ASSESSING THE PERFORMANCE OF A TWO-STEPS DYNAMIC DEMAND ESTIMATION MODEL ON LARGE SCALE CONGESTED NETWORKS

Guido Cantelmo
Faculty of Science, Technology and Communication
Research Unit in Engineering Science
University of Luxembourg
Rue Coudenhove-Kalergi 6, L1359 Luxembourg city, Luxembourg;
Email: guido.cantelmo@uni.lu

Francesco Viti
Faculty of Science, Technology and Communication
Research Unit in Engineering Science
University of Luxembourg
Rue Coudenhove-Kalergi 6, L1359 Luxembourg city, Luxembourg
Email: francesco.viti@uni.lu

Word count: 6750 words text + 3 tables/figures x 250 words (each) = 7500 words

Submission Date: 13/07/2017

ABSTRACT

This paper assesses the performance of the Two-Steps approach on a large-scale network. This model, recently proposed by Cantelmo et al. [1], sequentially calibrates generations and distribution values of the dynamic demand matrix. The authors initially applied this model to a simple motorway, showing its capability of producing more reliable results. This paper moves one-step forward, by applying this methodology with real data on the network of the Grand Duchy of Luxembourg.

Traffic counts and average speeds have been used to compare results obtained through the proposed methodology with the ones obtained by using a standard bi-level formulation. Results show how the proposed model outperforms the standard ones, as breaking the optimisation process in two parts strongly reduces the localism of the problem. The contribution of this paper is twofold. First, we show that the Two-Step approach results are less impacted by the choice of the initial seed matrix with respect to classical. As the model reduces the number of variables, the overall reliability of the model increases. Second, as the model assumes a linear relation between time-dependent distributions and generated traffic volumes, we show that the objective function explicitly accounts for the structure of the historical o-d flows, capturing congestion dynamics at a network level and reducing the overfitting of data issue.

To support the claim that the model is practice ready, a Matlab package has been developed to interface the proposed framework with PTV Visum, a commercial software for static and dynamic traffic analysis.

Keywords: O-D estimation, Two-Steps approach, Dynamic Traffic Assignment, SPSA.

INTRODUCTION

Dynamic Traffic Assignment (DTA) models represent essential tools for managing transportation systems. DTA models take as input the demand from each origin and destination and at each time period, and in turn estimate and/or predict route and link flows on transportation networks.

In order to generate the mobility demand, usually represented in the form of Origin-Destination (OD) matrices, traditional approaches combine survey data and mathematical tools [2]. Additionaly, more recent works have done a significant progress into including new data sources, such as Call Detail Records (CDR), GSM data, sensing data and geospatial data [3]. Unfortunately, the estimated demand matrix is at most a concise representation of the systematic component of the demand – such as the typical behaviour during a working day. However, daily demand patterns can substantially differ from the systematic ones because of several elements, including weather conditions or road works, as well as because of the inherent stochasticity of the travel demand. Deviations between estimated and actual demand patterns can be mitigated by using traffic measurements, which can be used to update an existing (a-priori) OD matrix. This problem, which is known in the literature as the Dynamic Origin-Destination Estimation (DODE) problem, exploits a properly specified objective function for estimating the time-dependent OD flows.

While the DODE problem has been initially treated as an extension of its static counterpart [5-6], the last decades have witnessed to a considerable effort by researchers in order to develop methodologies able to deal with the dynamic case [6]. As DTA models are applied in both *offline* (medium-long term planning and design) and *online* (real-time management) contexts, DODE is commonly classified between sequential or simultaneous approaches, where usually the first is adopted for *online* while the second for *offline* applications [6]. By limiting our focus to the *offline* case, DODE is usually formulated as a bi-level optimisation problem. In the upper level, OD flows are updated by minimising the error between simulated and observed traffic data, while in the lower level the DTA solves the combined Route Choice (RC) and Dynamic Network Loading (DNL) problems [7].

Earlier DODE models explicitly accounted for the assignment matrix – i.e. the set of rules linking OD and link flows - for updating the demand vector. However, this matrix assumes a linear relation between demand and supply parameters, assumption that does not hold for congested networks [8]. In order to overcome this issue, Balakrishna et al. [9] proposed a bi-level formulation that does not rely on this information. Instead, the authors suggested using a simulation-based DTA model to generate traffic measures and to include additional information, such as link speed, within the objective function, in order to represent the congested/uncongested network conditions. Following this seminal work, many researchers developed new and more robust assignment-free algorithms able to properly capture the non-linearity between link-flow propagation and time-varying OD demand [10]–[13]. Despite this intense effort, the resulting optimisation problem remains highly non-linear and non-convex. To reduce the number of possible solutions, classical methods often include information about a reference OD demand matrix (usually known as historical or "seed" matrix) within the objective function. Therefore, if the structure of this seed matrix is different from the real one, this localism can lead to substantial errors [8].

Recently, Marzano et al. [14] pointed out that DODE is generally unable to provide an effective estimation when the ratio between unknown and known variables (OD flows and traffic measurements, respectively) is greater than one. Hence, the easiest solution is to reformulate the objective function in order to reduce the number of variables. This can be done, for instance, by using Principal Component Analysis (PCA) [15]. Alternatively, Cascetta et al. [16] introduced the so-called "quasi-dynamic assumption", which assumes that the generated demand for a certain OD pair is time dependent, while its spatial distribution is constant. Under this assumption, as demonstrated in [16], the DODE problem becomes less underdetermined and more likely to find more robust results. Nevertheless, the authors point out that the resulting matrix will be "intrinsically biased" since this assumption introduces an "intrinsic error". Similarly, Cantelmo et al. [1] proposed a Two-Steps procedure, which separates the DODE in two sub-optimization problems. The first step searches for generation values that best fit the traffic data while keeping spatial and temporal distributions constant. In the second step, the standard bi-level procedure searches for a more reliable demand matrix.

Although the model has been tested on a simple network, the Two-Step approach has three characteristics that make it an ideal candidate for applications on large-scale networks. First, as pointed out by Antoniou et al. [17], the starting matrix is still a key input for all state-of-the-art DODE models. The first step of this formulation improves the historical demand matrix by performing a broad evaluation of the solution space and estimating a "good" updated seed matrix to be used in the second step. Secondly, the proposed model reduces the number of variables in the first step, increasing the overall reliability of the results [1], [14]. On this point, the idea of performing successive iterations and linearizations has been already introduced and validated in [18] for the online DODE, showing that the reliability of the results generally increases.

Driven by these considerations, the contribution of this paper is twofold. First, we apply the Two-Steps approach to the real metropolitan network of Luxembourg. While the previous study [1] tested the algorithm on a simple motorway, this paper shows that the Two-Steps approach outperforms the standard formulation on a real-life network. We numerically demonstrate that properties of robustness and reliability hold for a general network, and that the localism of the model strongly decreases. The test-network represents most of the country of Luxembourg, including urban roads, motorways and primary roads. Real traffic counts extracted from loop detectors are used within the calibration process to update the demand.

Second, as speed profiles on the counting stations were not available, we extend the objective function by including the average speeds over the analysis period, which have been calculated through Floating Car Data (FCD). We show that, when combined with a standard DODE procedure, this information leads to a poor calibration of the demand, as the DODE overfits the data within the objective function. However, as the Two-Steps approach over-impose a linear relation between distribution and generation for a certain traffic zone, it is more likely to capture congestion dynamics at network level, such as the systematic overestimation or underestimating of the demand, thus to avoid this issue.

Finally, to support the claim that the model is ready for practical implementation, it is interfaced with PTV-Visum, one of the most widely adopted software tools for traffic analysis [19].

The paper is structured as follows. The next paragraph defines the methodology, including the "conventional" model (called Single-Step OD estimation in the rest of this paper) and the proposed Two-Steps approach. The paper then describes the case study, including the network, the dataset used for the experiments and the results. Finally, in the last section conclusions are drawn.

METHODOLOGY

The DODE is usually formulated as a constrained optimisation problem, which requires the formulation of:

i. An objective function, which is composed of variables and constraints related to routeing conditions and behavioural assumptions;
ii. An optimisation method, which can be classified in Path Search, Pattern Search or Random Search approaches [9];
iii. A parameter updating rule;

The remainder of this section describes the set of functions and algorithms used in the proposed paper for performing Single-Step and Two-Steps demand estimation.

Objective function

Simultaneous GLS Estimator

The most widely adopted goal function for solving the *offline* DODE is the Generalized Least Squared (GLS) proposed in [5]. Considering different types of measures and a *simultaneous* approach, the problem can be formulated as:

$$(\boldsymbol{d}_1^*, \dots, \boldsymbol{d}_n^*) = argmin \begin{bmatrix} z_1(\boldsymbol{l}_1, \dots, \boldsymbol{l}_n, \widehat{\boldsymbol{l}_1}, \dots, \widehat{\boldsymbol{l}_n}) + \\ +z_2(\boldsymbol{n}_1, \dots, \boldsymbol{n}_n, \widehat{\boldsymbol{n}_1}, \dots, \widehat{\boldsymbol{n}_n}) + \\ +z_3(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \widehat{\boldsymbol{x}_1}, \dots, \widehat{\boldsymbol{x}_n}) + \\ +z_4(\boldsymbol{r}_1, \dots, \boldsymbol{r}_n, \widehat{\boldsymbol{r}_1}, \dots, \widehat{\boldsymbol{r}_n}) + \end{bmatrix} \tag{1a}$$

Where $\mathbf{l}/\hat{\mathbf{l}}$ represent, respectively, simulated and measured link performances, $\mathbf{n}/\hat{\mathbf{n}}$ calibrated and observed values on the node, $\mathbf{x}/\hat{\mathbf{x}}$ indicate the estimated and historical value for the OD flows (seed matrix) and $\mathbf{r}/\hat{\mathbf{r}}$ the simulated and observed route performances. Finally, $\boldsymbol{d}_{\boldsymbol{n}}^*$ designates the estimated demand matrix for time interval n, while $z\colon\{z_1, z_2, z_3, z_4\}$ is the estimator of the error between simulated/estimated and measured/a priori values.

The dependence between supply and demand in Equation (1a) is obtained directly by simulation performing a dynamic traffic assignment (DTA), so that:

$$\begin{aligned} \boldsymbol{l}_1, \dots, \boldsymbol{l}_n &= \mathbf{F}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \boldsymbol{n}_1, \dots, \boldsymbol{n}_n &= \mathbf{F}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \boldsymbol{r}_1, \dots, \boldsymbol{r}_n &= \mathbf{F}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{aligned} \tag{1b}$$

with function **F** representing the Dynamic Traffic Assignment (DTA) function. The objective function presented in Equation (1a) presents a series of agreeable properties that make it an ideal candidate for assignment-matrix free algorithms. First, apart from the traffic counts, the function may account for different sources of information, such as link speeds and densities – which have been proved to capture the non-linear relation between demand and supply parameters [9], [20]. Moreover, recent works showed how more elaborate information, such as point-to-point data, can also be included in this function, largely improving the overall estimation accuracy [17], [21], [22]. An additional advantage of the simultaneous GLS presented in Equation (1a) with respect to the sequential case is that all variables are jointly estimated, which is formally more correct as OD flows over different time intervals are likely to be correlated [8]. However, for large networks, this approach becomes less reliable and, if not enough traffic data is available [14], the sequential approach is preferred.

Strict Quasi-Dynamic Simultaneous Generalized Least Squared

As suggested in [16], the objective function described in equation (1) can be enhanced by exploiting information on aggregated socio-demographic data such as generation data by traffic zones. The objective function (1a) can be then reformulated as:

$$(\boldsymbol{E}_1^*, \dots, \boldsymbol{E}_n^*) = argmin \begin{bmatrix} z_1'(\boldsymbol{l}_1, \dots, \boldsymbol{l}_n, \widehat{\boldsymbol{l}_1}, \dots, \widehat{\boldsymbol{l}_n}) + \\ +z_2'(\boldsymbol{n}_1, \dots, \boldsymbol{n}_n, \widehat{\boldsymbol{n}_1}, \dots, \widehat{\boldsymbol{n}_n}) + \\ +z_3'(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \widehat{\boldsymbol{x}_1}, \dots, \widehat{\boldsymbol{x}_n}) + \\ +z_4'(\boldsymbol{r}_1, \dots, \boldsymbol{r}_n, \widehat{\boldsymbol{r}_1}, \dots, \widehat{\boldsymbol{r}_n}) + \end{bmatrix} \tag{2a}$$

subject to

$$x_n^{OD} = E_n^O d_{D|O}^{Seed,n} \qquad \forall O, \forall D, \forall n \tag{2b}$$

Where:

- $E_n{}^O$ = generated flow from traffic zone O and time interval n;
- $\boldsymbol{E}_n^*$ = generation vector containing the generated flow from all zones in time interval n.
- x_n^{OD} = demand flow from origin zone O to destination zone D in time interval n.
- $d_{D|O}^{Seed,n}$ = seed matrix spatial/temporal distribution to move in traffic zone D from traffic zone O in time interval n.

Constraint (2b) is the main difference between the general quasi-dynamic formulation proposed in [16] and the one proposed in Equation (2). The former explicitly considers a probability function that captures the correlation between generation and distribution over a certain sub-period of time. As a consequence, $d_{D|O}^{Seed,n}$ is updated during the optimization process. Instead, constraint (2b) assumes a constant value of the distribution, resulting in a smoother objective function. Equation (2b) presents two major advantages with respect to the standard GLS. First, as the number of unknown variables strongly decreases, the simultaneous approach can be applied to larger networks. Second, this approach does not necessarily require to explicitly account for historical OD flows within the objective function. As pointed out in the introduction, historical OD flows are usually included within equation (1a) in order to reduce the number of possible solutions. However, this information is already considered within constraint (2b), that over-impose, to the estimated matrix, the spatial/temporal structure of the historical demand. However, a main drawback of this formulation is that it is likely to provide a poor fit of the traffic data with respect to equation (1) or to the general quasi-dynamic formulation, as pointed out in [1]. Thus, it is an ideal candidate for being used in the first phase of the Two-Step approach, where the main purpose is to have a broad evaluation of the solution space, rather than to best fit the observations.

Optimisation method: SPSA

The optimisation method adopted in this paper is the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm proposed in [23]. While we adopt the original model in this paper, as has been proven to be very effective for tacling the DODE problem, many authors proposed enhanced versions [11]–[13], which can also be combined with the proposed framework. The SPSA is a stochastic approximation of the deterministic finite difference gradient method, which has been proved to be very effective for tackling the DODE, but becomes computationally too expensive for large networks [10]. By assuming a one-sided perturbation [11], the SPSA computes the approximated gradient $\boldsymbol{G^i}$ at each iteration –*i* as:

$$\widehat{\boldsymbol{g}}_k(\boldsymbol{\theta}^i) = \frac{z(\boldsymbol{\theta}^i + c^i\Delta^k) - z(\boldsymbol{\theta}^i)}{c^i}\begin{bmatrix}(\Delta_1^k)\\ \vdots \\ (\Delta_r^k)\end{bmatrix} \quad (3a)$$

$$\boldsymbol{G^i} = \overline{\boldsymbol{g}}(\boldsymbol{\theta}^i) = \frac{\sum_{\boldsymbol{k=1}}^{\boldsymbol{Grad\_rep}} \widehat{\boldsymbol{g}}_k(\boldsymbol{\theta}^i)}{Grad\_rep} \quad (3b)$$

With $\boldsymbol{\theta}^i$ the vector with the estimated variables, $z(\boldsymbol{\theta}^i)$ the objective function value in $\boldsymbol{\theta}^i$, c^i the perturbation step, *Grad_rep* the number of replications to compute the average gradient and Δ is a vector with elements {-1,1}. Given the stochastic nature of the model, it is recommended to repeat the perturbation multiple times in order to obtain a good approximation. If only one replication is used, then $\boldsymbol{G^i} = \widehat{\boldsymbol{g}}_k$. In Equation (3a), the asymmetric design (SPSA-AD) model is showed. The main advantage of using this formulation is that it allows to reduce the number of simulations needed while still providing a proper approximation of the gradient [11].

Parameter updating rule

Given a properly specified objective function and a descent direction – the gradient $\boldsymbol{G}^i$– the parameters are updated at each iteration according to:

$$\boldsymbol{\theta}^{i+1} = \boldsymbol{\theta}^i - \alpha^i\boldsymbol{G}^i \quad (4)$$

Where α^i is the stepsize and $\boldsymbol{\theta}^i$ is again the vector of parameters to be updated, the OD or the Generation flows if we are minimizing, respectively, objective functions (1) or (2). Concerning the value of α^i, we proposed to use a line search to find the optimal value in order to reduce the overall computational time.

Single-Step and Two-Steps approach

The Single-Step OD estimation is formulated in this paper as a single constrained optimisation problem, which minimises Equation (1) according to a certain optimisation method, the SPSA, and the parameter updating rule showed in Equation (4). Results of this, quite general, optimisation framework depend on the overall quality of the initial seed matrix [17]. While a more elaborate algorithm may improve the performances of the standard SPSA when applied to large networks, this critical element still remains [12], [17]. The main contribution of breaking the optimisation problem in two phases is to relax this strong limitation.

The proposed Two-Steps approach combines the set of rules, functions and algorithms described in the previous sub-sections. Specifically, the first step minimises Equation (2) in order to optimise the generated demand flows for each zone in each time interval. Hence, in the first phase, the variables are the total generated demand flows, which reduces the dimension of the problem considerably. In the second step, the classical DODE procedure is performed by minimising Equation (1), improving temporal and spatial matrix distributions. Breaking the problem as such, one benefits of the right demand level identified in the first phase. As the objective function presented in (2) reduces the number of variables used, it becomes less sensitive to the network size. Thus, the estimated matrix can be used in the second level of the Two-Step approach, where the optimisation can exploit a better initial point in order to achieve overall better results.

The idea of updating the generation in the first step derives from the increasing attention received by this type of aggregated information in the literature [11], [16]. This high significance derives mainly by the following considerations:

- Total generated trips can limit a demand overestimation during the DODE, which is otherwise likely to occur when dealing with congested networks;
- As generation models are considered the most reliable models in transport engineering applications, total generated trips are more easily observable than OD trips;
- Adopting the generation values inside the DODE, as in (2), reduces the number of variables.

The goal of the first step is to act on the seed matrix in order to obtain a "right level of demand", then moving to the second step in order to optimise the dynamic distributions OD trips as in (1).

CASE STUDY: LUXEMBOURG

We now test both approaches to the real large-sized network of Luxembourg, showed in Figure (1). The Grand Duchy of Luxembourg is a small country placed in the heart of Europe, bordered by Belgium to the west, Germany to the east and France to the south. As most of the activities are located in the capital, Luxembourg City, the country is facing mobility challenges, which are being made worse by the 170.000 workers - about 43% of the commuting demand [24] - coming every day to Luxembourg City from the neighbouring countries.

The main goal of the case study proposed in this section is to model the complex interaction between the cross-borders – commuters coming from France, Germany and Belgium – and the road users living within the Grand Duchy's borders. This latter demand segment can be further divided into people living in the capital and people living in the countryside, where the second one is the predominant component of the commuting demand. The ring of Luxembourg City represents the bottleneck for this system, as its capacity is not sufficient to properly serve the high volume of demand moving to the capital during the rush hour, hence major congestion patterns are reported every day.

The network, showed in Figure (1), includes all national motorways, which go from the city of Ettelbruck to Luxembourg City in the north, and from the capital to the east, west and south borders. Additionally, the network includes also primary and secondary roads, as they are commonly used by commuters.

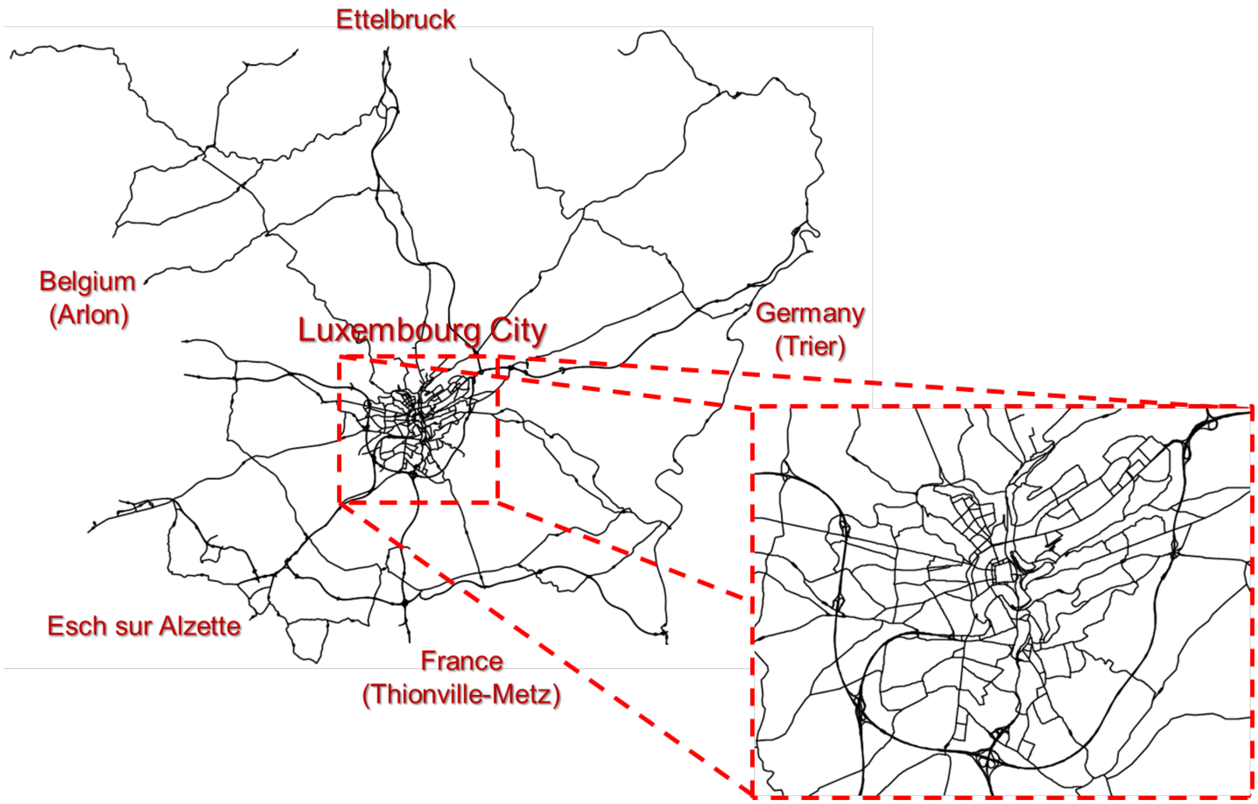

Figure 1: The Grand Duchy of Luxembourg network, with detail of Luxembourg City

Description of the Data

As part of the initiative "Digital Luxembourg", the Grand Duchy is developing a new open-data portal (https://data.public.lu), which gathers different sources of information including socio-demographic data. These data, collected by the National Institute of Statistics (STATEC), include the growth of the population for each year, the population for each canton and the number of cross-borders. Based on these statistics, a static matrix for the morning commute has been estimated through the classical Four-Step demand generation model. A departure time choice model based on the Vickrey/Small [25] formulation has been then used to derive a dynamic OD matrix from the static one. This dynamic matrix accounts for 46 traffic zones and represents the historical demand (Seed Matrix) for the experiments presented in the next sub-sections.

Concerning the supply side, the Luxembourgish Road Administration agency collects and provides traffic counts on most of the motorways and primary roads of the Grand Duchy. Unfortunately, these data present two major limitations. The first main limitation is that, based on the publicly available data, only three detectors are located inside the ring of Luxembourg. This means that we can expect to have a realistic representation of the demand on the regional network and on the ring, but it is not possible to validate the estimated solution inside the city. The second problem concerns the time interval aggregation for these data, as traffic counts are aggregated on an hourly basis. This time interval is clearly too large for a network with an average free-flow travel time of 20 minutes since basic congestion dynamics could not be properly captured. Finally, neither the open-data portal nor the Luxembourgish Road

Administration provides information on the speeds, which are an essential input when dealing with large congested networks such as the one proposed in Figure (1).

To compensate this lack of information, average speeds on the ring, which have been calculated by using Floating Car Data (FCD), have also been considered within the estimation process. The obtained information is based on the average of all available information and does not contain specifications about time and location.

FCD carry definitely more information than just the average speeds, as demonstrated in [26], but privacy laws do not allow sharing sensible data in Europe. Thus, the available average speed broadly captures, in this study, the congestion on the ringway at a network level. The downside is that many possible solutions exist, which can create congestion on the ring. As a consequence, the most logical solution for the DODE should be to keep the demand as close as possible to the historical demand, while at the same time reproducing the speed profile. However, as this information is strongly aggregate, the Single-Step approach has the tendency to over-fit the average speed, while the Two-Step approach manage to provide more reliable results by exploiting the Link-Flows as a constraint within the objective function. This claim is numerically illustrated in the next section.

Experiment Setup

The network introduced in the previous section consists of 3700 links and 1469 nodes. Luxembourg City, located in the heart of the system, represents the typical middle-sized European city in terms of network dimension and has the typical structure of a metropolitan area, composed of a city centre, the ring, and suburb areas. Considering the speed profile and that the infrastructure is composed of highways, primary roads and urban roads, we can classify this system as a large-sized heavily congested network. In this study, we consider the morning peak between 5 AM and noon (8 hours). After some data cleaning, 54 counting stations have been retained, all located on the main arterial roads going to Luxembourg City and on the ring. The seed-matrix accounts for 307.544 trips and 16928 time dependent OD pairs. Both traffic counts and the average speed are included in the objective function, where the Root Mean Squared Error (RMSE) is the chosen estimator $z: \{z_1, z_2\}$**:**

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{q}_i - q_i)^2}{N}} \tag{4}$$

Where N is the number of observations, $\hat{q}_i$ is the observed value for the measured data and q_i is the simulated one.

Finally, to be able to solve the DODE on the network of Luxembourg, we developed a Matlab package for solving the dynamic O-D estimation using PTV Visum as DTA model. The package, named MAMBA-DEV, allows performing assignment-free dynamic or static OD estimation, using a deterministic and/or stochastic approximation of the gradient. The package also includes the Two-Step approach discussed in this paper. While the MAMBA-DEV package has been designed for Luxembourg, it can work with any network in Visum, supporting the idea that the model is ready for practical implementation.

Results

Comparison between Single-Step and Two-Steps approach

The first experiment proposed in this section aims to numerically validate two properties of the Two-Steps approach formulated in the methodology section:

i. The Two-Steps approach outperforms the standard one on big sized networks;
ii. The first step is likely to find a good initial point to be updated through the objective function presented in Equation (1);

The starting point of this experiment is not a "good initial point", as it derives from a static matrix and has not previously been calibrated. The initial matrix provides in fact a rather poor fit with the traffic counts ($r^2 = 0.2686$ and $RMSE_{link-flows} = 452.61\,Veh/h$). In order to reduce this error, weights have been considered so that the traffic counts are responsible for 70% of the overall error within the objective function, while the average speed is responsible for the remaining 30%, thus a relatively poor representation of the average speed is expected.

As showed in Figure (2), the Two-Steps approach clearly outperforms the Single-Step in terms of estimation results, as the latter just collapses on the closest local minima. While the model reduces error on the traffic counts ($r^2 = 0.3194$), these results are far from being acceptable for any practical application ($RMSE_{Link-Flow}^{Single-Step} = 438.89\,Veh/h$).

By contrast, results from the Two-Steps approach seem more reasonable and similar to the expectations ($r^2 = 0.7097,\ RMSE_{Link-Flow}^{Two-Step} = 241.31\,Veh/h$). During the first phase, the model exploits Equation (2) to explore the solution space by updating only the generations. After finding a local minimum, the model switches to Equation (1) in order to find the best fit with the observations.

It should be pointed out that the second step of the model is basically adopting the same algorithm as the Single-Step approach. The only difference is the starting point, which has been updated during the first step of the algorithm. While this framework collapsed in a few iterations when coupled with the historical seed-matrix, exploiting the more reliable demand matrix estimated through Equation (2) gives a relevant contribution to the overall optimisation, stressing how both phases of the Two-Steps approach are complementary and, thus, necessary.

Figure (2e) depicts the Spider Chart plot of the estimation error for speeds, flows and seed-matrix – i.e. the initial point. For each measure, this relative error has been calculated as:

$$Rel_{error} = \frac{RMSE^{Measures}}{\max\left(RMSE^{Two-Step}, RMSE^{Single-Step}\right)} \tag{5}$$

Figure (2e) intuitively shows the dynamics behind the optimization. The Single-Step approach does not manage to move from the initial point, thus to reduce the error on the Link Flows. As the Two-Steps approach moves to a new solution during the first phase of the optimization, the distance with respect to the initial matrix is larger, while the error on the link flows is two times smaller than the one for the Two-Steps. However, the Two-Steps also increases the error on the speeds, which was expected as this information has a low weight in the goal function. Thus, in the section we introduce a second experiment, which aims at finding a consistent solution for both counts and speeds.

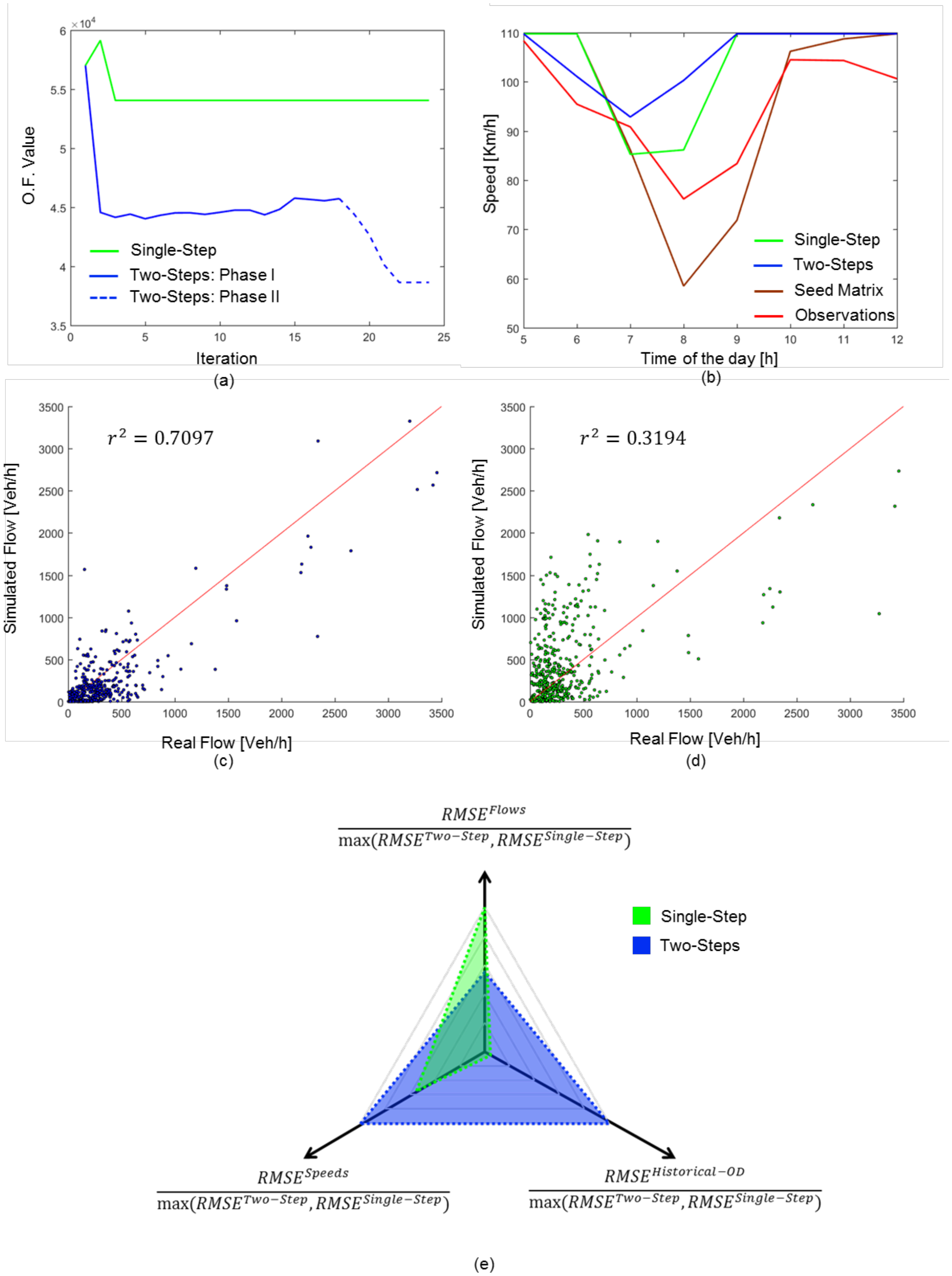

Figure 2: (a) Objective Function trend; (b) Estimated and Observed Average Speed; (c) Scatter Estimated and Observed Link Flows for the Two-Steps; (d) Scatter Estimated and Observed Link Flows for the Single-Steps; (e) Spider Chart of the relative error for the estimated matrix in terms of Link-Flows, Distance from the Historical OD flows (Seed Matrix) and Average Speed;

Improving the results obtained: Good Starting Matrix

The second experiment presented in this section aims at demonstrating that, even when a "good" a prior demand matrix is available, the Single-Step approach is more likely to over-fit the data with respect to the proposed methodology. Results illustrated in Figure (2) show how using the Two-Steps approach reduces the localism of the standard single-step DODE, relaxing the dependency on the starting matrix. However, although the model outperformed the Single-Step formulation, the overall estimation is still unsatisfactory. While the model largely reduced the error on the link flows, increasing the r^2 from 0.2686 to 0.7097, the estimated OD matrix significantly underestimates the congestion on the ring. Thus, we performed a second experiment to correct this error. The OD matrix obtained through the Two-Steps approach in the previous estimation is now used as initial point for this second experiment, simulating the situation for which a "good" a priori OD matrix is available. The objective function still accounts for both traffic counts and average speed, but this time the latter accounts for 70% of the error, while former are mostly used as a constraint to reduce the search space, avoiding the model to move too far from the current solution.

Results, shown in Figure (3), prove that both Two-Steps and Single-Step methods estimate a reasonable approximation of the congestion pattern. While congestion between 8 AM and 9 AM is still slightly underrepresented, the average speed on the ring seems more realistic, as the congestion period begins and terminates approximately at the same time for both models. However, the Single-Step clearly approximates the average speed on the ring better than the Two-Step approach. By contrast, the error on the link flows clearly shows that the Single-Step is overfitting the speed data, which was expected given the aggregate nature of this information, while strongly increasing the error with respect to the link flows ($RMSE_{Link-Flow}^{Single-Step} = 338.61\ Veh/h$).

Instead, the Two-Steps approach manages to provide a realistic fitting for both traffic counts and speed. Although the error on the Link Flows increases with respect to the starting point ($RMSE_{Link-Flow}^{Two-Step} = 291.10\ Veh/h$), the difference is not as big as for the Single-Step approach, as the r^2 shows in Figure (3). This brings to a second important consideration. In this second experiment, no improvement is observed in the second step of the Two-Steps approach.

Constraint (2b) imposes a linear relation between temporal and spatial distribution, meaning that the spatial and temporal structure of the demand is constant during the first step of the optimisation. The direct consequence of that is that the matrix estimated through Experiment II keeps the same structure as the one obtained through Experiment I, while the total demand is different. Although the real OD matrix is not available, as we are dealing with real traffic information, we can easily calculate the error in terms of Euclidean distance with respect the initial matrix, as we would like to keep the distance with respect to the "good" historical OD flows as small as possible. While the Euclidean distance between the estimated matrix and the initial one is only 718 trips for the Two-Step approach, this error increases up to 6449 trips when using the Single-Step approach as optimization framework.

In essence, we may argue that the Two-Steps approach kept the structure of the demand from the Seed-Matrix, but sensed and increased the demand in order to move the traffic state on the ring from the uncongested to the congested branch of the fundamental diagram. This suggests that the Two-Steps approach is more likely to exploit aggregate data, without altering the structure of the demand in order to overfit the available data.

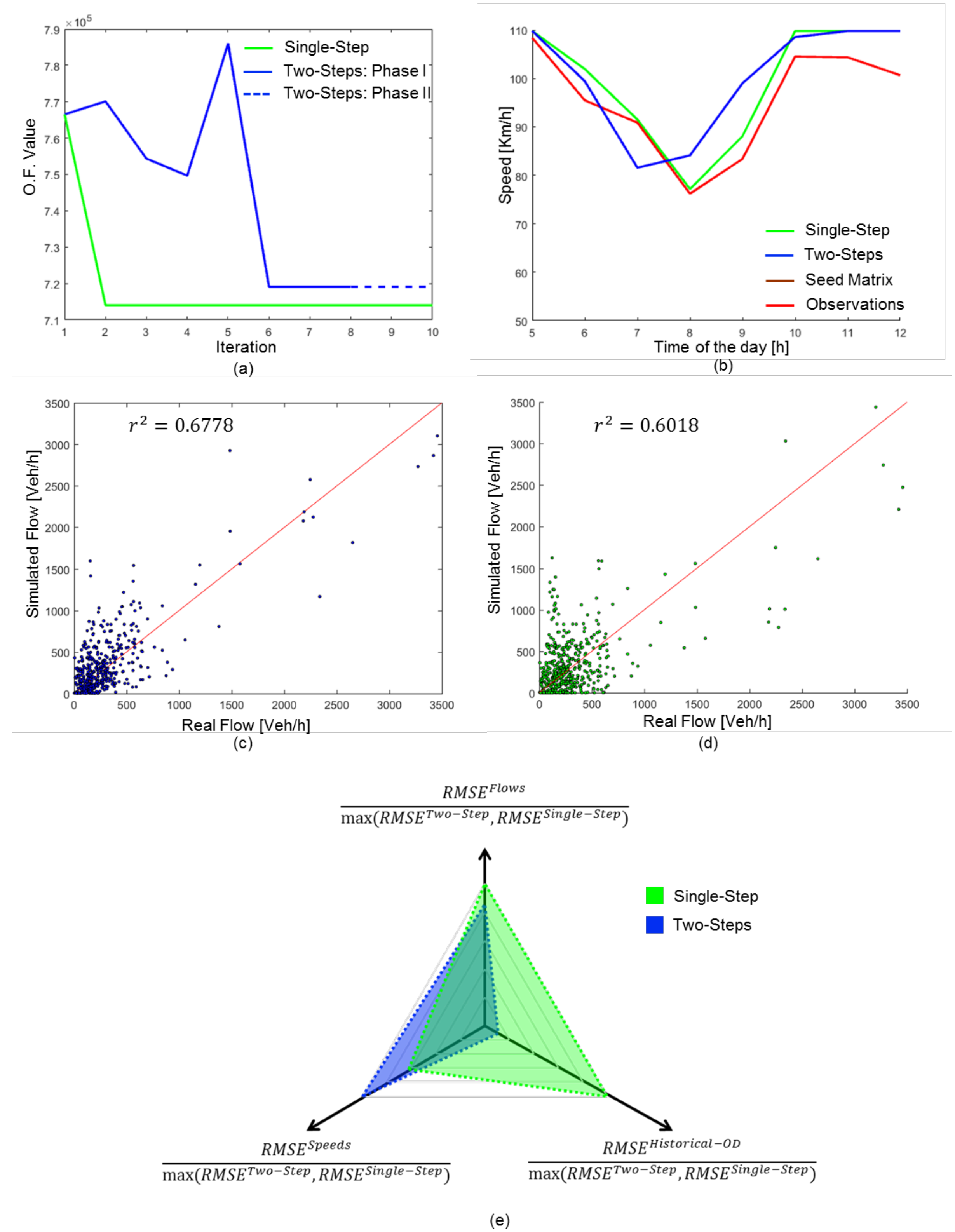

Figure 3: (a) Objective Function trend; (b) Estimated and Observed Average Speed; (c) Scatter Estimated and Observed Link Flows for the Two-Steps; (e) Spider Chart of the relative error for the estimated matrix in terms of Link-Flows, Distance from the Historical OD flows (Seed Matrix) and Average Speed;

This is further illustrated in the Spider Chart (Figure (3e)). While the Single-Step provides a substantial improvement with respect to the Two-Steps in terms of speeds, Figure (3e) shows that it clearly alters the structure of the demand, moving to a new local minimum and increasing the error on the link flow. Instead, the Two-Step estimation seems more robust. Although it does not provide an extremely accurate fit of the speed, it keeps the original structure of the demand and provides a reasonable approximation for both speeds and traffic counts, which is in line with the expectations.

CONCLUSIONS

In the previous work, the Two-Step approach showed to be able to provide more reliable estimation with respect to the Single-Step case. However, these results were based on a small network. This paper analyses the properties of using the Two-Steps approach on large-scale congested networks. The contribution of this study is both methodological and practical.

From the methodological point of view, the proposed approach relaxes the strong limitation of having a good starting demand matrix. As reported in [17], the capability of the DODE solution algorithm to correct the biases within the temporal and spatial structure of the demand is a strict requirement for having robust results. New data sources can improve the performances of the Single-Step, however, if such information is not available, then the Single-Step approach is likely to estimate a wrong structure of the demand. On the other hand, the Two-Steps approach showed that even considering only traffic counts and average speed, the model is capable of modifying the structure of the OD matrix in order to achieve more consistent results.

Following this procedure, the first step of the model estimates the total flow generated for each traffic zone, while keeping constant the distributions, thus using them as an indirect constraint for the demand. The assumption of having a linear relation between distributions reduces the number of possible solutions for the DODE. At the same time, as aggregate data works as an indirect constraint, the demand term can be removed from the goal function. The combination of these two effects creates a smoother objective function, with less local minima with respect to the classical sequential GLS proposed in [5]. Additionally, we showed in this paper that even when a "good" starting matrix is available, the Single-Step approach has the tendency of overfitting the available data, specifically those having a higher weight within the objective function. Although in this condition the estimated matrix results acceptable, the Two-Step approach seems likely to provide results that are more robust.

From a practical point of view, the model has been integrated within a Matlab package for dynamic demand estimation (MAMBA-DEV), which exploits PTV Visum as traffic assignment module. While the case study shows the network of Luxembourg, it can be easily implemented with any network in Visum, thus we can claim that the model is ready for practice.

Supporting these points, the paper introduces an experiment on a large-scale heavily congested network accounting for real traffic data. The main limitation of the work presented in this paper regards indeed the data used for the DODE. As the amount of information was limited, the calibrated matrix cannot be used yet for practical operations, such as long term planning. However, it represents the first attempt to have a dynamic matrix which covers a large part of the Grand Duchy of Luxembourg, which can be used as a starting point for further optimisation if a larger data set is available, as showed in the second experiment. Nevertheless, the test cases presented in this work support the idea that the model can handle real data and large networks. Future work will focus on validating the results through a larger database and implement more elaborate models within MAMBA-DEV package, which may include smarter optimisation methods [12] [15]. Specifically, the authors aims at comparing the proposed model with the Quasi-Dynamic model proposed in [16], as this model it is also expected to reduce the localism of the model with respect to the classical GLS.

Acknowledgment

The authors acknowledge for financing grant: AFR-PhD grant 6947587 IDEAS (Fonds National de Recherche FNR). The authors would also like to thank Motion-S for providing the average speed data.

REFERENCES

[1] G. Cantelmo, F. Viti, E. Cipriani, and M. Nigro, "Improving the Reliability of a Two-steps Dynamic Demand Estimation Approach By Sequentially Adjusting Generations and Distributions," presented at the Transportation Research Board 94th Annual Meeting, 2015.

[2] M. G. McNally, "The Four-Step Model," in *Handbook of Transport Modelling*, vol. 1, 0 vols., Emerald Group Publishing Limited, 2007, pp. 35–53.

[3] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, "The path most traveled: Travel demand estimation using big data resources," *Transp. Res. Part C Emerg. Technol.*, vol. 58, Part B, pp. 162–177, Sep. 2015.

[4] E. Cascetta and S. Nguyen, "A unified framework for estimating or updating origin/destination matrices from traffic counts," *Transp. Res. Part B Methodol.*, vol. 22, no. 6, pp. 437–455, Dec. 1988.

[5] E. Cascetta, D. Inaudi, and G. Marquis, "Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts," *Transp. Sci.*, vol. 27, no. 4, pp. 363–373, Nov. 1993.

[6] K. Ashok and M. E. Ben-Akiva, "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows," *Transp. Sci.*, vol. 36, no. 2, pp. 184–198, 2002.

[7] H. Tavana, "Internally-consistent estimation of dynamic network origin-destination flows from intelligent transportation systems data using bi-level optimization," PhD Thesis, 2001.

[8] R. Frederix, F. Viti, and M. J. Tampère, "Dynamic origin–destination estimation in congested networks: theoretical findings and implications in practice," *Transp. Transp. Sci.*, vol. 9, no. 6, pp. 494–513, Jul. 2013.

[9] R. Balakrishna, M. Ben-Akiva, and H. Koutsopoulos, *Offline Calibration of Dynamic Traffic Assignment: Simultaneous Demand-and-Supply Estimation*, vol. 2003. 2007.

[10] R. Frederix, F. Viti, R. Corthout, and C. M. J. Tampère, "New gradient approximation method for dynamic origin-destination matrix estimation on congested networks," *Transp. Res. Rec.*, no. 2263, pp. 19–25, 2011.

[11] E. Cipriani, M. Florian, M. Mahut, and M. Nigro, "A gradient approximation approach for adjusting temporal origin-destination matrices," *Transp. Res. Part C Emerg. Technol.*, vol. 19, no. 2, pp. 270–282, 2011.

[12] C. Antoniou, C. Lima Azevedo, L. Lu, F. Pereira, and M. Ben-Akiva, "W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models," *Transp. Res. Part C Emerg. Technol.*, vol. 59, pp. 129–146, Oct. 2015.

[13] A. Tympakianaki, H. N. Koutsopoulos, and E. Jenelius, "c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation," *Transp. Res. Part C Emerg. Technol.*, vol. 55, pp. 231–245, Jun. 2015.

[14] V. Marzano, A. Papola, and F. Simonelli, "Limits and perspectives of effective O–D matrix correction using traffic counts," *Transp. Res. Part C Emerg. Technol.*, vol. 17, no. 2, pp. 120–132, 2009.

[15] T. Djukic, J. Van Lint, and S. Hoogendoorn, *Application of principal component analysis to predict dynamic origin-destination matrices*. 2012.

[16] E. Cascetta, A. Papola, V. Marzano, F. Simonelli, and I. Vitiello, "Quasi-dynamic estimation of o–d flows from traffic counts: Formulation, statistical validation and performance analysis on real data," *Transp. Res. Part B Methodol.*, vol. 55, pp. 171–187, Sep. 2013.

[17] C. Antoniou *et al.*, "Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation," *Transp. Res. Part C Emerg. Technol.*, vol. 66, pp. 79–98, May 2016.

[18] K. Ashok, "Estimation and prediction of time-dependent origin-destination flows," Massachusetts Institute of Technology, 1996.

[19] P. GROUP, *VISUM 14 User Manual*. epubli, 2014.

[20] R. Frederix, F. Viti, and C. M. J. Tampère, "A density-based dynamic OD estimation method that reproduces within-day congestion dynamics," in *13th International IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 694–699.

[21] J. Barceló and L. Montero, "A Robust Framework for the Estimation of Dynamic OD Trip Matrices for Reliable Traffic Management," *Transp. Res. Procedia*, vol. 10, pp. 134–144, 2015.

[22] E. Mitsakis, J.-M. S. Grau, E. Chrysohoou, and G. Aifadopoulou, "A Robust Method for Real Time Estimation of Travel Times for Dense Urban Road Networks Using Point-to-Point Detectors," presented at the Transportation Research Board 92nd Annual Meeting, 2013.

[23] J. C. Spall, "Stochastic Optimization," in *Handbook of Computational Statistics*, J. E. Gentle, W. K. Härdle, and Y. Mori, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 173–201.

[24] F. Sprumont, P. Astegiano, and F. Viti, "On the consistency between commuting satisfaction and traveling utility: the case of the University of Luxembourg," *Eur. J. Transp. Infrastruct. Res.*, vol. 17, no. 2, Mar. 2017.

[25] K. A. Small, "The bottleneck model: An assessment and interpretation," *Econ. Transp.*, vol. 4, no. 1–2, pp. 110–117, Mar. 2015.

[26] M. Nigro, E. Cipriani, and A. Del Giudice, "Exploiting floating car data for time-dependent O-D matrices estimation," *J. Intell. Transp. Syst. Technol. Plan. Oper.*, vol. under review, 2017.