



A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks

K. Aboudolas *, M. Papageorgiou, A. Kouvelas, E. Kosmatopoulos

Dynamic Systems and Simulation Laboratory, Technical University of Crete, GR-73100 Chania, Greece

ARTICLE INFO

Article history:

Received 23 December 2008

Received in revised form 5 June 2009

Accepted 9 June 2009

Keywords:

Traffic signal control

Traffic congestion

Store-and-forward modeling

Rolling-horizon (model-predictive) control

Fundamental diagram of networks

ABSTRACT

The paper investigates the efficiency of a recently developed signal control methodology, which offers a computationally feasible technique for real-time network-wide signal control in large-scale urban traffic networks and is applicable also under congested traffic conditions. In this methodology, the traffic flow process is modeled by use of the store-and-forward modeling paradigm, and the problem of network-wide signal control (including all constraints) is formulated as a quadratic-programming problem that aims at minimizing and balancing the link queues so as to minimize the risk of queue spillback. For the application of the proposed methodology in real time, the corresponding optimization algorithm is embedded in a rolling-horizon (model-predictive) control scheme. The control strategy's efficiency and real-time feasibility is demonstrated and compared with the Linear-Quadratic approach taken by the signal control strategy TUC (Traffic-responsive Urban Control) as well as with optimized fixed-control settings via their simulation-based application to the road network of the city centre of Chania, Greece, under a number of different demand scenarios. The comparative evaluation is based on various criteria and tools including the recently proposed fundamental diagram for urban network traffic.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Urban road network congestion has been a problem of most municipalities around the world for several decades. Several measures have been proposed and partly implemented to reduce the traffic demand in urban areas, such as road pricing, access restrictions of various kinds, dedicated lanes and signal priority of public transport vehicles, bicycle lanes, etc. On the supply side, there is usually hardly any possibility (or political support) for road infrastructure extension; this calls for operational signal control strategies that exploit the available infrastructure in the best possible way, particularly under peak period congestion.

It is generally believed that real-time signal control systems responding automatically to the prevailing traffic conditions, are potentially more efficient than clock-based fixed-time control settings. On the other hand, the development of optimal network-wide real-time signal control strategies using elaborated network models is deemed infeasible due to the combinatorial nature of the related optimization problem (see e.g. Papageorgiou et al., 2003); as a consequence, any real-time feasible signal control strategy design must include some simplification, either in its modeling approach, or in its optimization algorithm, or in its extent of network coverage.

SCOOT (Hunt et al., 1982; Bretherton et al., 2004) and SCATS (Lowrie, 1982) are two well-known and widely used traffic-responsive strategies that function effectively when the traffic conditions in the network are undersaturated, but their

* Corresponding author. Tel.: +30 28210 37289; fax: +30 28210 37584.

E-mail addresses: aboud@dssl.tuc.gr (K. Aboudolas), markos@dssl.tuc.gr (M. Papageorgiou), tasos@dssl.tuc.gr (A. Kouvelas), kosmatop@dssl.tuc.gr (E. Kosmatopoulos).

performance was reported to deteriorate under congested conditions. Other field-operational elaborated model-based traffic-responsive strategies such as PRODDYN (Farges et al., 1983) and RHODES (Mirchandani and Head, 1998; Mirchandani and Wang, 2005) employ dynamic programming while OPAC (Gartner, 1983) employs exhaustive enumeration; due to the exponential complexity of these solution algorithms, the basic optimization kernel is not real-time feasible for more than one (or few) junctions and hence, interconnections between junctions must be addressed separately. More recently, a number of further research approaches have been proposed employing various computationally expensive numerical solution algorithms, including genetic algorithms (Abu-Lebdeh and Benekohal, 1997; Lo et al., 2001), multi-extended linear complementary programming (De Schutter and De Moor, 1998), and mixed-integer linear programming (Lo, 1999; Beard and Ziliaskopoulos, 2006); in view of the high computational requirements, the network-wide implementation of these optimization-based approaches might face some difficulties in terms of real-time feasibility.

A different design avenue for network-wide signal control is based on the store-and-forward modeling paradigm. Store-and-forward modeling of traffic networks was first suggested by Gazis and Potts (1963) and has since been used in various works, notably for road traffic control. This modeling philosophy describes the network traffic flow process so as to circumvent the inclusion of discrete variables and hence it allows for efficient optimization and control methods with polynomial complexity to be used for signal control of large-scale congested urban networks. On the other hand, the introduced modeling simplification allows only for split optimization, while cycle time and offsets must be delivered by other control algorithms, see Diakaki et al. (2003). A recently developed strategy of this type is the signal control strategy TUC (Diakaki et al., 2002) that has been successfully field-implemented in large networks of five cities in four different countries, see Kosmatopoulos et al. (2006) for recent field results.

TUC is based on a very convenient and simple Linear-Quadratic (LQ) multivariable regulator design approach with a posteriori consideration of the cycle-time and minimum-green constraints which is likely to reduce the achievable control performance. An extended approach that incorporates the constraints in the optimal control problem formulation was shown to lead to an open-loop quadratic programming problem (Aboudolas et al., 2009) with potential benefits over the simpler LQ control. For online application, the quadratic programming problem must be cast in a rolling-horizon framework, similarly to other aforementioned strategies (PRODDYN, OPAC, RHODES). The purpose of this paper is to investigate the efficiency of the rolling-horizon quadratic programming control (QPC) and to compare it with TUC and with optimized fixed-time control via simulation-based application to the road network of the city centre of Chania, Greece, under a number of different scenarios. The comparative evaluation is based on a number of criteria including the recently developed notion of a fundamental diagram for urban road networks.

2. Fundamental diagram of two-dimensional networks

The notion of a fundamental diagram (e.g. in the form of a flow-density curve) for freeways was recently found to apply (under certain conditions) to two-dimensional urban road networks as well; see Gartner and Wagner (2004) for simulation-based experiments; Geroliminis and Daganzo (2008) for real-data based investigations; Daganzo and Geroliminis (2008) and Farhi (2008) for analytical treatments. In fact a fundamental-diagram-like shape of measurement points was first presented by Godfrey (1969), but also observed in a field evaluation study by Dinopoulou et al. (2005), see Fig. 6 and the related comments therein.

Fig. 1 displays the typical shape of a fundamental diagram (FD) for urban road networks, where the y-axis reflects the total network flow (i.e. the sum of flows exiting the network links) or the total flow of vehicles reaching their respective des-

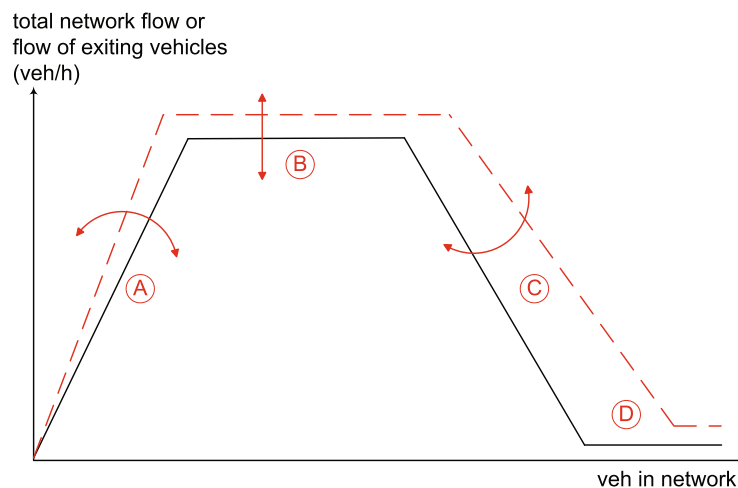


Fig. 1. Fundamental diagram for urban road networks.

tinations, while the x-axis reflects the number of vehicles present in the network. In the case of freeways, the FD is a result of the road infrastructure, the vehicle capabilities and the driver behaviour, but can also be influenced by some control actions such as variable speed limits (Papageorgiou et al., 2008) or other measures. In the case of urban road networks, the FD may also depend on the traffic pattern (origin–destination and routing of vehicles) as well as on the traffic signal operations. Thus, assuming that the traffic pattern at specific time-periods is comparable from day to day, the FD of urban road networks may be used for the comparative evaluation of different signal control strategies as attempted in Section 6.3.

Returning to Fig. 1, the traffic states on the rising line A reflect undersaturated traffic conditions (whereby vehicles waiting at signalized junctions are served during the next green phase), with green times being partially wasted due to lack of demand. Note that the slope of line A is proportional to the average speed in the network. This average speed may be changed (as indicated by the arrows in Fig. 1) via different traffic signal operations (splits, cycle, offsets). The traffic states on the horizontal line B reflect the network flow capacity that may also be subject to change via different signal settings. Note that capacity flow in urban road networks may be observed over a range of vehicle-numbers (hence the horizontal line) in contrast to freeway traffic where capacity flow is deemed to occur for a (more or less) specific density value. Traffic states along the line B are characterized by partial saturation, i.e. most network links experience saturation flow during the whole respective green phases, but no significant queue spillback to upstream junctions takes place.

When link queues start spilling back and blocking upstream junctions (leading to a waste of green time there), we enter the oversaturated region C; increasing vehicle-numbers within this region may lead to accordingly extended queue spillback occurrences and even partial gridlocks, increased waste of green times and, hence, lower total network flow. Better adapted signal control strategies may alter the region C in two possible ways: first, by increasing the vehicle-number at which region C starts, i.e. by extending the saturation region B towards higher vehicle-numbers; second, by increasing the (negative) slope of region C; both impacts lead to increased network flows at high vehicle-numbers. Finally, region D is characterized by a complete network-wide gridlock with very high vehicle-numbers and virtually zero flows, a situation that, once occurred, can hardly be alleviated by signal control.

A variety of real-time traffic signal control strategies for urban networks has been developed during the past decades, responding to the needs of individual cities/countries, the existing research and development base and advances in detection, communications and control technology. Without attempting a survey of this vast research area, see Papageorgiou et al. (2007) for an up-to-date account, we may distinguish two principal classes of signal control strategies. In the first class, strategies are only applicable to (or more efficient for) networks with undersaturated traffic conditions (regions A and partly B in Fig. 1). In the second class, we have strategies applicable to networks with saturated or oversaturated traffic conditions, whereby queues may grow in some links with an imminent risk of spillback and eventually even of gridlock in network cycles (regions B and C).

In principle, when traffic conditions are undersaturated (region A), the optimum signal control settings are determined from a knowledge of traffic demand and the saturation flows, aiming at minimizing the delay time at individual junctions as well as along arterials (via appropriate progression schemes). When the traffic network moves to state B, it appears appropriate for a split control strategy to balance the link queues so as to reduce the risk of queue spillback. Finally, when traffic conditions are entering region C, signal control strategies may need to apply gating so as to protect downstream links from overload. Note, however, that balancing the link queues may also be viewed as a way of gating when reaching region C.

The control strategies investigated in this paper attempt a balancing of link queues and are therefore most suitable for regions B and C. In fact, Farhi (2008) has recently shown that TUC improves the traffic conditions significantly when operating in regions B and C.

3. The investigated control strategies

3.1. Modeling

The urban road network is represented as a directed graph with links $z \in Z$ and junctions $j \in J$. For each signalized junction j , we define the sets of incoming I_j and outgoing O_j links. It is assumed that the offset and the cycle time C_j of junction j are fixed or calculated in real time by another algorithm. In addition, to enable network offset coordination within the present setting, we assume that $C_j = C$ for all junctions $j \in J$, which is a quite usual assumption. Furthermore, the signal control plan of junction j (including the fixed lost time L_j) is based on a fixed number of stages that belong to the set F_j , while v_z denotes the set of stages where link z has right of way (r.o.w.). Finally, the saturation flow S_z of link $z \in Z$ and the turning rates $t_{w,z}$, where $w \in I_j$ and $z \in O_j$, are assumed to be known and constant for LQ control but may be time-varying for the QPC approach.

By definition, the constraint

$$\sum_{i \in F_j} g_{j,i} + L_j = (\text{or } \leq) C \quad (1)$$

holds at junction j , where $g_{j,i}$ is the green time of stage i at junction j . Inequality in (1) may be useful in cases of strong network congestion to allow for all-red stages (e.g. for strong gating). In addition, the constraint

$$g_{j,i} \geq g_{j,i,\min}, \quad i \in F_j \quad (2)$$

where $g_{j,i,\min}$ is the minimum permissible green time for stage i at junction $j \in J$, is introduced to guarantee allocation of sufficient green time to pedestrian phases.

Consider a link z connecting two junctions M and N such that $z \in O_M$ and $z \in I_N$ (Fig. 2). The dynamics of link z are given by the conservation equation

$$x_z(k+1) = x_z(k) + T[q_z(k) - s_z(k) + d_z(k) - u_z(k)] \quad (3)$$

where $x_z(k)$ is the number of vehicles within link z (for the sake of brevity sometimes called queue in the following) at time kT , $q_z(k)$ and $u_z(k)$ are the inflow and outflow, respectively, of link z in the sample period $[kT, (k+1)T]$; T is the discrete-time step and $k = 0, 1, \dots$ the discrete-time index; d_z and s_z are the demand and the exit flow within the link, respectively. For the exit flow we set $s_z(k) = t_{z,0}q_z(k)$, where the exit rates $t_{z,0}$ are assumed to be known. The inflow to the link z is given by $q_z(k) = \sum_{w \in I_M} t_{w,z} u_w(k)$, where $t_{w,z}$ with $w \in I_M$ are the turning rates towards link z from the links that enter junction M . Queues are subject to the constraints

$$0 \leq x_z(k) \leq x_{z,\max}, \quad \forall z \in Z \quad (4)$$

where $x_{z,\max}$ is the maximum admissible queue length. These constraints may automatically lead to a suitable upstream gating in order to protect downstream areas from oversaturation during periods of high demand.

We now introduce a critical simplification for the outflow u_z that characterizes the utilized modeling approach. Provided that space is available in the downstream links and that x_z is sufficiently high (which is surveilled by constraints (4)), the outflow (real flow) u_z of link z is equal to the saturation flow S_z if the link has r.o.w., and equal to zero otherwise. However, if the discrete-time step T is equal to C , an average value for each period (modeled flow) is obtained by

$$u_z(k) = G_z(k)S_z/C \quad (5)$$

where G_z is the green time of link z , calculated as $G_z(k) = \sum_{i \in v_z} g_{j,i}(k)$. The consequences of this simplification for the derived signal control strategies are discussed in detail in Aboudolas et al. (2009).

3.2. Linear-Quadratic (LQ) optimal control (the TUC strategy)

Replacing (5) in (3) leads to a linear state-space model for road networks of arbitrary size, topology, and characteristics which is given by the linear state equation

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{B}\Delta\mathbf{g}(k) + T\Delta\mathbf{d}(k) \quad (6)$$

where $\mathbf{x}(k)$ is the state vector (consisting of the number of vehicles x_z of each link z); $\Delta\mathbf{g}(k) = \mathbf{g}(k) - \mathbf{g}^N$ and $\Delta\mathbf{d}(k) = \mathbf{d}(k) - \mathbf{d}^N$ are the control and demand deviations, respectively; $\mathbf{g}(k)$ is the control vector (consisting of all the green times $g_{j,i}$); \mathbf{g}^N is a nominal control vector (consisting of the nominal green times $g_{j,i}^N$) which corresponds to a pre-specified fixed signal plan; $\mathbf{d}(k)$ is the disturbance vector (consisting of the demand flows d_z of each link z); \mathbf{d}^N is a nominal disturbance vector, whereby $\mathbf{B}\mathbf{g}^N + T\mathbf{d}^N = \mathbf{0}$ holds for the nominal (e.g. steady-state) values. Finally \mathbf{B} results from (3) and (5) as a constant matrix of appropriate dimensions reflecting the network characteristics (topology, saturation flows, turning movement rates).

A quadratic criterion has the general form

$$\mathcal{J} = \frac{1}{2} \sum_{k=0}^{\infty} \left(\|\mathbf{x}(k)\|_{\mathbf{Q}}^2 + \|\Delta\mathbf{g}(k)\|_{\mathbf{R}}^2 \right) \quad (7)$$

where \mathbf{Q} and \mathbf{R} are diagonal weighting matrices. The diagonal elements of \mathbf{Q} are set equal to $1/x_{z,\max}$ in order to minimize and balance the occupancies $x_z/x_{z,\max}$ of the network links; while $\mathbf{R} = r\mathbf{I}$ (where \mathbf{I} is the unity matrix) with a very low value (e.g. 10^{-4}) given to the scalar weight r .

Minimization of the cost criterion (7) subject to (6) (assuming $\Delta\mathbf{d}(k) = \mathbf{0}$) leads to a linear multivariable feedback regulator given by

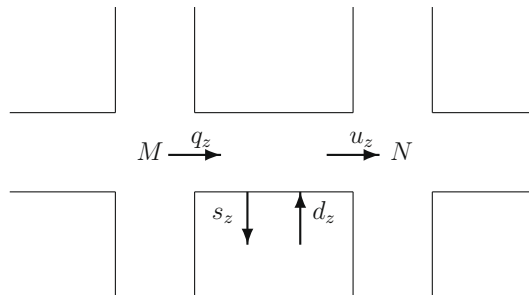


Fig. 2. An urban road link.

$$\mathbf{g}(k) = \mathbf{g}^N - \mathbf{L}\mathbf{x}(k) \quad (8)$$

where the feedback gain matrix \mathbf{L} results as a straightforward solution of the corresponding algebraic Riccati equation. This is the multivariable regulator approach taken by the signal control strategy TUC (Diakaki et al., 2002) to calculate in real time the network splits, while cycle time and offsets are calculated by other parallel algorithms (Diakaki et al., 2003).

Note that the LQ control theory does not allow for direct consideration of the constraints (1) and (2). For this reason, a suitable real-valued quadratic knapsack algorithm is applied after the application of (8) to modify the calculated $g_{j,i}$ green times of each junction so as to satisfy the constraints (1) and (2), see Aboudolas et al. (2009) for more details.

3.3. Quadratic-programming control (QPC)

In contrast to other store-and-forward based approaches (see for instance Singh and Tamura (1974)), we will now introduce the green times G_z of each link z as additional independent variables. The reason behind this modification is that we want to increase the control flexibility and potential efficiency while explicitly considering the queue constraints (4) (Papageorgiou, 1995; Aboudolas et al., 2009). The introduced link green times G_z are constrained as follows:

$$0 \leq G_z(k) \leq \sum_{i \in v_z} g_{j,i}(k), \quad \forall j \in J. \quad (9)$$

In view of this modification, replacing (5) in (3) for all z and organizing all resulting equations in one single vector-based equation leads to a linear state-space model for road networks of arbitrary size, topology, and characteristics which is given by

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \bar{\mathbf{B}}(k)\mathbf{G}(k) + T\mathbf{d}(k) \quad (10)$$

where $\mathbf{G}(k)$ is the link control vector (consisting of the green times G_z of each link z); $\bar{\mathbf{B}}$ results from (3) and (5) as a matrix of appropriate dimensions containing the network characteristics (topology, saturation flows, turning rates). Note that $\bar{\mathbf{B}}$ may be time-variant, if the involved saturation flows or turning rates are time-variant.

A suitable control objective under congested traffic conditions is to minimize the risk of oversaturation and spillback of link queues. To this end, one may attempt to minimize and balance the links' occupancies $x_z/x_{z,\max}$ via the following finite-horizon quadratic criterion

$$\mathcal{J} = \frac{1}{2} \sum_{k=0}^K \sum_{z \in Z} \frac{x_z^2(k)}{x_{z,\max}} \quad (11)$$

which is identical to (7) for $r = 0$ in (7) and $K \rightarrow \infty$ in (11). Alternatively, one may minimize the total time spent (which corresponds to minimization of the sum of x_z) but this would lead to a linear programming problem with vertex solutions (e.g. $x_z = 0$ for some links and $x_z = x_{z,\max}$ for others) that would increase the risk of link queue spillback.

On the basis of the linear model (10); the linear constraints (1), (2), (4) and the additional constraints (9); and the quadratic cost criterion (11); a (dynamic) optimal control problem may be formulated over a finite time-horizon K , starting with the known initial state $\mathbf{x}(0)$ in the state Eq. (10). This quadratic programming (QP) problem (with very sparse matrices) may be readily solved by use of broadly available codes or commercial software within few CPU-seconds even for large-scale networks and long time-horizons (Aboudolas et al., 2009).

4. The rolling-horizon (model-predictive) framework

For the application of the proposed QPC methodology in real time, the corresponding algorithm is embedded in a rolling-horizon (model-predictive) scheme. More precisely, the optimal control problem is solved on-line once per cycle using the current state (current estimates of the number of vehicles in each link) of the traffic system as the initial state $\mathbf{x}(0)$ as well as predicted demand flows over the finite horizon K . The optimization yields an optimal control sequence for K cycles, but only the first control (signal control plan) in this sequence is actually applied to the signalized junctions of the traffic network. More specifically, the rolling-horizon framework is as follows:

At time step k_0 , the QP problem is solved, based on a measured (or estimated) initial condition $\mathbf{x}(k_0)$ and on available demand predictions $\mathbf{d}(k)$, $k = k_0, \dots, k_0 + K - 1$, to obtain the controls $\mathbf{g}^*(k)$ and states $\mathbf{x}^*(k+1)$, $k = k_0, \dots, k_0 + K - 1$. However, only a part of the control trajectory is actually applied to the process, namely $\mathbf{g}^*(k)$, $k = k_0, \dots, k_0 + k_R - 1$, where $k_R \ll K$ (e.g. $k_R = 1$). Then, at time step $k_0 + k_R$, based on the new measured initial condition $\mathbf{x}(k_0 + k_R)$ (feedback) and updated demand predictions $\mathbf{d}(k)$, $k = k_0 + k_R, \dots, k_0 + k_R + K - 1$, the QP problem is solved again to obtain the controls $\mathbf{g}^*(k)$ and states $\mathbf{x}^*(k+1)$, $k = k_0 + k_R, \dots, k_0 + k_R + K - 1$, but only $\mathbf{g}^*(k)$, $k = k_0 + k_R, \dots, k_0 + 2k_R - 1$, is actually applied to the process, and so forth.

There are several important issues that are associated with the rolling-horizon framework just described:

- The saturation flows S_z and the turning rates $t_{w,z}$, may be time-variant, e.g. estimated or predicted in real time by well-known recursive estimation schemes (Cremer, 1991); in addition, the predicted demand flows $\mathbf{d}(k)$ may be calculated by use of historical information or suitable extrapolation methods (e.g., time series or neural networks).

- A satisfactory optimization horizon K should be in the order of the time needed to travel through the network. A much shorter optimization horizon may lead to “myopic” control actions.
- The computation time needed for the numerical solution of the QP problem must be short enough to permit the outlined repetitive on-line solution of the optimization problem. This is guaranteed for the present optimization method.
- The state variables \mathbf{x} (the number of vehicles in each link) must be measurable or be estimated in real time. Occupancy measurements collected via traditional detector loops may be utilized to estimate the numbers of vehicles within links via suitable nonlinear functions (Diakaki, 1999). The detector locations within links may be arbitrary, although the quality of estimation is best if the detectors are located around the middle of the link.

5. Application set-up

To demonstrate the efficiency and real-time feasibility of the proposed approach to the problem of urban signal control, the road network of the city centre of Chania, Greece, is considered. For this network, we compare the closed-loop behaviour of the LQ approach with the behaviour of the proposed QPC approach when embedded in a rolling-horizon control scheme as well as with optimized fixed signal plans. To ensure fair and comparable results, the three methodologies are evaluated by use of the same simulation model that is outlined in the next section.

5.1. The simulation model

The simulation model is simple but more accurate than the linear model (10) thanks to a nonlinear link outflow function that models the intra-cycle traffic flow process more accurately than (5). More precisely, we assume that the model time step is $T \ll C$ while the control time step T_c remains equal to C , i.e. control decisions are taken at each cycle. Then the outflow $u_z(k)$ is given by

$$u_z(k) = \begin{cases} 0 & \text{if any } x_{d,z}(k) \geq c x_{d,\max} \\ \min \left\{ \frac{x_z(k)}{T}, \frac{G_z(\kappa) S_z}{C} \right\} & \text{else} \end{cases} \quad (12)$$

where the index d refers to a downstream link of link z with turning rate $t_{z,d} \neq 0$, and we have the parameter c that should be selected close to 1; note that k is now the model discrete-time index (with time step $T \ll C$) while κ is the control discrete-time index (with time step $T_c = C$) and $\kappa = \text{int}(kT/T_c)$. Typical discrete-time model steps T for the traffic flow model (3) using (12) may be in the order of 5 s, while the control variables change their value in discrete-time control steps T_c , e.g. at each cycle. Note that, when using (12), the queue constraints (4) are considered indirectly and may hence be dropped; indeed the link outflow in (12) becomes zero if there is no vehicle in the link or if a downstream link is full. Note also that the basic simplification of store-and-forward modeling, i.e. a continuous link outflow (rather than zero flow during red and free flow during green), is still maintained in this approach.

Despite its relative simplicity, the simulation model reflects the essential phenomena of urban network traffic flow and is more accurate than the design models of the control strategies. In any case, our main interest here is in comparing different control approaches under the same simulated conditions rather than in deriving accurate absolute values of the control performance indexes for the specific network. To this end, the proposed simulation model appears to be appropriate.

5.2. Network and scenario description

The urban network of the city centre of Chania consists of $|J| = 16$ signalized junctions and $|Z| = 60$ links (Fig. 3). We omit the details on the (constant) turning rates $t_{w,z}$, lost times L_j , staging v_z and saturation flows S_z . The (fixed) cycle time in the network is $C = 90$ s, and $T_c = C$ is taken as a control interval for all strategies. Finally, for the simulation model we consider $T = 5$ s and $c = 0.85$ (i.e., overloaded links in (12) are considered the links z for which $x_z \geq 0.85 x_{z,\max}$).

Several scenarios were defined in order to investigate the behaviour of the control methodologies under different conditions. The simulation horizon for each scenario is 1 h (40 cycles). Fixed-time signal control, the linear multivariable feedback regulator (LQ), and the QPC approach embedded in a rolling-horizon scheme were applied and tested with the simulator for five demand scenarios with the following characteristics:

- (1) Very low demand in the network origins.
- (2) Low demand (25% higher than scenario 1).
- (3) Medium demand (50% higher than scenario 1).
- (4) High demand; in this scenario the network faces serious congestion for some 20 cycles (1/2 h) with some link queues spilling back into upstream links.
- (5) High demand presenting strong time fluctuations.

In addition to the demands at the network origins, moderate demands are also generated in some internal network links for all scenarios. In the first four scenarios, the shape of all demand trajectories is trapezoidal, thus featuring a constant-de-

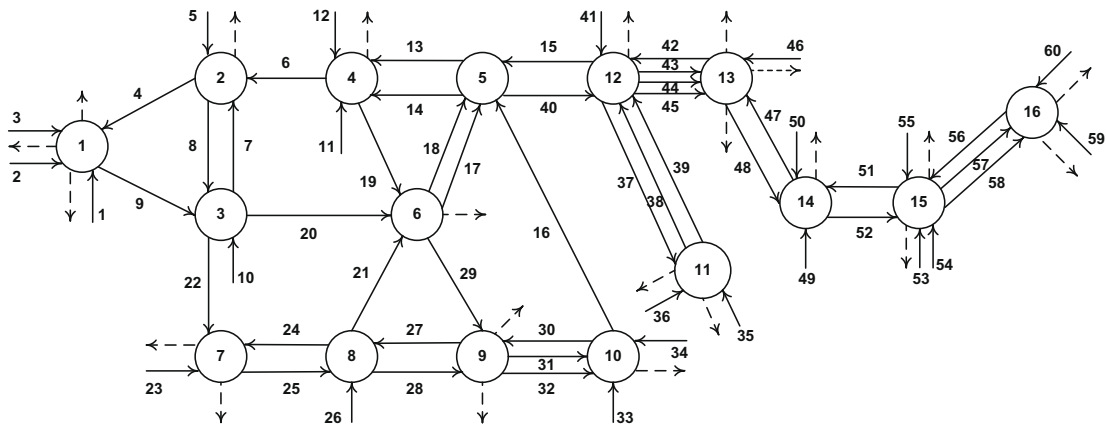


Fig. 3. The Chania urban road network.

mand peak period. In the fifth scenario, the peak demand varies quite strongly over time in order to enable an evaluation under more dynamically changing conditions. The scenarios include a final phase of zero demands which is long enough for the network to be completely emptied by all investigated signal control strategies, thus producing comparable results across the strategies.

5.3. Assessment criteria

For each of the five scenarios and for each control approach, two evaluation criteria are calculated for comparison via the simulation model. The total time spent

$$\text{TTS} = T \sum_{k=0}^{K_s} \sum_{z \in Z} x_z(k) \quad (\text{in veh} \cdot \text{h})$$

and the relative queue balance

$$\text{RQB} = \sum_{k=0}^{K_s} \sum_{z \in Z} \frac{x_z^2(k)}{x_{z,\max}} \quad (\text{in veh})$$

where K_s is the scenario time horizon.

5.4. Control strategy application

We consider the following signal control cases:

- Fixed-time signal control (FT-A) with a field-applied plan \mathbf{g}^N that is not fully adapted to the demand scenarios outlined above.
- Fixed-time signal control (FT-B) with plans \mathbf{g}^N optimized individually for each demand scenario. These signal plans were calculated for the specific demand scenarios using the QPC strategy as an off-line network optimization tool with $T_c = K_s T$, i.e. to specify optimal fixed greens over the whole scenario duration.
- Strategy LQ-A based on the linear multivariable feedback regulator (8) and the nominal plan \mathbf{g}^N of FT-A.
- Strategy LQ-B based on the linear multivariable feedback regulator (8) and the optimized nominal plan \mathbf{g}^N for each demand scenario (the FT-B optimized fixed plans).
- Rolling-horizon control scheme QPC-A without demand information; the QP problem is solved at each cycle (every $k_R = 1$) with investigated optimization horizons $K = 1, 2, \dots, 6$, i.e. 90 s, 180 s, ..., 540 s, respectively. In this case, it is assumed that there are no predicted demand flows $\mathbf{d}(k)$ available within the rolling-horizon control scheme, hence $\mathbf{d}(k) = \mathbf{0} \forall k$ is set (although non-zero demands are actually present in the simulations).
- Rolling-horizon control scheme QPC-B with perfect demand flow information; the QP problem is solved at each cycle (every $k_R = 1$) with investigated optimization horizons $K = 1, 2, \dots, 5, 9, 10, 20, 22, 25$. In this case, it is assumed that accurate demand flow predictions are available for the whole optimization horizon.

The real-time control strategies are applied with control interval $T_c = C$. All strategies are fed with the simulated \mathbf{x} -values (feedback) to make their decisions in real time. The QPC strategy is run with different optimization horizons K in order to investigate the impact of K on the control performance. It is expected that the results should tend to be better for greater

values of K (due to less myopic control actions); on the other hand, the lack of demand predictions in QPC-A may render the calculated controls increasingly outdated for long time-horizons K . Note also that the LQ approach does not involve any $\mathbf{d}(k)$ prediction by its design.

6. Simulation results

6.1. Evaluation based on global criteria

Table 1 displays the obtained results for the FT and LQ-variants. The optimized fixed plan control FT-B is seen to strongly outperform the FT-A signal plan for all scenarios. Thus, FT-B may be considered as a challenging (albeit rather idealistic) basis for the assessment of real-time control schemes. Table 1 shows that both LQ variants lead to significant reductions of both evaluation criteria compared to FT-B, which underlines the superiority of appropriate real-time decision-making even in case of optimized fixed control. The use of optimized nominal plans in LQ-B leads to better performances than LQ-A; nevertheless, the huge performance differences between FT-A and FT-B shrink to small or moderate differences between LQ-A and LQ-B, which indicates a relatively low sensitivity of the LQ approach to non-optimal \mathbf{g}^N plans in (8).

Table 2 displays the obtained results for the rolling-horizon QPC-A approach for different optimization horizons K . It can be seen that for $K \geq 2$ there are no significant deviations of the evaluation criteria for different optimization horizons K even for the high-demand scenario 4 and the strongly fluctuating and high-demand scenario 5; this is attributed to the complete lack of demand information ($\mathbf{d} \equiv \mathbf{0}$) that affects longer-term decisions. Since the required computational effort increases with increasing K , $K = 2$ seems to be a good choice. Comparing with Table 1, QPC-A is seen to be always better than FT-A, FT-B and LQ-A. Compared to LQ-B, QPC-A is seen to be better for scenarios 1–3 but slightly inferior in the heavy scenarios 4 and 5; this is attributed to the fact that the optimal \mathbf{g}^N incorporated in LQ-B is calculated with full demand knowledge, which is utterly missing in the QPC-A approach.

Table 3 displays the obtained results for the rolling-horizon QPC-B approach for different optimization horizons K . In this case, the availability of demand flow predictions allows for more reliable information about the evolution of the network traffic state in the future; as a consequence both evaluation criteria are seen to slightly improve as K is increased in some scenarios. In particular, for the high-demand and high-fluctuation scenario 5, the most satisfactory results with respect to both evaluation criteria are obtained with $K = 20$ (1/2 h). Comparing with Table 2, QPC-B outperforms QPC-A for all scenarios which underlines the utility of demand predictions compared to the zero-demand assumption. Comparing with Table 1, QPC-B outperforms FT-A, FT-B, LQ-A and LQ-B for all scenarios, as it incorporates all necessary information (demands and constraints) in its optimization problem.

Table 4 displays some percentage changes of both evaluation criteria for the six methodologies. LQ-A vs. FT-B demonstrates the superiority of appropriate real-time control over optimized fixed control. LQ-B vs. LQ-A indicates possible

Table 1

Assessment criteria for FT-A, FT-B, LQ-A and LQ-B strategies.

Scenario	Strategy							
	FT-A		FT-B		LQ-A		LQ-B	
	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB
1	21	1679	15	1411	13	561	14	702
2	99	36,403	56	13,532	44	7341	41	7187
3	231	97,947	201	79,578	164	53,993	146	48,282
4	481	218,565	353	142,363	285	100,136	250	89,946
5	594	259,101	412	170,329	348	130,891	311	120,760
Average	285	122,739	207	81,443	171	58,584	152	53,375

Table 2

Assessment criteria for the rolling-horizon QPC-A approach.

Scenario	K											
	1		2		3		4		5		6	
	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB
1	13	568	13	578	13	576	13	578	13	580	13	579
2	36	5827	36	5844	36	5838	36	5832	36	5811	36	5792
3	145	44,056	145	44,286	144	44,259	144	44,292	144	44,236	144	44,262
4	284	97,398	283	97,391	283	97,096	283	97,793	284	97,970	283	97,753
5	317	113,221	318	113,325	318	113,364	318	113,446	317	113,158	317	112,835
Average	159	52,214	159	52,285	159	52,227	159	52,388	159	52,351	159	52,244

Table 3

Assessment criteria for the rolling-horizon QPC-B approach.

Scenario	K													
	1		2		3		5		9		20		25	
	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB
1	11	281	11	280	11	280	11	280	11	280	11	279	11	279
2	30	4116	30	4087	30	4067	30	4043	30	4040	30	4030	30	4032
3	132	39,447	133	39,463	132	39,413	132	39,407	132	40,079	132	40,372	132	40,688
4	248	83,183	248	83,144	249	82,828	246	81,182	252	84,246	253	85,072	253	85,112
5	282	97,630	281	97,236	280	96,292	278	94,940	270	89,694	266	87,251	266	87,795
Average	141	44,931	141	44,842	140	44,576	139	43,970	139	43,668	138	43,401	138	43,581

Table 4

Comparison of assessment criteria.

Scenario	Strategy							
	LQ-A vs. FT-B		LQ-B vs. LQ-A		QPC-A (K = 2) vs. LQ-A		QPC-B (K = 9) vs. QPC-A (K = 2)	
	TTS (%)	RQB (%)	TTS (%)	RQB (%)	TTS (%)	RQB (%)	TTS (%)	RQB (%)
1	–7	–60	0	25	–7	3	–15	–52
2	–21	–46	–7	–2	–18	–20	–17	–31
3	–18	–32	–11	–11	–12	–18	–9	–9
4	–19	–30	–12	–10	–1	–3	–11	–13
5	–16	–23	–11	–8	–9	–13	–15	–21
Average	–17	–28	–11	–9	–7	–11	–13	–16

improvements within the LQ-approach by use of better nominal plans. QPC-A vs. LQ-A indicates the achievable improvements if the constraints (1), (2), (4) are incorporated in the optimization problem and link green times G_z are introduced; clearly, a more complex QP problem solution in real-time (compared with the simpler LQ regulator) is the price to pay for this particular improvement. Finally, QPC-B vs. QPC-A demonstrates the maximum range of achievable improvements via the use of demand predictions within the QPC-approach.

6.2. Some detailed results

In this section we report on some more detailed illustrative results focussing on the city's main shopping district (junctions 2, 4–6, 12 in Fig. 3). This area of the network features serious congestion during the morning, afternoon and evening peak hours with link queues that may spill back into upstream links.

For example, serious congestion develops in links 6 and 8 at junctions 2 and 3 in demand scenarios 4 and 5. This congestion spills back through links 13 and 15, and reaches link 38 (junction 12). Fig. 4a and b display the time evolution of the occupancy $x_z/x_{z,\max}$ within link 38 under FT-B, the LQ-variants, and the QPC-variants, for demand scenario 5. Moreover, Fig. 4c and d display the time evolution of the green time of stage 3 at junction 12 where link 38 has r.o.w., for the same strategies. These figures demonstrate that:

- Link 38 saturates for FT-B, LQ-A, LQ-B over a considerable period of time (Fig. 4a), that is longest for FT-B due to the rigid (constant) green time, while LQ-A and LQ-B apply higher green times during saturation (Fig. 4c). Note that the green times for both LQ-A and LQ-B take their respective nominal values g^N according to (8) when queues are close to zero, i.e. at the start and end periods of the simulation horizon.
- In the QPC-B strategy (for $K = 9$ or $K = 20$) the shape of the control trajectory (Fig. 4d) is shifted (horizontally) to the left by 4–5 cycles, when compared with QPC-A. By close examination, this is because the availability of accurate demand flow predictions in QPC-B allows for better prediction of the traffic network state in the near future (in contrast to QPC-A where $\mathbf{d} = \mathbf{0}$) and accordingly anticipated control actions. For the same reason, the link occupancies are slightly lower and smoother in the case of perfect demand flow predictions (Fig. 4b).
- The control trajectories in the LQ-variants (Fig. 4c) are smoother than in the QPC-variants (Fig. 4d). This is a general observation that is attributed to the infinite horizon in the LQ objective criterion (7) as opposed to the finite horizon of the QPC approach in (11), but also the a posteriori consideration of the constraints in the LQ approach.

To enable a more general evaluation with regard to the number of overloaded links, we define a network link as overloaded if its occupancy $x_z/x_{z,\max}$ is higher than 0.8; let $m(k)$ denote the number of overloaded links at the simulation cycle k for a specific control strategy; and let

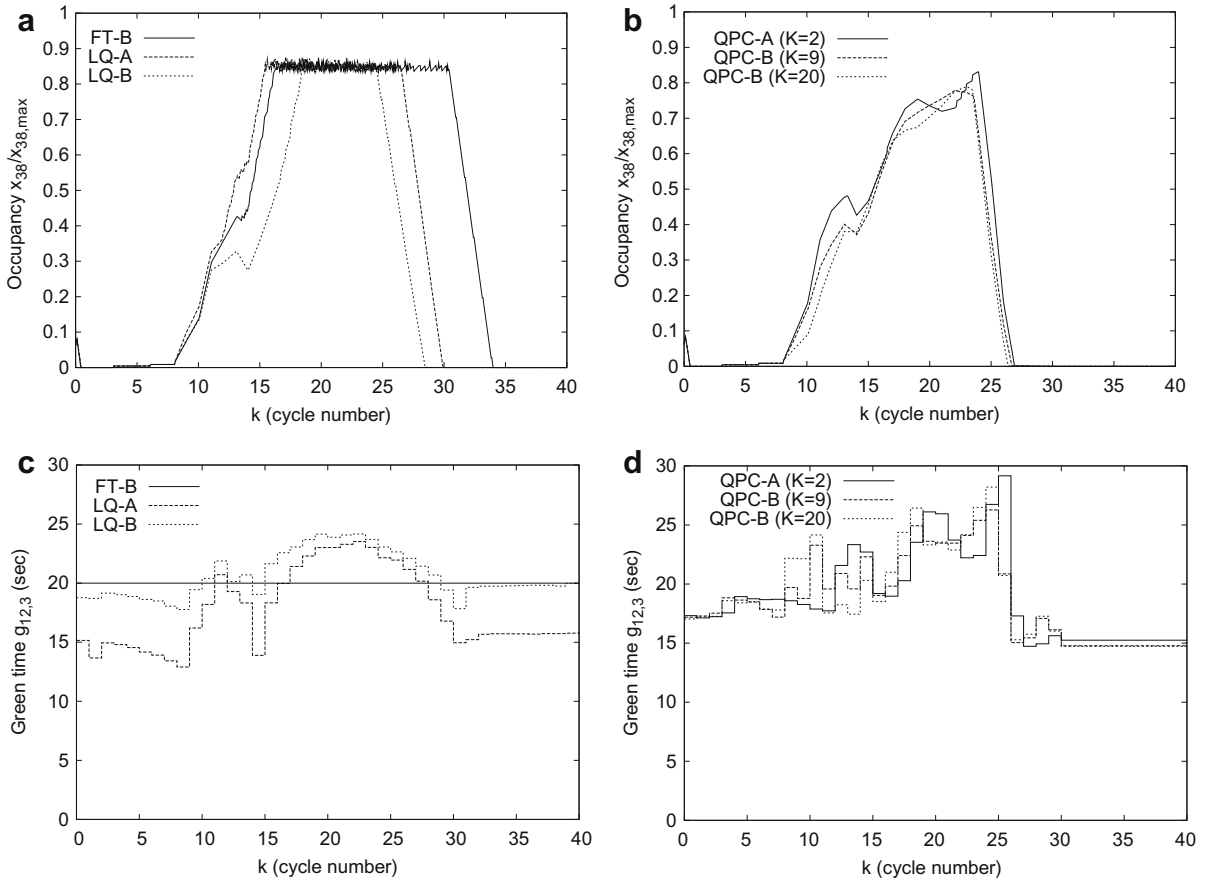


Fig. 4. Comparison of LQ-variants, QPC-variants, and FT-B for Scenario 5: (a), (b) occupancy of link 38 at junction 12; (c), (d) green time of stage 3 at junction 12 where link 38 has r.o.w.

$$M(k) = \sum_{\kappa=0}^k m(\kappa) \quad (13)$$

denote the accumulated number of overloaded link-cycles up to cycle k . Clearly, $M(K_s)$ then denotes the total number of overloaded link-cycles at the end ($K_s = 40$) of the simulation horizon. Fig. 5 displays the $M(k)$ quantities for each investigated signal control strategy for scenario 5. As expected, the ranking of the strategies with respect to this criterion is in agreement with the findings of Section 6.1. In particular, at the end of the simulation (cycle $K_s = 40$), the total numbers of overloaded link-cycles are 167 for FT-A, 147 for FT-B, 88 for LQ-A, 80 for LQ-B, 63 for QPC-A, and 51 for QPC-B.

Fig. 5 underlines the clear superiority of appropriately designed real-time signal control strategies over (even optimized) fixed-time control to handle urban network congestion. The figure also indicates the improvement of the new QPC strategy over the previous LQ approach and the value of demand information.

6.3. Fundamental diagrams

The fundamental diagram of the application road network under different signal control strategies may be used in order to extract additional useful insights on the related control impact and performance. Fig. 6 displays the fundamental diagrams resulting for scenario 5 when the six different signal control strategies are applied. Each measurement point in the diagrams corresponds to one 90 s-cycle.

As a first remark, Fig. 6 confirms the existence of a fundamental diagram for urban road networks, whose exact shape is seen to depend on the utilized signal control strategy. Remarkably, the diagrams indicate a hysteresis, i.e. a different path of measurement points (black circles) when filling the network than when emptying the network (where measurements are marked with stars). A hysteresis indicates that, for the same vehicle-number in the network, we may have different total flows when the network is filled than when it is emptied. This difference is attributed to partly different traffic patterns prevailing during the filling and emptying of the network and the absence of a traffic assignment module that would fill the network links more homogeneously in the simulation. Although a similar phenomenon might be observable in real traffic data as well, it is believed that the hysteresis is much more pronounced here due to the final phase of zero demands, the

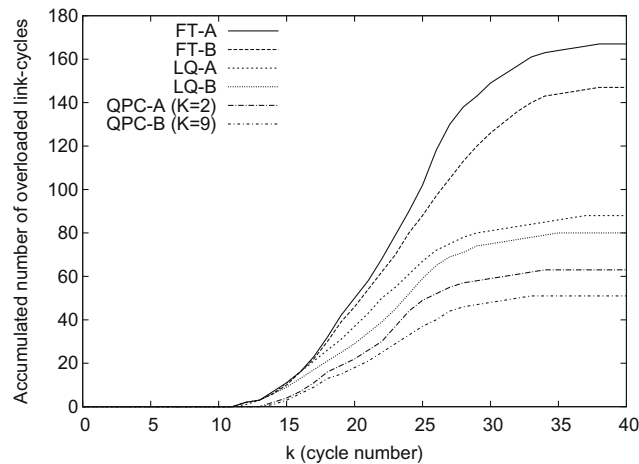


Fig. 5. Accumulated number of overloaded link-cycles for each signal control strategy.

complete emptying of the network at the end of the simulation and the mentioned absence of traffic assignment. For the following comments, we focus on the FD shape during the network filling phase only.

Regarding the FD region A (Fig. 1), which is seen to prevail in the application network for vehicle-numbers up to 200, the resulting slope (average speed) is highest for QPC-B; followed by LQ-B, QPC-A and FT-B; while LQ-A and FT-A are slightly worse since their signal plan is not fully adapted to the demand scenario.

For region B, we notice that the network flow capacity is lowest for FT-A (around 30,000 veh/h); it increases to some 32,000 veh/h for FT-B; it increases even more for LQ-A, LQ-B and QPC-A to reach some 34,000 veh/h for QPC-B. Note that, in absence of real-time cycle and offset control, this capacity increase is attributed mainly to less link-queue spillback, i.e. less wasting of green times due to overloaded downstream links.

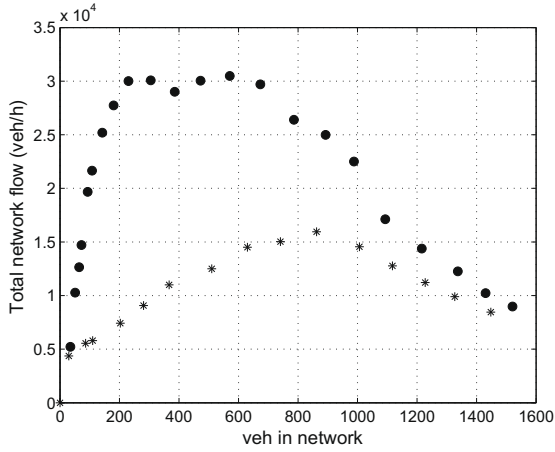
The difference in flow levels during the saturated traffic conditions of region B imply corresponding differences of the highest vehicle-number that is reached by each signal control strategy. Thus, FT-A is seen to reach up to 1500 veh in the network, with a fully formed region C of oversaturated traffic conditions and accordingly low flows (less than 10,000 veh/h) due to blocked links and partial gridlocks. FT-B reaches up to 1000 veh with total flows reducing to 17,000 veh/h, i.e. only a part of region C is actually visited; under LQ-A, the maximum vehicle-number does not exceed 900; while LQ-B and QPC-A reach 800 veh at the maximum and QPC-B even less. The total flow for all real-time control strategies is seen to only slightly reduce to 28,000 veh/h during the filling phase.

These results are quite conform with the performance criteria of Tables 1–3, along with providing more insights on the ways and reasons of improvements by each signal control strategy investigated. In particular, the FD results demonstrate that real-time control strategies designed to remain efficient under saturated traffic conditions may considerably improve the network performance and retard or avoid the detrimental effects of link queue spillback and gridlock.

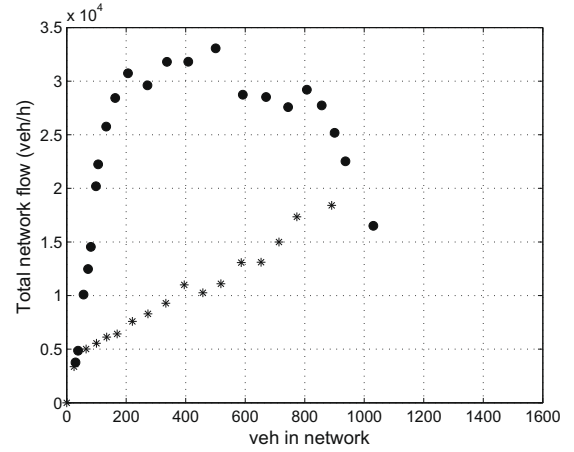
6.4. Sensitivity investigations and the impact of feedback

There is a fundamental difference in the use and significance of network traffic flow models in Transportation Planning (TP) as compared to traffic control (operations). In TP exercises, the typical goal is to predict (via a mathematical model) the potential behaviour of transportation infrastructures that are not existing so as to assess different planning alternatives, etc.; for this endeavour, the predicted infrastructure behaviour (and hence the utilized mathematical model) should be as accurate as possible. In the case of traffic control, the infrastructure under control is there and delivers (via appropriate sensors) valuable information on its current state; this information is used as a feedback loop (e.g. by the LQ regulator (8) or in the rolling-horizon frame of Section 4) in order to frequently update the control strategy decisions on the basis of the current real network traffic state. In other words, the impact of model inaccuracies (e.g. of inaccurate model parameter values) on the quality of the traffic control decisions is strongly reduced thanks to the real-time feedback. For example, a model may predict an increasingly wrong evolution into the future of the link queues if the utilized turning rates $t_{w,z}$ are different than the real ones (e.g. because turning rates may change in dependence of the control actions due to adapted driver routing); however, any past model-prediction inaccuracies are essentially nullified when new information (feedback) arrives that reflects the current real link queues as a basis for updated control decisions. Thus, some related concerns regarding the sensitivity of the control performance in case of moderate model-versus-reality inaccuracies (in particular of inaccurate turning rates), should be viewed under this perspective.¹ To demonstrate the low sensitivity of real-time control thanks to feedback, a specific investigation is proposed in this section.

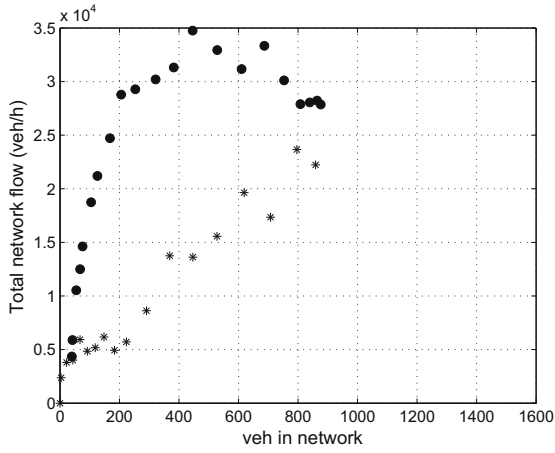
¹ The authors would like to thank an anonymous reviewer whose remarks triggered these additional sensitivity investigations.



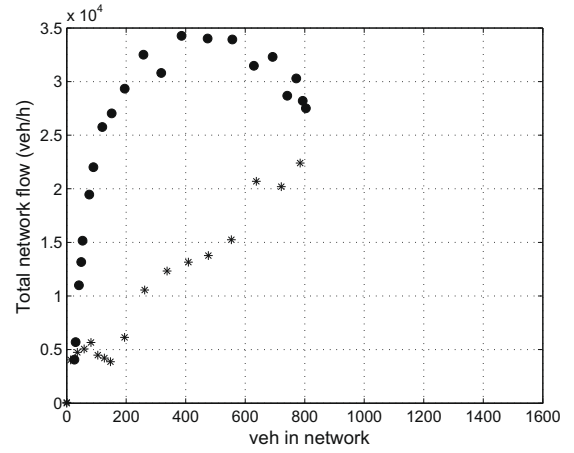
(a) FT-A



(b) FT-B



(c) LQ-A



(d) LQ-B

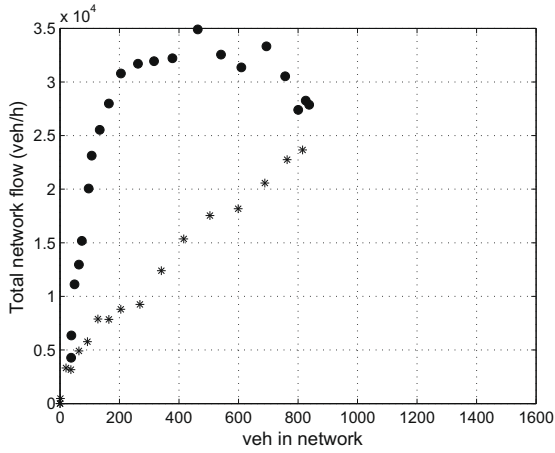
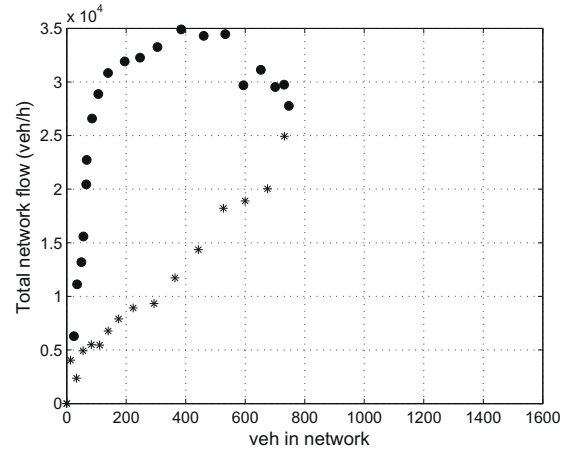
(e) QPC-A ($K = 2$)(f) QPC-B ($K = 9$)

Fig. 6. Fundamental diagrams for each signal control strategy.

The LQ-regulator 8 includes the gain matrix \mathbf{L} whose entries depend on matrices \mathbf{B} from (6) and \mathbf{Q} , \mathbf{R} from (7). The entries of matrix \mathbf{B} depend on the turning rates $t_{w,z}$. Since \mathbf{B} must be constant for this method, the turning rates must be pre-selected and may deviate from the real turning rates. If the real turning rates are strongly different on different times of the day

(e.g. for the p.m. peak period compared to the a.m. peak period), one may use accordingly different matrices \mathbf{L} ; but the following investigations demonstrate that this may only be worthwhile if the differences are really very strong.

The QPC approach employs the model (10) with the matrix \mathbf{B} also depending on the turning rates. But, because the model is used in real-time (according to Section 4), matrix \mathbf{B} may be time-varying and, moreover, may be updated from time-to-time based on real-time estimates of the turning rates. Again, the following investigations demonstrate that real-time estimates for $t_{w,z}$ may only be worthwhile if the related variations are really very strong.

The investigations in this section aim to test the sensitivity of the LQ and rolling-horizon QPC approaches to inaccuracies of the turning rates. To this end, the following cases are investigated via simulation for the five demand scenarios:

- Application of the LQ approach using the gain matrix \mathbf{L} obtained for the initial turning rates (LQ-A and LQ-B as described in Section 5.4).
- Application of the LQ approach using the gain matrix \mathbf{L}^* obtained for significantly different turning rates for all network junctions. These turning rates are modified randomly quite significantly, namely by some 40% with respect to initial turning rates (e.g. for a bifurcation with initial turning rates 0.5, 0.5 we may now have 0.3, 0.7 or 0.7, 0.3).
- Application of the rolling-horizon QPC approach using the initial turning rates in the linear model (10) (QPC-A and QPC-B as described in Section 5.4).
- Application of the rolling-horizon QPC* approach using significantly different turning rates as above for all network junctions in the linear model (10).

In all these cases, the simulation model (3) using (12) uses the initial turning rates. The QPC-A strategy is run with optimization horizon $K = 2$ as suggested in Section 6.1. The QPC-B strategy is run with two investigated optimization horizons, namely $K = 3$ and $K = 9$.

Table 5 summarizes the simulation results for the LQ-variants. The obtained results indicate that, even if the turning rates used for the calculation of the gain matrix are not accurately defined, the performance of the LQ-variants is not seriously affected. Similar results were reported based on extensive simulation investigations of the TUC strategy to inaccuracies of the traffic parameters, such as turning rates and saturation flows (Diakaki, 1999). Note that, in some scenarios, the inaccurate turning rates may even happen to lead to very slight improvements; this is explained by the fact that the LQ strategy is sub-optimal and hence its performance may even slightly improve in case of moderate inaccuracies.

Table 6 summarizes the obtained sensitivity results for the QPC-variants; it may be seen that the performance of the QPC-variants may be affected more strongly than the LQ-variants by the turning rates being used which is attributed to the explicit use of the model in the rolling-horizon scheme employed by the QPC approach. For QPC-B*, it can be seen that the use of smaller optimization horizon $K = 3$ (compared to $K = 9$) leads to lower sensitivity of both evaluation criteria in terms of

Table 5

Assessment criteria of sensitivity investigations of the LQ approach, using different gain matrices in (8) due to significantly different values of turning rates.

Scenario	Strategy							
	LQ-A		LQ-A*		LQ-B		LQ-B*	
	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB
1	13	561	14	554	14	702	14	700
2	44	7341	45	7720	41	7187	40	6829
3	164	53,993	170	56,442	146	48,282	152	51,561
4	285	100,136	295	103,942	250	89,946	253	90,420
5	348	130,891	339	122,754	311	120,760	315	119,270
Average	171	58,584	173	58,282	152	53,375	155	53,756

* The LQ approach uses the modified gain matrix \mathbf{L}^* .

Table 6

Assessment criteria of sensitivity investigations of the rolling-horizon QPC approach, using significantly different turning rates at all network junctions.

Scenario	Strategy									
	QPC-A ($K = 2$)		QPC-A* ($K = 2$)		QPB-B ($K = 3$)		QPB-B* ($K = 3$)		QPB-B ($K = 9$)	
	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB	TTS	RQB
1	13	578	13	686	11	280	12	423	11	280
2	36	5844	40	7099	30	4067	37	5767	30	4040
3	145	44,286	152	46,226	132	39,413	149	47,158	132	40,079
4	283	97,391	277	92,451	249	82,828	272	94,498	252	84,246
5	318	113,325	305	103,113	280	96,292	261	84,659	270	89,694
Average	159	52,285	157	49,915	140	44,576	146	46,501	139	43,668

* The QPC approach uses the modified turning rates.

Table 7

Comparison of assessment criteria of sensitivity investigations of the LQ and rolling-horizon QPC approaches.

Scenario	Strategy									
	LQ-A vs. LQ-A		LQ-B* vs. LQ-B		QPC-A* vs. QPC-A ($K = 2$)		QPC-B* vs. QPC-B ($K = 3$)		QPC-B* vs. QPC-B ($K = 9$)	
	TTS (%)	RQB (%)	TTS (%)	RQB (%)	TTS (%)	RQB (%)	TTS (%)	RQB (%)	TTS (%)	RQB (%)
1	8	−1	0	0	0	19	9	51	9	51
2	2	5	−2	−5	11	21	23	42	23	42
3	4	5	4	7	5	4	13	20	13	18
4	4	4	1	1	−2	−5	9	14	8	11
5	−3	−6	1	−1	−4	−9	−7	−1	7	12
Average	1	−1	2	1	−1	−5	4	4	9	13

* The LQ approach uses the modified gain matrix L^* ; the rolling-horizon QPC approach uses the modified turning rates.

average values because longer-term model predictions ($K = 9$) tend to be increasingly less accurate. Thus, the optimization horizon K should be selected so as to strike a balance between performance sensitivity with respect to inaccuracies of the traffic parameters versus myopic control actions. Comparing with Table 5, QPC-A* and QPC-B* are seen to be still better than LQ-A and LQ-B, respectively, and, of course, much better than fixed-time control, even though FT-B was obtained on the basis of accurate turning rates.

Table 7 displays some percentage changes of both evaluation criteria for the four cases. LQ-A* vs. LQ-A and LQ-B* vs. LQ-B demonstrate the low sensitivity of the LQ approach with respect to inaccuracies of the turning rates. QPC-A* vs. QPC-A indicates the low sensitivity of the rolling-horizon QPC-A approach for $K = 2$. QPC-B* vs. QPC-B demonstrates the moderate sensitivity of the QPC-B approach particularly for increasing optimization horizon K .

In general, the LQ approach seems less sensitive than the rolling-horizon QPC approach with respect to inaccuracies of traffic parameters; this might be attributed to the analytical feedback law derivation of (8) as opposed to the implicit feedback of the rolling-horizon approach (Dreyfus, 1964).

7. Conclusions

The paper investigated a computationally feasible quadratic-programming control (QPC) methodology for real-time network-wide signal control in large-scale urban networks that is appropriate also for congested traffic conditions. This methodology combines the traffic flow store-and-forward modeling paradigm with quadratic-programming optimization embedded in a rolling-horizon control scheme. A simulation-based investigation of the signal control problem for a realistic example aimed at demonstrating the strategy efficiency and feasibility, as well as its comparison with optimized fixed-time (FT) settings and the LQ approach taken by the signal control strategy TUC.

Two evaluation criteria have been used for strategy comparison via simulation. It was demonstrated that the LQ variants lead to significant reductions of both evaluation criteria compared to optimized fixed-time plans, which underlines the superiority of appropriate real-time decision-making even in case of optimized fixed control. Furthermore, it was shown that the LQ approach has a relatively low sensitivity to non-optimal nominal plans utilized within the strategy. The QPC approach was shown to outperform the FT and LQ approaches, particularly under full demand provision. Furthermore the impact of all strategies with regard to the network flow capacity and the retarding or avoiding of link queue spillback and gridlock was investigated, besides the traditional evaluation ways, also by use of the currently proposed fundamental diagram for urban road networks. Finally, the real-time control strategies were demonstrated to have low (LQ) or moderate (QPC) sensitivity with respect to model parameter (turning rate) inaccuracies thanks to their feedback character.

Future work will deal with the comparison of the proposed QPC rolling-horizon approach with other strategies (e.g. TUC) in more elaborated (e.g. microscopic) simulation as well as in real-life conditions.

Acknowledgements

This paper is part of the 03ED898 research project, implemented within the framework of the “Reinforcement Programme of Human Research Manpower” (PENED) and co-financed by National and Community Funds (75% from EU – European Social Fund and 25% from the Greek Ministry of Development-General Secretariat of Research and Technology).

References

- Aboudolas, K., Papageorgiou, M., Kosmatopoulos, E., 2009. Store-and-forward based methods for the signal control problem in large-scale congested urban road networks. *Transportation Research* 17C (2), 163–174.
- Abu-Lebdeh, G., Benekohal, R.F., 1997. Development of traffic control and queue management procedures for oversaturated arterials. *Transportation Research Record* (1603), 119–127.
- Beard, C., Ziliaskopoulos, A., 2006. A system optimal signal optimization formulation. In: *Proc. of the 85th TRB Annual Meeting*, Washington, DC, USA.
- Bretherton, D., Bodger, M., Baber, N., 2004. SCOOT – the future. In: *Proc. of the 12th IEE International Conference on Road Transport Information and Control*, London, UK, pp. 301–306.

- Cremer, M., 1991. Origin–destination matrix: dynamic estimation. In: Papageorgiou, M. (Ed.), *Concise Encyclopedia of Traffic and Transportation Systems*. Pergamon Press, Oxford, pp. 310–315.
- Daganzo, C.F., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research* 42B (9), 771–781.
- De Schutter, B., De Moor, B., 1998. Optimal traffic light control for a single intersection. *European Journal of Control* 4 (3), 260–276.
- Diakaki, C., 1999. Integrated control of traffic flow in corridor road networks. Ph.D. Thesis, Technical University of Crete, Chania, Greece.
- Diakaki, C., Papageorgiou, M., Aboudolas, K., 2002. A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engineering Practice* 10, 183–195.
- Diakaki, C., Dinopoulou, V., Aboudolas, K., Papageorgiou, M., Ben-Shabat, E., Seider, E., Leibov, A., 2003. Extensions and new applications of the traffic-responsive urban control strategy: coordinated signal control for urban networks. *Transportation Research* (1856), 202–211.
- Dinopoulou, V., Diakaki, C., Papageorgiou, M., 2005. Application and evaluation of the signal traffic control strategy TUC in Chania. *Journal of Intelligent Transportation Systems* 9 (3), 133–143.
- Dreyfus, S.E., 1964. Some types of optimal control of stochastic systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control* 2 (1), 120–134.
- Farges, J.L., Henry, J.J., Tufal, J., 1983. The PROLYN real-time traffic algorithm. In: *Proc. of the 4th IFAC Symposium on Transportation Systems*. Baden–Baden, Germany, pp. 307–312.
- Farhi, N., 2008. Modélisation minplus et commande du trafic de villes régulières. Ph.D. Thesis, Université de Paris I-Panthéon-Sorbonne, Paris, France.
- Gartner, N.H., 1983. OPAC: a demand-responsive strategy for traffic signal control. *Transportation Research Record* (906), 75–84.
- Gartner, N.H., Wagner, P., 2004. Analysis of traffic flow characteristics on signalized arterials. *Transportation Research Record* (1883), 94–100.
- Gazis, D.C., Potts, R.B., 1963. The oversaturated intersection. In: *Proc. of the 2nd International Symposium on Traffic Theory*, London, UK, pp. 221–237.
- Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. *Transportation Research* 42B (9), 759–770.
- Godfrey, J.W., 1969. The mechanism of a road network. *Traffic Engineering and Control* 11 (7), 323–327.
- Hunt, P.B., Robertson, D.I., Bretherton, R.D., Royle, M.C., 1982. The SCOOT on-line traffic signal optimization technique. *Traffic Engineering and Control* 23, 190–192.
- Kosmatopoulos, E., Papageorgiou, M., Bielefeldt, C., Dinopoulou, V., Morris, R., Mueck, J., Richards, A., Weichenmeier, F., 2006. International comparative field evaluation of a traffic-responsive signal control strategy in three cities. *Transportation Research* 40A (5), 399–413.
- Lo, H.K., 1999. A novel traffic signal control formulation. *Transportation Research* 33A (6), 433–448.
- Lo, H.K., Chang, E., Chan, Y.C., 2001. Dynamic network traffic control. *Transportation Research* 35A (8), 721–744.
- Lowrie, P.R., 1982. SCATS: the Sydney co-ordinated adaptive traffic system—principles, methodology, algorithms. In: *Proc. of the IEE International Conference on Road Traffic Signalling*, London, England, pp. 67–70.
- Mirchandani, P., Head, L., 1998. RHODES—a real-time traffic signal control system: architecture, algorithms, and analysis. In: *TRISTAN III (Triennial Symposium on Transportation Analysis)*, San Juan, Puerto Rico, vol. 2.
- Mirchandani, P., Wang, F.-Y., 2005. RHODES to intelligent transportation systems. *IEEE Intelligent Systems* 20 (1), 10–15.
- Papageorgiou, M., 1995. An integrated control approach for traffic corridors. *Transportation Research* 3C (1), 19–30.
- Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y., 2003. Review of road traffic control strategies. *Proceedings of the IEEE* 91 (12), 2043–2067.
- Papageorgiou, M., Ben-Akiva, M., Bottom, J., Bovy, P., Hoogendoorn, S., Hounsell, N., Kotsialos, A., McDonald, M., 2007. ITS and traffic management. In: *Barnhart, C., Laporte, G. (Eds.), Handbook in Operations Research & Management Science: Transportation*, vol. 14. Elsevier, North Holland, pp. 715–774.
- Papageorgiou, M., Kosmatopoulos, E., Papamichail, I., 2008. Effects of variable speed limits on motorway traffic flow. *Transportation Research Record* (2047), 37–48.
- Singh, M.G., Tamura, H., 1974. Modelling and hierarchical optimization of oversaturated urban traffic networks. *International Journal of Control* 20 (6), 913–934.