verification information was available (i.e. no measured travel times, but solely information indicating congestion).

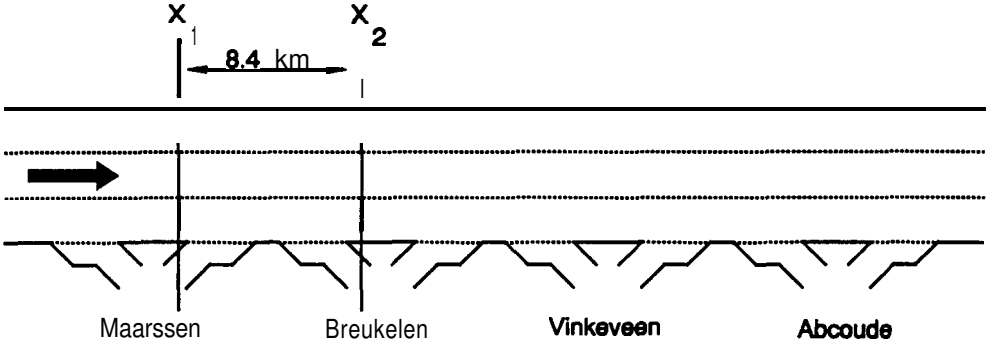

Figure 2.5 Test-site used for verification of the COMETT determined travel times

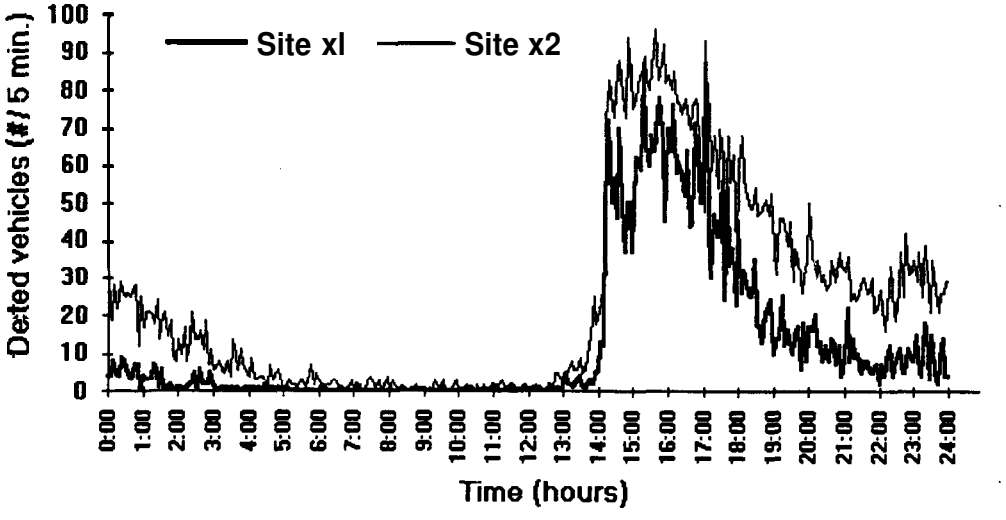

Figure 2.6 Detected vehicles at x_1 and x_2 per 5 minutes

In figure 6 the number of detected vehicles passing the measuring sites x_1 and x_2 has been plotted per time window of 5 minutes. The verification concerns the locations Maarssen (x_1, also denoted *A)* and Breukelen (x_2, also denoted *B*) with a mutual distance of 8.4 kilometers and one on/off ramp in between. Figure 7 shows the cumulative curves for these two locations *A*

(Maarssen, also denoted x_1) and *B* (Breukelen, also denoted x_2), for a time slice starting at 16.59[2] and ending at 18.39 and so including the evening commute.

When testing the COMETT computer model with these measured data, it appeared that re-calibrations could be performed only sporadically when using all detected vehicles. Solely from about *20:00* until about 04:00, multitudinous re-calibration could uncompromisingly be performed. However, when we applied COMETT to the category of heavy vehicles (i.e. all vehicles except regular passenger vehicles) re-calibrations during regular day-times were possible frequently.

The latter can logically be understood as heavy vehicles are less ambulatory, so clusters (groups) remain better preserved and finding correlations is simpler. Therefore, this complication is presumably caused by the way of finding correlations between the cumulative flow curves (by means of the least squared method) and should not be ascribed to principal causes. A more sophisticated approach would be by regarding the two cumulative curves as two (linear) signals, e.g. *h(t)* and *u(t)*, and computing their convolution signal, *y(t)*, by (Ziemer, Tranter and Fannin, 1983)

$$y(2) = \int_{-\infty}^{\infty} u(\lambda)\, h(t-\lambda)\, d\lambda$$

The convolution signal *y(t)* contains more information about the relation between both cumulative curves, but is also more difficult to compute. This latter is considered to be a serious demerit for real-time and network-wide applications. Accordingly, we will use the simple but easy to compute least squares method and the COMETT figures in this section concern heavy vehicles only. The vertical lines in figure 7 illustrate moments in time at which re-calibrations took place.

In figure 8, the COMETT determined travel times at 16 November 1989, from the location Maarssen *(A)* to Breukelen *(B)* from 16.59 to 18.39 are denoted as a function of time. On the right vertical axis in figure 8 the mean speeds in km/h are given, calculated by

$$v(t) = \frac{(x_2 - x)}{\tau(t)}$$

These computed mean speeds provided the first verification criterion, as these- converted speeds (of heavy vehicles) vary from approximately 90 km/h, during regular day time periods, to approximately 20 km/h, during the period of the day that would normally be the peak of the evening rush hour. Ergo, in any case, these speeds, and hence the computed travel times, are plausible.

---

[2] The input time period in COMETT is expressed in time windows of which we selected the %-minute time windows ranging from 50 to 250. This corresponds to 16:34 (begin of measurement) + 50 * ½ = 16:59.

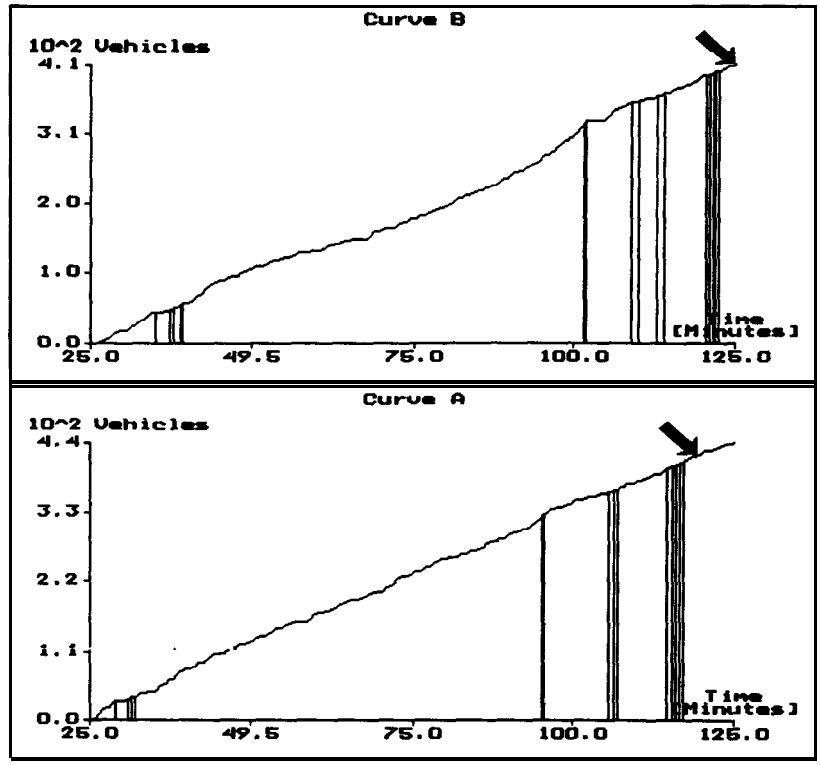

Figure 2.7 Output of COMETT: cumulative flow curves (curve A is Maarssen, curve B is Breukelen)

The second verification criterion was provided by the already mentioned Dutch Motorway , Control and Signalling System (MCSS; the present Motorway Traffic Management system, MTM) with induction loops placed every circa 500 meters at critical intersections. This system has specifically been designed for warning drivers for slowly moving traffic and has (automatically) been activated between the locations Maarssen and Breukelen during the period from 17:20 to 18:15.

The third and last verification criterion resulted from a study commissioned by the Transportation and Traffic Research Division of the Dutch Ministry of, Transport (Rijkswaterstaat), dealing with the same subject of determining travel times and using the signal-analysis package PRIMAL (Van Der Linden and Valk, 1990), accommodated for finding impulse responses between fluctuations in measured traffic flows. This study made use of the same data (or rather, we were kindly allowed to use the PRIMAL traffic data) and showed deviant responses, and hence extremely unrealistic travel times, between the positions Maarssen ($\equiv$ A $\equiv x_1$) and Breukelen ($\equiv B \equiv x_2$) during the period from 17: 15 to 18:05. This is an indication for congested traffic.

From the lasting high travel times in figure 8 it can be derived that, according to the results of COMETT, there was congestion during 17:17 and 18:14. These similarities with the second and third verification criteria, together with the first verification criterion, indicate that the COMETT

determined travel times are probably correct.

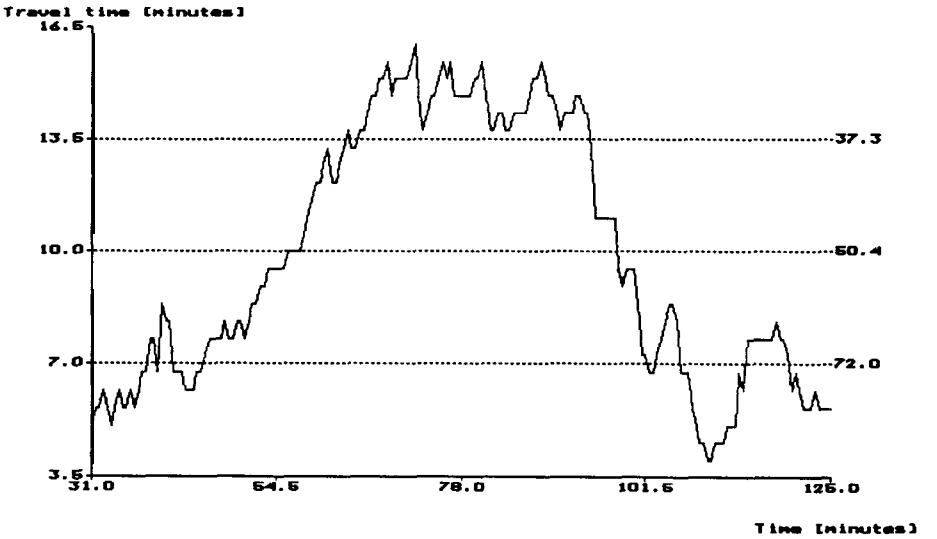

Figure 2.8 Output of COMETT: travel times as function of time

In order to further analyze the quality of the COMETT determined travel times, we were able to use a database with measured traffic data and measured travel times of the DRIVE II project GERDIEN V2044 (General European Road Data Information Exchange Network). This calibration will be discussed in the following section.

2.4 Calibration

The DRIVE II project GERDIEN (General European Road Data Information Exchange Network) is "*... aimed at the implementation of an open road data exchange network in inter-urban motorway networks . ..*" (Van Arem, Van Der Vlist, De Ruiter, Muste and Smulders, 1994). To demonstrate the functioning of this communication network, several sub-models have been implemented and tested using road traffic data (viz. speed, volume and occupancy on a l-minute basis per lane) measured by the traffic monitoring system incorporated in the GERDIEN communication network. In order to evaluate the travel times estimated by one of these sub-models (the Travel time and Congestion Monitoring module, TCM), field observations by means of drivers of vehicles participating in the traffic flow and manually registering their experienced travel times between predefined locations were conducted on parts of the Dutch motorways Al2 and A20. Figure 9 illustrates this measuring site, including the locations of the induction loop detectors which provided the raw traffic data. We refer to (Van Arem, Van Dijk and Rooymans, 1994) for a more detailed overview concerning these field observations.

In order to calibrate the results of the developed COMETT computer model we used both the

induction loop data (volume and speed aggregated over 1 minute) and the travel times registered during the GERDIEN field study. As these data concern a considerably large number of very large files (each file containing several megabytes of data) while the original COMETT computer model has basically been designed as a graphical tool illustrating the working of the developed algorithm, we have redesigned the COMETT computer model using the same algorithm discussed in section 2.2, but employing the dedicated scientific data processing software package IDL (Interactive Data Language (IDL, 1994A; IDL, 1994B)). An example of the output of this redesigned COMETT computer model (on a per minute basis) is pictured in the graphs of figure 10, where it has been applied to the locations $A \equiv$ Moordrecht and $B \equiv$ Zevenhuizen, with a mutual distance of amply 3 kilometers, from 00:00 to 23:59 at April 21 1994.
The seven graphs of the COMETT output, as illustrated in figure 10, are consecutively discussed below.

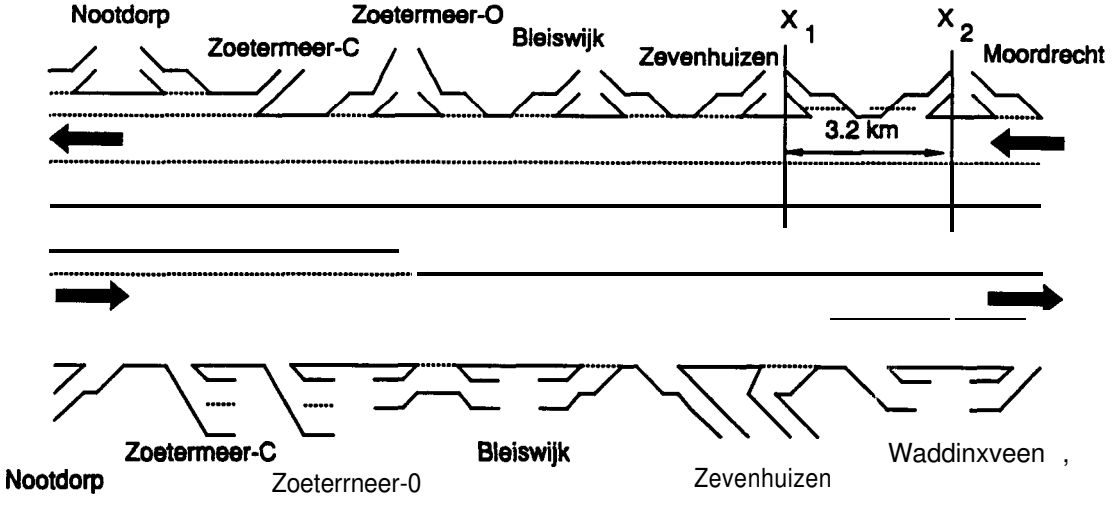

Figure 2.9 Test-site used for calibration of the COMETT determined travel times

Graph 1: Indicative 'speed travel times' τ_{speed}
The first graph shows the indicative travel times between the locations $A \equiv$ Moordrecht and $B \equiv$ Zevenhuizen computed using the speeds (of heavy vehicles) measured at both locations, according to

$$\tau_{AB}(t) = \tfrac{1}{2} L_{AB} / v_A(t) + \tfrac{1}{2} L_{AB} / v_B(t),$$

with L_{AB} the length of the road link between locations A and B.

From this first graph it seems that these indicative travel times adequately designate time periods

during which deviating traffic conditions are encountered, but that these travel times tend to fluctuate considerably. Therefore, the indicative travel times τ_{speed} are only applied qualitatively and not quantitatively (see also section 2.2 and figure 11).

Graph 2: Indicative 'correlation travel rimes' $\tau_{correlation}$

The second graph shows the indicative travel times between the locations $A \equiv$ Moordrecht and $B \equiv$ Zevenhuizen computed by means of correlation between the cumulative flow curves I_A and I_B (of heavy vehicles) at both locations. This graphs shows many indicative travel times that approximate the free-flow travel time (on the road link under consideration the free-flow travel time is about 2 minutes). These travel times can be surmised to represent instants of time at which a proper correlation between both cumulative flow curves was found. This same graph also shows many deviating indicative travel times (which often amount to circa 25 minutes). To all appearances, these deviating indicative travel times represent instants of time at which no appropriate correlation between both cumulative flow curves could be established and hence the best fitting (expressed in minimal surface between both curves; see below) points in time of both cumulative curves were interrelated. These interrelations may be based on coincidental resemblances. That means that under these conditions (i.e. no appropriate correlation) the results of this step do not exemplify a legitimate travel time estimation whatsoever and these indicative travel times should not be used in the remaining steps of the travel time estimation method. We would like to remark that during the time period that normally corresponds to the morning rush hour no proper correlations could be established, which is consistent with the theory described in section 2.2.

Graph 3: Level of correlation or minimized surface

This third graph indicates the 'level of correlation' between both cumulative flow curves, expressed in the least surface (least squared error) between $[\alpha_{x1}(t-i),\alpha_{x1}(t+i)]$ and $[\alpha_{x2}(t-j),\alpha_{x2}(t+j)]$, with $i=2j$. Among other things, this graph shows that almost during the entire day-time the minimal surface remains relatively high. This affirms that the used correlation technique (applying the basic least squares method) is indeed not very optimal (see section '2.3). The irregularity at the end of the time period in the second graph occurs while the end of the cumulative curve is reached (and $[\alpha_{x1}(t-i),\alpha_{x1}(t+i)]$ starts to approach zero; see figure 4).

Graph 4: Re-calibrations $t_{re\text{-}calibration}$

The fourth graph illustrates the time instants at which a re-calibration took place. To determine whether a re-calibration is allowed to be performed we used (see also section 2.2):

the difference between τ_{speed} (first graph) and $\tau_{surface}$ (second graph) should be sufficiently small (i.e not more than 2 minutes. This value of 2 has been determined empirically), and

the correlation between both curves should be sufficiently high (i.e. the relative surface between both curves (third graph) should be sufficiently small (i.e not more than 2½. This value of 2% has been determined empirically)). This relative surface is given by the computed absolute surface between the concerning parts of the cumulative curves in proportion to the size of the concerning part of the cumulative curve of measuring site B, i.e. $[\alpha_{x2}(t-d_{x2}),\alpha_{x2}(t+d_{x2})]$.

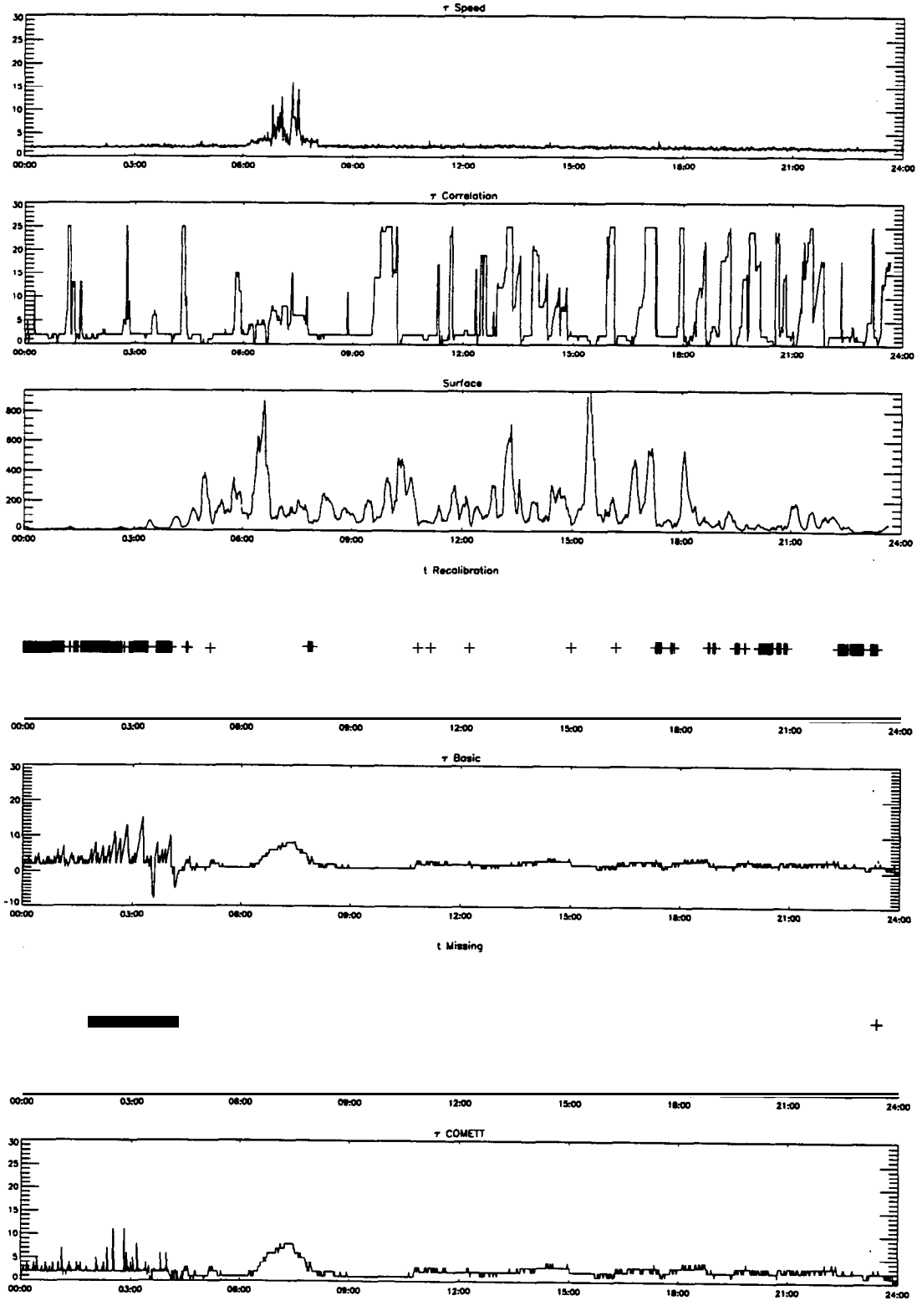

Figure 2.10 Output of redesigned COMETT from Moordrecht to Zevenhuizen

The fourth graph shows that in particular during the morning peak hour no re-calibrations were allowed to take place, which again is consistent with the theory.

Graph 5: Computed 'interim travel times' $\tau_{interim}$
In the fifth graph (which concerns all vehicles) the travel times computed by means of the basic principle in combination with the above established re-calibrations (based on merely heavy vehicles) are depicted. In particular, during time periods with extremely low flow rates (during the night time, until about 05:00) we see rather deviating travel times. This can easily be perceived as, at very low flow rates, a small measuring error has a large relative effect on the travel time determination using the basic principle and using the correlation technique (see also graph 2). This can easily be corrected as follows.

Graph 6: Corrections for missing measurements $t_{missing}$
The sixth graph denotes the periods in time at which no vehicles were detected passing the induction loop detectors. In these cases, the travel time computed in the previous minute is used, in order to exclude the travel time estimates of both the basic principle and the correlations and so to compensate for the effects mentioned before.

Graph 7: COMETT computed travel rimes τ_{COMETT}
Finally, graph number seven illustrates the ultimate travel times, that constitute the output of COMETT.

From the discussed output results of redesigned COMETT it becomes clear that the regular travel time level on the concerned road link amounts to approximately (rounded) 2 minutes, which is in accordance with reality ($\tau_{free\text{-}flow}=D_{AB}/speed\ limit \approx 3.2km/120kmh^{-1} \approx 1.6$ *minutes).* The rather large fluctuations (most within a range of circa 5 minutes, some slightly higher and some even negative) in the determined travel time that happen during the night period, are characterized in that they are brought about by extremely low flow rates and that they have an extremely short duration (typically of 1 minute). To a certain extent, these fluctuations have been compensated for by filtering out time instants with no detected vehicle passings. At the most, such fluctuations will engender a momentarily intensified state of standing by in the Traffic Management Center (TMC) and need under no circumstances induce a false alarm, i.e. do not need to lead to the conclusion that congestion is present on the concerning road link. These specific fluctuations have been smoothed in the graphs used for comparing the COMETT computed travel times with the GERDIEN field observations. The fluctuations during the day time, outside the peak periods, are less severe and more persistent. These are probably correct and may correspond to regular travel time fluctuations and are therefore not smoothed.

To our opinion (Westerman, 1992), the results of the redesigned COMETT computer model in their present form are most suitable and valuable in an ATMIS traffic monitoring system to convey to an operator of a Traffic Management Center, as they adequately signify the actual behaviour of the traffic process. In (De Ruiter, 1995) it is argued that permanent human supervision on automatically obtained travel times before these are further processed or communicated to users will remain recommendable. In any case, automatically obtained travel times should be treated with certain reservations. Under non-recurrent conditions of (practically) stationary traffic, which are mostly due to a temporal and partial obstruction (such as road works

or incidents), the obtained travel times are often meaningless as they only hold for the vehicles that have just passed this obstruction or that have just left the concerning road section. Hence, the determined travel times are only correct for a specific and very small amount of vehicles and only for a specific and very short period of time. For the other vehicles in the traffic flow as well as for ensuing periods of time the travel times will be significantly different (Stoelhorst and De Ruiter, 1995). Within this respect, we would like to remark that for the above-mentioned reason we did not aim at producing smooth and flawless output (which, as a matter of fact, would have been rather easy by adding extra filters or criteria). Conversely, we attempted to demonstrate the exact performance of the developed algorithm and how and in which steps the ultimate results are accomplished. When, by any chance, the output of the described travel time determination algorithm should be directly and automatically processed, the algorithm can be easily adapted for this purpose based on the results given in this chapter.

When we consider the first graph of the COMETT output (comprising seven graphs), representing the indicative travel times determined using only the speed measured at the boundaries *A* and *B* of the road link under consideration, we notice a conspicuous similarity between the ultimate output depicted in the seventh graph. In order to further analyze this similarity, the indicative 'speed travel times', τ_{speed}, and the ultimate COMETT determined travel times, τ_{COMETT}, have been enlarged in figure 11.

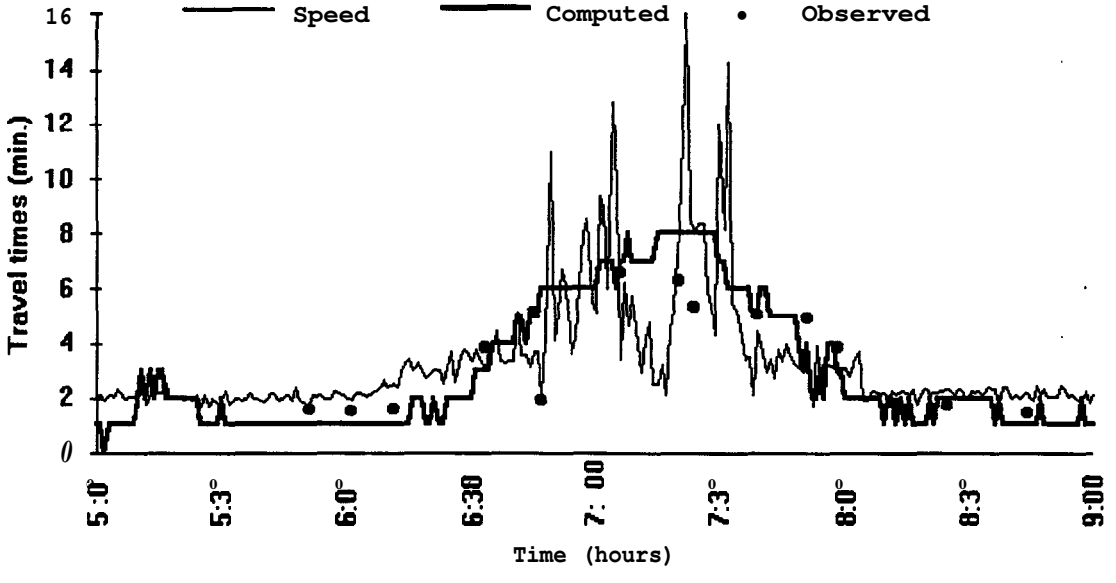

Figure 2. II Evaluation of 'speed travel times' τ_{speed}

Figure 11 concerns the road link from location Zevenhuizen to location Moordrecht during the morning peak hour from 05:00 to 09:00 at April 21, 1994. Furthermore, this figure contains the travel times from the GERDIEN field observations, denoted by the solid circles. In accordance with expectations, the indicative 'speed travel times' τ_{speed} are qualitatively correct but tend to fluctuate extremely. Therefore, they are quantitatively unreliable and inaccurate. Moreover, under

conditions of instable traffic, non-recurrent congestion, very slowly moving or stationary traffic as well as at an increasing length of the concerning road link the inaccuracies and unreliabilities of the 'speed travel times' can be shown to increase further. The latter has already been established in (Brocken and De Haes, 1990).

The redesigned COMETT computer model has been applied to the road links of the GERDIEN pilot area sketched in figure 9. The measuring sites containing the location Zoetermeer-Centrum were excluded from the calibration as the induction loop data measured at this location was erroneous. For all remaining combinations of measuring sites and measuring days where observed travel times from the GERDIEN field study were available, the ultimate COMETT determined travel times were compared to these field observations. In these comparisons, only the momentary and strong fluctuations (more than 2 minutes higher or lower than the previous and the subsequent value) in the COMETT output were smoothed. Such fluctuations appeared to only occur during the night (from around 00:00 until around 04:00) under conditions with very low flow rates.

We will discuss three combinations of measuring sites and measuring days, namely from Moordrecht to Zevenhuizen at April 21, 25 en 27, respectively, in the remainder of this section. These combinations are chosen as on this road link during these days significant (recurrent) congestion occurred.

Outside the regular peak hours the computed travel time depicted in each of the three graphs (in figure 12, 13 and 14) approaches 2 minutes, which is in accordance with the real free-flow travel time ($\tau_{free\text{-}flow}=D_{AB}/speed\ limit \approx 3.2km/120kmh^{-1} \approx 1.6$ minutes). The computed travel times are expressed in units of 1 minute. We realize that a difference of 1 minute in the travel time over a road link of circa 3 kilometers implies a significant difference in the road link mean speed. Accordingly, a practical application of the COMETT computer model would require a refinement of the time scale of the computations (e.g. expressed in units of half a minute). As the iteration-interval of the input variables (flow rate for the basic principle and speed for the re-calibration) is 1 minute, this refinement would introduce a number of extra and more complex computational steps. The results of the analyses with the existent model are amply sufficient for the purpose of calibration, so we have refrained from such a refinement as yet.

During the evening peak hour, from circa 16:00 to circa 19:00, on the concerning road link congestion occurred on neither of the regarding measuring days. This is confirmed by both the COMETT computed travel times and by the GERDIEN field observations.

During the morning peak hour, from circa 07:00 to circa 10:00, congestion occurred on each of the measuring days. This congestion on the distinct days differed in beginning/ending point in time, height and development throughout the time period. The graphs show that nearly all this time the COMETT computed travel times closely correspond with the field observations. However, after a period of several hours without re-calibrations, a deviation between the COMETT computed travel times and the GERDIEN field observations emerges. During congestion of a relatively short duration (e.g. 1.5 hours at April 21) this effect does not occur, during congestion with a slightly longer duration (e.g. 2 hours at April 25) this effect becomes perceptible and after relatively long lasting congestion

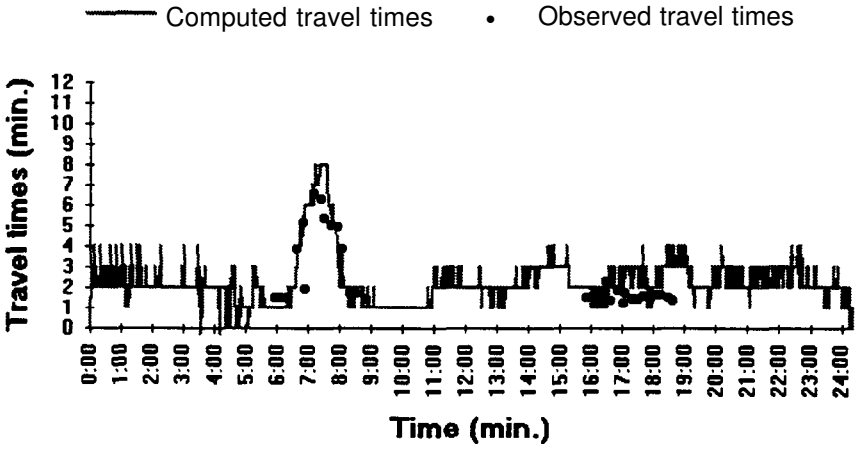

Figure 2.12 Computed & observed travel times from Moordrecht-Zevenhuizen at April 21

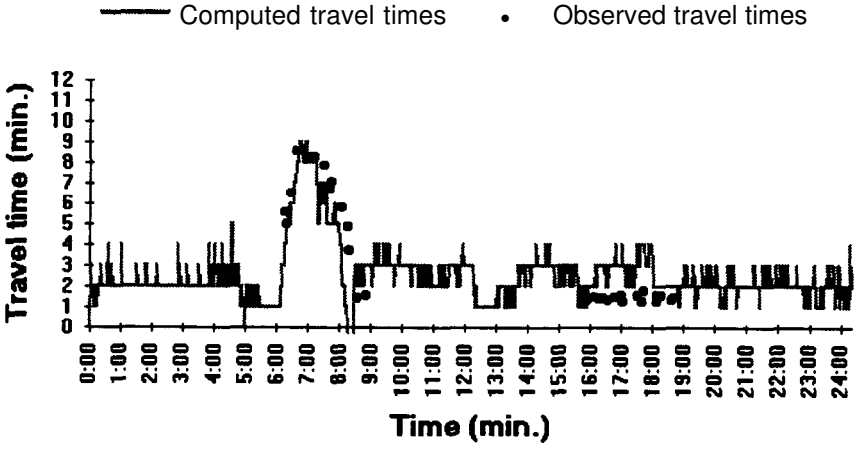

Figure 2.13 Computed & observed travel times from Moordrecht-Zevenhuizen at April 25

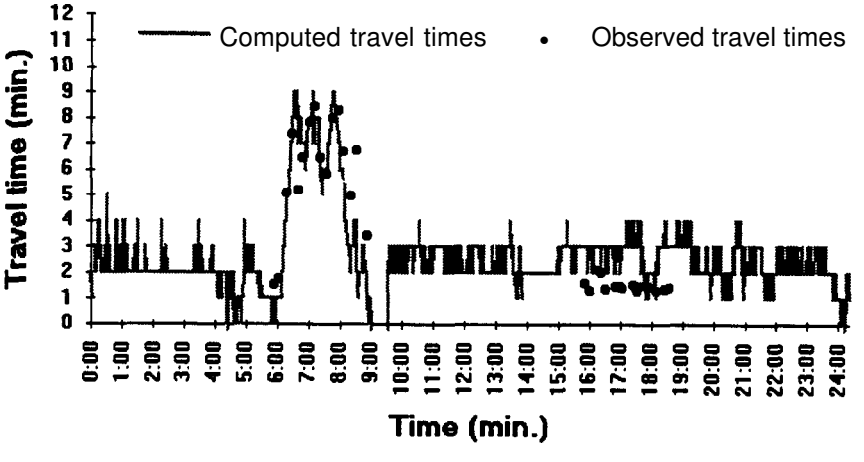

Figure 2.14 Computed & observed travel times from Moordrecht-Zevenhuizen at April 27

(e.g. 2.5 hours at April 27) the effect is somewhat larger. This can easily be explained as follows.

Already in section 2.2, we argued that the basic principle produces perfect travel times provided that the input data (flow rates) are faultless. However, under practical conditions, the induction loop detectors suffer from measuring errors to which the basic principle is rather susceptible. During non-congestion periods these measuring errors are compensated by means of re-calibrations, although particularly during these periods the relative effect of measuring errors is fairly high, and so still fluctuations (of 1 or 2 minutes) in the computed travel times occur. During congestion periods the relative effect of measuring errors is initially limited due to the large numbers of regarded vehicles. As the measuring errors accumulate and no re-calibrations take place during congestion, the effect of measuring errors on the computed travel times during congestion gradually increases and becomes noticeable after a certain period of time.

The reliability of the COMETT determined travel times has been shown to be very high since the presence and absence of each congestion is effectively signified. The accuracy of the COMETT determined travel times has been shown to be reasonably high: outside' congestion periods the regular fluctuations amount to circa 2 minutes, while during congestion the determined travel times are initially extremely accurate while the accuracy gradually decreases (up to around 3 minutes) when congestion persists for longer than circa 1.5 hours. These findings are supported by the theoretical observations. A more sophisticated comparison between the COMETT determined travel times and the field observations is considered to be not useful, since the error in the field observations compared to the actual travel times is rather large (Van Arem, Van Der Vlist, De Ruiter, Muste, Smulders, Dougherty, Cobett and Kirby, 1994).

When we compare the results of the previous investigations to global ATMIS requirements with respect to real-time travel times, it can be concluded that, in general, the developed algorithm is adequate and provides travel times that are sufficiently accurate and reliable for the designated ATMIS purposes. Except for the very short term ATMIS application Automatic Incident Detection, and also except from persistent congestion conditions. The first difficulty is in accordance with earlier expectations and findings (see commencement of this chapter). The latter difficulty might be overcome to some small extent by enhancing the re-calibration technique.

However, both difficulties will remain when utilizing infrastructure based traffic detectors for ATMIS purposes. For this reason, in chapter 4, we will further address determining real-time travel times and performing Automatic Incident Detection for ATMIS purposes, but now utilizing a combination of infrastructure based and non-infrastructure based traffic detectors (Westerman, 1995A).

In (Westerman, 1990) the influence of the distance between the concerning measuring sites has been analyzed. This influence appeared to be rather small for distances up to ample 10 kilometers. For larger distances, the accuracy of the determined travel times would considerably decline, seeing that the meso-level traffic flow characteristics are not preserved over such distances any more. Hence, for an ATMIS traffic monitoring system deploying induction loop detectors the inter detectors spacings should be about 10 kilometers at the most.

On and off ramps in between the measuring sites influence both the basic principle and the re-calibration process. Especially when an on/off ramp is heavily occupied, this will dislocate the developed algorithm. For this reason, in an appropriate ATMIS monitoring system, preferably all on and off ramps, but certainly all heavily occupied ramps, should be equipped with induction loop detectors.

For the calibration process we have used time intervals of 1 minute. For the ultimate ATMIS time discretization intervals of 1 to about 5 minutes may seem appropriate. Time intervals smaller than 1 minute makes ATMIS susceptible to micro-fluctuations that are not characteristic for the traffic flow. Time intervals larger than 5 minutes makes ATMIS insensitive to meso-level fluctuations, that are required for the re-calibration process.

2.5 Discussion

In this chapter we have investigated the utilization of infrastructure based induction loop detectors for obtaining the real-time road traffic information that is required for ATMIS purposes. For ATMIS, their area-oriented characteristic necessitates deployment of large distances between the fixed detectors.

We have assumed that for the class of ATMIS applications four macroscopic traffic flow parameters are of particular interest, viz. flow q, density k, link speed v or travel time τ, and the prevailing traffic regime (i.e. free flow or congested). Of these, the flow q and the density k *can* directly be obtained from induction loops.

With respect to detection of disturbances in the traffic flow (selection of prevailing traffic regime or Automatic Incident Detection for ATMIS) numerous research has been conducted in the past (see chapter 4, for an extensive literature review). The general conclusion of these studies is that appropriate detection of disturbances in the traffic flow, using infrastructure based traffic detectors with large inter-detector spacings (i.e. 5 to 10 kilometers) is virtually unfeasible. For this reason, this subject is deferred to chapter 4, in which infrastructure based traffic detectors are combined with non-infrastructure based traffic detectors in order to perform this task. Accordingly, in the remainder of this chapter, we have concentrated on estimating real-time link travel time from road traffic data measured by infrastructure based traffic detectors. An algorithm for estimating real-time travel times from road traffic data obtained from infrastructure based inductive loop detectors has been developed. This method is based on simple counting in combination with correlation of fluctuations in the traffic flow. The counting takes place by matching cumulative distributions of successive measuring sites, while the correlation (the re-calibration process) is used to guarantee a continuing reliability and accuracy of the travel times.

The developed method has been implemented into a computer model, COMETT (a Computer Model for Estimation of Travel Times using induction loop detectors), and has been applied to factual freeway data. Verification and calibration of the COMETT determined travel times has shown that the developed algorithm is capable of producing reliable and accurate real-time travel times in the case of free flow traffic as well as in the case of congested traffic. The latter is a significant improvement compared to other recent methods for which congestion caused major

complications. Application of COMETT to various configurations of measuring sites, with short, large and very large mutual distances and with and without on/off ramps in between, has revealed that road links with lengths of more than about 10 kilometer should be avoided and that heavily occupied ramps should be equipped with induction loops.

Although minor irregularities in the estimated travel times are expected to be eliminated by refining the model (especially the re-calibration process), the algorithm remains susceptible to the significant stochastic errors in the data measured by the loop detectors. In particular under conditions of small flow rates as well as in the case that congestion persists longer than about two hours, the computed travel times become inaccurate. To a certain extent, these errors are compensated by the re-calibration process, but additional re-calibrations are required for achieving a more persistent performance.

To this end, among other things, we will study monitoring road traffic flows by means of non-infrastructure based traffic detectors in the next chapter. Seeing that the properties of non-infrastructure based traffic detectors entire differ from the properties of infrastructure based traffic detectors, and are in more accordance with the characteristics of ATMIS, we surmise that this novel road traffic monitoring concept will be very appropriate for complying with the specific ATMIS demands.

With respect to the latter issue, we will combine infrastructure based and non-infrastructure based traffic detectors in chapter 4 and demonstrate how non-infrastructure based traffic detectors can be utilized for complying with the difficulties that have been come across in this chapter, i.e. to achieve additional re-calibrations for estimating real-time travel times as well as to support in performing automatic incident detection.

3. Non-Infrastructure Based Traffic Detectors

The purpose of this chapter is to analyze the possibilities for extracting information about the actual status of the traffic flow (macroscopic), from data comprised in the traffic messages received from the individual probe vehicles participating in this traffic flow (microscopic).

In a previous PATH research project, MOU 107, the issue of communicating traffic messages from roving probe vehicle to a traffic center has been studied extensively. The computer model PROMOT (PRObe vehilce concept for Monitoring road Traffic) has been developed and applied in the freeway network of the San Francisco Bay Area. In this chapter, we will build on the results achieved in that previous research and assume that probe vehicle messages have become available in the traffic center.

In (May, 1990) a framework for fundamental characteristics of traffic flow is provided. This framework distinguishes flow, speed and density on both microscopic (individual vehicles) and macroscopic level (groups of vehicles) as depicted in table 3.1.

Traffic characteristics	**Microscopic**	**Macroscopic**
Flow	Time headways	Flow rates
Speed	Individual speeds	Average speeds
Density	Distance headways	Density rates

Table 3.1 Framework for fundamental characteristics of traffic jlow (May, 1990)

From table 3.1 distinct approaches to meet the purpose of this chapter can be derived.

Approach 1 Direct Measurement of Time and Distance Headways
Systems like Intelligent Cruise Control (ICC) (Smith and Starkey, 1991), Adaptive Cruise Control (ACC) (Steinlecher, 1993) and Autonomous Intelligent Cruise Control (AICC) (Alleman, 1993) in which the position of the gas or brake pedal of a vehicle is adjusted according to the distance to the vehicle in front all require almost continuous measuring of the time and distance headways and have already been realized in the European PROMETHEUS program (see for instance (PROMETHEUS, 1992)). A deficiency of using time and distance headways measured in this way for obtaining macroscopic traffic information would be that the behaviour of these vehicles will not be representative for all vehicles in the traffic flow, as the on-board equipment will keep the speed and the time and distance headways stable and well-balanced and that a penetration grade of practically 100% would be required for accurately obtaining macroscopic traffic information.

Approach 2 Vehicle Trajectories

A second possible approach for deducing information about the actual status of the road traffic from received probe vehicle data would be by deploying the commonly used technique of vehicle trajectories for obtaining several microscopic traffic flow parameters, such as time and distance headway, mean speed, travel time and instantaneous location. This approach is already used in the probe vehicles themselves, where each probe vehicle constructs its individual trajectory in order to establish a probe message to be transmitted. For obtaining macroscopic traffic information the trajectory technique has the shortcoming that the reports transmitted by the probe vehicles do preferably not contain an identifier and that the time and distance headway between successive probe vehicles at a penetration grade that is significantly lower than 100%, is not useful.

Approach 3 Fundamental Diagrams

A third approach for estimating traffic parameters would be by utilizing the fundamental diagrams, according to which, at a given traffic flow rate q from the speed/flow-diagram it would be possible to estimate the mean traffic speed v. Estimation of the flow rate q or the density k would require a similar approach. However, in traffic engineering it is generally accepted that these fundamental diagrams should only be applied indicatively and do not accurately represent actual traffic flow behaviour (see e.g. (Banks, 1995)).

Approach 4 Statistical Techniques

A fourth approach is to focus on the macroscopic level of traffic and to analyze how information about these macroscopic traffic characteristics can be extracted from received probe vehicle data. This approach thus departs from the given standard possibilities of the on-board equipment of probe vehicles with respect to the ability to measure only speed, location and elapsed time and requires the development of statistical techniques in order to estimate the macroscopic traffic flow parameters from the probe vehicle samples.

From the above concise survey of possible approaches for deducing general (macroscopic) ATMIS information from individual (microscopic) probe vehicle data, we conclude 'that only statistical techniques are eligible. In section 3.1, we will briefly focus on the issue of statistically estimating traffic density and traffic flow from probe vehicle data only. Although this information can be obtained from infrastrucrue based traffic detectors (see chapter 2), it might contribute to a better insight into the actual traffic conditions if accurately obtained. For estimating the road link mean travel time or the road link mean speed , we will formulate a simple and straightforward method in section 3.2. Since the accuracy and reliability of this method are rather low when only few probe reports are received, we will develop a more advanced procedure in section 3.3, that is based on prior historic knowledge regarding the behaviour of the traffic flow, and will show that this novel estimator considerably outperforms the initial estimator with respect to accuracy and reliability. This chapter closes with conclusions in section 3.4.

3.1 Estimation of Traffic Density and Traffic Flow

3.1.1 Problem Definition

The traffic density k on a road link (here: with a length of 5 to 10 kilometers) during a certain time period (here: with a length of 5 to 15 minutes) is defined as the total number of vehicles that has occupied that road link during that time period. The traffic flow q is defined as the total number of vehicles that has passed a particular crosscut of a road link during a certain time period. These definitions imply that the traffic density depends on the inflow (q) and on the link travel time (τ) during a particular time period. Traditional techniques for obtaining link density and flow depart from infrastructure based traffic detectors, which measure all passing vehicles. For our purpose, only floating car data from probe vehicles is surmised to be available. This introduces the difficulty that the proportion of probe vehicles on a particular road link during a particular time period will presumably be unequal to the network probe vehicle penetration grade (which can be assumed to be proportional to the number of probe equipments or Dynamic Route Guidance systems sold). Hence, even when it is optimistically presumed that the network penetration grade ζ_N of probes is exactly known (typically, ζ_N is in the order of 0.1% to 10%, with low ζ_N prevailing during the early introduction phases), due to fluctuations in the composition of the traffic flow, the actual number of probes present on a road link in a certain short time period will vary to such an extent that it is expected that no reliable conclusions with respect to the actual traffic density can be drawn from received probe vehicle messages. In order to numerically verify or falsify the expectation stated above a simple statistical approach for computing road link density and flow from received probe vehicle messages will be formulated and assessed in the next sub-section.

3.1.2 Problem Assessment

In the following, we will denote with N_p the number of observed probe reports in a ceratin time interval T and with ζ_N the (given) network probe vehicle penetration grade for the entire road network. In order to estimate the density denoted k, *we* assume a Poisson distribution for the total number of vehicles on a road link during a certain time period (Gross and Harris, 1985). Accordingly, also the number of ***probe*** vehicles on this road link during this period is Poisson distributed, with a parameter λT that is equal to the number of observed probe vehicles N_p. So, the actual number of probe vehicles on a road link during a certain time interval, the probe vehicle density denoted k_{probe}, can be estimated by means of a distribution with mean $m=N_p$ and a deviation that is proportional to N_p (for small N_p obtained from a Poisson-nomogram and for large N_p approximated by the square root of N_p). Subsequently, the actual density of all vehicles is estimated by $k=100/\zeta_N\, k_{probe}$, with ζ_N expressed in a percentage between $[0,100]$.

We will illustrate this by means of an example. Assume that during a time period of $T=5$ minutes the density on a road link amounts to $k=200$ $(k=100)$ vehicles. At a network probe vehicle penetration grade of $\zeta_N=1\%$ we would expect a number of $N_p=10$ probe reports $(N_p=5)$, while from a Poisson-nomogram we find that N_p is likely to vary between 7 and 13 (3 and 7).

So, the density estimation in this example is likely to lie between *[140,260]* (*[60,140]*). It should be apparent that the accuracy of these estimations is too low for practical purposes (such as model computations), while this accuracy range comprises both free-flow and congested traffic conditions. This accuracy enhances when the number of involved probe messages (NJ increases. Since a combination of a relatively low value of ζ_N[3] and a high value of N_p implies a long time interval *T,* we conclude that for real-time or short term density estimations probe vehicles are not appropriate, but that for longer term purposes they seem very suitable.

A similar analysis with similar conclusions holds for estimating the actual traffic flow q through probe vehicles.

3.1.3 Conclusions

Information about real-time road traffic density and flow obtained from probe vehicles might be useful for ATMIS purposes, yet it is not essential since these types of data can be obtained from infrastructure based traffic detectors. The results of the analysis show that this information can only be obtained from probe vehicles alone with substantial inaccuracy. Even at high probe vehicle penetration grades, the estimations are shown to remain unreliable. This affirms practical perceptions that real-time density and flow are basic traffic characteristics that can only be determined by measuring (nearly) all vehicles in the traffic flow. From this, we conclude that probe vehicles are not appropriate for collecting real-time road link densities or flows as they only constitute a very small sample with an unknown composition. However, for longer term purposes, such as constructing road maps with annual traffic loads, probe vehicles seem very appropriate (Westerman and Hamerslag, 1993).

3.2 Estimation of Road Link Mean Speed and Travel Time

3.2.1 Problem Definition

Based on the research performed in the previous PATH research project MOU 107, in (Westerman, 1995) the exact contents of the probe messages have been appointed .in view of conveying the best possible information. According to this specification, the probe messages contain a location code, denoted l_i, of crosscuts where significant deviations in the traffic conditions (defined as an increase or decrease in the speed of the probe vehicle of at least $\alpha\%$) were experienced (related to specific pre-defined locations), together with the corresponding time, denoted τ_{ij}, needed to traverse the distance from the former registered location, denoted l_j. In this respect the location codes of pre-defined locations, denoted *ISRN* (International Standard Road Number (ALERT, 1990; CEN, 1994)), that are passed as well as the location code at the

[3] In (Cattling, 1994) the penetration grade of dynamic route guidance systems in the year 2000 is expected to amount to circa 5%.

moment of transmission, denoted l_{now}, are also registered. Furthermore, the (absolute) point in time, denoted t_1, at which the concerning probe vehicle started registering the traffic conditions contained in this particular messages is included in the message. Hence, the message that all probe vehicles transmit every circa T minutes contains the following data:

$$[\ ISRN^k, t_1, l_1, \tau_{12}, l_2, \tau_{23}, l_3, ..., l_{last}, \tau_{last\ now}, l_{now}\ ]$$

Combining[4] the messages from all probes that are present during each time period T, the entire road link (extending from one specific pre-defined location to the next one), is sub-divided into various road segments with constant (probably dissimilar) speeds for each probe vehicle traversing that road link during the time period under consideration. In order to illustrate this procedure, we will discuss an example.

Consider a road link that is encoded by $ISRN^k$ and suppose that during the past time period of T seconds, three probes have utilized (parts of) the regarding road link.

Let the transmitted messages be like this:

Probe 1

$$ISRN^k, t_1(1), l_1(1), \tau_{12}(1), l_2(1), \tau_{23}(1), l_3(1), \tau_{3k+1}(1), ISRN^{k+1}, \tau_{k+1now}(1), l_{now}(1)$$

Probe 2

$$ISRN^{k-1}, l_1(2), t_1(2), \tau_{1k}(2), ISRN^k, \tau_{k2}(2), l_2(2), \tau_{2k+1}(2), ISRN^{k+1}, \tau_{k+1now}(2), l_{now}(2)$$

Probe 3

$$ISRN^k, l_1(3), t_1(3), \tau_{12}(3), l_2(3), \tau_{2now}(3), l_{now}(3)$$

Note that the distances $l_i(1)$, $l_i(3)$ and $l_2(2)$ are measured from the beginning of the road link encoded by $ISRN^k$, while $l_1(2)$ is measured from the beginning of $ISRN^{k-1}$ and $l_{now}(2)$ is measured from the beginning of $ISRN^{k+1}$. Figure 3.1 illustrates the contents of the probe messages and the way they are processed for a specific road link denoted $ISRN^k$. For example, in the past T seconds, probe 1 changed his speed twice on the concerning road link. Its travel time between the points marked $Z,(l)$ and $l_2(1)$ equals $\tau_{12}(1)$ and the (approximately constant) speed on this road link segment is $(l_2(1) - l_1(1)) / \tau_{12}(1)$.

The transmitted locations of significant velocity changes of one probe vehicle divide the road link into disjunct road segments. On each of these road segments each probe vehicle has traveled with a constant speed. In this particular example, see also figure 3.1, the road link encoded by $ISRN^k$ is divided into eight segments, denoted 6, through δ_8. In this way, for each road segment

[4] In this respect, we assume that all probe messages that are successfully received in the TIC contain correct traffic and location data, although we realize that this might be slightly unrealistic (Ran, 1993).

a mean speed and travel time can be estimated and AID type decisions can be made.

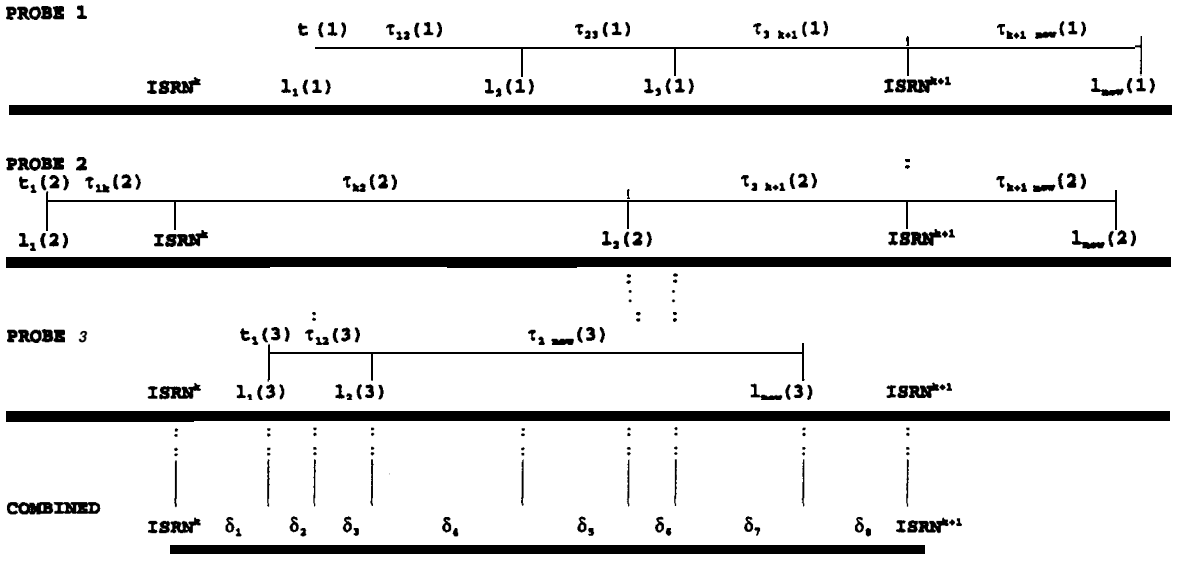

Figure 3.1 Illustration of subdividing a road link in road segments

Recognizing that the speed of each probe vehicle on each of these segments is constant by definition and that the exact length of each segment is known, the (mean) segment speed and the (mean) segment travel time are directly interchangeable. This fact enables us to compute the road link mean speed $\bar{v}$ instead of the road link mean travel time $\bar{\tau}$, which is constructively employed to design an advanced estimator incorporating prior, historic information in section 3.3. Furthermore, since the road link mean speed and the road link mean travel time are direct substitutes, we prefer using the road link mean speed as this enables us to directly combine the data measured by both discerned types of traffic detectors, viz. the speed obtained from the non-infrastructure based detectors (probe vehicles) and the flow obtained from the infrastructure based detectors (inductive loops). Combining both data sources will be done in chapter 4.

Hence, the road segments with piecewise constant speeds for all probe vehicles are taken as separate entities for the final estimation of the road link mean speed $\bar{v}$ and the road link mean travel time $\bar{\tau}$. In order to restrict the number of road segments that constitute a road link, we introduce a minimum length for each road segment, for instance of 100 meters. This discretization can be done in the probe vehicles themselves, or in the traveler center.

Figure 3.2, showing probe vehicle trajectories, indicates the sub-division of a road link into road segments with constant probe vehicle speeds. The horizontal axis denotes the expired time and the vertical axis denotes the traveled distance. We consider a road link bounded by the pre-

defined locations $ISRN^k$ (upstream) and $ISRN^{k+1}$ (downstream) (of, for instance 5 kilometers) and consider a time period from AT (of, for instance 5 minutes). Based on the data comprised in the probe vehicle reports the road link is sub-divided into segments δ_i with variable lengths (with a certain minimum of, for instance, 100 meters), which depend on experienced significant deviations in the traffic conditions (i.e. an alteration of the speed of more than 01%). This has the advantage that the location of a possible disturbance in the traffic flow can be determined accurately. Again, note that the speed of each probe vehicle on each segment δ_i is constant.

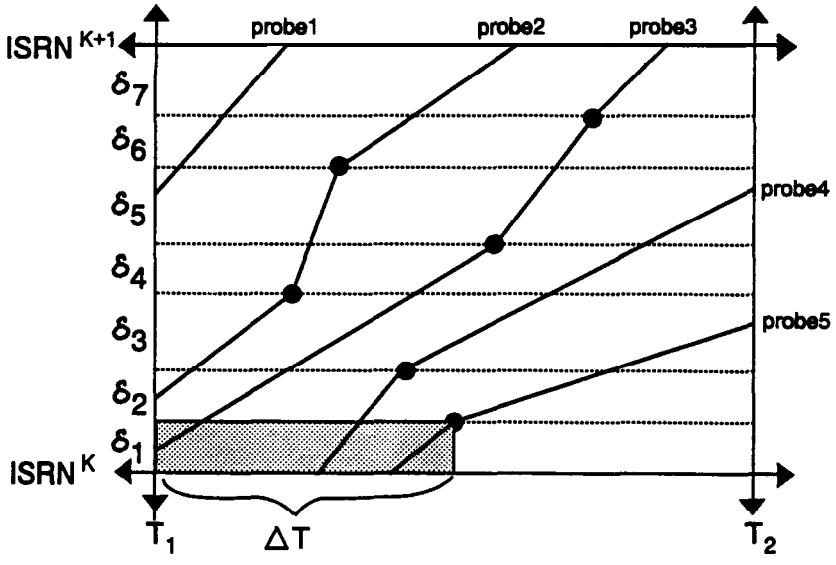

Figure 3.2 Sub-division of road link into segments with constant probe vehicle speeds

In figure 3.3 the contents of road segment δ_1, bounded by $ISRN^k$ and $ISRN^k + \delta_1$, is reproduced on a larger scale. Three probe vehicles were present on this segment during the relevant period from T_1 to T_1 + AT . From the travel times and the locations reported by these probe vehicles, the individual (constant) probe vehicle speed $v_i, i\varepsilon\{3,4,5\}$ on this particular road segment can be computed as illustrated in table 3.2. Note that, due to the way the probe reports are constructed, the travel time on a road segment, τ_{δ_i} , is never equal to zero, and the individual (constant) probe vehicle speed, $v_{\delta|}$, is always defined.

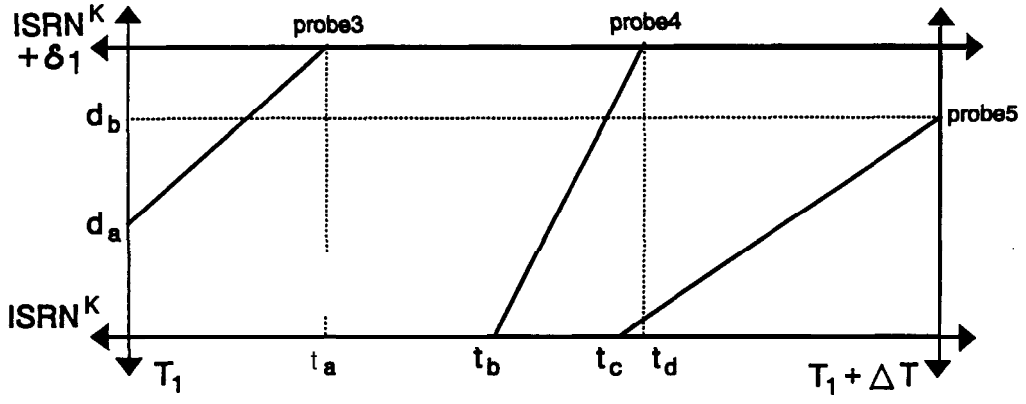

Figure 3.3 Computation of individual (constant) probe vehicle speeds on road segment δ_1

	Travel time on segment ($\tau_{\delta_1}(i)$)	Traveled part of segment ($d_{\delta_1}(i)$)	Individual probe speed $v_{\delta_1}(i) = d_{\delta_1}(i) / \tau_{\delta_1}(i)$
Probe 1	-	-	-
Probe 2	-	-	-
Probe 3	$\tau_{\delta_1}(3) = t_a - T_1$	$d_{\delta_1}(3) = (ISRN^k + \delta_1)$ $-d$	$v_{\delta_1}(3) = \frac{(ISRN^k + \delta_1) - d}{t_a - T_1}$
Probe 4	$\tau_{\delta_1}(4) = t_d - t_b$	$d_{\delta_1}(4) = \delta_1$	$v_{\delta_1}(4) = \frac{\delta_1}{t_d - t_b}$
Probe 5	$\tau_{\delta_1}(5) = (T_1 + \Delta T) - t_c$	$d_{\delta_1}(5) = \delta_1$	$v_{\delta_1}(5) = \frac{\delta_1}{(T_1 + \Delta T) - t_c}$

Table 3.2 Illustration of computing the individual probe vehicle speeds on road segment δ_1

From the probe vehicle speeds shown in table 3.2, the road segment mean speed on road segment δ_1, is estimated. The exact estimation procedures will be explored later on in this section. The estimation procedures divide a road link into small road segments and computes a mean speed for each of these segments. Mutually interrelating these computed mean speeds would require modeling techniques (including dynamic traffic assignment) which are‘ considered to be principally beyond the scope of this research.

3.2.2 Problem Assessment

In traffic engineering two distinct methods for averaging individual vehicle speeds are commonly distinguished and each of these leads towards a different mean, namely the time-mean-speed and the space-mean-speed (see for.instance (May, 1990)). The time-mean-speed at a certain location

during a certain time period is obtained by averaging the individual vehicle speeds measured at that particular location over that selected time period, viz.

$$\bar{v}_{TMS} = \frac{\sum_{i=1}^{n} v_i}{n}$$

The space-mean-speed at a certain point in time over a certain (part of a) road link is obtained by first converting the measured individual vehicle speeds into individual travel times (in the expression below expressed in hours per kilometer), then computing an average travel time by averaging the individual travel times, and finally computing the average speed at that particular point in time over that particular (part of a) road link, viz.

$$\bar{v}_{SMS} = \frac{1}{1/n \sum_{i=1}^{n} 1/v_i}$$

The space-mean-speed reveals spatial fluctuations in the traffic flow over an entire road link and is specifically convenient for localization of disturbances in the traffic process. However, obtaining the space-mean-speed is difficult. One possible way of measuring $\bar{v}_{SMS}$ would be by means of consecutive aerial photographs, which of course is not very realistic on a network-wide scale. The time-mean-speed is measured at one specific location and is therefore more convenient. However, $\bar{v}_{TMS}$ has the disadvantage that it only reveals temporal fluctuations at the concerning crosscut and neglects differences in the traffic conditions at the remaining' part of the road link. Furthermore, fast vehicles tend to predominate the average. The latter is often compensated by harmonically averaging the individual vehicle speeds measured by infrastructure based traffic detectors (and thus approaches the space-mean-speed), but this still has the drawback of just concerning the location of the measuring site. This is particulary critical under congested traffic conditions.

ATMIS purposes require real-time insight in the spatial as well as the temporal fluctuations in the actual traffic process. In theory, this could be achieved by continuously (in order to obtain temporal fluctuations) taking aerial photographs of each road link in the road network (in order to obtain $\bar{v}_{SMS}$ indicating spatial fluctuations) or by installing infrastructure based traffic detectors (in order to obtain $\bar{v}_{TMS}$ indicating temporal fluctuations) with very short mutual spacings on each road link of the road network (in order to obtain spatial fluctuations). Both approaches are obviously far from realistic. Fortunately, the nature of the probe vehicle concept makes this novel monitoring technique preeminently suitable for obtaining both temporal and spatial fluctuations in the traffic flow over a complete (longitudinal section of a) road link (Linnartz, Westerman and Hamerslag, 1994). For this purpose we define the road link mean speed $\bar{v}$ as the speed obtained by averaging over all individual probe vehicles, averaging over the complete road link with length D (of 5-10 kilometers) as well as averaging over the entire time period with length T (of l-5 minutes). Subsequently, the road link mean travel time τ is defined as the time needed to traverse the complete road link under consideration with a constant

speed that equals the road link mean speed $\bar{v}$. Thus, this road link mean speed can be regarded as a specific combination of both the time-mean-speed and the space-mean-speed.

In appendix A we have established that the relation between the road link mean speed obtained from the probe vehicles and the time- and space-mean-speed is given by

$$\bar{v}_{probe} = \frac{D\,\bar{v}_{SMS} + T\,(\bar{v}^2_{SMS} + \sigma^2)}{D + T\,\bar{v}_{SMS}}$$

From this direct relation it becomes clear that when we would compute the road link mean speed from the probe vehicles for an infinitely small time interval (i.e. ***T-O),*** the space-mean-speed would result. Likewise, when we would compute the road link mean speed for an infinitely small space interval (i.e. D-O), the time-mean-speed would result. Hence, what has been formulated as ideal for ATMIS purposes, namely a continuous determination of the space-mean-speed or a determination of the time-mean-speed at a string of locations, can to a large extent be achieved by the probe vehicle concept.

When the mean speed on each road ***segment*** with length l_δ has been estimated, the mean speed on the entire road ***link*** with length $\sum_\delta l_\delta$ (the road link mean speed) can be computed by[5]

$$\bar{v}^* = \frac{\sum_\delta l_\delta}{\sum_\delta l_\delta / \bar{v}_\delta^*}$$

The reliability of this ultimate estimate is defined as the probability that the actual (unknown) value of the road link mean speed lies within an accuracy range around this estimate, whereas the (relative) accuracy is defined as the (relative) radius of this accuracy range. Assuming that the road segments are mutually independent, the accuracy (range) of the road link mean speed estimation is given by

$$\sigma_{road\ link} = \sqrt{\sum_{all\ segments\ a} \sigma_a^2}$$

We remark that when a road link consists of a large number of segments, there will be a large number of segment-inaccuracies that contribute to the final inaccuracy in the link estimate, but each of the regarding road segments will be relatively small so that each segment-inaccuracy will also be relatively small. Conversely, when a road link consists of a small number of segments, there will be few segment-inaccuracies contributing to the final link-inaccuracy, but the inaccuracies in each segment estimate will be larger. In other words, the inaccuracy of the road link mean speed estimate can be regarded to be independent of the number of constituting road segments, whereas the standard deviation of both the factual and the estimated road

[5] Throughout this dissertation we will use the superscript * to denote the estimator of a random variable.

segment speed depends on the length of the concerned road segment. On that account, we will develop and analyze an estimator departing from an entire road link instead of from a single road segment. It goes without saying that this will not effect the theoretical findings, which will be appropriate for both segment- and link-level. This will only influence the values of the different parameters to be used for presenting illustrative results.

The most simple and straightforward approach for estimating the road link (segment) mean speed from the sample of the N_p probe vehicle speeds comprised in the probe reports that are received in the traffic center, is simply averaging the individual probe vehicle speeds v_i. This method is the most obvious and simple approach and is commonly referred to as the Moment estimator (see for instance (Grimmett and Stirzaker, 1982)). Here, we will refer to this estimator *as* the ***direct estimator.*** This statistical estimator for estimating the mean of a distribution from a sample of this distribution basically estimates the road link mean speed directly with the sample mean of the received individual probe vehicles. Hence[6]

$$\bar{v}^*_M = \frac{1}{N_p} \sum_{i=1}^{N_p} v_i$$

3.2.3 Conclusions

Obviously, the discussed estimation procedure is quite precarious, in particular when it concerns very small sample sizes (of typically 1 or 2 probe reports per time period per road segment), while the mutual (probe) vehicle speeds may considerably diverge and may be not representative of the entire flow. Furthermore, it neglects (empirical) traffic engineering experience about the common behaviour of (the speed of) traffic flows.

Not all speeds of the macroscopic traffic flow have equal probability of occurrence. In daily practice, either the flow of traffic is undisturbed (free-flowing) and propagates with a regular speed (e.g. of circa 110 km/h, with a certain dispersion) or the flow of traffic is' detained (disruption or congestion), the circulation stagnates and the traffic flow propagates with a disrupted speed (e.g. of circa 35 km/h with a certain dispersion which is, due to for instance stop-and-go phenomena, larger than the above-mentioned free-flow dispersion). Speeds in between (e.g. of circa 60-70 km/h) are unusual and have a much lower probability of occurrence. These customarily occur when the traffic transits from one state (i.e. free-flow or congestion) to the other.

In the suggested direct (Moment) estimator, we have disregarded these existing interrelationships between the actual (mean) speed $\bar{v}$ and the historically determined behaviour of this speed. Hence, the straightforward direct estimator can be improved by incorporating knowledge about the fundamental behaviour of (the speed) of the traffic flow. We will denote this historic mean

[6] Note that a clear distinction between the individual probe vehicle speeds v_i that are used in this formula and concern distinct locations and distinct time instants, and the individual vehicle speeds used in the formulae for the time- and the space-mean-speed, which concern either equal locations or **equal** time instants.

speed by $\bar{\bar{v}}$. This novel approach leads to the design of a multi-layered (Bayesian) speed model that represents the incorporated historic data for estimating the actual road link mean speed, that we will develop and analyze in the next section.

3.3 Estimation of Road Link Mean Speed and Travel Time using Historic Data

From the previous section, we have argued that estimating the road link mean speed and road link mean travel time using the simple direct estimator is rather arduous when the number of received probe vehicle samples during a certain time period is low. In this section, we will develop and analyze a more advanced procedure for estimating the road link mean speed from received probe vehicle samples, that incorporates prior knowledge about the (historic) behaviour of the (speed of the) traffic flow. Before formulating this (Bayesian) estimating procedure in sub-section 3.3.2 and 3.3.3, we will formulate a traffic flow model in sub-section 3.3.1, that constitutes the theoretical setting.

3.3.1 Traffic Flow Model

In order to accomplish the estimation of the road link mean speed $\bar{v}$ and the mean travel time? we define a traffic flow model to describe the coherence between the speeds of the individual (probe) vehicles v_i, which are comprised in the probe messages that are received in the traffic center, and the actual (mean) speed of the traffic flow on one road link, which is to be estimated from the probe samples. This traffic flow model is based on one of the fundamental relationships between the main traffic parameters, density *k, flow q* and speed v, namely the speed/flow diagram (vq-diagram). This fundamental q-diagram shows the progression of the speed on a (crosscut of a) road link influenced by the road link traffic flow. In this fundamental .vq-diagram three regimes can be distinguished: free-flow, instability and congestion. Since a, state of instability actually can be regarded as a transition from free-flow to congested flow or vice versa, our traffic flow model focuses on the two stable regimes only, each including a branch of the third, instable, regime.

The individual probe vehicle speeds v_i on a road link are assumed to be independent random variables following a certain distribution F around the road link mean speed $\bar{v}$ with variance σ^2:

$$v_i \sim F(\bar{v}, \sigma^2)$$

This distribution F can be obtained through extensive measurements and thus constructing a historic database.

The assumption that the individual probe vehicle speeds are independently distributed is realistic when only probe vehicles are considered at a low penetration grade. At an increasing probe vehicle penetration grade, the independence of the individual probe speeds diminishes, but under these conditions the number of received probe messages will be sufficiently high to deploy the

direct estimator described in the previous section.

The variance σ^2 of the distribution of the random variables v_i depends on the regime in the *vq*-diagram (i.e. free-flow or congested flow). A conceivable refinement would be to assume this variance to be dependent on the actual value of the traffic flow *q,* but this has not been put into effect as yet we have assumed that only probe vehicle data is available from which accurate flow information can not be obtained. When road segments would have been considered in this analysis instead of road links the variance of the random variables v_i would also depend on the length of the regarding road segment.

Furthermore, the actual road link mean speed $\bar{v}$ is a random variable which is distributed around a historic mean $\bar{\bar{v}}$. As already argued, we distinguish between a free-flow regime and a congested regime, hence

$$\bar{v}_f \sim G(\bar{\bar{v}}_f, \sigma_f^2)$$

for free-flow traffic, and

$$\bar{v}_c \sim H(\bar{\bar{v}}_c, \sigma_c^2)$$

for congested traffic.

We further provisionally assume that the values of $\bar{\bar{v}}_f$, $\bar{\bar{v}}_c$, σ^2_f and σ^2_c are fixed constants that are empirically known, for instance kept in a historic database in the traffic center. For an example of such a database and its construction we refer to (Heidemann, 1988) and (Heidemann, 1990) in which an empirical register has been established with historic speeds. For our purposes this register should be extended by differentiating between distinct months of the year (e.g. winter/summer), days of the week (e.g. working days/week-end), times of the day (e.g. rush hours) etcetera and also dissimilar weather conditions. So, the distributions G and *H are* gradually obtained.

The latter assumptions could further be improved by modelling the historic mean speeds $\bar{\bar{v}}_f$ and $\bar{\bar{v}}_c$ as variables in our traffic flow model, whose value depends on the actual traffic flow *q. This* might be accomplished by using the fundamental vq-diagram. This will be done in the next chapter. In this chapter, the flow and density parameters have been assumed unavailable. Hence, the traffic flow model sketched in figure 3.4, is simplified by representing the historic means by a historic mean for free-flow traffic, $\bar{\bar{v}}_f$, and a historic mean for congested traffic, $\bar{\bar{v}}_c$ (Westerman, 1995).

Another approach to enhance the performance of our traffic flow model with respect to the historic speeds $\bar{\bar{v}}_f$ and $\bar{\bar{v}}_c$ is by incorporating the road link mean speed estimation in the previous time period(s). In this way, the steady historic means becomes more flexible and adaptive to the prevailing traffic conditions. This can be accomplished by employing some form of smoothing (e.g. single exponential smoothing), of course provided that the prevailing traffic

regime remains the same. Hence, the historic mean speed, $\bar{\bar{v}}(t)$, to be used for the estimate at time instant t becomes a function of t and is given by

$$\bar{\bar{v}}_{c,f}(t) = \gamma \bar{\bar{v}}_{c,f} + (1\text{-}y) \bar{v}^*(t-1)$$

where y is a smoothing constant with range [0,1] and $\bar{\bar{v}}_{c,f}$ denotes the historic free-flow or congested mean speed. When we should want to attach considerable significance to the theoretical prior knowledge about the behaviour of the traffic flow, y is preferably kept high (e.g. > 0.8). When we have only little confidence in the practical correctness of these historic values, γ is preferably kept low (e.g. < 0.2). By using such a smoothing average rather than fixed values for the expected mean of the road link mean speed, a system results that continuously adapts to the actual traffic process. Another advantage is that when few probe reports are received, this prior knowledge, that comprises also the previous estimation, has a relatively significant contribution. Accordingly, at low N_p also the probe reports that were received in the previous time period are automatically taken into account. Conversely, at a high N_p the influence of both the historic prior and the previously estimated prior take hardly any share in the new estimation. This makes this traffic model preeminently suitable for estimating the road link mean speed from probe vehicle samples, both in situations with few and in situations with many received reports.

The traffic flow model specified above can be exemplified as a three-layered micro/meso/macro-model, consisting of the following three layers (see figure 3.4):

Macro-level: At the macro-level we discern the historic road link mean speed for the free-flow ($\bar{\bar{v}}_f$) and the congested ($\bar{\bar{v}}_c$) regimes. These historic means have been taken as endogenous constants (the two horizontal lines labelled $\bar{\bar{v}}_f$ and $\bar{\bar{v}}_c$), but can also be modelled dependent on the actual traffic flow q when this information would be available (sketched ellipse; see also chapter 4).

Meso-level: At the meso-level we perceive the actual road link mean speed $\bar{v}$.. This is a value with a distribution of which the expected mean is given by the higher macro-level.

Micro-level: At the micro-level we find the speeds of the individual (probe) vehicles v_i on the various segments of the considered road link. These stochastic variables are distributed with a stochastic location parameter given by the higher meso-level.

The multi-layered traffic flow model discussed above is the theoretical setting for designing a method to reliably estimate the value of the meso-level variable, i.e. the road link mean speed $\bar{v}$ which in turn is composed of a sequence of estimated speeds that are constant per road segment for all probe vehicles, from a quantity of micro-level observations, i.e. the number N_p of observed (individual) probe vehicle speeds $v_1,\ldots, v_{Np}$. This analysis basically consists of the following two steps, which are dealt with in the ensuing two sub-sections:

firstly, selection of the prevailing traffic regime (free-flow or congestion) (analyzed in sub-section 3.3.2);

secondly, estimation of the road link mean speed $\bar{v}$, using the outcome of the first step (analyzed in sub-section 3.3.3).

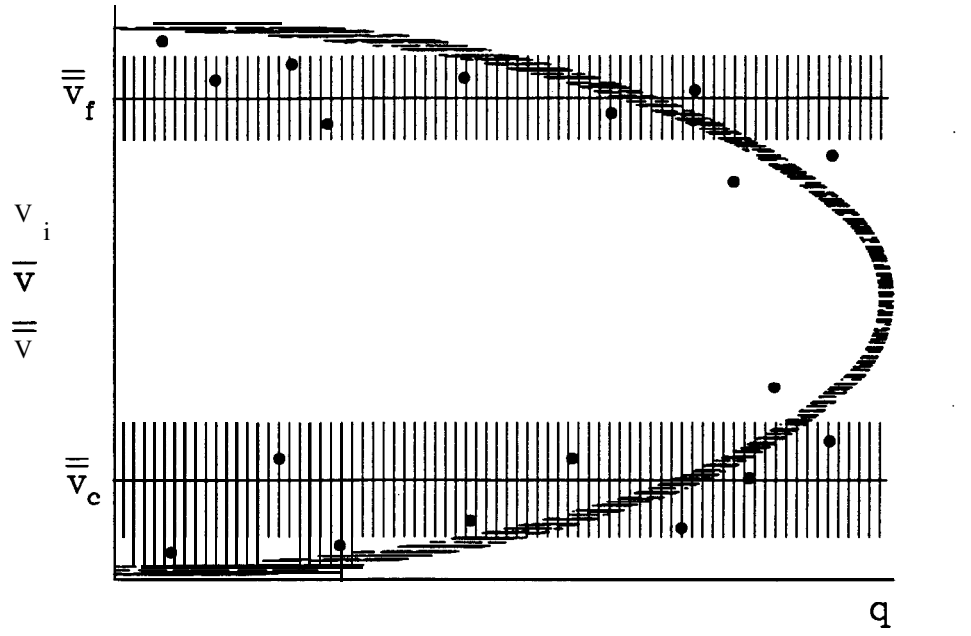

Figure 3.4 Multi-layered speed model for estimating road link mean speed from probe vehicle observations for ATIS

This process is illustrated in figure 3.5[7], which denotes a crosscut of the hybrid traffic flow model in figure 3.4, at one particular value of the traffic volume *q*. In this chapter, where we have aasumed that only probe vehicle data is available, this crosscut is equal for all values of *q* as we have modelled the historic mean speeds as constants. Figure 3.5 shows that the probability density function (pdf) of the road link mean speed is distributed with two local maxima, meaning that the prevailing traffic regime can be estimated with a certain reliability, depending on the (sample mean of the) speed values comprised in the received probe vehicle reports.

Besides major freeways, also information about the actual traffic conditions on road networks other than the primary freeway network are of importance. For secondary and urban road networks an approach that is similar to the one previously discussed for freeways can be followed, but now the shape of and the absolute values of the historic mean speeds in the hybrid micro/meso/macro-traffic flow model will differ. For this reason, the theories that we will develop in this section are generically applicable for all kinds of primary, secondary and urban road networks. With respect to the illustration of these theories (in sub-sections 3.3.2.3 and 3.3.3.3 presenting results) we will principally be orientated towards the primary freeway network.

[7] In the denotive illustration in figure 3.5, the probability that the traffic flow is in a state of congestion is presumed to amount to 1/3 (*P(congestion)=0.33*) and the probability that it is free-flowing is presumed to amount to 2/3 (*P(free-flow)=0.67*). We remark that these values of 1/3 and 2/3 are only used for drawing up this illustration. For further parameters that have been used in figure 3.5 we refer to sub-section 3.3.2.3.

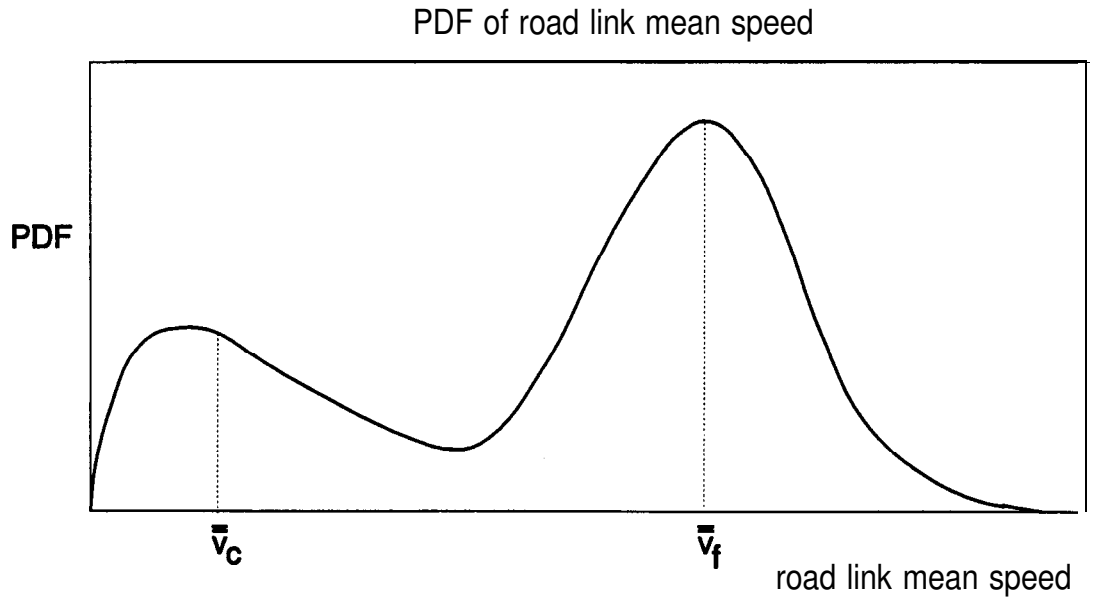

Figure 3.5 Probability-density function of road link mean speed at a given traffic volume

3.3.2 Step 1: Prevailing Regime Selection

We define congestion as the status of the road traffic where the actual traffic demand exceeds the available road capacity, either locally or network-wide. Congestion leads to delay; at first only the travel times, as well as their variability, over the congested (part of the) road link increase, later queues will build up leading to an increase in the travel times over adjacent upstream links too. This definition implies that congestion may occur due to changes in the actual traffic demand or due to changes in the actual road capacity. The first type of congestion occurs when the volume of traffic temporarily surpasses the capacity, which we will therefore denote with 'demand congestion'. These situations customarily arise day-by-day during peak hours at many locations and are then referred to as recurrent congestion. The second type of congestion arises when the actual road capacity is temporarily diminished due to a specific occurrence, which we will therefore denote with 'supply congestion'. As these situations happen exceptionally they are often referred to as non-recurrent congestion'.

The delay induced by inter-urban congestion is appropriately addressed by (Hall, 1993), where

[8] Note that in these definitions the terms 'demand' and 'supply' refer to the cause of the congestion, while the terms 'recurrent' and 'non-recurrent' refer to the emergence of the congestion. This implies that both demand and supply congestion can be recurrent as well as non-recurrent, although demand congestion is usually recurrent and supply congestion is mostly non-recurrent. Non-recurrent demand congestion may occur after the closing of a special event, an example of recurrent supply congestion may be the scheduled opening of a bridge.

travel time on a highway is represented as the sum of two factors: a free-flow time (travel time in the absence of congestion) and a congestion delay. Where the free-flow travel time τ_f is primary a function of design geometrics and associated speed limits, congestion delay is defined as the sum of a recurrent component τ_{cr} and a non-recurrent component τ_{cn}, viz.

$$\tau = \tau_f + [\ \tau_{cr} + \tau_{cn}\ ]$$

One major focus of ATMIS is to communicate an accurate and reliable value of the actual total travel time τ (on a road link or of a (part of a) trip) to the (equipped) road user. For this ATMIS purpose it is important to timely observe deviations of this total travel time τ from the free-flow travel time τ_f. It is remarked that τ_f is not necessarily equal to the 'normal' or historic travel time, τ_{hist}. In front of notorious bottlenecks, such as bridges and tunnels, a certain level of (recurrent) congestion during peak hour will occur day-by-day and can thus be considered 'normal' (i.e. part of τ_{hist}).

In order to further analyze supply congestion, we define an incident as any non-periodic event that causes a reduction in the actual road capacity. Analogous to (Van Vuren and Leonard, 1994) we distinguish between predictable and unexpected incidents; the former consist of, for example, planned road works and regular parking infringements in urban areas, whereas the latter include accidents and vehicle breakdowns. However, it is often neglected that not every incident will have a same impact on the traffic flow. One possible way to classify prevailing traffic conditions after the occurrence of an incident is based on the effect the incident has on the traffic flow.

This leads to the following categorization of incidents:

Category I: Incidents where the reduced capacity due to the occurrence of the incident is equal to or smaller than the actual traffic flow, or

Category ZZ: Incidents where the reduced capacity is larger than the actual traffic flow.

In the first situation the behaviour of the traffic flow is considerably affected which results in non-recurrent supply congestion. Shockwaves will arise which will propagate both upstream and downstream of the location of the incident. Consequently, alterations will occur in the fundamental traffic flow parameters. These alterations are directly experienced and registered by the probe vehicles participating in this traffic flow and can also be measured by fixed traffic detectors provided that the detector spacings are relatively short. As this class of incidents brings about a significant increase in the travel time τ_{cn} across the involved road link, quick and reliable detection of these incidents is of major importance for ATMIS.

In the second situation the occurrence of the incident has no significant impact on the behaviour of the traffic flow and no (non-recurrent supply) congestion originates. Only very local disturbances are likely to occur, but shock-waves will propagate neither upstream nor downstream of the incident. As a result, no significant and sudden changes in the macroscopic traffic parameters can be observed and most existing automatic incident detection algorithms that are based on data from fixed traffic detectors are not capable of detecting this class of incidents (unless the location of the incident and the location of the fixed traffic detectors happen to coincide). As probe vehicles are assumed to register the location and the point in time at which

deviant conditions are experienced (which includes these very local disturbances), it is likely that also this class of incidents can be detected. Although such conditions will not result in congestion and the travel times for the road users will not be affected quick and reliable detection of this class of incidents is desired for ATMIS purposes, since another major task is to clear the incident as soon as possible (possibly by bringing into action police or fire department).

The Automatic Congestion Detection (ACD) algorithm is the first step in the process of estimating the road link mean speed from received probe vehicle samples and comprises determining in which of the two considered regimes of the adopted fundamental vq-diagram the actual traffic process is most likely to be. To make ***this prevailing regime selection, we*** determine how justified it is to hold on to the selection we based on the set of data in the previous time interval. In general, this selection will be the most reliable and will function as the null hypothesis in the statistical test.

Suppose that before processing the current speed measurements of the probe vehicles, the traffic was assumed to be in a state of free-flow. The regime selection then is directly based on the outcome of the following statistical test:

$$H_0 : \textit{free-flow traffic}$$
$$H_1 : \textit{congested traffic}$$

In order to carry out the test, we use the sample mean speed (of the N_p probe messages) as an appropriate test statistic. We will reject the null hypothesis when this sample mean speed is contained in a certain critical region denoted by $[0,K)$ in which the parameter K is a distinctive 'policy threshold', viz.

$$\frac{1}{N_p} \sum_{i=1}^{N_p} v_i \in [0,K)$$

When the traffic would have been assumed to be in a state of congestion, the regime selection would be based on the outcome of the following statistical test:

$$H_0 : \textit{congested traffic}$$
$$H_1 : \textit{free-flow traffic}$$

while the null hypothesis would be rejected when the sample mean speed would be contained in the critical region denoted by (K, ∞), viz.

$$\frac{1}{N_p} \sum_{i=1}^{N_p} v_i \in (K, \infty)$$

In statistical tests, two error types can be distinguished: ***error of commission,*** when the null hypothesis is unjustly rejected, and the ***error of omission,*** when the null hypothesis is unjustly not rejected (H_1 is true, but H_0 is chosen). Given the faith generally placed on the null

hypothesis, the error of commission is the most serious of the two. Hence the test's critical region, or in our case critical value (*K*), is chosen such that the error of commission is as small as possible (within a domain of acceptable values).

Whichever of the two tests is used, there are four alternatives (see also (Stephanedes and Chassiakos, 1993)):

1. actual traffic is in ***free-flow,*** and ***free-flow*** is the selected regime (correct NO CONGESTION decision);
2. actual ***traffic*** is in ***free-flow, while congestion*** is the selected regime (false alarm);
3. actual ***traffic*** is ***congested,*** **while** ***free-flow*** is the selected regime (missed congestion);
4. actual ***traffic*** is ***congested,*** and ***congestion*** is the selected regime (correct CONGESTION decision);

From these alternatives it becomes clear that the performance of the prevailing regime selection for Automatic Congestion Detection can be characterized by two so-called Measures of Effectiveness (MOE's), namely (see also chapter 4):

- ***False alarm rate (far)***; we define the false alarm rate ***(far)*** of our ACD algorithm as the probability that the actual traffic state is one of free-flow, while the observations lead us to decide that it is one of congestion.
- ***Detection rate (dr)***; we define the detection rate ***(dr)*** of our ACD algorithm as the probability that the actual traffic state is one of congestion and the observations lead us to decide that it is such (for AID the detection rate ***dr*** is commonly defined as the number of detected incidents in proportion to the number of occurred incidents).

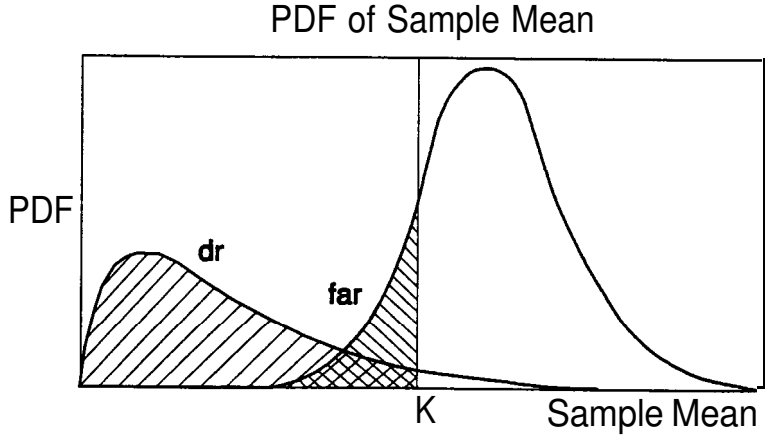

Figure 3.6 Effect of a high threshold K on false alarm rate (far) and detection rate (dr)

Furthermore, for conventional Automatic Incident Detection (AID) another MOE is commonly considered, namely:

- ***Time to detect (ttd);*** we will not use this MOE in our study concerning ACD, as detection of congestion is only an instrument that is needed to estimate the road link mean speed and evaluate the reliability and accuracy of this estimation (for AID, the ***ttd*** is commonly defined as the time needed to detect an incident).

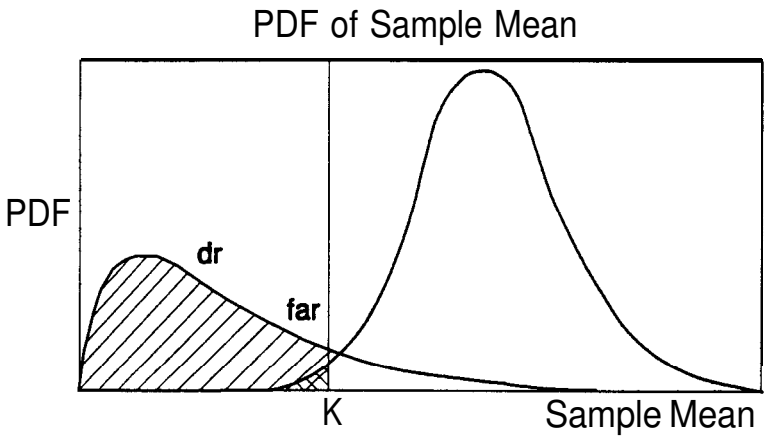

Figure 3.7 Effect of a low threshold K on false alarm rate (far) and detection rate (dr)

As we will see, these MOE's are highly dependent and conflicting; enhancing one MOE dilapidates the others.

The decision rules formulated in the foregoing and the resulting false alarm rate *(far)* and detection rate *(dr)* are illustrated in the following two figures, which denote the probability density function (pdf) of the sample mean. The vertical dash in both figures represents the critical value or 'policy threshold' ***K***. In figure 3.6 this policy threshold ***K*** has been taken relatively high, resulting in a high detection rate *(dr)* but also in a high false alarm rate *(far)*. *In* figure 3.7 a lower value has been chosen for the policy threshold ***K***. This decreases the false alarm rate but decreases the detection rate as well.

Once again we would like to point out that a clearly perceptible distinction between Automatic Congestion Detection (ACD) and Automatic Incident detection (AID) exists that is reflected in the prevailing traffic regime selection. Where one ATMIS is to to deploy measures' after any disruption in the traffic process (e.g. by means of traffic calming, ramp metering, road blocks or even dispatching emergency vehicles in the case of an accident), another ATMIS task is to only inform the road users (e.g. by means of congestion warning or recommending diversions and other routes to road users) when the additional travel time that results from the congestion necessitates. This means that ACD exclusively concentrates on those disturbances in the traffic flow that cause a significant increase in the travel times for the road users, while AID should detect any disruption of the traffic process. Therefore, for ACD a very low false alarm rate (far) is required as by no means misleading information should be disseminated. It is apparent that also the ultimate detection rate should be sufficiently high, but the time before traveler information needs to be distributed is not of such determinative importance (this period is likely to amount 5-15 minutes), so the time to detect may be rather long compared to AID.

If we recall the statistical test in case the previous regime selection procedure resulted in a free-flow decision, to be

$$H_0 : \textit{free-flow traffic}$$
$$H_I : \textit{congested traffic}$$

the critical value K for the sample mean speed follows from minimizing the probability of an error of commission over a limited domain of acceptable values. Note that an error of commission in the test above is simply a false alarm. The detection rate here equals the probability that H_0 is rejected justly. The following optimization problem leads to the optimal (and acceptable) value of the threshold K

$$min,\ far(K) \text{ subject to } dr(K) \geq D,$$

wherein the false alarm rate is minimized under the constraint that the detection rate is greater than or equal to D (e.g. $D=75\%$). The characteristics that the two rate functions are strictly increasing in K indicate that in the optimal situation, the constraint will be satisfied with equality. Thus, the optimization problem will lead to an optimal K_D^* such that $dr(K_D^*)=D$. The actual (minimal) false rate is then $far(K_D^*)$.

In the alternative statistical test, where the previously selected regime was one of congestion,

$$H_0: \textit{congested traffic}$$
$$H_1: \textit{free-flow traffic}$$

the critical value K is determined analogously, by solving the optimization problem

$$max_K\ dr(K) \text{ subject to } far(K) \leq F.$$

Note that in this test, the error of commission (to be minimized) equals one minus the detection rate, so that the detection rate is to be maximized. Again, in the optimal situation, the constraint will be satisfied with equality. Thus, the optimization problem will lead to an optimal K_F^* such that $far(K_F^*) = F$. The actual (minimal) detection rate is then $dr(K_F^*)$.

Note that it is clearly impossible to simultaneously optimize the detection and false alarm rate, since these have conflicting relations with control variable K. The obvious approach to follow then is to optimize one of both measures over a restricted domain, defined by the other' measure. This is exactly what is done in the statistical tests. It will appear that in some cases, especially when the sample size N_p is extremely small, the minimal false alarm rate (first test) remains relatively high (in certain *cases* the *far* will even remain higher than desired, which,. in general, means higher than an exogenous value, say F of, for instance, 10%). This is simply an indication that the sample size is to small to make a responsible ACD decision. A similar argument can be given for the second test type.

The result of this first step in the process of estimating the road link mean speed $\bar{v}$ from probe vehicle samples is a judgement about the present status of the traffic, viz. either in free-flow or congested state. This judgement is required as the means $\bar{\bar{v}}_f$ and $\bar{\bar{v}}_c$ as well as the variances σ^2_f and σ^2_c differ considerably.

Aside from using the sample mean of the probe reports for detecting congestion, also the 'amount of contents' of the probe vehicle messages might be used. As the probe vehicles are assumed to register the location at which a significant deviation in the experienced speed (i.e. an

alteration of the speed of more than $\alpha\%$) occurs, the probe vehicle messages will contain more locations (and corresponding travel times between these locations) under conditions of high traffic loads (congestion) than under conditions of low traffic loads (free-flow). A similar approach has first been recognized by (Greenshields, 1955), in which a so-called Quality of Flow parameter Q was defined *'...intended to reflect the aspects of traffic flow that require the driver's response...'*. This Q parameter is determined for a number of vehicles participating in the traffic flow according to $Q=KS/\Delta_s\sqrt{f}$ in which K is a constant (with a value of 1000), S in the average speed, Δ_s is the sum of speed changes (of 2 miles per hour or more per mile), and f is the number of such speed changes per mile. Also the Acceleration Noise parameter, defined in (Chandler, Herman and Montroll, 1958) to study vehicle interactions, or the Coefficient of Variation of Speed, expressing the ratio of speed noise (standard deviation of time-averaged speeds) to average speed as defined in (Chang and Herman, 1978), could be applied in order to enhance the process of determining the prevailing traffic conditions as described in the preceding sub-sections. This has however not been utilized in our analyses as the applicability of this approach can only be investigated with extensive empirical traffic data (see for instance (Sermons and Koppelman, 1995)).

In order to examine whether this process of Automatic Congestion Detection (ACD) can be practically performed, we will further investigate the required values of the policy threshold K in order to achieve a certain level of performance of the ACD procedure. This level of performance can be expressed in a low false alarm rate (e.g. 10% or less) and a high detection rate (e.g. 75% or more). In the ensuing two sub-sections we will analyze the relation between the policy threshold and the false alarm rate *(far)* and detection rate *(dr)* and determine practical values for the parameter K.

3.3.2.1 False Alarm Rate

The false alarm rate *(far)* has been defined as the probability that the actual traffic state is one of free-flow, while the observations lead us to decide that it is one of congestion. So,' the false alarm rate is simply the value of the cumulative probability function of the average of the observed probe vehicle speeds for the value of the 'distinctive policy threshold' K, and indicates the probability that the sample mean following the free-flow speed distribution is smaller than the policy threshold K (see also the surface with the 'left hatches' (i.e. from bottom right to top left) in figure 3.6 and 3.7):

$$far(K) := P_f\{\frac{1}{N_p}\sum_{i=1}^{N_p} v_i < K\}$$

3.3.2.2 Detection Rate

We have defined the detection rate *(dr)* as the probability that the actual traffic state is one of congestion, while the observations lead us to decide that it is such. So, the detection rate departs

from a congested state of traffic and is a function of the policy threshold K, for the average of the observed probe vehicle speeds and indicates the probability that the sample mean following the congestion speed distribution is smaller than the policy threshold K *(see also the* surface with the 'right hatches' (i.e. from bottom left to top right) in figure 3.6 and 3.7):

$$dr(K) := P_c\{\frac{1}{N_p}\sum_{i=1}^{N_p} v_i < K\}$$

3.3.2.3 Results

In order to illustrate the interpretation and implications of the established relations between the number N_p and the value $v_i, i = 1,...,N_p$, of the individual probe vehicle speeds received in the traffic center, as well as to determine the possibilities for performing Automatic Congestion Detection (ACD) quantitatively, the detection rate *(dr)* and the false alarm rate *(far)* have been computed. We have chosen realistic parameter values for the assumed probability distributions.

First of all, we assume that the speeds under free-flow conditions are normally distributed, hence (all speed values are in km/h)[9]:

$$v_i \sim N(\bar{v}, 15^2) \quad \text{and} \quad \bar{v} \sim N(110, 15^2)$$

The above assumptions are consistent with empirical findings (e.g. (BGC, 1990)).

Under conditions of congestion the variance of the road link mean speed will be larger than under free-flow conditions. Moreover, the probability of occurrence of negative speeds will be positive when using a normal distribution. For the latter reason we assume a gamma distribution for the speeds when there is congestion, although sofar little consistent empirical confirmation has been found concerning speed distributions under conditions of congestion (BGC, L989; Hall and Brilon, 1994). Hence:

$$v_i \sim \Gamma(\frac{\bar{v}}{625}, \frac{\bar{v}^2}{625}) \quad \textit{such that} \quad \frac{\rho}{\lambda} = \bar{v} \textit{ and } \frac{\rho}{\lambda^2} = 25^2$$

and

$$\bar{v} \sim \Gamma(\frac{7}{125}, \frac{49}{25}) \quad \textit{such that} \quad \frac{\rho_c}{\lambda_c} = \bar{v}_c = 35 \textit{ and } \frac{\rho_c}{\lambda_c^2} = \sigma_c^2 = 25^2$$

By presuming these parameter values for the equations derived in the previous two sub-section, we are able to compute both the false alarms rate and the detection rate for various values of N_p (the number of received probe vehicle messages) and the policy threshold (control variable) K.

---

[9] These parameter values concern an entire road link. Since the variance in the individual probe vehicle speeds is assumed to depend on the length of the regarded stretch of road, the variances will be smaller in the case a single road segment is concerned.

The presumed values for the parameters listed in the above have also been used for drawing up the probability density function of the road link mean speed in figure *3.5.* Hence, the ***pdf*** of the road link mean speed (pictured in figure 3.5) is

$$\bar{v} \sim \frac{1}{3}\Gamma(\frac{7}{125}, \frac{49}{25}) + \frac{2}{3}N(110,225)$$

Given the above assumptions, we will now further elaborate the formulas for the false alarm rate (*far*) and the detection rate *(dr).*

False Alarm Rate

In order to derive a (relatively simple) formula for the false alarm rate, we note that if the individual probe vehicle speeds are normally distributed around the road link mean speed $v_i \sim N(\bar{v}, \sigma^2)$, then the sample mean is also normally distributed:

$$\frac{1}{N_p}\sum_{i=1}^{N_p} v_i \sim N(\bar{v}, \frac{\sigma^2}{N_p})$$

so that

$$\frac{1}{N_p}\sum_{i=1}^{N_p} v_i = \left(\frac{1}{N_p}\sum_{i=1}^{N_p}(v_i - \bar{v})\right) + \bar{v} \sim N(0, \frac{\sigma^2}{N_p}) + N(\bar{\bar{v}}_f, \sigma_f^2) = N(\bar{\bar{v}}_f, \frac{\sigma^2}{N_p} + \sigma_f^2)$$

which gives an independent distribution for the sample mean $\frac{1}{N_p}\sum_{i=1}^{N_p} v_i$.

In this way, the false alarm rate defined in sub-section 3.3.2.1 is given by:

$$far(K) := P_f\{\frac{1}{N_p}\sum_{i=1}^{N_p} v_i < K\} = \Phi_{\bar{\bar{v}}_f, \frac{\sigma^2}{N_p} + \sigma_f^2}(K)$$

with Φ_{μ,σ^2} the cumulative density function of a normal distribution with mean μ and variance σ^2. Note that the false alarm rate is strictly increasing in K with $far(0)=0$ and $\lim_{K\to\infty} far(K) = 1$.

Detection Rate

In order to derive a (relatively simple) formula for the detection rate *(dr),* we note that if

$v_i \sim \Gamma(\lambda, \rho)$, then the sample mean is distributed according to

$$\frac{1}{N_p}\sum_{i=1}^{N_p} v_i \sim \Gamma(N_p\lambda, N_p\rho)$$

As the distribution of the sample mean is dependent on $\bar{v}$ and is not directly known, conditioning on $\bar{v}$, given the detection rate defined in sub-section 3.3.2.2, gives

$$\int_{0+}^{\infty} P_c\{\frac{1}{N_p}\sum_{i=1}^{N_p} v_i < K \mid \bar{v}=a\}\ P_c\{\bar{v}=a\}\ da = \int_{0+}^{\infty} F_{\frac{N_p a}{\sigma^2},\frac{N_p a^2}{\sigma^2}}(K)\ f_{\frac{\bar{v}_c}{\sigma_c^2},\frac{\bar{v}_c^2}{\sigma_c^2}}(a)\, da$$

with $F_{\frac{N_p a}{\sigma^2},\frac{N_p a^2}{\sigma^2}}$ the cumulative (gamma) density function of the sample mean and $f_{\frac{\bar{v}_c}{\sigma_c^2},\frac{\bar{v}_c^2}{\sigma_c^2}}$ the probability (gamma) density function of the road link mean speed v. Note that the detection rate is strictly increasing in K with $dr(0)=0$ and $\lim_{K\to\infty} dr(K) = 1$.

Figure 3.8 shows the computed false alarm rates (*far's*) and detection rates *(dr's)* for $N_p \in \{1,2,3,5,10\}$, while figures 3.9 and 3.10 show the relevant parts of *the far-* and *dr-curves,* grabbed from figure 3.8. What is relevant, is defined by what is considered acceptable (minimum value of *far(K))* or desirable (maximum value of *dr(K)*).

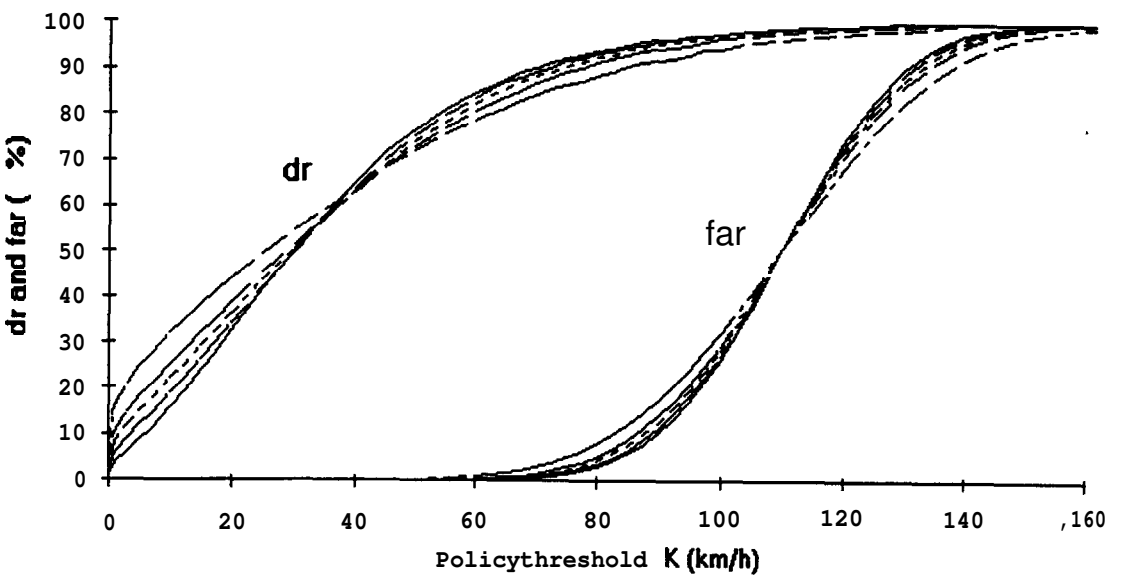

Figure 3.8 Computed false alarm rates (far) and detection rates (dr) for $N_p \in \{1,2,3,5,10\}$ (see figures 7.9 and 7.10 for a legend)

In our illustrative example, based on the values of the detection and false alarm rates and the acceptable levels of significance of F=10% and D=90%, we find that for the statistical test

$$H_0: \textit{free-flow traffic}$$
$$H_1: \textit{congested traffic}$$

the applied critical K-value for the sample mean (with e.g. N_p=3) equals K^*_D=74.1.

Alternatively, in the test

$$H_0: \textit{congested traffic}$$
$$H_1: \textit{free-flow traffic}$$

K=K^*_F=87.8, again for N_p=3. Note that the range of sample mean speeds that lead us to conclude that the actual traffic process is in a state of congestion, is larger in case congestion was also the prevailing regime in the previous ACD procedure. This is obvious, considering the faith we have (and should have) in our previous decision(s) (null hypothesis). A similar argument holds for free-flow assumption / selection.

Table 3.3 gives the threshold values K^*_D (with a minimum detection rate of D=90%) and a maximum false alarm rate of K^*_F (with F=1%, 5% and 10% respectively), along with the according detection and false alarm rates, for values of $N_p \in \{1,2,3,5,10\}$. For instance, from the second and third column we learn that for N_p=2 given a required detection rate of D=90%, the minimal possible false alarm rate is 3.7% (at K=77.2 km/h). At a maximum required false alarm rate of F=5% the corresponding detection rate is dr=92.8% (at K=81.5 km/h) for N_p=3. When N_p=1, the detection rate amounts to dr(86.4%) with a value of the policy threshold K of 75.1 km/h.

N_p	$K^*_{D=0.90}$ (km/h)	far($K^*_{D=0.90}$) (%)	$K^*_{F=0.01}$ (km/h)	dr($K^*_{F=0.01}$) (%)	$K^*_{F=0.05}$ (km/h)	dr($K^*_{F=0.05}$) (%)	$K^*_{F=0.10}$) (km/h)	dr($K^*_{F=0.10}$) (%)
1	84.4	11.4%	60.7	79.0%	75.1	86.4%	82.8	89.5%
2	77.2	3.7%	67.3	85.2%	79.8	91.1%	86.5	93.3%
3	74.1	1.9%	69.7	88.0%	81.5	92.8%	87.8	94.6%
5	71.6	1.0%	71.8	90.1%	83.0	94.0%	88.9	95.3%
10	70.0	0.6%	73.4	91.4%	84.1	94.8%	89.8	96.1%

Table 3.3 Applicable values for policy threshold K for different values of N_p and accompanying false alarm rates (far) and detection rates (dr)

From table 3.3 we learn that performing Automatic Congestion Detection with a low false alarm rate of 1% is possible, also when only 1 probe report is successfully received in the traffic center. In this situation the detection rate amounts to 79%.

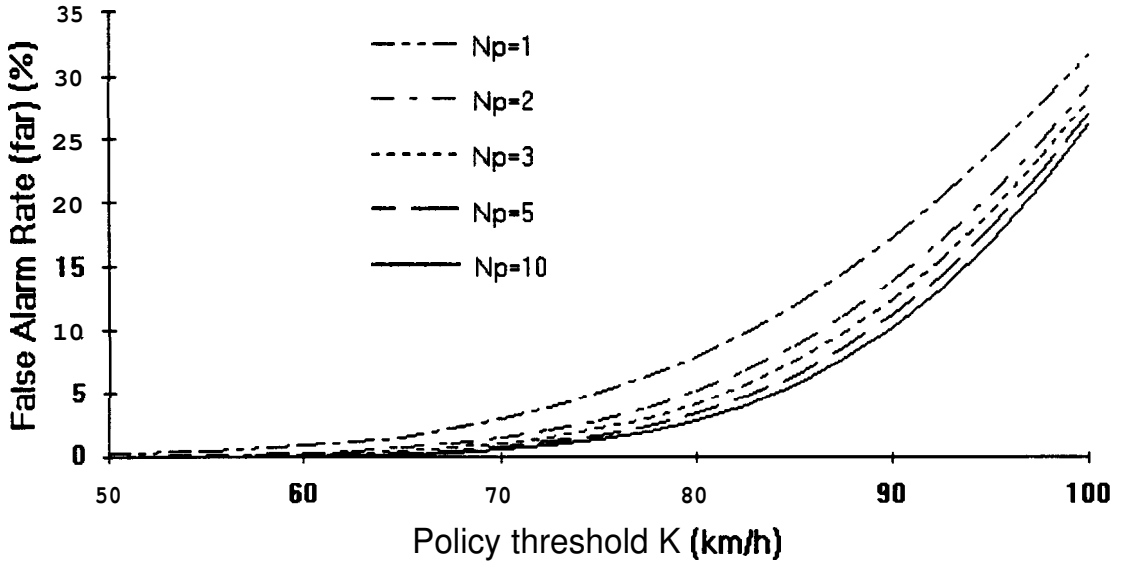

Figure 3.9 Relevant part of computed false alarm rates as function of policy threshold K

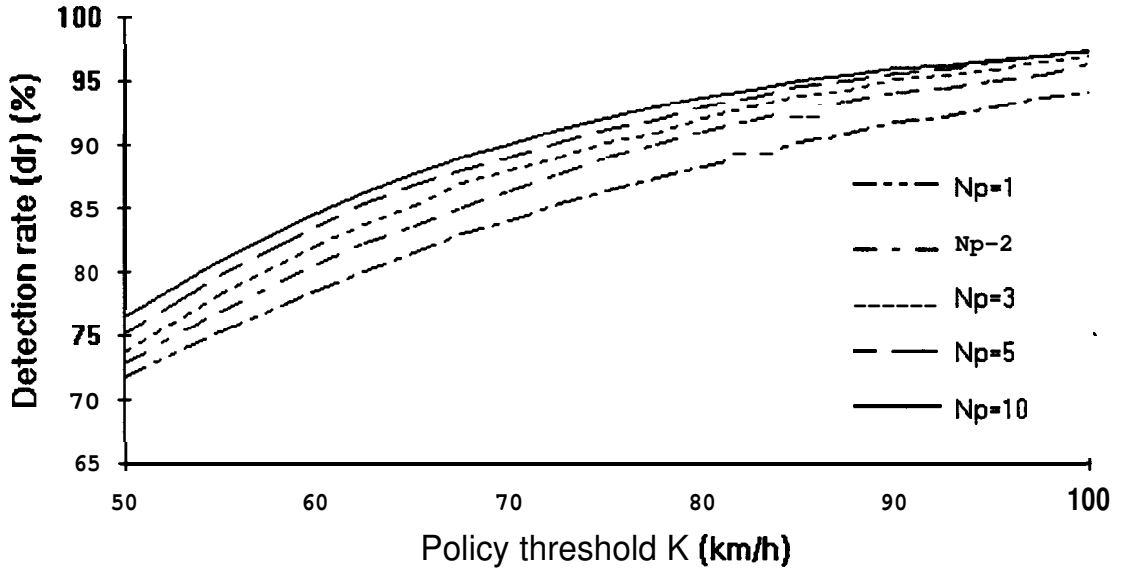

Figure 3.10 Relevant part of computed detection rates as function of policy threshold K

3.3.2.4 Conclusions

From the results in the figures 3.8, 3.9 and 3.10 and in table 3.3 it becomes clear that performing Automatic Congestion Detection (or, worded differently, selecting the prevailing traffic regime free-flow c.q. congestion) is reliably possible, even when only 1 probe report is received during a certain time period. When the required false alarm rate is rather high (*F=10%*) and N_p=2, the legitimate values of the policy threshold *K* range from 77.2 km/h to 86.5 km/h,

while the detection (false alarm) rates for the K values are contained in the interval [90.0% , 93.3%] ([3.7% , lO.O%]). Performing ACD with lower false alarm rates is also possible (e.g. far=1%), although this decreases the detection rate a little (dr=79% at N_p=1 and dr=85% at N_p=2). These values are considered to be acceptable, seeing that the time scale of the concerning ATMIS measures ranges from 5 to 15 minutes. Even though at a detection rate of 79% it will take some minutes longer to detect prevailing congestion beyond all doubt, it will be detected sufficiently swiftly for the concerning ATMIS purposes. The latter implies that the first step of our process of estimating the road link mean speed $\bar{v}$ can acceptably be performed when few probe reports are received (i.e. N_p=1 or N_p=2).

Reconsidering the results we have found in the previous PATH research project MOU107 6 we observe that at a moderate probe vehicle penetration grade of 1% (2%) the average number of probe reports that is received per kilometer from an average road link (i.e. an average traffic load (during peak hours) of circa 55 vehicles per kilometer road link and an average distance from the receiving base station of circa 6 km) approximately 1 or 2 (3 or 4) probe reports are successfully received per minute. This means that ATMIS has the possibility to reliably perform Automatic Congestion Detection on such road links during peak hours with time intervals of 1 minute as from a probe vehicle penetration grade of 1%. On road links that are more distant from a base station or contain less traffic, or during time periods outside peak hours the reliability of the ACD outcomes decreases below what can be considered acceptable. In this respect it is worthwhile mentioning that although the regular number of (probe) vehicles on a road link outside peak hours is about half of that during peak hours, it will increase when congestion occurs. When larger time intervals (e.g. of 5-15 minutes) are applied, which is reasonable for the concerning ATMIS purposes, a probe vehicle penetration grade of 1% will be amply sufficient in order to reliably perform ACD on any regular road link of a freeway during any time of the day. Moreover, for such longer time intervals (e.g. of 5-15 minutes) it becomes possible to estimate the actual traffic flow and density on a road link from probe vehicle samples with an acceptable accuracy and reliability. By also taking into account this type of actual traffic information, the process of selecting the prevailing traffic regime and the process of estimating the road link mean speed can be enhanced. This issue is addressed in the ensuing chapter, where we extent both processes by incorporating this type of information.

It is important to note that only a very small number of probe vehicles is sufficient 'for reliably selecting the prevailing traffic regime.

The cause of ***this*** is that the ***far's*** and ***dr's we*** have defined and utilized in this sub-section, differ from the terms commonly used in Automatic Incident Detection (see chapter 4). In our ACD ***case,*** the ***far*** and the ***dr*** denote the probability that a substantially congested traffic state is properly detected, that is, that traffic conditions which cause delays which are considered to be significant for the concerning ATMIS purposes are observed. The difference between significant and insignificant delays is expressed in the utilized prior values of the historic congestion speed and historic free-flow speed and their standard deviations, viz. $\bar{\bar{v}}_c$ = 35 and $\bar{\bar{v}}_f$ = 110 km/h, and σ_c=25 and σ_f=15. Accordingly, the developed ACD algorithm will not detect disturbances in the traffic flow which only lead to small travel time increases. For certain ATMIS applications, these fluctuations are part of the regular variability in the travel time along a route and should

not lead to actions, such as distinct route advisories or dissemination of corresponding traffic information. When, after a certain time period, the effects of this traffic flow disturbance should have resulted in a significant reduction in the average speed of the traffic flow, this will be detected positively by the developed ACD algorithm. This time delay is no decisive restraint for the concerning ATMIS purposes, as the acceptable duration before deployment of these ATMIS actions is relatively large.

The ***consequence*** of this is that the first step of the advanced estimator for the road link mean speed $\bar{v}$ can be performed properly when only few probe reports are received. This is advantageous, since we have introduced this advanced estimator to overcome the unreliabilities and uncertainties induced by straightforwardly estimating the road link mean speed under similar conditions by means of the direct (Moment) estimator.

In the second step of our analysis, we will estimate the road link mean speed $\bar{v}$, based on the observed probe vehicle data $v_i, i=1,..., N_p$ *and* the result of the prevailing regime selection step. In the next sub-section a (Bayes) estimator for estimating the road link mean speed $\bar{v}$ from probe vehicle samples v_i is considered. This estimate equals the expected value of $\bar{v}$'s posterior distribution, the update of the prior distribution, given the sample of speed data. For this estimator the selection of the prevailing traffic regime is important as it uses the prior, historic knowledge about the behaviour of the road link mean speed $\bar{v}$, which is distributed with mean $\bar{\bar{v}}_f$ (f) in the case of free-flow (congestion), the variances in this road link mean speed (σ^2_f for free-flow and σ^2_c for congestion) as well as the variances in the individual probe vehicle speeds σ^2.

3.3.3 Step 2: Estimation of Road Link Mean Speed and Travel Time

In this sub-section we will formulate and analyze an estimator for estimating the road link mean speed $\bar{v}$ and the road link mean travel time $\bar{\tau}$ using historic traffic data. This estimator, $\bar{v}^*_B$, is preferred over the direct estimator presented in section 3.2, if we have reliable prior knowledge regarding the distribution of $\bar{v}$. This prior knowledge has been specified in the defined three step speed model. In the following two sub-sections we will develop such an estimator for both the free-flow and the congestion macroscopic traffic regime of this speed model. In formulating this estimator, we make use of the assumptions regarding the behaviour of $\bar{v}$ under free-flow and under congested conditions. We will see that for the free-flow (normal distributed) regime the formulae and the computations are relatively simple, while for the congested (gamma distributed) regime they are somewhat more complex. Similar derivations can be made for any distribution of $\bar{v}$, which accordingly will hold for any concretization of the functions G and *H*, which represent the historically observed relation between $\bar{v}$ and $\bar{\bar{v}}$ (G for free-flow and *H* for congestion).

3.3.3.1 Free-flow Regime

For the free-flow regime, the prior distribution of $\bar{v}$ is assumed to be normal with a certain (historic) mean $\bar{\bar{v}}_f$, and variance σ_f^2. Accordingly, its posterior distribution, given the observed speeds $v_1, \ldots v_{Np}$ equals (see also (Raiffa and Schlaiffer, 1961)):

$$\bar{v} \mid v_1,\ldots,v_{N_p} \sim N\left(\frac{\eta \bar{\bar{v}}_f + \sum_{i=1}^{N_p} v_i}{\eta + N_p}, \frac{\sigma^2}{\eta + N_p}\right),$$

with $\eta = \sigma^2/\sigma^2_f$.

Note that according to this formula the sum of the individual probe vehicle speeds and the number of received probe messages give sufficient information to estimate the road link mean speed. The (Bayes) estimator equals the expected value of $\bar{v}$, given the N_p observed individual probe vehicle speeds. Since the posterior distribution is uni-modular and symmetric, the posterior mean is the best estimator under all commonly used evaluation criteria (e.g. MSE, absolute error):

$$\bar{v}^*_B = E_f\{\bar{v} \mid v_1,\ldots,v_{N_p}\} = \frac{\eta \bar{\bar{v}}_f + \sum_{i=1}^{N_p} v_i}{\eta + N_p}$$

Note that this estimator is analogous to the direct estimator for the road link mean speed formulated in sub-section 3.2.2, but incorporates the prior knowledge about the behaviour of this road link mean speed obtained from the first step (ACD).

3.3.3.2 Congestion Regime

For the congestion regime the estimator incorporating historic data is defined assuming a gamma distribution for the individual probe vehicle speeds. Also for this non-symmetrical distribution the posterior mean is the best estimator under the mean squared error criterion. Hence:

$$\bar{v}^*_B = E_c\{\bar{v} \mid v_1,\ldots,v_{N_p}\}$$

In appendix B this is shown to be equal to

$$\left| \int_{0^+}^{\infty} f_{\frac{\bar{\bar{v}}_c}{\sigma_c^2}, \frac{\bar{\bar{v}}_c^2}{\sigma_c^2}}(\bar{v}) \prod_{i=1}^{N_p} f_{\frac{\bar{v}}{\sigma^2}, \frac{\bar{v}^2}{\sigma^2}}(v_i)\, d\bar{v} \right|^{-1} \left| \int_{0^+}^{\infty} \bar{v}\, f_{\frac{\bar{\bar{v}}_c}{\sigma_c^2}, \frac{\bar{\bar{v}}_c^2}{\sigma_c^2}}(\bar{v}) \prod_{i=1}^{N_p} f_{\frac{\bar{v}}{\sigma^2}, \frac{\bar{v}^2}{\sigma^2}}(v_i)\, d\bar{v} \right|$$

with $f_{\lambda,p}$ the probability density function of I'(λ, p). Note that according to this formula the sum and the product of the individual probe vehicle speeds as well as the number of received probe messages give sufficient information to estimate the road link mean speed.

3.3.3.3 Stability of Estimator

The stability of the estimator using historic traffic data under both free-flow and. congestion conditions can be enhanced by employing (single exponential) smoothing. In this way, the final estimation of the road link mean speed $\bar{v}$ becomes

$$\bar{v}^*(t) = y\ \bar{v}^*_B(t) + (1-y)\ \bar{v}^*_B(t-1)$$

where y is a smoothing constant with range [0, 1]. When we should want to attach considerable significance to the present estimate, y is preferably kept high (e.g. > 0.8). When we have only little confidence in the practical correctness of this present estimate, y is preferably kept low (e.g. < 0.2) so that the previous estimates more heavily act upon the conclusive value.

From the structure of the estimator for the free-flow regime it becomes apparent that the contribution of the received individual probe vehicle speeds v_i, increases with the sample size N_p and, hence, that the contribution of the historic mean speed $\bar{\bar{v}}_f$ decreases. This is easily explained: since a large value of N_p corresponds with a large sample size and thus with a relatively large fraction of the vehicles on the considered road link (or segment) whose velocities are known, the sample becomes more representative of the traffic flow. Under these conditions the stabilizing effect of the prior knowledge (the historic mean speed $\bar{\bar{v}}_f$) becomes of subordinate relevance. Accordingly, the estimator (for the free-flow regime) employs the prior knowledge mainly in situations when the number of received probe vehicle reports is small. When the number of received probe reports increases the influence of the prior knowledge. diminishes while eventually the estimator incorporating historic data converges to the direct estimator when the value of N_p becomes large. Although it is not directly shown, the same holds for the estimators under congested traffic conditions. In this way, the observed drawback of the direct estimator, i.e. its problematic performance at low N_p, is effectively compensated.

3.3.3.4 Comparative Performance of Estimators

Under all circumstances, the performance of a statistical estimator that takes into account prior knowledge about the behaviour of the variable that is to be estimated is higher than, or at least equal to, the performance of an estimator that does not use this prior knowledge, provided that this prior knowledge is reliable. We refer to standard statistical and Bayesian literature for a more detailed treatment of this issue (for instance (Berger, 1985), (Ross, 1987) and (Bemando and Smith, 1994)).

However, the direct or Moment estimator has the advantage that the procedure for estimating the road link mean speed from received probe vehicle samples is extremely simple, while the Bayes estimator using historic data is somewhat more complicated. In order to decide which estimator should be preferred, the quality of both estimators has to be compared.

Departing from the multi-layered speed model that we have designed in sub-section 3.3.1, the individual probe vehicle speeds and the road link mean speed in the case of free-flow traffic are distributed with

$$v_i \sim N(\bar{v}, \sigma^2) \text{ and } \bar{v} \sim N(\bar{\bar{v}}_f, \sigma_f^2).$$

The direct (Moment) estimator for the road link mean speed from the received probe vehicle reports that has been formulated in sub-section 3.2.2 is expressed by

$$\bar{v}_M^* = \frac{1}{N_p}\sum_{i=1}^{N_p} v_i \sim N(\bar{v}, \frac{\sigma^2}{N_p})$$

Since $E_f\{\bar{v}_M^* - \bar{v}\} = 0$ the Moment estimator is unbiased, while its variance is $Var\{\bar{v}_M^*\} = \frac{\sigma^2}{N_p}$

So, the mean squared error of the direct (Moment) estimator $MSE(\bar{v}_M^*)$ is defined by

$$E_f\{(\bar{v}_M^* - \bar{v})^2\} = (E_f\{\bar{v}_M^* - \bar{v}\})^2 + Var\{\bar{v}_M^*\} = 0 + \frac{\sigma^2}{N_p} = \frac{\sigma^2}{N_p}$$

For the more advanced (Bayes) estimator using historic data, that takes full advantage of **all** information comprised in he multi-layered speed model, the posterior distribution of the road link mean speed is expressed by (**with** $\eta = \sigma^2/\sigma_f^2$)

$$\bar{v} \mid v_1,\ldots,v_{N_p} \sim N(\frac{\eta\bar{\bar{v}}_f + \sum_{i=1}^{N_p} v_i}{\eta + N_p}, \frac{\sigma^2}{\eta + N_p}),$$

For the Bayes estimator, its formula and probability distribution are defined by:

$$\bar{v}_B^* = E_f\{\bar{v} \mid v_1,\ldots,v_{N_p}\} = \frac{\eta\bar{\bar{v}}_f + \sum_{i=1}^{N_p} v_i}{\eta + N_p} \sim N(\frac{\eta\bar{\bar{v}}_f + N_p\bar{v}}{\eta + N_p}, \frac{N_p\sigma^2}{(\eta + N_p)^2})$$

Since

$$E_f\{\bar{v}_B^*\ (-3)\} = E_f\{\frac{\eta\bar{\bar{v}}_f + N_p\bar{v}}{\eta + N_p} - \bar{v}\} = \frac{\eta\bar{\bar{v}}_f + N_p\bar{\bar{v}}_f}{\eta + N_p} - \bar{\bar{v}}_f = 0$$

also this estimator is unbiased, while its variance is

$$Var\{\bar{v}_B^*\} = \frac{N_p\sigma^2}{(\eta + N_p)^2}$$

So, the mean squared error of this estimator $MSE(\bar{v}_B^*)$ is defined by

$$E_f\{(\bar{v}_B^* - \bar{v})^2\} = (E_f\{\bar{v}_B^* - \bar{v}\})^2 + Var\{\bar{v}_B^*\} = 0 + \frac{N_p\sigma^2}{(\eta + N_p)^2} = \frac{N_p\sigma^2}{(\eta + N_p)^2}$$

Having found the above characteristics, both estimators can be easily compared. Since both estimators are unbiased, the mean squared error (MSE) criterium equals the estimator's variance in both cases. The following formula shows that the (Bayes) estimator using historic data outperforms the direct (Moment) estimator for any value of the number of successfully received probe messages (N_p) and η:

$$\frac{MSE(\bar{v}_M^*)}{MSE(\bar{v}_B^*)} = \frac{\sigma^2}{N_p}\frac{(\eta + N_p)^2}{N_p\ \sigma^2} = (\frac{\eta}{N_{\bar{P}}} + 1)^2 > 1$$

Note that the MSE ratio converges to unity as N_p grows to infinity. This corresponds with the earlier remark that the estimators themselves converge to each other as N_p increases. Also, the formula above indicates that the extent to which the Bayes estimator using historic data outperforms the Moment estimator decreases with decreasing η, which in turn decreases with an increase in the flow rate q (see next chapter).

When we substitute the values of these variables that have used throughout this chapter, σ=*15 km/h* and σ_f=*15 km/h*, than we find that for N_p=1, 2 and 3, the MSE of the Moment estimate is 4, 2.25 and 1.78 times the MSE of the Bayes estimate. Ergo, the slight increase in **complexity** due to incorporating historic data is indisputably worthwhile, seeing that this considerably increases the accuracy and the reliability of the ultimate estimate of the road link mean speed, in particular for small N_p. For the congestion traffic regime similar derivations can be obtained which lead to analogous conclusions.

We will further analyze the reliability of this (Bayes) estimator incorporating historic traffic data. A common approach to give an indication of the reliability or accuracy of the estimate, is to present a confidence set for $\bar{v}$. The Bayesian analog of a classical confidence set is called a ***credible set*** and is fully based on $\bar{v}$'s posterior distribution (see for instance (Bemando and

Smith, 1994) and (Ross, 1987)). This is in contrast to classical confidence procedures, where the set is based on the probability distribution of the estimator, dependent on the unknown parameter. Here we also deviate from the conventional way of determining an interval with a prechosen coverage probability, by calculating the coverage probability of a prechosen interval instead. The interval, symmetrical around the estimate, is chosen given an accuracy parameter α, such that its coverage probability equals the probability that the actual road link mean speed deviates no more than a fraction a from its estimate.

The formula for the credible set can be derived easily from the posterior distribution of $\bar{v}$ (note that in this formula $\bar{v}_B^*$ is an estimate and not an estimator):

$$P_f\{ \mid \bar{v} - \bar{v}_B^* \mid \leq \alpha \bar{v}_B^* \mid v_1,\ldots,v_{N_p} \} =$$

$$= P_f\{ \mid \frac{\bar{v} - \bar{v}_B^*}{\sigma / \sqrt{\eta + N_p}} \mid \leq \frac{\alpha \bar{v}_B^*}{\sigma / \sqrt{\eta + N_p}} \mid v_1,\ldots,v_{N_p} \} = 2\,\Phi_{0,1}\left(\frac{\alpha \bar{v}_B^*}{\sigma / \sqrt{\eta + N_p}}\right) - 1$$

3.3.3.5 **Results**

The performance of the estimator using historic data has been computed with respect to its accuracy and reliability. These computations have been achieved by implementing the presented credible sets for several numbers $N_p \in \{1,2,3,5,10\}$ and values of probe vehicle samples. The probability mass of the estimator (under free-flow conditions) contained in the credible sets is shown in table 3.4. We have assumed the standard deviation of the individual probe vehicle speeds to amount to $\sigma=15$ km/h, for the prior distribution of $\bar{v}$ we take $\bar{\bar{v}}_f=110$ km/h and $\sigma_f=15$ and use different values for the involved number of successfully received probe vehicle speeds N_p and the sample mean speed. The required quality of the speed estimations (the level of accuracy a) has been taken 10% ($\alpha=0.10$) and 15% ($\alpha=0.15$).

The results in the table show that, for example, if the number of probe messages received from a particular road link in a particular time period is 2 ($N_p = 2$) and the sample mean of these messages is 95 km/h, the estimation of $\bar{v}$ equals 100 km/h and the probability that $\bar{v}$ 'lies in [85, 115] (i.e. α=0.15) is 91.67%.

For the congestion traffic regime we have performed similar computations by implementing the presented credible sets for congested traffic for $N_p \in \{1,2,3,5,10\}$. The probability mass contained in the credible sets for the estimator under congested conditions is shown in table 3.5. We have assumed the standard deviation of the individual probe vehicle speeds to amount to $\sigma=25$ km/h, for the prior distribution of $\bar{v}$ we take $\bar{\bar{v}}_f=35$ km/h and $\sigma_f=25$ and use different values for the involved number of successfully received probe vehicle speeds N_p and the sample mean speed. Again, the required quality of the speed estimations (the level of accuracy a) has been taken 10% ($\alpha=0.10$) and 15% ($\alpha=0.15$).

		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 95$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 110$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 125$	
$N_p = 1$	$\bar{v}^*_B$	102.50		110		117.50	
(α=0.10)	(α=0.15)	66.61%	85.28%	70.03%	88.02%	73.21%	90.34%
$N_p = 2$	$\bar{v}^*_B$	100		110		120	
(α=0.10)	(α=0.15)	75.18%	91.67%	79.60%	94.33%	83.41%	96.23%
$N_p = 3$	$\bar{v}^*_B$	98.75		110		121.25	
(α=0.10)	(α=0.15)	81.21%	95.17%	85.75%	97.22%	89.41%	98.47%
$N_p = 5$	$\bar{v}^*_B$	97.50		110		122.50	
(α=0.10)	(α=0.15)	88.87%	98.31%	92.76%	99.29%	95.45%	99.73%
$N_p = 10$	$\bar{v}^*_B$	96.36		110		123.64	
(α=0.10)	(α=0.15)	96.69%	99.86%	98.50%	99.97%	99.37%	100.00%

Table 3.4 Credible sets for road link mean speed estimations for free-flow traffic

The results in the table show that, for example, if the number of probe messages received from a particular road link in a particular time period is 3 ($N_p = 3$) and the sample mean of these messages is 25 km/h, the Bayes estimation of $\bar{v}$ equals 33.78 km/h and the probability that $\bar{v}$ lies in [28.71, 38.841 (i.e. α=0.15) is 67.78%. Note that in table 3.5 the reliability of the estimate decreases with increasing sample mean, while in table 3.4 the reliability of the estimate increases with increasing sample mean. This is obvious since table 3.5 concerns congestion estimations, which can be performed more reliably when the speeds of the probe vehicle are low. Furthermore, note that the reliabilities in table 3.5 are lower than those in table 3.4. This is provoked by the difference in the behaviour of the speeds of the traffic under congested and free-flow conditions. When the traffic is congested, the range of possible speeds is large compared to free-flowing traffic. This produces an elongated shape of the congested speed distribution (represented by a Gamma distribution in this particular example).

Two further remarks need to be made with respect to the results in tables 3.4 and 3.5.

Firstly, note that the results concern one single estimate only, while in practice estimations will be performed continuously and each new estimate incorporates the previous estimates in its prior knowledge. That means that the results for the estimator using historic traffic data in the table are comparatively more dependent on the fixed prior assumptions of the historic mean speeds $\bar{\bar{v}}_f$, in particular when a small sample size (low N_p) is concerned. In a practical organization this prior knowledge will be outweighed by the previous estimation(s) resulting from the preceding time step(s), so that under these conditions the downright correctness of the historic prior knowledge will become of subordinate importance.

		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 25$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 35$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 45$	
$N_p = 1$	$\bar{v}^*_B$	34.85		42.69		50.23	
(α=0.10)	(α=0.15)	38.98%	48.19%	38.75%	47.85%	38.21%	47.03%
$N_p = 2$	$\bar{v}^*_B$	34.21		41.52		48.55	
(α=0.10)	(α=0.15)	54.11%	56.61%	53.18%	55.12%	49.84%	53.67%
$N_p = 3$	$\bar{v}^*_B$	33.78		40.15		46.39	
(α=0.10)	(α=0.15)	65.98%	67.78%	64.03%	67.01%	62.75%	65.15%
$N_p = 5$	$\bar{v}^*_B$	33.06		40.09		46.18	
(α=0.10)	(α=0.15)	74.32%	77.03%	72.34%	73.20%	67.41%	69.87%
$N_p = 10$	$\bar{v}^*_B$	32.89		37.80		45.91	
(α=0.10)	(α=0.15)	78.42%	84.53%	77.14%	78.19%	73.41%	73.90%

Table 3.5 Credible sets for road link mean speed estimations for congested traffic

Secondly, note that the obtained percentages for the credible sets presume that the used prior knowledge is fully reliable. However, this prior knowledge has been chosen by means of the Automatic Congestion Detection (ACD) procedure, which is characterized by a certain false alarm rate and a certain detection rate. This means that the ultimate size of the credible set in the second step of the estimator is a weighted mean of the credible free-flow set (see table 3.4) and the credible congestion set (see table 3.5), as follows[10]

Credible Set = β {Free-flow Credible Set) + (1-β) {Congested Credible Set}

with β the probability of free-flow given the received probe vehicle samples.

3.3.3.6 Conclusions

[10] More formally, the ultimate credible set is given by

$$P\{cf \mid v_1,\ldots,v_{N_p}\}\, P_c\{\,|\bar{v}-\bar{v}^*_B| \le \alpha \bar{v}^*_B \mid v_1,\ldots,v_{N_p}\} + P\{ff \mid v_1,\ldots,v_{N_p}\}\, P_f\{\,|\bar{v}-\bar{v}^*_B| \le \alpha \bar{v}^*_B \mid v_1,\ldots,v_{N_p}\}$$

where $P\{ff \mid v_1,\ldots,v_{N_p}\} = \dfrac{P_f\{v_1,\ldots,v_{N_p}\}\, P\{ff\}}{P_f\{v_1,\ldots,v_{N_p}\}\, P\{ff\} + P_c\{v_1,\ldots,v_{N_p}\}\, P\{cf\}}$

with $P_f\{v_1,\ldots,v_{N_p}\} = \int_{-\infty}^{\infty} P_f\{v_1,\ldots,v_{N_p} \mid \bar{v}=a\}\, P_f\{\bar{v}=a\}\, da = \int_{-\infty}^{\infty} \left(\prod_{i=1}^{N_p} \phi_{a,\sigma^2}(v_i)\right) \phi_{v_f,\sigma_f^2}(a)\, da$

and $P\{ff\}$ ($P\{cf\}$) the probability of free-flow (congestion), which is equal to the fraction of the time that the traffic on the concerning road link is in state of free-flow (congestion).

In this sub-section we have developed and analyzed a procedure for estimating the road link mean speed, making use of prior knowledge regarding the (historic) behaviour of (the speed of) the traffic flow. This prior knowledge stems from the Automatic Congestion Detection (ACD) routine, developed in the previous sub-section, which leads to an initial assessment of the prevailing speed of the traffic flow, hence the free-flow mean $\bar{\bar{v}}_f$ or the congestion mean $\bar{\bar{v}}_c$. As a result, the value of the received probe vehicle speeds is less prominent in this estimator than in the direct or Moment estimator, that was described in section 3.2.

The (Bayes) estimator incorporating historic traffic data should be preferred over the more simple direct (Moment) estimator when only few probe reports are received. It has been shown that under this condition the reliability of the first estimator considerably outperforms the reliability of the latter. When more probe vehicles are present on a road link and a larger number of probe reports can be used by the estimators, the performance of both estimators has been shown to converge. This is easily explained since, under this condition, it is permissible to assume that the probe samples accurately represent the actual traffic flow, while the prior historic knowledge exerts only relatively little influence on the Bayes estimate when the number of samples is large.

In conclusion we state that the estimator incorporating historic traffic data should indisputably be preferred for estimating the road link mean speed from probe vehicle samples, at the expense of a slight increase in complexity. In addition, it is important to note that it should only be applied when the prior information is considered to be correct. If this is not the case, the estimator will drag the resulting estimate to a possibly incorrect value. For this reason, we have suggested to use an aggregate of both the historically determined mean value ($\bar{\bar{v}}_f$ or $\bar{\bar{v}}_c$) and the previously estimated value of $\bar{v}$ as the conjoint prior information.

3.4 Discussion

In this chapter we have addressed utilization of probe vehicles for ATMIS. We have‘established that accurate real-time traffic flows and traffic densities can not be reliably obtained exclusively from probe vehicles. However, both parameters are appropriately obtained from infrastructure based traffic detectors (see chapter 2). For obtaining real-time speeds and travel times from probe vehicle samples we have first of all studied a simple procedure, which however appeared to be rather arduous when only few probe vehicle messages are successfully received. Subsequently, we have designed a more advanced procedure for estimating road link mean speed and travel time. This approach takes into account historic traffic data and is based on a multi-layered speed model, which mutually relates the individual probe vehicle speeds, the actual road link mean speed and the historic mean speed, and consists of two steps. In the first step, Automatic Congestion Detection (ACD), the prevailing traffic regime (i.e. free-flow or congested) is selected. The outcome of this first step is used in the second step to estimate the actual road link mean speed. We have shown that when few probe vehicle messages are available this advanced estimator considerably outperforms the simple estimator. When more probe reports are available, the performance of both estimators has been shown to converge.

The issue of what probe vehicle penetration grade would be required for accurately and reliably detecting congestion and obtaining real-time speeds and travel times (Boyce, Hicks and Sen, 1991), depends on the road network (type and shape) and its traffic load (peak hour traffic or off-peak traffic), the communication infrastructure (transmission medium and protocol), the exact contents of the probe reports (what is transmitted by the probes and when) and the operating procedures for deducing relevant information from received probe data. In the foregoing research, including the previous PATH MOU 107 research project, we have addressed each of these topics. Based on the established results, we conclude that for obtaining real-time travel times (or mean speeds) on a free-way with an accuracy of about 85% and a reliability of about 85%, approximately 3-5 probe reports per time period are required, provided that a database with proper historic traffic data is available. The demands for reliably performing automatic congestion detection are lower and call for 1 to 2 probe reports per time period that are successfully received. From the study into communication aspects of probe vehicle data collection by means of the PROMOT computer model reported in the PATH MOU 107 research report, we learn that these requirements correspond to a probe vehicle penetration grade of about 1% when time periods of circa 10-15 minutes are concerned and a penetration grade of about 5% when time periods of l-5 minute are required. These figures depart from a freeway network during peak hours. Lower classes of roads or times of the day outside peak hours impose higher probe vehicle penetration grades.

4. Integration of Infrastructure Based and Non-Infrastructure Based Traffic Detectors

The principal objective of the research presented in this chapter is to establish techniques for fusing data from both sources, in other words, to focus on new theoretical research of how data delivered from the mobile probe vehicles concerning spatial fluctuations in the traffic process may be included to enhance the reliability and accuracy of estimates from data measured by traditional infrastructure based traffic detectors concerning temporal fluctuations in the traffic process (Cremer, 1991; TRANSCOM, 1993; Philipps, Hoops, Ibbeken and Riegelhuth, 1994).

When data obtained through non-infrastructure based probe vehicles is included into the data measured by infrastructure based induction loop detectors, probe vehicles can:
- enhance the quality of information obtained through infrastructure based detectors,
- accurately determine the exact location of a disturbance in the traffic flow,
- quicken the process of observing irregularities in the traffic process, and
- provide updates in order to adjust the measuring errors of infrastructure based detectors.

In chapter 2 we have addressed the issue of utilization of infrastructure based traffic detectors for obtaining real-time road traffic information for ATMIS purposes. Since this type of traffic detectors is specifically designed for measuring cross-section data, i.e. registering all vehicles that pass the detector, the flow q and (to a certain extent) the density k are directly obtained. Also in that chapter, we have developed an algorithm for estimating real-time travel times from induction loop data. From the calibration, it became clear that under certain conditions it is rather arduous to base the estimate of the actual travel time over a rather long stretch of freeway (typically 5 to 10 kilometers) exclusively on fixed detector data, installed at the boundaries of this stretch. For this reason, in *sub-section 4.1.1* we will incorporate probe vehicle data into the COMETT computer model. In chapter 3 of this research report we have developed an algorithm for performing Automatic Congestion Detection and extracting road link mean speeds and travel times from only probe vehicle data. In *sub-section* 4.1.2 we will extend this algorithm by incorporating data from fixed detectors. With respect to (automatic) detection of disturbances in the traffic flow, numerous research has been conducted *in* the past. *Section 4.2* gives an overview of the significance of adequate Incident Management for ATMIS as well as of proficiencies of existing algorithms for automatically detecting incidents. The general conclusion of these studies is that appropriate Automatic Incident Detection, using infrastructure based traffic detectors with large inter detector spacings (i.e. several kilometers) is virtually' unfeasible. Hence, in *section* 4.3, we will present principles of Automatic Incident Detection using combined data from infrastructure based and non-infrastructure based traffic detectors. Conclusions and references to this chapter are given in *section 4.4* and *section 4.5.*

4.1 Estimation of Road Link Mean Speed and Travel Time

In this section we will address the issue of estimating the real-time (mean) speed and travel time of the traffic flow using a combination of probe vehicle and induction loop data. In sub-section

4.1.1 we will depart from induction loop data and analyze how probe vehicle data can be included to enhance the estimations, using the algorithm for only induction loop data that has been described in chapter 2. In sub-section 4.1.2 we will depart from probe vehicle data and analyze how induction loop data can be included, using the algorithm for only probe vehicle data that has been described in chapter 3.

4.1.1 Utilization of Induction Loop Data and Additional Probe Vehicle Data

4.1.1.1 Problem Definition

The reliability and the accuracy of the real-time travel times that have been estimated by means of the COMETT computer model developed in chapter 2, have been shown to be susceptible to (stochastic) errors in the data measured by infrastructure based traffic detectors. In particular under conditions of small flow rates as well as in the case that congestion persists for a long period of time, the computed travel times become increasingly inaccurate. To a certain extent, these errors are compensated by a re-calibration process, but additional re-calibrations would be required in order to achieve a more persistent performance. We will extend the COMETT algorithm to incorporate probe vehicle data in order to achieve the required re-calibrations (Westerman, 1995).

4.1.1.2 Problem Assessment

According to the re-calibration process that has been described in more detail in sub-section 2.2, certain so-called 'starting points in time', t_{x1} and t_{x2}, are repeatedly redefined using a 'reference' travel time between the boundaries of the road link under consideration, τ_{cor}. The determination of this reference travel time τ_{cor} is based on correlation of (meso-level) fluctuations in the flow rates at both measuring sites, viz. measuring site x_1 and measuring site x_2, according to

$$\tau_{cor}(t) = t - t''$$

in which t is the point in time at which the travel time determination takes place, and t'' is the point in time at which the characteristic flow pattern observed at measuring site x_2 at t, was observed at the preceding measuring site x_1 (in sub-section 2.2 this re-calibration process in described in more detail). Finally, the starting points in time are redefined or re-calibrated using

$$t_{x1} = t$$
$$t_{x2} = t''$$

The principal contribution of the re-calibration process is to compensate for measuring errors and to prevent the algorithm from digressing, and is based on the availability of a reference travel time. In the case this reference travel time complies with certain check-criteria, its correctness is confirmed and a re-calibration is allowed to take place. When we assume that the elementary road traffic data for ATMIS applications is predominantly being obtained from infrastructure based traffic detectors, the process for estimating real-time travel times can easily be enhanced by incorporating probe vehicles, since the reference travel time can obviously also be obtained

from present probe vehicles.

From the traffic messages that are received from the probe vehicles, a road link mean speed $\bar{v}(t)$ at point in time t is computed. For this, we use a method that is analogous to the method that has been developed in the previous chapter for the 'probe only' case, but now, we also incorporate the actual value of the flow rate $q(t)$ into this computation. This flow data is directly obtained from the infrastructure based detectors. A description of exactly how this estimation of the road link mean speed using also flow data takes place, is deferred to section 4.2.

Hence, the starting points in time t_{x1} and t_{x2} are computed from

$$t_{x2} - t_{x1} = \bar{v}(t_{x2})$$

in which t_{x2} is the point in time at which the travel time determination takes place, and t_{x1} is the corresponding point in time of measuring site x,. The remainder of the travel time estimation procedure is equal to the COMETT algorithm described in section 2.2.

For using the probe vehicle data to perform additional re-calibrations, it is important to ascertain that the road link mean speed obtained from the probe vehicle samples is correct. First of all, we have shown in the previous chapter that when we use the Bayes estimator to compute the road link mean speed, the accuracy and the reliability of this estimate is rather large. Furthermore, the correctness of the estimated road link mean speed can be confirmed by comparing this estimate to the travel time obtained from the fixed detectors. Since frequent re-calibrations can be performed when employing the probe determined road link mean speeds, it is allowed to have this fixed travel time serve as a representative reference travel time obtained from an independent data source.

Ergo, we can employ the probe vehicle data to achieve additional re-calibrations, independent of whether there is congestion, in order to enhance the reliability and accuracy of the ATMIS travel time estimates. When we would expect a rather large probe vehicle penetration grade, the original re-calibration process, using cross-correlation techniques, could even be abolished. This would make the ultimate procedure as well as the hard- and software requirements for obtaining real-time travel times for ATMIS purposes extremely simple.

4.1.13 Results

In order to examine how the enhancement of the combined travel time determination algorithm that has been described in the previous sub-section contributes to a higher level of reliability and accuracy of the resulting travel times, we will use the GERDIEN[12] field measurements that

[12] GERDIEN (General European Road Data Information Exchange Network) is a DRIVE II project for which small scale field measurements have been performed which provided induction loop data of volume and speed (aggregated over 1 minute) as well as accomplished travel times over the same part of the road network.

have also been used for calibrating the COMETT model outcomes in sub-section 2.4. Since no factual probe vehicle data are available, we will generate these probe vehicle updates by randomization of the field measurements.

We will presume that each time period of 5 minutes, a road link mean speed estimate becomes available from probe vehicle samples. Using the layered speed model presented in the previous chapter, the probe vehicle updates are given by

$$v \sim N(\bar{\bar{v}}_f, \sigma_f^2)$$

for free-flow traffic, with $\bar{\bar{v}}_f$ = 120 *km/h* and σ_f = *15 km/h*,

and by $$\bar{v} \sim \Gamma(\lambda_c, \rho_c) \quad such\ that\ \frac{\rho_c}{\lambda_c} = \bar{\bar{v}}_c\ and\ \frac{\rho_c}{\lambda_c^2} = \sigma_c^2$$

for congested traffic, with $\bar{\bar{v}}_c$ = *(extrapolated) field measurements* and a, = 25 *km/h.*

In figure 4.1, the travel times computed by means of the enhanced COMETT algorithm, using regular induction loop data as well as the generated probe vehicle updates (one every 5 minutes), are depicted as a function of time (ranging from 00:00 to 23:59) for the stretch of road from Moordrecht (measuring site x,) to Zevenhuizen (measuring site x_2) at April 27, 1994. For a (graphical) description of this test-site we refer to chapter 2, figure 2.9. What is free-flow and what is congestion in figure 4.1 (used for the exact distribution of the road link mean speed as indicated in the above) is determined by the GERDIEN reference congestion indicator, described in (Van Arem, et. al., 1994). Furthermore, the solid circles in figure 4.1 represent the authentic GERDIEN field measurements.

From figure 4.1 we see that the advanced COMETT computed travel times are respectably accurate (the discrepancy between 'reality', i.e. the authentic GERDIEN data denoted by the solid circles, and the COMETT outcomes during the evening peak hour are ascribed to the fact that COMETT only produces integer travel time values in unities of 1 minute). Moreover, the advanced COMETT computed travel times are persistent under congestion conditions. Were the travel times determined from only induction loop data (see figure 2.14 in chapter 2) increasingly deteriorated when congestion lasted for a rather long period of time, the travel times determined using both data sources (infrastructure based induction loop detectors and non-infrastructure based probe vehicles) remain correct. This means that incorporating traffic data obtained from probe vehicles can significantly enhance the reliability and the accuracy of real-time travel times that are required for ATMIS purposes. In this respect, we would like to emphasize that, since no factual probe vehicle data could be used, the generated probe vehicle updates are only used for the purpose of illustrating the combined travel time determination method.

For a more extensive description of GERDIEN we refer to chapter 3. sub-section 3.3.4, or to (Van Arem, et. al., 1994).

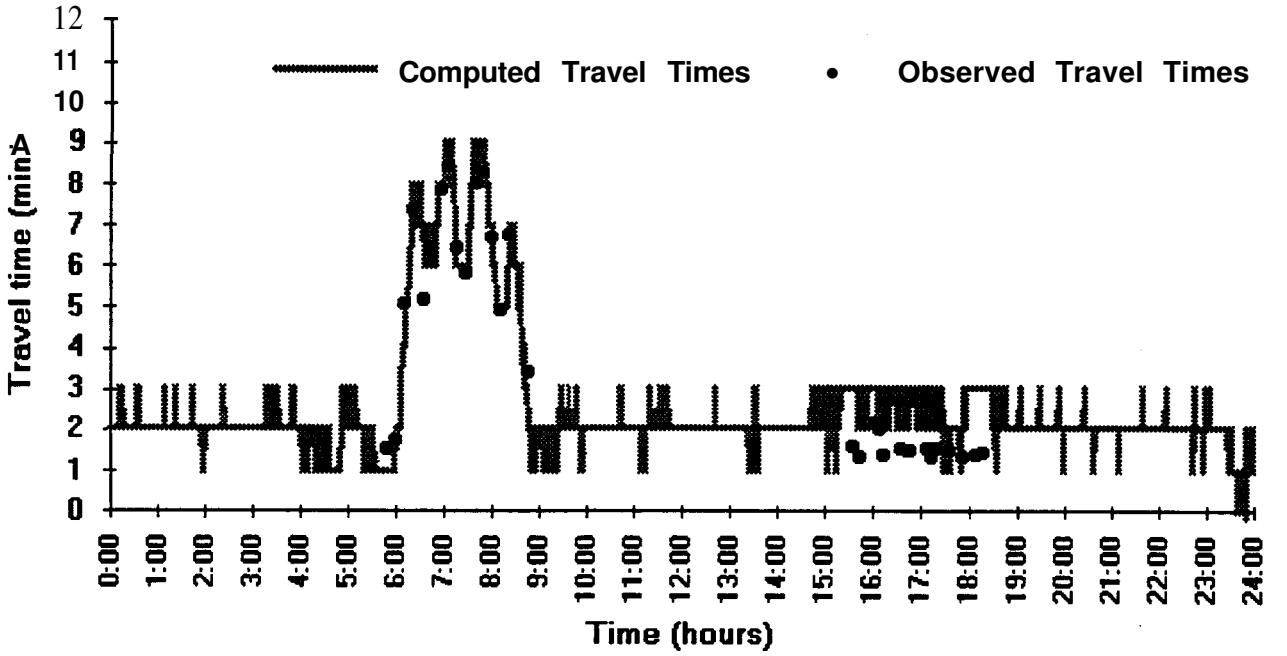

Figure 4.1 Travel times from the enhanced COMETT algorithm (using regular induction loop data and generated probe vehicle updates) compared with field observations

4.1.2 Utilization of Probe Vehicle Data and Additional Induction Loop Data

4.1.2.1 Problem Definition

The methodology that we have developed in the previous chapter to estimate the road link mean speed from received probe vehicle samples for ATMIS purposes was composed of two steps, namely *Prevailing Regime Selection,* also referred to as *Automatic Congestion Detection,* (step 1) and *Estimation of Road Link Mean Speed and Travel Time* (step 2). In chapter 3 'only probe vehicle data was assumed to be available, so these procedures were based on probe vehicle data only. In this chapter, also data from infrastructure based induction loop detectors is available. This means that the procedures that were developed in chapter 3 are to be extended to incorporate induction loop data. This brings about several complexities since probe data is received continuously, concerns a longitudinal section of a road link and concerns measurements from a small fraction of all vehicles in the traffic flow. Conversely, induction loop data is obtained once every time interval (e.g. 1 or 5 minutes), concerns a cross-section of .a road link and concerns measurements from all vehicles in the traffic flow.

In the ensuing sub-sections we will study how data from both dissimilar data sources can be fused in order to determine the real-time road link mean speed v or real-time travel timer more accurately and reliable than has been achieved by using only induction loop data (in

chapter 2) or only probe vehicle data (chapter 3).

4.1.2.2 Step 1: Prevailing Regime Selection

The procedure for Automatic Congestion Detection (ACD) using a combination of probe vehicle and induction loop data is analogous to the procedure using only probe vehicle data described in chapter 3, except for the issue of prior knowledge. When data from both data sources is fused, the prior knowledge can be enhanced by using measured local-related quantities, which are available through induction loop measurements. More specifically, we incorporate flow data in order to model the historic mean speeds and their related variances for both the free-flow and the congested traffic flows on the considered road link and for the actual time period. So, instead of assuming the historic mean speeds and their variances to be constants we assume them to depend on the actual flow rate *q*. It goes without saying that this will make the. underlying model as well as he resulting estimates more realistic. This refinement requires a modification of the macro-level of the multi-layered speed model that was defined in the previous chapter. The meso- and micro-level of this speed model remain unmodified.

By incorporating the actual flow rate measured by the induction loop detectors, the actual road link mean speed v becomes a random variable that is distributed around a historic mean $\bar{\bar{v}}(q)$. Distinguishing between a free-flow and a congested regime, this macro-level is defined according to

$$\bar{v}_f(q) \sim G(\bar{\bar{v}}_f(q), \sigma_f^2(q))$$

for free-flow traffic, and

$$\bar{v}_c(q) \sim H(\bar{\bar{v}}_c(q), \sigma_c^2(q))$$

for congested traffic, with G and ***H*** certain distribution functions.

Analogous to the conditions discussed in the previous chapter, we assume that the values of $\bar{\bar{v}}_{c,f}(q)$ and $\sigma_{c,f}^2(q)$ are empirically known, for instance kept in a historic database in the traffic center. So, the distributions G and ***H*** are gradually obtained. In addition, also the variance in the road link mean speed will depend on the actual traffic volume. Also this relation is assumed to be empirically obtained.

Furthermore, the road link mean speed estimations in the previous time period(s) can be incorporated into this prior knowledge. Hence, the historic mean speed, F(t), to be used for the estimate at time instant ***t*** is given by

$$\bar{\bar{v}}_{c,f}(t) = \gamma\, \bar{\bar{v}}_{c,f}(q) + (1-\gamma)\, \bar{v}^*(t-1)$$

where γ is a smoothing constant with range [0,1]. This smoothing process has also been applied for probe vehicle data only in the previous chapter.

The multi-layered micro/meso/macro speed model to be used when combined probe vehicle and induction loop data is available for ATMIS purposes is sketched in figure 4.2[13]. Note that when only probe vehicle would be available (as has been assumed in chapter 3) the prior knowledge and its accuracy range are modelled as constants, while in figure 4.2 the prior knowledge is a function of the actual flow rate *q* and so is adjusted to the actual traffic conditions.

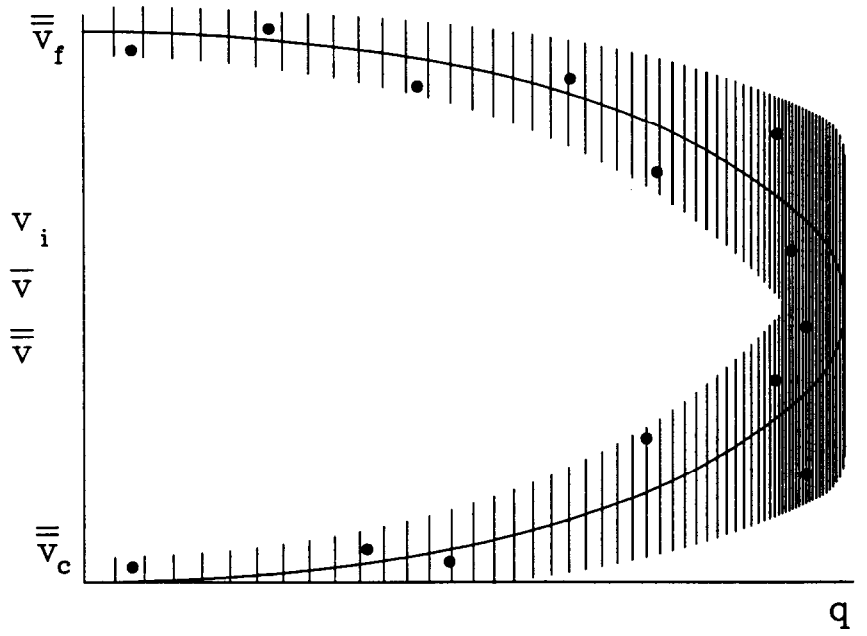

*Figure 4.2 Multi-layered speed model incorporating measured **flow** rates **for** estimating road link mean speed from probe vehicle observations **for ATMS***

Figure 4.3 illustrates the influence of incorporating measured flow data on the shape of the probability density function of the road link mean speed. In this figure three cross-cuts of the hybrid traffic flow model in figure 4.2 are denoted, at distinct values of the traffic volume *q:* a low, a medium and a high value of *q,* respectively. Figure 4.3 shows that the probability density functions corresponding to each of these cross-cuts are distributed with two local optima, but with considerably different shapes for each cross-cut. This means that selecting the prevailing traffic regime (i.e. performing ACD) for ATMIS not **only** depends on (the sample mean of) the speeds comprised in the received probe vehicle reports, but also to a large extent on the value of the actual flow rate.

Based on the sample mean of the received individual probe vehicle speeds and the involved probability distribution function of the road link mean speed as sketched in figure 4.3, we

[13] We refer to sub-section 8.1.2.3 for the exact parameters that are used for drawing up figure 8.2 and 8.3.

determine how justified it is to hold on to the currently selected regime.

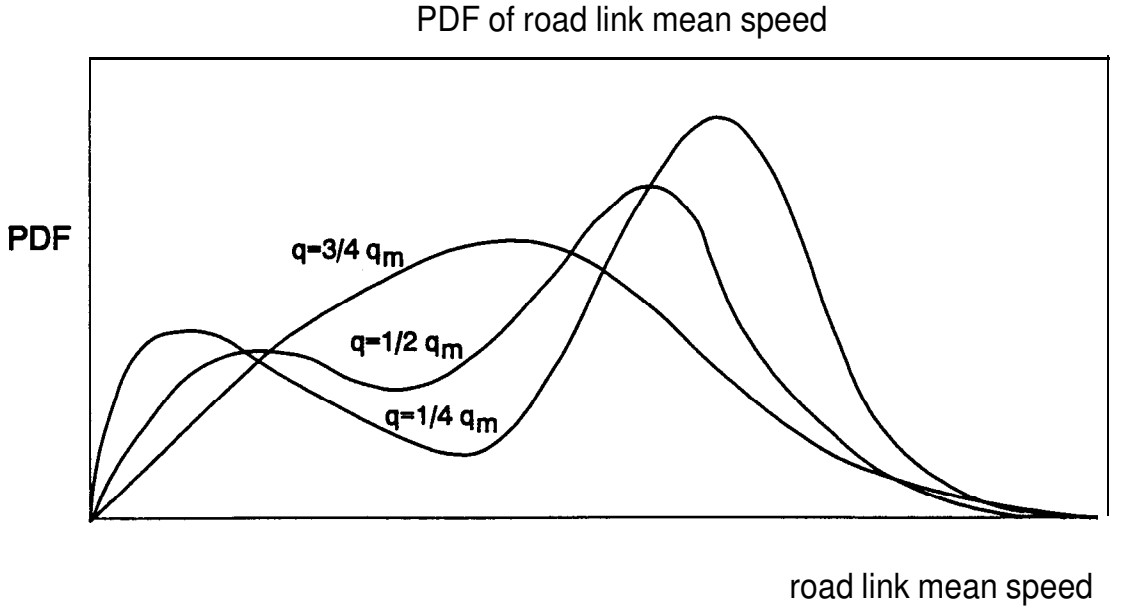

Figure 4.3 Probability-density functions of the road link mean for ATMS using different flow rates

For instance, for the currently selected traffic regime being free-flow, we have shown in the previous chapter that the ACD procedure departs from the following statistical test:

$$H_0 : \textit{free-flow traffic}$$
$$H_1 : \textit{congested traffic}$$

in which the null hypothesis will be rejected when the sample mean speed is contained in a certain critical region denoted by $[0,K)$, viz.

$$\frac{1}{N_p} \sum_{i=1}^{N_p} v_i \in [0,K)$$

A similar approach is followed when the currently selected traffic regime is congestion (i.e. H_0: *congested traffic).* We refer to the previous chapter for a more detailed description of this basic ACD procedure.

In order to further analyze the performance of the combined probe vehicle induction loop ACD-algorithm, without having available an extensive database with historic speed data and its variance relative to flow data, we will provisionally model the relation between the historic

mean speed $\bar{\bar{v}}$ and the actual traffic volume q (measured by the induction loop detectors) using the linear model of Greenshields (Greenshields, 1935). We realize that this model does not exactly represent reality (particularly for very low values of the flow q this model tends to overrate the value of the speed $\bar{\bar{v}}_{c,f}(q)$), but still adopt this particular model for its simplicity.

Hence, for our analyses we will model the historic mean speed $\bar{\bar{v}}(q)$ as a function of the traffic flow q measured by the induction loop detectors, using

$$q(\bar{\bar{v}}) = \bar{\bar{v}} k_j - \frac{k_j}{v_m} \bar{\bar{v}}^2$$

with v_m the maximum speed, the speed that exists under nearly vacant flow conditions, and k_j *the* jam density, the density that occurs when both flow and speed approach zero.

This relation[14] leads to the following values for $\bar{\bar{v}}_{c\ (q)}$ and $\bar{\bar{v}}_f(q)$:

$$\bar{\bar{v}}_c(q) = \frac{1}{2} v_m \left(1 - \sqrt{1 - \frac{4q}{v_m k_j}}\right) \quad \text{and} \quad \bar{\bar{v}}_f(q) = \frac{1}{2} v_m \left(1 + \sqrt{1 - \frac{4q}{v_m k_j}}\right)$$

In addition, we will provisionally model the relation between the variance in the road link mean speed and the actual traffic volume q by assuming a linear increase in standard deviations from $q=O$ to $q=q_m$ (q_m is the maximum flow or capacity) and by assuming that for $\bar{v}$ variances in congested traffic situations are larger than in cases of free-flow. We introduce the following formulae for $\sigma_c(q)$ and $\sigma_f(q)$ respectively

$$a,(q) = A_c \frac{q}{q_m} + B_c \text{ and } \quad \sigma_f(q) = A_f \frac{q}{q_m} + B_f$$

In these formulae, $A, _f + B_{c,f}$ comprehends the maximum standard deviation, which is actual at maximum flow conditions, with $B_{c,f}$ the fixed part and $A, _f$ the part actually depending on the current traffic *flow q.* Note that equality $A_c+B_c=A_f+B_f$ has to be satisfied, since this is the value of both $o,(q)$ and $\sigma_f(q)$ at $q=q_m$.

We have used the above assumptions for drawing up the speed-flow ($\bar{\bar{v}}q$) model as illustrated figure 4.2. We have used parameter values of $q_m=6600\ veh/h, k_j=750\ veh/km$ and $v_0=145\ km/h$ (remember that the Greenshields diagram tends to overrate the value of the speed at low flow rates). This figure also includes the range of speeds between $\bar{v}_c(q)-\sigma_c(q)$ and $\bar{v}_c(q)+\sigma_c(q)$ c.q.

---

[14] Note that any relation between the historic mean speeds $\bar{\bar{v}}_c(q)$ and $\bar{\bar{v}}_f(q)$ and the actual flow rate q will lead to two distinct solutions for one value of q.

between $\bar{v}_f(q)-\sigma_f(q)$ and $\bar{v}_f(q)+\sigma_f(q)$, using $A_c=15, A_f=25, B_c=20$ and $B_f=10$. The three probability distribution function (*pdf's*) of the road link mean speed that are sketched in figure 4.3 correspond to values of $q=¼q_m$, $q=½q_m$ and $q=¾q_m$.

Since the probability distribution function of the road link mean speed that is used in the procedure for Automatic Congestion Detection using combined probe vehicle and induction loop data depends on the actual measured flow rate, also the performance of this ACD procedure will depend on the value of the flow q. This means that the false alarm rate and the detection that have been computed in the previous chapter for probe vehicle data only for the entire ACD process, will be dissimilar and flow rate dependent when a combination of probe vehicle and induction loop data can be used. The three distinct *pdf's* that have been sketched in figure 4.2 for $q=¼q_m$, $q=½q_m$ and $q=¾q_m$ roughly correspond with values of $\bar{\bar{v}}_c(q)$ and $\bar{\bar{v}}_f(q)$ of 35 and *110,* 45 and 100, and 60 and 85 *km/h*, respectively. This implies that a value of the flow rate of $q=¼q_m$ for flow rate depending prior knowledge corresponds with the 'probe only' situation in chapter 3. So, the performance of ACD for combined probe vehicle and induction loop data at a flow rate of $q=¼q_m$, i.e. using the values of $\bar{\bar{v}}_c(q)=35$ *km/h* and $\bar{\bar{v}}_f(q)=110$ *km/h,* has been studied in the previous chapter. Here, we will investigate the performance of the designed ACD procedure using $\bar{v}_c(q)=60$ *km/h* and $\bar{v}_f(q)=85$ *km/h*, i.e. for $q=¾q_m$. For this, the formulae that have been derived in the previous chapter are used, with the above-mentioned parameters of $A_{,}$, and $B_{c,f}$ (in the formulae for σ_c^2 and σ_f^2) and a standard deviation in the individual probe vehicle speeds of $\sigma=35$ *km/h* in the case of congestion and $\sigma=30$ *km/h* in the case of free-flow.

N_p	$K^*_{F=0.01}$ (km/h)	$dr(K^*_{F=0.01})$ (%)	$K^*_{F=0.05}$ (km/h)	$dr(K^*_{F=0.05})$ (%)	$K^*_{F=0.10}$ (km/h)	$dr(K^*_{F=0.10})$ (%)
1	11.4	9.5%	33.0	28.4%	44.5	40.0%
2	13.3	11.2%	34.3	29.6%	45.5	42.0%
3	13.9	12.1%	34.7	30.0%	45.8	44.2%
5	14.4	13.0%	35.1	30.8%	46.1	45.7%
10	14.8	13.8%	35.4	31.7%	46.3	47.9%

Table 4.1 Applicable values for policy threshold K^*_F for different values of N_p and accompanying false alarm rates (far) and detection rates (dr)

Tables 4.1 presents the computed false alarm rates, $far(K^*_F)$, and the accompanying detection rates, $dr(K^*_F)$. From this table it becomes clear that performing ACD becomes more complicated when the flow rate increases. In order to achieve a low false alarm rate (i.e $far=1\%$) the value of the policy threshold K^*_F at $N_p=2$ amounts to 60.7 *km/h* (table 3.3) when $q=¼q_m$ and amounts to 13.3 *km/h* (table 4.1) when $q=¾q_m$. *This* is easily explained since the historic mean' speeds for congestion and free-flow, $\bar{\bar{v}}_c(q)$ and $\bar{\bar{v}}_f(q)$, approach each other when the flow rate q increases.

This causes the detection rate to deteriorate: $dr=85.2\%$ when $q=¼q_m$ (table 3.3) and $dr=11.2\%$ when $q=¾q_m$ (table 4. 1), at $N_p=2$ and $far=1\%$. However, this also causes the absolute estimation error in the estimate for the road link mean speed that results when using the wrong prior knowledge (i.e. using $\bar{\bar{v}}_c(q)$ instead of $\bar{\bar{v}}_f(q)$) diminishes when the flow rate increases. This is intuitively clear and can be seen from the formulae of the estimators.

We will also illustrate the above by means of the following example.
Consider a situation with a relatively low flow rate of $q=¼q_m$, where $\bar{\bar{v}}_c(q)=35$ *km/h* and $\bar{\bar{v}}_f(q)=110$ *km/h*. When 2 probe reports are received ($N_p=2$) with speeds of $v_1=80$ *km/h* and $v_2=85$ *km/h*, the estimate using the free-flow prior knowledge amounts to *91.67 km/h*, while the estimate using the congested prior knowledge amounts to 72.96 *km/h*. For a situation with a relatively high flow rate of $q=¾q_m$, where $\bar{\bar{v}}_c(q)=60$ *km/h* and $\bar{\bar{v}}_f(q)=85$ *km/h*, the estimate using the free-flow prior knowledge amounts to 82.63 *km/h*, while the estimate using the congested prior knowledge amounts to 79.83 *km/h*.

In conclusion we can state, that on the one hand, performing ACD is increasingly complicated using the refined multi-layered speed model based on a combination of probe vehicle and induction loop data, at high flow rates but that, on the other hand, the resulting estimate is decreasingly sensitive to a misselection (i.e. a false alarm or a m&detection). Accordingly, the absolute requirements for the entire procedure, expressed in the parameters F (maximum *fur)* and D (minimum *dr),* are lower and higher respectively, so adequately performing ACD remains possible at low N_p and high q. In this respect, the performance of the ACD procedure is considered adequate when its outcome supports a proper estimate of the road link mean speed, that is, an estimate whose value resembles the actual value of the road link mean speed. This implies that a high flow rates a correct ACD outcome is only of minor importance when the actual value of the road link mean speed fluctuates around the optimum speed v,. When the actual value of the road link mean speed substantially deviates from this optimum speed v,, a correct ACD outcome remains necessary, but is also possible very well.

4.1.2.3 Step 2: Estimation of Road Link Mean Speed and Travel Time

The process for estimating the road link mean speed and travel time from a combination of received individual probe vehicle speeds and measured induction loop data is analogous to the 'probe only' estimation process already described in the previous chapter, except for the issue of prior knowledge. When probe vehicle and induction loop data are integrated; the prior knowledge is enhanced by using measured local-related quantities, which are available through induction loop measurements. Exactly how this loop data is incorporated into the prior knowledge has already been described in sub-section 4.1.2.2.

The resulting (Bayesian) estimator for combined probe vehicle and induction loop data, $\bar{\bar{v}}^*_B(q)$, is given by (see chapter 3 for a derivation)

$$\bar{\bar{v}}^*_B(q) = E_{c,f}\{\bar{\bar{v}} \mid v_1,\ldots,v_{N_p}\}$$

For free-flow traffic, this can be written

$$\bar{\bar{v}}^{*}_{B}(q) = \frac{\eta(q)\,\bar{\bar{v}}_{f}(q) + \sum_{i=1}^{N_p} v_i}{\eta(q) + N_p}$$

For congested traffic this formula is somewhat more complex and we refer to chapter 3 for an exact specification (here $\bar{v}_c(q)$, $\sigma_c(q)$ and $\sigma(q)$ are flow rate dependent). Also the formula for the credible set of the Bayes estimator which has been presented in chapter 3 for a situation where only probe vehicle data was assumed to be available should be adjusted accordingly when probe vehicle and induction loop data are combined. For the ultimate estimation process, it depends on the outcome of the first step (Automatic Congestion Detection) whether the free-flow or the congested estimator with corresponding prior knowledge is used.

From the formula of the (Bayesian) estimator for free-flow it becomes clear that, since an increasing flow rate q causes the variance in the road link mean speed $\sigma_f(q)$ to increase, the parameter $\eta(q)=\sigma^2(q)/\sigma_f^2(q)$ diminishes. So, the role of the prior knowledge $\bar{\bar{v}}_f(q)$ becomes less significant at increasing q. Therefore, in the case where probe vehicle and induction loop data is integrated by incorporating flow rates in the (Bayes) estimator of the road link mean speed, this estimator (for the free-flow regime) converges to the direct or Moment estimator at high flow rates. Through the dependence of the speed variances on q in the combined case, this convergence process takes place more rapidly than in the 'probe only' case, for which the speed variances were modeiled as constants. Since the estimations for $\lim q \downarrow q_m$ and $\lim q \uparrow q_m$, where both regimes 'meet', should in practice be equal, this convergence process should also hold for the congestion regime, although this is not directly shown.

Table 4.2 and 4.3 present results of computations by implementing the presented formulae for the estimator and credible sets for several numbers of $N_p \in \{1,2,3,5,10\}$ and values of probe vehicle samples. The tables contain the estimate as well as the probability mass of the estimator contained in the credible set. We have used a high flow rate of $q=¾q_m$. The obtained results are to be compared to the results presented in the previous chapter (table 3.4 and 3.5) where only probe vehicle data was assumed to be available, since a value of the flow rate of $q=¼q_m$ (tables 4.2 and 4.3) corresponds with the results in tables 3.4 and 3.5 (see earlier). Furthermore, for $q=¾q_m$, the standard deviation in the individual probe vehicle speeds amounts to $o(q)=35$ km/h for congestion and to $\sigma(q)=30$ km/h for free-flow, the standard deviations in the road link mean speed amount to $\sigma_c(q)=31.25$ km/h and $\sigma_f(q)=28.75$ km/h (see formulae) and the historic mean speeds are $\bar{\bar{v}}_c(q)=60$ km/h and $\bar{\bar{v}}_f(q)=85$ km/h. Again, the required quality of the speed estimations (the level of accuracy a) has been taken 10% ($\alpha=0.10$) and 15% ($\alpha=0.15$).

		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 70$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 85$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 100$	
$N_p = 1$	$\bar{v}^*_B$	77.82		85		92.18	
(α=0.10)	(α=0.15)	33.46%	48.34%	35.94%	51.64%	38.40%	54.82%
$N_p = 2$	$\bar{v}^*_B$	75.29		85		94.71	
(α=0.10)	(α=0.15)	40.78%	57.83%	44.65%	62.58%	48.39%	76.00%
$N_p = 3$	$\bar{v}^*_B$	73.99		85		96.00	
(α=0.10)	(α=0.15)	46.63%	64.95%	51.38%	70.39%	55.92%	75.24%
$N_p = 5$	$\bar{v}^*_B$	72.68		85		97.32	
(α=0.10)	(α=0.15)	55.65%	74.96%	61.44%	80.68%	66.73%	85.38%
$N_p = 10$	$\bar{v}^*_B$	71.47		85		98.53	
(α=0.10)	(α=0.15)	70.32%	88.23%	76.73%	92.65%	82.08%	95.60%

Table 4.2 Credible sets for road link mean speed estimations for-free-flow traffic

		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 50$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 60$		$\frac{1}{N_p}\sum_{i=1}^{N_p} = 70$	
$N_p = 1$	$\bar{v}^*_B$	57.33		63.97		72.45	
(α=0.10)	(α=0.15)	31.84%	40.67%	30.22%	39.04%	29.07%	38.85%
$N_p = 2$	$\bar{v}^*_B$	56.79		63.35		72.04	
(α=0.10)	(α=0.15)	48.51%	49.53%	43.12%	46.25%	36.52%	41.89%
$N_p = 3$	$\bar{v}^*_B$	56.21		62.70		71.75	
(α=0.10)	(α=0.15)	56.30%	58.96%	50.04%	56.80%	47.13%	55.03%
$N_p = 5$	$\bar{v}^*_B$	55.12		62.16		71.20	
(α=0.10)	(α=0.15)	62.03%	67.42%	58.48%	63.69%	56.40%	60.98%
$N_p = 10$	$\bar{v}^*_B$	53.48		61.78		70.92	
(α=0.10)	(α=0.15)	65.84%	71.00%	62.91%	66.54%	61.30%	62.04%

Table 4.3 Credible sets for road link mean speed estimations for congested traffic

The results in the tables show that, for example, if the number of probe messages received from a particular road link in a particular time period is 3 (N_p = 3) and the sample mean of these messages is 70 *km/h* the Bayes estimation of $\bar{v}(q)$ for free-flow (congestion) equals 73.99 *km/h* (71.75 *km/h*) and the probability that $\bar{v}(q)$ lies in *[62.89,* 85.091 (*[60.99, 82.52]*) (i.e. α=0.15) is 64.95% (55.03%).

From these tables it turns out that estimating the actual road link mean speed at high flow rates is more difficult than at low flow rates. This is caused by both regimes, i.e. free-flow and

congestion, converging with a increasing variance at an increasing flow rate. As a result, the accuracy and the reliability of the estimates decrease. However, since both regimes converge, also the historic and the free-flow prior knowledge converges. This means that the distinction between an estimator using $\bar{\bar{v}}_c(q)$ and an estimator using $\bar{\bar{v}}_f(q)$ eventually becomes insignificant, and so the required demands with respect to reliability and accuracy of the ultimate road link mean speed estimate should also decrease with increasing flow rate.

4.1.3 Conclusions

In this section we have studied how data from non-infrastructure based probe vehicles and data from infrastructure based induction loop detectors can be fused in order to enhance the reliability and accuracy of real-time road link mean speed and travel time estimations for ATMIS. We have distinguished between two approaches: utilization of induction loop data and additional probe vehicle in sub-section 4.1 .1, and utilization of probe vehicle data and additional induction loop in sub-section 4.1.2.

With respect to the first approach, we have extended the COMETT algorithm for estimating real-time travel times from induction loop data alone that was developed in chapter 2, by incorporating probe vehicle data. Probe vehicle updates serve as additional re-calibrations in order to guarantee a persistent performance of the travel time determination process, even in the case of protracted congestion. We have simulated probe vehicles traveling on a road link equipped with induction loop detectors and applied the extended COMETT algorithm. The obtained results have been shown to be more accurate and, above all, more reliable compared to the results when only induction loop data was available.

With respect to the second approach, we have extended the algorithms for estimating the real-time road link mean speed and travel time from probe vehicle data alone that were developed in chapter 3, by incorporating induction loop data. Induction loop measurements of the flow rate contribute to a more realistic, actual prior knowledge that is used in both the selection of the prevailing traffic regime as in the estimation of the actual road link mean speed. In the case where **only** probe vehicle data is available we have presumed this prior knowledge, comprising of a historic free-flow mean speed and a historic congested mean speed, to be constant. Although this simplification is not very realistic, in particular at high flow rates the actual difference between the speed of free-flowing and congested traffic is much smaller that presumed by modeling these as two constants, for automatically detecting delaying congestion (ACD) it is sufficiently accurate since this exclusively concerns significant delays, i.e. a road link mean speed that is significantly lower than regular, is of concern. For automatically detecting incidents (AID) however, an accurate estimation of the value of the actual road link mean speed is of concern, so the utilized historic prior knowledge has to be accurate, i.e. in accordance with reality. This is done by modeling this historic knowledge depending on the actual flow rate measured by the induction loop detectors. In this way, for low flow rates, the model studied for the 'probe only' case in the previous chapter and the corresponding appropriate results eventuate. For high flow rates, selecting the prevailing traffic regime becomes increasingly complicated, but the resulting outcome is decreasingly sensitive to a misselection.

When a similar number of probe messages is successfully received as under conditions of low flow rates, the accuracy of the estimated road link mean speed under conditions of high flow rates is lower. However, since the free-flow and congested traffic regimes converge under conditions of high flow rates, the distinction between utilizing the free-flow or congested prior knowledge becomes increasingly insignificant and the demands with respect to reliability and accuracy of the estimate should also decrease with increasing flow rate.

From the results obtained in this section we learn that fusing probe vehicle data and induction loop data can considerably enhance the quality of derived ATMIS traffic information. Hence, constructing a combined induction loop and probe vehicle monitoring system for ATMIS is highly advisable. On road links where relatively many infrastructure based induction loop detectors are present, such as most major freeways, the first approach for obtaining real-time road link mean speeds or travel times should be applied. On road links with relatively few infrastructure based induction loop detectors are available, such as small highways and urban roads, the second approach for obtaining real-time road link mean speeds or travel times should be applied.

4.2 Incident Management and Existing Methods for Automatic Incident Detection

In chapter 2 it has been concluded that determination of reliable and accurate real-time travel times and adequately performing Automatic Incident Detection for ATMIS purposes using exclusively infrastructure based traffic detectors with rather large mutual spacings, is extremely arduous. The first issue has efficaciously been addressed in the previous section by incorporating probe vehicle data into the formerly developed COMETT algorithm and by incorporating induction loop data into the formerly developed 'probe only' procedures. The latter issue, Automatic Incident Detection for ATMIS, is the subject of the next section. Before we will formulate principles for fusing traffic data from infrastructure and non-infrastructure based traffic detectors in order to perform appropriate AID for ATMIS, we will outline the significance of proper Incident Management for ATMIS and give a concise overview of existing methods for automatically detecting disturbances in the traffic flow in this section.

4.2.1 Incident Management"

According to (FHWA, 1987) incidents and the non-recurrent congestion they induce accounted for 64 percent of all traffic congestion in the USA in 1987 and is expected to exceed 70 percent by 2005. For the Netherlands, in 1994, more than 30 percent of all congestion was unpredictable (i.e not caused by regular bottlenecks) (RWS, 1994).

---

[15] The following overview of incident management for ATMS is encapsulated from a proposal concerning incident detection issues (Virginia Tech, 1992) submitted to FHWA by the ITS Center of Excellence Virginia Polytechnic Institute and State University in association with among others TU Delft, in which the author of this dissertation participated on behalf of TU Delft.

Delays caused by traffic congestion exacerbate air pollution and engender wasted time and fuel, while in particular unpredictable non-recurrent congestion adversely impacts economy and the reliability of transportation processes. Incident management refers to the coordination of activities undertaken to restore traffic to normal operation after an incident has occurred. Figure 4.4 illustrates the effect that a traffic incident has on traffic flow. Traffic flow past the incident is a fraction of the demand flow, and it is a function of the number of lanes available, and the type of disturbance at the incident site. The displacement along the ordinate from the line representing the flow past the incident to the demand flow is the length of the queue. The queue takes a remarkably long time to dissipate after the traffic incident is cleared from the roadway. The shaded area between the demand flow line of the supply flow line in figure 4.4 represents the total delay incurred, measured in vehicle hours.

The goal of Incident Management (IM) is to minimize the delays caused by traffic incidents. This can be achieved through swift incident detection, quick response, efficient clearance, and appropriate (on-scene) traffic management, resulting in reduced queue build ups while road users can realize significant reductions in delay times. From figure 4.4 we learn that incident management involves five major tasks:

incident detection and verification,
incident response,
on-scene management,
incident clearance, and
deploying traffic management measures.

Incident Detection

Incident detection is the process of identifying the spatial and temporal coordinates of an incident occurrence. Traditionally, traffic monitoring and surveillance systems play an important role in incident detection, but are only operated at critical and very local sites in a road network. Wide-area incident detection mostly relies on highway service personnel, police patrol or road users reporting an encountered incident, for instance using (cellular) car phones. However, in many cases the incident is detected only after precious time has been lost and long queues have developed.

Incident Verification

Verification of an incident occurrence is an important method to eliminate false alarms, no matter what the detection method might be. Currently, much of the verification is done through police and service patrols or video cameras, where the camera is directed from the traffic center at the specified location following from an incident report. The reliability of the verification is a function of two factors, namely the incident type and the source reporting. Proper incident verification can eliminate the dispatch of inappropriate equipment for incident clearance and minimizes losses from false alarms.

Incident Response

Incident response involves the mobilization of the personnel, equipment, and materials that are needed to clear the incident. Response to an incident may involve several agencies, both private and public, to clear the incident and to provide assistance to traffic and to the passengers involved in the incident. One of the major components of a successful response is the presence

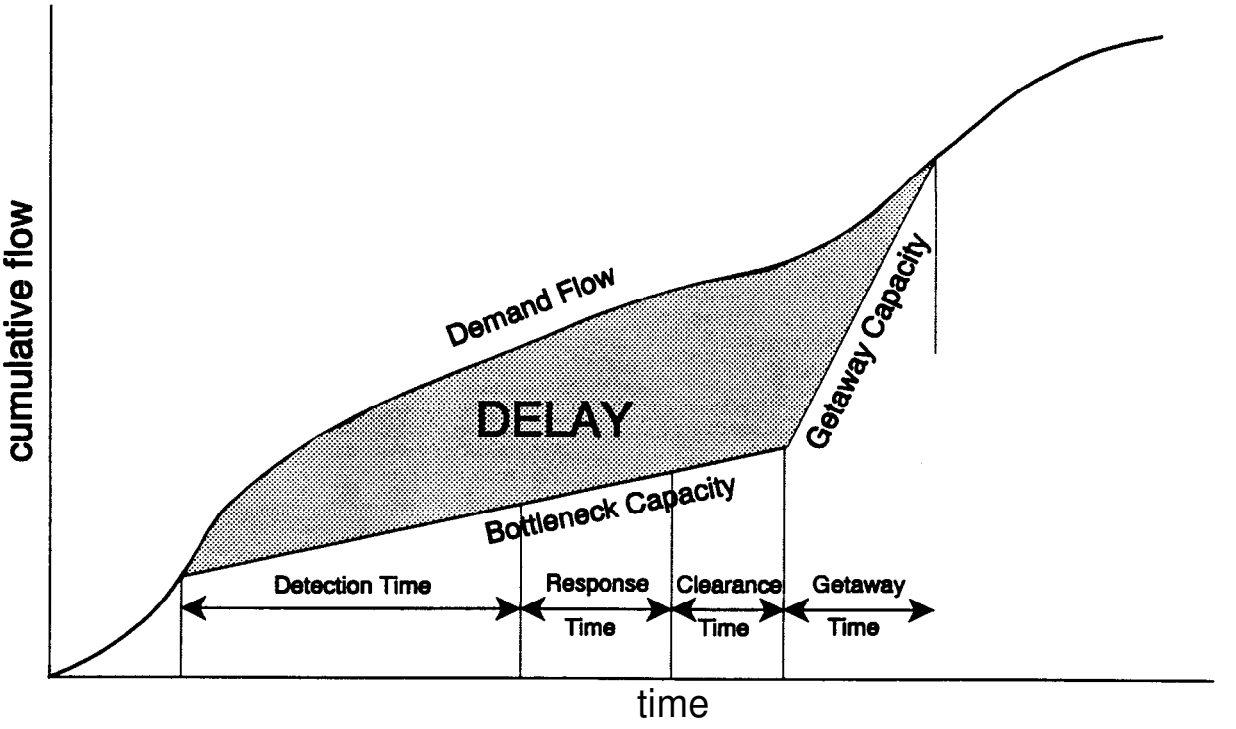

Figure 3.4 ***Traffic flow*** *during incidents (Source: (Virginia Tech, 1992))*

of an efficient communication network connecting the various agencies.

On-Scene Management
On-scene management refers to the management of personnel involved in traffic control and operations during the clearance process. Personnel management deals with the issue of leadership during the clearance process, the chain of command, and managing the aspect of working within the overlapping responsibilities and authorities that exist.

Incident Clearance
Incident clearance forms the central objective of incident management. Clearance of an incident involves the coordinated functioning of several agencies. The various incident management factors that affect the clearance process are the availability of resources, operational procedures, training of personnel, and administrative coordination.

Deploying ***Traffic*** *Management Measures*
The road user is one of the most important participants in traffic management during an incident. It is the road users' role to make adjustments to his/her travel patterns in response to the information provided or the measures deployed by the traffic center. This role includes diverting around the site of the incident and changing departure times at trips origins. Deploying traffic management measures and proving information to road users involves two major components: collecting and further processing of reliable traffic data, and deliverance of this information by means of traffic management measures or information dissemination in (more or less) real-time.

4.2.2 Incident Detection Issues

Incident detection is one of the most crucial components of Incident Management. Many incidents are only detected after previous time has been lost and long queues have been developed, while without adequate information about the occurrence of the incident it is not possible to initiate any management actions. The process of incident detection encompasses several issues which contribute significantly to the process involved in the confirmation of the incident occurrence:

algorithmic issues,
surveillance issues, and
verification issues.

Algorithmic Issues

The Automatic Incident Detection (AID) system comprises of an algorithm that uses input from a traffic surveillance system to detect a sudden alteration in the flow of incoming traffic data, and flags an alarm when there is symptomatic evidence of an abrupt traffic disruption. Typically, traffic flows, occupancies, and speeds form the inputs to AID systems. Several algorithms currently exist that perform this function (see also the next sub-section) and the logic behind these algorithms varies considerably. But, the central objective is to identify the location, time and the nature of the disruption. The most common Measures of Effectiveness (MOEs) of these AID algorithms are (see also chapter 3):

detection rate *(dr),*
false-alarm rate *(far),* and
time-to-detect (*ttd*).

These MOEs play an important role in the successful implementation of any AID algorithm. False alarms are very expensive due to unnecessary dispatches and block resources from reaching real incident occurrences. The detection rate reflects the number of incidents that are correctly detected by the algorithm, where undetected incidents are of particular concern due to greater delays and the safety of road users involved. A quick detection time is one of the most important requirements for efficient decision making for traffic management after an incident.

Most incident detection algorithms do not possess the reliability levels required for network-wide application with the large inter-detector spacings that are characteristic for ATMIS. As we have already seen in chapter 3, the three MOEs are greatly inter-related and constitute an optimization problem with a non-inferior solution (Payne and Tignor, 1978). An increase in the detection rate is always accompanied with an increase in the false alarm rate, while a reduction in the false alarm rate is at the expense of increased time-to-detect. The influence of external factors, such as additional surveillance data and modification of the algorithm by adopting combination of multiple approaches to build one system, on this relationship is not clear yet. Hence, the simultaneous improvement of all the three MOEs has not yet been attained thus far.

Surveillance Issues

Traffic surveillance is the process of employing hardware and software to sense the traffic flow and relay it to the traffic center. Thus a surveillance system is the source of all traffic flow data that algorithms use to detect an incident. The various aspects of surveillance that reflect upon the

data are: reliability, completeness, detector density, data accuracy, sparsity, real-time characteristics and performance under different environments (such as weather conditions). Traditionally, during the 60's and the 70's, traffic surveillance was employed to obtain data for purposes that included traffic counts and speeds for planning and safety studies. Later, local traffic control originated, requiring constant and comprehensive information on the traffic flow at one specific location. With the advent of central traffic control and traffic management (ATMIS), requiring real-time traffic information from a entire road network, the traditional and local surveillance systems were viewed as potential data sources. But, they are sparse and not frankly adequate to perform this task.

Surveillance issues play a decisive role in the quality and characteristics of data available to AID systems. The AID algorithm deployed, thus, draws close relationships with the deployed surveillance system. Since most of current algorithms rely only on values of flows, speed and occupancies, additional surveillance systems, such as probe vehicles, could be incorporated towards enhanced algorithm performance.

Verification Issues

Incident verification forms the basis for minimizing false alarms and responding adequately to the incident site. Most AID systems at present verify incidents through video camera surveillance (if available) or through the police. Thus verification bears heavily on the surveillance system available at the location, while the extent of manual verification required depends on the reliability of the algorithm to a large extent. Manual verification through police patrols consumes precious time and resources. Though incident reports are often reported by road users passing by, experience indicates that relying completely on such information is not desirable. Improving incident detection algorithms shall greatly minimize the need for manual verification, since none of the currently used AID algorithms provide a level of performance required to eliminate the manual verification process.

4.2.3 Existing Methods for Automatic Incident Detection

The origination of (automatic) incident detection goes back to the 1960s where, in Chicago, loop detector occupancies (greater than 40%) were used to denote a high reduction in capacity and an alert of a potential incident. When such values were observed at non-lane drop locations, emergency patrol vehicles were requested to investigate the situation. With the increase in number of incidents such approaches were inefficient. This resulted in the urge for developing Automatic Incident Detection (AID) systems.

The earliest exploration of automatic detection and location of incidents was probably performed in Detroit, by (Courage and Levin, 1968). While these studies were limited in nature, they demonstrated some ability to detect incidents. This was followed by the development of detection models based on occupancy changes at the upstream and downstream station, for a Los Angeles freeway (Schaefer, 1969). A detection model using volumes as a controlling parameter was suggested by (Whitson et al. 1969). (Cook and Cleveland, 1974) conducted a study and analyzed the California model (Schaefer, 1969) and noted specific weaknesses. Later, (Dudek, Messer and Friebele, 1973) developed a control logic for automatic operation of safety warning

systems. (Dudek, Messer and Nuckles, 1974) proposed an automatic incident detection model based on statistical principles. A detailed development and testing of detection algorithms was conducted by (Payne, Helfenbein and Knobel, 1976), which resulted in further development and refinement of the California algorithm and several variations of it. This algorithm became popular, was extensively tested and deployed, and is currently the most widely used incident detection algorithm in the USA.

In the Netherlands, in 1972, about 40 to 50 percent of all accidents on highways appeared to find its cause in the origination of incidents and the congestion these induced. The majority of these secondary incidents happened within 10 minutes of the initial incident. This was the main occasion for developing the Motorway Control and Signalling System (MCSS) which was introduced in 1981 between the Hague and Rotterdam, with the principal objective to warn upcoming traffic (see also (Westerman, 1995)). The algorithm for automatically detecting such stagnations in the traffic flow used speeds measured by induction loop data of measuring sites with a mutual distance of about 500 meters. This simple and rather basic algorithm has hardly been changed since and is still incorporated in the present Dutch Motorway Traffic Management system (MTM).

Despite their general use, the California Department of Transportation (Caltrans) has- been using the family of California algorithms for many years, these algorithms had several deficiencies that limited its complete acceptance for automatic incident detection. Unacceptable false alarm rates, calibration difficulties and data requirements were the major issues. To overcome these deficiencies, other types of algorithms were studied, including different varieties in statistical type. While these possessed certain distinct advantages, they could not fulfil all the expectations of AID systems. In the 1980s, researchers realized that loop detector based algorithms have certain limitations. These limitations were due to the reliability of detector data, the very local orientation of this detector data, and due to the fact that these algorithms did not detect incidents directly, but through the effects of incidents. This not only led to the investigation of other types of surveillance, especially devoted to video image processing systems although no such system supports AID through image recognition yet, but also to improvements in detector technology, automatic checking of data accuracy (Jacobson, Nihan and Bender, 1990; Cleghom et al., 1991), and algorithmic developments. Newer concepts such as catastrophe theory (Persaud, Hall and Hall, 1989; Gall and Hall, 1989; Persaud, Hall and Hall, 1990; Forbes, 1992), artificial intelligence (Han and May, 1989) and neural networks (Cheu et al., 1988; Ritchie; Cheu and Recker, 1992; Ritchie and Cheu, 1993) were recently proposed. Recently, other new technologies have been envisaged for application in incident detection, such as Automatic Vehicle Identification (AVI) and Automatic Vehicle Location (AVL), cellular phone surveillance systems, and, among others in this chapter, probe vehicles.

Before advancing on indicating how a probe vehicle surveillance system and obtained probe vehicle traffic data can be incorporated into traditional AID systems, we will briefly examine existing AID algorithms in more detail, and determine their particular abilities and weaknesses. Based on the compilation above and following (Bush and Fellendorf, 1990) and (Bush, 1991) we adopt a distinction between existing automatic incident detection algorithms based on basic principles, leading towards the classification below.

Direct Comparison
The basic principle of this type of AID systems comprises directly comparing measured values or values derived from them to previously fixed threshold values. Smoothing techniques are often used to eliminate the influence of stochastic variations in the measured values. In general, occupancy and mean speed, as well as appropriate combinations of both, are utilized. The best known example of AID systems using direct comparison is the family of California algorithms (Schaefer, 1969). Also the AID algorithm used in the Dutch signalling system MTM belongs to this type (Klijnhout and Langelaar, 1987).

Temporal Forecast
This type of AID systems departs from values of traffic parameters measured at a certain site, forecasts their value in time and compares the forecasted values with the measurements at that particular time period. It is assumed that an incident produces a significant deviation between the forecasted and the measured values. For practical applications with small time periods, methods using a simple smoothed mean with linear trend corrections have shown equal performance compared to advanced time-series analysis methods. In general, occupancy and mean speed, as well as appropriate combinations of both, are utilized. Examples of time-series analysis can be found in (Dudek and Messer, 1974) concerning a standard deviation algorithm, (Cook and Cleveland, 1974) concerning a double exponential algorithm) and (Ahmed and Cook, 1979; Ahmed and Cook, 1980) and (Masters, Lam and Wong, 1991).

Spatial Forecast
This type of AID systems makes a forecast of the value of certain traffic parameters at a site downstream on the basis of values of traffic parameters measured at a site upstream. The spatially forecasted values are compared to measurements at the site downstream. Again, it is assumed that an incident produces a significant deviation between the forecasted and the measured values. In general, occupancy and mean speed, as well as appropriate combinations of both, are utilized. Examples of spatial forecast AID algorithms can be found in (Payne, et al., 1976; Levin and Krause, 1979).

Filtering Techniques
This type of AID systems tries to eliminate a serious deficiency of, in particular, the first three types of AID systems, namely that alarm notifications often correspond to sudden and unexpected traffic disturbances that are not caused by an incident. In this way, filtering techniques aim at reducing the likelihood of false incident decisions by smoothing raw detector data. The pre-processed data is used as input for other AID algorithms. Tests of freeway incident detection through filtering can be found in (Stephanedes and Chassiakos, 1991) and (Stephanedes and Chassiakos, 1993).

Model-based Estimation
This type of AID systems uses a (simplified) mathematical model representing the traffic system situated between the measurement cross-sections and is permanently updated using traffic measurements. By means of this traffic model the occurrence of an incident becomes known from a significant discrepancy between model outcomes and measured parameters, whereafter the location and the severity of the conjectured incident can be assessed. A comprehensive model-based AID system has been developed in (Willsky et al., 1980).

Pattern Recognition
This type of AID systems is based on a similar assumption that has been utilized for formulating the travel time estimation algorithm presented in chapter 3 of this dissertation, namely that during non-incident conditions the traffic flow propagates undisturbed and values of traffic parameters observed at an site upstream can be observed at a site downstream delayed by a certain period of time. Good results are obtained with this type of AID systems, particularly at low traffic volumes. Examples can be found in (Payne, Helfenbein and Knobel, 1976; Payne and Tignor, 1978; Tsai and Case, 1979).

Cluster Analysis
This type of AID systems analyses off-line sets of traffic parameters for disturbed and undisturbed situations, and transforms these data into distinct clusters of different traffic conditions. Values of measured parameters are then compared to those of the clusters in order to classify the prevailing traffic conditions. In (Hilgers, 1980) such a cluster technique has been presented.

Probabilistic Approach
This type of AID systems takes into account the statistical (un)certainty of the occurrence of an incident, by determining the probability of an incident on the monitored section as well as the detection rates of the algorithm for disturbed and undisturbed traffic situations, and continuously updating these. In this way, the significance of an incident alarm can be determined. In (Levin and Krause, 1978; Levin and Krause, 1979) such a probabilistic method has been proposed, which however has appeared to be not very suitable for practical implementation.

Image Processing
This type of AID systems produces images of traffic situations with small time intervals, analyses these images and compares characteristic properties, either to characteristics of earlier images or to pre-defined standards. At the moment, the technique of using video image processing systems for automatic incident detection has only been applied in pilot systems and no surveillance system is available that supports AID through image recognition yet. However, they will become more suitable for practical application with the increasing capacity of micro-processors. Examples of video image processing systems are (see also chapter 3, sub-section 3.1) AUTOSCOPE (Michalopoulos, 1991), TRISTAR (Motyka et al., 1991) and IMPACTS (Hoose et al., 1991).

Many studies have dealt with comparing AID algorithms. For instance, (Cook and Cleveland, 1974), (Payne, et al., 1976) and (Levin and Krause, 1979) compared numerous direct comparison and temporal and spatial forecast AID algorithms. The studies of (Leutzbach and Bush, 1983) and (Busch, 1986) also included other types of AID algorithms. The general conclusions from these studies can be summarized as follows (see also (Busch, 1991)):

- none of the tested methods proved to be superior in all situations,
- at high traffic volumes and small detector spacings (up to 1 kilometer) most algorithms produced a high detection rate and a short time-to-detect,
- at low traffic volumes most of the algorithms fail to produce satisfactory results,
- at increased detector spacings (more than about 2 kilometers) the performance of all algorithms severely diminishes; none of the algorithms appeared to be suitable for inter-

detector spacings of more than 5 kilometers (see also (Bush and Fellendorf, 1990)).

The above concise overview of Incident Management in general and Automatic Incident Detection in particular, has revealed that none of the existing AID algorithms is able to fulfil the fundamental requirements of IM within ATMIS. Especially, the high false alarm rate, the long time-to-detect and the impossibility of adequate automatic verification are unacceptable for ATMIS purposes. Based on the results gained in this dissertation with respect to non-infrastructure based probe vehicles and their complementary characteristics compared to infrastructure based induction loop detectors, we surmise that the combination of probe vehicles with loop detectors allows the development of novel traffic surveillance and data' processing techniques that substantially outperform existing AID systems. However, the fundamental properties and behavior of these new techniques are not yet known. The ensuing section addresses the specification of principles of automatic incident detection using combined probe vehicle and induction loop data using principles, techniques and particular experiences described in the concise overview in this sub-section. The intention of the ensuing section is not to design a AID algorithm that is directly suitable for practical implementation, but deals with formulating how data from both data sources can be fused in order to enhance the performance of present AID systems. In particular, we will focus on ways to decrease the false alarm rate. Furthermore, although complete elimination of false alarms will still not be guaranteed, we will show that it will be possible to introduce a high level of automatic (non-human interfered) verification of automatic incident detection by incorporating data obtained through probe vehicles.

4.3 Principles of Automatic Incident Detection using Combined Probe Vehicle and Induction Loop Data

In the previous section we have shown that adequate Incident Management is one of the core activities of ATMIS and that existing Automatic Incident Detection systems and algorithms are not capable of dealing with the extended road links (of about 5 to 10 kilometers). that are required for managing road traffic on a entire network by means of ATMIS.

The foremost difficulty of the examined Automatic Incident Detection algorithms when applied on extended road links is that, due to the large detector spacings, they are not able to appropriately detect disturbances in the traffic flow and to distinguish between traffic conditions which are and which are not representative of the traffic flow on the entire road link under consideration. Hence, they suffer from a rather low detection rate, a rather high false alarm rate or a rather large time to detect. Recognizing the characteristics of probe vehicles observed in the previous chapter about utilization of probe vehicles we surmise that the difficulties with respect to Automatic Incident Detection can be resolved by incorporating non-infrastructure based probe vehicles into the infrastructure based induction loop detectors. We will further investigate this issue in this section and we will formulate fundamental principles of such a compound algorithm for Automatic Incident Detection using data from both infrastructure and non-infrastructure based traffic detectors.

In addition to the above-mentioned data sources in ADVANCE also anecdotal data sources are

employed (Bhandari, Koppelman, Schofer, Sethi and Ivan, 1995). These anecdotal sources comprise reports of particular events affecting traffic flow, provided by people traveling on or monitoring the road network, such as emergency service workers. In this section we will not consider anecdotal data sources, nor will we consider exceptional probe vehicle reports such as emergency messages, as this type of traffic data is considered to be very divergent from the other two classes of data sources and is not directly appropriate for automatic processing.

Since, in particular during the introduction phase, the probe vehicle penetration grade ζ will be rather low, a combined Automatic Incident Detection algorithm should be capable of properly operating also when no probe reports are received from a certain road link during a certain time interval. For this reason, an infrastructure based or fixed detector algorithm should constitute the basis of a combined algorithm and a non-infrastructure based detector or probe vehicle algorithm basically should serve to enhance the performance with respect to reducing the false alarm rate, increasing the detection rate and restraining the time-to-detect. When no probe data is available in a particular time interval for a particular road link, the outcome of the fixed detector algorithm is the outcome of the compound algorithm. Hence, the performance of such a compound algorithm is always at least equal to the performance of the incorporated fixed detector algorithm.

Fixed Detector Algorithm and Measured Local-Related Quantities

The fixed detector algorithm deals with local-related quantities, which are measured at the crosscut where the regarding induction loops are located. For a road link bounded by measuring sites ***A*** and ***B*** the local-related quantities may concern measured speeds, $v_{A,B}$, flows $q_{A,B}$ and occupancies $o_{A,B}$ occupancies (defined as the time the loops are occupied by a vehicle in relation to the total time). Based on the analysis in the previous chapter, we claim that the probe vehicle data can not improve the measured local-related quantities $o_{A,B}$ and $q_{A,B}$. Nor can probe vehicle data contribute to the determination of the quantities $v_{A,B}$ since these are also exclusively local-related.

The review of existing AID algorithms in the previous section has shown that the performance of comparative algorithms still outperforms other approaches for automatically detecting traffic flow disturbances using large inter-detector spacings. For this reason, we propose to adopt a comparative algorithm to serve as fixed detector algorithm in a compound probe vehicle and induction loop system. The traditional California Algorithm is the most commonly used of this class of AID algorithms. From many theoretical and practical experiences (see among others (Stephanedes, Chassiakos and Michalopoulos, 1991) and (Bush and Fellendorf, 1990)) the tests within this algorithm have revealed to produce a high detection rate but also to suffer from a rather high false alarm rate and a long time to detect when applied on extended road links. Hence, the property of a rather high detection rate is very appropriate for a trigger purpose while probe vehicle data should be incorporated in such a way that it compensates for the other deficiency, that is lowering the time-to-detect. Furthermore, probe vehicle data should ought to provide automatic verification.

Probe Vehicle Algorithm and Measured Section-Related Quantities

The probe vehicle algorithm deals with measured section-related quantities, viz. the individual probe vehicle speeds over (a longitudinal section of) the road link. This data collection comprises the raw traffic data that is transmitted by the probe vehicles and successfully received in the traffic center.

In the previous chapter we have developed an Automatic Congestion Detection (ACD) methodology to process the raw probe vehicle data in order to determine whether the traffic flow is in state of free-flow or congestion. In sub-section 4.1.2 this ACD algorithm has been extended to incorporate also induction loop data in order to enhance the performance of estimating the real-time road link mean speed and travel time. The probe vehicle algorithm within the compound AID algorithm comprises applying this extended ACD procedure.

For detecting incidents for ATMIS the final outcome of the probe vehicle algorithm should be treated with certain reservation, since the ACD procedure is basically intended to detect delaying congestion. Subsequently, this information was used to estimate the actual road link mean speed. On the one hand, this means that when the probe vehicle algorithm has observed an delaying congestion this obviously involves a 'proper' incident. On the other hand, this means that when the probe vehicle algorithm has not observed delaying congestion this not necessarily excludes an incident. So, also the probe vehicle algorithm is very appropriate for a trigger purpose, when the utilized policy threshold K is chosen appropriately. This however implies that its outcome should be verified in an additional step.

Trigger Mechanism and Verification Mechanism

The above-mentioned properties of the fixed detector algorithm and the probe vehicle algorithm call for a two step structure of the compound AID algorithm: a trigger mechanism which indicates a suspicion of occurrence of an incident and a verification mechanisme in which this suspicion of occurrence of an incident is (automatically) verified.

The trigger mechanism is aimed at swift detection of disturbances in the traffic flow which might give occasion to or might be brought about by the occurrence of an incident on the regarding road link. On that account, the parameters in this first trigger step are chosen such that this step is characterized by a high probability of detecting these (initial) disturbances (i.e. a high detection rate). This however, will imply a high probability of detecting non-incidents (i.e. a high false alarm rate). Hence, the second step (the verification mechanism) is predominantly added to compensate for this high false alarm rate and verifies the occurrence of an incident. In order to enhance the last Measure of Effectiveness (MOE), i.e. time-to-detect, we propose to employ a probabilistic instead of a deterministic approach. By repeatedly performing the verification step, an overall likelihood that an incident has occurred on a particular road link can be- determined. In this way, in relation to each type of incident and each traffic management purpose (such as deploying optional ATMIS measures, deploying compulsory ATMIS measures and dispatching emergency vehicles), a desired time-to-detect can be achieved, obviously with a

corresponding level of certainty (or vice versa)[16].

The utilization of two (more or less) independent data sources, i.e. from infrastructure based induction loop detectors and from non-infrastructure based probe vehicles, points towards a three branch approach of the verification mechanism: one module based on solely fixed detector data (a hybrid fixed detector algorithm), one module based on solely probe vehicle data (a hybrid probe vehicle algorithm), and one module in which data from fixed detectors and probe vehicles is directly interrelated.

Request more Probe Transmissions

When the outcome of the fixed detector algorithm or the probe vehicle algorithm. (conjointly constituting the trigger mechanism) indicates a suspicion of occurrence of an incident on a particular road link, all probes on the regarding road link are instructed to transmit experienced traffic conditions to the traffic center more frequently. In order to prevent the increase in the number of transmitted probe vehicle messages (from this particular road link due to the adjusted transmission scheme and due to the higher traffic density resulting from the incident) to eventuate in manifold harmful message collisions (Linnartz, Westerman and Hamerslag, 1994), the capacity of the uplink probe vehicle communication channel is first of all reserved for these particular transmissions by restraining the transmission rate of probes on other (neighbouring) road links. For this, the System Control Protocol as discussed in our previous PATH project MOU107 (Linnartz, Westerman and Hamerslag, 1994) is used. In this way, the verification mechanism has more section-related measurements at its disposal.

Furthermore, for ATMIS purposes, the time needed to detect an incident with a certain certainty should be as short as conceivable. Hence, the Application Data Protocol discussed in our previous PATH project MOU107 (Linnartz, Westerman and Hamerslag, 1994) could be adjusted in such a way that when a probe vehicle experiences an extremely significant deviation, i.e. an increase or decrease in the speed of the probe vehicle of at least a%, with say α=25, it instantaneously dispatches its traffic message. In this way, the mean time-to-detect can be kept as short as possible.

Hybrid Fixed Detector Algorithm

This algorithm comprises a more comprehensive interpretation of the rather basic fixed detector algorithm in the trigger step. The main purpose of this hybrid fixed detector algorithm is to

[16] For instance: After 1 minute the probability that an incident has occurred on a particular road link is 15%. This will intensify the stand by in the Traffic Management Center. After 2 minutes the probability of occurrence of an incident on the regarding road link has run up to 45% and diversion scenarios are computed and prepared. After 3 minutes the probability that an incident has actually occurred has increased to 70% and optional ATMS measures ate deployed. When after 5 minutes the probability amounts to 90%, compulsory ATMS measures are brought into action. Eventually, after 7 minutes, complete certainly can be assumed to have been achieved that there has occurred an incident on the particular road link and emergency vehicles are dispatched.

verify the occurrence of the incident indicated by this trigger mechanism.

The hybrid fixed detector algorithm may consist of a repetition of the California algorithm, since a 'real' incident is characterized by the relative spatial occupancy difference to remain large for some time. In this way the persistent version of the California algorithm results (Stephanedes, Chassiakos, and Michalopoulos, 1991). Besides the occupancies $o_{A,B}$, the infrastructure based traffic detectors measure the speeds $v_{A,B}$ as well as the flows $q_{A,B}$. The hybrid fixed detector should also comprise tests using these local-related quantities. For this, we the Speed-Density Difference formulated for the Siemens fuzzy logic AID system (Busch and Ghio, 1994) might be adopted. We would like to emphasize that the fixed detector AID tests to be used within the hybrid probe vehicle algorithm should be capable of recognizing a persisting incident, which is not a customary characteristic of tests within conventional AID algorithms.

Hybrid Probe Vehicle Algorithm

The purpose of the hybrid probe vehicle algorithm is to (automatically) verify the occurrence of an incident on the regarding road link by computing the road link mean speed using the increased amount of transmitted probe vehicle data together with the measured fixed detector data, and evaluating it. According to the relation between the prior knowledge and the actual flow rate described in sub-section 4.1.2, one measured flow value q produces two possible values for the historic mean speed $\bar{\bar{v}}(q)$, viz. $\bar{\bar{v}}_f(q)$ and $\bar{\bar{v}}_c(q)$. As the verification mechanism is only activated when congestion is suspected we obviously will use $\bar{\bar{v}}_c(q)$ in this procedure. Subsequently, the computed road link mean speed is evaluated. The objective of this evaluation is to determine whether veritably a significant decrease in the road link mean speed is likely to have occurred (using a time-series comparative approach) and subsequently to determine whether veritably a non-recurrent incident is likely to be involved (using a pattern recognition comparative approach), since the definition of an incident that we have adopted excludes recurrent congestion. In the case the outcome of the first test is positive, an incident is surmised to have occurred and the second test is performed. In the case this second test is also positive, a non-recurrent incident is surmised. In the other cases a false alarm is signified. Additionally, the exact contents of the probe vehicle reports are examined in order to determine whether these give cause for the occurrence of an incident too.

Time-Series Approach

The time-series comparative approach compares the estimated value of the road link mean speed, $\bar{v}^*(t)$, to the estimation from the time period just before the incident alarm was signified, $\bar{v}^*(t_0 - 1)$. The purpose of this comparative test is to check whether the estimated road link mean speed is significantly lower than the road link mean speed at the moment just before the trigger mechanism observed the incident. For this reason, the comparison concerns the random variables $\bar{v}(t)$ and $\bar{v}(t_0-1)$ more than their expected values and the comparative test involves checking whether it is significantly possible that $\bar{v}(t)$ is smaller than $\bar{v}(t_0-1)$. Since we know through our speed

model the (posterior) probability (and cumulative) distribution functions of $\bar{v}(t)$ and $\bar{v}(t_0-1)$, we can derive the concerned probability as follows:

$$P\{\, \bar{v}(t) < \bar{v}(t_0-1) \,\} =$$

$$= \int_0^\infty P\{\ V(f)'\ \bar{v}(t_0-1) \mid \bar{v}(t_0-1) = a \,\}\, P\{\, \bar{v}(t_0-1) = a \,\}\, da =$$

$$= \int_0^\infty P\{\ \bar{v}(t) < \text{cl} \,\}\, P\{\, \bar{v}(t_0-1) = \text{a} \,\}\, da.$$

In this derivation it is assumed that the random variables $\bar{v}(t)$ and $\bar{v}(t_0-1)$ are stochastically independent. This is plausible since the state of traffic after the trigger alarm differs significantly from that before the alarm, since the occurrence of the incident will perturb the traffic flow.

Direct comparison of the probability calculated above, with a prechosen level of significance α leads to the desired conclusion: if the probability is greater (less) than α, the current road link mean speed is considered (not) to be significantly smaller than the one just before the incident alarm was signified. Generally, threshold a will be rather large, for instance *90%,* thus guaranteeing that if it is decided that $\bar{v}(t) < \bar{v}(t_0-1)$, this is significantly true and the chances of making an incorrect decision (error of commission) are low".

Pattern Recognition Approach

When the time-series comparative approach has indicated that the actual (estimated) value of the road link mean speed is significantly higher than before the trigger alarm, the pattern recognition comparative approach is activated, which compares the estimated value of the road link mean speed, $\bar{v}^*(t)$, to the expected, historic value of this road link mean speed, $\bar{\bar{v}}(t)$. The difference between this expected, historic road link mean speed $\bar{\bar{v}}(t)$ and the prior knowledge $\bar{\bar{v}}_c(t)$ used for the estimation itself is that the latter represents the expected mean speed given the fact that congestion is currently present, while the first-mentioned value represents the normal, every day mean speed. This normal mean speed may either arise due to every day congestion, in which case $\bar{v}^*(t)$ and $\bar{\bar{v}}_c(t)$ are equal and indicate a recurrent incident, or may arise due to regular free-flow traffic, in which case $\bar{v}^*(t)$ and $\bar{\bar{v}}_c(t)$ are different and indicate a non-recurrent incident. So, the pattern recognition comparative test is performed in order to classify the conjectured incident as a recurrent or non-recurrent incident. More specifically, it is checked whether the actual road link mean speed $\bar{v}(t)$ is significantly smaller than its 'normal' level ($\bar{\bar{v}}(t)$).

---

[17] Note that the test analogy to the proposed procedure is that of testing $H_0: \bar{v}(t) = \bar{v}(t_0-1)$ versus $H_1: \bar{v}(t) < \bar{v}(t_0-1)$.

The check can be carried out by constructing a (single-sided) confidence interval for $\bar{v}(t)$ of the form $(-\infty ; U]$: given the desired level of significance α, the interval's upper bound U is determined from the formula

$$P_c\{\bar{v}(t) \leq U \mid v_1,...,v_{N_p}\} = \int_{0^+}^{U} P_c\{\bar{v} \mid v_1,...,v_{N_p}\}\, d\bar{v} = \alpha$$

(the formula for the posterior probability density $P_c\{\bar{v} \mid v_1, \ldots, v_{N_p}\}$ is given in appendix B). Hence, the chosen value of parameter a (for instance $\alpha = 0.90$) is the probability that the actual road link mean speed is less than or equal to U. The resulting decision follows from a direct comparison of U and $\bar{\bar{v}}(t)$: if $\bar{\bar{v}}(t)$ lies in the constructed confidence interval ($\bar{\bar{v}}(t) \leq U$), it is not significantly larger than the current road link mean speed and therefore the conjectured incident is classified as **recurrent**[18]. Since the described test is performed only after a trigger alarm, the traffic is always assumed to be in a state of congestion. At high values of $\bar{\bar{v}}(t)$, the conjectured incident is generally concluded to be a non-recurrent circumstance. However, when the historically expected mean $\bar{\bar{v}}(t)$ is low, the traffic on the considered road link at **the** considered time is 'normally' congested and the decision to be made here, is whether the seriousness of the (conjectured) incident is around (recurrent incident) or above (non-recurrent incident) the usual level.

Composition of Probe Reports

The third test of the hybrid probe vehicle algorithm is activated when the actual (estimated) value of the road link mean speed indicates that non-recurrent congestion is involved, that is when both the time-series comparative test and the pattern recognition comparative test were **affirmative,** and is based on the composition of the probe reports.

The probe vehicles are assumed to register the location at which a significant deviation in the experienced speed (Linnartz, Westerman and Hamerslag, 1994) occurs. Hence, a probe report will contain more locations (and corresponding travel times between these locations) under conditions of unstable, congested, traffic than during undisturbed, free-flow, traffic. In the previous chapter we have mentioned several approaches dealing with this subject, such as the Quality of Flow parameter Q (Greenshields, **1955),** the Acceleration Noise parameter (Chandler, Herman and Montroll, 1958) and the Coefficient of Variation of Speed (Chang and Herman, 1978). Departing from the exact contents of the probe vehicle reports, we learn that a modification of the Q parameter best fits our purpose. In order to remain in accordance with the specified exact contents of the probe reports, we prefer to use alterations in the speed of the probe vehicles that are relative to the present speed instead of absolute speed changes as were used for the original Q factor. The eventual value of this Q factor is to be compared with a pre-determined threshold, distinguishing between regular and deviant fluctuations in the individual

[18] Note that the test analogy to the proposed procedure is that of testing $H_0: \bar{v}(t) = \bar{\bar{v}}(t)$ (recurrence) versus $H_1: \bar{v}(t) < \bar{\bar{v}}(t)$ (non-recurrence).

(probe) vehicle speeds. The practical value of this threshold needs to be determined by extensive empirical field testing, for instance as is described in (Sermons and Koppelman, 1995).

When also this test is affirmative, the outcome of the hybrid probe vehicle algorithm confirms the occurrence of a (non-recurrent) incident.

Compare Local-Related Estimations and Section-Related Estimutions

The third module of the proposed verification mechanism is based on directly comparing the local-related quantities measured by the infrastructure based traffic detectors (induction loops) and the section-related quantities measured by the non-infrastructure based traffic detectors (probe vehicles). For this, we *estimate local-related* quantities from *measured section-related* quantities and vice versa.

Local-Related Estimations

In chapter 2 we have established that fluctuations in the traffic flow occur at various levels of scale and that meso-level fluctuations tend to be preserved over several kilometers when the traffic is undisturbed. This property can be deployed for directly comparing section- and local-related quantities in order to observe an incident.

Considering a road link bounded by measuring sites *A* and *B*. Every time interval t the fixed induction loop detectors at these measuring sites measure the occupancy $o_A(t)$ and $o_B(t)$. The traffic flow propagates with a speed that is assumed to be equal to the mean speed $\bar{v}(t)$ on the regarding road link (with length L), obtained from the probe vehicles. From this, the travel time $\bar{\tau}(t)$ on the regarding road link is directly given, using $\bar{\tau}(t) = L/\bar{v}(t)$. This means that the occupancy that is measured at measuring site *A* at point in time $t - \bar{\tau}(t)$ will also be measured at measuring site *B* at point in time t, provided that the traffic flow can propagate undisturbedly. When a discontinuity in the traffic flow occurs (for instance due to an incident), both occupancies will differ significantly.

Hence, two test can be formulated which check the absolute and relative difference between both occupancies[19]

$$o_A(t-\bar{\tau}_{AB}(t)) - o_B(t) \geq X_1 \quad \textit{and} \quad \frac{o_B(t)}{o_A(t-\bar{\tau}_{AB}(t))} \geq X_2$$

When both relations exceed the thresholds X_1 and X_2 (that should be determined empirically) an incident is likely to have occurred.

[19] Similar tests can be performed using the density $k_A(t)$ and $k_B(t)$ and the flow $q_A(t)$ and $q_B(t)$ measured by the induction loop detectors at both measuring sites.

A similar approach has first been formulated in the stream-discontinuity-model in (Sakasita and May, 1975), using only induction loop data and short inter-detector spacings. Here, we wish to relate induction loop data and probe vehicle data, so we use the travel time obtained from the probe vehicles instead. In this way, the' local-related quantities that are measured by the fixed detectors located at two successive measuring sites are directly related by means of the section-related quantities obtained through the probe vehicles.

Section-Related Estimations

Also in chapter 2 we have established that the mean speed $v_{AB}(t)$ over a road link with a considerable length L_{AB} (of 5 to 10 kilometers) determined from the local-related speeds $v_A(t)$ and $v_B(t)$ measured by the fixed detectors that are located at sites *A* and *B* at the boundaries of this road link, and is computed by

$$v_{AB}(t) = \frac{v_A(t) + v_B(t)}{2}$$

is an appropriate *qualitative* indication of the actual (mean) speed on the regarding road link. In the case of free-flowing traffic the actual (mean) speed and the speed computed by means of the above formula are almost equal. In the case of congestion both speeds will differ. Hence, when the above formula is compared with road link mean speed $\bar{v}(t)$ it can be established whether the status of the traffic flow on the regarding road link is free-flow.

So, the outcome of the test

$$| \bar{v}(t) - v_{AB}(t) | > Y_1$$

with Y_1 a threshold that should be determined empirically, indicates whether congestion is likely to have occurred on the regarding road link. In this way, the local-related quantities measured by the fixed detectors and the section-related quantities obtained through the probe vehicles are directly compared.

When both tests (i.e. section-related and local-related estimations) are positive, the outcome of this third component of the verification mechanism is also positive.

Combined Probability of Occurrence of an Incident

The last step of the proposed verification mechanism conjoins the three components (i.e. the hybrid probe vehicle algorithm, the hybrid fixed detector algorithm and the comparison process, which can be considered to be mutually independent) in order to determine the overall likelihood that an incident has occurred on a particular road link. This overall likelihood is used to verify the suspicion of occurrence of an incident that has resulted from the first step (the trigger) and may be repeated several times in order to reach a desired level of certainty about the occurrence of an incident on the regarding road link, each with a corresponding time-to-detect.

From the (deterministic) outcomes of the three components, a probability of occurrence of an incident can be obtained by taking into account which outcomes are positive (i.e. indicate an incident) and which are negative, as well as how many times the verification step has already been performed

$$P\ (Incident) = \frac{i}{a}\ [x_1 C_1 + x_2 C_2 + x_3 C_3]$$

with C_k the outcome of component k, C_k=1 (0) when the outcome is positive (negative), x_k the mutual weight-factor of component k, $\sum_{k=1}^{3} x_k = 1$, i an index variable indicating how many times the verification step has already been performed, and a the verification-factor, indicating the maximum number of verification steps (after α times affirmative verifications the probability of occurrence of an incident is regarded to be 1).

In figure 4.5 it is indicated how the probability of occurrence of an incident resulting from each verification step can be operated with.

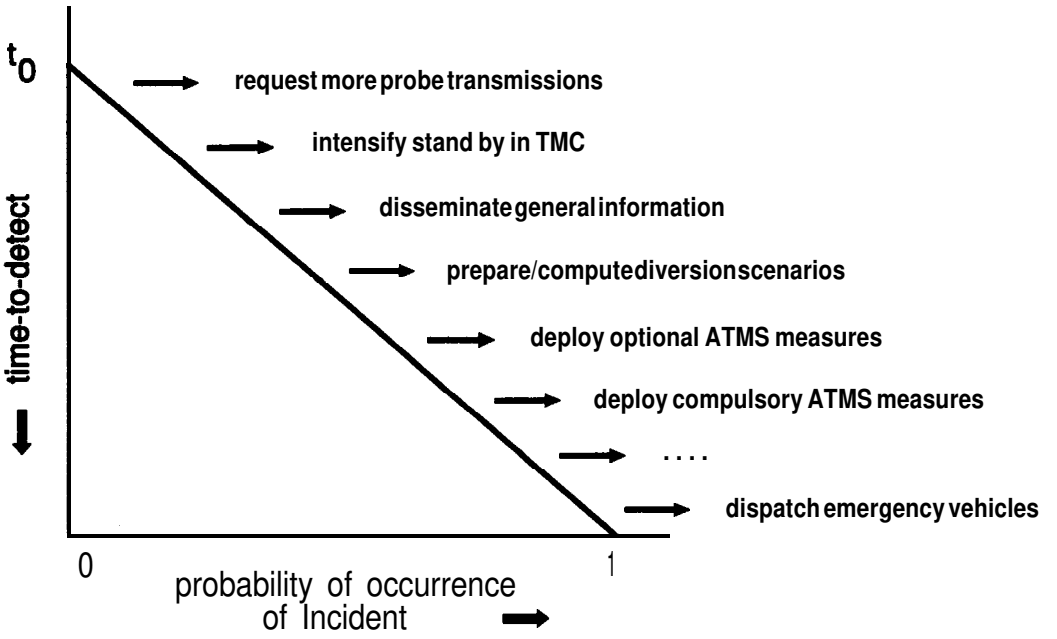

*Figure 4.5 Processing the outcome **of** the verification mechanism*

When the trigger mechanism has indicated that an incident might have occurred on a certain road link in a certain time period, more frequent probe transmission are requested from the concerning road link and the transmission rate of probes on neighbouring links can be restrained if necessary. Subsequently, the verification mechanism is performed (utilizing the more frequent probe transmissions) and in the case its outcome does not contradict the trigger outcome,

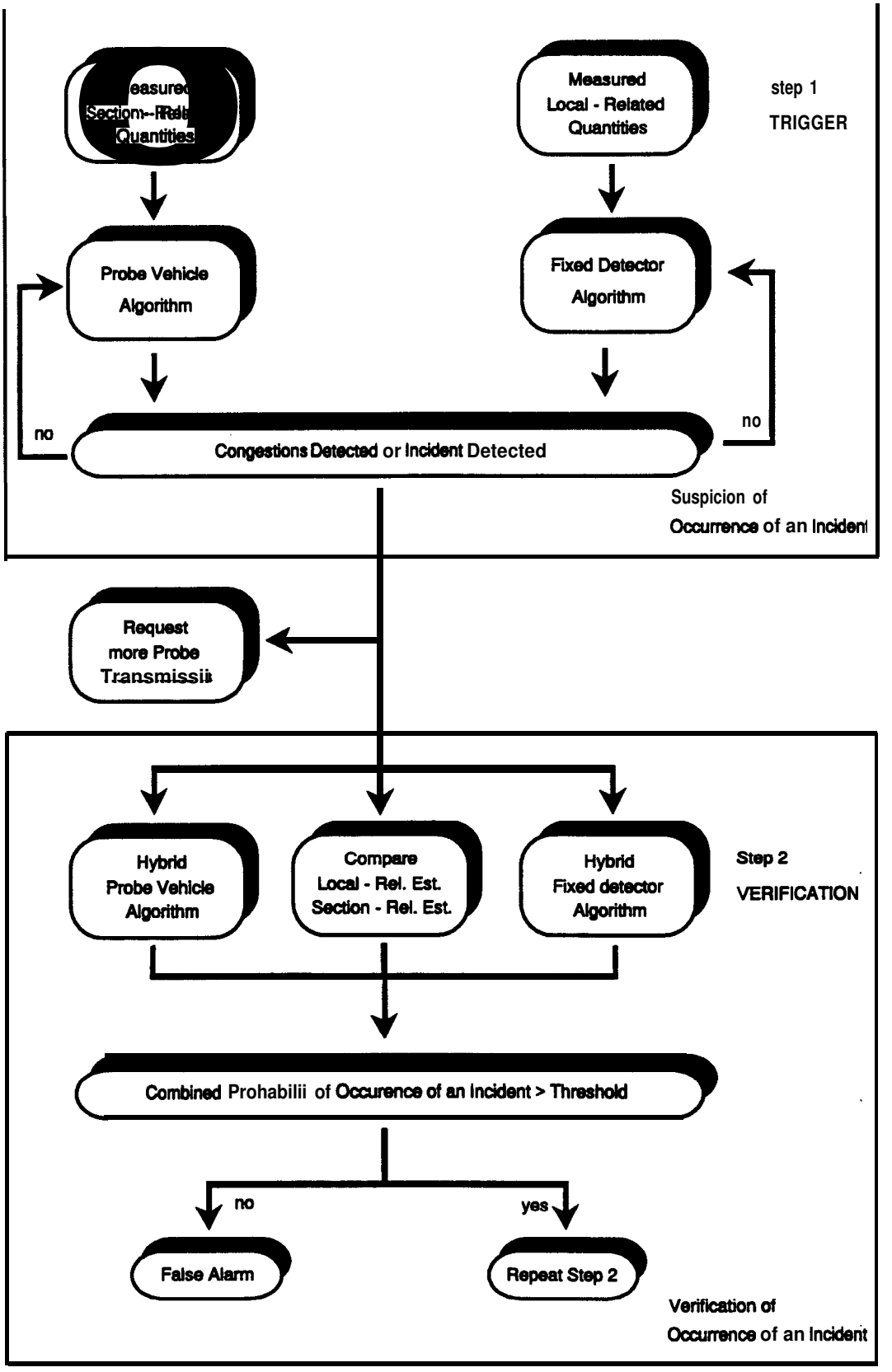

Figure 4.6 Generic structure of a compound probe vehicle and induction loop AID system

equipment and personnel in the traffic center should intensify their stand by. When the next that time the verification mechanism is performed it still affirms the occurrence of an incident, the probability that veritably an incident has occurred increases and general information can be disseminated, for instance by means of radio traffic information. After the next time of performing the verification mechanism with a affirmative outcome, diversion scenarios can be prepared and computed. In order to gain more certainty about the occurrence of an incident the verification mechanism can be performed more times, and optional (e.g. variable message signing) and compulsory ATMIS measures (e.g. route diversions) can be deployed. Eventually, when the verification mechanism keeps conforming the persistence of the conjectured incident, emergency vehicles can be dispatched. It should be apparent that this strategy requires a deliberation between a high (low) probability of occurrence of an incident and a rather long (short) time-to-detect.

In this sub-section we have presented fundamental principles and sketched a generic structure of a possible compound probe vehicle and induction loop monitoring system for Automatic Incident Detection, indicating how data from non-infrastructure based probe vehicle can be incorporated in order to enhance the performance of AID based on data from infrastructure based induction loop detectors alone. Since information deduced by considering several independent data sources is likely to be more accurate and reliable than information from one single source (Luo and Kay, 1991), fusion of probe and loop data is critical for such a compound AID system. The proposed generic structure comprises a trigger mechanism and a verification mechanism, and is sketched in figure 4.6. Although different structures are possible, from the ATMIS requirements with respect to AID and the potentialities of fixed detector and probe vehicle data, the proposed outline is considered the most obvious. With respect to processing data from fixed traffic detectors we propose to make use of existing procedures as much as possible as these have often been utilized and have demonstrated their capabilities on short road sections.

The probe vehicle components are incorporated in order to enhance their performance at long road sections. Most of these probe vehicle components were already developed and analyzed in previous parts of this dissertation, while we have addressed novel components of the presented compound algorithm in more detail in this sub-section.

Based on the presented principles of Automatic Incident Detection using combined probe vehicle and induction loop data, a compound algorithm should be implemented which should be tested and evaluated, either by means of a simulation study or by means of a real-life pilot. Simulation of incidents and their microscopic effects on the propagation of the traffic flow requires such an extensively and detailed study that this is regarded to be beyond the scope of this research. This applies for performing a pilot study in an increasing degree. For this reason, we have restricted the research described in this section to indicating the general principles of such a compound AID concept.

4.4 Discussion

In particular obtaining reliable and accurate real-time travel times as well as performing adequate incident management and automatic incident detection becomes considerably

complicated when infrastructure based traffic detectors with rather large mutual inter-detector spacings (of about 5 to 10 kilometers) are deployed. In this chapter we have addressed these two issues and have analyzed how data obtained through probe vehicles can be included into the induction loop data in order to enhance the reliability and accuracy of the process of deriving relevant information for ATMIS.

In the first part of this chapter we have formulated two approaches for fusing probe vehicle and induction loop data for estimation real-time travel times or road link mean speeds.

The first approach assumes the availability of a reasonable dense network of fixed (loop) detectors. Departing from this assumption, an algorithm has been formulated for estimating real-time travel times and road link mean speeds using induction loop data and additional probe vehicle data. The performance of the developed compound algorithm has been analyzed by simulating probe vehicles and the experienced traffic data obtained by them driving over a road link bounded by two induction loop measuring sites. This performance analysis has shown that the obtained real-time travel times are more accurate and, above all, more reliable compared to similar results using only induction loop data.

The second approach assumes that only sporadically fixed (loop) detectors are available in the road network. Departing from this assumption, an algorithm has been formulated for estimating real-time road link mean speeds and travel times using probe vehicle data and additional induction loop data. The performance of the developed compound algorithm has been analyzed by assuming realistic parameter values. This performance analysis has shown that the quality of the obtained estimations depends on the actual traffic volume, and that at high flow rates achieving a certain accuracy and reliability would require more probe messages to be successfully received than at low flow rates. This is due to the fact that the free-flow and congested traffic regimes converge under conditions of high flow rates. However, this convergence process also implies that the distinction between utilizing the free-flow or congested prior knowledge in the estimator becomes increasingly insignificant and the demands with respect to reliability and accuracy of the estimate should also decrease with increasing flow rate. Taking into account these assertions the results of the algorithm using probe vehicle and additional loop data are more in correspondence with reality than when only probe data is used and still, only few probe reports (2 to 3) are sufficient for accurately and reliably estimating road link mean speeds and travel times.

In the second part of this chapter we have indicated the importance of appropriate Incident Management and Automatic Incident Detection for ATMIS, and have concisely reviewed existing methods and their properties for automatic incident detection. Based on these experiences and on the obtained properties of non-infrastructure based probe vehicles and infrastructure based induction loop detectors throughout this dissertation, we have formulated fundamental principles of fusing both data sources in order to perform adequate AID for ATMIS. This has resulted in a generic structure of a compound probe vehicle and induction loop AID system. We have deliberately excluded implementing and testing the formulated generic AID system in the second part of this chapter, since this would require extensive microscopic simulation or a real-life pilot project, which are both considered to be principally beyond the scope of this dissertation.

From the results obtained in this chapter we learn that fusing probe vehicle data and induction loop data can considerably enhance the quality of derived ATMIS traffic information. Hence, constructing a combined induction loop and probe vehicle monitoring system for. ATMIS is indisputably highly advisable.

5. Conclusions and Recommendations

5.1 Conclusions

In this research report we have addressed the issue of obtaining relevant ATMIS information from measured road traffic data. We have discerned three possible approaches for monitoring road traffic flows for ATMIS, namely infrastructure based induction loop detectors with large mutual distances, probe vehicles acting as non-infrastructure traffic detectors, and a combination of probe vehicles and induction loop detectors. The motivation for studying methods for processing induction loop and probe vehicle data alone before addressing the issue of fusing both data sources, is that few sound theories exist for processing elementary into relevant ATMIS information from each of these single data sources. Integration of both data sources is only useful provided that their collective performance exceeds the performance of both data collection techniques on their own.

Infrastructure based induction loop detectors
For measuring real-time flows and time-mean speeds an infrastructure based traffic monitoring system based on induction loop detectors is preeminently suitable. For obtaining real-time travel times and performing adequate Automatic Incident Detection (AID) their network-wide characteristics, and thus the elongated road links, are less commendatory.

We have formulated an algorithm for estimating real-time travel times using an infrastructure based induction loop monitoring system with large detector spacings. This algorithm is based on simple counting (by matching cumulative flow distributions of successive measuring sites) in combination with correlation of fluctuations in the traffic flow (the re-calibration process). Calibration of the implemented computer model COMETT using factual traffic data, has shown to produce reliable and accurate real-time travel times both in the case of free flow traffic and congested traffic, for elongated road links (typically 5 to 10 kilometers) including several on/off ramps. Road links with lengths of more than about 10 kilometers should be avoided and heavily occupied ramps should be measured. Although the computed travel times are suitable for ATMIS, the algorithm remains somewhat susceptible to the significant loop measuring errors, and when congestion persists longer than about two hours the accuracy of the computed travel times starts to diminish. To a certain extent, these errors are compensated for by (refinements in) the re-calibration process, but additional re-calibrations would be required for achieving a more persistent performance.

From literature regarding existing AID algorithms we learn that performing reliable AID for ATMIS is unfeasible when using infrastructure based traffic detectors with mutual distances longer than a few (about 5) kilometers. For this reason we have not attempted to address the subject of performing automatic incident detection using infrastructure based traffic detectors with large mutual spacings.

Non-infrastructure based probe vehicles

In order to obtain real-time travel times or road link mean speeds from received probe vehicle reports, the most straightforward technique, that is simply averaging the probe reports, is shown to be very arduous when only few reports are received. For this reason, we have designed a method based on a multi-layered speed model comprising a micro-level representing the individual probe vehicle speeds, a meso-level representing the actual road link mean speed, and a macro-level representing the historic mean speed. The comprehensive estimator consists of two steps: Automatic Congestion Detection (ACD) and estimation of the road link mean speed. Quantitative analyses showed that when only 1 (2) probe reports are received the probability of detection delaying congestion amounts to 79% (85%), both at a probability of mis-detection of only 1%. Based on established results for the second step, we conclude that for obtaining real-time travel times (or mean speeds) on a freeway with an accuracy of about 85% and a reliability of about 85%, approximately 3-5 probe reports per period are required, provided that a database with proper historic traffic data is available. From the study of communication aspects we have learned that these requirements correspond to a probe vehicle penetration grade of about 1% when periods of about 10-15 minutes are concerned and a penetration grade of about 5% when periods of l-5 minute are required (departing from a freeway network during peak hours; lower classes of roads or times of the day outside peak hours impose higher probe vehicle penetration grades).

We found that accurate real-time traffic flows and traffic densities can not be reliably obtained exclusively from probe vehicles. For determining both traffic variables for longer time periods (e.g. several hours or days) probe vehicles are preeminently suitable, in particular for secondary and urban roads where, in general, no infrastructure based traffic detectors are present.

Integration of induction loop and probe vehicle data

We have addressed two approaches for fusing probe vehicle and induction loop data in order to estimate real-time travel times.

On the one hand we have developed a method for estimating real-time travel times from induction loop data and additional probe vehicle data. This method is suitable for road links where relatively many infrastructure based induction loop detectors are present, such as most major freeways. This method is constructed by incorporating additional probe vehicle data into the COMETT algorithm that has been designed for induction loop data alone. By means of simulated probe vehicle updates the extended COMETT model produces travel times that are more accurate and, above all, more reliable compared to similar results using only loop data.

On the other hand, when no reasonably dense network of fixed (loop) detectors is available, real-time travel times for ATMIS can be estimated using probe vehicle data and additional induction loop data. For this, we have extended the algorithm formulated for probe vehicle data alone to incorporate flow rates obtained through induction loop measurements. This provides travel time estimates that are more realistic and more appropriate compared than the ones obtained from probe vehicle data alone, still requiring only few probe vehicle samples (2 to 3).

With respect to performing Automatic Incident Detection using combined probe vehicle and

induction loop data, we have formulated fundamental principles of a compound probe vehicle and induction loop AID system, in which incorporation of probe vehicle data lowers the false alarm rate and the time to detect. This has resulted in a generic structure comprising several components. We have extensively addressed and specified each of these components.

Overall Conclusion
The research described in this PATH report has yielded theoretical and practical insight into the issues of how data delivered from non-infrastructure based probe vehicles concerning spatial fluctuations in the traffic process can be included to enhance the reliability and accuracy of travel time estimates and AID performance based on data measured by traditional infrastructure based (induction loop) detectors concerning temporal fluctuations in the traffic process.

From the results obtained in this PATH research report we learn that fusing probe vehicle data and induction loop data can considerably enhance the quality of derived ATMIS traffic information. By implementing the data processing techniques described in this PATH research report adequate real-time road traffic information for ATMS can be attained. Hence, constructing a combined induction loop and probe vehicle monitoring system for ATMIS is indisputably recommended.

5.2 Suggestions for Further Research

This section proposes further research that either logically follows from the established findings in this report and concerns issues that require additional research or that have been not been specifically addressed here but deserve consideration.

Construction of a Historic Database
The construction of an appropriate historic database with characteristic traffic parameters, such as link travel times or speeds and link densities, deserves to receive proper attention in prospective traffic engineering applications. We have shown that producing estimates of traffic flow parameters is considerably more accurate and reliable when correct historic traffic data can be incorporated. Moreover, such a database can be exploited for making swift predictions, seeing that under regular conditions an extrapolation of the historic means is generally more precise than an advanced prediction tool.

Probe vehicles are preeminently suitable for constructing such an historic database as, in principle, they achieve continuous road network coverage and accurately register both temporal and spatial fluctuations in the traffic process. With sufficient prorogation even the dynamics in the infrastructure can be observed accurately. This is in contrast to traditional fixed detectors that only register traffic at one specific location. In the Berlin LISB one of the earliest experimental examples of such a database with historic probe vehicle data was constructed.

Profound Traffic Engineering Theories
Traditional traffic engineering modeling, techniques and theories are more based on pragmatism and realism than on theoretical correctness. Yet, various emerging advanced telematics appli-

cations, such as processing probe vehicle data into relevant traffic flow inforrnation and detecting traffic flow disturbances, require profound traffic flow analyses for which commonly established and correct theories are essential. Hence, the goal of achieving an optimal traffic process, e.g. by utilizing transportation telematics, demands more fundamental and theoretical traffic engineering research. This research should at least address the behavior of and the inter-relationships between the fundamental traffic flow parameters of flow, density and speed (Banks, 1995).

*Data Processing as Part **of** Field Trials*

In current field trials the concept of probe vehicles constitutes part of a more comprehensive dynamic route guidance system. This causes the attention in such field trials to be predominantly focused on technical issues, while data processing is only simplistically addressed. Based on the (theoretical) research in this PATH research report we propose to consider probe vehicle data processing as a more intrinsic component and to evaluate various data processing techniques, including those based on combined probe vehicle and induction loop data. The capability of probe vehicle data to enhance information estimates from induction loop data should be assessed as well.

Executing a Probe Vehicle Concept

Implementation of a probe vehicle concept is technologically feasible. Except, perhaps, for a suitable data communication infrastructure, all technological components are practical. In order to break through the current impasse of organization (who will serve as operator, administrator and financier) and introduction (road users will only purchase when they regain correct and relevant information while obtaining such information requires sufficient operational systems) a cooperative initiative should be established, consisting of, for instance, a car manufacturer (to supply data collectors and to raise sales), a service provider (to supply a communication infrastructure and to be able to provide a profitable service), and an electronic industry (to supply equipment and to obtain appropriate information). In order to encourage introduction and utilization as well as to avoid undesirable developments the government should actively participate, by means of providing financial or strategic commitment or additional (traffic) data and a technological infrastructure.

*Organizational and Institutional Aspects **of** a Probe Vehicle Concept*

In this dissertation we have demonstrated how ATMIS and road users in general can benefit when data from non-infrastructure based probe vehicle and data from infrastructure based traffic detectors are combined. Exactly how such an organizational or institutional integration between data from infrastructure based and non-infrastructure based traffic detectors can be established requires more research and elaborate consultations, while in particular consensus' should be reached about issues of who will become 'owner' of the probe vehicle data and how to deal with the possibly conflicting objectives of ATMIS and the road users. Above all, it should be realized that a constructive cooperation of organizations involved in ATMIS is vital for introduction of integrated ATMIS and is in the interest of all.

6. Literature

Ahmed, S.A. and A.R. Cook, "Analysis of Freeway Traffic Time-Series Data by using Box-Jenkins Techniques", *Transportation Research Record, No. 722, 1979, pp. 1-9.*

Ahmed, S.A. and A.R. Cook, "Time Series Model for Freeway Incident Detection", *ASCE Journal of Transportation Engineering, Vol. 106, No. 6, November 1980, pp. 731-745.*

Albright, D., "Development of National Highway Traffic Monitoring Standards", *Presented at the 70th Transportation Research Board, Washington D.C., January 1991, Preprint. Paper No. 910050.*

Arem, B. van, M.J.M. van der Vlist, J.C.C. de Ruiter, M. Muste and S.A. Smulders, "Design of the Procedures for Current Capacity Estimation and Travel Time and Congestion Monitoring", *GERDIEN, DRIVE-II Project Number: V2044, Deliverable 9, Workpackage SP6.Wp2 & SP6. WP3, Technical Annex 2 (revised), September, 1994.*

Arem, B. van, M.J.M. van der Vlist, J.C.C. de Ruiter, M. Muste, S.A. Smulders, M.S. Dougherty, M.R. Cobett and H.R. Kirby, "Evaluation Results of the Models for Monitoring the Current and Expected Traffic State", *GERDIEN, DRIVE-II Project Number: V2044, Deliverable 16, Workpackage SP6.Wp3, September, 1994.*

Bakkenist, "Project Definition Monitoring System" (In Dutch: Projectdefinitie Monitoringsysteem), *Bakkenist Management Consultants, 1989.*

Banks, J.H., "Another Look at A Priori Relationships among Traffic Flow Characteristics", *Presented at the 74th Annual Meeting of the Transportation Research Board, Washington D.C., January 1995, Preprint Paper No. 9501 IO.*

Bhandari, N., F.S. Koppelman, J.L. Schofer, V. Sethi and J.N. Ivan, "Arterial Incident Detection Integrating Data from Multiple Sources", *Presented at the 74th Transportation Research Board, Washington D.C., January 1995, Preprint Paper No. 950122.*

Blosseville, J.M., V. Motyka and S. Espie, "Video sensors", *In: Papageorgiou, M. (eds.), "Concise Encyclopedia of Traffic and Transportation Systems", Pergamon Press, ISBN O-08-036203-6, 1991, pp. 611-616.*

Böhnke, P. and E. Pfannerstill, "A System for the Automatic Surveillance of Traffic Situations", *ITE Journal, Vol. 1, 1986.*

Bolte, F., "Monitoring Traffic in the Whole Motorway Network: What Effects can be Expected?", *Federal Highway Research Institute (BASt), Research Report, Germany, 1991.*

Brandston, D., "Link Capacity Functions: a Review", *Transportation Research, Vol. 10, 1976.*

Brocken, M.G.M. and F. de Haes, "Estimating Travel Times on Highways" (In Dutch: Schatten van Reistijden op Autosnelwegen), *Proceedings of Colloquium Vervoersplanologisch Speurwerk, The Hague, November, 1990.*

Bush, F., "Automatische Stiirungserkennungen auf Schnellverkehrsstrassen - ein Verfahrensvergleich", *Ph.D Thesis, Technical University of Karlsruhe, 1986.*

Bush, F., "Incident Detection", *In: Papageorgiou, M. (eds.), "Concise encyclopedia of traffic and transportation systems", Pergamon Press, 1991, ISBN o-08-036203-6, pp. 219-225.*

Bush, F,. and M. Fellendorf, "Automatic Incident Detection on Motorways", *Traffic Engineering and Control, April 1990, pp. 221-227.*

Bush, F. and A. Ghio, "Automatic Incident Detection on Motorways by Fuzzy Logic", *Paper presented at the International Conference Traffic and Transport Solutions, Amsterdam, 1994.*

Chandler, R.E., R. Herman and E.W. Montroll, "Traffic Dynamics: Studies in Car Following", *Operations Research, Vol. 6, March/April, 1958, pp. 165-184.*

Chang, M. and R. Herman, "An Attempt to Characterize Traffic in Metropolitan Areas", *Transportation Science, Vol. 12, No. 1, 1978, pp. 59-79.*

Cheu, R.L., S.G. Ritchie, W.W. Recker and B. Bavarian, "Investigation of a Neural Network Model for Freeway Incident Detection", *California PATH Research Report, UCB-ITS-RR-88-10, Institute of Transportation Studies, University of California, Berkeley, 1988.*

Cleghom, D., F.L. Hall and D. Garbuo, "Improved Data Screening Techniques for Freeway Traffic Management Systems", *Transportation Research Record, No. 1320, 1991.*

Cook, A.R. and D.E. Cleveland, "Detection of Freeway Capacity Reducing Incidents by Traffic-Stream Measurements", *Transportation Research Record, No. 495, 1974, pp. l-11.*

Courage, K.G. and M. Levin, "A Freeway Corridor Surveillance Information and Control System", *Texas Transportation Institute, Texas A&M University, College Station, Research Report 488-8, December, 1968.*

Cremer, M., "Flow variables: Estimation", *In: Papageorgiou, M. (eds.), "Concise encyclopedia of traffic and transportation systems", Pergamon Press, 1991, ISBN o-08-036203-6, pp. 143-148.*

Dailey, D.J., "Travel-Time Estimation using Cross-Correlation Techniques", *Transportation Research B, Vol. 27, No. 2, 1993.*

Donk, P. and A. Versluis, "Quality Survey of the ES06-Traffic Registration" (In Dutch: Kwaliteitsonderzoek van de ES06-Verkeersregistratie), *Transportation and Traffic Research Division, Rijkswaterstaat, 1985.*

Duckworth, G.L., J. Bing, S.H. Carlson, M.L. Frey, M.A. Hamilton, J.C. Heine, S.D. Milligan, R. Mlawski, C.E. Remer, S. Ritter, L.R. Warner, B.G. Watters, R.H. Welsh and D. Whittemore, "A Comparative Study of Traffic Monitoring Sensors", *Proceedings of the IVHS America 1994 Annual Meeting, Atlanta, Georgia, April, 1994, pp. 283-293.*

Dudek, C.L., C.J. Messer and J.D. Friebele, "Investigation of Lane Occupancy as a Control Variable for a Safety Warning System for Urban Freeways", *Texas Transportation Institute, Texas A & M University, College Station, March, 1973.*

Dudek, C.L., C.J. Messer and N.B. Nuckles, "Incident Detection on Urban Freeways", *Transportation Research record, No. 495, 1974, pp. 12-24.*

Federal Highway Administration (FHWA), "Future National Highway Program Task Force", *Urban and Suburban Highway Congestion, Working Paper No. 10, Washington D.C., 1987.*

FHWA, "Traffic Monitoring Guide", *Federal Highway Administration, Washington, D. C., 1985.*

Forbes, G.J., "Identifying Incident Congestion", *Institute of Transportation Engineers Journal, Vol. 6, No. 62, 1992, pp. 17-22.*

Ford, L.R., "Network Flow Theory", The *Rand Corporation, 1956.*

Fumess, K.P., "Time Function Iteration", *Traffic Engineering and Control, Vol. 7, 1965.*

Gall, A.I. and F.L. Hall, "Distinguishing between Incident Congestion and Recurrent Congestion: a Proposed Logic" *Transportation Research Record, No. 1232, 1989, pp. l-8.*

Greenshields, B .D., "A Study in Highway Capacity", *Proceedings of the Highway Research Board, Vol. 14, 1935, pp. 448-477.*

Greenshields, B., "Quality of Traffic Transmission", *Highway Research Board Proceedings, Vol. 34, 1955, pp. 508-522.*

Grontmij, "Investigation Travel Times on Highways" (In Dutch: Onderzoek Reistijden op Autosnelwegen), *Department of Infrastructure, The Netherlands, February, 1990.*

Gunter, M. and F. Hall, "Transitions in the Speed-Flow Relationship", *Transportation Research Record, Vol. 1091, 1986.*

Hall, F. and M. Gunter, "Further Analysis of the Flow-Concentration Relationship", *Transportation Research Record, Vol. 1091, 1986.*

Hamerslag, R. and W.J. Kribbe, "A Proposal for a Travel-Time-Prediction-System", *5th Word Conference on Transport Research, Yokohama, Japan, 1989.*

Han, L.D. and A.D. May, "Artificial Intelligence Approaches for Urban Network Incident

Detection and Control", *Proceedings of the Engineering Foundation Conference on Traffic Control Methods, Santa Barbara, 1989.*

Hilgers, C.J., "A Method for Calculating a Group of Algorithms for Automatic Incident Detection", *Rijkswaterstaat, The Hague, The Netherlands, 1980.*

Hoffman G. and J. Janko, "Travel Times as a Basic Part of the LISB Guidance Strategy", *Paper presented at the IEEE Road Traffic Control Conference, London, 1990.*

House, N., M.A. Vicencio and F. Bessaguet, "INVAID Type B Processor: The Use of Rules Based Techniques to Detect Traffic Incidents from Qualitative Traffic Data", *Advanced Telematics in Road Transport, Proceedings of DRIVE Conference, Vol. 1, Brussels, Belgium, February, 1991.*

Hounsell, N. and S. Ishtiaq, "Journey Time Forecasting for Dynamic Route Guidance Systems", *Proceedings of the Second DRIVE-II Workshop on Short-Term Traffic Forecasting, Delft, November, 1994, pp. l-18.*

IDL - Interactive Data Language, *Mathematics Guide Version 3.6.1, Research Systems Inc., Boulder, Colorado, May, 1994A.*

IDL - Interactive Data Language, *User's Guide Version 3.6.1, Research Systems Inc., Boulder, Colorado, May, 1994B.*

Jacobson, L.N., N.L. Nihan and J.D. Bender, "Detecting Erroneous Loop Detector Data in a Freeway Traffic Management System", *Transportation Research Record, No. 1287, 1990.*

Kell, J.H., I.J. Fullerton and M.K. Mills, "Traffic Detector Handbook", *U.S. Department of Transportation, Federal Highway Administration, Office of Research and Development, Report No. FHWA-IP-90-002, July, 1990.*

Kirschfink, H., "SIC - Service & Information Center: Collective and Individual Traffic Information", *Brochure Heusch Boesefeldt, Traffic Control Engineering/Systems - Traffic Modelling Division, Aachen, 1994.*

Klijnhout, J.J. and H.C.G. Langelaar, "Motorway Control and Signalling", *Traffic Engineering and Control, 1987.*

Kroes, J.L. de and J.J. Reijmers, "Precise Measurements of Speed and Length of Vehicles using Induction Loops: Trapezium Method", *Journal A, Vol. 21, No. 3, 1980, pp. 104-I 11.*

Koutsopoulos, H.N. and H. Xu, "An Information Discounting Routing Strategy for Advanced Traveler Information Systems", *Transportation Research C, Vol. 1, No. 3, 1993, pp. 249-264.*

Leutzbach, W. and F. Bush, "Störfallentdeckkungen", *Research Project, Federal Ministry of Transport, Karlsruhe, Germany, 1983.*

Levin, M. and G.M. Krause, "Incident Detection: A Bayesian Approach", *Transportation Research Record, No. 682, 1978, pp. 52-58.*

Levin, M. and G.M. Krause, "A Probabilistic Approach to Incident Detection on Urban Freeways", *Traffic Engineering and Control, Vol. 20, No. 3, March 1979, pp. 107-109.*

Levine, S.Z. and W.R. McCasland, "Monitoring Freeway Traffic Conditions with Automatic Vehicle Identification Systems", *Journal of the Institute of Transportation Engineers, March, 1994, pp. 23-28.*

Linden, R.J.P. van der and J.A. Valk, "Detection of Traffic-Jams" (In Dutch: File-Detectie), *Institute Information-Technology for Production Automation (TNO-ITP)/Technical University of Eindhoven, The Netherlands, 1990.*

Linden, R.J.P. van der and R.M.F. Driessen, "Methods for Detection of Situations of Constrained Traffic Flow" (In Dutch: Methoden voor Detectie van Situaties van Gedwongen Verkeersafwikkeling), *TNO-IPL, The Netherlands, June, 1991.*

J.P.M.G Linnartz, M. Westerman, R. Hamerslag, "Monitoring the San Francisco Bay Area Freeway Network using Probe Vehicles and Random Acces Radio Channel", *California PATH Research Report UCB-ITS-PRR-94-23, October, 1994.*

Luo, R.C. and M.G. Kay, "Multisensor Integration and Fusion in Intelligent Systems", *IEEE Transactions on Systems, Man, and Cybernetics", Vol. 19, No. 5, 1991, pp. 901-931.*

Masters, P.H., J.K. Lam and K. Wong, "Incident Detection Algorithms for COMPASS - An Advanced Traffic Management System", *Proceedings of the Vehicle Navigation and Information Systems Conference, Dearborn, Michigan, October 1991, pp. 295-310.*

MARGOT, "Modelling and Assessing Route Guidance of Traffic: Improved Journey Time Prediction", *MARGOT Deliverable 6014, DRIVE-II Project V2033, 1994.*

Michalopoulos, P.G., "Incident Detection Through Video Image Processing", *Proceedings of the Second International Conference on Applications of Advanced Technologies in Transportation Engineering, Minneapolis, Minnesota, 1991.*

Mohamed, S.A. and A.R. Cook, "Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques", *Transportation Research Record, No. 722, 1979, pp. l-9.*

Motyka, V. F. Lenoir, J.M. Blosseville and N. Djemame, "Motorway Type A processor", *Final Report V1026 INVAID - Integration of Computer Vision Techniques for Automatic Incident Detection, INRETS, France, September 1991.*

OECD, "Intelligent Vehicle Highway Systems: Review of Field Trials", *Road Transport research, 1992.*

Okutani, I. and Y.J. Stephanedes, "Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory", *Transportation Research B, Vol. 18, No. 7, 1984, pp. l-1 1.*

Papageorgiou, M., "Freeway", *In: Papageorgiou, M. (eds.), "Concise Encyclopedia of Traffic and Transportation Systems", Pergamon Press, ISBN o-08-036203-6, pp. 148, 1991.*

Pfannerstill, E., "Automatic Monitoring of Traffic Conditions by Re-Identification of Vehicles", *Proceedings of the Second IEEE International Conference on Road Traffic Monitoring and Control, London, February, 1989, pp. 172-l 75.*

Payne, H.J., E.D. Helfenbein and H.C. Knobel, "Development and Testing of Incident Detection Algorithms: Vol. 2 - Research Methodology and Detailed Results", *Report No. FHWA-RD-76-20, Federal Highway Administration, Washington D. C. April 1976.*

Payne, H.J and S.C. Tignor, "Freeway Incident - Detection Algorithms Based on Decision Trees with States", *Transportation Research Record, No. 682, 1978, pp. 30-37.*

Persaud, B.N., F.L. Hall and L.M. Hall, "Catastrophe Theory and Patterns in 30-Second Freeway Traffic Data Implication for Incident Detection", *Transportation Research A, Vol. 23, No. 2, 1989, pp. 103-113.*

Persaud, B.N., F.L. Hall and L.M. Hall, "Congestion Identification Aspects of the McMaster Incident Detection *Algorithm", Transportation Research Record, No. 1287, 1990.*

Philipps, P., M. Hoops, K. Ibbeken and G. Riegelhuth, "The RHAPIT Field Trial and Evaluation of SOCRATES Applications", *Proceedings of the First World Congress on ATT and IVHS, Paris, December, 1994, pp. 2297-2304.*

Pursula, M. and I. Kosonen, "Microprocessor and PC-Based Vehicle Classification Equipments using Induction Loops", *Journal of the Institute of Transportation Engineers, January, 1987, pp. 24-28.*

Reijmers, J.J., "Incident Detection on Highways", *Department of Electrical Engineering, Telecommunications and Traffic-control, TU Delft, The Netherlands, 1980A.*

Reijmers, J.J., "On-Line Vehicle Classification", *IEEE Transactions on Vehicular Techno-logy, Vol. VT-29, No. 2, 1980B.*

Ritchie, G.S, R.L. Cheu and W.W. Recker, "Freeway Incident Detection using Artificial Neural Networks", *Engineering Foundation Conference, Ventura, California, June 1992.*

Ritchie, S.G. and R.L. Cheu, "Simulation of Freeway Incident Detection using Artificial Neural Networks", *Transportation Research C, Vol. 1, No. 3, 1993, pp. 203-217.*

Ruiter, J.C.C. de, "The Highway Monitoring System" (In Dutch: Het HWN-Monitoringsysteem), *Transportation and Traffic Research Division, Rijkswaterstaat, 1989.*

Ruiter, J.C.C. de, "Travel Time Determination in Road Traffic: Overview and Assessment" (In Dutch: Reistijdbepaling in het Wegverkeer: Overzicht en Beoordeling), *Internal concept note, October, 1994.*

RWS, "Verkeersgegevens: Jaarrapport 1994" (In Dutch: Traffic Data: Annual Report 1994), *Ministry of Transport and Public Works, DG Rijkswaterstaat, 1994.*

Sakasita, M. and A.D. May, "Development and Evaluation of Incident-Detection Algorithms for Electronic-Detector Systems on Freeways", *Transportation Research Record, No. 533, 1975, pp. 48-63.*

Schaefer, W.E., "California Freeway Surveillance System", *California Division of Highways, Department of Public Highways, November 1969.*

Schuurman, H. and R.G.G.M. Vermijs, "Level of Service at Freeways" (In Dutch: Verkeersafwikkeling bij Discontinuiteiten), *Transportation Research Laboratory, Report TU Delft, VK 2205.309, ISSN LW 0920-0592, Delft, 1993.*

Sermons, M.W., F.S. Koppelman, "Assessing the Effectiveness of Individual Vehicle Movement Characteristics in the Detection of Roadway Incidents", *Presented at the Transportation Research Board, Washington D.C., January 1995, Preprint Paper No. 950892.*

Stephanedes, Y.J. and A.P. Chassiakos, "A Low Pass Filter for Incident Detection", *Proceedings of the Second International Conference on Applications of Advanced Technologies in Transportation Engineering, Minneapolis, Minnesota, 1991, pp 378-382.*

Stephanedes, Y.J. and A.P. Chassiakos, "Freeway Incident detection Through Filtering", *Transportation Research C, Vol. 1, No. 3, 1993, pp. 219-233.*

Stephanedes, Y.J., A.P. Chassiakos and P.G. Michalopoulos, "Comparative Performance Evaluation of Incident Detection Algorithms", *Transportation Research Record, No. 1360, 1991, pp. 50-57.*

Stoelhorst, H. and J.C.C. de Ruiter, "Memo: Travel Time Information Road Traffic Freeway Network - Concept Note" (In Dutch: Memo: Reistijdinformatie Wegverkeer HWN - Concept *Notitie), Internal RWS Memo, October, 1994.*

TRANSCOM, "TRANSCOM ETTM IVHS Operational Field Test - Final Feasibility Report", *TRANSCOM Contract No. XCM 92-090.01, prepared by Farradyne Systems, Inc., Edwards and Kelcey, Inc., Eng-Wang, Taub and Associates and Castle Rock Consultants, January, 1993.*

Transpute, "Continuation of the Research for Estimation of Delay Times and Congestion Size with Traffic Monitoring" (In Dutch: Vervolgonderzoek voor Schatten van Vertragingstijden en File-Omvang bij Verkeersmonitoring), *Report Prepared for the Transportation and Traffic Engineering Division, Ministry of Transport and Public Works, Report No. CXT91068.RAP, April, 1992.*

Tsai, J. and E.R. Case, "Development of Freeway Incident-Detection Algorithms by Using Pattern Recognition Techniques", *Transportation Research Record, No. 722, 1979, pp. 113-l 16.*

Velzen, G.A. van, "Dynamic Monitoring of the Freeway Network: Automatic Incident Detection" (In Dutch: Dynamische Monitoring van het Hoofdwegennet: Automatische Incident Detectie), *Grontmij Advies & Techniek, Zeist, October, 1993.*

Virginia Tech, "Incident Detection Issues", *Research Proposal DTFH61-92-R-00122 submitted to Federal Highway Administration by Virginia Polytechnic Institute and State University in association with among others Delft University of Technology, 1992.*

Wardrop, J.G., "Some Theoretical Aspects of Road Traffic Research", *Proceedings of the Institute of Civil Engineers, Vol. II, No. 1, 1952, pp. 325-378.*

Weits, E.A.G., "A Stochastic Heat Equation for Freeway Traffic Flow", *Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1990.*

Westerman, M. and L.H. Immers, "A Method for Determining Real-Time Travel Times on Motorways", *Proceedings of the 25th ISATA Conference, Florence, Italy, June, 1992, pp. 221-*228.

Westerman, M., "Estimation of Travel Times for ATMS and ATIS", *Proceedings of the 24th UTSG Conference, Leeds, England, 1993.*

Westerman, M., "Real-Time Data Collection for Dynamic Road Traffic Management", *Proceedings of the 1st TRAIL Ph.D Congress, Rotter&m, May, 1995.*
This paper has been awarded as best conference paper.

Westerman, M., "Inwinning en Verwerking van Dynamische Verkeersgegevens" (In Dutch: Collecting and Processing Dynamic Traffic Data), *Delft University of Technology, Faculty of Civil engineering, forthcoming, 1995.*

Willigen, D. van and J.J. Reijmers, "Traffic Guidance Systems" (In Dutch: Verkeersbegeleidingssystemen), *Lecture Notes, Delft University of Technology, Faculty of Electrical Engineering, Department of Telecommunication and Traffic Guidance Systems, 1992.*

Willsky, AS, E.Y. Chow, S.B. Gershwin, C.S. Green, P.K. Houpt and A.L. Kurkjian, "Dynamic Model-Based Techniques for the Detection of Incident Freeways", *IEEE Transactions on Automatic Control, Vol. AC-25, No. 3, June 1980, pp. 347-359.*

Whitson, R.H., J.H. Buhr, D.R. Drew and W.R. McCasland, "Real Time Evaluation of Freeway Quality of Traffic Service", *Highway Research Record, Vol. 289, 1969.*

Ziemer, R.E, H.W. Tranter and D.R. Fannin, "Signals and Systems: Continuous and Discrete", *Macmillan Publishing Co., Inc, New York, ISBN o-02-431650-4, 1983.*

Appendix A Definition of Road Link Mean Speed

In the main document of this PATH research report the issue of deducing relevant traffic information from received probe vehicle data is addressed. With respect to obtaining information about the actual (mean) speed of the traffic flow we have developed an estimator for estimating the so-called 'road link mean speed' from probe vehicle samples. In this appendix we will investigate the relation between this road link mean speed and the mean speeds that are traditionally discerned in traffic engineering.

In traffic engineering two distinct methods for averaging individual vehicle speeds are commonly distinguished and each of these leads towards a different mean, namely the time-mean-speed and the space-mean-speed (see for instance (May, 1990)). The time-mean-speed at a certain location during a certain time period is obtained by averaging the individual vehicle speeds measured at that particular location over that selected time period, viz.

$$\bar{v}_{TMS} = \frac{\sum_{i=1}^{n} v_i}{n}$$

The space-mean-speed at a certain point in time over a certain (part of a) road link is obtained by first converting the measured individual vehicle speeds into individual travel times (in the expression below expressed in hours per kilometer), then computing an average travel time by averaging the individual travel times, and finally computing the average speed at that particular point in time over that particular (part of a) road link, viz.

$$\bar{v}_{SMS} = \frac{1}{1/n \sum_{i=1}^{n} 1/v_i}$$

The space-mean-speed reveals spatial fluctuations in the traffic flow over an entire road link and is specifically convenient for localization of disturbances in the traffic process. However, obtaining the space-mean-speed is difficult. One possible way of measuring $\bar{v}_{SMS}$ would be by means of consecutive aerial photographs, which of course is not very realistic on a network-wide scale. The time-mean-speed is measured at one specific location and is therefore more convenient. However, $\bar{v}_{TMS}$ has the disadvantage that it only reveals temporal fluctuations at the concerning crosscut and neglects differences in the traffic conditions at the remaining part of the road link. Furthermore, fast vehicles tend to predominate the average. The latter is often compensated by harmonically averaging the individual vehicle speeds measured by infrastructure based traffic detectors (and thus approaches the space-mean-speed), but this still has the drawback of just concerning the location of the measuring site. This is particular-y critical under congested traffic conditions.

ATMIS applications require real-time insight in the spatial as well as the temporal fluctuations

in the actual traffic process. In theory, this could be achieved by continuously (in order to obtain temporal fluctuations) taking aerial photographs of each road link in the road network (in order to obtain $\bar{v}_{SMS}$ indicating spatial fluctuations) or by installing infrastructure based traffic detectors (in order to obtain $\bar{v}_{TMS}$ indicating temporal fluctuations) with very short mutual spacings on each road link of the road network (in order to obtain spatial fluctuations). Both approaches are obviously far from realistic. Fortunately, the nature of the probe vehicle concept makes this novel monitoring technique preeminently suitable for obtaining both temporal and spatial fluctuations in the traffic flow over a complete (longitudinal section of a) road link. For this purpose we define the road link mean speed $\bar{v}$ as the speed obtained by averaging over all individual probe vehicles, averaging over the complete road link with length D (of 5-10 kilometers) as well as averaging over the entire time period with length T (of l-5 minutes). The road link mean travel time τ subsequently is defined as the time needed to traverse to complete road link under consideration with a constant speed that equals the road link mean speed $\bar{v}$. Thus, this road link mean speed can be regarded as a specific combination of both the time-mean-speed and the space-mean-speed.

Having introduced the traditional traffic engineering speed concepts, we will further study the relation between the road link mean speed obtained by the probe vehicles, the space-mean-speed and the time-mean-speed by defining $k(v)$ to denote the expected number of probe vehicles in an area of unity size driving at speed v. Hence, the dimension of $k(v)$ is number of vehicles per kilometer. The total number of probe vehicles in this area, the total probe vehicle density k_{tot}, is given by

$$k_{tot} = \int_0^\infty k(v)\, dv$$

As defined earlier, the space-mean-speed at a certain time (in an unity space area, e.g. of 1 kilometer) conveys the averaged speed of all probe vehicles that are present in this area at that time. The space-mean-speed can be expressed in terms of $k(v)$ by aggregating the product of all probe vehicle densities k(v) and their speed v over all possible speeds, and dividing (averaging) this aggregate by the total number of probe vehicles present in this space area (k_{tot}), viz.

$$\bar{v}_{SMS} = \frac{\int_0^\infty v\, k(v)\, dv}{\int_0^\infty k(v)\, dv} \quad \left(= \int_0^\infty v \left(\frac{k(v)}{\int_0^\infty k(v)\, dv} \right) dv = \int_0^\infty v\, p(v)\, dv = E\, v \right)$$

In the same way, the time-mean-speed at a certain location during time interval T can be expressed in terms of $k(v)$. For this, we write the total number of probe vehicles driving at speed v and passing the considered location during time period T as

$$N(v) = k(v)\, v\, T$$

The time-mean-speed conveys the product of the number of these probe vehicles $N(v)$ and their speed v, aggregated over all possible speeds v, and divided by (averaged over) the total number of probe vehicles that passes the location during time period T, viz.

$$\bar{v}_{TMS} = \frac{\int_0^\infty v\, N(v)\, dv}{\int_0^\infty N(v)\, dv}$$

Substituting $N(v) = k(v)\, v\, T$ gives

$$\bar{v}_{TMS} = \frac{T \int_0^\infty v\, k(v)\, v\, dv}{T \int_0^\infty k(v)\, v\, dv} = \frac{\int_0^\infty v^2\, k(v)\, dv}{\int_0^\infty v\, k(v)\, dv}$$

Using statistical equality $E\{v^2\} = \{Ev\}^2 + \sigma_v^2$ and dividing both numerator and denominator by k_{tot}, this relation can be rewritten as

$$\frac{\int_0^\infty v^2\, k(v)\, dv}{\int_0^\infty v\, k(v)\, dv} = \frac{\int_0^\infty v^2\, k(v)\, dv \;\; \int_0^\infty k(v)\, dv}{\int_0^\infty k(v)\, dv \;\; \int_0^\infty v\, k(v)\, dv} = \frac{E\{v^2\}}{Ev} = \frac{\{Ev\}^2 + \sigma_v^2}{Ev} = Ev + \frac{\sigma_v^2}{Ev}$$

Hence, using the previously deduced formula of the space-mean-speed, the time-mean-speed is given by

$$\bar{v}_{TMS} = \bar{v}_{SMS} + \frac{\sigma_v^2}{\bar{v}_{SMS}}$$

in which σ_v^2 is the variance in the individual probe vehicle speeds. From this relation it becomes clear that the space-mean-speed and the time-mean-speed can only be equal then and only then when σ_v^2 is equal to zero, that is, only when all probe vehicles in the traffic stream travel at the same speed. This relation between the time-mean-speed and the space-mean-speed has also been shown in (Wardrop, 1952).

Now, we will express the road link mean speed, obtained from the probe vehicles, in terms of $k(v)$. For this, we define a time and space window as sketched in figure A. 1.

Every probe vehicle that occupies this window, i.e. travels on a part of the road link with length *D* during time period *T,* contributes to the road link mean speed. Hence, the expected total number of such probe vehicles that occupy this time-space window and are traveling with speed v is given by *k(v) [D + v T] (see* figure A.l).

Consequently, the road link mean speed is computed using

$$\bar{v}_{probe} = \frac{\int_0^\infty v\ k(v)[D + vT]\ dv}{\int_0^\infty k(v)\ [D + vT]\ dv}$$

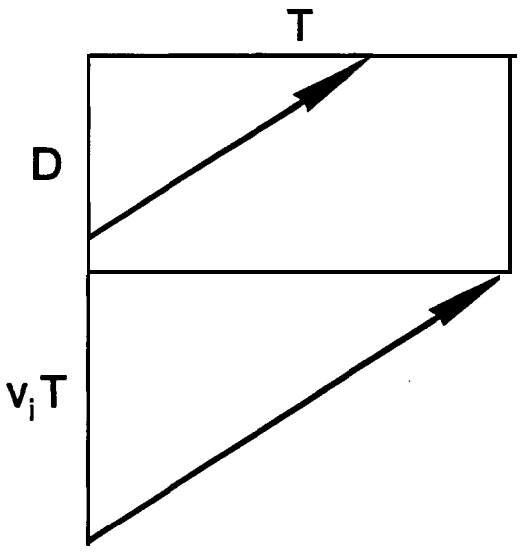

Figure A.1 Time-space window for computing the road link mean speed

This relation can be rewritten by substituting the formerly derived formula for the space-mean-speed and dividing each term by the total probe vehicle density k_{tot}, viz.

$$\bar{v}_{probe} = \frac{D\int_0^\infty v\,k(v)\,dv + T\int_0^\infty v\,k(v)\,v\,dv}{D\int_0^\infty k(v)\,dv + T\int_0^\infty k(v)\,v\,dv} = \frac{D\frac{\int_0^\infty v k(v) dv}{\int_0^\infty k(v)\,dv} + T\frac{\int_0^\infty v^2 k(v)\,dv \int_0^\infty v\,k(v)\,dv}{\int_0^\infty k(v)\,dv \int_0^\infty v\,k(v)\,dv}}{D\frac{\int_0^\infty k(v)\,dv}{\int_0^\infty k(v)\,dv} + T\frac{\int_0^\infty v\,k(v)\,dv}{\int_0^\infty k(v)\,dv}}$$

Hence, the road link mean speed obtained from the probe vehicles is given by

$$\bar{v}_{probe} = \frac{D\ \bar{v}_{SMS} + T\,\bar{v}_{TMS}\ \bar{v}_{SMS}}{D + T\ \bar{v}_{SMS}}$$

When the derived formula for the time-mean-speed is substituted in the above formula, the following direct relation between the road link mean speed and the time-mean-speed results

$$\bar{v}_{probe} = \frac{D\,\bar{v}_{SMS} + T\,(\bar{v}^2_{SMS} + \sigma^2)}{D + T\,\bar{v}_{SMS}}$$

From this direct relation it becomes clear that when we would compute the road link mean speed from the probe vehicles for an infinitely small time interval (i.e. $T \approx 0$), the space-mean-speed would result. Likewise, when we would compute the road link mean speed for an infinitely small space interval (i.e. $D \approx 0$), *the* time-mean-speed would result. Hence, what has been formulated as ideal for both ATIS and ATMS purposes, namely a continuous determination of the space-mean-speed or a determination of the time-mean-speed at a string of locations, can, to a certain extent, be achieved by the probe vehicle concept.

References

May, A.D., "Traffic Flow Fundamentals", *Prentice Hall, New Yersey, 1990, ISBN 0-13-926072-2.*

Wardrop, J.G., "Some Theoretical Aspects of Road Traffic Research", *Proceedings of the Institution of Civil Engineers Vol. II, No. 1, 1952, pp. 325-378.*

Appendix B Derivation of Estimator using Historic Data for Congested Traffic

In the main document of this PATH research report the derivation of the (Bayes) estimator for estimating the road link mean speed from received probe vehicle samples incorporating historic traffic data under conditions of congested traffic has been presented concisely for reasons of comprehensibility. This appendix presents the derivation of this estimator for congested traffic more extensively.

Posterior Distribution for Congested Traffic

Here we will determine $\bar{v}$'s posterior distribution, the update of the prior distribution, given the sample of received individual probe vehicle speed data, needed in order to derive the formula for the estimator and its confidence interval.

For a given value of $\bar{v}$, the posterior probability density (in the congested regime) equals:

$$P_c\{\bar{v} \mid v_1,...,v_{N_p}\} = \frac{P_c\{\bar{v}\}\, P_c\{v_1,...,v_{N_p} \mid \bar{v}\}}{P_c\{v_1,...,v_{N_p}\}} =$$

$$= \frac{P_c\{\bar{v}\}\, P_c\{v_1,...,v_{N_p} \mid \bar{v}\}}{\int_{0^+}^{\infty} P_c\{\bar{v}\}\, P_c\{v_1,...,v_{N_p} \mid \bar{v}\}\, d\bar{v}} = \frac{P_c\{\bar{v}\} \prod_{i=1}^{N_p} P_c\{v_i \mid \bar{v}\}}{\int_{0^+}^{\infty} P_c\{\bar{v}\} \prod_{i=1}^{N_p} P_c\{v_i \mid \bar{v}\}\, d\bar{v}} =$$

$$= \frac{f_{\bar{v}_c, \sigma_c^2/N_p}(\bar{v}) \prod_{i=1}^{N_p} f_{\bar{v}, \sigma^2}(v_i)}{\int_{0^+}^{\infty} f_{\bar{v}_c, \sigma_c^2/N_p}(\bar{v}) \prod_{i=1}^{N_p} f_{\bar{v}, \sigma^2}(v_i)\, d\bar{v}}$$

with $f_{\lambda,\rho}$ the probability density function of I'(I, ρ) . Here it is assumed that all v_i's are independent random variables. Used is Bayes' rule for conditional probabilities.

Because of the structure of the probability density function of the Gamma distribution, we can simplify the calculations involved somewhat by using:

$$P_c\{v_1,...,v_{N_p}|\bar{v}\} = \prod_{i=1}^{N_p} f_{\lambda,\rho}(v_i) = \left(\frac{\lambda^\rho}{\Gamma(\rho)}\right)^{N_p} \left(\prod_{i=1}^{N_p} v_i\right)^{\rho-1} e^{-\lambda \sum_{i=1}^{N_p} v_i},$$

(with $\lambda = \frac{\bar{v}}{\sigma^2}$ and $\rho = \frac{\bar{v}^2}{\sigma^2}$) and further by cancelling out identical factors in the numerator and the denumerator of the posterior distribution formula.

Estimator for Congested Traffic

The (Bayes) estimation of $\bar{v}, \bar{v}^*_B$, equals the expected value of $\bar{v}$'s posterior distribution. This estimator is preferred over the direct or Moment estimator if we have reliable prior knowledge regarding the distribution of $\bar{v}$. This prior knowledge has been specified in the defined multi-layered speed model. For this estimator the selection of the prevailing traffic region is important as it assumes a different probability distribution of $\bar{v}$, regarding both its shape and its first and second moment.

For the congestion regime the estimator is defined assuming a gamma distribution for the individual probe vehicle speeds as follows:

$$\bar{v}^*_B = E_c\{\bar{v}|v_1,...,v_{N_p}\} = \int_{0'}^{\infty} \bar{v} P_c\{\bar{v}|v_1,...,v_{N_p}\} d\bar{v}.$$

Using the formula for the posterior probability density function derived above we find:

$$\left[\int_{0^+}^{\infty} f_{\frac{\bar{v}_c}{\sigma_c^2},\frac{\bar{v}_c^2}{\sigma_c^2}}(\bar{v}) \prod_{i=1}^{N_p} f_{\frac{\bar{v}}{\sigma^2},\frac{\bar{v}^2}{\sigma^2}}(v_i)\, d\bar{v}\right]^{-1} \left[\int_{0^+}^{\infty} \bar{v} f_{\frac{\bar{v}_c}{\sigma_c^2},\frac{\bar{v}_c^2}{\sigma_c^2}}(\bar{v}) \prod_{i=1}^{N_p} f_{\frac{\bar{v}}{\sigma^2},\frac{\bar{v}^2}{\sigma^2}}(v_i)\, d\bar{v}\right]$$

again with $f_{\lambda,\rho}$ the probability density function of I'(λ, ρ) .