REAL-TIME ESTIMATION OF CRITICAL VALUES OF THE MACROSCOPIC FUNDAMENTAL DIAGRAM FOR MAXIMUM NETWORK THROUGHPUT

Konstantinos Ampountolas†
University of Glasgow
School of Engineering
Rankine Building, Oakfield Avenue
Glasgow G12 8LT, United Kingdom
Tel: +44-(0)141-330-5343
E-mail: Konstantinos.Ampountolas@glasgow.ac.uk

Anastasios Kouvelas
École Polytechnique Fédérale de Lausanne (EPFL)
School of Architecture, Civil & Environmental Engineering
Urban Transport Systems Laboratory
GC C2 390 (Bâtiment GC), Station 18
1015 Lausanne, Switzerland
E-mail: Tasos.Kouvelas@epfl.ch

† Corresponding Author
Number of words: 6,000 (including abstract and references) + 1,500 (6 Figures) = 7,500

ABSTRACT

Perimeter flow control or gating has recently been found to be a practical and efficient control scheme in mitigating traffic congestion in urban road networks. This control scheme aims at stabilising the accumulation of vehicles (or a proxy of accumulation, e.g. average occupancy or density) of the macroscopic or network fundamental diagram near critical accumulation to achieve maximum network throughput. Nevertheless, the maximum throughput (capacity flow) in urban road networks may be observed over a range of accumulation-values. In this work, an extension of a previously proposed real-time feedback perimeter flow control strategy is proposed that allows the automatic monitoring of the critical accumulation to help maintain the accumulation near the optimal range of accumulation-values, while network’s throughput is maximised. To this end, we design a Kalman filter-based estimation algorithm that utilises real-time measurements of circulating flow and accumulation of vehicles to produce estimates of the currently prevailing critical accumulation. The developed strategy may be valuable whenever the network fundamental diagram is not well defined and the critical accumulation cannot accurately be specified or is subject change due to traffic-responsive signal control, traffic composition (e.g. cars versus buses), or non-recurrent day-to-day traffic patterns. We use real experimental data from an urban area with 70 sensors and show that the area exhibits a fundamental diagram with low scatter. We demonstrate that the fundamental diagram is reproduced under different days but its shape and critical occupancy depend on the applied semi-real-time signal control and the distribution of congestion in the network. Preliminary results from the application of the estimation algorithm to the experimental data indicate good estimation accuracy and performance, and rapid tracking behaviour.

Keywords: Macroscopic or Network Fundamental Diagram; Urban traffic; Critical accumulation; Critical occupancy; Recursive estimation; Kalman filter
INTRODUCTION

Urban road networks without effective means of flow control have been shown to exhibit the throughput-load relationship illustrated in Figure 1 (e.g. in the form of a flow-vehicle accumulation curve). This is the so-called Macroscopic or Network Fundamental Diagram (MFD or NFD) of two-dimensional urban road networks, which provides for network regions under certain regularity conditions (mainly homogeneity in the spatial distribution of congestion and the network topology), a concave, low-scatter relationship between network vehicle accumulations $n$ (veh) and network outflow $q$ (veh/h). According to the figure, the typical characteristic of two-dimensional urban road networks is that as the density is increased from zero, the network throughput (circulating flow or trip completion rate) increases to a maximum (flow capacity) and then turns down and decreases sharply to a low value possibly zero (in case of gridlock). Local-area gridlock phenomena can indeed occur but a global gridlock is mostly preventable due to the self-healing (recovery) mechanism of the real transportation networks (as revealed in studies with empirical data). Note that capacity flow in urban road networks may be observed over a range of vehicle-numbers in contrast to freeway traffic where capacity flow is deemed to occur for a (more or less) specific density value. Physically, when an urban network is congested, some regions may be blocked and significant queue spillback to upstream junctions takes place, which leads to flow degradation and hence infrastructure underutilisation. Hence, flow control mechanisms are necessary to prevent throughput degradation and maintain maximum flow.

The idea of a network fundamental diagram with an optimum (critical) accumulation $\tilde{n}$ at which flow capacity $q_{\text{cap}}$ is reached belongs to (1) and has been re-initiated later in (2, 3, 4, 5, 6) and others. However, the verification of its existence with real-data and dynamic features is very recent (7). This property of two-dimensional urban networks is important for modelling purposes, as details in individual links are not needed to describe the congestion level of cities and its dynamics. It can also be utilised to introduce perimeter flow control policies to improve mobility in single-region homogeneous networks (4, 8, 9) and multi-region heterogeneous networks (10, 11, 12). The general idea of a perimeter flow control policy is to “meter” the input flow to the system and to hold vehicles outside the controlled area if necessary, so as to maximise the throughput.

Despite these findings for the existence of a fundamental diagram, these curves should not be a universal law. Recent works (13, 14, 15) have identified the spatial distribution of congestion in the network as one of the key components that affect the scatter of the fundamental diagram and its shape. Other works (16) have noticed that flow capacity in urban networks may be observed over a range of accumulation-values (region B in Figure 1), and thus the critical accumulation $\tilde{n}$ cannot accurately be specified or is subject to change due to adaptive signal control (16, 17, 18) or routing (19). On the other hand, recently developed feedback perimeter flow control (8, 12) relies on constant pre-specified set values $\hat{n} \approx \tilde{n}$.

These shortcomings of the fundamental diagram call for new developments aiming at extending previously proposed feedback perimeter flow control strategies, so as to improve their performance. This paper suggests an adaptive version of the strategy in (8) and (12) that allows for the automatic tracking of the critical accumulation to help maintain the accumulation near the optimal range of accumulation-values, while network’s throughput is maximised. To this end, we propose a Kalman filter-based estimation algorithm that utilises real-time measurements of flow and accumulation of vehicles to produce estimates of the currently prevailing critical accumulation (similar to (20) for motorway traffic). The paper reports some preliminary results from the
evaluation of the estimation algorithm with real experimental data from an urban area (21).

PERIMETER AND BOUNDARY FLOW CONTROL STRATEGIES

The objective of perimeter flow control or gating is to protect urban regions from over-saturation in the sense of limiting the entrance in the network when it is close to overload. In the sequel, we briefly describe three recently proposed perimeter flow control strategies of homogeneous and heterogeneous networks and identify the role and need of pre-specified critical values.

Homogeneous single-region cities

In the case of a homogeneous single-region city (Figure 2(a)) an optimal gating policy is to allow as many vehicles to enter the network as possible without allowing the accumulation to reach states in the congested regime. This policy can be formulated as follows (4): when the network operating in the uncongested regime ($n < \tilde{n}$), vehicles are allowed to enter the perimeter of the network as quickly as they arrive with respect to the critical accumulation $\tilde{n}$; once accumulation reaches $\tilde{n}$ (i.e. $n \geq \tilde{n}$) entrance to the network is limited to the minimum entrance flow. This policy corresponds to the so-called “bang-bang control” in control theory given by

$$q_{in}(k) = \begin{cases} q_{\text{max}} & \text{if } n(k) < \tilde{n} \text{ and } n(k+1) < \tilde{n} \\ q_{\text{min}} & \text{else} \end{cases}$$

where $q_{in}$ is the flow that is allowed to enter the network from the perimeter; and $q_{\text{min}}$, $q_{\text{max}}$ are the minimum and maximum entrance flows, respectively. The bang-bang controller (Equation 1) works well when the system under consideration has relatively slow dynamics, but tends to oscillate between the extremes $q_{\text{min}}$ and $q_{\text{max}}$ as demonstrated in (12).

Alternatively the following Proportional-Integral (PI) feedback controller is well suitable for smooth and efficient operations (8)

$$q_{in}(k) = q_{in}(k-1) - K_p \left[ \text{TTS}(k) - \text{TTS}(k-1) \right] - K_I \left[ \text{TTS}(k) - \hat{\text{TTS}} \right]$$

where TTS (veh·h) is the Total Time Spent (a proxy of accumulation); $\hat{\text{TTS}}$ is a desired value of TTS where the Total Traveled Distance (TTD) is maximised (according to the fundamental
FIGURE 2 A network modelled as (a) single-region system and (b) multi-region system.

diagram, see Figure 2(a). TTS and TTD are proxies of accumulation and circulating flow, respectively); and, $K_p$ and $K_I$ are design parameters. This controller aims at stabilising TTS around the selected set point TTS. A well-known feature of the PI regulator is that the regulator error becomes automatically zero, i.e. $TTS = \hat{TTS}$ under stationary conditions. For the design of the controller (Equation 2) in (8) the corresponding fundamental diagram is constructed in terms of TTS and TTD as it has been observed in a field evaluation study, see Figure 6 in (22) and the related comments therein.

**Heterogeneous multi-region cities**

In the case of multi-region cities (Figure 2(b)), a single-region policy may induce uneven distribution of vehicles in the regions, and, as a consequence, may invalidate the homogeneity assumption of traffic loads within the urban regions and degrade the total network throughput. In this case, a PI multivariable feedback regulator, which is suitable for smooth and efficient operations in multi-region networks reads (12)

$$\beta(k) = \beta(k - 1) - K_p [n(k) - n(k - 1)] - K_I [n(k) - \hat{n}]$$

(3)

where $\beta$ is a vector with elements $\beta_{ij}$ the fraction of the gated flow that enters region $j$ from region $i$; $n$ is a vector with elements $n_i$, the accumulation of vehicles in region $i$; $\hat{n}$ is a set (desired) value for the accumulation of vehicles in each region $i$; $K_p$ and $K_I$ are the proportional and integral gains (design parameters), respectively. This controller aims at stabilising $n(k)$ around the selected set value $\hat{n}$. Typically, but not necessarily, $\hat{n} \approx \tilde{n} = (\tilde{n})_i$ may be selected for each region $i$, in which case the individual regions’ output is maximised (become close to flow capacity), see Figure 2(b) for a sketch.

**METHODOLOGY**

**Motivation**

The feedback controller in Equation 3 (or Equation 2) aims at stabilising the accumulation of vehicles (or a proxy of accumulation, e.g. TTS, average occupancy or density) of the network fundamental diagram around a pre-specified set point $\hat{n} \approx \tilde{n}$ to achieve maximum network throughput,
where $\tilde{n}$ should be chosen according to Figure 2(b) (respectively, Figure 2(a)). The design of the feedback controller involves two stages (see (8) and (12)): (a) linearisation of the original nonlinear traffic dynamics around $\hat{n} \approx 90\% \tilde{n}$, to allow specification of a nonzero derivative that prevents traffic to visit the congested regime; and (b) specification of control gains $K_p$ and $K_I$ (or $K_p$ and $K_I$ in (8)). Thus the value of $\tilde{n}$ should be known beforehand and for practical reasons it is estimated visually from the fundamental diagram of one representative day. Note that this value is also required for the real-time execution of Equation 3 with given parameters $K_p$, $K_I$ and state measurements $n(k)$. In this work, we develop an estimation algorithm that allows the automatic monitoring of the critical accumulation and propose an adaptive perimeter control strategy that relies on real-time estimates of critical accumulation to help maintain the accumulation near the optimal range of accumulation-values, while network’s throughput is maximised. The following observations furthermore support the need for an estimation algorithm of critical accumulation and adaptive perimeter (and boundary) flow control:

- The flow capacity in urban road networks may be observed over a range of accumulation-values (see Figure 1; (8, 12, 16) and the related comments therein) in contrast to motorway traffic where capacity flow is deemed to occur for a (more or less) specific density value; Physically speaking, flow capacity may be observed for different accumulation-values due to the spatiotemporal distribution of congestion in the two-dimensional network.
- The spatial distribution of congestion in the network is one of the key components that affect the scatter of an MFD and its shape. Recent works (13, 14, 15) have observed that the average network flow is consistently higher when link density variance is low for the same network density, but higher densities can create points below an MFD (analogous to capacity drop in motorways) when they are heterogeneously distributed.
- The critical accumulation $\tilde{n}$ cannot accurately be specified or is subject change due to traffic-responsive signal control (see e.g. (16, 17, 18)), traffic composition (e.g. cars versus buses, see (23)), or non-recurrent day-to-day traffic patterns (as shown later in the paper).
- Recent studies (24, 25) have shown that the location of loop detectors can affect the shape of the fundamental diagram and the value of critical accumulation, as the occupancy value is representative in the proximity of the detector and not for the whole link.
- In the case of heterogeneous networks with multiple regions of attraction the specification of set points (critical accumulation for each region) would require some care to prevent regions with high density of destinations from entering the congested state. For example if heavily directional flows from the periphery of a network pass through a small region to enter the center, the set point for the small region should be smaller than the set point of the periphery (16).

**An adaptive perimeter flow control strategy**

The aforementioned shortcomings of the fundamental diagram call for the development of an adaptive perimeter flow control strategy that relies on real-time estimates of critical accumulation $\tilde{n}$ (or the critical value of any other parameter of the fundamental diagram). To produce real-time estimates of $\tilde{n}(k)$ we utilise measurements of circulating flow $q(k - 1)$ and accumulation of vehicles $n(k - 1)$. Then the estimate $\hat{n}(k)$ is used as set value in Equation 3 or 2 (we may accordingly estimate occupancy, density or $\tilde{TTS}$), i.e. $\hat{n} = \tilde{n}(k)$. The corresponding adaptive control scheme is shown in Figure 3. Next section presents the estimation algorithm for $\hat{n}(k)$. 

---

TRB 2015 Annual Meeting  
Paper revised from original submittal.
Estimation of Critical Accumulation

The basic idea behind the proposed Kalman filter-based (KF) estimation algorithm is to utilise real-time measurements of circulating flow $q(k-1)$ and accumulation of vehicles $n(k-1)$ to produce estimates of the derivative of the fundamental diagram. The derivative of the fundamental diagram gives us valuable information about the operation of the network to the stable (positive derivative) or unstable regimes (negative derivative). These estimates may then be used to produce estimates of the currently prevailing critical accumulation $\hat{n}(k)$, for which the throughput of the network is maximised. Our approach, with some modifications, is similar to (20) and (26) for the case of fundamental diagrams of one-dimensional uninterrupted motorway traffic.

To start with, we assume that the evolution of the derivative $\delta(k)$ (state equation of the KF) and its observed output (output equation of the KF) are described by a random walk

\begin{align}
\delta(k) &= \delta(k-1) + \gamma(k) \\
\delta^o(k) &= \delta(k) + \zeta(k)
\end{align}

where $\delta^o(k)$ is the observed output; $\gamma(k)$ and $\zeta(k)$ are independent, zero-mean, Gaussian processes (white-noise) whose variance should be chosen so as to reflect the typical time variation of the corresponding variables. The derivative $\delta(k)$ in Equation 5 may be calculated at each time step $k$ by utilising real-time measurements of $q(\kappa)$ and $n(\kappa)$, $\kappa = k-2, k-1$ (first-order backward difference), as follows

\begin{align}
\delta(k) &= \frac{\Delta q(k-1)}{\Delta n(k-1) + \varepsilon}
\end{align}

where $\Delta q(k-1) = q(k-1) - q(k-2)$ and $\Delta n(k-1) = n(k-1) - n(k-2)$; and $\varepsilon \neq 0$ is a small parameter in the denominator so that possible high-frequency oscillations of the derivative be suppressed.
The resulting Kalman filter reads (exponential smoothing)
\[
\hat{\delta}(k) = K \delta^\alpha(k) + (1 - K)\hat{\delta}(k - 1), \quad \hat{\delta}(0) = \hat{\delta}_0
\]
with \( K \in [0, 1] \) typically specified according to the Kalman filter theory, although may manually specified according to the physics of the problem, i.e. if we trust the filter \( \hat{\delta}(k) \) (or the output \( \delta^\alpha(k) \)) then \( K < 0.5 \) (respectively, \( K > 0.5 \)) must be selected. This simple exponential filter indicates that the filtered measurement \( \hat{\delta}(k) \) is a weighed sum of the current measurements \( \delta^\alpha(k) \) and the filtered value at the previous sampling instant \( \hat{\delta}(k - 1) \).

A more accurate and detailed Kalman filter can be derived by calculating the derivative of the fundamental diagram around the critical accumulation (or occupancy) \( \hat{n} \) where approximately flow capacity is observed (20). In this case, \( \delta(k) \) may be written as
\[
\delta(k) = \frac{q(k) - Q}{n(k) - \hat{n}}
\]
where \( Q \) is the flow around the current estimate \( \hat{n} \). Note that, if \( \delta(k) \approx 0 \) then flow \( q = Q \) may be viewed as the capacity of the fundamental diagram. Considering now a filter with two state variables \( \mathbf{x} = [\delta \quad Q]^T \) that is described by a random walk
\[
\mathbf{x}(k) = \mathbf{x}(k - 1) + \mathbf{\gamma}(k)
\]
and one output equation (rearranging in Equation 8)
\[
q^\alpha(k) = \varphi(k)\mathbf{x}(k) + \zeta(k)
\]
where \( q^\alpha(k) \) is the observed output and \( \varphi = [(n - \hat{n}) \quad 1] \) is the output vector; \( \mathbf{\gamma}(k) \) and \( \zeta(k) \) are independent, zero-mean, Gaussian processes (white-noise) with covariance \( \Gamma \) and variance \( Z \), respectively. The estimate \( \hat{\mathbf{x}}(k) \) that minimises the conditional expectation, given past observation
\[
\mathbf{\Pi}(k) = \mathbf{E}\{ (\hat{\mathbf{x}}(k) - \mathbf{x}(k))(\hat{\mathbf{x}}(k) - \mathbf{x}(k))^T \}
\]
is given by the Kalman filter
\[
\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k - 1) + \mathbf{K}(k - 1)\left[q^\alpha(k) - \varphi(k)\hat{\mathbf{x}}(k - 1)\right], \quad \hat{\mathbf{x}}(0) = \hat{\mathbf{x}}_0
\]
where \( \hat{\mathbf{x}} = [\hat{\delta} \quad \hat{Q}]^T \) and \( \mathbf{K} \) is the filter gain. The Kalman filter gain \( \mathbf{K} \) depends upon \( \Gamma \) and \( Z \) and is calculated from (27, 28)
\[
\mathbf{K}(k - 1) = \left[ \mathbf{P}(k - 1) + \Gamma \right] \varphi^\tau(k) \left\{ \varphi(k)\left[ \mathbf{P}(k - 1) + \Gamma \right] \varphi^\tau(k) + Z \right\}^{-1}
\]
while the error covariance matrix \( \mathbf{P} \) is updated according to
\[
\mathbf{P}(k) = \mathbf{P}(k - 1) + \Gamma - \mathbf{K}(k - 1)\varphi(k)\left[ \mathbf{P}(k - 1) + \Gamma \right], \quad \mathbf{P}(0) = \mathbf{P}_0
\]

This problem of parameter estimation can be seen as a special case of recursive system identification (27, 29). The recursive filter in Equations 12–14 is executed starting from a given initial state \( \hat{\mathbf{x}}(0) = \mathbf{E}\{\mathbf{x}(0)\} \) and \( \mathbf{P}(0) \).
Our challenge is to devise a scheme for the real-time estimation of the critical accumulation based on the sign (positive or negative) of the estimated derivative $\delta$ from Equation 7 or 12 (20). In the simplest scheme an initial critical occupancy $\tilde{n}(0)$ is chosen; then in each time step $k$ of the algorithm if the corresponding derivative (KF in Equation 7 or 12) is sufficiently positive (respectively, negative), the new estimate of $\tilde{n}(k)$ is produced by adding (subtracting) an increment $\Delta^+$ (respectively, $\Delta^-$) to the current estimate $\tilde{n}(k-1)$. Nominal values are $\Delta^+ = 1\%$ and $\Delta^- = 1.2\%$ of normalised vehicle accumulation or occupancy. On the other hand, if the value of the derivative $\delta(k)$ is found to be around zero, the new estimate $\tilde{n}(k)$ is set equal to $\tilde{n}(k-1)$. The reason of introducing two increments $\Delta^+$ and $\Delta^-$ (in contrast to (20)) is to allow a different rate of increase (decrease) of the critical accumulation during the onset (respectively, offset) of congestion. In this way, the algorithm allows a fast recovery of the critical accumulation during abnormal traffic conditions in the network (or whenever the fundamental diagram is not well-defined). The algorithm could run in the background and should be activated at each control interval $T_E$ and only within specific time windows, e.g. by use of two thresholds $n_{act}$ and $n_{stop}$. The interval $T_E$ is a multiple of the time period $T$ of the perimeter control strategy. The algorithmic scheme of the estimation algorithm is illustrated in Algorithm 1 below.

**Algorithm 1: Estimation Algorithm for Critical Accumulation $\tilde{n}$**

**Data:** Initial value of the KF $\tilde{n}(0)$; $\tilde{n}_{min}$; $\tilde{n}_{max}$; $\Delta^+ \approx 1\%$, $\Delta^- \approx 1.2\%$; $\Delta\delta$; $D^-; D^+$

**Result:** Estimate of the critical $\tilde{n}(k)$ at each $k$

1. Enter new measurements $q(k-1)$ and $n(k-1)$ (aggregated to 5-10 min);
2. Calculate the derivative $\delta(k)$ according to Equation 6 or 8
3. Apply a rate-of-change filter to the calculated derivative:

$$
\delta(k) = \begin{cases} 
\hat{\delta}(k), & \text{if } |\hat{\delta}(k) - \hat{\delta}(k-1)| \leq \Delta\delta \\
\delta(k-1) - \Delta\delta, & \text{if } \delta(k-1) - \hat{\delta}(k) > \Delta\delta \\
\delta(k-1) + \Delta\delta, & \text{if } \hat{\delta}(k) - \delta(k-1) > \Delta\delta
\end{cases}
$$

(15)

4. Calculate the derivative $\hat{\delta}(k)$ according to Equation 7 or 12;
5. Set $\tilde{n}(k) = \Pi\{\tilde{n}(k-1) + s(k)\}$, where $\Pi$ is the projection of the calculated critical occupancy onto $[\tilde{n}_{min}, \tilde{n}_{max}]$ and $s$ is a saturation function as follows

$$
s(k) = \begin{cases} 
\Delta^+, & \text{if } \delta(k) > D^+ \\
-\Delta^-, & \text{if } \delta(k) < D^- \\
0, & \text{otherwise}
\end{cases}
$$

(16)

6. Set $\tilde{n}(k) = \tilde{n}(k)$ for the perimeter flow control strategy;
7. Set $k := k + 1$; go to step 1

Step 3 (Equation 15) of the algorithm checks if the calculated derivative changes suddenly by a large amount and then returns close to the original value at the next time instant. To this end, a rate-of-change filter is used to limit the maximum allowable change of the filtered derivative to $\Delta\delta > 0$. Given that our experimental data are very noisy (as we will see later) step 3 is very crucial for the proper function of the overall adaptive scheme. If noise derivatives are not removed
by filtering before the calculated derivative is sent to Step 4, the algorithm will produce large, sudden changes in estimated critical accumulation. Step 3 may be omitted or relaxed (via suitable \( \Delta \delta \)) if the available data are aggregated to a low resolution (e.g. 10-20 min values).

**EXPERIMENTAL DATA ANALYSIS AND PRELIMINARY RESULTS**

**Site and data description**

The test site is the central business district (CBD) of Chania (Greece), including about 24 closely spaced signalised junctions and 71 links with lengths varying from 50 to 500 m (Figure 4). The number of lanes for through traffic varies from 2 to 5 lanes and the free flow speed is 45 km/h. Traffic signals are all multiphase operating under the commercial semi-real-time signal control strategy “Traffic-Actuated Signal plan Selection” (TASS) by Siemens. TASS selects, every 15 min, one out of six fixed pre-defined network signal plans (each with different cycle times of 60-100 s, splits, and offsets), depending on the current traffic conditions in the network, as reflected by the measurements of 17 “strategic” detectors placed at appropriate network locations. Typical loop-detector locations within the Chania urban network links are either around the middle of the link or some 40 m upstream of the stop line.

The CBD is congested during the weekday’s peaks (especially in the summer due to tourism), with average speeds dipping below 10-15 km/h, which may sometimes lead to partial gridlock situations. It should be noted that traffic conditions in Chania are quite different, even among weekdays, due to differences in shop opening times. Real flow-occupancy data from 70 loop detectors (some detectors covering more than one link) and spanning one week, were available for the testing of the proposed estimation algorithms. The real data were available from a field evaluation of the TUC/HYBRID signal control strategy and the TASS strategy in May-June 2006 (21). The data were available in 1.5-min samples (every 90 s) of flow (vehicles per 1.5 min) and occupancy (%) and were aggregated to 6-min all-lanes data via averaging of lane occupancies and summing of flows. The next section analyses experimental data collected from the operation of the TASS strategy from June 5 to June 11, 2006. These data are then used in an offline mode to test the proposed estimation algorithm. Evaluation results of the coupled parameter estimation and adaptive perimeter flow control problem will be reported in an extended version of this paper.

**Real data analysis and some remarks on hysteresis loops**

We now show with experimental data that the CBD of Chania exhibits a network fundamental diagram with low scatter. We demonstrate that the fundamental diagram is reproduced under different traffic conditions (different days) but its shape and critical occupancy depend on the applied semi-real-time signal control and the distribution of congestion in the network.

Figure 5 depicts experimental data (flow-occupancy time series) for three representative weekdays and a weekend day, and the corresponding fundamental diagrams (circulating flow versus occupancy). Each measurement point on the fundamental diagram corresponds to 90 s. Remarkably, Figure 5 (see fundamental diagrams on the right) confirms the existence of a fundamental diagram for the CBD urban area of Chania with moderate scatter across different days (aggregated data may produce lower scatter, not shown here). It can be seen that flow capacity

---

1Note that proxies of occupancy (e.g. density or accumulation) may be obtained by use of suitable transformations, if the measured time-occupancy is assumed to approximately reflect the link’s space-occupancy.
(around 800 veh/cycle) occur in different occupancies across different days and only a part of region C is formed (congested regime in Figure 1). After careful inspection of Figure 5, the following observations may be highlighted:

- On a regular weekday (Monday, see Figure 5(a)), flow capacity is observed at $\hat{\delta} \approx 20\%$ and a part of the congested regime C (see Figure 1) of the fundamental diagram is actually visited (notice the negative slope).
- On Tuesday and Friday (Figure 5(b)-(c)) where market and shops are open in the evening (see occupancy 18:00-22:00), flow capacity is observed over a range of occupancies (18-35% and 20-25% respectively); this is attributed to the distribution of congestion (cf. flow and occupancy at 18:00-22:00) and the applied semi-adaptive signal control. The spatiotemporal distribution of congestion may be seen from the average flow from 18:00 to 20:00 on both weekdays where the average occupancy is sub-maximal on Thursday. The impact of applied semi-real-time signal control (TASS strategy) to the fundamental diagram on Tuesday and Friday may be inspected from the difference in occupancy levels during the saturated traffic conditions of region B in Figure 1, i.e. from the highest occupancy that is reached in each day by different signal control plans. The fundamental diagrams in Figure 5(b)-(c) indicate that the TASS strategy maintains the overall throughput to high values (flow capacity) during the heart of the rush on Tuesday (it is seen to reach up to 35% occupancies in the network, with a fully formed region B, see Figure 1), but it fails to maintain the same high values of throughput on Friday (although with lower traffic demand) and a part of the congested regime C of the fundamental diagram is visited (notice the negative slope, which is slightly higher than the corresponding slope on Monday). This is attributed to the spatiotemporal distribution of congestion in the network.
- On Saturday (Figure 5(d)) with low demand a part of the fundamental is formed, and flow capacity is deemed at $\hat{\delta} \approx 20\%$.

From the above observations, it is evident that although the CBD exhibits a fundamental diagram the critical accumulation cannot accurately be specified or is subject change due to semi-adaptive signal control and non-recurrent day-to-day traffic patterns. This calls for real-time estimation of the critical occupancy that exploit the available real-time loop detector measurements in the best possible way, particularly under different traffic conditions. Next section presents some preliminary results on the performance of the proposed estimation algorithm in monitoring the prevailing critical occupancy in the CBD.
FIGURE 5 Experimental data for three representative weekdays and a weekend day. Each measurement point on the fundamental diagram corresponds to 90 s. (a) Regular weekday, flow capacity is observed at $\hat{o} \approx 20\%$; (b, c) Market is open on Tuesday and Friday evenings (see occupancy 18:00-22:00), with congested traffic conditions and due to adaptive signal control flow capacity is observed over a range of occupancies (18-35\% and 20-25\%, respectively); (d) With low demand a part of the fundamental diagram is formed, flow capacity is deemed at $\hat{o} \approx 20\%$. 

TRB 2015 Annual Meeting

Paper revised from original submittal.
We close this section with an important remark on the hysteresis phenomenon in two-dimensional urban networks. It can be seen from Figure 5 that the fundamental diagram of CBD (remarkably with high data resolution) does not indicate a hysteresis loop, i.e. a different path of measurement points when filling the network than when emptying the network, which contradicts earlier simulation-based studies for the same network (8, 16). To verify this statement, we plotted and inspected the fundamental diagram of each day (one week) with different colours in the onset and offset of congestion (not shown here). The authors in (16) attributed the hysteresis phenomenon in a simulation-based study of the CBD to non-adaptive drivers (effect of dynamic traffic assignment) and the non-conservative nature of flow in two-dimensional networks. The fact that the hysteresis loop disappears in the real network of Chania supports this interpretation.

Our conjecture is that hysteresis phenomena may occur in real two-dimensional urban networks in response to the non-conservative nature of flow, even if the network operates under Wardrop equilibria. Note that during light traffic (uncongested network) user equilibrium tends to the system optimum (see e.g. (30)); on the other hand during congested traffic conditions it is essential to influence user traveling behaviour to reach the system optimum. This may be only effectuated through information provision and combined traffic assignment and signal control, which is still an open problem in transportation research. It is also our opinion that the fundamental diagram is a property of the infrastructure and thus hysteresis phenomena may occur in off-grid networks due to capacity bottlenecks that are the most vulnerable points in a network. Other culprits include arterial networks with limited number of alternative routes and special (e.g. with short links) grid structures, in this case the two-dimensional traffic may be viewed as one-dimensional motorway traffic, and thus the network fundamental diagram virtually reduces to an one-link fundamental diagram (see (31) for the hysteresis phenomenon in individual links).

FIGURE 6 Derivative estimate (with and without filtering) of the fundamental diagram on Tuesday, June 6, 2006; Real-time critical occupancy estimate for flow-occupancy data of Figure 5(b).
Preliminary evaluation of the estimation algorithm

In this section we present some preliminary results on the performance of the proposed estimation algorithm in monitoring the prevailing critical occupancy $\tilde{o}_{cr}$ in the CBD. Although the general algorithmic scheme 1 concerns the real-time estimation of critical accumulation the same algorithm may be used in estimating critical occupancies. For the application of the estimation algorithm to the data in Figure 5 the following initial values were selected: $o_{act}(0) > 10\%$, $\tilde{o}_{cr}(0) = 20\%$, $[\tilde{o}_{min}, \tilde{o}_{max}] = [10\%, 30\%]$, $\Delta^+ = 1\%$, $\Delta^- = 1.1\%$; data were aggregated to 6-min. The limits of the derivative were chosen $D^+ = 100$ and $D^- = -60$ given that only a part of region C is actually visited at the heart of rush (notice the negative slope of the fundamental diagram in Figure 5, which is smaller (absolute value) than the corresponding positive slope in region A).

Figure 6 depicts the results of algorithm in Equation 7 for the flow-occupancy data in Figure 5(b) where the critical occupancy is seen to depend on the applied semi-real-time signal control (the CBD is seen to reach up to 35% occupancies, with a fully formed region B). The first subfigure displays the obtained derivative (Equation 6) before the actual filtering. It can be seen that the derivative tends to oscillate (with high-frequency) between extremes (negative and positive), which is reasonable given that our experimental data are noisy (for this reason the obtained data were aggregated to 6-min all-lanes). The second subfigure depicts the estimated derivative after the application of the rate-of-change filter in step 3 of the algorithm. It is evident that the high-frequency oscillations of the derivative are suppressed after the application of the filter. This allows us to produce very smooth critical occupancy estimates. The third subfigure displays the obtained estimates of critical occupancy. In the morning peak (8:00-12:00) where real occupancy is around 20% (see Figure 5(b), occupancy subfigure) the algorithm is seen to estimate a critical occupancy around 21% (no risk for over saturation). In the evening peak (18:00-22:00) occupancy increases to values around 30%; as can be seen the algorithm quickly increases $\tilde{o}_{cr}$ during the evening peak and reach a value of around 25%. Thus the algorithm is seen to estimate quickly and quite accurately the new actual critical occupancy. It should be noted that during the off-peak period the algorithm quickly decreases the critical occupancy to 15%. The goal of the estimation algorithm is to estimate the critical occupancy (given the derivative of the fundamental diagram) thus it does not track the actual occupancy. Note that relocating critical occupancy outside region B (see Figure 1) will jeopardise the performance of the proposed adaptive perimeter flow control strategy ($\tilde{o}$ is an input for the regulator in Equation 3, see Figure 3).

CONCLUSIONS

In this work, we proposed a real-time feedback perimeter flow control strategy that allows the automatic monitoring of the critical accumulation to help maintain the accumulation near the optimal range of accumulation-values, while network’s throughput is maximised. To this end, we designed a Kalman filter-based estimation algorithm that utilise real-time measurements of circulating flow and accumulation of vehicles to produce estimates of the currently prevailing critical accumulation. The developed estimation algorithm coupled with the proposed adaptive perimeter flow control strategy may be valuable whenever the network fundamental diagram is not well defined and the critical accumulation cannot accurately be specified or is subject change due to traffic-responsive signal control, traffic composition (e.g. cars versus buses), or non-recurrent day-to-day traffic patterns. We used real-data from a CBD with 70 sensors and showed that the urban area exhibits a fundamental diagram with low scatter. We demonstrated that the fundamental diagram...
is reproduced under different days but its shape and critical occupancy depend on the applied semi-
real-time signal control and the distribution of congestion in the network. Preliminary evaluation
results of the estimation algorithm with the experimental data indicated good estimation accuracy
and performance.

On-going work considers the calibration and evaluation of the estimation algorithm and
comparison of the proposed adaptive perimeter flow control strategy with previously developed
strategies.

REFERENCES

2. Ardekani, S., and Herman, R. Urban network-wide traffic variables and their relations. *Trans-
   International Symposium on Transportation and Traffic Theory*, Elsevier, Amsterdam, The
   Netherlands, 1987, N. Gartner and N. Wilson, Eds.
4. Daganzo, C. F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Trans-
5. Daganzo, C. F., and Geroliminis, N. An analytical approximation for the macroscopic funda-
   mental diagram of urban traffic. *Transportation Research Part B*, Vol. 42, No. 9, 2008,
   pp. 771–781.
7. Geroliminis, N., and Daganzo, C. F. Existence of urban-scale macroscopic fundamental dia-
   pp. 759–770.
8. Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., and Papageorgiou, M. Exploiting the
   fundamental diagram of urban networks for feedback-based gating. *Transportation Research
9. Haddad, J., and Shraiber, A. Robust perimeter control design for an urban region. *Transpor-
11. Geroliminis, N., Haddad, J., and Ramezani, M. Optimal perimeter control for two urban
    regions with macroscopic fundamental diagrams: A model predictive approach. *IEEE Trans-
12. Aboudolas, K., and Geroliminis, N. Perimeter and boundary flow control in multi-reservoir
    as determinant of urban network capacity. *Philosophical Transactions of the Royal Society A*,
15. Daganzo, C., Gayah, V., and Gonzales, E. Macroscopic relations of urban traffic variables:


