Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Part C 16 (2008) 18–35
TRANSPORTATION RESEARCH PART C
www.elsevier.com/locate/trc

Real-time estimation of vehicle-count within signalized links

Georgios Vigos [a], Markos Papageorgiou [a,*], Yibing Wang [b,1]

[a] Dynamic Systems and Simulation Laboratory, Technical University of Crete, 73100 Chania, Greece
[b] Institute of Transport Studies, Department of Civil Engineering, Monash University, Victoria 3800, Australia

Abstract

The number of vehicles included in a metered motorway ramp or an urban signalized link at any time is valuable information for real-time control. A Kalman-Filter is employed to produce reliable estimates of this quantity based on real-time measurements of flow and occupancy provided by (at least) three loop detectors. The resulting vehicle-count estimator is tested via microscopic simulation for a variety of metered ramp scenarios and traffic conditions. Several related fundamental issues are addressed: the effects of loop density, update period, downstream signal cycle, vehicle length and link length. The simulation investigations indicate a robust estimation performance with low calibration effort needed, which facilitates easy applicability of the method.

Keywords: Link queue; Ramp queue; Kalman-Filter; Estimation of vehicle-count

1. Introduction

Traffic-responsive control systems require reliable real-time information on the prevailing traffic conditions to make sensible control decisions. More particularly, the number of vehicles included in signalized links (such as urban road links or metered motorway ramps) is a valuable information for urban signal control and motorway ramp metering systems. Direct measurements of this quantity may in principle be collected by video sensors, but this is often difficult due to visibility obstacles, limited visible link length, image processing algorithm accuracy and, last but not least, cost.

As an alternative possibility, traditional low-cost loop detectors or other emerging magnetic sensors (Cheung et al., 2005) measuring time-occupancy may be employed. The difficulty faced by this approach is due to the strongly inhomogeneous character of the traffic state in signalized links caused by the frequent switching of the upstream and, most importantly, downstream traffic lights. Thus, a detector station (across all link lanes) positioned at a specific link location (e.g., at the signal stop line or in the middle or at the upstream end of the link) delivers (local) occupancy information that is not representative for the whole link. In other words, local

* Corresponding author. Tel.: +30 28210 37289; fax: +30 28210 37584.
E-mail addresses: gvigos@dssl.tuc.gr (G. Vigos), markos@dssl.tuc.gr (M. Papageorgiou), Yibing.Wang@eng.monash.edu.au (Y. Wang).
[1] Tel.: +61 3 9905 9339.

0968-090X/$ - see front matter
doi:10.1016/j.trc.2007.06.002

time-occupancy measurements collected by loop detectors need to be translated into space-occupancy or link vehicle-count estimates for further use.

A relatively high-cost approach, that is, e.g., practiced in some ramp metering installations in the UK, is to install a large number of loop detectors along the link (e.g., one detector every 50 m in the UK ramps). Yet another (theoretical) possibility would be to install two flow-measuring detectors at the respective extreme points of the link and deduce vehicle-counts in the link from the conservation equation; however, this approach would soon lead to unacceptably high estimation errors due to time-accumulation of the inevitable flow measurement errors.

The estimation method presented in this paper is partly based on some relationships between time- and space-occuppancy in signalized links that were derived in a preceding paper (Papageorgiou and Vigos, 2008); in addition, a Kalman-Filter is employed here to deliver reliable real-time estimates of the vehicle-count (i.e. number of vehicles included) in signalized links based on measurements of (at least) three loop detector stations located at both extreme points and at the middle of the link, respectively. This Kalman-Filter turns out to be similar to the one proposed in Bhouri et al. (1988) for the estimation of traffic density in short non-signalized motorway links.

An interesting question addressing the degree of hardware cost (for detectors, communications and maintenance) savings thanks to exploitation of low-cost algorithmic intelligence is: How many (additional) loop detectors would be necessary to reach the estimation accuracy of the proposed method? This question as well as a thorough assessment of the proposed method under several different conditions are treated via microscopic simulation.

Although the presented method can be applied equally well to urban signalized links, the reported simulation investigations were chosen to resemble to typical metered motorway ramps. Vehicle-count estimates for motorway ramps are required within ramp metering systems for:

- Efficient ramp queue control to avoid spillback in the adjacent street network in a more appropriate way than the ordinary occupancy-based queue override (Smaragdis and Papageorgiou, 2003; Sun and Horowitz, 2005).
- Efficient and equitable ramp metering coordination that may call for real-time information on the currently available ramp storage space (Kotsialos and Papageorgiou, 2005).

A different approach to vehicle-count estimation for motorway ramps using speed measurements and curve fitting to a high number of simulation data (that may be difficult to collect in the field) was proposed in Sun and Horowitz (2005).

2. Time-occupancy and space-occupancy

Papageorgiou and Vigos (2008) presented a detailed analysis and simulation-based illustration of the relationship between the easily measurable time-occupancy and the space-occupancy that will be summarized in this section as it provides a first step towards the pursued Kalman-Filter development.

The time-occupancy is measured at specific highway locations (e.g., by inductive loop detectors); more specifically, a passing vehicle j with speed y_j activates a detector for a time $t_j = L_j/y_j$, where L_j is the effective vehicle length that may be decomposed $L_j = L_j^{\mathrm{Ph}} + \varepsilon_j$, with L_j^{Ph} the physical vehicle length and ε_j the effective detector length. The time-occupancy $o_{\mathrm{t}}(k) \in [0,1]$ which is collected over an update period (or sampling time) T (e.g., $T = 10$ s) (see Papageorgiou and Vigos, 2008), verifies the relationship

$$o_{\mathrm{t}}(k) = \sum_{j=1}^{N(k)} L_j/y_j T \tag{1}$$

where $N(k)$ is the number of passed vehicles during the time-period $[kT,(k+1)T)$ and $k = 0,1,2,\ldots$ is the discrete time index. The detector also measures the traffic volume (or flow) $q(k) = N(k)/T$ (in veh/h).

A space-occupancy $o_{\mathrm{s}}(k) \in [0,1]$ is an instantaneous (for time instant kT) space-extended quantity that reflects the portion of highway-lane length that is covered by vehicles. It is shown in Papageorgiou and

Vigos (2008) that the time-occupancy is a bias-free estimate of the space-occupancy in a sufficiently small space–time window, if the effective vehicle lengths L_j equal the corresponding physical vehicle lengths L_j^{Ph}, in which case the effective detector length ε_j becomes zero, i.e. the loop detector effectively shrinks to a line; and the time-occupancy signal is activated for as long as that line is covered by a passing vehicle.

The measured time-occupancy (with $L_j = L_j^{\mathrm{Ph}}$) delivers a sufficiently accurate estimate of the space-occupancy in a reasonably dimensioned space–time window around the measurement location if traffic conditions change moderately over space and time, as e.g., on freeways. This condition, however, does not hold for signalized links where traffic conditions vary strongly in space and time due to the frequent switching of the traffic lights. In other words, the space-occupancy (or the vehicle-count) on (even short) signalized links cannot be reflected reliably via a single time-occupancy measurement at some link location.

Consider a single-lane link of length Δ and its space-occupancy o_s. Define $o_s(x)$, $0 \leqslant x \leqslant \Delta$, to be the space-occupancy at the link location x such that

$$o_s(x) = \begin{cases} 1 & \text{if } x \text{ is covered by a vehicle} \\ 0 & \text{else.} \end{cases}$$

We then have

$$o_s = \frac{1}{\Delta} \int_0^{\Delta} o_s(x)\,\mathrm{d}x. \tag{2}$$

Note that $o_s(x)$ is defined in an infinitesimal space–time window, hence $o_s(x) = o_t(x)$ (if $L_j = L_j^{\mathrm{Ph}}$) where $o_t(x)$ is the time-occupancy at x with infinitesimal update period. Consider M link detectors measuring time-occupancy $o_t(x^i)$, $i = 1, \ldots, M$, in distances d from each other, where $M = \Delta/d$, according to Fig. 1. Then, relationship (2) may be approximated via

$$o_s \approx \sum_{i=1}^{M} o_s(x^i) d/\Delta = \sum_{i=1}^{M} o_t(x^i)/M \tag{3}$$

i.e. via an average of the measured time-occupancies. Papageorgiou and Vigos (2008) analyse in some detail the level of accuracy of (3) in dependence of several contributing factors.

To start with, the approximation (3) will be increasingly accurate for decreasing update periods T and detector spacing d. This, however, may call for unreasonably high measurement cost, particularly for signalized links, therefore our attention is turned to cases of low M or even $M = 1$. Papageorgiou and Vigos (2008) argue that the placement of M (even $M = 1$) detectors according to Fig. 1 is well-suited to deliver low-biased estimates by use of (3).

Regarding the impact of the update period T, Papageorgiou and Vigos (2008) conclude that, as mentioned above, short T contribute to a lower estimation error variance if M is high; in contrast, if M is low or 1, longer update periods T contribute to a lower estimation error variance.

A so-called ZSZO (zero-speed zero-occupancy) phenomenon was identified by Papageorgiou and Vigos (2008) which occurs when two link vehicles are standing still (due to a red traffic signal) just upstream and downstream, respectively, of a detector, without activating the detector. In this randomly appearing case, the detector measures a zero value for time-occupancy while the real traffic situation is quite the opposite. It was found that the ZSZO-phenomenon does not affect the expected value (bias) of the error resulting from (3) but may increase the error variance, particularly for $M = 1$.

The Kalman-Filter development pursued in this paper for estimating the space-occupancy and vehicle-count on signalized links starts with the approximate relationship (3) with $M = 1$. The aim of the Kalman-Filter is to reduce the estimation error variance that would result from a direct application of (3).

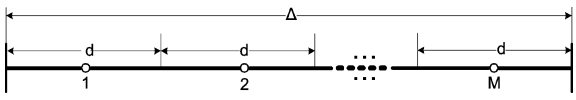

Fig. 1. Placement of M internal detectors for time-occupancy.

3. Kalman-filter development

3.1. Problem statement

The vehicle-count estimation problem is illustrated in Fig. 2. Fig. 2a depicts the relevant signal and detector configuration on the signalized link. Two traffic signals are located, respectively, just upstream of the upper boundary of the link and at the downstream end of the link. The upstream signal (if it exists) determines the traffic demand approaching the link while the downstream signal controls the vehicle flow exiting the link. Obviously, whenever the link demand is larger than the link outflow, a queue is built. It is also shown in Fig. 2a that three detectors are installed: at the upstream end of the link, at the downstream end of the link, and in the middle of the link. Both boundary detectors provide flow measurements, while the middle detector provides time-occupancy measurements. The basic structure of the ramp queue estimation is shown in Fig. 2b:

- The estimator is fed in real-time (every T) with the flow and time-occupancy measurements from the link detectors.
- The estimator delivers in real-time (every T) the estimated number of vehicles in the link (between the two boundary detectors).

3.2. Modelling for estimation

An appropriate state-space model and a measurement model are needed for application of the Kalman-Filter. The state-space model is the conservation-of-vehicles equation in the link while the measurement model is based on the insights gained in Papageorgiou and Vigos (2008) and summarized in Section 2.

The evolution of the number of vehicles in a link obeys the following conservation equation:

$$N(k) = N(k-1) + T[q_{\text{in}}(k-1) - q_{\text{out}}(k-1)] \tag{4}$$

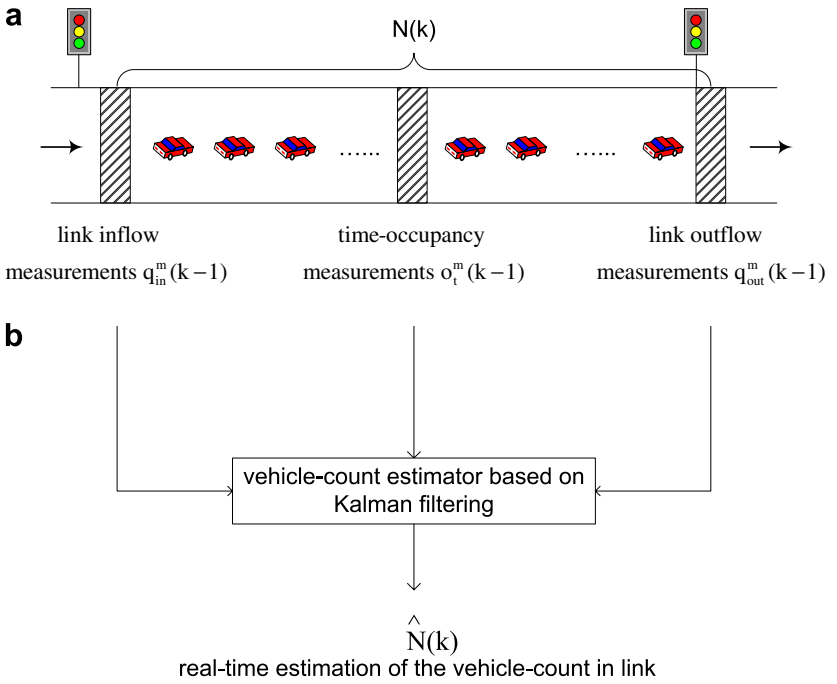

Fig. 2. Vehicle-count estimation: (a) the signal and detector configuration on the link; (b) the link vehicle-count estimation.

where $N(k)$ denotes the number of vehicles in the link at time kT and T is the measurement and estimation update period (or sampling time); $q_{\text{in}}(k-1)$ and $q_{\text{out}}(k-1)$ are the flows entering and exiting, respectively, the link during the period $[(k-1)T, kT)$. These flows are measured by the boundary detectors

$$q_{\text{in}}^{\text{m}}(k-1) = q_{\text{in}}(k-1) + \gamma_{\text{in}}(k-1) \tag{5}$$

$$q_{\text{out}}^{\text{m}}(k-1) = q_{\text{out}}(k-1) + \gamma_{\text{out}}(k-1) \tag{6}$$

where q_{in}^{m}, $q_{\text{out}}^{\text{m}}$ are the related measurements while $\gamma_{\text{in}}^{\text{m}}$, $\gamma_{\text{out}}^{\text{m}}$ denote the corresponding measurements noise and are assumed to be zero-mean stochastic variables (see Section 3.3 for biased measurements). The measurement noise may realistically be modeled in proportion to the related flow value rather than independent thereof; this, however, would render the resulting state equation nonlinear and would call for application of the more complex Extended Kalman-Filter. Preliminary simulation investigations indicated that the higher level of modeling realism does not lead to substantially higher estimation accuracy; hence it was decided to consider $\gamma_{\text{in}}^{\text{m}}$ and $\gamma_{\text{out}}^{\text{m}}$ in (5) and (6) as random variables with constant variance that is independent of the value of the measured flows. Replacing (5), (6) in (4) yields

$$N(k) = N(k-1) + T[q_{\text{in}}^{\text{m}}(k-1) - q_{\text{out}}^{\text{m}}(k-1)] + T\gamma(k-1) \tag{7}$$

where $\gamma = \gamma_{\text{out}} - \gamma_{\text{in}}$. Although, in principle, the conservation equation with inflow and outflow measurements could be directly used for estimating the vehicle-count $N(k)$, Eq. (7) reveals that such a procedure would accumulate the unavoidable measurement noise γ leading to increasingly inaccurate estimates. Therefore, more information is necessary to counter the accumulation of measurement noise in the state equation (7).

The required additional information may be provided by the middle detector in Fig. 2a. According to Eq. (3) for $M = 1$, the time-occupancy $o_{\text{t}}^{\text{m}}(k-1)$ collected by this detector during $[(k-1)T, kT)$ may be related to the link's space-occupancy $o_{\text{s}}(k-1)$ at time $(k-1)T$

$$o_{\text{t}}^{\text{m}}(k-1) = o_{\text{s}}(k-1) + \zeta_0(k-1) \tag{8}$$

where the random variable ζ_0 incorporates several potential sources of error:

- Detector measurement noise: As for the flow measurement noise above, this noise will be considered independent of the measured occupancy value in order to obtain a simpler estimation algorithm.
- Modelling error due to the approximative character of (3), particularly for small M. Note that this error may also include some (small) bias. Finally, this error also includes the impact of the update period T and of the ZSZO-phenomenon outlined in Section 2.
- Error due to the effective vehicle lengths being different from the physical vehicle lengths while measuring time-occupancy (see Section 2).

In order to relate the time-occupancy and the vehicle-count N in the link, we observe that, by definition

$$o_{\text{s}} = (NL^{\text{Ph}})/\Delta\lambda \tag{9}$$

where λ is the number of lanes in the link. In the light of (8), (9) we may define the "measured" vehicle-count

$$N^{\text{m}} = \frac{\Delta\lambda}{L^{\text{Ph}}} o_{\text{t}}^{\text{m}} \tag{10}$$

and replacing (9), (10) in (8) we finally obtain

$$N^{\text{m}}(k-1) = N(k-1) + \zeta(k-1) \tag{11}$$

where $\zeta = \zeta_0 \Delta\lambda / L^{\text{Ph}}$. Note that the transformation of the collected time-occupancy measurement o_{t}^{m} into a "measured" vehicle-count N^{m} via (10) involves the (arithmetic) average physical vehicle length L^{Ph} which may not be accurately known or may fluctuate over time; this introduces a further potential source of bias in the measurement equation. Note also that the term $N_{\max} = \Delta\lambda / L^{\text{Ph}}$ appearing in (10) actually corresponds to the maximum number of vehicles that can be accommodated in the link in a bumper-to-bumper manner.

The state equation (7) and measurement equation (11) have the appropriate form for Kalman-Filter application. To this end we should consider the system noise γ and measurement noise ζ to be independent

zero-mean white gaussian random variables. These assumptions may not be verified fully in practice but the Kalman-Filter may nevertheless deliver practically useful, suboptimal estimates. Of particular importance is the possible appearance of biased measurements (i.e. non-zero-mean error) because biased measurements cannot be rejected by the Kalman-Filter and will lead to accordingly biased estimates (see Section 3.3).

Despite the various sources of (partly non-zero-mean) errors, it is expected that the measurement equation contains a sufficient level of reliable information that may be exploited by the Kalman-Filter in order to reduce the accumulated error that would result from the usage of the conservation equation (7) alone.

The quality of the measurement equation may be improved if more internal detectors are used to produce an appropriate average measurement signal according to (3). Based on the general scheme of Fig. 1, one may employ M (instead of one) internal detectors (separated by equal distances $d = \Delta/M$), in which case o_t^m may be calculated at each period as the arithmetic average of time-occupancies collected by the individual internal detectors according to (3).

3.3. Kalman-Filter estimator

Based on the state equation (7) and measurement equation, a Kalman-Filter estimator (Jazwinsky, 1970) for the number of vehicles in the link may be immediately derived

$$\widehat{N}(k) = \widehat{N}(k-1) + T[q_{in}^m(k-1) - q_{out}^m(k-1)] + K[N^m(k-1) - \widehat{N}(k-1)] \tag{12}$$

where $\widehat{N}(k)$ is the delivered estimate of vehicle-count and K is the (stationary) gain parameter of the filter. The produced estimate $\widehat{N}$ is truncated if it exceeds the range $[0, N'_{max}]$ where N'_{max} is the maximum number of vehicles that can be accommodated in the link at standstill including the usual safety distance D among vehicles (e.g., $D = 1$ m); N'_{max} may be calculated, similarly to (10), from $N'_{max} = \Delta\lambda/(L^{Ph} + D)$. The filter (12) consists of a system model (the conservation equation comprising the first two terms on the r.h.s. of (12)) and a correction term that attempts to reduce the estimation error resulting from the system noise γ in (7).

The filter equation (12) may be re-arranged

$$\widehat{N}(k) = K \cdot N^m(k-1) + (1-K) \cdot \widehat{N}(k-1) + T\left[q_{in}^m(k-1) - q_{out}^m(k-1)\right] \tag{13}$$

in which case a further interpretation may be given. The produced estimate $\widehat{N}(k)$ results from the combination of

- An exponential smoothing (first two terms on the r.h.s. of (13)) based on the arriving measurements N^m.
- A prediction term that involves the most recent flow measurements.

According to the Kalman-Filter theory (Jazwinsky, 1970), the value of the gain K should be selected

$$K = \Pi/(\Pi + Z)$$

where Π satisfies

$$\Pi = (1-K)\Pi + T^2\Gamma.$$

From these equations we obtain

$$K = 0.5\left(-\alpha + \sqrt{\alpha^2 + 4\alpha}\right) \tag{14}$$

where $\alpha = T^2\Gamma/Z$; and Γ, Z are the variances of the system noise γ and measurement noise ζ, respectively. While the system noise variance Γ could be approximately determined based on the typical flow measurement errors, the value of the measurement noise variance Z is related to many different sub-processes and hence difficult to derive. However, (14) suggests that the value of K depends only on the ratio α, not on the explicit values of Γ and Z. Hence, rather than attempting to derive appropriate values for Γ, Z, one may attempt to fine-tune the ratio α, or, even more directly, the value of K to be used in the filter (12). Note that for $\alpha \to 0$ (i.e.

zero system noise or infinite measurement noise), (14) yields $K = 0$ which means that the estimation (12) makes no use of the measurements N^{m}; on the other hand, for $\alpha \to \infty$ (i.e. zero measurement noise or infinite system noise) it may be shown from the K, Π equations above that $K = 1$, i.e. the exponential smoothing in (13) is based on the latest measurements N^{m} only. Thus, potential K-values are included in the range $[0,1]$ as K is indeed monotonically increasing with α.

The aforementioned quantity Π represents the variance of the estimation error $\widehat{N} - N$ and is calculated

$$\Pi = 0.5Z\left(\alpha + \sqrt{\alpha^2 + 4\alpha}\right) \tag{15}$$

Thus, for $\alpha \to 0$ (and finite Z) we obtain $\Pi = 0$ (due to perfect model) while for $\alpha \to \infty$ we obtain $\Pi = T^2\Gamma$ (i.e. the estimation error in (13) results only from the prediction term due to perfect measurement N^{m}). The variance Π may be easily shown to be a monotonically increasing function of both the measurement error variance Z and the system error variance $T^2\Gamma$.

If the measurement N^{m} were sufficiently accurate, it might be better to use directly $\widehat{N} = N^{m}$ (as in Papageorgiou and Vigos, 2008) rather that the Kalman-Filter (12). For this to be true, we should have the variance Z of the error $\widehat{N} - N^{m}$ being smaller than the filter error variance Π, i.e. we should have $\Pi \geqslant Z$. After some calculations we derive from (15) the equivalent inequality $Z \leqslant 2T^2\Gamma$ which is quite unlikely to hold in practice as Z is usually much higher than Γ.

Eqs. (14) and (15) also contain valuable information on the role of the update period T. As mentioned in Section 2, the measurement variance Z obtains a minimum value for some value of T. Note that smaller Z can be easily shown to lead to smaller Π in (15). On the other hand, for higher T the system noise $T^2\Gamma$ increases, and consequently also α and the estimation error variance Π increase. In conclusion, the optimal update period T for best Kalman-filter estimates may not coincide with the optimal update period for least measurement error in (11).

If the measurement noise ζ contains a bias $E\{\zeta\} = b$, i.e., from (11), $E\{N^{m}\} = E\{N\} + b$; and if $E\{q_{in}^{m}\} = E\{q_{out}^{m}\}$, i.e. there is no bias in the traffic volume measurement; then we obtain from (12)

$$E\{\widehat{N}\} = E\{N^{m}\} = E\{N\} + b \tag{16}$$

i.e. the measurement bias cannot be rejected by the Kalman-Filter and is fully transmitted to the estimates $\widehat{N}$. If the flow measurements are also biased, i.e. $E\{q_{in}^{m}\} = E\{q_{out}^{m}\} + B$, we obtain from (12)

$$E\{\widehat{N}\} = E\{N\} + b + T \cdot B/K \tag{17}$$

which shows that flow measurement bias is also transmitted to the estimates but does not accumulate over time as it would in absence of the KF correction term in (12). The flow-induced bias of the estimates $\widehat{N}$ is seen in (17) to increase for higher T and lower K.

4. Simulation results

The developed Kalman-Filter estimator will be tested in this section by use of microscopic simulation. Although simulation of the traffic flow process may not reflect fully the real traffic phenomena, it may nevertheless provide valuable insights for a number of crucial questions that should be answered before actual field implementation:

- Does the Kalman-Filter estimator improve over the direct use of measurements? How many more internal detectors would be required to reach the estimation quality of the Kalman-Filter?
- What is the range of suitable K-values for the Kalman-Filter? What is the sensitivity of the estimation quality around the optimal K-value?
- What is the quantitative impact of the estimation period T and what value is recommended for the field implementation?
- How does the Kalman-Filter estimator perform under different conditions regarding N_{max}, traffic load, traffic light signaling?
- What is the impact of various measurement bias?

4.1. Simulation description

A self-developed microscopic simulator was used to describe the traffic phenomena on a single-lane 194-m long ramp as the one displayed in Fig. 2, with both downstream and upstream traffic signals. Vehicles are generated by the simulator far upstream of the upstream signal. The vehicles are moving on one lane according to appropriate Bando-type (Bando et al., 1995) discrete-time car-following equations detailed in Papageorgiou and Vigos (2008) based on a simulation time step $T_{\text{sim}} = 0.25$ s.

When a vehicle passes a detector, a time-occupancy signal is produced, whereby the duration of the detector occupancy depends on the vehicle physical length, the vehicle speed and the detector's effective length. The flow and time-occupancy measurements for each estimation period are calculated according to Section 2 and are eventually perturbed with random noise Δq and Δo, respectively, given by

$$\Delta q = 0.2q\psi, \quad \Delta o = 0.05o\psi$$

where ψ is a white random variable with the unit normal distribution. Thus the magnitude of the simulated measurement error depends on the current value of the measured quantity.

A standard simulation scenario with a duration of 1.38 h will be defined next to be used in the simulation investigations. Modifications of this scenario will be produced later as appropriate. For the standard scenario the simulation starts with an empty link. The upstream traffic signal is operated with a cycle of 90 s, while for the downstream signal we have a cycle of 20 s. The fixed green/red phases of both signals are chosen appropriately so as to create all possible values of vehicle-counts in the link (see Fig. 3). The detector effective lengths are zero for the standard scenario while the physical vehicle lengths are uniformly distributed in the range [3 m, 5 m]. One single internal detector is included for occupancy measurements in the middle of the link and the estimation time step equals 20 s. All simulations start with an erroneous initial estimate

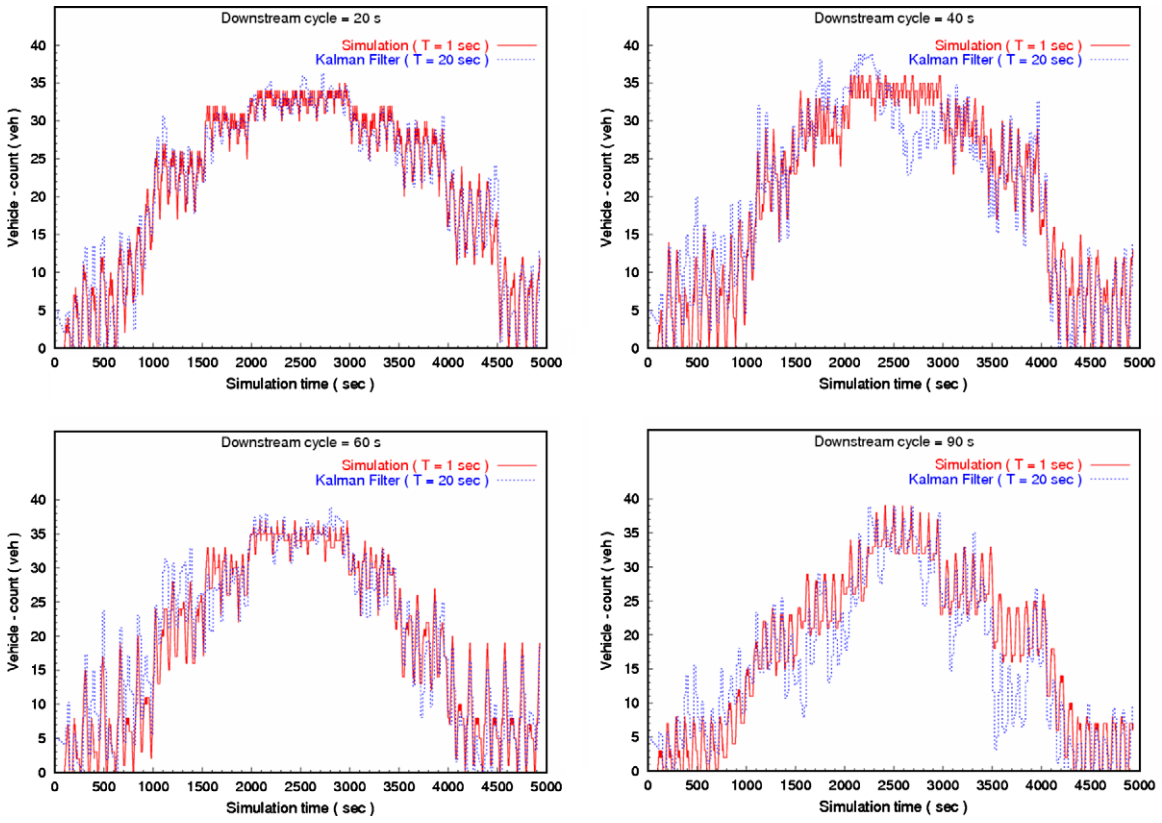

Fig. 3. Real and KF estimated vehicle-counts over time for four basic scenarios.

$\widehat{N}(0) = 5$ veh while the real number is zero. The value of $N'_{\max}$ used to truncate the Kalman-Filter estimates is 39 veh; this was obtained by dividing the link length Δ by the average physical vehicle length (4 m) augmented by $D = 1$ m which separates vehicles at standstill. Finally, $L^{\mathrm{Ph}} = 4$ m is used in (10) to produce the measured N^{m}.

Three further basic traffic scenarios were created by changing the traffic light settings of the standard scenario. More specifically, the three additional scenarios have different cycle times of 40 s, 60 s and 90 s, respectively, at the downstream traffic signal while the upstream signal is still operated with a cycle of 90 s. The green/red phases of both signals were again selected appropriately so as to create all possible values of vehicle-counts in the link (Fig. 3). A fifth stochastic scenario was created with the downstream traffic signal cycle changing stochastically (with uniform distribution) between 10 s and 90 s during the simulation (and green/red phases changing accordingly).

The evaluation of all scenarios is based on the following Relative Mean Square Error criterion:

$$\mathrm{RMSE} = 100\% \sqrt{K \sum_{k=1}^{K} [\widehat{N}(k) - N(k)]^2} \Big/ \sum_{k=1}^{K} N(k)$$

where $N(k)$ is the real and $\widehat{N}(k)$ the estimated vehicle-counts in the link. Note that the estimates $\widehat{N}$ are produced every T (e.g., $T = 20$ s); the produced estimate $\widehat{N}$ is then contrasted in the RMSE criterion to the latest simulated N value; in this sense, the discrete time index k in the RMSE criterion relates to the estimation period T (not to the simulation period T_{sim}).

4.2. Basic estimation results

Fig. 3 displays the actual and Kalman-Filter (KF) estimated vehicle-counts N and $\widehat{N}$, respectively, for the four basic scenarios described in Section 4.1, with appropriate K-values in (12) that will be detailed later. Note that the N-curve is displayed with a time-resolution of 1 s (i.e. every forth simulated N is actually displayed) while the estimates $\widehat{N}$ are updated every $T = 20$ s. It may be seen that the N-trajectory is subject to two kinds of time-variation; a high-frequency variation due to the periodical traffic signal switchings; and a low-frequency variation due to changing demand. Table 1 (column $L_j = L_j^{\mathrm{Ph}}$) displays the corresponding RMSE values and average errors (bias) $E\{N - \widehat{N}\}$ (the other columns of Table 1 are addressed in Section 4.5). It may be seen that

- Despite the availability of only one link-internal detector, $\widehat{N}$ follows N reasonably well (Fig. 3) for all scenarios and for all ranges of link occupancy. The displayed results are clearly better than the corresponding measurement-only estimates reported in Papageorgiou and Vigos (2008), which justifies the introduction of the Kalman-Filter. In particular the ZSZO-phenomenon that was clearly visible in Papageorgiou and Vigos (2008) is much less visible here thanks to the KF-imposed smoothing.
- The RMSE is seen in Table 1 to increase with increasing downstream signal cycle; this is due to the corresponding increase of the measurement error variance that was attributed by Papageorgiou and Vigos (2008) to the accordingly more frequent occurrence of the ZSZO-phenomenon. In addition, longer cycles with correspondingly longer green times create stronger density waves that render traffic conditions in the link more inhomogeneous; this contributes also to an increasing RMSE.

Table 1
RMSE values and $E\{N - \widehat{N}\}$ for five scenarios and various cases of effective vehicle length

Scenario cycle	$L_j = L_j^{\mathrm{Ph}}$		$L_j = L_j^{\mathrm{Ph}} + 1$ m		Corrected	
	RMSE (%)	$E\{N - \widehat{N}\}$ (veh)	RMSE (%)	$E\{N - \widehat{N}\}$ (veh)	RMSE (%)	$E\{N - \widehat{N}\}$ (veh)
20 s	9.8	0.59	16.8	−2.83	9.4	0.33
40 s	17.6	−0.04	24.6	−4.07	10.7	−0.50
60 s	14.8	−0.59	22.3	−3.5	10.9	−0.64
90 s	27.5	1.25	24.4	−2.43	19.0	0.17
Stochastic	22.8	0.29	23.3	−1.76	17.3	0.41

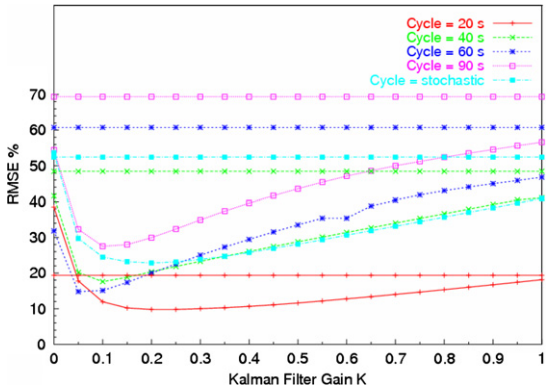

Fig. 4. Measurements and estimation RMSE in dependence of the KF gain K for various traffic scenarios.

- The average error (bias) $E\{N - \widehat{N}\}$ is seen in Table 1 to be close to zero (less than 1 veh) thanks to a similarly low bias of the measurement error (Papageorgiou and Vigos, 2008).
- The initial estimation $\widehat{N}(0) = 5$ veh is rapidly reduced in all cases thanks to the correction term of the Kalman-Filter equation (12).

The various traffic scenarios may also be investigated in order to assess the robustness of the KF estimator, the range of appropriate K-values in (12), as well as the sensitivity of the KF estimates to different K-values. To this end, Fig. 4 displays the KF RMSE for each scenario in dependence of different K-values; Fig. 4 displays for comparison also the corresponding measurement-only RMSE (with measurement noise ψ) which are of course independent of K and are therefore traced as horizontal lines. On the other hand, for $K = 0$ the Kalman-Filter (12) exploits the boundary flow measurement (conservation equation) only, i.e. the internal occupancy measurement is not used; thus, the value of each KF RMSE curve on the y-axis (i.e. for $K = 0$) reflects exactly this case.

The results displayed in Fig. 4 give rise to the following comments:

- The optimal gains K for all scenarios are in the range $[0.05, 0.25]$. In particular, the optimal K-values of the various scenarios are smaller if the measurement error variance is bigger, according to (14).
- The sensitivity of the RMSE within the mentioned range of K-values is rather low. This means that no elaborate fine-tuning of K is needed in practice; this is a significant property of the developed Kalman-Filter because the exact fine-tuning of K would require exact vehicle-counts $N(k)$ that are quite cumbersome and costly to obtain in the field.
- The KF estimates (for optimal K) are much better than the conservation equation by itself ($K = 0$) or the internal measurement by itself (horizontal lines) for all scenarios, which demonstrates the added value of the KF estimator.

4.3. Impact of the update period T

Fig. 5 displays the RMSE of the standard and stochastic scenarios in dependence of the estimation/measurement period T for the following cases:

- $\widehat{N} = N^{m}$ where N^{m} is produced according to (10) from occupancy measurements o_{t}^{m} stemming from one single internal detector without measurement noise Δo.
- $\widehat{N}$ produced from the KF estimator (12) based on perturbed (i.e. $\psi \neq 0$) flow $(q_{in}^{m}, q_{out}^{m})$ and occupancy (o_{t}^{m}) measurements, the latter from one single internal detector. The also displayed (optimal) gain parameter K was roughly fined-tuned for each individual KF case.

Note that the displayed measurement RMSE reflects the corresponding values of the variance Z of the measurement error in dependence of T, while the displayed KF RMSE reflects the corresponding values of the

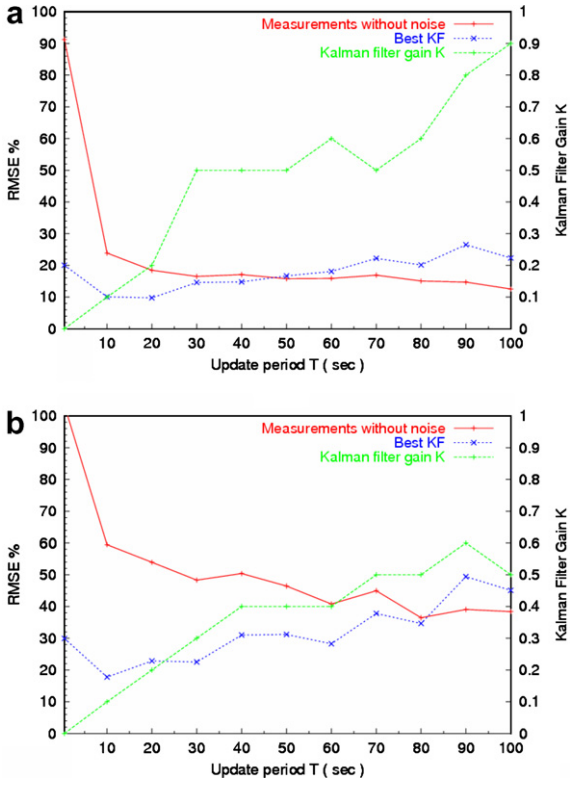

Fig. 5. Measurements and estimation RMSE in dependence of update period T for (a) the standard and (b) the stochastic scenarios.

variance Π of the KF estimation error in dependence of T. Fig. 5 indicates that, as expected from Papageorgiou and Vigos (2008), the measurement RMSE (and hence Z) is quite high for very small T; it reaches a plateau around $T = 20$ s (Fig. 5a) or $T = 80$ s (Fig. 5b); and then it increases slightly as T increases further. Clearly, the system error variance $T^2\Gamma$ increases quadratically with T since Γ is independent of T. Thus, the value of $\alpha = T^2\Gamma/Z$, appearing in (14) and (15), depends on the update period T in a corresponding way. Indeed, the following observations are made from Fig. 5:

- In accordance with (15) the KF RMSE is slightly increasing with increasing T, although for $T < 20$ s (Fig. 5a) or $T < 40$ s (Fig. 5b) the increase is partly compensated by the continuously improved measurement (which leads to lower Z-values).
- In accordance with (14), the optimal filter gain value is increasing with increasing T. More specifically, when T is small and the measurement noise variance Z is large, the optimal gain is close to zero. As T increases and Z decreases, α and consequently the optimal gain K are increasing, the latter reaching a value around 0.5 for $T > 30$ s.
- For moderate T-values, the KF produces better estimates than the measurements by themselves; but for $T > 50$ s (Fig. 5a) or $T > 80$ s (Fig. 5b) this situation is reversed due to increasing system error variance $T^2\Gamma$ while the measurement error variance Z is virtually constant.

4.4. Impact of the number of detectors

Fig. 6 displays the RMSE of the standard and stochastic scenarios in dependence of the number of internal detectors (measuring time-occupancy) with $T = 20$ s for the following cases:

- $\widehat{N} = N^{\mathrm{m}}$ produced from all available occupancy measurements without measurement noise Δo.
- $\widehat{N} = N^{\mathrm{m}}$ as above but with occupancy measurements perturbed with measurement noise Δo.

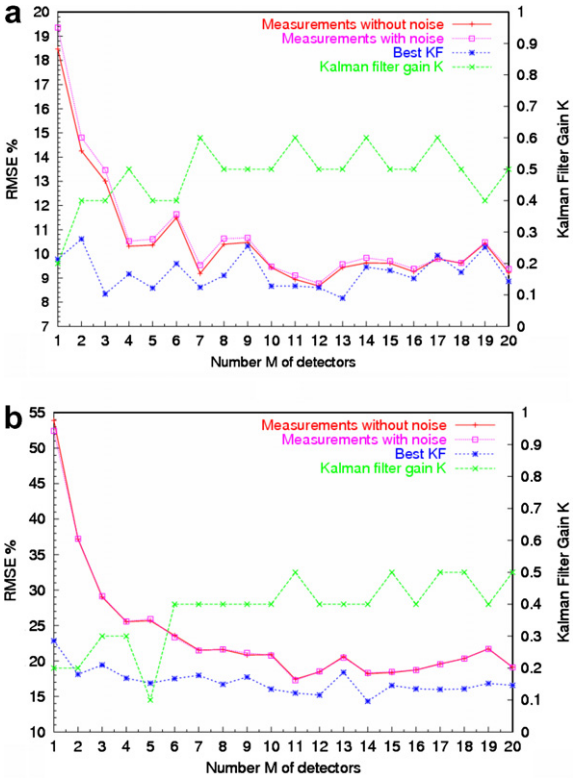

Fig. 6. Measurements and estimation RMSE in dependence of the number of internal detectors for $T = 20$ s for (a) the standard and (b) the stochastic scenarios.

- $\widehat{N}$ produced from Kalman-Filter (12) based on all available measurements (perturbed with noise) and roughly fine-tuned gains K.

The following observations are made from Fig. 6:

- The impact of the occupancy measurement noise is quite negligible compared to the other error sources in the measurement equation.
- The measurement RMSE of 20% (Fig. 6a) or 50% (Fig. 6b) for one internal detector is reduced to around 10% (standard scenario) or 20% (stochastic scenario) for a sufficient number of internal detectors. The inclusion of more than ten internal detectors does not improve the quality of the measurements further.
- The KF RMSE is pretty stable around 9% (for the standard scenario) or 18% (for the stochastic scenario) for any detector number. Note that the optimal gain K takes values around 0.5 for $M > 5$ due to better measurement quality.
- Since the RMSE of 9% or 18% is reached with measurements only, by use of ten or more internal detectors while the KF needs only three detectors to reach this quality (two boundary and one internal detector), it may be concluded that, based on these scenarios, the application of the Kalman-Filter allows for the cost of roughly seven detectors to be saved.

4.5. Non-zero effective detector length

If the effective detector length is non-zero while measuring time-occupancy (or, equivalently, if the effective vehicle length is different than the physical vehicle length), then the measured time-occupancy is not a bias-free representation of the space-occupancy (Section 2) and hence the bias of the errors ζ_0 in (8) and ζ in increases accordingly. If the non-zero effective detector length ε is known, one may convert the collected time-occupancy

measurements o_t^m into bias-free representations of the space-occupancy by multiplying them with $L^{Ph}/(L^{Ph}+\varepsilon)$ (Papageorgiou and Vigos, 2008), in which case one may partly recover the results presented earlier. On the other hand, if this transformation is not performed and the biased measurements are used to feed the Kalman-Filter, then the produced estimates will be accordingly biased as the Kalman-Filter has no means to reject the measurement bias according to (16).

Fig. 7 presents the measurements and KF estimation results for different K values for the unbiased (standard scenario), the biased (with $\varepsilon = 1$ m) and the corrected (as mentioned above) cases. Table 1 displays the RMSE and average error of the five scenarios for the cases without L_j-bias (column $L_j = L_j^{Ph}$), with L_j-bias (column $L_j = L_j^{Ph} + 1$ m) and for the corrected case. It may be seen in Fig. 7 that:

- The RMSE of the (biased) measurement of N^m is almost doubled compared to the RMSE of the measurement of the standard case as already known from Papageorgiou and Vigos (2008).
- Due to the lower quality of the measurement variance (higher Z), the optimal KF gain K is smaller (around 0.05) than in the non-biased case. The RMSE of the resulting KF estimates when using biased measurements is more than doubled compared to the unbiased case.
- In the corrected case (where measurements o_t^m are multiplied with $L^{Ph}/(L^{Ph}+\varepsilon)$), the original KF performance is virtually recovered for a gain K-value similar as in the unbiased case.

Indeed, Table 1 reveals (as already observed in Papageorgiou and Vigos, 2008) that, for some scenarios, the estimation RMSE of the corrected case may become lower than in the unbiased (column $L_j = L_j^{Ph}$) case, because a non-zero effective detector length decreases the probability of the ZSZO-phenomenon thus reducing the measurement and eventually the estimation error variance. These results suggest that it may be more beneficial for the estimation accuracy to employ detectors with non-zero (but approximately known) effective detector length ε and to proceed to a correction of the obtained measurements; than to adjust the detectors such that $L_j = L_j^{Ph}$.

4.6. Uncertain physical vehicle length average

In all previous investigations we assumed that the average physical vehicle length L^{Ph} used to transform the time-occupancy measurement o_t^m into a vehicle-count N^m in (10) is accurately known (equal to $L^{Ph} = 4$ m). If this value is not accurately known in practice, then an additional measurement bias will result whose effects are similar in nature as those of Section 4.5.

To investigate this issue we have considered two variations of the standard scenario whereby the vehicles created in the microscopic simulation include 10% and 30% of trucks, respectively; trucks were assumed to have a length in the range [8 m, 10 m] with uniform distribution. Note that in these two cases we still use $L^{Ph} = 4$ m in (10) (as in the standard case) thus creating biased measurements. Fig. 8 displays the results for both cases, along with the standard case for comparison. It may be seen that:

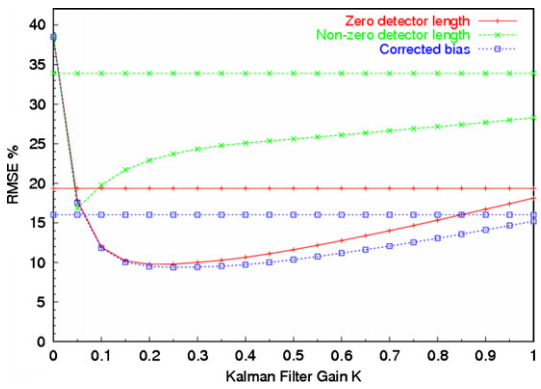

Fig. 7. The case of non-zero effective detector length for standard scenario.

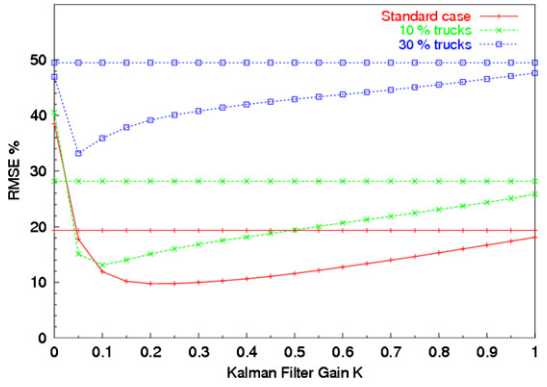

Fig. 8. The cases of uncertain average physical vehicle length.

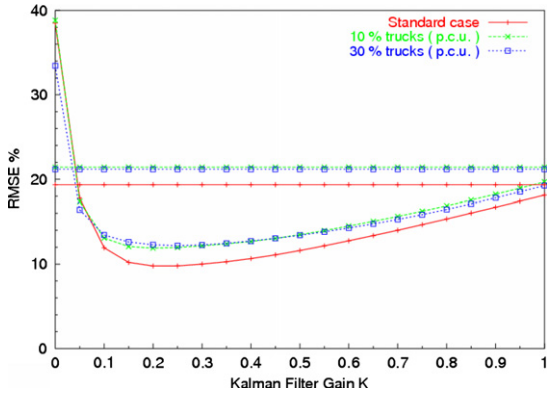

Fig. 9. The cases of 10% and 30% trucks with N in p.c.u. and 1 truck = 2.25 p.c.u.

- The measurement RMSE of both cases increases with increasing percentage of trucks due to increasing bias.
- Due to the lower quality of the measurement (higher variance Z), the optimal KF gain K is accordingly smaller than in the standard case.
- The RMSE of the resulting KF estimates is increasingly higher for increasing bias, but the accuracy is still quite good for the 10%-trucks case.

It should be emphasized that the above increased errors occur only if the evaluation of the results is based on real vehicle numbers N, whereby 1 truck = 1 vehicle. If the real quantity N is measured in p.c.u. (passenger car units) and we assume here 1 truck = 2.25 p.c.u., then the RMSE for the measurement and the estimates are similar to the standard case as Fig. 9 indicates. In fact, the measurements in this case may be deemed unbiased while the variance of the state error increases slightly because the flow measurements and hence the conservation equation address vehicles, not p.c.u.; since the state error variance increases while Z is virtually the same as in the standard scenario, the resulting optimal K are slightly greater than in the standard case, but the sensitivity of the results is rather low in a broad range of K-values. Note that estimates $\widehat{N}(k)$ in p.c.u. may be more useful for signal control or ramp metering applications (where avoidance of queue spillback upstream of the link is a major concern).

4.7. Longer link

To check the KF efficiency in the case of links with different geometry, the link length of the standard case was doubled to 394 m. Note that a larger link length Δ increases the variance Z of the measurement as

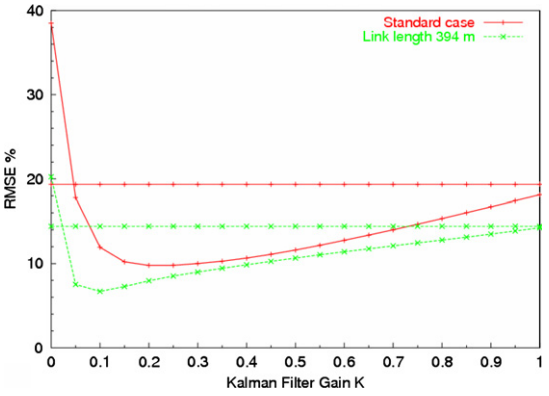

Fig. 10. Estimation results for a longer link.

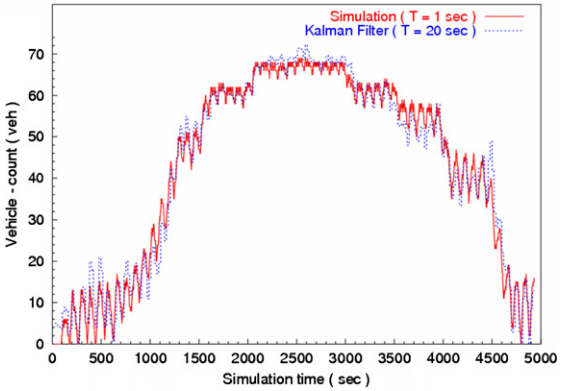

Fig. 11. Real and KF estimated vehicle-counts over time for longer-link scenario.

evidenced by the definition of the measurement error ζ in (11) while the variance Γ of the system error remains unchanged. Hence it is expected that the resulting optimal value of the KF gain K will be smaller than in the standard case and the results displayed in Fig. 10 actually confirm this conclusion. Since RMSE is a relative error and the absolute values of the vehicle-count N are higher for a longer link, the relative estimation accuracy of both the measurement and the Kalman-Filter in Fig. 10 is slightly better than in the standard case. Fig. 11 displays the time trajectories of the real and KF-estimated vehicle-counts N and $\widehat{N}$, respectively, that confirm the excellent performance of the filter also for a longer link.

4.8. Large initial estimation error

In all previous investigations it was assumed that the initial estimate is $\widehat{N}(0) = 5$ veh while the real vehicle-count is $N(0) = 0$ veh, i.e. an initial estimation error of 5 veh was imposed. The importance of the correction term in the filter equation (12) may be appreciated in a last scenario where an even larger initial estimation error is assumed with $\widehat{N}(0) = 20$ veh. Without the correction term, this initial estimation error would not be reduced and, indeed, Fig. 12 indicates that the KF RMSE for $K = 0$ is much higher for $\widehat{N}(0) = 20$ veh than in the standard case. Moreover, the optimal KF gain K is seen to be slightly larger than in the standard case due to the larger initial state error. Fig. 13 shows that, thanks to the correction term, this large initial error is rapidly reduced within a few estimation time-steps.

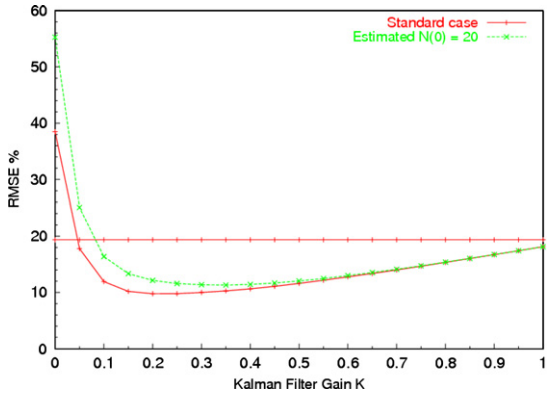

Fig. 12. The case of larger initial estimation error ($\widehat{N}(0) = 20$ veh).

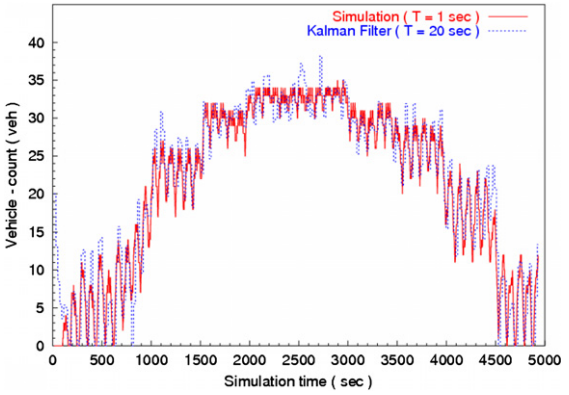

Fig. 13. Real and KF estimated vehicle-counts over time for the scenario with $\widehat{N}(0) = 20$ veh.

5. Guidelines, recommendations and conclusions

5.1. Guidelines and recommendations for practical application

We will here summarize the necessary steps, equations and recommended parameter values required if the described estimation scheme for vehicle-counts in signalized links is to be adopted and implemented.

To start with, a basic prerequisite is the availability of two boundary detectors measuring flows q_{in}^{m} and $q_{\text{out}}^{\text{m}}$, respectively; and one internal detector (preferably located around the middle of the link) measuring time-occupancy o_{t}^{m}, see Fig. 2. In case of availability of multiple internal detectors, they should be preferably placed as Fig. 1 indicates and the overall occupancy measurement o_{t}^{m} feeding the Kalman-Filter should be the average of all available time-occupancies according to (3). All occupancies in this paper are assumed to take values within the range $[0, 1]$.

In real-time operation, the filter is fed with the latest available measurements $q_{\text{in}}^{\text{m}}(k-1)$, $q_{\text{out}}^{\text{m}}(k-1)$, $o_{\text{t}}^{\text{m}}(k-1)$, collected during the time-period $[(k-1)T, kT)$, where $k = 1, 2, \ldots$ is the discrete time index, to produce the estimated vehicle number $\widehat{N}(k)$, valid for the time instant kT. The usually required range of values for T is [10 s, 30 s].

The measured occupancy o_{t}^{m} should ideally be collected with an effective detector length ε equal to zero. If the corresponding adjustment of the detectors is not possible, the collected occupancy should be multiplied with $L^{\text{Ph}}/(L^{\text{Ph}} + \varepsilon)$ before further use (see Section 4.5), where L^{Ph} is the average physical vehicle length (around 4 m). Typical ε-values are $1 \ldots 2$ m depending on the employed loop detectors. Note that, according

to Section 4.5, the case $L_j \neq L_j^{\mathrm{Ph}}$ with correction of the time-occupancies may lead to better results than the $L_j = L_j^{\mathrm{Ph}}$ case.

Subsequently, the o_t^{m} measurement must be converted into a corresponding N^{m} measurement by use of (10). Note that the term $\Delta\lambda/L^{\mathrm{Ph}}$ appearing in (10) is equal to (and could be replaced by) $N_{\max}$, the maximum number of vehicles that could be accommodated in the link in a bumper-to-bumper manner. In case of a non-negligible percentage of trucks, there are two options:

- Truck length is not considered in the utilised average physical length L^{Ph}, which remains equal to around 4 m; in this case, the Kalman-Filter estimates $\widehat{N}$ will be automatically delivered in p.c.u., where L^{Ph} corresponds to 1 p.c.u. (this option is recommended).
- Truck length is considered (according to their usual percentage) when selecting the value of L^{Ph} (which would then be naturally higher than 4 m); in this case, the Kalman-Filter estimates $\widehat{N}$ will be delivered in veh, that may be either passenger cars or trucks in the pre-specified proportion.

After these arrangements the Kalman-Filter estimate $\widehat{N}(k)$ may be produced by use of (12) and subsequent possible truncation if the estimate exceeds the range $[0, N'_{\max}]$. The recommended range of values for the filter gain K is $[0.05, 0.3]$, but if no fine-tuning is carried out, a value of $K = 0.1$ is deemed quite appropriate and robust. Finally the $N'_{\max}$ value needed to truncate the filter results, corresponds to the maximum number of vehicles that can be accommodated in the link at standstill, including the usual safety distance D, i.e. $N'_{\max} = \Delta\lambda/(L^{\mathrm{Ph}} + D)$.

5.2. Conclusions

A rather simple Kalman-Filter estimator was designed for the number of cars included in a signalized link. The estimator was found in manifold simulation investigations to be quite efficient and robust. The algorithmic intelligence of the estimator was shown to replace for several additional loop detectors that would be required in order to reach an equivalent accuracy without the estimator. Several issues and options were investigated in detail in order to come up with suitable conclusions and recommendations with respect to various aspects including:

- The impact of measurement/estimation update period T.
- The impact of multiple internal detectors for time-occupancy.
- The proper values of the filter gain parameter K.
- The impact (and countermeasures) of non-zero effective detector length in the occupancy measurements.
- The impact of trucks.
- The impact of various traffic conditions and the ramp geometry.

Field testing is the next step in the development of this algorithm.

Acknowledgements

This work was supported in part by the European Commission's IST (Information Society Technologies) Programme under the project EURAMP (IST-2002-23110). The contents of the paper are under the sole responsibility of the authors and do not necessarily reflect European Commission's views.

References

Bando, M., Hasebe, K., Nakayama, A., Shibata, A., Sugiyama, Y., 1995. Dynamical model of traffic congestion and numerical simulation. Physical Review E 51, 1035–1042.

Bhouri, N., Hadj-Salem, H., Papageorgiou, M., Blosseville, J.M., 1988. Estimation of traffic density on motorways. Traffic Engineering and Control 29, 579–583.

Cheung, S.Y., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W., Varaiya, P., 2005. Traffic measurement and vehicle classification with a single magnetic sensor. 84th Annual Meeting of the Transportation Research Board, Washington DC.

Jazwinsky, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

Kotsialos, A., Papageorgiou, M., 2005. A hierarchical ramp metering control scheme for freeway networks. In: Proc. 2005 American Control Conference, Portland, Oregon, pp. 2257–2262.

Papageorgiou, M., Vigos, G., 2008. Relating time-occupancy measurements to space-occupancy, density and link vehicle-count. Transportation Research Part C 16 (1), 1–17.

Smaragdis, E., Papageorgiou, M., 2003. A series of new local ramp metering strategies. Transportation Research Record No. 1856, pp. 74–86.

Sun, X., Horowitz, R., 2005. A localized switching ramp-metering controller with a queue length regulator for congested freeways. In: Proc. 2000 American Control Conference, Portland, Oregon, pp. 2141–2146.