Data fusion solutions to compute performance measures for urban arterials

Hans Van Lint

Robert. L. Bertini

Serge P. Hoogendoorn

Symposium Celebrating 50 Years of Traffic Flow Theory

2014 TFT Summer Meeting - August 11-13, Portland, Oregon USA

March 25, 2014

INTRODUCTION

One of the key problems faced by traffic management operators of large urban traffic networks is the lack of sufficient data to compute performance indicators. These indicators, such as travel time, queue length, loss hours, total time spent, are useful for both offline evaluation purposes, as well as online traffic control applications. In the latter case, such data is particularly of use in coordination algorithms that require information on the number of vehicles present or queuing in certain areas. This information in turn is used for example to assess the amount of buffer space available to temporarily store or reroute vehicles from more densely used parts of the network. Computing the amount of vehicles present or queuing in a certain area requires, of course, counting the number of vehicles that enter or exit that area. In this extended abstract we show how through fusing vehicle counts and travel times (measured by any means available), the well-known drift-error can be reduced to virtually zero. In the complete paper we show how this algorithm fits in a wider suite of data fusion tools to compute urban traffic performance indicators on the basis of multiple sources of data.

CUMULATIVE DRIFT

Denote $N(t)$ as the number of vehicles present on the road stretch at a certain time intant t. Figure 1 (top) illustrates that the number of vehicles on a simple hypothetical roadstretch of 1 km can be derived directly using cumulative inflow and outflow curves $Q_1(t)$ and $Q_2(t)$. The horizontal distance between these curves depicts the travel time of the n^{th} vehicle entering ths roadstretch, whereas the vertical distance depicts the number of vehicles $N(t) = Q_1(t) - Q_2(t)$. It is straightforward to compute $N(t)$ numerically (we consider discrete time steps k of size Δt):

$$N(k+1) = N(k) + \Delta t(q_1(k) - q_2(k)) \tag{1}$$

Unfortunately, this approach is highly sensitive to counting errors. Suppose both detectors make a 1% counting error, that for the sake of argument is assumed normally distributed around the actual number of vehicles passing both locations

$$q_i^{obs}(k) = q_i(k) + \eta(k) \tag{2}$$

with

$$\eta_i(k) \sim N(0, 0.01 \times q_i(k))$$

Substituting (2) into (1) of course results in

$$N(k+1) = N(k) + \Delta t(q_1(k) - q_2(k)) + \zeta(k) \tag{3}$$

with

$$\zeta(k) = \Delta t(\eta_1(k) - \eta_2(k))$$

which describes a so-called random walk (the sum of two independent error terms) that may result in values anywhere between minus or plus infinity.

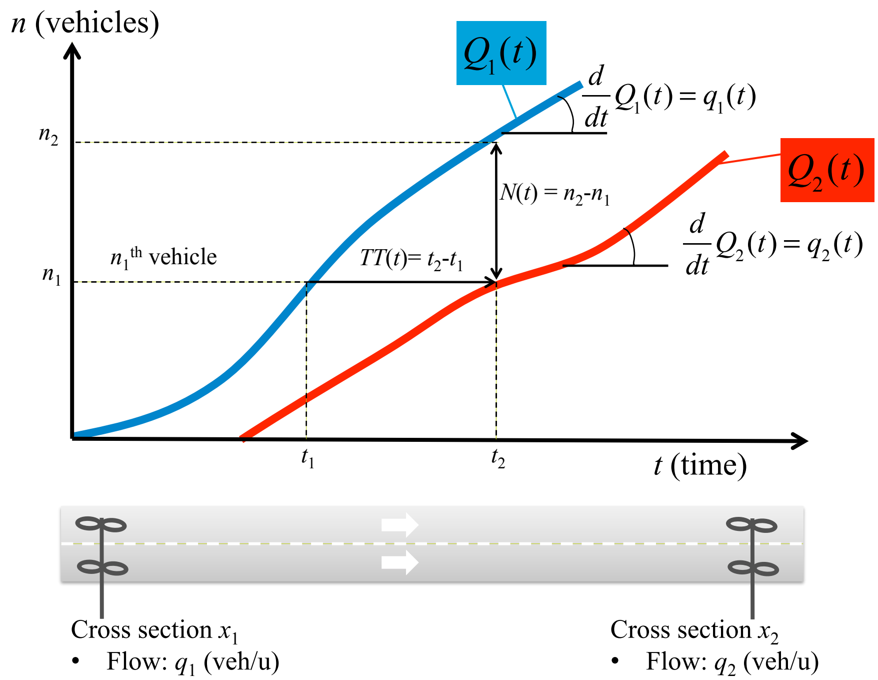

FIGURE 1 Cumulative inflow and outflow curves for a straight road stretch

Figure 2(a) and (b) (top) show the measured in- and outflows $q_1(k)$ and $q_2(k)$ on the same hypothetical roadstretch of 1 km, in which the errors of both 1% and 5% are hardly visible, whereas Figure 2 (a) and (b) (bottom) show the resulting error in the estimated number of vehicles $N(k)$ using equation (1). Even in case of a small error of 1% the error is sizeable—8 vehickles per km may make the difference between free-flowing traffic or oversaturated conditions.

FUSING FLOW OBSERVATIONS AND MEASURED TRAVEL TIMES

One approach to correct this so-called "drift" error is to use measured (realised) travel times $TT_r^{obs}(t)$ along the same road stretch. The underlying assumption is that travel time measurements are more accurate than loop counts[1]. Recall that the horizontal distance between the cumulative curves allows us to estimate this travel time of the n^{th} vehicle entering this roadstretch. Figure 3 illustrates the basic idea. Suppose at some time instant t_2 we obtain a "travel time measurement" $TT_r^{obs}(t_2)$. At that same time instant we can also compute an estimate for the realised travel time using the cumulative curves:

$$TT_r^{est}(t_2) = t_2 - t_1$$

We can now use the error in this travel time estimate

$$\epsilon_{TT}(t_2) = TT_r^{obs}(t_2) - TT_r^{est}(t_2)$$

to correct both cumulative curves. We define n^* as the factor with which we correct the curves. It turns out this correction factor is proportional to the travel time error, that is, $n^* \propto$

[1]This is true under the condition that outliers are properly removed before computing the average (or median). An example of such an outlier removal technique is the following: Remove all travel times observations in a certain time period further away from the median travel time than $\alpha \times TT^{75} - TT^{25}$, where TT^{XX} represents the XX^{th} travel time percentile. The outlier removal procedure should be applied strictly for this purpose

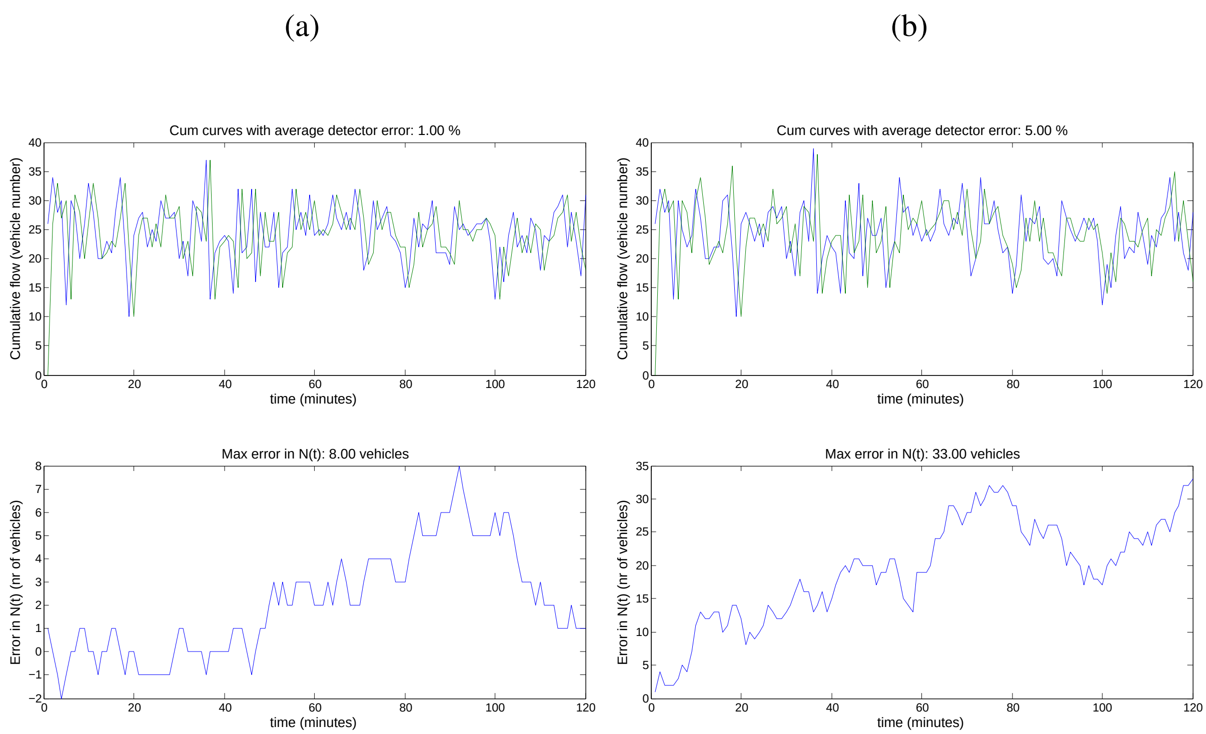

FIGURE 2 Cumulative drift. Top (a) and (b): measured in- and outflows flows $q_1(k)$ and $q_2(k)$ for detectors with a 1% and 5% error rate respectively; bottom (a) and (b): resulting errors in the estimated number of vehicles $N(k)$

$\epsilon_{TT}(t_2)$—details will be provided in the full paper. The rationale of course makes sense. If we overestimate the outflow (Q_2), we overestimate the rate at which vehicles are able to depart from this road stretch and the result is that we underestimate the travel time these vehicles have incurred. The opposite occurs when we overestimate the inflow (Q_1); in that case we exaggerate the amount of vehicles accumulating on the road stretch and as a result we overestimate travel time. In Figure 3 the direction of this correction n^* is indicated in case the observed travel time is *larger* than the estimated travel time ($\epsilon_{TT}(t_2) > 0$). It turns out that this approach works remarkably well and is capable of eradicating the error in $N(t)$ virtually completely, not just in case of random errors, but also in case of structural errors (bias).

Figure 4 gives an example of in case of both 5% structural and random errors—again on the hypothetical 1 km road stretch introduced earlier. Figure 4 (a) shows a detail of the corrected cumulative curves, the erroneous cumulative curves and the ground-truth cumulative curves. In absolute sense, the algorithm does correct the errors, but relative to one another the algorithm does, which is demonstrated by time series of the erroneous versus the corrected number of vehicles estimated in Figure 4(b)—the latter lies very close to the ground-truth.

OUTLOOK FULL PAPER

This data fusion algorithm appears to "fix" cumulative drift errors by combining counts with travel times that may come from AVI systems, floating car data, bluetooth sniffers or any type of sensors. In the full paper we will discuss the algorithm in-depth and demonstrate it on various real-life examples in both The Netherlands and the US.

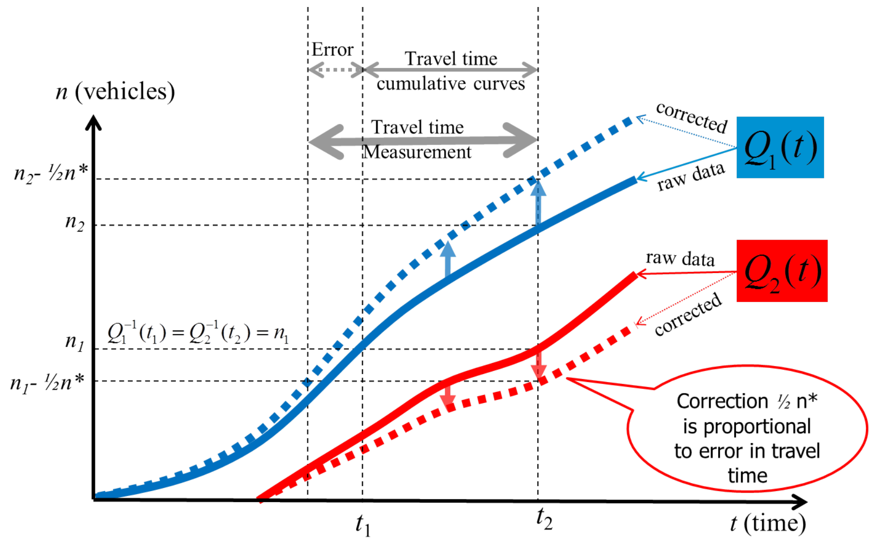

FIGURE 3 Basic idea of using measured travel times to correct for the cumulative errors due to "drift"

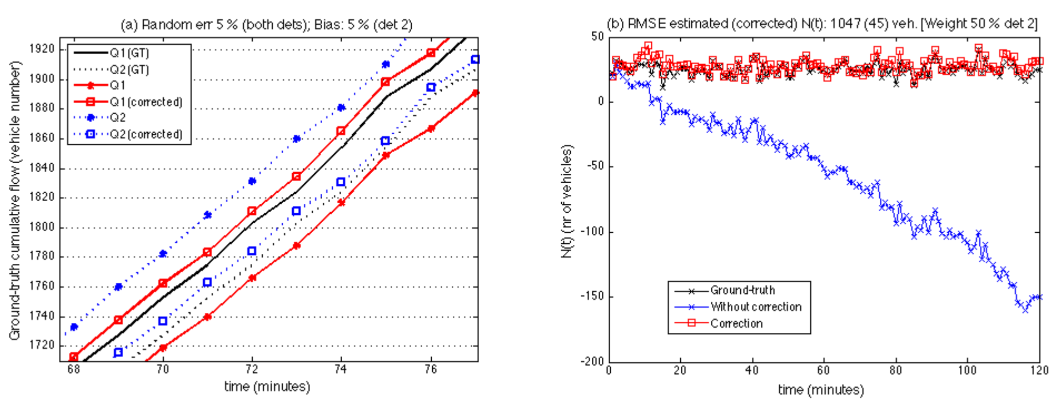

FIGURE 4 Example of error correction algorithm in case of both random errors and structural errors. (a) a detail of the corrected cumulative curves, the erroneous cumulative curves and the ground-truth cumulative curves; (b) time series of the erroneous versus the corrected number of vehicles estimated—the latter lies very close to the ground-truth.