Delft University of Technology

Delft Center for Systems and Control

Technical report 13-033

Optimal dynamic route guidance: A model predictive approach using the macroscopic fundamental diagram*

M. Hajiahmadi, V.L. Knoop, B. De Schutter, and H. Hellendoorn

If you want to cite this report, please use the following reference instead:
M. Hajiahmadi, V.L. Knoop, B. De Schutter, and H. Hellendoorn, "Optimal dynamic route guidance: A model predictive approach using the macroscopic fundamental diagram," *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, The Hague, The Netherlands, pp. 1022–1028, Oct. 2013.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.51.19 (secretary)
fax: +31-15-278.66.79
URL: http://www.dcsc.tudelft.nl

Optimal Dynamic Route Guidance: A Model Predictive Approach Using the Macroscopic Fundamental Diagram

Mohammad Hajiahmadi, Victor L. Knoop, Bart De Schutter, and Hans Hellendoorn

Abstract— **Since centralized control of urban networks with detailed modeling approaches is computationally complex and inefficient, hierarchical decentralized methods based on aggregate models are of great importance. In this paper, we use an aggregate modeling approach based on the macroscopic fundamental diagram (MFD), in order to find dynamic optimal routing strategies. An urban area can be divided into homogeneous regions each modeled by a (set of) macroscopic fundamental diagrams. Thus, the problem of route guidance can be solved in a regional fashion by using model predictive control and the novel high-level MFD-based model used for prediction of traffic states in the urban network. The optimal routing advices obtained from the high-level controller can be used as references (to track) for lower-level local controllers installed at the borders of the regions. Hence, the complexity of solving the routing problem will be decreased significantly. The performance of the proposed approach is evaluated using a multi-origin multi-destination grid network. Further, the obtained results show significant performance of the optimal dynamic route guidance over other static routing methods.**

I. INTRODUCTION

The Macroscopic Fundamental Diagram (MFD) (or network fundamental diagram) has been first introduced in [1] and it has been investigated recently in [2]–[4]. Basically, the MFD relates the accumulations, i.e. the number of vehicles, in a network to the so-called production, defined as the average flow in the network. For a network in which the congestion is homogeneously distributed, a low-scatter and well-defined MFD can be extracted [5]. The literature has studied the properties of the MFD [6], and the ways to split a network into homogeneous subnetworks [7].

The MFD has provided the possibility for macroscopic modeling and control of large networks. This is not possible with sufficiently detailed modeling of individual elements of a large-scale network, since centralized control based on such a detailed model is not computationally efficient. Therefore, modeling and control using the MFD have attracted attention of researchers in recent years. Among them is the perimeter control, i.e. limiting the inflow of a network to ensure a high production [2], [8]–[10]. Further, the MFD can be used to design advanced control schemes in order to decrease delays and to increase accessibility in large urban networks. Nevertheless, it should be noted that changes in the network topology, the signal timing plans of the signalized intersections, and the infrastructure characteristics affect the shape of the MFD. In [11] this issue has been addressed and the authors have introduced another level of control inside urban networks with taking into account the variability of the MFD.

As mentioned before, the MFD can be exploited for high-level and aggregated traffic state modeling. Moreover, with MFD-based modeling, the number of parameters that need to be calibrated is significantly lower. Also, the computational complexity is much lower in the high-level modeling using the MFD. Therefore, in this paper we aim at using the MFD-based modeling approach for solving another challenging problem in the traffic networks, viz., dynamic route guidance.

Dynamic route assignment has been an interesting topic for researchers [12]–[16]. The main concept of dynamic routing is to guide the traffic toward alternative routes in the network in order to reduce the imbalance in the distribution of traffic flows, to improve the overall travel time and/or to minimize other traffic objectives such as emissions or total fuel consumption. In this paper, we address the route guidance problem using a high-level scheme. The aim is to use the aggregate modeling approach based on the MFD for describing the flow of vehicles traveling in a multi-region urban network. The high-level MFD-based traffic flow model is then utilized in an optimal dynamic route guidance framework. The framework is developed based on the theory of model predictive control (MPC) [17], [18] and its main goal is to determine optimal references for guiding the traffic of vehicles between urban regions in order to achieve minimum delays in reaching the destinations. Basically, the proposed route guidance scheme consists of two levels. At the higher level, a central controller uses the MFD-based traffic flow model in order to find optimal splitting rates for traffic flows heading specific destinations. A major advantage of this approach is that the necessity of having pre-defined routes in the network and searching for the optimal ones is relaxed by finding the destination-dependent splitting fractions towards the neighboring regions of a region. Hence, we lifted from the link-level splitting rates to region-level splitting of traffic flows. Another main advantage of this scheme is that the computational complexity is much less than the usual route assignment problems that deal with a huge collection of roads and intersections.

The obtained optimal splitting rates will be communicated to the lower level controllers that are installed at the borders of urban regions. These local controllers have the task to realize the reference splitting rates by manipulating the

This work is supported by the European 7th Framework Network of Excellence HYCON2 and the European COST Action TU1102.

M. Hajiahmadi, B. De Schutter and Hans Hellendoorn are with the Delft Center for Systems and Control (DCSC), Delft University of Technology, Delft, The Netherlands. m.hajiahmadi@tudelft.nl

V. Knoop is with the Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands. v.l.knoop@tudelft.nl

signaled intersections and/or by adopting the dynamic route guidance information panels.

The rest of the paper is organized as follows. In Section II, the high-level modeling of multi-region urban networks is presented based on the concept of MFD and macroscopic traffic flow modeling. Next, in Section III, a multi-level scheme for optimal dynamic route guidance is introduced and the optimization problem for the high-level is formulated. Section IV presents the set-up of a case study in order to illustrate the proposed route guidance approach and to evaluate and compare the performance of the proposed scheme with other simple routing methods. The paper ends with conclusions and ideas for future research.

II. MULTI-REGION MACROSCOPIC MODELING

The modeling method starts by splitting the network into several regions, which are as homogeneous as possible in the sense of congestion distribution. For heterogeneous networks, it might be possible to partition them into more homogeneous regions such that each region has a well-defined MFD, see [7]. The dynamics of traffic are modeled in these regions, using the extracted MFD for each region (as depicted in Fig. 1).

In each region $i \in \mathcal{R}$, with $\mathcal{R}$ the set of all regions, the accumulation is defined as the weighted density of all links in region i and is formulated as follows:

$$n_i(k) = \frac{\sum\limits_{\lambda \in \Lambda_i} \big(\kappa_\lambda \cdot L_\lambda \cdot \rho_\lambda(k)\big)}{\sum\limits_{\lambda \in \Lambda_i} \big(\kappa_\lambda \cdot L_\lambda\big)}, \qquad (1)$$

where Λ_i contains all links in region i and κ_λ, L_λ, and ρ_λ are the number of lanes, the length, and the density of link λ, respectively.

The set of neighboring regions of region i is defined as $\mathcal{J}_i$. The flow from region i to region $j \in \mathcal{J}_i$ is determined by the minimum of three elements:

1) The capacity of the boundary between region i and region j, $C_{i,j}$.
2) The demand from region i to region j, $D_{i,j}$.
3) The supply in region j, S_j.

The demand from region i to region j is determined based on the MFD, a function we indicate as $P_i(n_i)$. In fact, we can construct a demand and supply scheme similar to the cell transmission model [19]. The supply can be determined in the same way as in the cell transmission model; the supply is equal to the critical production $P_{j,\text{crit}}$ for accumulations lower than the critical accumulation n_{crit} and is equal to the MFD for higher accumulations:

$$S_j(k) = \begin{cases} P_{j,\text{crit}} & \text{if } n_j(k) \le n_{j,\text{crit}} \\ P_j(n_j(k)) & \text{if } n_j(k) > n_{j,\text{crit}} \end{cases} \qquad (2)$$

Contrary to the cell transmission model, the demand in a region decreases with the accumulation exceeding the critical accumulation. This is because there might be internal traffic jams in the region, limiting the potential outflow. This fact is shown graphically in Fig. 2.

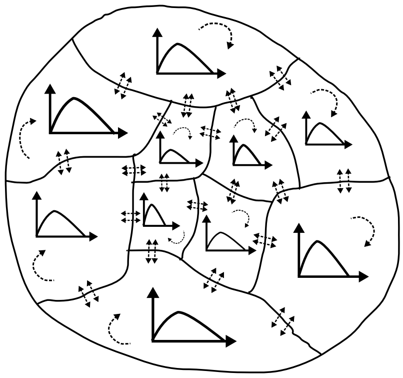


Fig. 1. Schematic multi-region urban network.

The part of accumulations in each region i heading towards destination $d \in \mathcal{D}$ is known and is denoted by $n_{i,d}$. Moreover, the routing from region i to a destination d is coded by the next neighboring region j in the so called destination-specific splitting rates $\alpha_{i,j,d}$. Therefore, the total demand from region i towards region j is formulated as:

$$D_{i,j}(k) = \sum_{d \in \mathcal{D}} \left(\alpha_{i,j,d}(k) \cdot \frac{n_{i,d}(k)}{n_i(k)} \cdot P_i\big(n_i(k)\big) \right), \qquad (3)$$

where $\mathcal{D}$ is the set of all destinations. This demand is limited by the capacity of the boundary of regions i and j, giving the effective demand $\tilde{D}_{i,j}$ as:

$$\tilde{D}_{i,j}(k) = \min\{D_{i,j}(k), C_{i,j}\} \qquad (4)$$

The fraction of traffic allowed over the boundary between i and j is indicated by $\frac{\tilde{D}_{i,j}(k)}{D_{i,j}(k)}$. As an intermediate step, we now have the demand from region i to destination d via region j:

$$\tilde{D}_{i,j,d}(k) = \alpha_{i,j,d}(k) \cdot \frac{n_{i,d}(k)}{n_i(k)} \cdot P_i\big(n_i(k)\big) \cdot \frac{\tilde{D}_{i,j}(k)}{D_{i,j}(k)} \qquad (5)$$

The total demand towards region j is determined by adding up all effective demands towards region j

$$D_j(k) = \sum_{i \in \mathcal{J}_j} \tilde{D}_{i,j}(k) \qquad (6)$$

This value is compared with the supply in region j. If the supply is larger, the flow is unrestricted. However, if the supply is lower, the fraction of the flow that can travel into region j is determined as:

$$\psi_j(k) = \min\left\{ \frac{S_j(k)}{D_j(k)}, 1 \right\} \qquad (7)$$

If the supply restricts the flow, the actual flow to cell j is proportional to the demands towards the cell. Now, the flow is set as the minimum of demand and supply. This flow is assumed to be constant between two consecutive time steps.

Now, for all regions $j \in \mathcal{J}_i$ that have a demand $\tilde{D}_{i,j}(k) > 0$, the minimum of the outflow fractions calculated in (7) is determined:

$$\phi_i(k) = \min_{j \in \mathcal{J}_i} \{\psi_j(k)\} \qquad (8)$$

Since traffic cannot go independently to any destination (congestion and blocking back will occur within region i),

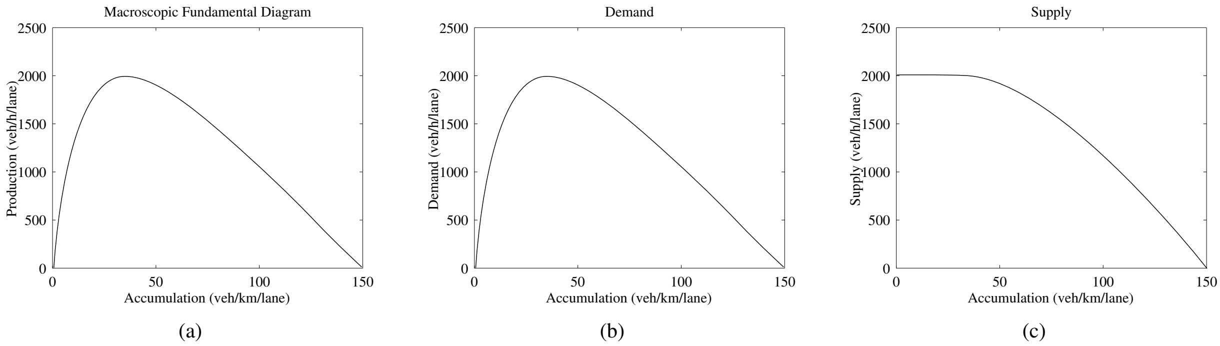


Fig. 2. The factors determining the flow, (a) Macroscopic Fundamental Diagram, (b) demand, and (c) supply.

for all the demands from region i to *every* neighboring region j, the same fraction (8) will be applied. Hence, the outflow from region i to region $j \in \mathcal{J}_i$ is formulated as:

$$q_{i,j}(k) = \phi_i(k) \cdot \tilde{D}_{i,j}(k) \tag{9}$$

The flow can be separated per destination. So, similar to reducing the overall flow (9), we can modulate the flow per destination (5) as:

$$q_{i,j,d}(k) = \phi_i(k) \cdot \tilde{D}_{i,j,d}(k) \tag{10}$$

Therefore, the accumulation in any region i towards destination d can now be updated as follows:

$$n_{i,d}(k+1) = n_{i,d}(k) + \frac{T_{\mathrm{s}}}{\sum\limits_{\lambda \in \Lambda_i} \kappa_\lambda L_\lambda} \Big(\sum_{j \in \mathcal{J}_i} q_{j,i,d}(k) - \sum_{j \in \mathcal{J}_i} q_{i,j,d}(k) \Big), \tag{11}$$

with T_{s} the sample time. Hence the total accumulation in region i will be:

$$n_i(k+1) = \sum_{d \in \mathcal{D}} n_{i,d}(k+1) \tag{12}$$

In the next section, we use the presented model for prediction of accumulations in the network in order to determine optimal routes.

III. High-level optimal route guidance

In this section, we develop a route guidance scheme based on the high-level MFD-based model derived in the previous section. In the proposed framework, we solve the dynamic routing problem on a macroscopic level. This means that instead of taking into account individual roads and intersections, we deal with regional destinations and the way that traffic flow should be split towards the neighboring regions in order to avoid congestion in the intermediate regions, to decrease the overall travel time and consequently, to improve the arrival rates at the destinations. We assume a two-level structure as depicted in Fig. 3. At the top level, the optimal route guidance problem is solved based on the aggregate model presented in the previous section. At the lower level, the optimal variables (the splitting rates) that are obtained from the high-level optimization problem are taken as references, i.e. local controllers in the lower level aim at realizing the optimal splitting rates for (destination dependent) flows of vehicles that want to travel across the regions. In the following, we elaborate on the type of optimization problem that has to be solved in the highest level in order to achieve the aforementioned goals.

A. Objective function

In order to formulate the routing problem, an objective needs to be defined. The major aim in an urban network could be maximizing the arrival rate, i.e. the number of vehicles that complete their trips and reach their destinations, or similarly minimizing the total travel delays. Over the (discrete) simulation interval $[0, \cdots, K-1]$, the total delay criterion J_{TD} (veh·s) is formulated as:

$$J_{\mathrm{TD}} = T_{\mathrm{s}} \cdot \sum_{i \in \mathcal{R}} \sum_{k=0}^{K-1} \bigg(\big(\sum_{\lambda \in \Lambda_i} \kappa_\lambda L_\lambda \big) \cdot n_i(k) \bigg). \tag{13}$$

Moreover, one can introduce a penalty term on the differences between average speeds of all regions as follows:

$$J_{\mathrm{v}}(K-1) = \sum_{i,j \in \mathcal{R}} \big(\bar{V}_i(K-1) - \bar{V}_j(K-1) \big)^2, \tag{14}$$

with $\bar{V}_i(K-1)$ the average speed in region i determined from the MFD of that region at the end of simulation period (note that one can calculate the differences between speeds of all regions for all time steps, but it may increase the computation time of the corresponding optimization problem. Therefore, we try to normalize the speeds only at the end of the time horizon). Basically, with the values of the accumulations in each region, one can estimate an average speed for that region. Assuming an exponential function for the MFD, the average speed can be determined as follows:

$$\bar{V}_i(k) = V_{\mathrm{free},i} \cdot \exp\Big(-\frac{1}{2} \Big(\frac{n_i(k)}{n_{\mathrm{crit},i}} \Big)^2 \Big), \tag{15}$$

with $V_{\mathrm{free},i}$ the free-flow speed and $n_{\mathrm{crit},i}$ the critical accumulation corresponding to the maximum production. Essentially when there is no congestion, the average speeds in the regions are high. But in case of congestion in a region, the average speed will decrease and consequently, the travel time for vehicles inside that region will increase. By minimizing (13), the overall travel delay in the network will decrease, but it might be possible that the traffic is not distributed evenly and in some regions the average speed will be high while in others we observe low speeds. The objective function (14)

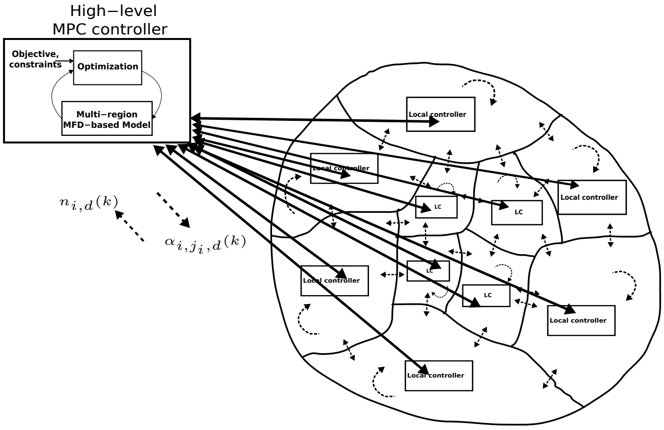


Fig. 3. High-level optimal dynamic route guidance.

tries to avoid the congestion to build up in some regions until those regions are blocked while other regions receive very little flow. On the other hand, minimizing only (13) might lead to keeping the flows in some regions and this will again reduce the speed in those regions. Therefore, minimizing the objective (14) will help to normalize the average speed and travel times of all regions.

The total objective function can be defined as a weighted sum of the criteria (13) and (14):

$$J_{\text{total}} = \frac{J_{\text{TD}}}{J_{\text{TD,nom}}} + \omega \cdot \frac{J_{\text{v}}(K-1)}{J_{\text{v,nom}}}, \tag{16}$$

where $J_{\text{TD,nom}}$ and $J_{\text{v,nom}}$ denotes the nominal values of the objective functions that can be obtained using (uncontrolled) simulation. Moreover, the weight ω enables us to devote more weight to either one of the objectives. Also note that one can prioritize access to some regions or it might be the case that reducing the accumulations of certain regions is more important and therefore additional weights can be assigned to the accumulations of those regions in (13).

B. Model predictive control for high-level route guidance

Model Predictive Control (MPC) [17], [18] is an advanced control method originally developed for industrial processes and now for broader applications such as traffic networks. In the traffic control framework, the main idea is to use a prediction model of the network (e.g. the aggregate model derived in Section II) and an objective function assessing the desired performance of the urban traffic network, in order to find the optimal inputs through an optimization algorithm. In our case, the optimization variables are the optimal splitting rates for flows of vehicles heading multiple destinations. The overall optimization variables include splitting rates of all destination-dependent flows in all regions of the network. The optimization algorithm assumes a prediction horizon N_{p} for evolution of the network variables and minimizes the objective function over the horizon. The obtained optimal variables constitute a sequence of optimal splitting rates for the whole prediction horizon. In the MPC context, only the first sample of the obtained values is used and the prediction horizon is shifted one step forward, and the prediction and optimization procedure over the shifted horizon are repeated using new observations from the network. Moreover, to reduce the number of optimization variables, usually a control horizon $N_{\text{c}} < N_{\text{p}}$ is introduced and from the control step $k_{\text{c}} + N_{\text{c}} - 1$ onward, the control inputs (splitting rates) are taken to be constant.

Furthermore, the optimal routes or the corresponding splitting rates are communicated to the lower level local controllers as references to track. Basically, the local controllers try to achieve the optimal splitting rates by manipulating the timing plans of the signalized intersections placed at the borders of regions. Communication and coordination between the local controllers placed on different borders of a region is crucial. Note that control is carried out only at the borders and thus the MFDs of regions are expected to be unchanged. Nevertheless, we can take advantage of the approach proposed in [11] in order to extend the control to inside regions and hence to distribute the congestion in a more uniform way. This can be done by defining several timing plans for intersections inside each region and hence having a set of MFDs obtained for that region. By proper switching between the pre-defined timing plans, we will be able to normalize the congestion inside regions in addition to the determination of splitting rates for flows traveling to neighboring regions.

In order to formulate the problem of finding optimal splitting rates in the MPC framework, we define k_{c} and T_{c} as the control time step and the control sample time. Here we assume that the control sample time is an integer multiple of the simulation sample time, i.e. $T_{\text{c}} = M \cdot T_{\text{s}}$. If we define $J_{\text{TD}}^{\text{MPC}}$ as

$$J_{\text{TD}}^{\text{MPC}} = T_{\text{s}} \cdot \sum_{i=1}^{R} \sum_{k=M\cdot k_{\text{c}}}^{M\cdot(k_{\text{c}}+N_{\text{p}})-1} \Big(\big(\sum_{\lambda\in\Lambda_i} \kappa_\lambda L_\lambda \big) \cdot n_i(k) \Big), \tag{17}$$

the overall optimization problem will be formulated as follows:

$$\min J_{\text{total}}(k_{\text{c}}) = \min_{\tilde{\alpha}_{i,j,d}(k_{\text{c}})} \frac{J_{\text{TD}}^{\text{MPC}}}{J_{\text{TD,nom}}^{\text{MPC}}} + \omega \cdot \frac{J_{\text{v}}(M\cdot(k_{\text{c}}+N_{\text{p}})-1)}{J_{\text{v,nom}}} \tag{18}$$

subject to:

$$\text{model equations (11), (12)},$$

$$0 \le \alpha_{i,j,d}(k) \le 1, \tag{19}$$

$$\sum_{j\in\mathcal{J}_i} \alpha_{i,j,d}(k) \le 1, \quad \forall i\in\mathcal{R}, \quad \forall d\in\mathcal{D}, \tag{20}$$

$$\alpha_{i,j,d}(k) = \alpha_{i,j,d}^{\text{c}}(k_{\text{c}}), \quad \text{if } k\in\{M\cdot k_{\text{c}},\ldots,M\cdot(k_{\text{c}}+1)-1\}, \tag{21}$$

for all $i\in\mathcal{R}, j\in\mathcal{J}_i$, and $d\in\mathcal{D}$. The optimization variables defined over the prediction horizon N_{p} are $\tilde{\alpha}_{i,j,d}(k_{\text{c}}) = [\alpha_{i,j,d}^{\text{c}}(k_{\text{c}}),\ldots,\alpha_{i,j,d}^{\text{c}}(k_{\text{c}}+N_{\text{p}}-1)]^{\text{T}}$, where $\alpha_{i,j,d}^{\text{c}}(k_{\text{c}}+l)$ for $l = 0,\ldots,N_{\text{p}}-1$, is the splitting rate corresponding to the fraction of the flow towards destination d that travels from region i to region j at control time step $k_{\text{c}}+l$. Furthermore, by defining the inequality (20), we let a fraction of vehicles travel inside of regions until they get the opportunity to pass the borders. This will help to prevent the congestion from accumulating in the successor regions. Note that the optimization algorithm will determine whether the sum of the splitting rates for the flows heading towards a certain destination should be equal to one or otherwise less than one.

The optimization problem (18)–(21) is a nonlinear optimization problem that can be solved using either global optimization algorithms or multi-start local optimization methods.

In the next section, the proposed optimal route guidance approach is implemented on an urban network case study.

IV. Case Study

This section describes modeling and optimal routing of an urban network case study. The aim is to show the performance of the proposed high-level modeling and optimal dynamic route guidance approach. In the first part, the set-up of the case study is described and in the second part, the obtained results together with the discussions are presented.

A. Set-up

In order to implement the model presented in Section II, we consider a grid network. The network is a 4x4 regional network, with regions of 5x5 km, as shown in Fig. 4. The regions are homogeneous, with a critical accumulation $n_{\rm crit}$=25 veh/km and 10 km of road length in the region. The free flow speed is assumed to be $V_{\rm free}$=100 km/h. The capacity of the borders is set to 2000 veh/h/km. For each region, an MFD is assumed and it is approximated with an exponential function as follows;

$$P_i = n_i \cdot V_{\rm free} \cdot \exp\Big(-\frac{1}{2}\Big(\frac{n_i}{n_{\rm crit}}\Big)^2\Big) \qquad (22)$$

For each region i, the neighboring regions are defined as the ones that are in the horizontal or vertical directions with respect to the location of the region i. For instance, for region 7 the set of neighboring regions is $\{3, 6, 8, 11\}$. As illustrated in Fig. 4, the origins are indicated by blue squares and the destination are marked as red circles. The demand (veh/h) for each origin-destination pair is selected as in Table I.

TABLE I

Origin-destination demands (veh/h)

	Region 2	Region 8	Region 9	Region 14
Region 1	1000	1800	1750	3000
Region 4	1900	1400	1000	1400
Region 11	1700	1200	1300	1300
Region 16	2000	1000	1000	1800

The demand values are multiplied by time-varying factors in order to consider the uncertainty in the demand profiles and also to make it more realistic. At each time step, the demand values in Table I are multiplied by a uniformly distributed random number with mean value 1 and variance 0.1.

The optimal route assignment is carried out first by a static shortest-path algorithm and next by the model predictive

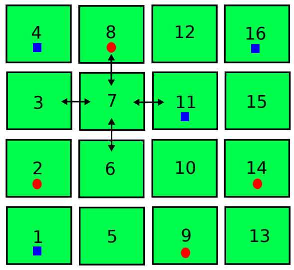


Fig. 4. Layout of the 4x4 urban network.

scheme described in Section III. The shortest-path algorithm determines the shortest routes (in time) based on the average speeds of all regions. Traffic will avoid the regions with low speed. First the costs of traveling between neighboring regions are obtained based on the current state of the network. Next the shortest path between each pair of regions is calculated using the Floyd-Warshall algorithm [20].

For the MPC scheme, we choose the prediction and control horizons $N_{\rm p} = 6$ and $N_{\rm c} = 2$, respectively. Using simulations for different horizons, these values have proved to be sufficient for our case (in general, the prediction interval should be long enough to include important dynamics of the system under control) while the computation time of the optimization algorithm is acceptable. It has been observed that with increasing prediction and/or control horizons, there are small improvements in the results while the computational complexity will grow exponentially.

The simulation sample time is chosen as $T_{\rm s} = 10$ s. The control sample time (for both static and dynamic schemes) is selected to be $T_{\rm c} = 60$ s. Between the two consecutive control calls, the control inputs (optimal splitting rates) are assumed to be constant. As mentioned in Section III, instead of searching for optimal routes from a set of pre-defined routes, we find the splitting rates towards the neighboring regions.

Furthermore, in order to take into account the uncertainty in modeling, the updated accumulations in the network model (simulation model) are corrupted with additive white Gaussian noises that have zero mean value and 2% of the measured accumulations as variance. Note that the prediction model in the MPC framework is assumed to be free of noise, but it is supplied with actual accumulations as initial values for the prediction model.

The nonlinear optimization inside the MPC scheme is solved using the *snopt* algorithm integrated in the *Tomlab* toolbox of MATLAB. This optimization algorithm tries to find the (global) optimal value for the objective function (16) subject to the linear constraints on the splitting rates. In order to escape from the local optima, we use a multi-start technique with random initialization.

B. Results

Results for simulation of the urban network for a period of 3000 s are depicted in Fig. 5. In the first column, the time evolution of the fixed-routing case is presented. By fixed-routing we mean that the routes are determined using the shortest-path algorithm and are fixed during the simulation period. As time progresses, congestion builds up in the regions that are located in the center. This is due to the fact that the center regions are the intermediate regions for many routes between the introduced origin and destination regions, and if no routing policy is considered, the accumulation would grow especially in these regions till it reaches the critical point. From then, the inflow to these regions is constrained and instead the congestion forms upstream of these regions. The total delay in the network for the whole simulation interval is $3315 \cdot 10^4 \pm 4.32\%$ (veh·s) (for 5 times running with the same initial conditions).

In column (b) of Fig. 5, results of utilizing the shortest-path algorithm are presented (every 6 simulation time steps, the shortest-path algorithm recalculate the shortest routes). As can be observed, the congestion level is less than for the fixed routing case. However, the route advices in this approach are determined based on the current situation of network. Therefore, this approach is unable to take into account the future impacts of the trip demands on the accumulation and hence it cannot prevent the congestion from occurring in the intermediate regions. Nevertheless, by rerouting the traffic, the level of congestion reduces a bit (as can be observed from columns (a) and (b) of Fig. 5) as a result of preventing the traffic from entering the congested regions. The overall delay in the network for the whole simulation period is $2431 \cdot 10^4 \pm 8.45\%$ (veh·s) (again for 5 times running with the same initial conditions).

The best performance is achieved by the MPC scheme as shown in the third column of Fig. 5. The total delay in the network for the whole simulation interval is $1820 \cdot 10^4 \pm 9.87\%$ (veh·s). The achieved number is again the average over 5 simulations of the whole system.

The congestion level is significantly reduced in the destination and intermediate regions. The total delay obtained using MPC is much lower than the two other approaches meaning that the arrival rates are high in the proposed framework.

Note that another major advantage of using the proposed high-level routing scheme is that the computation time is reasonable compared to other routing methods in the literature which are based on detailed modeling.

V. CONCLUSIONS AND FUTURE RESEARCH

A high-level scheme for optimal dynamic route guidance in urban traffic networks using the macroscopic fundamental diagram (MFD) has been presented. On the high level, the dynamics of the urban regions and the flows of vehicles traveling towards multiple destinations in the network were described using an aggregate traffic flow model developed based on the MFD. The presented model enables us to efficiently model and control urban networks that can be partitioned into a number of homogeneous regions. Next, the route guidance problem was solved on the high level and as finding the optimal splitting rates towards neighboring regions.

Taking into account that the modeling approach does not depend on the shape of the regions, we have developed the model for a grid network and solved an optimization problem inside the MPC framework in order to find the optimal splitting rates. The obtained results showed significant performance of the proposed scheme over an existing shortest-path method.

Note that the optimal splitting rates are realized using local controllers installed at the borders of regions and therefore the MFDs will not be altered. However, as an extension to the current work and to improve the route guidance, we aim at utilizing the idea of having multiple timing plans inside regions (and consequently defining multiple MFDs for each region) to provide the opportunity for having control inside the urban regions. Further, since the routing problem is solved on a high level, the computational complexity of the proposed scheme is expected to be low compared to other existing approaches that are based on detailed modeling. This should be investigated using extensive numerical experiments. Also, in order to reduce the computation time even more, we aim at approximating and reformulating the model in order to achieve mixed integer linear optimization problems.

Furthermore, more extensive tests and validation of the model along with the performance evaluation of the proposed multi-level route guidance scheme using real networks' layouts and empirical data are included in the future research.

REFERENCES

[1] J. W. Godfrey, "The mechanism of a road network," *Traffic Engineering and Control*, vol. 11, no. 7, pp. 323–327, 1969.

[2] C. F. Daganzo, "Urban gridlock: Macroscopic modeling and mitigation approaches," *Transportation Research Part B*, vol. 41, no. 1, pp. 49–62, 2007.

[3] C. F. Daganzo and N. Geroliminis, "An analytical approximation for the macroscopic fundamental diagram of urban traffic," *Transportation Research Part B*, vol. 42, no. 9, pp. 771–781, 2008.

[4] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: some experimental findings," *Transportation Research Part B*, vol. 42, no. 9, pp. 759–770, 2008.

[5] A. Mazloumian, N. Geroliminis, and D. Helbing, "The spatial variability of vehicle densities as determinant of urban network capacity," *Philosophical Transactions of the Royal Society A*, vol. 368, no. 1928, pp. 4627–4647, 2010.

[6] N. Geroliminis and J. Sun, "Properties of a well-defined macroscopic fundamental diagram for urban traffic," *Transportation Research Part B*, vol. 45, no. 3, pp. 605–617, 2011.

[7] Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transportation Research Part B*, vol. 46, no. 10, pp. 1639–1656, December 2012.

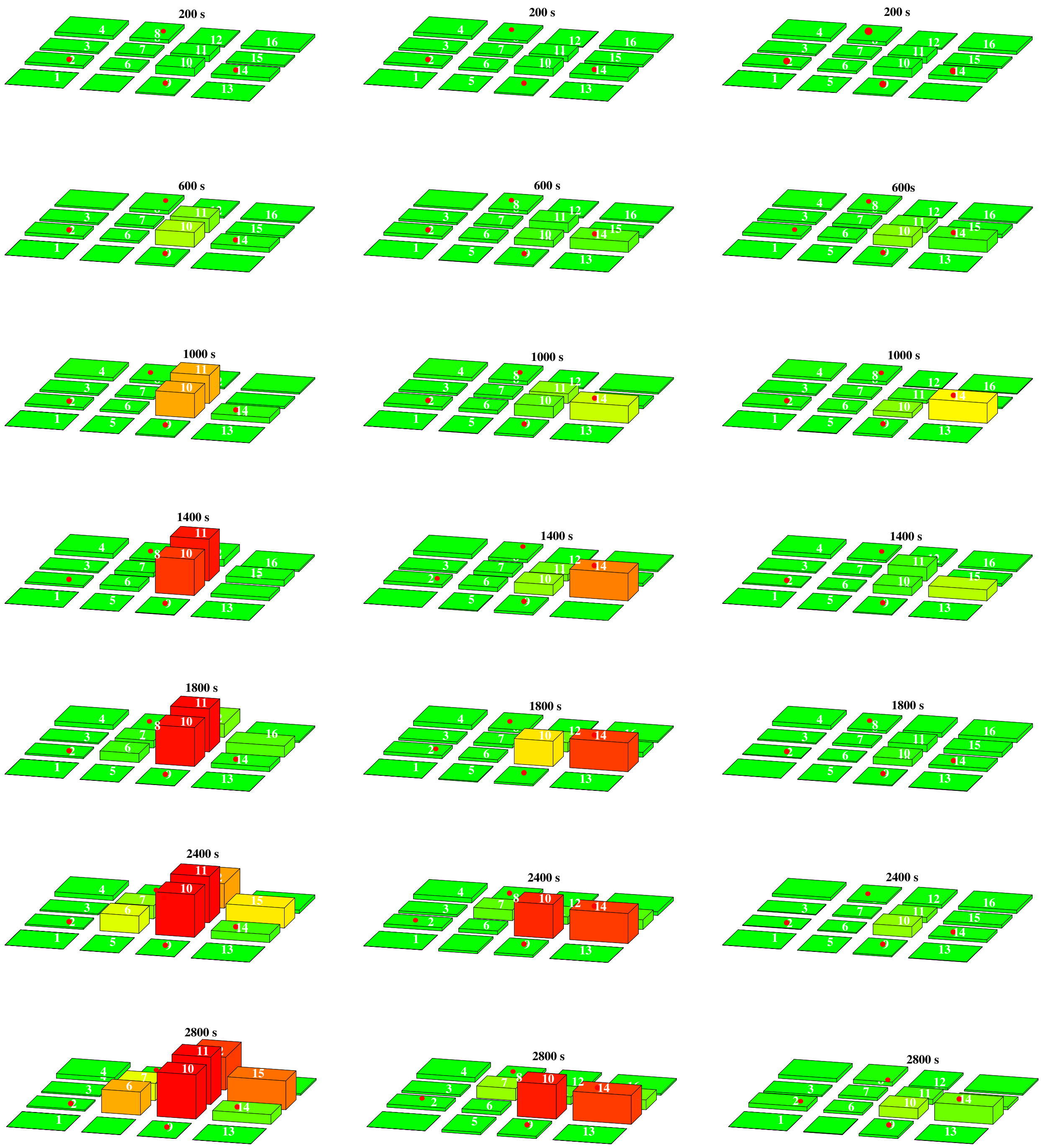
[8] T. Yoshii, Y. Yonezawa, and R. Kitarmura, "Evaluation of an area metering control method using the macroscopic fundamental diagram," in *Proceedings of the 12th World Conference of Transportation Research*, Lisbon, Portugal, 2010, pp. 1–12.

[9] N. Geroliminis, J. Haddad, and M. Ramezani, "Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 348–359, 2013.

[10] M. Keyvan-Ekbatani, A. Kouvelas, I. Papamichail, and M. Papageorgiou, "Exploiting the fundamental diagram of urban networks for feedback-based gating," *Transportation Research Part B*, vol. 46, no. 10, pp. 1393–1403, 2012.

[11] M. Hajiahmadi, J. Haddad, B. De Schutter, and N. Geroliminis, "Optimal hybrid macroscopic traffic control for urban regions: Perimeter and switching signal plans controllers," in *Proceedings of the 12th European Control Conference*, Zurich, Switzerland, 2013.

[12] S. Peeta and A. Ziliaskopoulos, "Foundations of dynamic traffic assignment: The past, the present and the future," *Networks and Spatial Economics*, vol. 1, no. 3, pp. 233–265, 2001.



(a) (b) (c)

Fig. 5. Results for 4×4 network, (a) Uncontrolled (fixed routes), (b) Shortest-path algorithm, (c) Optimal dynamic routing using MPC

[13] B. N. Janson, “Dynamic traffic assignment for urban road networks,” *Transportation Research Part B*, vol. 25B, no. 2/3, pp. 143–161, 1991.

[14] T. Saumtally, J. Lebacque, and H. Haj-Salem, “Static traffic assignment with side constraints in a dense orthotropic network,” *Procedia- Social and Behavioral Sciences*, vol. 20, no. 1, pp. 465–474, 2011.

[15] M. Papageorgiou, “Dynamic modeling, assignment, and route guidance in traffic networks,” *Transportation Research Part B*, vol. 24B, no. 6, pp. 471–495, 1990.

[16] V. L. Knoop, S. Hoogendoorn, and J. W. C. van Lint, “Routing strategies based on the macroscopic fundamental diagram,” in *Proceedings of the 91th Annual Meeting of the Transportation Research Board*, Washington, DC, United States, 2012.

[17] J. Maciejowski, *Predictive Control with Constraints*. Harlow, England: Prentice Hall, 2002.

[18] J. Rawlings and D. Mayne, *Model Predictive Control: Theory and Design*. Madison, WI, United States: Nob Hill Publishing, 2009.

[19] C. Daganzo, “The cell transmission model, part ii: Network traffic,” *Transportation Research B*, vol. 29, no. 1, pp. 79–94, 1995.

[20] S. Ray, *Graph Theory with Algorithms and Its Applications*. New Delhi, India: Springer, 2013.