Transportation Research Part C 134 (2022) 103477

Contents lists available at ScienceDirect

Transportation Research Part C

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

A practical approach to assignment-free Dynamic Origin–Destination Matrix Estimation problem

Xavier Ros-Roca [a,b,*], Lídia Montero [a], Jaume Barceló [a], Klaus Nökel [b], Guido Gentile [c]

[a] Departament d'Estadística i Investigació Operativa (DEIO), Universitat Politècnica de Catalunya, Barcelona, Spain
[b] PTV Group, Karlsruhe, Germany
[c] Dipartamento di Ingegneria Civile, Edile e Ambientale (DICEA), Sapienza Università di Roma, Roma, Italy

ARTICLE INFO

Keywords:
Dynamic origin–destination matrices
Dynamic traffic assignment
ICT traffic data
Nonlinear optimization

ABSTRACT

Dynamic traffic models require dynamic inputs, one of the main ones being the Dynamic Origin–Destinations (OD) matrices describing the variability over time of the trip patterns across the network. The Dynamic OD Matrix Estimation (DODME) is a challenging problem since no direct observations are available, and therefore one should resort to indirect estimation approaches. Among the most efficient approaches, the one that formulates the problem in terms of a bi-level optimization problem has been widely used. This formulation solves at the upper level a nonlinear optimization problem that minimizes some distance measures between observed and estimated link flow counts at certain counting stations located in a subset of links in the network, and at the lower level a traffic assignment that estimates these link flow counts assigning the current estimated matrix. The variants of this formulation differ in the analytical approaches that estimate the link flows in terms of the traffic assignment and their time dependencies. Since these estimations are based on a traffic assignment at the lower level, these analytical approaches, although numerically efficient, imply a high computational cost. The advent of ICT applications has made available new sets of traffic-related measurements enabling new approaches; under certain conditions, the data collected allows to estimate the most likely used paths, from which a *de facto* assignment matrix can be computed. This allows extracting empirically similar information to that provided by the dynamic traffic assignment that is used in the analytical approaches. This paper explores how to extract such information from the recorded commercial data, proposes a new constrained non-linear optimization model to solve the DODME problem, with a reduced number of variables linearly depending on network size instead of quadratically. Moreover, the bilevel iterative process and the traffic assignment need are avoided. Validation and computational results on its performance are presented.

1. Introduction

Trip patterns in terms of origin-to-destination (OD) traffic flows are a key input to traffic assignment models, namely to dynamic traffic assignment models. The trip patterns must also be dynamic – or at least time discretized – to properly approximate the time variability of the demand. However, OD matrices are not yet observable; in the best case, the measurements from Information and Communications Technologies (ICT) such as GPS vehicle tracking or mobile phone call detail records (CDR) can be used to collect samples that must then be used to infer or estimate the whole population. In order to perform this estimation, indirect processes that are usually based on mathematical models that take as input traffic counts are most commonly used.

* Corresponding author at: Departament d'Estadística i Investigació Operativa (DEIO), Universitat Politècnica de Catalunya, Barcelona, Spain.
E-mail address: xavier.ros.roca@upc.edu (X. Ros-Roca).

https://doi.org/10.1016/j.trc.2021.103477
Received 30 December 2020; Received in revised form 8 November 2021; Accepted 9 November 2021
Available online 10 December 2021

The problem of OD matrix estimation has been interesting for both researchers and practitioners for decades, (Van Zuylen and Willumsen, 1980). Furthermore, the aim of capturing the congestion effects of the changing OD flows during the OD estimation process conducted the research to an appealing mathematical formulation as a bi-level optimization problem. Moreover, in the recent years, the dynamic traffic assignments arise in order to capture the evolution of the traffic system according its dynamicity, leading to the dynamic OD matrix estimation problem.

1.1. The OD matrix estimation problem

The OD estimation problem in terms of the bi-level optimization problem is shown in Eq. (1), aimed at adjusting an initial target OD, $\mathbf{X}^H$, so that it could explain the observed link flow counts $\mathbf{Y}$ at counting stations in the network. Lundgren and Peterson (2008) give the followed formulation for the static situation.

$$\begin{aligned} &\min Z(\mathbf{X}) = w_1 F_1(\mathbf{X}, \mathbf{X}^H) + w_2 F_2(\mathbf{Y}, \hat{\mathbf{Y}}) \\ &\quad s.to \quad \mathbf{Y} = Assignment(\mathbf{X}) \\ &\qquad\quad \mathbf{X} \geq 0 \end{aligned} \tag{1}$$

where F_1 and F_2 are suitable distance functions between estimated and observed values; while w_1 and w_2 are weighting factors reflecting the uncertainty of the information contained in $\mathbf{X}^H$ and $\mathbf{Y}$, respectively. The underlying hypothesis is that $\mathbf{Y} = Assignment(\mathbf{X})$ are the link flows predicted by assigning the demand matrix $\mathbf{X}$ onto the network, that will be notated as $\mathbf{Y}(\mathbf{X})$. This mathematical model is highly undetermined since the number of variables of the problem, the OD flows, is much larger than the number of the available link traffic counts. Therefore, the resolution of the optimization problem can lead to different solutions, even in the case when the seed OD matrix is proper to the solution, (Yang et al., 1992). It is well known that even a full covered network with traffic sensors on each link does not ensure a determined problem, (Bierlaire, 2002), therefore, an appealing research topic has been to explore new approaches including further information, such as link speeds or travel times (Cantelmo et al., 2014a; Nigro et al., 2018; Kostic and Gentile, 2019; Behara et al., 2020b), aimed at reducing such underdetermination.

Traffic modelling for transportation analysis has evolved to the dynamic traffic assignment (DTA) models, which are able to include the time dependencies on the traffic system, overcoming in this way the main drawbacks of static assignment of not accounting for the congestion generation and its dynamic propagation across the network. Dynamic models require then dynamic inputs, which means dynamic OD matrices that are represented as a time series of sequentially OD matrices. Therefore, the dynamic OD matrix estimation problem (DODME) becomes more complex, with more variables and time dependencies across the time periods of the traffic simulation, (Frederix et al., 2011b,a).

Cantelmo et al. (2014b) have considered DODME problem based on the bi-level approach and include a utility-based DTA models in the lower level relying on activity location and trip duration information. They demonstrate that, extending the bi-level approach by taking into account such information, the number of free parameters in the DODME problem systematically decreases, improving the reliability of the estimated dynamic OD matrices by reducing the underdetermination of the solution.

The relevant question of reliability is also addressed by other authors as (Yang et al., 1992) for the static OD matrix estimation problem. In this context, the analysis in Djukic (2014) can be taken as a reference, where it is proven that traditional distances are not able to capture the structural similarity between two matrices. Behara et al. (2020a) adopt Levenshtein distance, traditionally used to compare sequences of strings, and extends it to quantify the structural comparison of OD matrices and Ruiz de Villa et al. (2014) use Wasserstein distance to quantify structural similarity and address matrix estimation reliability. Djukic et al. (2013), Behara et al. (2018) use the Mean Structural Similarity Index (MSSIM), which is also adopted and enhanced by the authors on this paper.

Many different techniques have been used in the literature to solve the bi-level OD estimation problem on the offline context. On the one hand, the simultaneous perturbation stochastic approximation (SPSA) (Antoniou et al., 2015; Balakrishna, 2006; Kostic and Gentile, 2019; Cantelmo et al., 2014a) among others) and Osorio (2019) and her consequent research on the use of metamodels in simulation optimization-based methods that aim to find a derivative-free estimated descent direction of the objective function with a low number of evaluations, each of which requiring a full assignment. On the other hand, other approaches use an analytical form of the objective function and apply different methods to solve the problem, for instance GLS approaches for the offline problem (Cascetta et al., 2013) or Kalman filter for the dynamic online problem (Ashok and Ben-Akiva, 2002; Barceló et al., 2013).

Recent literature is addressed to include ICT measures into the DODME problem to reduce the underdetermination of the underlying problem. Mo et al. (2020) propose a two-step ordinary least squares (OLS) OD estimation model, which incorporates the output from a Bayesian path reconstruction model developed to cope with insufficient coverage rate of ICT data from Licence Plate Recognition and coestimates the dynamic OD demand and assignment matrix without any historical matrix need. Finally, Yang et al. (2017) and Krishnakumari et al. (2019) use the geopositioning data of probe vehicles from an ad hoc experiment designed by the authors to obtain an *a priori* dynamic OD matrix and the reconstructed paths are included into the OD estimation process.

1.2. Motivation

The linearization of the relationship between traffic counts and OD flows is one way of solving these problems in a more computationally efficient way. This can be achieved by using the proportion of the OD demand flows passing through the count location at a certain link. In these terms, the dynamic assignment matrix $\mathbf{A}(\mathbf{X})$ is the result of the mapping and can be rewritten as:

$$y_{lt} = \sum_{i\in I}\sum_{j\in J}\sum_{r\in \mathcal{T}} a_{ijr}^{lt} x_{ijr} \rightarrow \mathbf{Y} = \mathbf{A}(\mathbf{X})\mathbf{X} \tag{2}$$

where a_{ijr}^{lt} represents the proportion of the OD flow that departs from origin i at time period r and goes to destination j that crosses link $l \in \hat{L} \subseteq L$ at time period $t \geq r$. I, J and $\mathcal{T}$ stands for the set of origin zones, destination zones and the simulation time periods, respectively.

However, this linearization cannot properly account for the impacts of the traffic dynamics, time dependencies and route choice alternatives induced by congestions, because, as (Frederix et al., 2011a, 2013) highlight, in this case, counts in or downstream of congestion are not informative of demand, but of (discharge) capacity. Indeed, this linear mapping between the link flows and the OD flows is the first term in the Taylor expansion of the relationship between link flows and OD flows, at an OD matrix in the neighbourhood of **X**, (Toledo and Kolechkina, 2013). The additional terms of its Taylor expansion would capture the assignment matrix's sensitivity to such mentioned changes and its effects. However, how this paper proposes to calculate a_{ijr}^{lt} accounts implicitly for these effects as far as the time-dependent link travel times are obtained from the actual traffic conditions and therefore the estimation of a time-dependent assignment matrix is addressed. OD path proportion use is not constant across the optimization since OD time-depending travel costs are considered to derive OD path proportions and thus, OD path flows.

Then, the resulting bi-level optimization problem solves (at the upper level) the nonlinear optimization problem by substituting the estimated flows **Y** in the objective function of (1) with the relationship (2):

$$\min Z(\mathbf{X}) = w_1 F_1(\mathbf{X}, \mathbf{X}^H) + w_2 F_2(\mathbf{A}(\mathbf{X})\mathbf{X}, \hat{\mathbf{Y}}) \qquad (3)$$
$$s.to \quad \mathbf{X} \geq 0$$

If F_1 and F_2 are quadratic distances, the problem stated on Eq. (3) has an analytical and differentiable objective function, which makes it possible to use iterative optimization methods that present nice properties of convergence and stability. One example is applying the maximum descent method, which leads to the dynamic version of Spiess (1990). Details of its implementation can be found in Ros-Roca et al. (2020a). However, a DTA at the lower level to calculate the dynamic assignment matrix $\mathbf{A}(\mathbf{X})$ is still required for the evaluation of the objective function.

These analytical approaches to DODME problem show that all of them rely on the availability of the Assignment Matrix $\mathbf{A} = \left[a_{ijr}^{lt}\right]$ for the various time intervals, calculated at the lower level of (3) by the Dynamic Traffic Assignment at each time interval.

The availability of the GPS generated data enables us to assume that, after a suitable data processing to find the empirical paths and the inference of path choice proportions, it is possible to estimate a dynamic assignment matrix that relies on the information regarding traffic conditions. Since it would play a similar role to that of the analytical assignment matrix obtained by a Dynamic Traffic Assignment (DTA), we focus our attention on how to efficiently estimate that assignment matrix, in practical terms, from available commercial data as discussed above.

1.3. Article's structure

Regarding the use of empirical data for the intended purpose, a key aspect is whether one can control the data collection process or, on the contrary, one depends on the commercial GPS traces as supplied by data providers. The first situation ensures the quality and reliability of the data collected and also its adequacy to make the necessary estimations in terms of complete reliable trajectories from origins to destinations (Yang et al., 2017; Krishnakumari et al., 2019).

In the second case, data are usually of two types, either non-processed waypoints or in-house processed information as, for instance, speed profiles. Non-processed waypoints are not directly useable for transportation analysis, and they must be processed before: they must be filtered, cleansed to remove outliers and correct errors, and suitably map matched to fit the transport network. Alternatively, data supplier companies also provide references to tools to extract additional information like speed profiles at link level from the waypoints, which can be used for transportation analysis to infer OD travel times among other applications. An example of such a tool would be OpenLR, (OpenLR, 2020).

The paper is organized in three sections, after this introductory section. The first one, Section 2, defines the methodological framework designed for the data-driven OD estimation method: Section 2.1 describes the basic estimation of most likely used paths from estimated travel times at link level followed by the process of obtaining the estimated dynamic assignment matrix estimation; and Section 2.2 resumes the statement of the optimization problem for DODME. Section 2.3 proposes an enhanced MSSIM structural similarity indicator to assess goodness of fit of DODME results

In Section 3, we outline the synthetic experimental design that aims to validate the proposed methodology in terms of consistency, robustness and sensitivity. We do this by discussing the experimental design on Section 3.1 and outline the data collection process to select synthetic GPS data that it is detailed in Appendix B. We validate the Methodological Approach and assess the observed assignment matrix properties for the synthetic experiment in Section 3.2 and Sections 3.3 and 3.4 discuss the quality of DODME results and their relationship to GPS penetration rate.

Section 4 presents a real data experiment; Section 4.1 illustrates the data-driven procedure to estimate an observed assignment matrix based from GPS traces on a larger and real network (Turin) and obtained DODME results are discussed in Section 4.2. We offer conclusions and a discussion on further research avenues in Section 5.

The type of information we assume in this paper is either available from the proposed processing, or generated by the analyst using other available tools. The estimation of link speeds, or equivalently link travel times, from the recorded data, is a common requirement to most of the data-driven approaches, as for Krishnakumari et al. (2019). Therefore, unless the already mentioned data collection hypothesis are satisfied (e.g., (Yang et al., 2017)), they are not directly measured but can be inferred from the available data (e.g., Lopez et al. (2017a,b)), where the assumption is that path travel times are obtained by license plate recognition

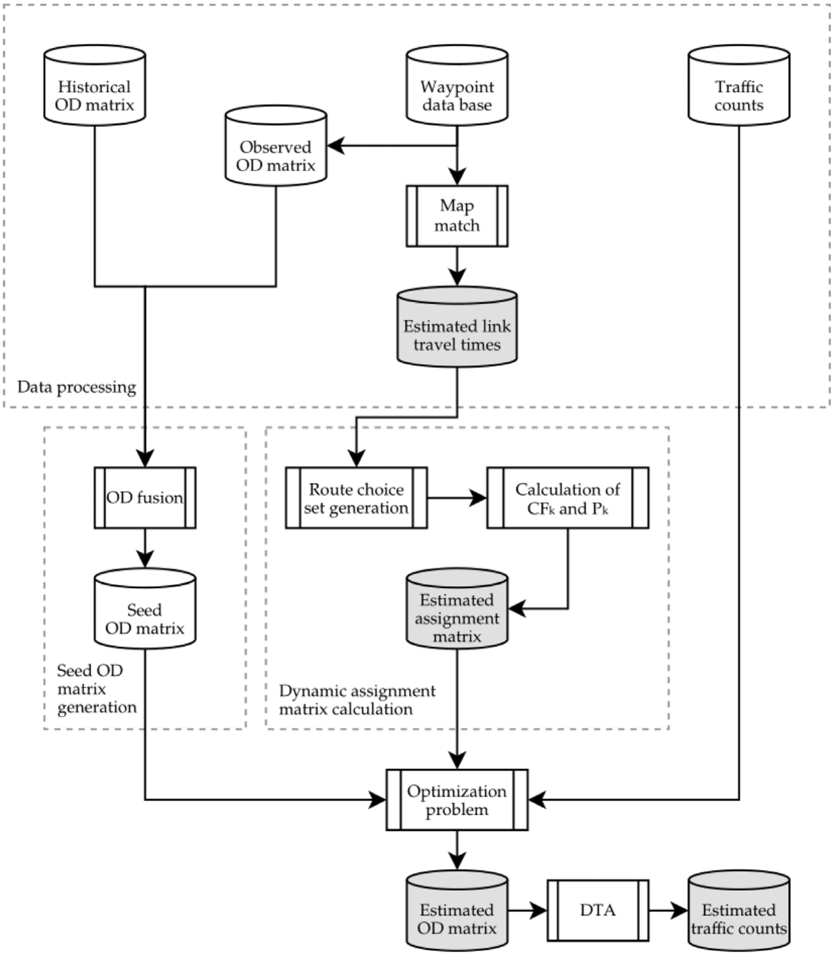


Fig. 1. The Data-driven Assignment-free DODME methodology.

and are available from shortest paths between origins and destinations by a simple heuristic approach. In our case, as path travel times could not be reliable for the mentioned reasons, a different approach is proposed based on the waypoints and the map-matching process, either with professional tools, like OpenLR, or those provided by transport planning platforms, like the ones that for practical purposes provide the utilities of software platforms for transportation analysis (as for instance the GPX matching utility in PTV Visum, (PTV Vissim, 2020b)). These practical details are addressed in Appendix A. In all cases, this paper assumes that the available data have already been filtered, cleansed and processed, and therefore, we focus the work on what can be done with the available link travel times estimations.

2. Methodological approach: A data-driven assignment-free DODME

Therefore, assuming that a set of estimated link travel times – obtained either from ICT providers, or processed by the user according to the process explained in Appendix A – and a set of traffic counts are available for a selected period of time; a specific purpose designed process produces route choice paths and proportions for generating an estimated assignment matrix. Then, the research question addressed in the following sections is to investigate whether it is possible to use such information to state a different formulation of the DODME problem, in terms of an optimization model, not requiring the execution of any dynamic traffic assignment procedure. The conceptual computational scheme of the proposed data-driven assignment-free DODME approach, powered by the ICT applications capturing GPS data trajectories and providing estimated link travel times, is summarized in Fig. 1.

2.1. Calculation of the dynamic assignment matrix

According to assignment-based methods, the paths used to travel between origins and destinations are provided by a user equilibrium assignment. In an assignment-free approach, we propose an alternative method relying on the available estimated link travel times to generate a plausible Route Choice Set, $\mathcal{K} = \left\{K_{ijr}, \forall i \in I, \forall j \in J, \forall r \in \mathcal{T}\right\}$, specifically from among the most likely used paths between each origin and each destination at each departure time. As already mentioned, when these estimated link travel times are not available, we suggest to obtain them from GPS tracking data, using the specific tools from providers for cleaning and filtering and obtaining the travel times at link level, or by using the described methodology in Appendix A.

The estimated link travel times for each link of the network l at each time period t, $\hat{tt}_{lt}$ are the main inputs for generating the route choice set. That is the set of most likely alternatives. This is usually done through a selective approach that identifies the routes based on some previously mentioned criteria (k shortest paths, path flows computation, etc.).

Many alternative approaches can be used for this, and all them are essentially based on variants of k shortest paths. Alternatives based on iteratively applying Dijkstra-based algorithms for similar purposes while explicitly accounting for overlapping penalties have been analysed by Janmyr and Wadell (2018) and Nassir et al. (2014). Other alternative procedures based on Chabini (1998) time dependent shortest paths or path search algorithms can be found in the literature. The option implemented in this paper is discussed in Section 3.1.

Once the candidate routes in the route choice set are specified, a key question in the route choice model is how to address the problem that the alternatives are usually not independent but correlated due to overlapping paths. From a theoretical point of view, Probit models are likely those who better account for these correlations, but the difficulties in practically implementing them led to search for other approaches. Cascetta et al. (1996), Ben Akiva and Bierlaire (1999) propose alternative models for capturing the correlation among alternatives by modifying the logit-based choice, specifically by measuring the degree of similarity between the alternatives and adding it to the utility's deterministic component in the corresponding discrete choice model. This term is usually called the "*commonality factor*", and its main role, (Cascetta, 2001), is to overcome the problems deriving from the basic hypothesis of independence of irrelevant alternatives, which discrete choice logit models assume and could otherwise lead to unrealistic results. This term reduces the systematic utility of a path in proportion to its level of overlapping with other alternative paths. The formulation adopted in this paper is Janmyr and Wadell (2018)'s proposed modification of the formulation of Bovy et al. (2008).

The paths in K_{ijr} are denoted by $k(i,j,r) \in K_{ijr}$ in order to explicitly show the dependence on (i,j,r). For a certain path $k(i,j,r)$, the sequence of links that compound it is the set $\Gamma_{k(i,j,r)} = \{e_1,\dots,e_{m_k}\}$. Then, the proportion of path choice for each path in the set K_{ijr} is calculated as a modified discrete logit-based choice model that uses the commonality factor within the OD pair and time period, CF_k. It further acts as an additive penalization factor on current travel times (Bovy et al., 2008). That is:

$$CF_{k(i,j,r)} = \frac{1}{\mu_{CF}} \sum_{a\in\Gamma_{k(i,j,r)}} \left(\frac{l_a}{L_{k(i,j,r)}} \log\left(\sum_{h\in K_{ijr}} (\delta_{ahr}+1) \right)\right) \tag{4}$$

$$P_{k(i,j,r)} = \frac{\exp[\mu_{P_k}(-\hat{tt}_{k(i,j,r)} - CF_{k(i,j,r)})]}{\sum_{h\in K_{ijr}} \exp[\mu_{P_k}(-\hat{tt}_{h(i,j,r)} - CF_{h(i,j,r)})]} \tag{5}$$

where δ_{ah} indicates whether path $h \in K_{ijr}$ uses link a; l_a is the length of link a; and $L_{k(i,j,r)}$ is the total length of path $k \in K_{ijr}$. In order to be unit's consistent, to adapt magnitudes for the discrete choice summation and permit more variability between paths in the case of short travel times, μ_{P_k} and μ_{CF_k} are parameters that were fixed as follows:

$$\mu_{P_k} = \mu_{CF_k} = \frac{1}{mean_{k\in K_{ijr}}\left(\hat{tt}_{k(i,j,r)}\right)} \tag{6}$$

These calculations obtain the flow distribution for each path on the basis of observed path travel times, which are the summation of the observed time-dependent link travel times, that as long as they are estimated from the actual traffic conditions from GPS data, they satisfy the conditions for Eq. (2). This implies that the arrival time, $t(k)$, at each link a, included in the path $k(i,j,r)$:

$$\hat{tt}_{k(i,j,r)} = \sum_{a\in\Gamma_{k(i,j,r)}} \hat{tt}_{at(k)} \tag{7}$$

Once $\mathbf{P}_k = \left[P_{k(i,j,r)}\right]$ is determined from the k shortest paths that were obtained from the travel times, which themselves were estimated from the GPS data for all OD pairs, we can then calculate the estimated time-dependent assignment matrix $\bar{\mathbf{A}} = \left[\bar{a}^{lt}_{ijr}\right]$:

$$\bar{a}^{lt}_{ijr} = \sum_{k\in K_{ijr}} \delta^{lt}_{k(i,j,r)} P_{k(i,j,r)} \quad , \quad \forall i,j,r,l,t \tag{8}$$

where $\delta^{lt}_{k(i,j,r)}$ is the estimated incidence indicator:

$$\delta^{lt}_{k(i,j,r)} = \begin{cases} 1 & \text{if path } k(i,j,r) \text{ uses link } l \text{ at time } t \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

This estimated assignment matrix is the outcome of Dynamic Assignment Matrix Calculation box included in Fig. 1 and constitutes an estimate of the dynamic assignment matrix that would be obtained by a DTA assignment based on the ground truth matrix. As a consequence, the estimated assignment matrix derived by the proposed process does not depend on a perfect calibrated model allowing to split the calibration process into the network supply calibration and the demand calibration stages.

2.2. Optimization procedure

The possibility of estimating an assignment matrix (8) allows reformulating DODME by replacing (2) (the assignment matrix provided by the DTA) with the estimated matrix. Then, (2) can be rewritten as:

$$\bar{y}_{lt} = \sum_{i\in I}\sum_{j\in J}\sum_{r=1}^{t} \bar{a}_{ijr}^{lt} x_{ijr} \tag{10}$$

where $\bar{y}_{lt}$ is the estimated flow in link l at time period t; x_{ijr} is the flow departing origin $i \in I$, with destination $j \in J$, at time interval $r \in \mathcal{T}$; and $\bar{a}_{ijr}^{lt}$ is the estimated assignment matrix, which is the fraction of trips from origin i with destination j, departing at time r reaching link l at time t.

As in all optimization methods that aim to find a solution, a seed OD matrix must be provided to the OD estimation process as a feasible starting point. In this described methodology, many alternatives arise. One common option is to use a reliable historical OD matrix as a suitable seed for the optimization algorithm (Cascetta et al., 2013; Kostic and Gentile, 2019; Cantelmo et al., 2014a; Nigro et al., 2018), specially, in those cases in which DODME is applied to not very long time periods to support dynamic traffic models in conditions where surveillance systems likely provide reliable historical OD estimates containing a wealth of structural information (Ashok and Ben-Akiva, 1993; Ben Akiva et al., 2001). In real life applications of these approaches for traffic management, the assumption of having a reliable reference OD matrix also holds, (Djukic et al., 2018; Aimsun, 2017). Another alternative, if data collected from a sample of GPS-tracked vehicles is available, is to create a discrete time estimate of the target OD matrix from it, which is the observed OD matrix, $\hat{\mathbf{X}} = [\hat{x}_{ijr}]$. In both cases, the target OD matrix can be expanded to estimate the OD matrix in terms of the scaling factors per origins, $\alpha_i, i \in I$, and per destinations $\beta_j, j \in J$, such that:

$$x_{ijr} = \alpha_i \beta_j \hat{x}_{ijr} \quad , \quad \forall i \in I, \forall j \in J, \forall r \in \mathcal{T} \tag{11}$$

The decoupling into independent scaling factors for origins and destinations is a simplifying assumption that will require a further analysis in more complex scenarios. The proposal is inspired by gravity models that set bi-dimensional constraints for rows and columns, as in the double-constrained models that are common for updating gravity distribution models, (Ortúzar and Willumsen, 2011). The inclusion of a third scaling factor γ_r, depending on the sliding time windows, could make sense for larger time periods, when the time variability of the demand can be influenced by other structural aspects, but not in the short term investigated in this paper.

However, when a historical OD is available from other sources, then a seed matrix x_{ijr}^0 can be generated combining the Historical OD matrix x_{ijr}^H with the observed OD matrix $\hat{x}_{ijr}$, which is obtained from GPS tracked trips. The seed OD matrix is denoted as $\mathbf{X}^0 = [x_{ijr}^0]$. This possibility consists of generating a proper seed OD matrix as a combination of the two different sources, Eq. (12). A specific proposal for fusing both OD matrices, discussing the functional form, is made in Section 3.3.

$$x_{ijr}^0 = \begin{cases} \hat{x}_{ijr} & \text{when only } \hat{x}_{ijr} \text{ is available} \\ f(\hat{x}_{ijr}, x_{ijr}^H) & \text{when both matrices are available} \\ x_{ijr}^H & \text{when only } x_{ijr}^H \text{ is available} \end{cases} \tag{12}$$

If $\hat{y}_{lt}$, $l \in \hat{L} \subseteq L$, $t \in \mathcal{T}$ are the link flows measured at the counting stations, in a subset $\hat{L} \subseteq L$ of the network links, the dynamic data-driven assignment-free OD matrix estimation problem can be formulated as an optimization problem for finding the values of the scaling factors $\alpha_i, i \in I$ and $\beta_j, j \in J$, without any need to conduct the traffic assignment at the lower level of (1). This is done by exploiting the estimated assignment matrix $\bar{a}_{ijr}^{lt}$.

From (10), (11) and (12), the proposed new formulation of the DODME problem including the estimated assignment matrix is:

$$\min_{\alpha_i,\beta_j} \left[w \left(\sum_{i\in I}\sum_{j\in J}\sum_{r=1}^{t} \left(x_{ijr}^H - \alpha_i\beta_j x_{ijr}^0 \right)^2 \right) + \sum_{l\in\hat{L}}\sum_{t\in\mathcal{T}} \left(\hat{y}_{lt} - \sum_{i\in I}\sum_{j\in J}\sum_{r=1}^{t} \alpha_i\beta_j \bar{a}_{ijr}^{lt} x_{ijr}^0 \right)^2 \right] \tag{13}$$
$$s.to \quad \alpha_i, \beta_j \geq LB$$

The problem variables are multiplicative scaling factors for each origin α_i and destination β_j, that have been chosen to drastically reduce the number of variables from $|I| \cdot |J| \cdot |\mathcal{T}|$ to $|I| + |J|$. Moreover, using the scaling factors as variables aims to preserve the structure of the seed OD matrix, as gravity models do. In this situation, the objective function is a quartic polynomic function with respect to the scaling factors, and it is convex as it is the sum of convex functions. The minimization problem is solved iteratively by means of the L-BFGS-B method, (Morales and Nocedal, 2011). It is a quasi-Newton method solved for constrained non-linear problems with a high number of variables that efficiently reduces the memory requirements and the computational burden. The available version in python package *scipy.optimize* has been used in this case.

Theoretically, LB should be a non-negativity constraint for all the scaling factors α_i, β_j. However, from a practical point of view, $\alpha_i = 0$ or $\beta_j = 0$ implies that a positive OD flow of the seed OD matrix from a certain origin or certain destination must be converted to 0. Therefore, considering that the seed OD matrix, in Eq. (12), comes from reliable information on mobility, the scaling factors cannot be 0 and the lower bound should therefore be $LB > 0$.

By the end of the optimization procedure, an estimation of the OD matrix, $\mathbf{X}^* = [x_{ijr}^*]$, is obtained. Therefore, as shown in Fig. 1, a dynamic traffic assignment is launched to obtain the corresponding simulated values for traffic counts, that are $\hat{\mathbf{Y}}^* = [y_{lt}^*]$.

2.3. MSSIM to measure OD matrix similarity

In a previous paper Ros-Roca et al. (2020a) that follows the trends of Djukic (2014) or Behara et al. (2018), we showed that the conventional indicators for measuring DODME quality are insufficient. Although they provide information on the DODME optimization problem's convergence of the objective function and how well the simulated flows fit the observed flows in terms of R^2, they do not pay any attention to the quality of the results from a structural point of view. From a traffic point of view, classic measures between OD matrices (MSE, MAE…) do not identify whether traffic OD patterns resulting from their adjustment approach exhibit an acceptable degree of structural similarity to the target matrix or, alternatively, if the approach provides a perturbed matrix that is structurally different.

The SSIM – the structural similarity index –, for a matrix of pixels that is the product of three different comparison components: luminance, contrast and structure, is a suitable measure to take into account the similarity in terms of magnitude, dispersion and structure. It is defined as follows:

$$SSIM(\mathbf{a},\mathbf{b}) = L(\mathbf{a},\mathbf{b})^{\alpha} C(\mathbf{a},\mathbf{b})^{\beta} S(\mathbf{a},\mathbf{b})^{\gamma} \tag{14}$$

where

$$\begin{cases} L(\mathbf{a},\mathbf{b}) = \dfrac{2\mu_a\mu_b + C_1}{\mu_a^2 + \mu_b^2 + C_1} \\ C(\mathbf{a},\mathbf{b}) = \dfrac{2\sigma_a\sigma_b + C_2}{\sigma_a^2 + \sigma_b^2 + C_2} \\ S(\mathbf{a},\mathbf{b}) = \dfrac{\sigma_{ab} + C_3}{\sigma_a\sigma_b + C_3} \end{cases} \tag{15}$$

Here, $\mu_a, \sigma_a, \mu_b, \sigma_b, \sigma_{ab}$ are the mean, standard deviation and covariance of the vectors $\mathbf{a}$ and $\mathbf{b}$, while $C_1 = C_2 = 2 \cdot C_3 = 10^{-6}$ are small stability constants for avoiding numerical problems. α, β, γ are weighting coefficients typically set to 1.

The idea, borrowed from Wang et al. (2004), applies $MSSIM$ (averaged $SSIM$ index for sliding windows covering the whole matrix) to measure the structural similarity between an OD matrix, $\mathbf{X}$, and an adjusted OD matrix $\mathbf{X}^*$. In Ros-Roca et al. (2020a), averages according to rows and columns are proposed, that is, by using rectangular sliding rules that correspond to either rows or columns in the OD matrix, which correspond to the trip distribution structure and therefore have a straightforward interpretation in terms of the underlying transportation system. Thus, SSIM will capture the similarity between these distributions by considering the mean, the variance and the structure of departing and arriving distributions, all of which correspond to the structural property of the trip patterns described by the OD matrix.

Furthermore, let us assume that there are N_S defined submatrices in $\mathbf{A}$ and $\mathbf{B}$. Then, if $MSSIM$ is $SSIM(\mathbf{a},\mathbf{b})$ is averaged over N_S sliding windows, a key question arises in regard to whether all windows have the same weight or their role in the total demand requires having different weights. In the case of OD matrices, it is obvious that not all origins or destinations are equivalent in a transport network. Therefore, a weighted MSSIM (as in Wang and Simoncelli (2008)) prioritizes those origins and destinations with more impact on the network. This proposed weighting average is defined as follows:

$$MSSIM(\mathbf{A},\mathbf{B}) = \frac{\sum_{i=1}^{N_s} W(\mathbf{a}_i,\mathbf{b}_i) SSIM(\mathbf{a}_i,\mathbf{b}_i)}{\sum_{i=1}^{N_s} W(\mathbf{a}_i,\mathbf{b}_i)} \tag{16}$$

where $\mathbf{a}_i, \mathbf{b}_i$ are respectively the ith windows of $\mathbf{A}, \mathbf{B}$, while the weight $W(\mathbf{a}_i,\mathbf{b}_i)$ is given by:

$$W(\mathbf{a}_i,\mathbf{b}_i) = \log\left[\left(1 + \frac{\sigma_{a_i}^2}{C_2}\right)\left(1 + \frac{\sigma_{b_i}^2}{C_2}\right)\right] \tag{17}$$

Weighting factors for the sliding windows, in the case of OD matrices, account for variances of the selected windows that, given how they are defined, represent the variance of trips from an origin (or to a destination) to all destinations (from all origins).

3. Synthetic experiment

Antoniou et al. (2016) established a framework for synthetic experiments for OD estimation algorithm, based on the synthetic generation of traffic counts and a historical OD matrix. This framework aims to test and validate any proposed methodology using controlled data sets that enable the evaluation of its consistency and performance. In this sense, the following section performs a synthetic data generation for the selected network that is used to evaluate the robustness of the presented method and validate its effectiveness on a potential real case under certain goodness conditions of the available data.

In Appendix B, the details of the data generation process are available. A framework is defined, which generates not only traffic counts and a historical OD matrix, but also a sampled GPS data base with independent trajectories from which travel times can be estimated, whether using commercial tools or using Appendix A as mentioned above. It should be highlighted that, since the synthetic generation allows to control the traffic conditions and data gathering, the final GPS data set is a data set without biases, filtered and cleansed to obtain reliable estimations of travel times at link level.

Regarding the testing of the methodology, the network used is shown in Fig. 2 and its characteristics are in Table 1. This network model is suitable for the test computational experiments, since all intersections are signed-controlled and the dynamic traffic

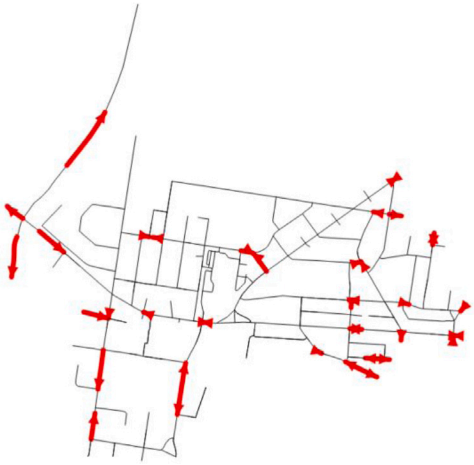


Fig. 2. The network used with the detection layout.

Table 1
Network and OD characteristics.

Time periods	4
Zones	114
Detectors	40
OD pairs X time	≈52k
Ground truth trips	8300
Ground truth positive OD	≈41k (78.49%)
Average number of paths per OD	7.27

assignment accounts specifically for queuing dynamics at controlled intersections. The model has been calibrated and the ground truth assignment shows the dynamics of the congestion spill backs through the network. Moreover, the time periods considered in the computational experiment is in the middle of a warm-up prior to the data collection and a discharging process to ensure that the corresponding trips have crossed the network. The simulation time interval is from 07:30AM to 08:30AM sliced into 15-minutes time periods. The detection layout was generated as explained in Appendix A. A total of 40 sensors were placed according to a certain level of desired coverage. They cover 97.27% of the total ground truth flow, which is 69.64% of the OD pairs totally captured and 13.07% of those partially captured. The traffic counts for each time interval at each sensor are produced. As shown in Table 1, the average number of alternative meaningful paths for each OD pair is approximately 7, which makes it a suitable network to study the assignment matrix calculation methodology. Further than the size of the network, what is relevant is its structure and the average number of meaningful routes between each OD pair.

3.1. Specific implementation of a proof of concept

Consequently with the identified drawbacks of using commercial data in Section 1.3, synthetic data sets have been generated by following the methodological framework defined in Appendix B, in order to test the assignment-free DODME described in Section 2 and its sensitivity and robustness. This is similar to some of the most current final tests of validation of different approaches, for example (Krishnakumari et al., 2019; Antoniou et al., 2015), etc. The mesoscopic model used is Visum-SBA (simulation based assignment) (PTV Vissim, 2020b), while PTV Vissim (PTV Vissim, 2020a) is the microscopic software used.

Finally, the historical OD matrix is generated from the ground truth OD matrix by following Antoniou et al. (2016)'s procedure, that is:

$$x_{ijr}^{H} = x_{ijr}^{GT}(p + q \cdot \varepsilon) \quad , \quad p = 0.75 \quad , \quad q = 0.15 \quad , \quad \varepsilon \sim N(0, 1/3) \quad , \quad \forall i, j, r \tag{18}$$

The historical OD matrix is on average decremented by 25% with a random perturbation. This perturbation tries to emulate a realistic historical OD matrix from surveys and past projects that represent similar traffic conditions.

In the case of the estimated link travel times, they are calculated after generating the GPS traces of different vehicles of the network. In this case, we emulated a controlled data collection, which consists of recording the waypoints sequences of vehicles on different days but on similar traffic conditions. We collected data during 200 days in similar conditions (independent replications by microsimulation, Appendix B), equivalent to an annual average working days, and fixed different penetration rates, which represent the number of random vehicles (among the total) that are recorded each day. In this study, a uniform penetration rate has been used for every day and also for all the OD pairs. Depending on the selected penetration rate, 5, 10 or 15%, these samples contain between 4.7 M and 14 M waypoints, which represent between 106 k and 318 k different vehicle trips circulating on the network.

Table 2
MSSIM values for the observed OD matrix.

	pen_rate	L	C	S	$MSSIM$
GT	5%	0.9753	0.9595	0.9046	0.8654
	10%	0.9856	0.9773	0.9357	0.9113
	15%	0.9894	0.9844	0.9548	0.9355
Hist	5%	0.9396	0.9059	0.9047	0.7884
	10%	0.9472	0.9284	0.9351	0.8320
	15%	0.9501	0.9375	0.9537	0.8550

The frequency of recording waypoints has been assigned differently to each vehicle, following an empirical distribution of latencies from an INRIX real GPS data set of another network. Improvements would be expected as penetration rates increase, since a higher penetration rate provides better information about the traffic conditions.

For processing the simulated GPS data generated by Vissim, we used the tool *Import GPX file*, which transforms the GPS data set as in Table 8 into paths using Visum links and interpolating travel times at the link level, as explained in Appendix A.

The estimated time-dependent link travel times are used to generate a route choice set (see Section 2.1) by using an independent tool, namely a path search algorithm available in Visum (PTV Vissim, 2020b, 6.18) that calculates specified sets of k shortest paths by perturbing the link's impedances with a normal distribution perturbation. The initial link costs are the estimated link travel times, and the maximum number of paths between connectors for each OD pair are set to $N_{max} = 5$. This parameter must be stated differently according to the network characteristics in order to generate sufficient number of paths but limiting the number of irrelevant paths.

The optimization procedure in Eq. (13) is set with two different stopping criteria: a maximum of 100 iterations; and a threshold for the relative error of the objective function, which is $thrsh = 0.005$.

3.2. Validation of the methodological approach

In order to computationally test the consistency and quality of the algorithmic framework that is defined conceptually in Fig. 1 and to analyse the quality of the partial results at each step, we have conducted a set of computational experiments based on the synthetic data generated by simulation. This allows further analysis of the quality of the methodology described above. Following Fig. 1, all the different sub-results of the steps of the algorithm are analysed.

3.2.1. Observed OD matrix analysis

The observed OD matrix is the mere counting of how many trips departs from each origin and arrive to each destination at each time period on the synthetic generated data according to the corresponding penetration rate, fully detailed on Appendix B. Based on the building process of the GPS tracking data, it is expected to obtain an observed OD matrix similar in the OD pattern structure to the ground truth OD matrix, given that the penetration rate of the GPS technology has been set homogeneous to all the OD pairs.

Instead of using the conventional MSE or similar indicators to compare OD matrices the $MSSIM$ and related metrics described in Section 2.3 are used.

The measure used to compare both OD matrices is the MSSIM and its components L, C and S (see Eq. (15)) that have been weight-averaged once calculated for each sliding window. In order to use L and C, the observed OD matrix must be scaled in order to have the same magnitude for both matrices in terms of total OD trips.

Table 2 shows the MSSIM values for the observed OD matrix with respect to the ground truth and historical OD matrices. The L values, which correspond to magnitude, are very high because of the previously mentioned scaling; therefore, no further analysis has to be made. High C values indicate similar dispersion of the values, and high S means that the observed OD matrix has a similar pattern, thus indicating that the GPS sample is not biased and the penetration rate is homogeneous for origins and destinations, as expected. Globally, MSSIM values are high, and they increase as the penetration rate increases, meaning that the built sample of GPS data presents the appropriate goodness to be used to estimate the link travel times. Furthermore, these results validate the use of the observed OD matrix as a seed OD matrix for the DODME method's optimization procedure.

3.2.2. The effect of the commonality factor

As mentioned in Section 2.1, the term CF_k acts as a penalization for the discrete choice. CF_k, P_k and $\bar{a}^{lt}_{ijr}$ are calculated by applying (4), (5) and (8). The commonality factor, CF_k, penalizes those paths that are similar to others in the same set K_{ijr}. Then, it reallocates the flows accordingly by increasing or decreasing the corresponding flow. In order to visualize the effect of CF_k on the path flow distribution, a positive OD flow is selected. Its corresponding route choice set is depicted in Fig. 3, where eight paths resulted from the route choice set calculations. Since there are three paths that are very similar, they were clustered by similarity into three sets (I, II, III) in order to better understand the role played by the commonality factor. Although link lengths are not large, network complexity in terms of number of OD paths per OD pair is high making the network suitable to address the consistency of the estimated assignment matrix appearing in the methodological proposal.

As shown in Fig. 3, Tables 3 and 4, set I is compounded by four very similar paths and without the penalization (using only travel times) results in more than one-half of the flow being assigned to I. On the other hand, when the penalization is applied,

Fig. 3. Path set (K_{ijr}) for a selected origin and destination in the network.

Table 3
Results of path distribution for a selected OD pair in the network.

Path	Length	Observed time	P_k without CF_k	P_k with CF_k	% of gain	Set
1	1.20 km	2 min 11 s	15.21%	15.18%	−0.03%	I
2	1.61 km	2 min 57 s	12.64%	11.63%	−1.01%	I
3	1.55 km	3 min 13 s	11.68%	13.33%	+1.65%	II
4	1.52 km	2 min 59 s	12.45%	13.29%	+0.84%	III
5	1.38 km	3 min 02 s	12.24%	11.15%	−1.09%	I
6	1.26 km	3 min 06 s	12.02%	10.54%	−1.48%	I
7	1.32 km	3 min 18 s	11.54%	12.69%	+1.15%	II
8	1.29 km	3 min 03 s	12.22%	12.19%	−0.03%	III

Table 4
Results of path distribution by sets for a selected OD pair in the network.

Set	Mean length	Mean Obs time	P_k without CF_k	P_k with CF_k	% of gain
I	1.36 km	2 min 49 s	52.11%	48.50%	−3.61%
II	1.43 km	3 min 15 s	23.22%	26.02%	+2.80%
III	1.41 km	3 min 01 s	24.67%	25.48%	+0.81%

48.50% of the flow is assigned to these four paths. This is because the four paths are very similar and capture the majority of the flow from the beginning.

On the other hand, this flow is assigned to the other sets, which are different and do not share very much with the other sets in Table 4. For instance, set II receives more flow after the commonality factor penalization, because it is the one that shares less with the other sets.

3.2.3. Qualitative analysis of the estimated assignment matrix

The estimated assignment matrix resulting from this methodology is used to capture the real mobility of the network, doing so by processing a large amount of GPS waypoints. While solving the optimization problem, in which an objective function minimizes the traffic flow differences detected by sensors, this assignment matrix remains invariant for each time interval, thus confirming the stability hypothesis formulated by Cascetta et al. (2013). Therefore, it is important to check the consistency between paths in the SBA (which is used to generate the reference data) and to identify the paths used to analyse the route choice set of paths (which are generated from the empirical data). Stated briefly, one should confirm that the theoretical assignment matrix from SBA is consistent with the estimated assignment matrix. Our qualitative proposal is to use the most relevant link in the network to compare OD flows.

The left-hand side of Fig. 4 shows the OD flows using a singular link (the most used link) in the network, according to the ground truth SBA assignment. On the right, the same picture is plotted by means of the estimated assignment matrix and the historical OD values. Both graphs are qualitatively similar, such that the downstream and upstream propagation of the flows circulating on this link indicate a similar assignment matrix.

Fig. 4. OD flows for the most used link using SBA assignment (left) and using estimated OD matrix for a selected experiment (right).

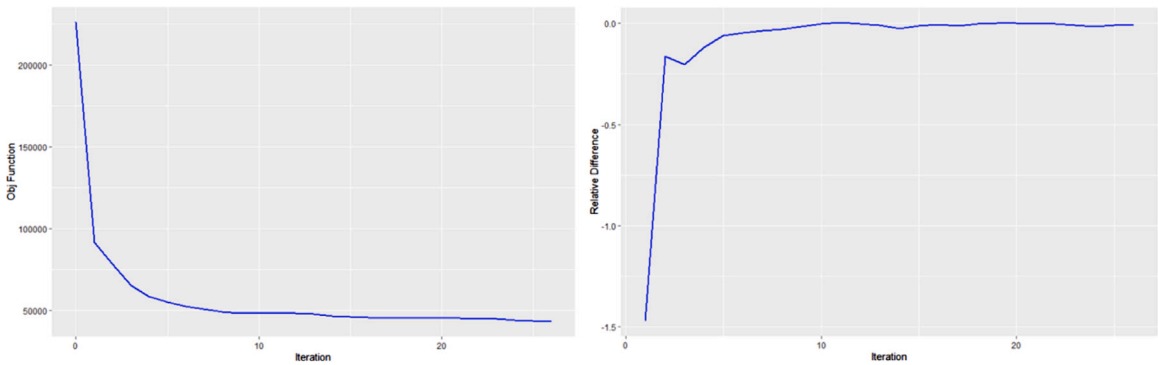


Fig. 5. Objective function and relative error for a selected experiment.

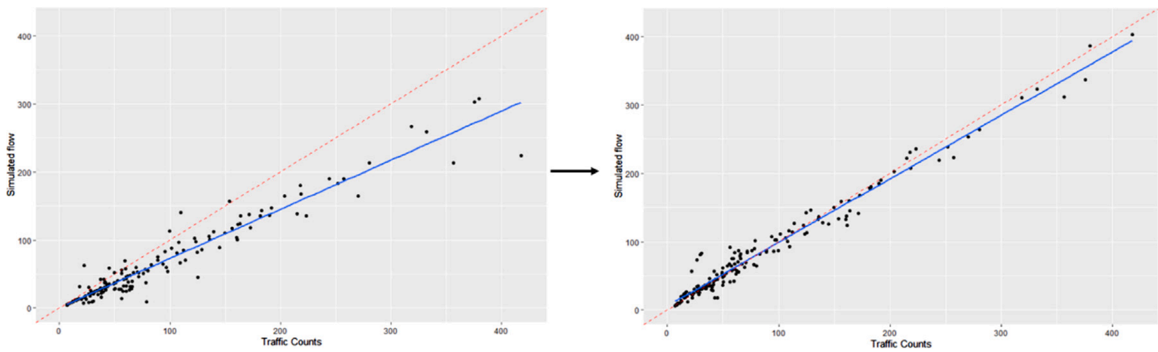


Fig. 6. R^2 improvement in a selected experiment.

3.2.4. Optimization procedure

As seen in Fig. 5, the convergence of the method is clear and fast. Moreover, in all of the 24 experiments we performed, the maximum number of iterations ($N_{\max} = 100$) was never reached.

Fig. 6 shows two versions of the linear regression between the traffic counts measured by sensors and their corresponding simulated values. On the left is the regression before optimization, when the estimated assignment matrix is already calculated but the OD values are not yet calibrated. On the right is the linear regression after computing the convergence of the minimization method and using Visum-SBA to incorporate the resulting assignment matrix from the assignment of the estimated OD matrix, $\mathbf{X}^*$.

The evolution shown in the figure validates the optimization method, which uses scaling factors instead of OD values as variables. The estimated scaling factors correctly adjust the flows in order to replicate the traffic counts measured in the network.

Finally, the use of scaling factors as optimization variables has also been studied. It has been done by comparing the estimated OD matrix with the one obtained by using the OD variables as variables on Eq. (13). That is, firstly scaling the OD flows to the Historical magnitude and then using a maximum descent method to solve the optimization problem. Concretely, the experiment performed used the historical OD matrix as seed, with a penetration rate of 10% and $w = 0$. The results of convergence and fitting are high, presenting a final $R^2 = 0.974$. However, and as expected, the MSSIM values with respect the ground truth OD matrix is $MSSIM(\mathbf{X}^{GT}, \mathbf{X}^*) = 0.5479$, significantly lower than the ones obtained by using the scaling factors, $MSSIM(\mathbf{X}^{GT}, \mathbf{X}^*) = 0.7980$. Therefore, we consider that the use of scaling factors not only reduces the number of variables but also permits to obtain a more stable solution in terms of similarity to the given seed.

3.3. Experimental design

Finally, a set of experiments using the synthetic network and data generated is used to assess the robustness and the sensitivity of the methodology described with regard to some factors. This design factors were:

- The penetration rate of the GPS technology, as a percentage of vehicles that are captured by the GPS sample: $5\%, 10\%$ and 15%
- The initial OD matrix for the minimization procedure $\left(\mathbf{X}^0 = [x^0_{ijr}]\right)$: As stated in Eq. (12), the seed OD matrix can be the historical ($\mathbf{X}^0 = \mathbf{X}^H$), the observed ($\mathbf{X}^0 = \hat{\mathbf{X}}$), or both OD matrices combined. In this case, the combination tries to fill in the empty cells of the observed OD matrix with information from a reliable historical OD matrix. The two tested combinations are the following:

$$x^0_{ijr} = f(\hat{x}_{ijr}, x^H_{ijr}) = \begin{cases} k \cdot \hat{x}_{ijr} & \text{when } \hat{x}_{ijr} > 0 \\ x^H_{ijr} & \text{otherwise} \end{cases} \tag{19}$$

$$x^0_{ijr} = f(\hat{x}_{ijr}, x^H_{ijr}) = \begin{cases} \hat{x}_{ijr} & \text{when } \hat{x}_{ijr} > 0 \\ x^H_{ijr}/k & \text{otherwise} \end{cases} \tag{20}$$

  where k is a factor that increases or reduces the number of trips in the seed OD matrix in order to approximate both magnitudes:

$$k = \frac{NT(\mathbf{X}^H(\hat{\mathbf{X}} > 0))}{NT(\hat{\mathbf{X}}(\hat{\mathbf{X}} > 0))} = \frac{NT(\mathbf{X}^H(\hat{\mathbf{X}} > 0))}{NT(\hat{\mathbf{X}})} \tag{21}$$

  These four matrices are named, respectively, *Hist, Obs, Comb1* and *Comb2*.
- The objective function for the minimization procedure. By using $w = 0$ or $w = 1$ in Eq. (13), the objective function may or may not include the discrepancy term regarding the historical OD matrix.

3.4. Results

The results of the full DODME procedure are summarized using different KPIs in Table 5. These four indicators are:

- R^2: The coefficient of determination of the regression line between traffic counts and the corresponding assigned traffic flows, after launching a DTA with the estimated OD matrix, $\mathbf{X}^*$.
- NT: The total number of trips in the estimated OD matrices.
- $MSSIM$ to GT: Measure of similarity between the ground truth OD matrix and the estimated OD matrix, calculated using Eq. (16).
- $MSSIM$ to Hist: Measure of similarity between the historical OD matrix and the estimated OD matrix, calculated using Eq. (16).

R^2 reaches very high values in all experiments, which means that the DODME procedure works very well as an optimization problem for adjusting traffic count measurements. R^2 is higher when $w = 1$ and as the penetration rate increases. However, depending on the seed OD matrix, there are no significant changes when $w = 1$.

The total number of trips (NT) in the estimated OD matrices is always near the ground truth total number of trips, which is $NT_{GT} = 8,300$ vehicles (see Table 1). Setting the seed OD matrix process seems to have no effect on the final number, so the scaling factors are able to adapt to the initial situation. What is more, the value of w has a negative impact on reducing the total number of trips when it is set to $w = 1$.

In terms of similarity, the contribution of $w = 1$ in obtaining better $MSSIM$ results is well known. Additionally, the seed OD matrix also impacts the final result. The best choice is the historical OD matrix, since it is the OD matrix with the largest MSSIM compared to the ground truth. However, *Comb1* and *Comb2* do not significantly improve these indicators. Generally, the estimated OD matrix presents a higher $MSSIM$ when comparing to the ground truth OD matrix, than when comparing to the historical. That means, especially when the seed is the historical OD matrix, that the resulting OD matrix has adapted its structure to the ground truth traffic conditions, by using the proposed methodology.

Table 5
Results of the experimental design.

Seed OD	Penetration rate	$w = 0$	$w = 1$	$w = 0$	$w = 1$	$w = 0$	$w = 1$	$w = 0$	$w = 1$
		R^2	R^2	NT	NT	$MSSIM$ to GT	$MSSIM$ to GT	$MSSIM$ to Hist	$MSSIM$ to Hist
Hist	5	0.9594	0.9870	8,601	8,037	0.7048	0.9521	0.7283	0.9114
	10	0.9808	0.9862	8,365	8,085	0.7980	0.9140	0.7928	0.8705
	15	0.9739	0.9867	8,280	8,005	0.8165	0.9397	0.8125	0.9032
Obs	5	0.9495	0.9846	8,446	8,019	0.5716	0.8318	0.5339	0.7802
	10	0.9734	0.9840	8,132	7,941	0.5974	0.8646	0.5671	0.8171
	15	0.9478	0.9866	8,302	8,027	0.5985	0.8461	0.5761	0.7926
Comb1	5	0.9401	0.9842	8,847	8,118	0.5940	0.8672	0.5859	0.8027
	10	0.9658	0.9810	8,407	8,100	0.6695	0.8694	0.6652	0.8096
	15	0.9658	0.9873	8,339	7,969	0.7605	0.9171	0.7474	0.8648
Comb2	5	0.9712	0.9845	8,257	8,117	0.7398	0.8712	0.6844	0.8066
	10	0.9741	0.9815	8,210	8,167	0.6122	0.8500	0.5747	0.7854
	15	0.9495	0.9855	8,230	8,001	0.7257	0.8877	0.6884	0.8275

Table 6
Results of dynamic Spiess and DDM algorithm.

		R^2	NT	$MSSIM$	Comp time
$w = 0$	DynSpiess	0.9992	8,380	0.8237	1 h 40 min
	DDAF 10%	0.9808	8,365	0.7980	1 h 15 min
$w = 1$	DynSpiess	0.9993	8,152	0.7658	1 h 30 min
	DDAF 10%	0.9862	8,085	0.9140	1 h 10 min

Globally, the 24 experiments summarized in this work show that the DODME procedure is able to calibrate the demand for a middle-sized network model in different situations. This new method requires combining two sources of data: traffic counts on certain links and GPS processed data, as estimated link travel times. However, in order to achieve better results, it is desirable to use a historical OD matrix as a seed OD matrix in the optimization step. The historical OD matrix should also be compared in the objective function by using the second term weighted with $w = 1$.

Moreover, when a reliable historical OD matrix is available, using it improves the results. It can be included as the seed OD matrix of the model; in the corresponding term can be added in the objective function; or both of these can be done. This is consistent with the hypothesis formulated by Ashok and Ben-Akiva (1993, 2002). They assume that when the historical OD matrix is at least reliable with respect to the structure of the mobility patterns, then it is worth information that it incorporates into the models.

3.4.1. Comparison to DODME analytical models

In Ros-Roca et al. (2020a), a dynamic version of the Spiess method was built and studied by the authors. This method is also applied to the currently discussed network, using the same historical OD matrix as the seed OD matrix in the optimization procedure. The main difference is that Dynamic Spiess requires a dynamic traffic assignment at each iteration, while the supplementary data, link travel times, is not needed. The quality of the results is compared to those obtained by DDAF.

The results of dynamic Spiess and DDAF at 10% of penetration rate are provided in Table 6:

While dynamic Spiess outperforms in R^2 and in the number of trips in the resulting OD matrices, the DDAF method improves the similarity of KPIs when $w = 1$, a situation where dynamic Spiess presents its worse results. Moreover, all the procedures have similar computational times, although DDAF takes advantage of available data while dynamic Spiess cannot include it in the OD matrix estimation.

Furthermore, it has not been taken into account that the main requirement to use a dynamic Spiess is that the network must be previously calibrated for launching a dynamic traffic assignment, from which the assignment matrix is calculated and is always a very time-consuming task. In contrast, the proposed DDAF methodology does not require a fully calibrated model since the necessary information, link travel times, can be estimated from the data.

4. Real experiment

Despite the already mentioned drawbacks of the available physical measurements in Section 1.3 and considering our experience with the synthetic data, we – for the sake of completeness – conducted a further test on a real network using commercial GPS data to infer estimated link travel times. The selected exercise is useful for practitioner's point of view since a ground truth OD matrix is not available, but only traffic counts, a historical OD matrix and estimated link travel times obtained from GPS commercial data. It is downtown Turin, in Italy.

Fig. 7. The Turin network used with the detection layout.

Table 7
Turin network and OD characteristics.

Time periods	4
Zones	221
Detectors	302
OD pairs X time	≈195k
Historical trips	129k
Historical positive OD	≈40k (39.71%)

4.1. Turin's network

The network used is shown in Fig. 7, with its characteristics presented in Table 7. The detection layout comprises 302 counting stations, situated in the network as shown and prioritizes traffic counts over main streets. Furthermore, there is a historical OD matrix with high level of confidence, since it is the estimated OD matrix of a previous study.

The available mobility data for this network is a sample of GPS tracking data provided by INRIX that contains one year of indistinguishable private and fleet vehicles circulating on labour days during the peak period in the morning. The sample contains 3.76M waypoints, which represents 232k different trips (partially) circulating on the network. The penetration rate is 1.32% if we compare yearly GPS trips to the historical OD matrix's for the selected period in terms of the number of trips. Moreover, this incompleteness reflects a mobility pattern that does not correspond to the network's OD pattern, as shown by further analysis of the matrix structure.

Since the data set does not provide information regarding the vehicle types, we cannot distinguish between fleet and regular vehicles nor filtering and cleansing the sample to obtain estimations of the link travel times with a controlled degree of confidence. However, an estimation of link travel times for each time period of study has been made using the methodology described in Appendix A.

4.2. Results

Because this is a real network experiment, the ground truth conditions are unknown and it is therefore impossible to compare the resulting OD matrix to the ground truth OD matrix. The reference OD matrix in this case is the available historical OD matrix, as it is a reliable OD matrix from a previous study.

The observed OD matrix is built from the GPS data set by aggregating the trips according to their origin and destination zones, as well as their departure times. The obtained OD matrix, $\hat{\mathbf{X}}$, has 35k positive OD values, which are 17.76% of the OD values. The term S of the $MSSIM$ index of Eq. (15) between the observed and the historical OD matrices is an appropriated indicator of whether

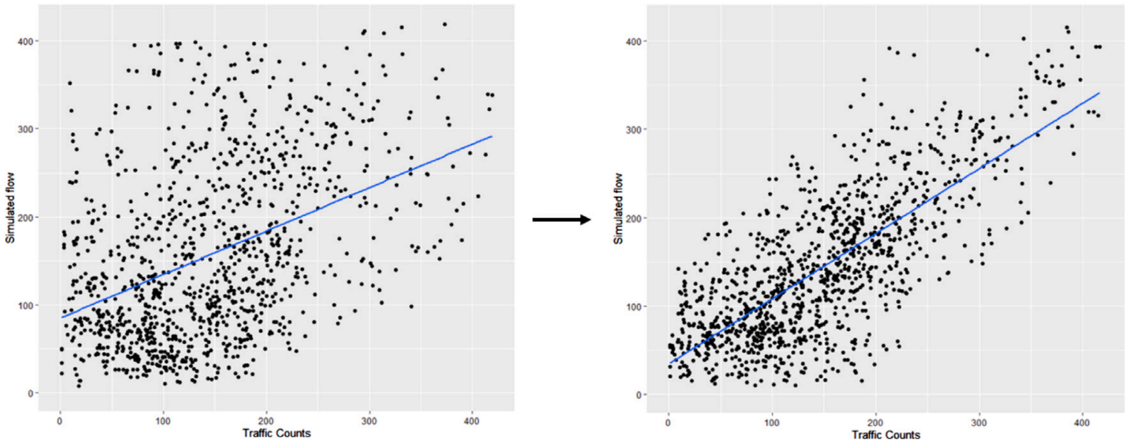


Fig. 8. R^2 improvement after DDAF method.

the observed OD matrix is a good seed choice for the optimization procedure, since it indicates how similar these matrices are. In this case, the term $S(\hat{\mathbf{X}}, \mathbf{X}^H) = 0.0287$, thus indicating it is not suitable as a seed OD matrix, since the captured mobility pattern is not similar to that of the more reliable historical OD matrix. The presence of fleet vehicles and the unknown percentage of them is the main cause for discarding this observed OD matrix as a seed, since fleet vehicles have neither a fixed origin nor destination.

The DDAF method converged after 45 iterations, and the linear regression between the traffic counts of sensors and the corresponding simulated values is included in Fig. 8. The corresponding simulated traffic counts have been obtained by launching a DTA with the resulting estimated OD matrix. As shown, the optimization procedure increases the fitting of these measurements, from $R_0^2 = 0.3261$ to $R_f^2 = 0.7266$.

The total number of trips resulting from the estimated OD matrix are $NT = 112931$ and the positive OD values are 71983 (36.85%). The similarity between the historical and the estimated OD matrix measured with $MSSIM$ is $MSSIM\left(\mathbf{X}^*, \mathbf{X}^H\right) = 0.5167$.

These figures show a real-life application of the proposed methodology when GPS data and traffic counts are available and produces an improved estimated OD matrix consistent to the historical matrix and satisfying observed counts. Observed OD matrices from commercial providers does not directly reflects OD pattern and direct use for transportation analysis after a DODME Spiess-based adjustment should be avoided.

5. Conclusions

The components shared by most of the DODME approaches are: an assignment matrix $\mathbf{A} = [a_{ijr}^{lt}]$ whose elements represent the proportion of the OD demand x_{ijr} travelling from origin i to destination j, departing the trip at time period r and reaching the counting station on link l at time t; a historical OD matrix $\mathbf{X}^H$ that provides additional information on the mobility patterns, namely their space–time structure; and link flow counts, $\hat{y}_{lt}$, for a subset of links $l \in \hat{L} \subseteq L$ where counting stations are located. The assignment matrix describing the dynamic structure of the temporal use of the network is usually costly to obtain using dynamic assignment procedures that depend on an initial OD matrix. This initial matrix is not always reliable and in addition, the DODME applications are limited by the significant computational effort that they require to rely on a good network calibration.

This paper is driven by a desire to investigate whether the data provided by new ICT sources (namely GPS data) could empirically provide better estimates in DODME practical applications. In other words, our general hypothesis assumes that the data contain information about the generating phenomenon, by which we aim to specifically find a suitable mean to process the data and incorporate the information into the DODME process.

Despite the results using a real network of the Turin downtown are very promising, there is still margin to improve. The exploratory analysis of the GPS tracking from the professional providers reveals that data collection methods are tailored to commercial goals and that nowadays this data is not able to generate a reliable initial OD matrix for a traditional DODME process. However, after the synthetic experiments, with a synthetic data generation process, exposed that validate the procedure, the authors think that a well-designed data collection from on-board GPS for the specific DODME purpose could generate a waypoints database that add enhanced information about the real mobility pattern by adding reliable estimated link travel times and an estimated dynamic assignment matrix.

The computational results prove that the proposed methodology is reliable for estimating the key component: the assignment matrix. The estimated assignment matrix is consistent, as well as the empirical route choice set approach to identify the most likely used paths, their travel times, and the path flow proportions from which the assignment matrix is derived.

The experimental design based on synthetic data allows conducting a sensitivity analysis of the equipped vehicles' penetration rate. Furthermore, this analysis could be difficult to perform when using only pure empirical data. The sensitivity analysis shows the robustness of the assignment-free approach, and, as expected, an increase in the quality of the results when the percentage of equipped vehicles increases.

The computational performance is very good, converging very quickly in few iterations to achieve good results. A relevant point to highlight is the consistent performance regarding the total number of trips and the structure of the final estimated matrix. Although Djukic (2014) and Behara et al. (2018) provide a remarkable improvement including the analysis of the structural similarity, we found that the number of trips should also be considered since in some cases, as the computational results show, the increase or decrease to fit the measured flows could be consistent with the underlying transportation phenomenon.

Quite frequently, R^2 is very good indicator for a simply meta-regression model, but at the price of increasing or decreasing the number of trips or destroying historical OD trip matrix pattern. In order to fit the observed flow count in a link, it pulls forwards to and backwards from the OD pairs whose paths use that link. However, considering the underlying physical system (i.e., the transportation system), the resulting estimated OD matrix may not be very realistic because some affected OD pairs are forced to generate or attract an unrealistic number of trips. The developed assignment-free DODME outperforms on matrix similarly, nevertheless is also based on obtaining a good quality of the R^2 that explains the link flows, but it exhibits remarkable stability in total number of vehicles when comparing the ground truth OD to the resulting estimated OD. A high degree of structural similarity also exists between both matrices. Therefore, we can conclude that the estimated OD is more reliable than those obtained by other approaches.

As with all data-driven approaches, the quality of the results strongly depends on the quality of the data used. The computational results reported in Table 5 lead us to conclude that the GPS-observed OD may be unreliable due to the flaws in and drawbacks of the empirical data. On the other hand, when the estimated link travel times are reliable (which is frequently the case, in practice, with professional data), then the quality of the estimated assignment matrix is good enough for the DDAF DODME approach to be applied soundly. Using only measured data is recommended when the data quality ensures observed OD matrices, which could perhaps be achieved when using purpose-oriented commercial data (per agreement with the data providers). This would thus overcome the mentioned drawbacks.

5.1. Future research

Data-driven approaches open the door to many new alternatives that can be explored, such as when penetration rates are not uniform for all TAZ may instead depend on socioeconomic characteristics. However, the designed methodology to estimate the assignment matrix use information from the vehicles circulating through the network, without an interest about their origin or destination.

Other formulations of DODME in an assignment-free approach can be applied where assignment matrices are still necessary, as in the case of Kalman Filtering, which may provide short-term real time estimates of the OD matrix based on observed data useful for traffic management purposes.

CRediT authorship contribution statement

Xavier Ros-Roca: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Lídia Montero:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Jaume Barceló:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Klaus Nökel:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – review & editing. **Guido Gentile:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – review & editing.

Acknowledgements

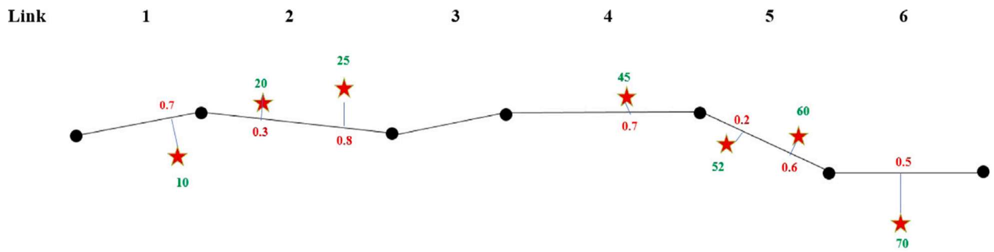
This research was funded by the PTV Group, Germany, by TRA2016-76914-C3-1-P Spanish R+D Programs, by the Secretaria d'Universitats i Recerca Generalitat de Catalunya, Spain 2017 SGR- 1749, and by the Industrial PhD Program, Spain 2017 DI 041. We also want to acknowledge the inspiring discussions with Frauke Petersen, from PTV AG, and Alessandro Attanasi, from PTV Sistema.

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors.

All authors approved the version of the manuscript to be published.

Table 8
GPS waypoints data sample.

ID	Date	Timestamp	Latitude	Longitude
4261353	2019-11–30	22:43:58	45.445988	9.1244048
4261353	2019-11–30	22:44:27	45.445496	9.1241952
....	....	....	....	....
4261353	2019-11–30	22:50:57	45.444767	9.1192517
4261355	2019-11–30	22:43:58	45.445980	9.1247048
4261355	2019-11–30	22:44:27	45.445574	9.1192821
....	....	....	....	....
4261355	2019-11–30	22:50:57	45.444767	9.1197541
....	....	....	....	....



$$tt_3 = \frac{l_3}{0.2l_2+l_3+0.7l_4}20\text{s} \quad \text{(a)} \qquad tt_4 = \frac{0.7l_4}{0.2l_2+l_3+0.7l_4}20\text{s} + \frac{0.3l_4}{0.3l_4+0.2l_5}7\text{s} \quad \text{(b)}$$

Fig. 9. Example of the interpolation of travel times according to the waypoint sequence.

Appendix A. Map matching for estimation of link travel times

When tracking the trajectories of GPS equipped vehicles across a network, the data collected by GPS devices are usually formatted as sets of trips, as shown in Table 8. They are detailed by recorded information for an ordered sequence of waypoints, $(ID_k, ts_{kl}, lat_{kl}, long_{kl})$, in which each trip k has a trip identity ID_k; the date is the time tag $ts_{(k,l)}$ for the lth observation; and the latitude and longitude of the current position are indicated.

This data as it is cannot be used for transportation analysis and should be map-matched onto the network of the scenario being analysed. The map-matching process transforms sequences of waypoints to paths onto the network, (PTV Vissim, 2020b). Firstly, each trajectory is map-matched using a Viterbi approach, (Viterbi, 1967), where each waypoint is assigned to a certain point on the nearest link of the network. Fig. 9 shows how this works, using an example in which the red stars are the waypoints and the red numbers near the links are the relative position of the waypoint projection onto the target link. Timestamps for waypoints are depicted in green. The first and last link are not fully covered by time information and are thus dropped from the link sequence resulting from the GPX trajectory mapping.

Next, link travel times are estimated from waypoint timestamps according to their sequence. For all the links in the sequence, the interpolated travel time for a link is the sum of the timestamp differences of two consecutive waypoints mapped in the target link. In the case of two consecutive waypoints that are not wholly projected within one link, the distance-based fraction within the link is taken (l_k is the length of link k in Fig. 9).

For instance, the travel time for link l_3 can be estimated taking into account that the travel time for trip between the 3rd and 4th waypoints is 20 s, and this time is the estimated travel time of the whole link l_3 plus a 0.2 fraction of l_2 and a 0.7 fraction of l_4 (equation (a) in Fig. 9). The estimated travel time of link l_4 is obtained by adding two parts, equation (b). The first part is the travel time proportion between the 3rd and 4th timestamps in link l_4 (adding 0.7 of l_4 to 0.2 of the length of link l_2 plus the entire length of link l_3). The second part is estimated directly from the proportion of link l_4 lying between 4th and 5th timestamps (a fraction of 7 s calculated as 0.3 of the l_4 distance within the total distance between 4th and 5th waypoints: $0.3l_4 + 0.2l_5$).

Finally, once all the waypoint sequences are converted to several paths with full details at the link level, the link travel times are averaged. The outcome of this process is the set of observed link travel times at each time period t: $\hat{tt}_{lt}, \forall l \in L, \forall t \in T$ for all links in the network that are used by the GPS tracking. This is the data set of observed link travel times.

Despite the huge quantity of trajectories introduced into the network, depending on the GPS data's penetration rate among the population, the GPS sample may uncover links. Moreover, the procedure that infers link travel times can produce non-feasible values when link travel times are below free-flow link travel time. In these situations, scaled travel time is used:

$$\hat{tt}_{l't} = R \cdot tt_{0l'} \quad , \quad R = mean_{l \in GPS}\left(\frac{\hat{tt}_{lt}}{tt_{0l'}}\right) \tag{22}$$

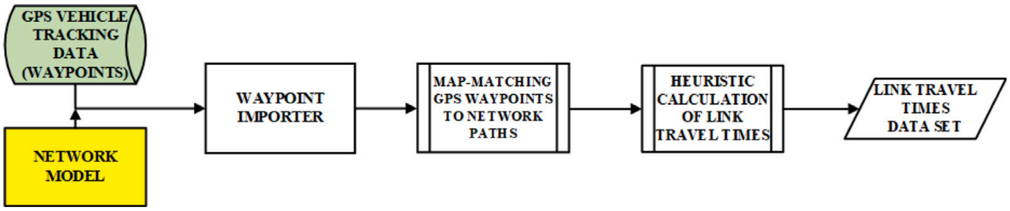


Fig. 10. Conceptual methodological approach to importing waypoints into a Visum model and using them to estimate link travel times.

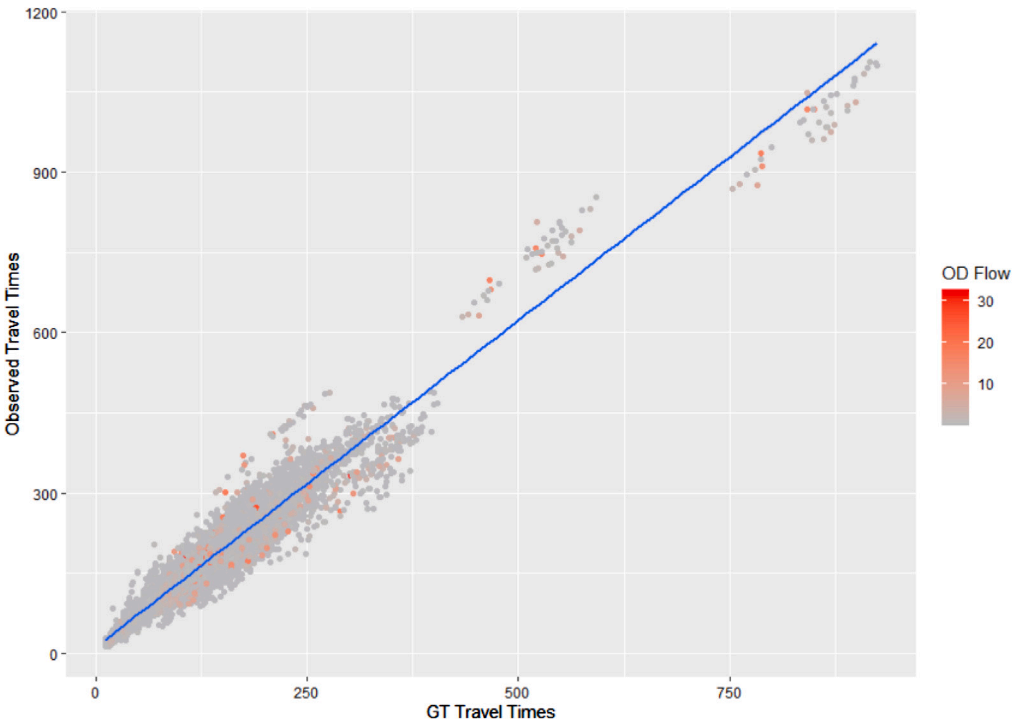


Fig. 11. Comparison of observed OD travel times with ground truth OD travel times.

where tt_{0l} is the free-flow travel time at each link and R is computed using all observed link travel times and their corresponding free-flow travel times. R is then the arithmetic mean of the expanding factors found for each link, which can be understood as a global expanding factor by the congestion effect.

The methodological process for generating the observed link travel times data set is summarized in Fig. 10.

Once the GPS tracking data collection process has gathered the waypoints and they are matched to paths in the target network through a suitable map-matching procedure, the results provide the input for the heuristic calculation of the time-dependent link travel times.

As a validation of the estimation of link travel times, the synthetic network used on the paper has also been used as a test site for this appendix. In this sense, once the GPS sample is map matched, the heuristic obtains estimations of link travel times at each time period.

As stated before, a large amount of data should ensure reliable estimations by averaging those travel times at each link and time interval.

In order to check their reliability, they must be compared with the path travel times obtained from the ground truth OD matrix. All the path travel times for each OD pair at each different time interval are collected and the mean by OD pair and time interval has been done, obtaining the OD travel times. The comparison is shown in Fig. 11.

The correspondence of both measurements is high, because the fit is $R^2 = 0.9445$ after removing some outliers that represent only 0.31% of total OD flow. These results ensure that, when the GPS sample is appropriately filtered, cleansed and does not present bias regarding the OD pairs, the proposed methodology to estimate link travel times is reliable, since using them to construct different OD paths present similar travel times as in ground truth conditions.

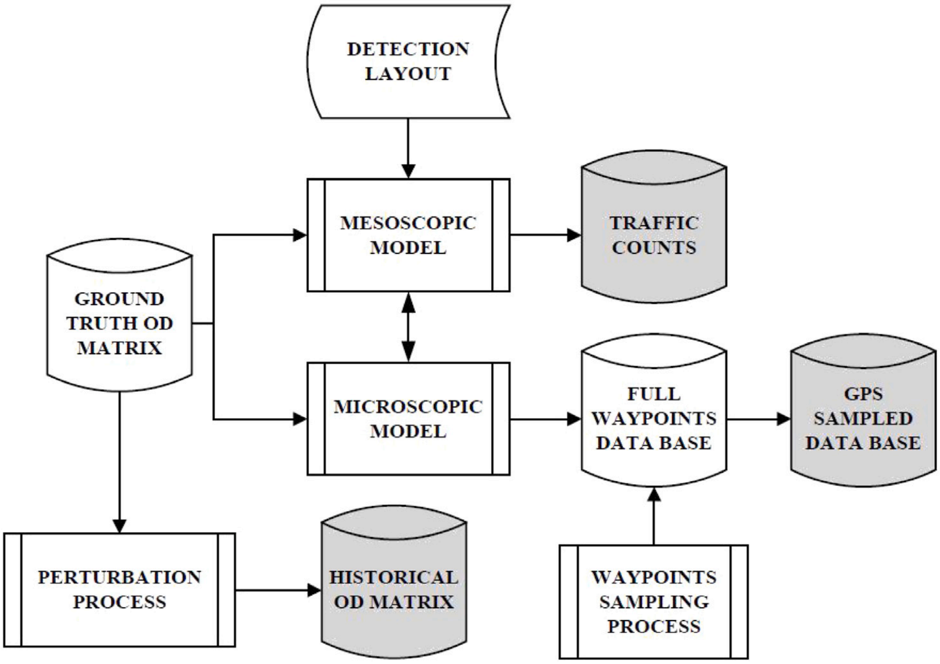


Fig. 12. Methodological scheme of the synthetic data generation for computational testing.

Appendix B. Data generation

In this appendix, the synthetic data generation for an OD estimation process with GPS data is detailed. Following the framework of Antoniou et al. (2016), this algorithmic methodology permit to obtain a set of traffic counts on determined points of the network, $\hat{y}_{lt},\ l \in \hat{L} \subseteq L,\ \ \forall t \in \mathcal{T}$, a reference OD matrix, $\mathbf{X}^H = [x_{ijr}^H]$, and filtered and cleansed GPS data, as formatted in Table 8, that using Appendix A or other tools, can be transformed to estimations of link travel times.

The methodology proposed is supported by a combined mesoscopic and microscopic model of the selected test network. A properly calibrated dynamic assignment model for a given OD matrix is used as a ground truth OD matrix, $\mathbf{X}^{GT}$, which generates the link flow counts that emulate the physical counting. The set of equilibrium paths that are imported into the microscopic model will enable tracking individual vehicles and the GPS data collection in terms of waypoints.

Counting stations that are suitably located in the network collect the observed traffic counts; therefore, the detection layout is another aspect that must be considered when generating the synthetic data. For the synthetic experiments conducted in this paper, the detection layout (where detectors are placed) adapts the first phase of Barceló et al. (2012)'s detection layout procedure, which consists of a greedy algorithm whose suboptimal solution identifies the detection layout while maximizing coverage of the OD demand in terms of link and path flows.

The methodological process for generating the synthetic data is conceptually summarized in Fig. 12. A microsimulation model of the selected site is run along with a vehicle tracking procedure for generating vehicle-tracking data that is similar to those physically collected from GPS devices, with the same format shown in Table 8.

Traffic counts are commonly obtained in practice from averaging traffic counts from several days showing the same traffic conditions: a short sample of days is necessary since they are fairly stable. GPS data has to be representative of the same traffic conditions accounting for a larger number of days to cope with low representativeness of GPS equipped data with respect the whole population of drivers. This process generates a waypoints database that is large enough to emulate reality (i.e., data collected for the average number of working days in a year). These data are collected for the whole population over very short time periods (i.e., 0.1 s), and they enable defining the sampling processes with a variety of design factors such as the GPS technology's penetration rate percentage and the recording latency for each vehicle, depending on the time distribution between successive draws that determines the sample size. The frequency of recording waypoints has been assigned differently to each vehicle, following an empirical distribution of latencies from an INRIX real GPS data set of another network.

Once these factors are set, the data set is reduced to a sample of waypoints for each computational experiment, which emulates the GPS data that is received from the GPS data provider.

Finally, the historical OD matrix is generated from the ground truth OD matrix by following Antoniou et al. (2016)'s procedure, that is:

$$x_{ijr}^H = x_{ijr}^{GT}(p + q \cdot \varepsilon) \quad , \quad p = 0.75 \quad , \quad q = 0.15 \quad , \quad \varepsilon \sim N(0, 1/3) \quad , \quad \forall i, j, r \tag{23}$$

The historical OD matrix is on average decremented by 25% with a random perturbation. This perturbation tries to emulate a realistic historical OD matrix from surveys and past projects that represent similar traffic conditions.

References

Aimsun, 2017. Aimsun Next 8.2 - User's Manual. Aimsun, Barcelona, Spain.

Antoniou, C., Azevedo, C., Lu, L., Pereira, F., Ben-Akiva, M., 2015. W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. Transp. Res. Procedia 59, 129–146.

Antoniou, C., Ciuffo, B., Montero, L., Casas, J., Barceló, J., Cipriani, E., Djukic, T., Marzano, V., Nigro, M., Bullejos, M., Perarnau, J., Breen, M., Punzo, V., Toledo, T., 2016. A framework for the benchmarking of OD estimation and prediction algorithms. Transp. Res. C http://dx.doi.org/10.1016/j.trc.2015.08.009.

Ashok, K., Ben-Akiva, M., 1993. Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems. In: Daganzo, C. (Ed.), Transportation and Traffic Theory. Proceedings of the 12th ISTTT. Elsevier Science Publishing Company Inc.

Ashok, K., Ben-Akiva, M., 2002. Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows. Transp. Sci. 36 (2), 184–198.

Balakrishna, R., 2006. Off-Line Calibration of Dynamic Traffic Assignment Models (Ph.D. thesis). MIT, USA.

Barceló, J., Gilliéron, F., Linares, M.P., Serch, O., Montero, L., 2012. Exploring link covering and node covering formulations of detection layout problem. Transp. Res. Rec. 2308, 17–26.

Barceló, J., Montero, L., Bullejos, M., Serch, O., Carmona, C., 2013. A kalman filter approach for exploiting bluetooth traffic data when estimating time-dependent od matrices. J. Intell. Transp. Syst. Technol. Plan. Oper. 17 (2), 123–141.

Behara, K.N.S., Bhaskar, A., Chung, E., 2018. Classification of typical bluetooth OD matrices based on structural similarity of travel patterns - case study on Brisbane. Transportation Research Record, 97th Annual Meeting.

Behara, K.N.S., Bhaskar, A., Chung, E., 2020a. A novel approach for the structural comparison of origin–destination matrices: Levenshtein distance. Transp. Res. C 111, 513–530.

Behara, K.N.S., Bhaskar, A., Chung, E., 2020b. A novel methodology to assimilate sub-path flows in bi-level OD matrix estimation process. IEEE Trans. Intell. Transp. Syst. 1–11.

Ben Akiva, M., Bierlaire, M., 1999. Discrete choice models and their application to short term travel decisions. In: Hall, R.W. (Ed.), Handbook of Transportation Science. Kluwer, ISBN: 0-7923-8587-X (Chapter 1).

Ben Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H.N., Mishalani, R., 2001. Network state estimation and prediction for real-time traffic management. Netw. Spat. Econ. 1, 293–318.

Bierlaire, M., 2002. The total demand scale: A new measure of quality for static and dynamic origin–destination trip tables. Transp. Res. B 36 (9), 837–850.

Bovy, P., Bekhor, S., Prato, C., 2008. The factor of revisited path size. Transp. Res. Board 2076, 132–140. http://dx.doi.org/10.3141/2076-15.

Cantelmo, G., Cipriani, E., Gemma, A., Nigro, M., 2014a. An adaptive Bi-level gradient procedure for the estimation of dynamic traffic demand. IEEE Trans. Intell. Transp. Syst. 15 (3), 1348–1361.

Cantelmo, G., Viti, F., Cipriani, E., Nigro, M., 2014b. A utility-based dynamic demand estimation model that explicitly accounts for activity scheduling and duration. Transp. Res. Procedia 23, 440–459.

Cascetta, E., 2001. Transportation System Engineering: Theory and Methods. Kluwer, ISBN: 0-7293-6972-8.

Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A., 1996. A modified logit route choice model overcoming path overlapping problems. In: Proceedings of the 13th International Symposium on the Theory of Road Traffic Flow.

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

Chabini, I., 1998. Discrete dynamic shortest path problems in transportation applications: Complexity and algorithms with optimal run time. Publication in Transportation Research Records 1997.

Djukic, T., 2014. Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management (Ph.D. thesis). Delft University of Technology (Chapter 6).

Djukic, T., Hoogendoorm, S., Van Lint, H., 2013. Reliability assessment of dynamic OD estimation methods based on structural similarity index. Transportation Research Board 92nd Annual Meeting.

Djukic, T., Masip, D., Breen, M., Perarnau, J., 2018. Heuristic-based framework for dynamic OD demand estimation in the congested networks. Transportation Research Board 97th Annual Meeting.

Frederix, R., Viti, F., Corthout, R., Tampère, C., 2011a. New gradient approximation method for dynamic origin-destination matrix estimation on congested networks. Transp. Res. Rec.: J. Transp. Res. Board 2263 (1), 19–25.

Frederix, R., Viti, F., Tampère, C., A density-based dynamic OD estimation method that reproduces within-day congestion dynamics. In: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, vol. 1, pp. 664–699.

Frederix, R., Viti, F., Tampère, C., 2013. Dynamic origin-destination estimation in congested networks: Theoretical findings and implications in practice. Transp. A: Transp. Sci. 9 (6), 494–513.

Janmyr, J., Wadell, D., 2018. Analysis of Vehicle Route Choice During Incidents (MSc thesis). University of Linköping, Department of Science and Technology, LiU-ITN-TEK-A–18/020-SE.

Kostic, B., Gentile, G., 2019. Using traffic data of various types in the estimation of dynamic O-D matrices. In: 2015 International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2015, pp. 66–73.

Krishnakumari, P., Lint, H.Van., Djukic, T., Cats, O., 2019. A data driven method for OD matrix estimation. Transp. Res. C http://dx.doi.org/10.1016/j.trc.2019.05.014.

Lopez, C., Krishnakumari, P., Leclercq, L., Chiabaut, N., Van Lint, H., 2017a. Spatiotemporal partitioning of transportation network using travel time data. Transp. Res. Record: J. TRB 2623, 98–107.

Lopez, C., Leclercq, L., Krishnakumari, P., Chiabaut, N., Van Lint, H., 2017b. Revealing the day-to-day regularity of urban congestion patterns with 3D speed maps. Sci. Resports 2 (7), 14029.

Lundgren, J.T., Peterson, A., 2008. A heuristic for the bi-level origin-destination matrix estimation problem. Transp. Res. B 42 (4), 339–354.

Mo, B., Li, R., Dai, J., 2020. Estimating dynamic origin–destination demand: A hybrid framework using license plate recognition data. Comput.-Aided Civ. Infrastruct. Eng. 35 (7), 734–752.

Morales, J.L., Nocedal, J., 2011. Remark on algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization. ACM Trans. Math. Software 38 (1), Article N.7.

Nassir, N., Ziebarth, J., Sall, E., Zorn, L., 2014. Choice set generation algorithm suitable for measuring route choice accessibility. Transp. Res. Rec. 2430 (1), 170–171.

Nigro, M., Cipriani, E., del Giudice, A., 2018. Exploiting floating car data for time-dependent origin–destination matrices estimation. J. Intell. Transp. Syst.: Technol. Plan. Oper. 22 (2), 159–174.

OpenLR, 2020. OpenLR White Paper. Version 1.5, revision 2. https://www.openlr-association.com/fileadmin/user_upload/openlr-whitepaper_v1.5.pdf.

Ortúzar, J., Willumsen, L., 2011. Modelling Transport. Wiley Ed.

Osorio, C., 2019. Dynamic origin–destination matrix calibration for large-scale network simulators. Transp. Res. C 98, 186–206.

PTV AG Vissim, 2020a. PTV Vissim 2020 - User's Manual. PTV Group, Karlsruhe, Germany.

PTV AG Visum, 2020b. PTV Visum 2020 - User's Manual. PTV Group, Karlsruhe, Germany.

Ros-Roca, X., Montero, L., Barceló, J., 2020a. Investigating the quality of spiess-like and SPSA approaches for dynamic OD matrix estimation. Transp. A: Transp. Sci. 17 (3), 235–257.

Spiess, H., 1990. A Gradient Approach for the OD Matrix Adjustment Problem. vol. 693, Centre for Research on Transportation, University of Montreal, Canada, pp. 1–11.

Toledo, T., Kolechkina, T., 2013. Estimation of dynamic origin-destination using linear assignment matrix approximations. IEEE Trans. Intell. Transp. Syst. 14 (2), 618–626.

Van Zuylen, H.J., Willumsen, L.G., 1980. The most likely trip matrix estimated from traffic counts. Transp. Res. B 14 (3), 281–293.

Ruiz de Villa, A., Casas, J., Breen, M., 2014. OD matrix structural similarity: Wasserstein metric. TRB 93rd Annual Meeting Compendium of Papers.

Viterbi, A.J., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans. Inform. Theory 13 (2), 260–269.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612.

Wang, Z., Simoncelli, E.P., 2008. Maximum differentiation (MAD) competition : A methodology for comparing computational model of perceptual quantities. J. Vis. 8 (12), 1–13.

Yang, Z., Lu, Y., Hao, W., 2017. Origin-destination estimation using probe vehicle trajectory and link counts. J. Adv. Transp. http://dx.doi.org/10.1155/2017/4341532.

Yang, H., Sasaki, T., Iida, Y., Asakura, Y., 1992. Estimation of origin–destination matrices from link traffic counts on congested networks. Transp. Res. B 26 (6), 417–434.