Transportation Research Part C 86 (2018) 245–262

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A resource allocation problem to estimate network fundamental diagram in heterogeneous networks: Optimal locating of fixed measurement points and sampling of probe trajectories

Ali Zockaie[a,*], Meead Saberi[b], Ramin Saedi[a]

[a] Department of Civil and Environmental Engineering, Michigan State University, USA
[b] Institute of Transport Studies, Civil Engineering Department, Monash University, Australia

ARTICLE INFO

Keywords:
Network Fundamental Diagram (NFD)
Macroscopic Fundamental Diagram (MFD)
Heterogeneous networks
Probe trajectories

ABSTRACT

Network Fundamental Diagram (NFD) or Macroscopic Fundamental Diagram (MFD) represents dynamics of traffic flow at the network level. It is used to design various network-wide traffic control and pricing strategies to improve mobility and mitigate congestion. NFD is well defined when congestion distribution in the network is homogenous. However, in real world networks traffic is often heterogeneously distributed and initiated from an asymmetric and time-varying origin-destination (OD) demand matrix. In this paper, we formulate a resource allocation problem to find the optimal location of fixed measurement points and optimal sampling of probe trajectories to estimate NFD accounting for limited resources for data collection, network traffic heterogeneity and asymmetry in OD demand in a real-world network. Data from probe trajectories are used to estimate space-mean speed while data from fixed detectors are used to estimate traffic flow. Thus, the proposed model does not require an aggregate penetration rate of probe vehicles to be known a priori, which is one of the main contributions of this study. The proposed model is a mixed integer problem with non-linear constraints known to be NP-hard. A heuristic solution algorithm (Simulated Annealing) is implemented to solve the problem. Using a calibrated simulation-based dynamic traffic assignment model of Chicago downtown network, we present successful application of the proposed model and solution algorithm to estimate NFD. The results demonstrate sensitivity of the NFD estimation accuracy to the available budget, namely number of fixed measurement points and probe trajectories. We show that for a fixed proportion of OD trajectories, the increase in the proportion of fixed detection points increases the accuracy of NFD estimation as expected. However, when the proportion of fixed detection points is set to be constant, the increase in the proportion of OD trajectories does not necessarily improve the estimated NFD. Results hold true when varying demand is used to emulate variation in day-to-day traffic patterns. The robustness of the proposed methodology to the initial solution and trajectory availability for each OD pair is demonstrated in the numerical results section. We also found that a uniform distribution of selected links and ODs for NFD estimation across the network may not necessarily result in an optimal solution. Instead, distribution of links and OD pairs should follow the same distribution of links and OD pairs in the network.

* Corresponding author.
E-mail address: zockaiea@egr.msu.edu (A. Zockaie).

https://doi.org/10.1016/j.trc.2017.11.017
Received 1 June 2017; Received in revised form 9 October 2017; Accepted 15 November 2017

Nomenclature

T	Number of time intervals over the horizon for NFD estimation
t	time interval index
ζ	weight factor in objective function for minimizing deviation of estimated average flow from the ground-truth average network flow
η	weight factor in objective function for minimizing deviation of estimated average density from the ground-truth average network density
Q_t	ground-truth average network flow at time interval t
$\widehat{Q}_t$	estimated average network flow by fixed detectors at time interval t
K_t	ground-truth average network density at time interval t
$\widehat{K}_t$	estimated average network density at time interval t
I	number of links in the network
i	Link number index
q_i^t	flow at link i at time interval t
l_i	lane-length of link i
s_i^t	space-mean speed at link i at time interval t
x_i	binary variable associated with fixed detection at link i
J	number of origin-destination pairs in the network
j	origin-destination pair index
$K(j)$	number of trajectories available for origin-destination pair j
k	Trajectory index for origin–destination pairs
p_{ijk}^t	binary parameter specifying if kth trajectory of origin-destination pair j includes link i at time interval t
$\widetilde{tt}_{ijk}^t$	experienced travel time at link i and time interval t by k^{th} trajectory of origin-destination pair j
$\widehat{tt}_i^t$	experienced average travel time at link i and time interval t by available trajectory of selected origin-destination pairs
$\widehat{s}_i^t$	Estimated space-mean speed from mobile probe trajectories at link i at time interval t
z_i^t	binary variable specifying if there is any trajectory from selected origin-destination pairs that includes link i at time interval t
M	a large number
w_i^t	binary variable specifying if estimated speed is available through at least one trajectory from selected origin-destination pairs including link i at time interval t, and estimated flow is available through the detector at link i
y_j	Binary variable associated with the probe trajectory data for origin–destination pair j
c_i	data collection or acquisition cost if there is a fixed detector at link i
f_j	data collection or acquisition cost for probe trajectory data of origin-destination pair j
B	total available budget for data collection or acquisition

1. Introduction

The network-wide relationship between average flow, average density, and average speed, known as Network Fundamental Diagram (NFD) or Macroscopic Fundamental Diagram (MFD), is a powerful tool for representing traffic dynamics in large-scale networks (Godfrey, 1969; Mahmassani et al., 1984, 1987; Geroliminis and Daganzo, 2008). NFD can be used to design and implement specific control and pricing strategies to improve mobility at the network level (Haddad and Geroliminis, 2012; Zheng et al., 2012; Geroliminis et al., 2012; Keyvan-Ekbatani et al., 2012; Ramezani et al., 2015; Yildirimoglu et al., 2015; Haddad and Mirkin, 2016; Mariotte et al., 2017). NFD is well defined and has low scatter when congestion distribution in the network is homogenous (Buisson and Ladier, 2009; Ji et al., 2010; Mazloumian et al., 2010; Daganzo et al., 2011; Gayah and Daganzo, 2011; Geroliminis and Sun, 2011; Saberi and Mahmassani, 2012; Mahmassani et al., 2013; Saberi and Mahmassani, 2013; Knoop et al., 2012; Zockaie et al., 2014b). Estimating NFD in real-world networks, when data collection budget is limited and network traffic is heterogeneous and initiated from an asymmetric and time-varying origin-destination (OD) demand matrix, is a challenging task.

Analytical methods to estimate NFD based on variational theory developed previously by Daganzo and Geroliminis (2008) and later refined by Geroliminis and Boyaci (2012) and Leclercq and Geroliminis (2013) are limited to urban corridors in stationary conditions and cannot be applied to large-scale heterogeneous networks. A recent study by Leclercq et al. (2014) evaluated existing estimation methods for NFD focusing only on homogenous network loading. They suggested that using the complete population of vehicle trajectories to estimate NFD is the only estimation method with no bias agreeing with recent findings of Saberi et al. (2014). However, availability of the entire population of trajectories is still limited in urban networks and will continue to be limited even when connected vehicles are deployed in near future. Gayah and Dixit (2013) proposed a method to estimate average network density using probe vehicles combined with NFD. Leclercq et al. (2014) suggested that combining information from probe vehicles and traffic loop detectors can also provide fairly accurate estimation of NFD in stationary conditions even for sample rates as low as 10%. Other studies by Ortigosa et al. (2014) and Nagle and Gayah (2014) estimate NFD using combined mobile probes and traffic loop detector data. Ortigosa et al. (2014) studied the optimal number and location of measurement points by minimizing the error in estimated average network density. However, they overlooked the potential of the application of probe trajectory data in NFD estimation. Nagle and Gayah (2014) proposed a method to estimate average network density and flow using data from mobile probes given a constant and known penetration rate of probes across a network. In a later study, Du et al. (2015) extended the method to varying penetration rates with heterogeneous demand in an idealized square grid network. A limitation of this method is that the penetration rates of probes must be known a priori. More recently, Ambuhl and Menendez (2016) proposed a fusion algorithm that decomposes the network into two sub-networks and uses both loop detector data and floating car data to estimate NFD.

Despite the growing number of studies on NFD estimation, there is still a need to further develop methods to properly estimate NFD when network loading is not homogenous, data from fixed detectors and mobile probes are combined, and data collection resources are constrained. Data from fixed detectors are not always available uniformly throughout an urban network. Similarly, availability of trajectory data from mobile probes is usually geographically limited and yet expensive to acquire. Therefore, in this paper we formulate a resource allocation problem as a mathematical model to estimate NFD using a combined sample of vehicle trajectories and fixed detector data in a large-scale real-world network accounting for traffic heterogeneity and asymmetry in OD demand. Information from mobile probes provides an accurate space-mean network speed while information from fixed detectors provides an accurate mean network flow as shown in Leclercq et al. (2014) and Ambuhl and Menendez (2016). Similarly, probe vehicles can be applied to estimate flow only when all trajectories are available or the penetration rate of probe vehicles are known a priori. Building upon recent studies by Leclercq et al. (2014) and Du et al. (2015), the proposed method finds the optimal location of measurement points and optimal set of trajectories associated to OD pairs in a heterogonous network required to estimate an NFD. The main contribution of this study is, therefore, formulating NFD estimation as a resource allocation problem and the presented solution algorithm that facilitates estimation of NFD by optimally locating fixed measurement points and sampling of probe trajectories. Note that unlike some of the previous studies, we take advantage of both fixed and mobile data sources without the need for an aggregate pre-known penetration rate for probe vehicles.

The proposed mathematical model minimizes the deviation between the NFD and estimated network average flows and densities using the optimal configuration of detectors and set of trajectories from OD pairs. The optimal configuration considers a budget constraint for detector installation and trajectory data acquisition. In the proposed formulation, flows are estimated from links selected to be equipped with the detectors. However, traffic flow can only be used in NFD estimation, if at least a single trajectory is available to measure the link space mean speed. The proposed methodology incorporates the ground-truth NFD as input to find a subset of links and trajectories to estimate NFD, and it does not rely on the ground-truth NFD calculation method. Evidently, the ground-truth NFD may not be available or might just be measured for certain conditions. Here, we focus on estimation of NFD in general conditions considering variability of demand and day-to-day traffic patterns. Therefore, a surrogate is necessary to provide this input such as analytical methods or traffic simulation. Assuming that this surrogate provides the ground-truth NFD, the proposed methodology finds the optimal configuration of links and trajectories to collect data and estimate the NFD considering the available budget for data collection. To demonstrate the performance of the proposed framework based on a certain ground-truth NFD, we compare the estimated NFD by the selected detectors and trajectories with the ground-truth NFD, under a range of different demands. This comparison confirms applicability of the model various conditions. Note that the presented approach also contributes to estimation of network traffic state any given time t. Even when the ground-truth NFD is known, we still need to estimate network traffic state for traffic control purposes; in other words, where on the NFD the network is at any time t.

The proposed model is a mixed integer problem with non-linear constraints. Mixed integer linear problems are known to be NP-hard, and non-linear constraints add more complexity by providing local optimal solutions for the formulated problem. As a result, no exact solution method exists for large-scale network applications. Therefore, a heuristic solution algorithm, Simulated Annealing (SA), is implemented to solve the problem. The method starts with an initial solution that is randomly selected as the current solution. Then, the current solution is perturbed iteratively to a neighbor solution (selected randomly) and is substituted by the neighbor solution, when the objective function is improved. The neighbor solution is also probabilistically accepted when the objective function is not improved to avoid local optimal solutions. This probability needs to be reduced by the iteration number to ensure convergence of the algorithm. We show that how different combinations and variations in the proportion of fixed detection points and OD trajectories affect the estimated NFD. The focus of the paper is not on the implemented solution algorithm. Therefore, no comparison with other possible solution methodologies is made. The paper mainly focuses on the problem formulation and numerical results from a real-world network. The successful implementation of the SA algorithm in the numerical results section for a case study with large-scale network application confirms suitability of this heuristic method for the proposed mathematical model. The SA algorithm is a proper methodology for problems in which evaluation of the objective function is intuitive for a given solution of decision variables. This is the case for the proposed model, where average flows and densities need to be calculated for a given detector and set of trajectories.

The remainder of the paper is organized as follows: Section 2 presents the mathematical problem formulation. Section 3 includes a discussion on the proposed solution algorithm. Section 4 presents numerical results from implementation of the formulated problem on a sub-network of downtown Chicago. Section 5 propose some practical insights on inquiry of probe trajectories from certain OD pairs and selection of links to be equipped with detectors. Section 6 concludes the paper and suggests directions for the future research.

2. Model formulation

Here we formulate the NFD estimation problem as a resource allocation problem to find the optimal locations of fixed measurement points and the optimal set of trajectories associated with OD pairs given a fixed data acquisition budget. For this purpose, the presented mathematical model finds a subset of candidate links for fixed detectors in order to estimate an NFD with the maximum possible precision. Moreover, the mathematical model finds a subset of OD pairs in which their associated trajectory data are needed to improve NFD estimation. The optimal combination of the fixed detectors and OD pairs provides an accurate estimation of both network-wide space-mean speeds and average flows, when a sufficient fixed data acquisition budget is provided. Using the space-mean speeds estimated by trajectories and estimated flows by fixed detectors, densities can be calculated using the traffic flow fundamental identity $q = k.v$. Both flow and space-mean speed are estimated at the link level to be later incorporated in estimation of

network-wide averages.

The constraints of the model limit the total number of detectors to be deployed and the total number of trajectories to be acquired. The optimal solution for the mathematical model includes the optimal measurement locations and OD pairs that provide an estimated NFD with minimum discrepancy from the case with no constraints on the number of loop detectors and trajectories. Here, we also present a heuristic solution algorithm to estimate the optimal solution in a reasonable computational time frame for large-scale applications considering heterogeneous traffic demand distribution and asymmetric OD demand matrix. The proposed problem formulation is as follows.

$$Min \quad \sum_{t=1}^{T} [\zeta(Q_t - \widehat{Q}_t)^2 + \eta(K_t - \widehat{K}_t)^2] \tag{1}$$

Subject to

$$Q_t = \frac{\sum_{i=1}^{I}(q_i^t l_i)}{\sum_{i=1}^{I} l_i} \quad \forall\, t = 1,...,T \tag{2}$$

$$K_t = \left[\sum_{i=1}^{I}(l_i q_i^t / s_i^t)\right] \Big/ \sum_{i=1}^{I} l_i \quad \forall\, t = 1,...,T \tag{3}$$

$$\widehat{tt}_i^{\,t} = \frac{\left[\sum_{j=1}^{J}\sum_{k=1}^{K(j)}(p_{ijk}^t \widetilde{tt}_{ijk}^t y_j)\right]}{\left[\sum_{j=1}^{J}\sum_{k=1}^{K(j)}(p_{ijk}^t y_j)\right]} \quad \forall\, t = 1,...,T, \forall\, i = 1,...,I \tag{4}$$

$$\widehat{s}_i^t = l_i / \widehat{tt}_i^{\,t} \quad \forall\, t = 1,...,T, \quad \forall\, i = 1,...,I \tag{5}$$

$$\sum_{j=1}^{J}\sum_{k=1}^{K(j)}(p_{ijk}^t y_j) \leqslant z_i^t M \quad \forall\, t = 1,...,T, \forall\, i = 1,...,I \tag{6}$$

$$z_i^t \leqslant \sum_{j=1}^{J}\sum_{k=1}^{K(j)}(p_{ijk}^t y_j) \quad \forall\, t = 1,...,T, \forall\, i = 1,...,I \tag{7}$$

$$z_i^t + x_i - 1 \leqslant w_i^t \quad \forall\, t = 1,...,T, \quad \forall\, i = 1,...,I \tag{8}$$

$$w_i^t \leqslant z_i^t \quad \forall\, t = 1,...,T, \quad \forall\, i = 1,...,I \tag{9}$$

$$w_i^t \leqslant x_i \quad \forall\, t = 1,...,T, \quad \forall\, i = 1,...,I \tag{10}$$

$$\widehat{Q}_t = \left[\sum_{i=1}^{I}(l_i q_i^t w_i^t)\right] \Big/ \left[\sum_{i=1}^{I} l_i w_i^t\right] \quad \forall\, t = 1,...,T \tag{11}$$

$$\widehat{K}_t = \left[\sum_{i=1}^{I}\left(\frac{q_i^t}{\widehat{s}_i^t} l_i w_i^t\right)\right] \Big/ \left[\sum_{i=1}^{I}(l_i w_i^t)\right] \quad \forall\, t = 1,...,T \tag{12}$$

$$\sum_{i=1}^{I} c_i x_i + \sum_{j=1}^{J} f_j y_j \leqslant B \tag{13}$$

$$x_i = 0 \text{ or } 1 \quad \forall\, i = 1,...,I \tag{14}$$

$$y_j = 0 \text{ or } 1 \quad \forall\, j = 1,...,J \tag{15}$$

$$z_i^t = 0 \text{ or } 1 \quad \forall\, i = 1,...,I \tag{16}$$

$$w_i^t = 0 \text{ or } 1 \quad \forall\, j = 1,...,J \tag{17}$$

In the formulation, x and y are binary decision variables, specifying if a detector is located on a link ($x_i = 1$) or a specific OD pair is selected to acquire data trajectories ($y_j = 1$). z and w are binary state variables. z specifies if there is any crossing trajectory on link i at time t based on the decision variable y over OD pairs. w specifies if the link should be considered for the network-wide average value calculation or not based on availability of detectors (specified by x) and trajectory data over the link (specified by z). The latter is one for a specific link and time interval when both x and z are equal to one for that link and time interval.

Eq. (1) lists the objective function as the weighted sum of squared deviations in network-wide average values of flow and density at different time intervals. In this equation, estimated network-wide average values by selected links and trajectories are compared with the NFD calculated by all links and trajectories. As the order of network-wide average values for flow and density are different, the two terms are weighted with different factors (ζ *and* η). Eqs. (2) and (3) present NFD using network-wide average flow and

density for each time interval based on all link flows and space-mean speeds. Note that the weighted average values are calculated considering the lane-length of each link. Eq. (4) calculates the average travel time on each link at each time interval using experienced travel times by crossing probe vehicles at the time interval of interest. The probe vehicles' travel times are considered in the average travel time, if the OD pair of the probe vehicle is selected in the optimal solution. This equation simply goes over all trajectories that can be observed (with y equal to 1) and adds the observed travel time for each link i along the trajectory with departure time interval of t to the total travel time for that link and time interval. Then, it divides the total observed travel time for each link and time interval by actual number of observed trajectories. Eq. (5) calculates the space-mean speed for each link and time interval. Eqs. (6) and (7) specify if there is any crossing probe vehicle available that its origin and destination are among the selected OD pairs in the optimal solution (specifying z based on y). When there is at least one trajectory that crosses link i at time interval t, which can be observed (with y equal to 1), z needs to be equal to 1 for that link and time interval to hold equation 6. In this case, equation 7 would not be binding. When there is no such trajectory, Eq. (7) sets z for that link and time interval to zero and Eq. (6) would not be binding. In this model, for simplicity of the presentation, we assume that when an OD pair is selected, all trajectories traveling between the origin and destination are available for NFD estimation. The model formulation can be easily adjusted to consider other assumptions such as partial trajectory availability for OD pairs by considering an extra binary variable for each trajectory in these equations (u_{kj}). Eqs. (8)–(10) determine if a link and time interval should be considered in the network-wide average calculation based on the availability of both detector and crossing probe vehicles at that link and time interval (specifying w based on z and x). Note that the three linear constraints in Eqs. (8)–(10) can be replaced by a single non-linear constraint as $w_i^t = x_i \times z_i^t$. Eqs. (11) and (12) estimate the average flow and density at each time interval over the selected links in the optimal solution with positive w. Eq. (13) enforces the budget constraint by limiting the number of selected links and OD pairs. Eqs. (14)–(17) are feasibility constraints for the binary variables. Note that the feasibility constraint for non-binary variables, which is positivity constraints, is not listed in the above formulation.

3. Solution algorithm

The proposed model, formulated as a Mixed Integer Non-Linear Program (MINLP), aims to minimize the deviation of the estimated NFD by a limited number of links and trajectories, due to the budget constraint for data collection or acquisition, and the calculated NFD using all the network links and trajectories. The location problems are known to be NP-hard, especially for the defined problem in this study where the non-linear constraints add to the problem complexity. Thus, we develop a heuristic algorithm to estimate NFD with optimal sampling of fixed detection points and OD pairs. Here, we select and implement a Simulated Annealing (SA) algorithm (e.g., Metropolis et al., 1953; Davidovich and Mikhailovich, 1980; Kirkpatrik et al., 1983; Cerny, 1985; Van Laarhoven and Aarts, 1987) to solve the proposed mathematical model. SA is a proper solving technique for the problem of interest in this study, as it can find the optimal solution without being trapped in local optimal solutions. Note that the problem of interest has non-linear constraints that might lead to including many local optimal solutions. SA is an appropriate methodology for problems in which the objective function can be evaluated by simple calculations when the decision variables are known. This is the case for the formulated problem where given x and y, the objective function can be simply calculated.

The implementation of the SA method in this paper is based on the Metropolis algorithm (Metropolis et al., 1953; Hejazi, 1999) and we follow the more recent implementation in Zockaie et al. (2016) and Ghamami et al. (2016). The algorithm derives its underlying idea from the heating and cooling phenomenon of solid materials. A solid at a liquid state cooled gradually will form a crystal, whereas rapid cooling from the same initial condition will lead to a frozen solid. Accordingly, the SA algorithm usually "simulates" the process of gradual temperature decreasing. At each temperature, an equilibrium state should be achieved before moving to a lower temperature. The final solution will be achieved at the minimum or final temperature.

The SA algorithm starts with a feasible initial solution. In each iteration, it picks a neighborhood solution through a local search and evaluates the objective values, $C(.)$, at the current, n, and previous, m, solutions. The algorithm will then decide whether to move to the new solution or to stay at the current solution as follows. It sets $\Delta C_{mn} = (C(n) - C(m))/C(m)$. If $\Delta C_{mn} \leqslant 0$, n is selected as the new solution with probability 1, and if $\Delta C_{mn} > 0$, n is selected as the new solution with the probability of $\exp\left(\frac{-\Delta C_{mn}}{Temp}\right)$, where *Temp* is a control parameter interpreted as the "temperature" in the cooling process. This implies that a solution worse than the current one may be adopted with a non-zero probability. This specification provides the opportunity of not being trapped in a local optimal solution. The above process continues until an equilibrium state is reached at a certain temperature such as *Temp*. In each step, *Temp* is reduced and the system is equilibrated under a new temperature. The algorithm terminates at a small *Temp* such that any increase in the objective function value is prohibited (Zockaie et al., 2016). A description of the basic SA algorithm follows.

3.1. Simulated annealing algorithm

Step 1: Consider an initial solution m (which assigns variable x and y to 0 or 1 for different sets of i and j considering the budget constraint to maintain feasibility)
Set iteration index $iter = 0$
Set an initial $Temp_0$, the number of main iteration *Iter* and the number of inner iteration *Inner-Iter*
Step 2: Repeat Steps 3–5 *Inner-Iter* times and afterwards go to Step 6.
Step 3: Select solution n close to m (finding a neighbor solution by changing one of x and y values for a particular i or j considering the budget constraint to maintain feasibility)

Step 4: Calculate $\Delta C_{mn} = (C(n) - C(m))/C(m)$ (where function $C(.)$ calculates the objective function)
Step 5: If $\Delta C_{mn} \leqslant 0$, then set $m = n$;
Otherwise, if $\exp\left(\frac{-\Delta C_{mn}}{Temp_k}\right) > Random\ number \in [0,1)$ then set $m = n$.
Step 6: iter = iter + 1
Step 7: Reduce *Temp* using $Temp_{iter+1} = f(Temp_{iter})$, if $iter \leqslant Iter$ go to Step 2, else stop.

Several remarks on implementation issues are in order here. The first has to do with the choice of control parameters, such as initial temperature, the number of inner iterations at each temperature, the temperature reduction function, and the number of main iterations. The second issue is related to the choice of initial solution and the third and perhaps the most important implementation question is how to get a neighbor solution by perturbing the current solution. In this study, following Zockaie et al. (2016), the initial temperature is assumed to be 0.05. The temperature is decreased by the rate of 0.85 for each outer loop iteration. Therefore, we use the following relation at Step 7 of the algorithm:

$$Temp_{iter+1} = 0.85 * Temp_{iter} \text{ or } Temp_{iter+1} = (0.85)^{iter} Temp_0 \tag{18}$$

To select the initial and neighbor solutions, we use random number generation, considering uniform distribution, to select specific x, and y to be included in the initial or neighbor solutions. The main challenge in the random selection strategy is how to move from current solution to a neighbor solution satisfying the budget constraint (Eq. (13)). For the simplicity, we assume that the costs associated with the fixed detection and trajectory data are uniformly distributed over the network and OD pairs. Therefore, we enforce the budget constraint through solving the problem for different proportions of both fixed detection points and OD trajectories. Furthermore, we assume $\zeta = 1$ and $\eta = 1$. Note that these assumptions do not affect the generality of the solution method and do not limit its applicability to the formulated problem. We also assume that the availability of trajectory data is linked to the OD pairs.

The algorithm uses proportion rates for available fixed detections, a, and trajectory OD pairs, b, as the main input. Note that parameters a and b do not represent the penetration rate of probe vehicles. They represent the proportion of employed fixed detection points and OD pairs, respectively. To further clarify, parameters a and b are only representative of the "estimation budget". Given the proportions, a set of links and OD pairs are randomly selected as the initial solution, and are stored as the current solution. To find the neighbor solution, first we randomly decide to perturb the current solution, by choosing between modifying the location of fixed measurement points and modifying OD trajectories, to move to a neighbor solution. Then, based on this decision, one of the links or OD pairs in the current solution is randomly selected and removed from the current solution. Also, one of the links or OD pairs that is not in the current solution is selected to be added in the neighbor solution. Note that by adding one link or OD pair, and removing another link or OD pair, the neighbor solution has exactly the same proportion rates as the current solution. Furthermore, for any given current and neighbor solutions, estimating the objective function is intuitive. At the selected locations of fixed measurements, the link flows are known, and crossing trajectories are used to estimate the space-mean speed. Knowing flow and space-mean speed, density can be calculated using the traffic flow fundamental identity. Finally, network-wide flow and density averages at each time interval can be calculated using Eqs. (11) and (12). Applying the random selection strategy to find the initial and neighbor solutions and estimating the objective function as explained earlier, the iterative steps of the algorithm can be implemented as discussed in Step 2 to Step 7 of the algorithm. In this study, we consider 100 inner loop iterations required iterations to reach an equilibrium state at

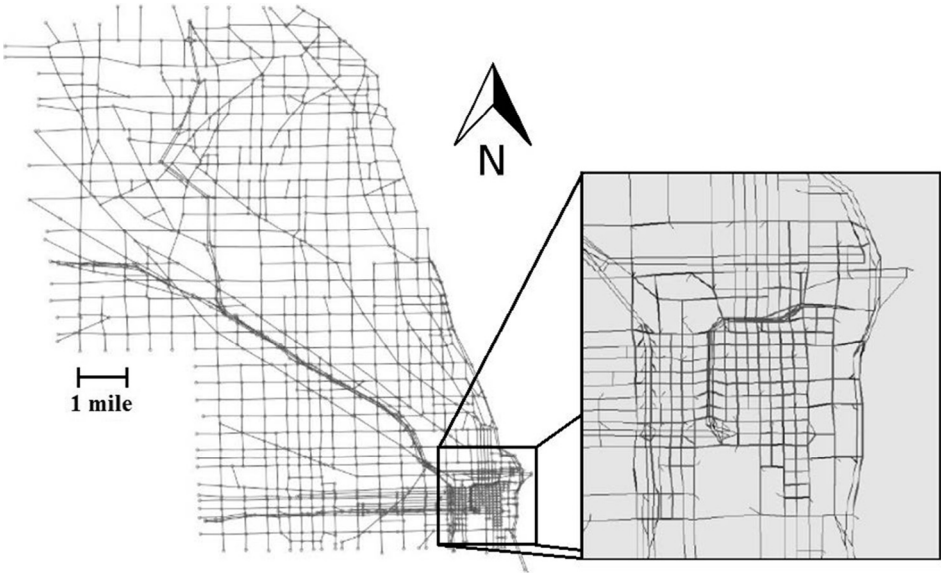

Fig. 1. Illustration of the Chicago network and its downtown sub-network.

each temperature for each of the outer loop iterations required iterations to decrease temperature and as a result decreasing probability of accepting worse neighbor solutions relative to the current solution. We also modify the temperature for 50 iterations as the outer loop.

4. Numerical results

Here, we apply the proposed mathematical model to a large-scale network of Chicago to estimate the NFD for downtown sub-network as illustrated in Fig. 1. In the proposed framework, the ground-truth NFD is an input. As mentioned earlier, the actual ground-truth might not be available at all or it might be measured under specific conditions. In this study, we have developed a methodology to estimate the ground-truth NFD using sampled observations from detectors and trajectories. Here, we use a surrogate of the ground-truth NFD as an input to optimize the sampling procedure. The surrogate NFD used here can be measured experimentally or analytically for small networks. However, for large-scale network applications, a mesoscopic traffic simulation model can alternatively be used. Here, we use a calibrated simulation-based dynamic traffic assignment (DTA) model of Chicago to estimate the surrogate for the ground-truth NFD. We will then find the optimal location of trajectories and detectors. Observations made by the located detectors and trajectories under general conditions provide proper estimation of NFD. To examine the applicability of the proposed methodology under general conditions, we also show performance of the estimation method for cases where ground-truth NFD is modified as a result of variation in demand and seed number for random number generations that are used for different purposes throughout the network simulation (demand generation, signal control, en-route users assignment, etc.).

4.1. Study network

We use DYNASMART-P to conduct the experiments and to obtain the ground-truth NFD using fixed measurements of density and flow at all links throughout the selected sub-network. DYNASMART-P is a simulation-based dynamic traffic assignment tool and it has the capability to find the dynamic user equilibrium in an iterative process. It also can distribute traffic over the network using the best

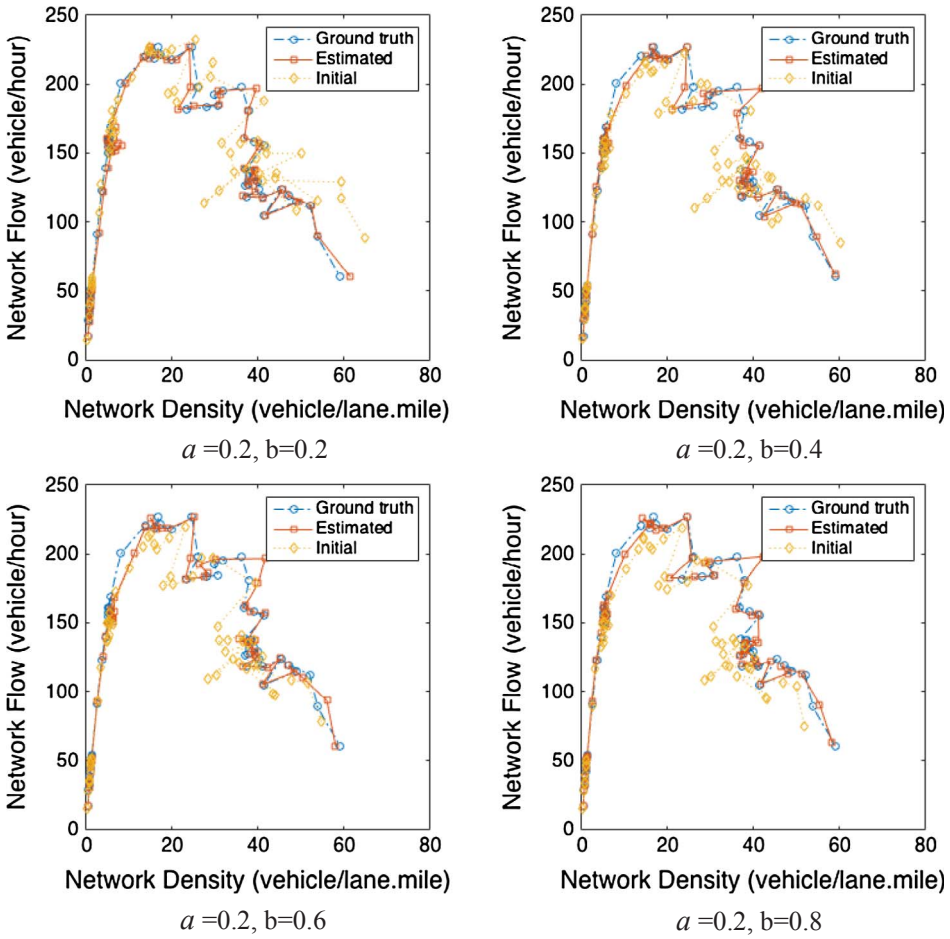

Fig. 2. Comparison of estimated NFD, initial NFD, and ground-truth NFD for different proportions of ODs b = {0.2, 0.4, 0.6, 0.8} with a constant proportion of fixed detectors a = {0.2}.

current path, which can be changing during the simulation time depending on the congestion distribution across the network. In this study, the best current path, often known as zero shot simulation, is used. Modeled trajectory of all vehicles in the network during the simulation horizon is available. The Chicago downtown network, which is bound from West and East by O'Hare airport and Lake Michigan respectively, is considered as the large-scale case study. It includes downtown Chicago and some Western and Northern suburban cities of Chicago, and contains 1578 nodes, 4805 links, and 218 zones. The simulation horizon is the AM peak period between 5:00 AM and 10:00 AM. The static hourly demand is provided by the Chicago Metropolitan Agency for Planning (CMAP) for the AM peak. The static demand is transformed into a time-dependent OD demand table through the OD-estimation techniques presented in Zockaie et al. (2014a) resulting in about 550,000 vehicles simulated in the network. In this case study, we are interested to estimate NFD for a sub-network located in Downtown Chicago. This sub-network, which is illustrated in Fig. 1, includes 921 links and 7212 OD pairs with positive demand. Chicago CBD is selected because it is the most congested part of the greater Chicago area network. The traffic distribution is more homogenous in this part of the network relative to the entire network. Also the NFD for this portion of the network is more difficult to estimate relative to the entire network, where the congestion regime is phased out in the average value calculation by many uncongested links in other parts of the network. Generally speaking, the proposed method can be applied to any network, larger or smaller in size without any computational complexity issue. In this study, we are interested to estimate the NFD for the CBD sub-network to be consistent with a previous study in (Mahmassani et al., 2013). Note that the dynamic simulation is conducted for the greater Chicago area network. Therefore, the OD pairs counted and reported in the paper relate to the larger network. We run the simulation for the larger network but only apply the proposed method to the CBD sub-network for the NFD estimation.

4.2. Optimal locating of fixed measurement points and sampling of probe trajectories

We first examine how different proportions of fixed detection points $a = \{0.2, 0.4, 0.6, 0.8\}$ and OD pairs $b = \{0.2, 0.4, 0.6, 0.8\}$ affect the initial and optimal objective values as a representation of discrepancy between the estimated and ground-truth NFD. We test all the 4×4 combinations and report selected numerical results to demonstrate the applicability of the proposed method in a

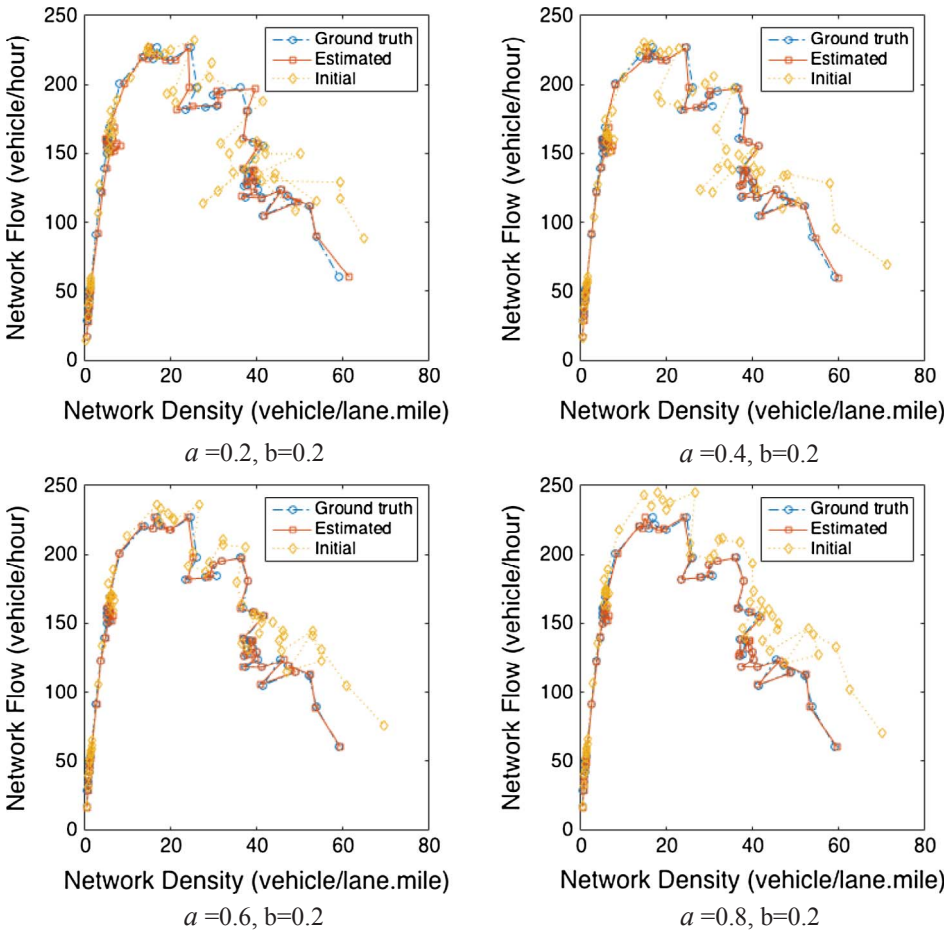

Fig. 3. Comparison of estimated NFD, initial NFD, and ground-truth NFD for different proportions of fixed detectors $a = \{0.2, 0.4, 0.6, 0.8\}$ with a constant proportion of ODs $b = \{0.2\}$.

large-scale heterogeneous network. Figs. 2 and 3 show a comparison of the ground-truth, initial, and estimated NFDs for a combination of a and b values. We first set the proportion of fixed detection points to a constant $a = 0.2$ and change the proportion of OD pairs $b = \{0.2, 0.4, 0.6, 0.8\}$ as illustrated in Fig. 2. The initial NFD is estimated from randomly selected links for fixed detection and randomly selected ODs for probe trajectories. The estimated NFDs obtained from optimally selected links and ODs are also shown. We then set the proportion of OD pairs to a constant $b = 0.2$ and change the proportion of fixed detection locations $a = \{0.2, 0.4, 0.6, 0.8\}$ as illustrated in Fig. 3. The comparison of NFDs demonstrates that the optimal estimated NFDs have smaller deviation from the ground-truth NFD compared to NFDs estimated from randomly selected initial solutions, as expected. Later, we also show that increasing the proportion of fixed detection points or increasing the proportion of OD pairs both improve the initial estimated NFD. Finally, we show that for a constant proportion of detection points, increasing the proportion of trajectories does not affect the quality of the estimated NFD. However, for a constant proportion of OD pairs, increasing the proportion of detection points improves the estimated NFD with small deviation from the ground-truth NFD.

Fig. 4 illustrates the location of the optimally selected fixed measurement points in the network for the same combination of a and b values. An interesting observation is the variation between the selected links across cases with the fixed $a = 0.2$. Although the proportion of fixed detection points is constant, the locations of the detection points vary because of the changing proportion of ODs b $= \{0.4, 0.6, 0.8\}$. Fig. 5 visualizes the selected OD pairs for the same combination of a and b values and for the entire network including all ODs using a schematic graph. Here, we draw a circular graph with nodes on the circumference of the circle in order of node degree and links connecting nodes. A node pair is connected with a link if the OD pair is selected for estimating NFD. The graph showing all ODs represents the entire OD matrix of the Chicago network. The circular illustration here does not represent the links entering or exiting the CBD area. All traffic analysis zones with positive demand crossing the CBD area are located on the circle to show the connectivity between different OD pairs in the network. We found that the selected ODs across the network vary even when the proportion of OD pairs is fixed (e.g. $b = 0.2$) because of varying proportion of fixed measurement points $a = \{0.2, 0.4, 0.6, 0.8\}$. Increasing detection points proportion for a fixed proportion of OD pairs results in more dispersed distribution of selected OD-pairs. This improves the network connectivity to cover the increased selected number of links.

Fig. 6 shows the convergence pattern for the same combination of a and b values. The initial objective value, representing the deviation between estimated and ground-truth NFD, is as large as 10,000. However, as the number of iterations increases, the SA algorithm significantly reduces the optimal objective value, demonstrating the convergence pattern of the algorithm. When the number of iterations is large enough, increasing the OD pair proportions does not improve the optimal objective value. In other words, when the link proportion is limited, there is no benefit in increasing the OD pair proportion to cover the entire network. However, increasing the OD pair proportion to a certain limit, ensures reliability of the optimal solution. Note that the number of iterations listed in this figure, includes both inner and outer loops. Iterations 1–100 are associated with the first outer loop and iterations 101–200 are associated with the second outer loop and so on.

Fig. 4. Selected locations for optimal estimation of NFD for different proportions of ODs and fixed detectors.

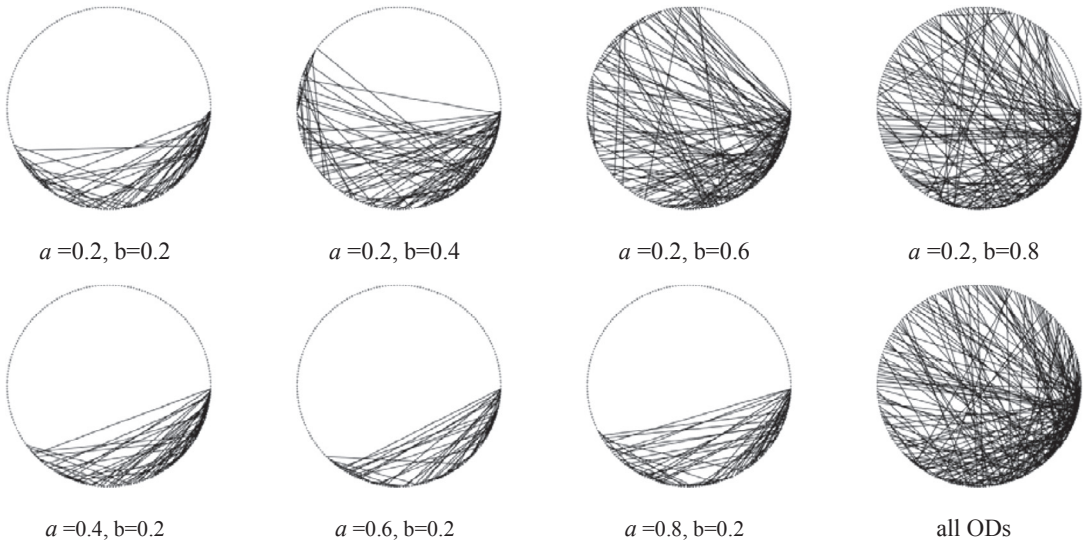

Fig. 5. Selected ODs for optimal estimation of NFD for different proportions of ODs b = {0.2, 0.4, 0.6, 0.8} with a constant proportion of fixed detectors a = {0.2} and for different proportions of fixed detectors a = {0.2, 0.4, 0.6, 0.8} with a constant proportion of ODs b = {0.2}; for all ODs.

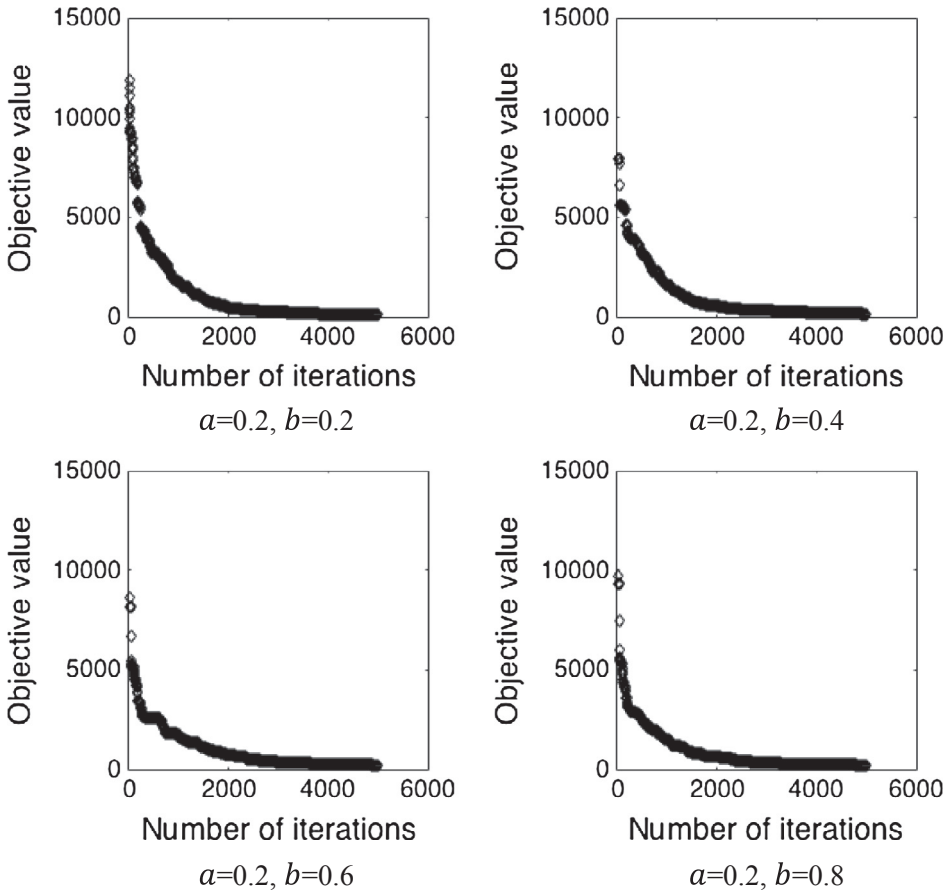

Fig. 6. Convergence pattern for different proportions of ODs a = {0.2, 0.4, 0.6, 0.8} with a constant proportion of fixed detectors b = {0.2}.

Fig. 7 shows the initial and optimal objective values as a function of a and b. In Fig. 7(a), we illustrate the change in the initial and optimal objective values for a set of constant proportions of fixed measurements a while changing b. The initial objective values are relatively large and do not follow a specific pattern given that they are mainly based on a random selection of links and OD pairs for

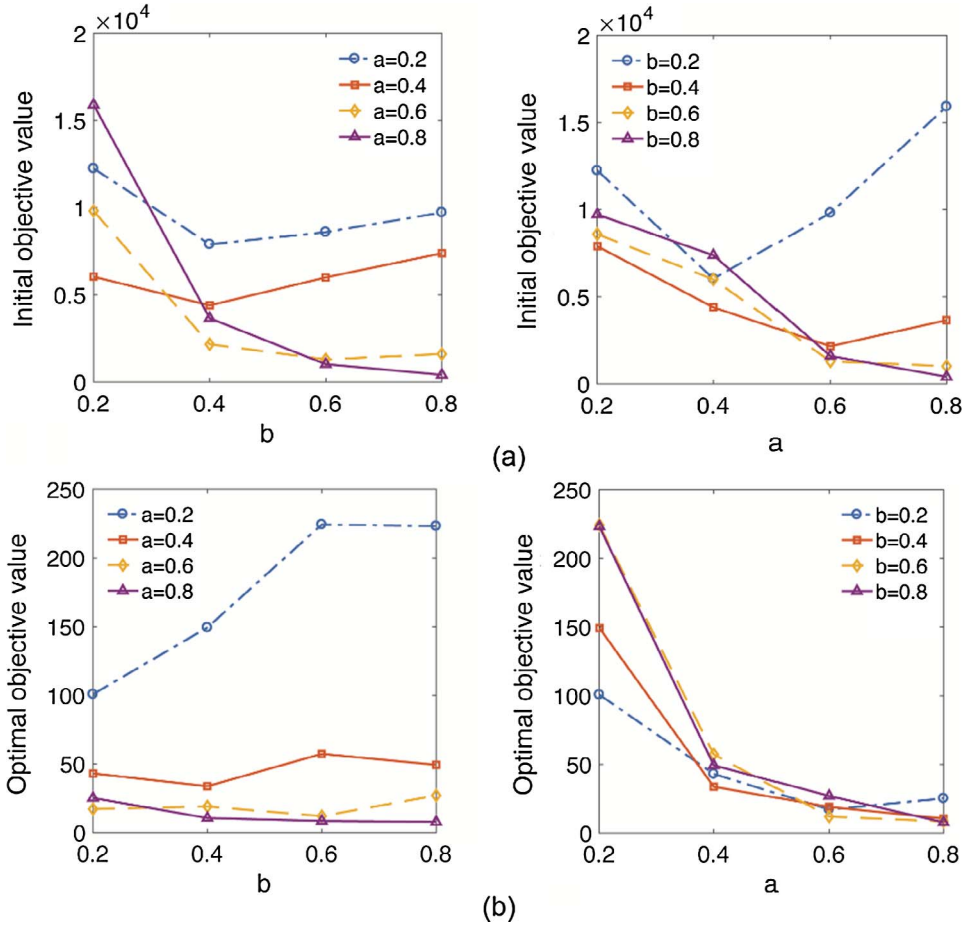

Fig. 7. Initial and optimal objective values as a function of a and b.

NFD estimation. Counter intuitively, the optimal objective value does not change monotonically with increase or decrease of b for constant link proportions a. However, the value of optimal objective function for different OD pair proportions decreases with increase of the link proportion. In Fig. 7(b), we study the change in the initial and optimal objective values for a set of fixed proportion of OD pairs b while changing a. The initial objective values follow a decreasing pattern for all b when a increases (except $b = 0.2$ due to the randomness of the methodology). The optimal objective values also follow a decreasing pattern when a increases.

Results suggest that for a fixed proportion of OD trajectories, the increase in the proportion of fixed detection points decreases the deviation from the ground-truth NFD as expected. However, when the proportion of fixed detection points is set to be constant, the increase in the proportion of OD trajectories does not necessarily improve the estimated NFD. This is due to the fact that for a certain link proportion, an optimal level of OD pair proportion can be found. Increasing or decreasing the level of OD pair proportion might lead to selection of some links in the optimal solution that increases the objective function value. To clarify why such pattern is observed in the numerical results, we present a simple example that explains how increasing the OD pair proportion might increase the objective function value. Let's say a certain link proportion results in selecting n links among the m links of the network. The OD pair proportion is set to a certain value that exactly results in crossing probe vehicles in $n - 1$ of the links. Let's say the left out link is the least similar link to the ground-truth NFD. Increasing the OD pair proportion results in selecting the left out link as the maximum link proportion is n. As this link has the minimum similarity to the ground-truth NFD, this addition would increase the deviation of the estimated NFD from the ground-truth NFD. Although increasing OD pair proportion might not improve the objective function value for a constant link proportion, it might improve the reliability of the estimated NFD under general conditions. In a rare case, there might be only a single link in the network with a fundamental diagram that matches the ground-truth NFD. Selecting this link alone, along with OD-pairs that include probe vehicles crossing this link, results in the objective function value of zero. However, this estimation is not reliable and with a minor change in the network conditions, the ground-truth NFD might deviate from the estimated NFD. Therefore, as far as the budget constraint allows we should increase the link and OD pair proportions.

To understand the reliability of the solution found by the proposed methodology, we study performance of NFD estimation by the selected detectors and trajectories under general conditions with different ground-truth NFDs relative to the one that is used as the input to the model. To this end, we consider three different ground-truth NFDs all derived from the simulation. The first ground-truth NFD is the base scenario that is used as the input for the mathematical model to find the optimal set of detectors and trajectories as

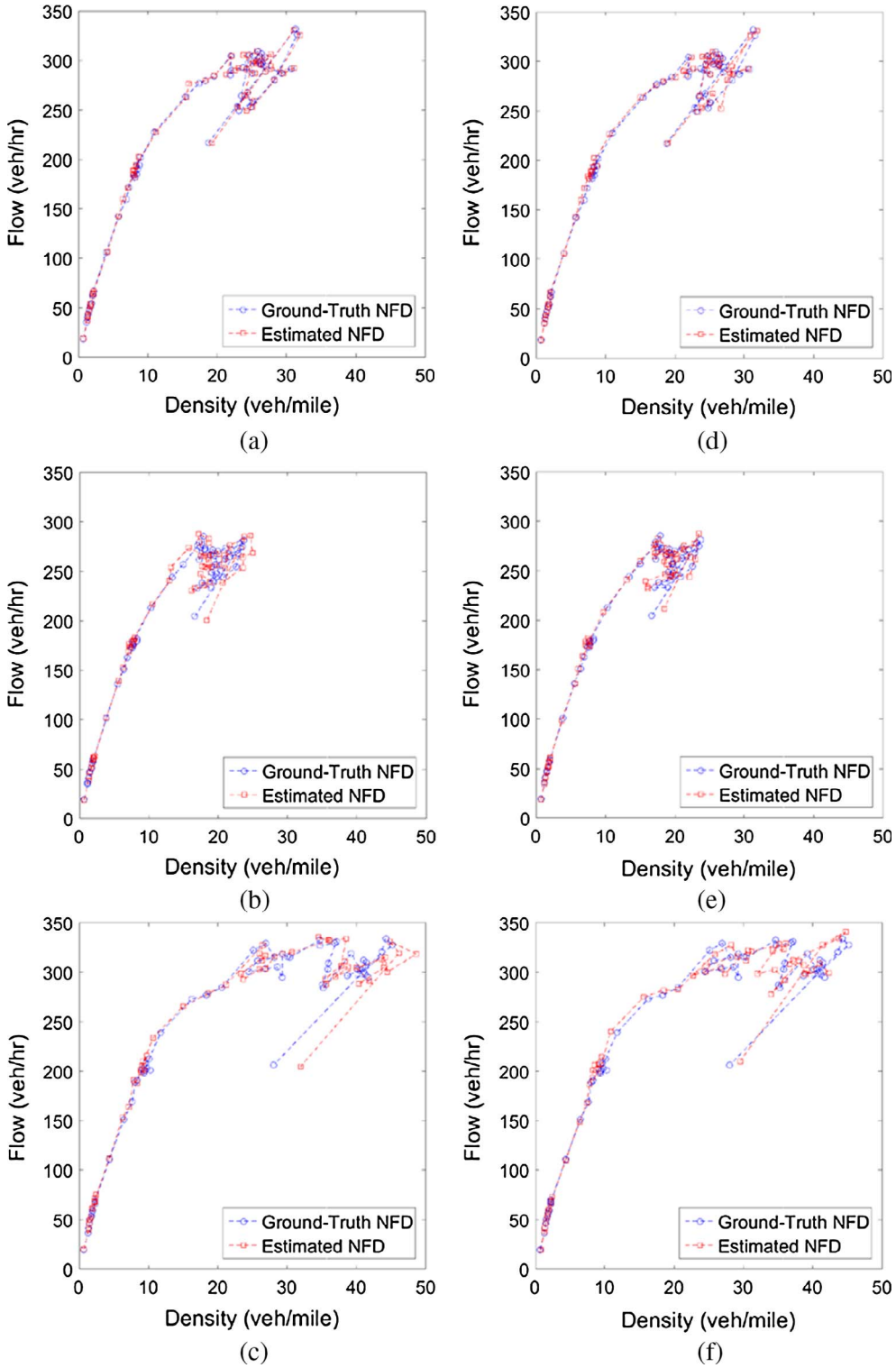

Fig. 8. Comparison of estimated and actual NFD (a) base scenario with $a = 0.6$, $b = 0.2$ (b) decreased-demand scenario with $a = 0.6$, $b = 0.2$ (c) increased-demand scenario with $a = 0.6$, $b = 0.2$ (d) base scenario with $a = 0.6$, $b = 0.6$ (e) decreased-demand scenario with $a = 0.6$, $b = 0.6$ (f) increased-demand scenario with $a = 0.6$, $b = 0.6$.

shown in Fig. 8(a) and (d). Decreasing the total simulated demand in the base scenario by 5 percent results in the second ground-truth NFD, shown in Fig. 8(b) and (e), and increasing the demand by 5 percent results in the third ground-truth NFD, shown in Fig. 8(c) and (f). In both scenarios, different seed numbers are used for generating random numbers to specify users who have access to en-route travel time information. Results show that changes in demand and route assignment create significant variations in the shape of the ground-truth NFD.

The proposed methodology is applied to the first scenario to find the optimal set of detectors and trajectories to estimate the NFD for two different cases. The proportion rate for detectors is considered as 0.6 for both cases ($a = 0.6$), while the proportion rate for OD pairs is 0.2 for the first case ($b = 0.2$) and 0.6 for the second case ($b = 0.6$). The actual NFD for the first scenario and the estimated NFDs by the proposed methodology for both cases are presented in Fig. 8(a) and (d). The OD pairs proportion of 0.2 seems to be as effective as of the OD pairs proportion of 0.6 to estimate the ground-truth NFD. However, the role of the additional proportion rate of OD pairs can be seen in the reliability of the NFD estimation in Fig. 8(b), (c), (e) and (f). In these figures, the selected detectors and subset of trajectories, as the optimal solution based on the ground-truth NFD in the first base scenario, are used to observe the NFD under decreased and increased demand scenarios, respectively. Fig. 8(b) compares the observed NFD with the ground-truth NFD of the decreased demand scenario, where the proportion rates of detectors and OD pairs are 0.6 and 0.2. Fig. 8(e) presents a similar comparison for the decreased demand scenario when the proportion rates of detectors and OD pairs are both equal to 0.6. Fig. 8(c) and Fig. 8(f) present similar comparisons for the increased demand scenario. In both scenarios, the observed NFDs better match with the ground-truth NFD in the second case, Fig. 8(e) and (f), where the proportion rate of the selected trajectories based on the first scenario is higher. This is unlike the similar performances of the two cases in the base scenario and demonstrates the importance of finding reliable optimal solutions. Therefore, we suggest extending the presented methodology to consider stochasticity of the travel demand and supply in a future study.

4.3. Sensitivity to the initial solution

The proposed methodology and the implemented SA algorithm are robust to the initial solution. To further demonstrate this, we use different initial solutions and present the estimated NFDs. In the SA algorithm, the initial solution can affect the number of required iterations to achieve an acceptable optimal solution. The number of iterations is related to the computational efficiency of the algorithm, which was not an issue for the large scale proposed case study. Therefore, for any initial solution with a proper number of iterations, the optimal solution can be achieved. Here, we apply the model for three different initial solutions with link and OD proportions of a = b = 0.6 as an example and compare the results to demonstrate the robustness of the methodology with respect to the initial solution. Note that the SA algorithm avoids local optimal solutions by accepting the worse solutions in the initial outer loop iterations.

The first and second initial solutions here are generated randomly with different seed numbers. The third initial solution is generated by considering a different distribution of link types relative to the other two initial solutions. Instead of considering a uniform distribution of link types, in this solution, all freeways and ramps are selected. The remaining links for the initial solution are selected randomly among the arterial links with a different seed number relative to the other two initial solutions. Results presented in Fig. 9 suggest that having a different initial solution does not significantly change the estimated optimal NFD. The values of the objective function for all three initial solutions and their associated optimal objective function values are shown in Table 1. This shows that although the initial solutions are different, the estimated NFDs in all three cases are comparable.

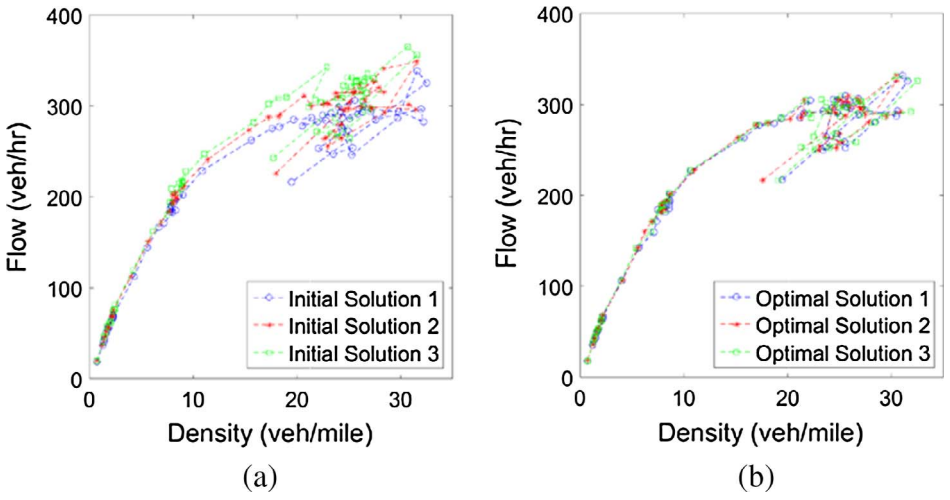

Fig. 9. (a) Initial and (b) optimally estimated NFDs for three different initial solutions.

Table 1
Objective function values of the initial and optimal solutions for three different initial solutions.

	Objective function value	
	Initial	Optimal
Initial solution 1	2074	13
Initial solution 2	6764	22
Initial solution 3	28,506	31

4.4. Sensitivity to the availability of trajectories

Here, we demonstrate that the proposed method is implicitly sensitive to the proportion of trajectories available for each OD pair. The proposed mathematical framework does not optimize the proportions of trajectories for each OD pair as decision variables. Instead, selecting trajectories is based on their availability. Note that not all trajectories may be available for each OD pair. Using a portion of trajectories from each OD pair instead of all trajectories can indeed affect the quality of the space mean speed estimates. Without loss of generality, we assume that all trajectories are available for each selected OD pair. As mentioned earlier, the proposed formulation can be easily modified and extended to cases with less than 100% availability of trajectories. To demonstrate the impact of trajectory availability, we randomly select a fraction of all trajectories for each OD pair in the network (10%, 25%, and 50%) with a fixed link and OD proportion rates of a = 0.6 and b = 0.6, and compare the results with the original case where 100% of trajectories are available (called base case). Note that here, for simplicity, a fixed proportion of trajectory availability is considered for all OD-pairs to demonstrate sensitivity of the method to this parameter. In reality, different proportions might be available for different OD-pairs.

Fig. 10 shows that the difference between the estimated NFDs for all examined cases is very small. The values of the objective function are significantly improved from initial values (see Table 2). This suggests robustness of the method to the proportion of available trajectories. Reducing the fraction of available trajectories has no impact on the number of selected OD pairs and link detectors selected in the estimation process as a = b = 0.6. The optimal objective function value increases with the reduction of the available trajectories, as expected. However, the methodology successfully decreases the objective function value to an acceptable level even for the 10% case. This clearly indicates that the model amends the optimal solution when the availability of trajectories for each OD pair is different. Nevertheless, the estimated NFDs have low errors and good agreement with the ground truth NFD.

In another experiment, we use the optimal solution (configuration of links and OD pairs) of the base case to evaluate the objective function for the cases with 50%, 25% and 10% of trajectories available for each OD pair without running the optimization algorithm. The objective function values for each case are also presented in Table 2. Results suggest that considering a proportion of trajectories for each OD pair does not affect the objective function, while the solutions (selected links and OD pairs) are different. The formulated mathematical problem aims to find the optimal location of fixed measurement points and set of OD pairs in estimating NFD. The model responds to the proportion of available trajectories for each OD pair implicitly and as a result, the optimal solution will differ when different proportions are used. However, the model doesn't explicitly find the optimal proportion of trajectories to be available. We actually showed that the resulting estimated NFD when a lower proportion of trajectories is available may not affect the estimated

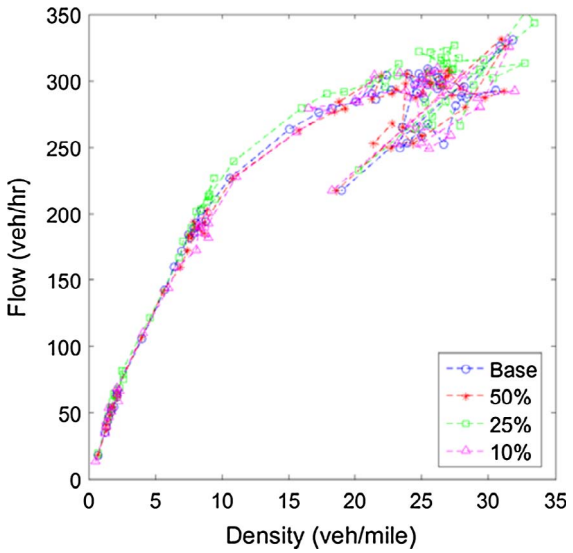

Fig. 10. Optimal estimation of NFD for the base case (all trajectories available), 50% case (50% of trajectories available), 25% case (25% of trajectories available) and 10% case (10% of trajectories available); when a = 0.6 and b = 0.6.

Table 2
Summary statistics of the base, 50%, 25% and 10% cases; when a = b = 0.6.

	Initial value of objective function	Optimal value of objective function	Shared OD Pairs with base case	Shared link detectors with base case
Base case	2074	29	100%	100%
50% case	3307	47	70%	64%
50% case using base solution	1654	–	100%	100%
25% case	17,329	73	69%	64%
25% case using base solution	16,043	–	100%	100%
10% case	129,905	137	67%	65%
10% case using base solution	131,262	–	100%	100%

NFD as long as there is at least one trajectory going through the links with detectors. Having said that, it is obvious that a greater availability of trajectories may result in higher quality and more reliable estimation.

5. Practical insights

Here, we present several insights on the distribution of links and OD pairs for a more accurate estimation of network traffic states (e.g. average network flow and speed). In real world networks, obtaining a ground truth NFD is almost impossible, given the complexity of network geometry, traffic control, turning movements, changing capacities and OD patterns. It is also almost infeasible to have loop detectors on all links and have access to detailed trajectory of every single vehicle in the network. Therefore, a more practical approach is to estimate NFD using a combination of loop detector and probe vehicle data.

We characterize the optimal solution with three measures: (i) distribution of OD pair distances, (ii) distribution of link types (arterial, freeway, and ramps), and (iii) spatial center of the links in the network. We also consider three different initial solutions. In

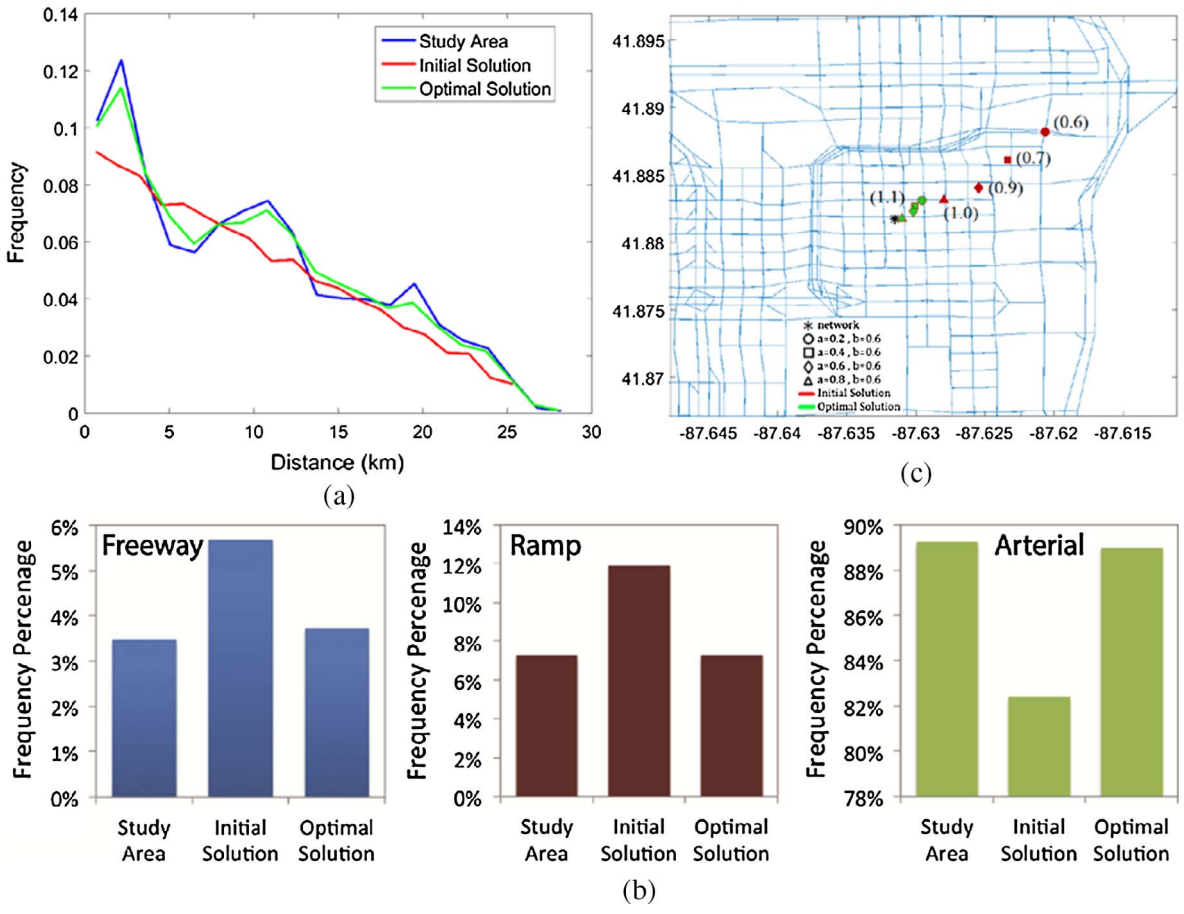

Fig. 11. (a) Normalized frequency of OD pair distances for a = b = 0.6, (b) Frequency percentage of link types for a = b = 0.6, (c) Evolution of the spatial center of the links in the initial and optimal solutions compared to the spatial center.

the first case, initial solution is obtained from a synthesized distribution of the OD pair distances, which is different from the overall distribution in the network (Fig. 11a). In the second case, the initial solution has a different distribution of link types relative to the overall link type distribution across the network (Fig. 11b). Finally, in the third case, initial solution is obtained such that the spatial center of the selected links is located furthest away from the spatial center of all links in the network (Fig. 11c). The spatial center of links is defined as the average of the center point of the links weighted by the link lane-length. This measure represents the extent that selected (optimal) links are spatially distributed. These initial solutions are selected to represent a different pattern relative to the original network. Following, we explore specifications of the optimal solution relative to the initial solutions and network overall specifications.

Fig. 11a shows the normalized frequency of OD pair distances in the network, the synthesized initial solution, and the optimal solution for a fixed link and OD-pair proportions ($a = b = 0.6$). While the frequency of OD pair distances in the initial solution follows an almost triangular distribution, the proposed optimization framework results in an optimal solution that follows the overall distribution of OD pair distances in network more closely. The objective function values for the initial and optimal solution are 2038 and 23, respectively. Results further suggest that the proposed method is robust to the initial solution, and reveal that the optimal OD pairs for estimating NFD should follow a similar distribution of OD pair distances in the network.

A similar pattern can also be observed from Fig. 11b where the frequency percentage of link types in the optimal solution almost matches the frequency percentage in the network while the initial solution follows a completely different distribution. Also, Fig. 11c illustrates the evolution of the spatial center of the links in the initial and optimal solutions compared to the spatial center of the links in the network when $b = 0.6$ and a varies between 0.2 and 0.8. Regardless of the location of the initial solution and parameters a and b, the optimal solution moves towards the spatial center of all links in the network. The numbers next to the center points on the figure show the average distance of the selected links to the spatial center of the network weighted by the link lane-length as a dispersion measure for selected links. This value is equal to 1.1 km for the network and optimal solutions in all cases. This means that the distribution of the selected links (optimal solution) around the spatial center is similar to the distribution of all the links around the spatial center of the network, which is different for the initial solutions.

Overall, results suggest that for an accurate estimation of network density and flow, when combined loop detector and probe vehicle data are used, selected links and OD pairs should be distributed over the network such that it follows a similar spatial pattern to the network original links and OD pairs distribution. A previous study by Courbon and Leclercq (2011) suggested that loop detectors should be uniformly distributed within links across the network. Our findings suggest that the selection of an appropriate set of links with loop detectors and OD pairs are also important. We found that a uniform distribution across the network may not necessarily result in optimal solution. Instead, distribution of links and OD pairs should follow the same distribution of links and OD pairs in the network.

6. Conclusions

This paper presents a mathematical model and solution algorithm to find the optimal location of fixed measurement points and sampling of probe trajectories in a resource allocation problem framework to estimate NFD in a large-scale heterogeneous network with asymmetric demand. The objective is to minimize the discrepancy between the estimated NFD and ground-truth NFD subject to a limited budget for data collection given that the availability of fixed detectors and probe trajectories are not always uniformly distributed across a network. The main contribution of the proposed method is that it does not require any priory known penetration rate for probe trajectories. Data from fixed detectors are used to estimate flows and data from probe trajectories passing through the links with fixed detectors are used to estimate space-mean speed. The problem is formulated as a mixed integer program with non-linear constraints, which is known to be NP-hard including possible local optimal solutions. Therefore, an SA algorithm is used to solve the proposed model using a dynamic network model of Chicago.

In real world networks, obtaining a ground truth NFD is almost impossible given the complexity of network geometry, traffic control, turning movements, changing capacities and OD patterns. It is also almost infeasible to have loop detectors on all links and have access to detailed trajectory of every single vehicle in the network. Therefore, a more practical approach is to empirically estimate NFDs using a combination of loop detector and probe vehicle data. The presented MINLP has two purposes: (a) finding the optimal location of fixed measurement points (link detectors) and OD pairs to estimate an NFD based on a ground-truth NFD which can be achieved either by simulation or any other methods such as the method of cuts, and (b) estimating network traffic state at any time *t*. Even when the ground-truth NFD is known, we still need to estimate network traffic state, or in other words, where on the NFD the network is at any time *t*. Given the budget and traffic monitoring limitations in real word, it is important to identify optimal locations in a network to monitor flows and speeds, so the resulting monitored/estimated NFD is consistent with the ground truth NFD. This has been the main motivation of our study, which is consistent and builds upon previous studies as explained earlier. The optimal location of detectors and trajectories can be estimated based on a surrogate ground truth NFD. Then, incorporating the detectors and trajectory data in real world, one can monitor the actual ground truth NFD instead of the surrogate measure and estimate network traffic state in real time.

The major findings from the numerical experiments are summarized below:

- A combination of fixed detectors and probe vehicles provides sufficient data to estimate NFD with minimal deviation from the ground-truth NFD in a heterogeneous network.
- When the proportion of fixed detection points is constant, increasing the proportion of probe trajectories does not necessarily improve the estimated NFD under the same conditions, however, it might improve reliability of NFD estimation under general

traffic conditions.

- When the proportion of probe trajectories is constant, increasing the proportion of fixed measurement points improves the estimated NFD.
- The optimal locations of fixed measurement points are not only a function of the proportion rate of fixed measurement points; rather, it also changes when the proportion of probe trajectories varies.
- The optimal set of OD pairs is not only a function of the proportion rate of OD pairs; rather, it also changes when the proportion rate of fixed measurements varies.
- A uniform distribution of links and OD pairs across the network may not necessarily result in an optimal solution. Instead, distribution of links and OD pairs, for NFD estimation, should follow the same distribution of all links and OD pairs in the network.

The proposed methodology incorporates the ground-truth NFD as an input to find a subset of links and trajectories to estimate NFD for traffic state monitoring. As the ground-truth NFD may not always be available, a surrogate needs to be obtained analytically or using simulation. Assuming that this surrogate provides the ground-truth NFD, the proposed methodology finds the optimal configuration of links and trajectories for data collection and to estimate the NFD given a limited budget. Here, we utilized a calibrated mesoscopic traffic simulation model of a large-scale network to obtain the surrogate NFD. Estimating NFD under more general conditions and evaluating reliability of a certain solution (selected links and OD pairs) are among possible future research directions. Furthermore, applying the same methodology when a heterogeneous network is clustered into multiple sub-networks with homogenous traffic distribution (Saeedmanesh and Geroliminis, 2016; Lopez et al., 2017) could provide further insights. Studying the impacts of network clustering on number of required fixed measurement points and probe trajectories is another direction for future research.

References

Ambuhl, L., Menendez, M., 2016. Data fusion algorithm for macroscopic fundamental diagram estimation. Transp. Res. Part C 71, 184–197.

Buisson, C., Ladier, C., 2009. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. Transport. Res. Record: J. Transport. Res. Board, No. 2124 127–136.

Cerny, V., 1985. Thermodynamical approach to traveling salesman problem: an efficient simulation algorithm. J. Operat. Theor. Appl. 45 (1), 41–45.

Courbon, T., Leclercq, L., 2011. Cross-comparison of macroscopic fundamental diagram estimation methods. Proc.-Soc. Behav. Sci. 20, 417–426.

Daganzo, C., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transp. Res. Part B 42 (9), 771–781.

Daganzo, C.F., Gayah, V., Gonzales, E., 2011. Macroscopic relations of urban traffic variables: bifurcations, multivaluedness and instability. Transp. Res. Part B 41 (1), 278–288.

Davidovich, L.L., Mikhailovich, L.E., 1980. Statistical Physics, 5 (3). Oxford: Pergamon Press, ISBN 0-7506-3372-7.

Du, J., Rakha, H., Gayah, V., 2015. Deriving macroscopic fundamental diagrams from probe data: issues and proposed solutions. Transp. Res. Part C 66, 136–149.

Gayah, V., Daganzo, C.F., 2011. Clockwise hysteresis loops in the macroscopic fundamental diagram: an effect of network instability. Transp. Res. Part B 45 (4), 643–655.

Gayah, V., Dixit, V., 2013. Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: tests using micro-simulation. Transport. Res. Record: J. Transport. Res. Board, No. 2390, 76–86.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transp. Res. Part B 9, 759–770.

Geroliminis, N., Sun, J., 2011. Properties of a well-defined macroscopic fundamental diagram for urban traffic. Transp. Res. Part B 45 (3), 605–617.

Geroliminis, N., Boyaci, B., 2012. The effect of variability of urban systems characteristics in the network capacity. Transp. Res. Part B 46 (10), 1576–1590.

Geroliminis, N., Haddad, J., Ramezani, M., 2012. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: a model predictive approach. IEEE Trans. Intell. Transp. Syst. 14 (1), 348–359.

Ghamami, M., Zockaie, A., Nie, Y.M., 2016. A general corridor model for designing plug-in electric vehicle charging infrastructure to support intercity travel. Transp. Res. Part C 68, 389–402.

Godfrey, J.W., 1969. The mechanism of a road network. Traffic Eng. Control 11 (7), 323–327.

Haddad, J., Geroliminis, N., 2012. On the stability of traffic perimeter control in two-region urban cities. Transp. Res. Part B 46 (9), 1159–1176.

Haddad, J., Mirkin, B., 2016. Adaptive perimeter traffic control of urban road networks based on MFD model with time delays. Int. J. Robust Nonlinear Control.

Hejazi, B., 1999. Facility location of bus terminals in urban areas using simulated annealing algorithm. Master Thesis. Sharif University of Technology.

Ji, Y., Daamen, W., Hoogendoorn, S., Hoogendoorn-Lanser, S., 2010. Investigating the shape of the macroscopic fundamental diagram using simulation data. Transport. Res. Record: J. Transport. Res. Board 2161, 40–48.

Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., Papagerogiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. Transp. Res. Part B 46 (10), 1393–1403.

Kirkpatrik, S., Gelaltt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Knoop, V., Van Lint, J.W.C., Hoogendoorn, S., 2012. Routing strategies based on the macroscopic fundamental diagram. Transport. Res. Record: J. Transport. Res. Board 2315, 1–10.

Leclercq, L., Geroliminis, N., 2013. Estimating MFDs in simple networks with route choice. Transp. Res. Part B 57, 468–484.

Leclercq, L., Chiabaut, N., Trinquier, B., 2014. Macroscopic fundamental diagrams: a cross-comparison of estimation methods. Transp. Res. Part B 62, 1–12.

Lopez, C., Krishnakumari, P., Leclercq, L., Chiabaut, N., Van Lint, H., 2017. Spatio-temporal partitioning of transportation network using travel time data. Transport. Res. Record: J. Transport. Res. Board http://dx.doi.org/10.3141/2623-11 (In press).

Mazloumian, A., Geroliminis, N., Helbing, D., 2010. The spatial variability of vehicle densities as determinant of urban network capacity. Philos. Trans. Roy. Soc. A 368, 4627–4647.

Mahmassani, H., Williams, J.C., Herman, R., 1984. Investigation of network-level traffic flow relationships: some simulation results. Transport. Res. Record, No. 971, 121–130.

Mahmassani, H.S., Williams, J.C., Herman, R., 1987. Performance of urban traffic networks. In: Proceedings of the 10th International Symposium on Transportation and Traffic Theory. Elsevier Science Publishing, pp. 1–20.

Mahmassani, H.S., Saberi, M., Zockaie, A., 2013. Urban network gridlock: theory, characteristics, and dynamics. Transp. Res. Part C 36, 480–497.

Mariotte, G., Leclercq, L., Laval, J.A., 2017. Macroscopic urban dynamics: analytical and numerical comparisons of existing models. Transp. Res. Part B 101, 245–267.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., 1953. Equation of state calculation by fast computing machines. J. Chem. Phys. 21 (6).

Nagle, A., Gayah, V., 2014. Accuracy of networkwide traffic states estimated from mobile probe data. Transport. Res. Record: J. Transport. Res. Board, No. 2421, pp. 1–11.

Ortigosa, J., Menendez, M., Tapia, H., 2014. Study on the number and location of measurement points for an MFD perimeter control scheme: a case study of Zurich. EURO J. Transport. Logist. 3, 245–266.

Ramezani, M., Haddad, J., Geroliminis, N., 2015. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. Transp. Res. Part B 74, 1–19.

Saberi, M, Mahmassani, H.S., 2012. Exploring the properties of network-wide flow-density relations in freeway networks. Transport. Res. Record: J. Transport. Res. Board, Transportation Research Board of the National Academies, Washington, D.C., 2315, 153–163.

Saberi, M., Mahmassani, H., Hou, T., Zockaie, A., 2014. Estimating network fundamental diagram using three-dimensional vehicle trajectories: extending Edie's definitions of traffic flow variables to networks. Transport. Res. Record: J. Transport. Res. Board, No. 2422, 12–20.

Saberi, M., Mahmassani, H.S., 2013. Hysteresis and capacity drop phenomena in freeway networks: empirical characterization and interpretation. Transport. Res. Record: J. Transport. Res. Board 2391, 44–55.

Saeedmanesh, M., Geroliminis, N., 2016. Clustering of heterogeneous networks with directional flows based on "Snake" similarities. Transp. Res. Part B 91, 250–269.

Van Laarhoven, P.J., Aarts, E.H.L., 1987. Simulated Annealing: Theory and Applications. D. Reidel Publishing-Company, Dordrecht, The Netherlands.

Yildirimoglu, M., Ramezani, M., Geroliminis, N., 2015. Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. Transport. Res. Procedia 9, 185–204.

Zheng, N., Waraich, R., Axhausen, K., Geroliminis, N., 2012. A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model. Transp. Res. Part A 46 (8), 1291–1303.

Zockaie, A., Chen, Y., Mahmassani, H.S., 2014a. Adaptive drivers and time-dependent origin-destination demand estimation: methodology and application to large-scale network. In: Presented in the 93rd annual meeting of Transportation Research Board of the National Academies, Washington, D.C.

Zockaie, A., Mahmassani, H., Saberi, M., Verbas, O., 2014b. Dynamics of urban network traffic flow during a large-scale evacuation. Transport. Res. Record: J. Transport. Res. Board 2422, 21–33.

Zockaie, A., Aashtiani, H.Z., Ghamami, M., Nie, Y., 2016. Solving detour-based fuel stations location problems. Comput.-Aid. Civil Infrastruct. Eng. 31 (2), 132–144.