Data fusion solution to fix the cumulative drift problem on urban arterials

Hans van Lint, Robert Bertini, Serge Hoogendoorn

Symposium Celebrating 50 Years of Traffic Flow Theory
2014 TFT Summer Meeting
August 11-13, Portland, Oregon USA
Context

Huge surge of monitoring projects in urban environments

- In the Netherlands
  - Virtually 100% vehicle actuated traffic controllers: inductive loops measuring flows (and in)
  - Last five years: huge investments in urban monitoring, particularly in AVI systems (cams, BT)
    - TRAVEL TIMES
    - REALISED ROUTES
    - PARTIAL OD RELATIONS
- Usefulness for urban traffic management debated ...
Overview

1. Deducing vehicle accumulation using vehicle counts (cum curves) is straightforward ...

2. Problem: cumulative drift due to errors in counts

3. Solution: (f)use counts (with) measured travel times

4. Results of this “simple trick” are rather good
\[
Q_1(t) = \frac{d}{dt} n_1(t) = q_1(t)
\]

\[
Q_2(t) = \frac{d}{dt} n_2(t) = q_2(t)
\]

\[
N(t) = n_1 - n_2
\]

Cross section \( x_1 \)
- Flow: \( q_1 \) (veh/u)

Cross section \( x_2 \)
- Flow: \( q_2 \) (veh/u)
The cumulative drift problem

Occurs when $q_1(t)$ and $q_2(t)$ contain errors

- Source errors (miscounts, double counts):
  - lane changes, power failure, etc.

- Errors may be random or structural (bias)

- Consequence:

\[
N(t) = \int_{q_1(s)} ds - \int_{q_2(s)} ds
\]

\[
q_i(t) = \hat{q}_i(t) + \epsilon_i(t)
\]

With e.g. $\epsilon_i(t) \sim N(\mu, \sigma)$
The cumulative drift problem

Occurs when $q_1(t)$ and $q_2(t)$ contain errors

- Source errors (miscounts, double counts):
  - lane changes, power failure, etc.

- Errors may be random or structural (bias)
- Consequence:

$$N(t) = \int_{t} q_1(s) \, ds - \int_{t} q_2(s) \, ds$$

$$q_i(t) = \hat{q}_i(t) + \varepsilon_i(t)$$

$$N(t) = \hat{N}(t) + \int_{t} \left( \varepsilon_1(s) - \varepsilon_2(s) \right) \, ds$$

This is a random walk!
(which means vehicle accumulation is practically unobservable using counts)
The cumulative drift problem

(a) Cum curves with average detector error: 0.00 %

(b) Cum curves with average detector error: 1.00 %

(c) Cum curves with average detector error: 5.00 %

(d) Cum curves with average detector error: 10.00 %
Solution

\[ n \text{ (vehicles)} \]

\[ \hat{TT}_r(t_2) \]

\[ n_0 \]

\[ n_2 \]

\[ t_0 \]

\[ t_1 \]

\[ t_2 \]

\[ t \text{ (time)} \]

\[ Q_1(t) \]

\[ Q_2(t) \]

raw data

raw data
Solution

\[ \varepsilon_{TT}(t_2) \quad \widehat{TT}_r(t_2) \]

\[ Q_1(t) \quad Q_2(t) \]

Implies we either
- Underestimated inflow
- Overestimated outflow
- Or both
Solution

\[ \epsilon_{TT}(t_2) \quad \widehat{TT}_r(t_2) \]

\[ n \text{ (vehicles)} \]

\[ n_1 \]

\[ n_2 \]

\[ n_0 \]

\[ t_0 \]

\[ t_1 \]

\[ t_2 \]

\[ t \text{ (time)} \]

The correction factor is proportional to \( \epsilon_{TT} \).

\[ TT^{obs}_r(t_2) \]

\[ Q_1(t) \]

\[ Q_2(t) \]

Corrected raw data
Data fusion on Urban Arterials - 2014 TFT Summer Meeting, Portland, Oregon
Solution turns out to be

A simple parameter-free correction algorithm

- Correction factor can be expressed as function of known quantities only

\[
\frac{\varepsilon_N(t_2)}{\varepsilon_{TT}(t_2)} = \frac{n_2 - n_0}{t^* - t_0}
\]

- Or more generally

\[
\varepsilon_N(t_i) = \varepsilon_{TT}(t_i) \frac{Q_i(t_i) - n_0}{Q_i^{-1}(n_i) - Q_i^{-1}(n_0)}
\]
Results

Rows (random errors): \{1\%, 5\%, 10\%\}
Columns (bias): \{-5\%, 0, 5\%\};
Results

Rows (random errors): \{1\%, 5\%, 10\%\}
Columns (bias): \{-5\%, 0\%, 5\%\};
Discussion

• Good news for urban traffic management agencies:
  • Algorithm works offline or online (although with a time lag of course)

• Quite a few puzzles to solve:
  • Limits algorithm (magnitude and nature of errors)
  • What to do when no closed counting situation?
  • What to do when no measured travel times?
  • How to incorporate travel time errors?
Next steps …

- Solve puzzles

- Pubs:
  - TRB2015 paper:
    - Basic idea + extension to multiple links
  - TFT50 / special issue jnl paper
    - Basics TRB Paper
    - + combination with additional methods
    - + real data case studies