AN ASSIGNMENT FREE DATA DRIVEN APPROACH TO THE DYNAMIC ORIGIN DESTINATION MATRIX ESTIMATION PROBLEM

X. Ros-Roca[a,b], J. Barceló[a,b], L. Montero[a],
[a]Department of Statistics and Operations Research, Universitat Politècnica de Catalunya
[b]PTV GROUP

Abstract

Dynamic traffic models require dynamic inputs, and one of the main inputs are 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 hard problem since no direct full 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 bilevel optimization problem has been widely used. This formulation solves at the upper level a nonlinear optimization 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 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 on used paths could be interpreted as an empirical assignment observed de facto. This allows extracting empirically the same information provided by an assignment that is used in the analytical approaches. This research report explores how to extract such information from the recorded data, proposes a new optimization model to solve the DODME problem, and computational results on its performance.

Key Words: Dynamic OD Estimation, Bi-level Optimization, Non-linear optimization, ICT Data

1. Introduction: Analytical Approaches to DODME

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, where they also must be dynamic, or at least time discretized, to properly approximate the time variability of the demand. OD matrices are not yet observable; in the best case, the measurements from Information and Communication Technologies (ICT), as GPS vehicle tracking, or mobile phones Call Detail Records (CDR), allow drawing samples that must be suitably expanded to provide estimates of the whole population. Therefore, their estimation must be done resorting to indirect process, usually based on mathematical models.

One of the most appealing mathematical formulations of the OD estimation problem is in terms of bilevel optimization problems (1), aimed at adjusting an initial target OD, $\boldsymbol{X^H}$, so that it could explain the observed link flow counts $\widehat{\boldsymbol{Y}}$ at counting stations in the network, Ros-Roca et al. (2018).

$$min\, Z(\boldsymbol{X},\boldsymbol{Y}) = w_1F_1\left(\boldsymbol{X},\boldsymbol{X^H}\right) + w_2F_2\left(\boldsymbol{Y},\widehat{\boldsymbol{Y}}\right) \quad (1)$$

$$\boldsymbol{Y} = Assignment(\boldsymbol{X})$$

$$\boldsymbol{X} \geq 0$$

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 $\boldsymbol{X^H}$ and $\widehat{\boldsymbol{Y}}$, respectively. The underlying hypothesis is that $\boldsymbol{Y}(\boldsymbol{X})$ are the link flows predicted by assigning the demand matrix $\boldsymbol{X}$ onto the network, which can be expressed by a proportion of the OD demand flows passing through the count location at a certain link. In terms of the assignment matrix $\boldsymbol{A}(\boldsymbol{X})$, which is the proportion of OD flow that contributes to a certain link traffic count, is:

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{X})\boldsymbol{X} \qquad (2)$$

Then the resulting bilevel optimization problem solves (at the upper level) the nonlinear optimization problem by substituting the estimated flows $\boldsymbol{Y}$ in the objective function of (1) with the relationship (2):

$$min\, Z(\boldsymbol{X},\boldsymbol{Y}) = w_1 F_1\left(\boldsymbol{X},\boldsymbol{X^H}\right) + w_2 F_2\left(\boldsymbol{A}(\boldsymbol{X})\boldsymbol{X},\widehat{\boldsymbol{Y}}\right) \qquad (3)$$

$$\boldsymbol{X} \geq 0$$

To estimate a new assignment matrix $\boldsymbol{X}$, while at the lower level, a static user equilibrium assignment is used to solve the assignment problem $\boldsymbol{Y} = Assignment(\boldsymbol{X})$ in order to estimate the assignment matrix $\boldsymbol{A}(\boldsymbol{X})$ induced by the new $\boldsymbol{X}$. (Spiess 1990) is a good example of a seminal model based on this approach.

This mathematical model is highly undetermined since the number of variables, OD pairs, is much larger than the number of equations. Link flow counts available at a subset of links in the network, along with the distance functions in the objective function, are the additional information aimed at reducing the degree of indetermination, in the most frequent implementations, using a static traffic assignment to solve the lower level problem, Spiess (1990), Codina and Montero (2006), Lundgren and Peterson (2008). The main reason for these implementations is that they are algorithmically efficient and present nice properties for convergence and stability. However, since static assignment models support them, they cannot properly account for the impacts of traffic dynamics and the induced congestions.

2. Extensions of Analytical Formulations to Account for Time Dependencies

Assuming that the functional dependency between the estimated flows $\boldsymbol{Y}$, the assignment matrix $\boldsymbol{A}(\boldsymbol{X})$ and the estimated matrix $\boldsymbol{X}$, set up in (2), allows a Taylor expansion around the current solution which provides a more detailed insight of the how the path flows contribute to the link flows, which is in essence the information provided by the assignment matrix. This improved approach was explored in Lundgren and Peterson (2008) still using a static assignment. Other researchers, Frederix, et al. (2013), Toledo and Kolechkina (2013), or Yang et al. (2017) proposed to use a Dynamic Traffic Assignment at the lower level to account for time dependencies, and therefore for congestion building processes. That allows a richer Taylor expansion also in terms of time, which captures these phenomena. To properly reformulate (2), let's assume that:

- I is the set of Origins, J the set of Destinations and $N \coloneqq I \times J$ the set of OD pairs.
- $\mathcal{T} = \{1, \dots, T\}$ is the set of time intervals.
- L is the set of links in the network. $\hat{L} \subseteq L$ is the subset of links that have sensors.
- $\hat{y}_{lt}$ are the measured flow counts at link l during time period t. y_{lt} are the corresponding simulated flow counts, $\forall l \in \hat{L} \subseteq L$ and $\forall t \in \mathcal{T}$. $\boldsymbol{Y} = (y_{lt})$ and $\widehat{\boldsymbol{Y}} = (\hat{y}_{lt})$ are link flow counts in vector form.
- x_{ijr} are the OD flows for n-th OD pairs departing during time period r, $\forall i \in I, \forall j \in J$ and $\forall r \in \mathcal{T}$. $\boldsymbol{X} = \left(x_{ijr}\right)$ are the OD flows in vector form.

- a_{ijr}^{lt} is the flow proportion of the (i,j)-OD pair departing at time period $r \in \mathcal{T}$ and captured by link $l \in \hat{L}$ at time period $t \in \mathcal{T}$. $\boldsymbol{A} = \left[a_{ijr}^{lt}\right]$ is the assignment matrix.

Then the DODME problem can be reformulated in the following terms: Given a network with a set of links L, a set $N := I \times J$ of OD pairs, and the set of time periods $\mathcal{T}$. The goal of the dynamic OD-matrix estimation problem is to find a feasible vector (OD-matrix) $\boldsymbol{X}^* \in \mathrm{G} \subseteq \mathbb{R}_+^{N \times \mathcal{T}}$, where $\boldsymbol{X}^* = \left(x_{ijr}^*\right),\ (i,j) \in N, r \in \mathcal{T}$, consists of the demands for all OD pairs. It can be assumed that, when assigning the time-sliced OD matrices onto the links of the network, it should be done according to an assignment proportion matrix $\boldsymbol{A} = \left[a_{ijr}^{lt}\right], \forall l \in L, \forall (i,j) \in N, \forall r,t \in \mathcal{T}$, where each element in the matrix is defined as the proportion of the OD demand x_{ijr} that uses link l at time period t. The notation $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{X})$ is used to indicate that, in general, these proportions depend on the demand. The linear relationship between the flow count on a link and the given OD pair is in matrix form, which thus sets the vector of detected flows as $\boldsymbol{Y} = (\boldsymbol{Y_1}, \dots, \boldsymbol{Y_T}) = (y_{e_1 1}, \dots, y_{e_L 1}, \dots, y_{e_1 T}, \dots y_{e_L T})$ and the vector of OD flows as $\boldsymbol{X} = (\boldsymbol{X_1}, \dots, \boldsymbol{X_T}) = (x_{i_1 j_1 1}, \dots, x_{i_I j_J 1}, \dots, x_{i_1 j_1 T}, \dots, x_{i_I j_J T})$. The relationship can be expressed as a matrix product, that is:

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{X}) \cdot \boldsymbol{X}$$

$$\text{with } \boldsymbol{A}(\boldsymbol{X}) = \begin{pmatrix} \boldsymbol{A^{11}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{A^{12}} & \boldsymbol{A^{22}} & \boldsymbol{0} & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{A^{1T}} & \cdots & \boldsymbol{A^{T-1\,T}} & \boldsymbol{A^{TT}} \end{pmatrix} \text{ where } \boldsymbol{A^{rt}} = \begin{pmatrix} a_{i_1 j_1 r}^{1t} & \cdots & a_{i_I j_J r}^{1t} \\ \vdots & \ddots & \vdots \\ a_{i_1 j_1 r}^{Lt} & \cdots & a_{i_I j_J r}^{Lt} \end{pmatrix} \quad (4)$$

where a_{ijr}^{lt} represents the proportion of OD flow departing at time r, x_{ijr}, passing through link l at time t, y_{lt}. $\boldsymbol{A^{rt}}$ represents the assignment matrix for the departing flows at time window r detected at time window t and, therefore, $\boldsymbol{A}$ is a lower-diagonal matrix, because OD flow departing at time r cannot pass through link l at time $t < r$.

This linear mapping between the link flows and the OD flows is indeed the first term in the Taylor expansion of the relationship between link flows and OD flows, where additional terms capture the assignment matrix's sensitivity to changes in the OD flows, path choice and congestion propagation effects (Frederix et al. 2011, 2013; Toledo and Kolechkina 2013). Let $\widetilde{\boldsymbol{X}}$ be in the neighbourhood of $\boldsymbol{X}$. Then, the Taylor expansion is:

$$\begin{aligned} y_{lt} &= \sum_{(i,j)\in N} \sum_{r=1}^{t} a_{ijr}^{lt}(\widetilde{\boldsymbol{X}}) \tilde{x}_{ijr} + \sum_{(i,j)\in N} \sum_{r=1}^{t} \frac{\partial y_{lt}(\widetilde{\boldsymbol{X}})}{\partial x_{ijr}} \left(x_{ijr} - \tilde{x}_{ijr}\right) = \\ &= \sum_{(i,j)\in N} \sum_{r=1}^{t} a_{ijr}^{lt}(\widetilde{\boldsymbol{X}}) \tilde{x}_{ijr} + \sum_{(i,j)\in N} \sum_{r=1}^{t} \left. \frac{\partial \left[\sum_{(i,j)\in N} \sum_{r=1}^{t} a_{ijr}^{lt}(\widetilde{\boldsymbol{X}}) x_{ijr}\right]}{\partial x_{ijr}} \right|_{\widetilde{\boldsymbol{X}}} \left(x_{ijr} - \tilde{x}_{ijr}\right) = \\ &= \sum_{(i,j)\in N} \sum_{r=1}^{t} a_{ijr}^{lt}(\widetilde{\boldsymbol{X}}) \tilde{x}_{ijr} + \sum_{(i,j)\in N} \sum_{r=1}^{t} \left(x_{ijr} - \tilde{x}_{ijr}\right) \left[\sum_{(\tilde{\imath},\tilde{\jmath})\in N} \sum_{r'=1}^{t} \left. \frac{\partial a_{\tilde{\imath}\tilde{\jmath}r}^{lt}(\widetilde{\boldsymbol{X}})}{\partial x_{\tilde{\imath}\tilde{\jmath}r}} \right|_{\boldsymbol{X'}} \tilde{x}_{\tilde{\imath}\tilde{\jmath}r} \right] \end{aligned} \quad (5)$$

2.1. Dynamic Spiess algorithm

This allows redefining Spiess' approach to the dynamic case, Ros-Roca et al. (2019), by simply using the first term in the above Taylor expansion. It does not account for the propagation effects, but it explicitly considers the time dependencies:

$$\min Z(\boldsymbol{X}) = \frac{1}{2}\sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}}\left(\left(\sum_{(i,j)\in N}\sum_{r=1}^{t} a_{ijr}^{lt} x_{ijr}\right) - \hat{y}_{ijt}\right)^2 \tag{6}$$

$$\boldsymbol{A} = Assignment(\boldsymbol{X})$$
$$x_{ijr} \geq 0$$

where $\boldsymbol{A} = \left[a_{ijr}^{lt}\right]$ is the assignment matrix described before. Therefore, the linear combination inside the brackets is the simulated flow y_{lt}, and

$$\frac{\partial y_{lt}}{\partial x_{ijr}} = a_{ijr}^{lt} \tag{7}$$

As in (Spiess 1990), the chain rule can be used to obtain the gradient of the objective function:

$$\frac{\partial Z}{\partial x_{ijr}} = \sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}}\frac{\partial y_{lt}}{\partial x_{ijr}}(y_{lt} - \hat{y}_{lt}) = \sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}} a_{ijr}^{lt}(y_{lt} - \hat{y}_{lt}) \tag{8}$$

In addition, to find the optimal step size by using the same procedure, we obtain similar equations:

$$y'_{lt} = \frac{d\, y_{lt}}{d\, \lambda} = \sum_{r=1}^{t}\sum_{(i,j)\in N}\frac{d\, x_{ijr}}{d\, \lambda}\frac{\partial y_{lt}}{\partial x_{ijr}} = \sum_{r=1}^{t}\sum_{(i,j)\in N} -x_{ijr}\frac{\partial Z}{\partial x_{ijr}}\frac{\partial y_{lt}}{\partial x_{ijr}} \tag{9}$$

The optimal step length λ can then be calculated solving a 1-dimensional optimization problem, whose solution is given by:

$$\mathrm{Z}'(\lambda) = \sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}} y'_{lt}(\tilde{y}_{lt} - \hat{y}_{lt} + \lambda y'_{lt}) = 0$$
$$\lambda^* = \frac{-\sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}} y'_{lt}(y_{lt} - \hat{y}_{lt})}{\sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}} {y'_{lt}}^2} \tag{10}$$

The Spiess like formulation can be improved adding a second term in the objective function, as in (1), in order to compare the estimated matrix to a historical OD matrix. Assuming a quadratic function to measure the distances between the estimated and the historical or target matrix:

$$\min Z = \frac{1}{2}\sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}}\left(\left(\sum_{(i,j)\in N}\sum_{r=1}^{t} a_{ijr}^{lt} x_{ijr}\right) - \hat{y}_{lt}\right)^2 + \frac{w}{2}\sum_{r\in\mathcal{T}}\sum_{(i,j)\in N}\left(x_{ijr} - x_{ijr}^{H}\right)^2 \tag{11}$$

In this case, Equation (8) is updated as:

$$\frac{\partial Z}{\partial x_{ijr}} = \sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}}\frac{\partial y_{lt}}{\partial x_{ijr}}(y_{lt} - \hat{y}_{lt}) + \frac{w}{2}x_{ijr} = \sum_{t\in\mathcal{T}}\sum_{l\in\hat{L}} a_{ijr}^{lt}(y_{lt} - \hat{y}_{lt}) + \frac{w}{2}x_{ijr} \tag{12}$$

Then the iterative Dynamic Spiess procedure would be as follows:

$$X_i^{(k+1)} = \begin{cases} X_i & for\ k = 0 \\ X_i^{(k)}\left(1 - \lambda^{(k)}\left[\frac{\partial Z(X)}{\partial X_i}\right]_{X_i^{(k)}}\right) & for\ k > 0 \end{cases}$$

3. A new approach

The analytical approaches to DODME problem discussed so far, show that all are based on the availability of the Assignment Matrix ***A*** and, in the case of the dynamic extensions, of its expansion a_{ijr}^{lt} for the various time intervals, and the travel times from the origin of the trip to the corresponding link *l*. And the main role of the Dynamic Traffic Assignment at the lower level of (1) is just to provide this estimate at each time interval.

The availability of the GPS tracking data enables us to assume that after a suitable data processing the empirical paths and the path choices can be interpreted in terms of an empirical dynamic assignment. Then if an empirical assignment matrix can also be estimated then it would play a similar role to that of the analytical assignment matrix estimated from the Dynamic Traffic Assignment.

Therefore, the research question addressed in the following sections is:

- Assuming that suitable GPS tacking data (e.g. waypoints) are available for a given period of time, and
- An *ad hoc* data processing generates an empirical assignment matrix of enough quality, and
- Additional traffic information is also available (e.g. link flow counts at a subset of links in the network)

Then, to investigate whether it is possible to use such information to find a new formulation of the DODME problem, in terms of an optimization model, not requiring the execution of any Traffic Assignment procedure.

3.1 From GPS Data to Link travel Times

GPS Data: Collected from GPS devices equipping fleets of vehicles when tracking their trajectories across the network. They are usually available in the format of datasets, as the one in Table 1, of trips detailed by an ordered sequence of waypoints, $(ID_k, ts_{k\,l}, lat_{kl}, long_{kl})$, for each trip k with a Trip Identity ID_k, the recording date, the recording time tag $ts_{k,l}$ when the *l-th* observation of trip *k* was recorded, and the latitude and longitude of the current position of the vehicle when the data were recorded.

Table 1. GPS WAYPOINTS DATA SAMPLE

ID	DATE	TIME TAG	LATITUDE	LONGITUDE
4261353	2015-11-30	22:43:58	45.445988	9.124048
4261353	2015-11-30	22:44:57	45.445496	9.121952
………..	……………	…………	…………..	…………….
4261353	2015-11-30	22:50:57	45.444767	9.119217
4261355	2015-11-30	22:43:58	45.44598	9.124048
4261355	2015-11-30	22:44:57	45.445496	9.121952
…………	…………….	…………….	…………..	……………..

4262355	2015-11-30	22:50:57	45.444767	9.119217
…………	……………	…………….	……………	…………….

These data should be map matched onto the map of the scenario being analyzed, the conceptual approach to the procedure developed in this work is depicted in Figure 2.

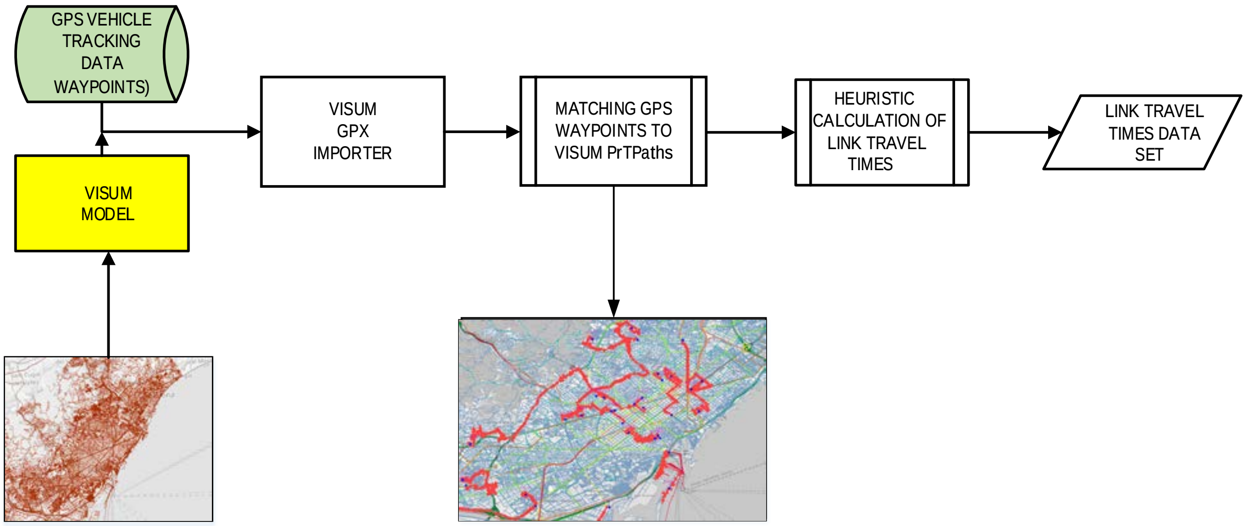


Figure 2. Conceptual methodological approach to the process of importing waypoints into a Visum model and their use to calculate link travel times.

The map matching process Waypojnts →GPX Importer →PrTPaths is illustrated in Figure 3.

Trip p starting at point $e_{p_1} = i$, and ending at point $e_{p_n} = j$, corresponding to the trajectory of a car from the "Selected Sample Database". The latitude and longitude of the starting point $e_{p_1} = i$, is the first waypoint $WP_{p0} = e_{p_1}$, and the latitude and longitude of the end point $e_{p_n} = j$, is the final waypoint $WP_{pn_k} = e_{p_n}$.

The tracking of the routes route between i and j, is defined as a sequence of waypoints $\{WP_{p1}, \dots, WP_{pn_k}\}$ defined by a triple, longitude, latitude and time tag: $WP_{ps} = \{X_{ps}, Y_{ps}, t_{ps}\}$, determined by the sampling rule.

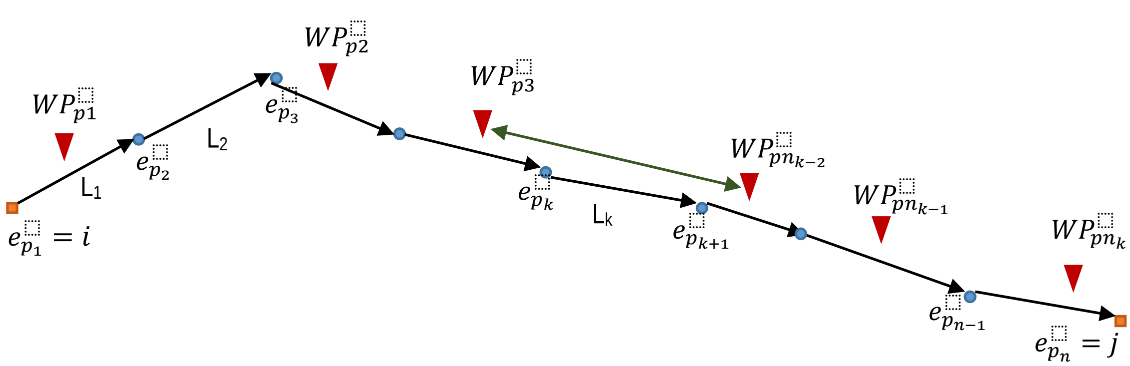
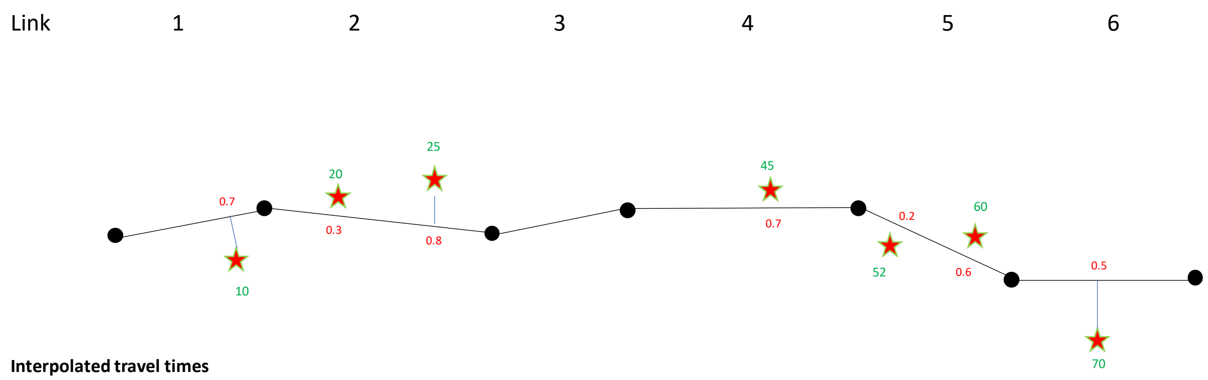


Figure 3. Hypothetical map matching of trajectory of trip p Waypoints of trip p from $e_{p_1} = i$ to $e_{p_n} = j$

The trajectory is identified by the route defined by the sequence of links {L_1, L_2,…..} and each link is defined by a pair of nodes, e.g. $L_k = (e_{p_k}, e_{p_{k+1}})$. Some of the links have associated a waypoint, the s-th waypoint of path p, WP_{ps}, is defined by a triple, longitude, latitude and time tag: WP_{ps}={X_{ps}, Y_{ps}, t_{ps}}, from which travel times between successive waypoints can be calculated: $tt_{ps,ps+1} = t_{ps+1} - t_{ps}$. This information can be used to heuristically estimate link travel times as follows:



Interpolated travel times

$tt_1 = tt_6$ = not covered completely > Undefined > not part of PrT Path

$$tt_2 = \frac{0.3l_2}{0.3l_1+0.3l_2}10s + 5s + \frac{0.2l_2}{0.2l_2+l_3+0.7l_4}20s$$

$$tt_3 = \frac{l_3}{0.2l_2+l_3+0.7l_4}20s$$

$$tt_4 = \frac{0.7l_4}{0.2l_2+l_3+0.7l_4}20s + \frac{0.3l_4}{0.3l_4+0.2l_5}7s \quad \text{etc.}$$

The interpolated travel time for Link i is stored in the PrT Path Element for the FromNode of Link i.

The first waypoint information and the interpolated travel time by link is used to propagate the time and the information is averaged by link and arrival time interval. The outcome of this process is the set of link travel times at time *t*: $\hat{tt}_{lt}, \forall l \in L, \forall t \in T$ for all links in the network used by the GPS traces. This is the Data Set of Observed Link Travel Times.

3.2 Route Choice Set Generation

Using the Observed Link Travel Times as input, k-shortest paths between each OD pair for each Departing Time interval must be calculated somehow. Many alternatives are considered in this step:

- *Dijsktra and CF_k penalization*: For each OD pair (i, j) and departing time interval r, a shortest path is calculated. In order to have many alternative routes, the first route is penalized by the Commonality Factor and a new shortest path is calculated again. This is done iteratively to determine different paths. The Commonality Factor for a given path k, it is calculated by, (Bovy et al., 2008) as modified by Janmyr and Wadell, 2018):

$$CF_{k(i,j,r)} = \frac{1}{\mu_{CF}} \sum_{a \in \Gamma_{k(i,j,r)}} \left(\frac{l_a}{L_{k(i,j,r)}} \log \left(\sum_{h \in K_{ijr}} (\delta_{ahr} + 1) \right) \right) \quad (13)$$

The Commonality Factor penalization to be applied to each link *a* belonging to path *k* is:

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

where K_{ijr} is the calculated set of paths for (i,j)-OD pair at r departing time interval, l_a is the length of the link and $L_{k(i,j,r)}$ is the total length of the shortest path for set K_{ijr}. $\Gamma_{k(i,j,r)}$ is the sequence of links for path k and $\delta_{ahr} \in \{0,1\}$ indicates whether link a does belong to path h included in the calculated path set K_{ijr} or not.

This alternative is computationally expensive because it requires $|\boldsymbol{I}| \times |\boldsymbol{J}| \times |\boldsymbol{T}| \times \boldsymbol{k}$ shortest path algorithm evaluations.

- *S-HEAP algorithm:* In order to reduce the computational effort. The algorithm can be changed to a *Shortest Path Tree* algorithm for each origin, using (Gallo and Pallotino, 1988) algorithm.
  - Similarly, to the first option, an iterative procedure must be used to compute k-shortest paths, using the same penalization.
  - **In this case, the computational effort is reduced to $|\boldsymbol{I}| \times |\boldsymbol{T}| \times \boldsymbol{k}$ evaluations of the algorithm. However, each individual step is a bit more expensive in time.**
- *Assignment:* This option is to use an auxiliary Dynamic Assignment in order to obtain the different paths for each OD pair and each time interval. The role of this auxiliary Dynamic Assignment in this case is not that of providing an t assignment matrix, as in the analytical formulation, it is only an auxiliary tool to generate the paths that are used at each time interval. This option is weak in the sense that it will be contradictory with the name "Assignment Free DODE", but is an acceptable option to produce a first "Proof of Concept" of the algorithm, due to its easy implementation.
- *k-Time Dependent Shortest Paths:* This was the main option because it is the most accurate since it considers the proper calculation of the k time dependent shortest paths (TDSP). The proposed algorithm to find the k-TDSP is (Chabini, 1997). **This option is the most complex, so it requires more implementation effort and also computational time.** However, from our previous experience, at the moment the increase in quality does not justify the implementation effort for the Proof of Concept purpose.

3.3. Calculation of $\boldsymbol{CF_k}$ and $\boldsymbol{P_k}$

At this step, the input is a set of calculated paths, K_{ijr}, for each origin i, destination j and departing time r. These paths are noted by $k(i,j,r) \in K_{ijr}$, to show explicitely the dependence to (i,j,r). For a certain path $k(i,j,r)$, the sequence of links is the set $\Gamma_{k(i,j,r)} = \{e_1, \dots, e_{m_k}\}$.

The path choice for each path on the set K_{ijr} is calculated as a discrete choice model that uses the commonality factor, CF_k, as a penalization factor on travel times, (Nassir et al., 2014). That is:

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

$$P_{k(i,j,r)} = \frac{\exp\left[\mu_p\left(-\widehat{tt}_{k(i,j,r)} - CF_{k(i,j,r)}\right)\right]}{\sum_{h\in K_{ijr}} \exp\left[\mu_p\left(-\widehat{tt}_{h(i,j,r)} - CF_{h(i,j,r)}\right)\right]} \quad (14)$$

These calculations permit to obtain the flow distribution for each path, based on observed path travel times. These observed path travel times are the summation of the observed link travel times, considering the arrival time, t(k), at each link a, included in the path k:

$$\widehat{tt}_{k(i,j,r)} = \sum_{a\in\Gamma_{k(i,j,r)}} \widehat{tt}_{at(k)} \quad (15)$$

3.4. Calculation of the Time Dependent Assignment Matrix

Once $\boldsymbol{P_k} = [P_{k(i,j,r)}]$ is calculated from the k-shortest paths calculated after the travel times estimated from the GPS data for all OD pairs, the empirical assignment matrix $\left[\bar{a}_{ijr}^{lt}\right]$ can be calculated:

$$\bar{a}_{ijr}^{lt} = \sum_{k\in K_{ijr}} \delta_{k(i,j,r)}^{lt} P_{k(i,j,r)} \qquad \forall i,j,r,l,t \quad (16)$$

where $\delta_{k(i,j,r)}^{lt}$ is the empirical incidence indicator:

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

3.5 An Assignment Free DODME Approach

The introduction in Section 1 to the analytical approaches to DODME problem, and the discussion of their extensions in Section 2, shows that all are based on the availability of the Assignment Matrix $\boldsymbol{A}$. And the main role of the Dynamic Traffic Assignment at the lower level of (1) is just to provide this estimate at each time interval.

However, the availability of the GPS tracking data, and the proposed methodological analysis, enables us to assume that the paths $\bar{K}_{ijr}$ calculated in 3.2, and the path choices (14) in 3.3 can be interpreted in terms of an empirical dynamic assignment. Then the empirical assignment matrix (16) plays the same role as the analytical assignment matrix estimated in Section 2 from the Dynamic Traffic Assignment.

Considering that the dynamic formulation of (1) model decomposes it into time intervals to emulate the flow propagation across the network, assuming that at each iteration of the optimization procedure the relationships between the estimated flows and the adjusted matrix can be set up in terms of a time dependent assignment matrix (2). Then stopping the Taylor expansion in the first order terms, as in (5), Toledo and Koleckhina (2013), Frederix et al (2013), Ros-Roca et al. (2019), and replacing the analytical by the empirical assignment matrix (16) the relationship between the link flows and the OD matrix can be restated as:

$$\bar{y}_{lt} = \sum_{i\in I}\sum_{j\in J}\sum_{r=1}^{t} \bar{a}_{ijr}^{lt} x_{ijr} \quad (17)$$

where $\overline{y}_{lt}$ is the estimated flow in link l at time period t, x_{ijr} is the flow departing origin $i \in I = \{set\ of\ origins\}$, with destination $j \in J = \{set\ of\ destinations\}$, at time interval r, and $\overline{a}_{ijr}^{lt}$, the empirical assignment matrix, is the fraction of trips from origin $i \in I$ with destination $j \in J$, departing from *i* at time *r*, that reach link *l* at time *t*.

Then, assuming that traffic counts measured by real detectors placed onto the network are available. They are measured for each time interval and denoted by: $\hat{y}_{lt}$, where $l \subseteq \hat{L}$ is the link with the detector of the network and t the time interval when it is measured.

The research question addressed in this research report can be formulated in the following terms: if the data collected from a sample of GPS tracked vehicles provide us with a discretized time estimate of the target OD matrix $\widehat{\boldsymbol{X}} = \left[\hat{x}_{ijr}\right]$, and a suitable processing, Janmyr and Wadell (2018), Krishnakumari et al. (2019), Nassir (2014), provides a sound empirical estimate of $\overline{a}_{ijr}^{lt}$, then the expansion of the sampled target matrix to estimate the OD matrix can be done in terms of the scaling factors per origins, $\alpha_i, i \in I$, and per destinations $\beta_j, j \in J$, such that:

$$x_{ijr} = \alpha_i \beta_j \hat{x}_{ijr}, \forall i \in I, \forall j \in J, \forall r \in T \qquad (18)$$

If $\hat{y}_{lt}, l \in \hat{L} \subset L, t \in T$ are the link flows measured at the counting stations, in a subset $\hat{L} \subset L$ of the network links, the Dynamic Data-Driven OD Matrix Estimation problem can be formulated as the following optimization problem of finding the values of the scaling factors $\alpha_i, i \in I$ and $\beta_j, j \in J$, without the need of conducting the traffic assignment at the lower level of (1), exploiting the empirical assignment matrix $\overline{a}_{ijr}^{lt}$. However, if an historical OD is available from other sources, then a seed matrix x_{ijr}^{0} can be generated combining that historical OD matrix x_{ijr}^{H} and the observed OD matrix $\hat{x}_{ijr}$ obtained from GPS tracked trips, and it is denoted as $\boldsymbol{X^0} = \left[x_{ijr}^{0}\right]$. From (17) and (18) the proposed new formulation of the DODME problem is:

$$\min_{\alpha_i, \beta_j} \left[\sum_{l,t} \left(\hat{y}_{lt} - \sum_{i,j,r} \alpha_i \beta_j \overline{a}_{ijr}^{lt} x_{ijr}^{0} \right)^2 \right] \qquad (19)$$

$$s.t. \quad \alpha_i, \beta_j \geq 0$$

The variables of the problem are multiplicative scaling factors for each origin, α_i, and destination, β_j. They are inspired on gravity models where bi-dimensional constraints for rows and columns are set. The minimization problem is solved iteratively with the L-BFGS-B method appropriated for constrained non-linear problems using the version available in python package *scipy.optimize*.

The conceptual computational scheme of the proposed Assignment Free DODME approach, powered by the ICT applications capturing GPS data trajectories is depicted in Figure 4.

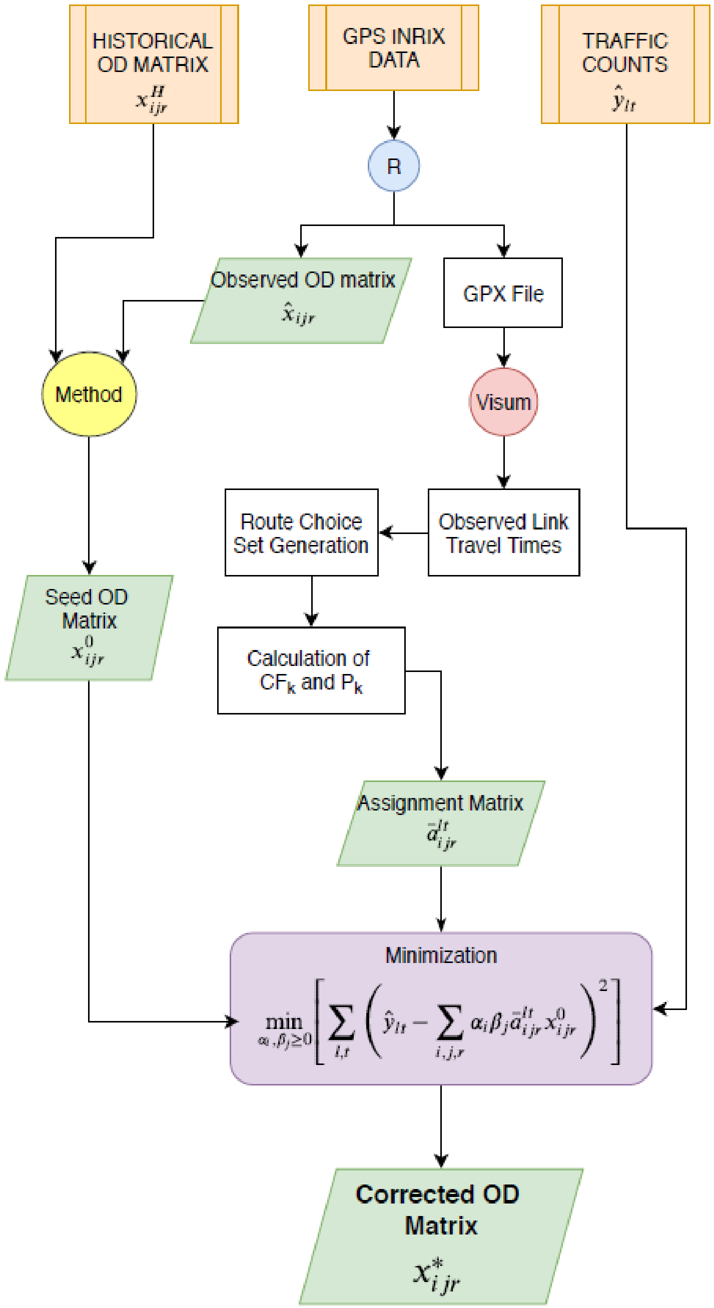


Figure 4. A Data Driven Assignment Free DODME

4. A computational proof of concept of the Assignment Free data Driven DODME

A first test with the Torino CDB subnetwork, Figure 5, has been conducted in order to check the functional feasibility of the algorithmic chain, Figure 4, of the new approach.

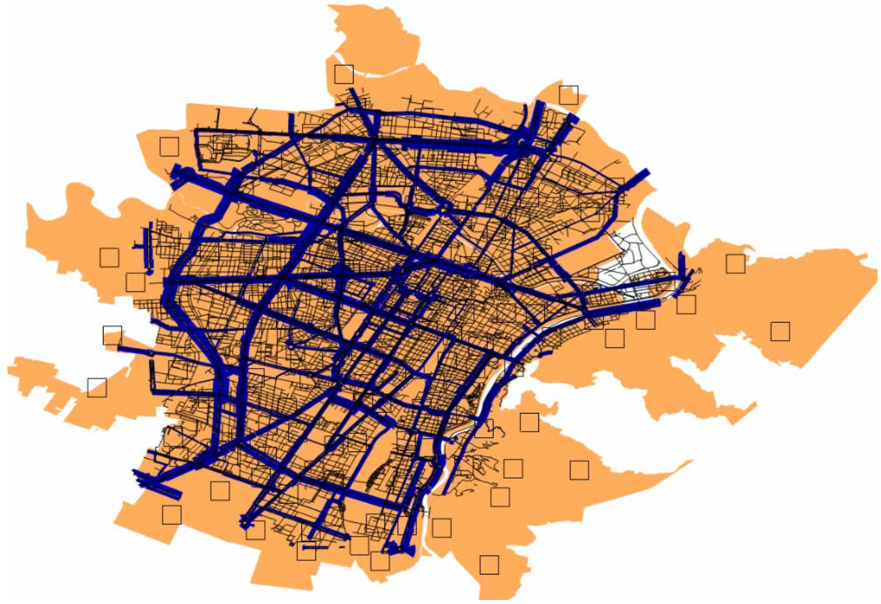


Figure 5: Torino CBD Subnetwork

A summary of the model and the real inputs is shown in the Table 2:

Table 2: Torino's CBD Experiment characteristics

TIME PERIODS	4
ZONES	268
DETECTORS	582
OD PAIRS X TIME	287k
INRIX WAYPOINTS	2.9M
INRIX TRIPS	166k
OBS OD > 0	32k (12%)
HIST OD > 0	120k (42%)

4.1. GPS Data: Building the Seed OD matrix [yellow circle Method in Figure 4]

6 months of GPS data coming from INRIX for all Piemonte region are available. These are 220M waypoints for 5.9M trips. There is no further information regarding the type of vehicle (fleet/private) and it points that there was no distinction between them. Consequently, the only filtering that can be done relies on zoning and weekday selection.

After the pre-processing and filtering steps, the final dataset of waypoints (GPX file) contains 2.9M of waypoints for 166k trips (reduction to the 1.31% of the raw data). The observed OD matrix, $\hat{x}_{ijr}$, has 32k positive values, which is the 12% of the matrix.

The GPS data have to be preprocessed in order to filter the useful data for the method proposed in this research paper. In a first step, only weekdays trips (Tuesday, Wednesday and Thursday) in the considered time interval have been selected.

After that, using a SHP file with the information of the Transportation Zoning System (TAZ) of the model, an Origin and a Destination Zone has been assigned to each trip. Those trips that begin and/or end out of the study area, are cut and considered to be started/ended on the first/last zone of the network where they are observed.

There are two outputs of this procedure:

- *GPX file*: This GPX file is generated with the filtered and processed trips. They are coded in a GPX file to be inserted in VISUM GPX import tool.
- *Observed OD matrix*: The filtered and processed trips are counted by Origin, Destination and Departing Time to be translated to an observed OD matrix, $\hat{x}_{ijr}$.

  Remark: The experience gained preprocessing the INRIX data shows they could be two questionable issues:

  a) The mix of data from commercial fleets and private vehicles that must be properly filtered.

  b) The bias that can be induced by the data collection process splitting long trips in shorter trips due to privacy issues, in which case a significant fraction of origins and destinations do not correspond to the underlying mobility pattern.

  These drawbacks are independent on the size of the sample, since they are inherent to the data collection procedure, designed for other purposes than the OD matrix estimation.

  <u>This is one of the reasons to consider this computational testing only as a proof of concept since we cannot ensure the quality of the inputs, and therefore assess the quality of the outputs.</u>

After the Map Matching into VISUM, the total number of PrTPaths is 130k, which means the 78% of the sample. This loss is acceptable due to the short trips.

The observed link travel times per time interval, $\widehat{tt}_{lt}$, computed by averaging the PrTPaths' Interpolated travel times at link level after propagating the time, covers the major part of network, resulting a connected network, Figure 6. Furthermore, the number of observations per link and time interval is right-skewed distributed as shown:

min	Q1	Median	Mean	Q3	max
1	7	24	50	69	658

We observed that the more observations a link has, the lower the standard deviation of its travel times is, which is consistent.

Figure 6: Links with observed link travel time.

Regarding the possibility of not finding shortest paths because there could be not connected zones, the preliminary test with a set of OD pairs showed that the connectivity is not a problem, since up to the 93% of the times, a shortest path is found.

Some tests using Shortest Path Visum show the goodness of these observed travel times. Figure 7 shows an example of the shortest path calculation between two nodes calculated using the free-flow travel time (left) and the observed travel time (right).

Figure 7: Comparison of shortest paths between t0 and observed travel time.

However, the use of the shortest path tool of Visum is not feasible because of its computational time. It takes 0.6 seconds for each path, which for each OD pair and time interval is approximately: 268·268·4·0.6sec = 143,468sec = 40h = 1day16h. Other alternatives as those listed before must be considered.

4.2. Dynamic Stochastic Assignment as Route Choice Set

For the proof of concept, the Route Choice Set Generation has been substituted by the Dynamic Stochastic Assignment (10 iterations) available in Visum as a shortcut to define the calculated paths for each OD pair at each time interval. For these paths, commonality factors and path proportions, CF_k and P_k, are calculated using the observed link travel times coming from the processing of the GPS data.

4.3. Calculation of CF_k , P_k and $\overline{a}_{nr}^{lt}$ and Optimization

The calculation of CF_k, P_k and $\overline{a}_{ijr}^{lt}$ has been conducted applying (14), (15) and (16). The objective function (19) shows a nice and fast decrease and convergence, see Figure 8.

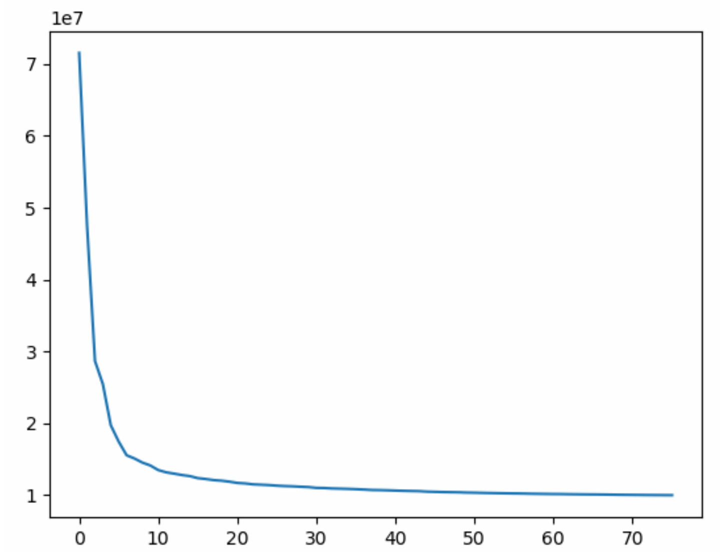


Figure 8: Descent of the Objective Function

The seed OD matrix, x_{ijr}^{0}, is in this case the Historical OD matrix for Torino, i.e., $x_{ijr}^{0} = x_{ijr}^{H}$. An analysis of the evolution of traffic counts is shown in Figure 9:

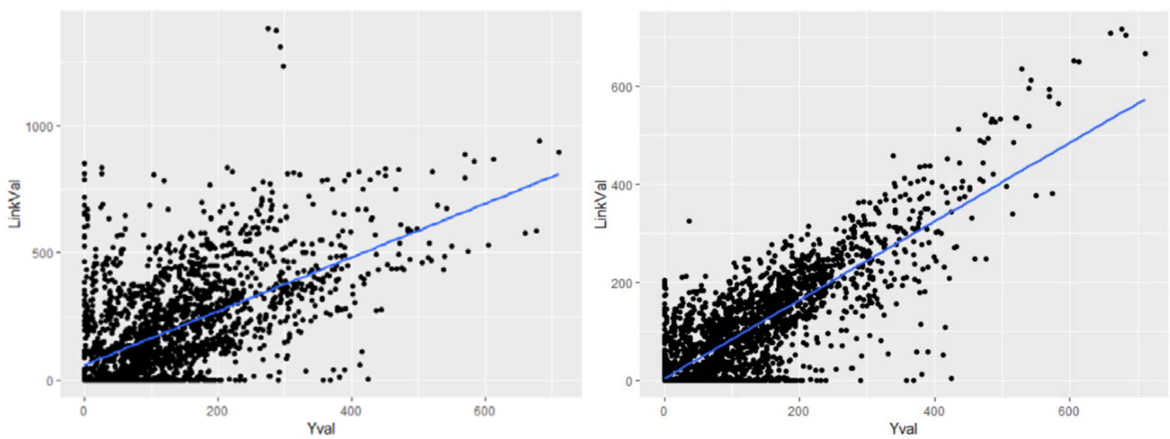


Figure 9: Traffic Flow counts before and after the optimization procedure.

The initial R^2 is 0.33 and the final R^2 is 0.66.

4.4. Summary of Computational Times

- Preprocess and filtering of GPS Data -> 50 minutes
- Import GPX tool -> 4h 21minutes
- Link Travel Times Computation -> 6minutes
- Dynamic Stochastic Assignment -> 13minutes
- Calculation of CF_k, P_k and $\bar{a}_{ijr}^{lt}$ -> 1minute
- Optimization -> 10minutes/iteration

5. Alternative Assigment Free DODME proposal

- Assuming that a time sliced historical OD matrix $\boldsymbol{X^H}$ is available
- The expansion of the sample x_{ijr}^0 is bounded from below by lb_{ijr} and from above by ub_{ijr}, accounting explicitly for the socioeconomic characteristics of the TAZ
- Weights w_1 and w_2 on the distance functions

$$\min_{\alpha_i,\beta_j}\left\{w_1\left[\sum_{i\in I}\sum_{j\in J}\sum_{r\in T}\left(x_{ijr}^H-\alpha_i\beta_j x_{ijr}^0\right)^2\right]+w_2\sum_{t\in T}\sum_{l\in\hat{L}}\left[\hat{y}_{lt}-\sum_{i\in I}\sum_{j\in J}\sum_{r=1}^{t}\bar{a}_{ijr}^{lt}\alpha_i\beta_j x_{ijr}^0\right]^2\right\} \quad (20)$$

$$lb_{ijr}\le\alpha_i\beta_j x_{ijr}^0\le ub_{ijr},\forall i\in I,\forall j\in J,\forall r\in T$$

$$\alpha_i\ge 0,\forall i\in I$$

$$\beta_j\ge 0,\forall j\in J$$

5.1 Variant 1

However, while enabling to capture the underlying physical characteristics of the underlying transport system, this formulation has the drawback of the high number of non-linear constraints:

$$lb_{ijr}\le\alpha_i\beta_j x_{ijr}^0\le ub_{ijr},\forall i\in I,\forall j\in J,\forall r\in T$$

$|I|\times|J|\times|T|$. Summing for *i* and *j* the lower and upper bounds respectively, one can obtain aggregated bounds that instead bounding each cell (i,j) bound, respectively, the generation and attraction characteristics of each TAZ:

$$LB_{ir}=\sum_{j\in J}lb_{ijr}\quad and\ \ UB_{ir}=\sum_{j\in J}ub_{ijr}\quad\forall i\in I,\forall r\in T$$

$$LB_{jr}=\sum_{i\in I}lb_{ijr}\quad and\ \ UP_{jr}=\sum_{i\in I}ub_{ijr}\quad\forall j\in J,\forall r\in T$$

Then the set of bounding constraints can be reformulated as:

$$LB_{ir} \leq \alpha_i \sum_{j \in J} \beta_j x_{ijr}^0 \leq UB_{ir} \ \ \forall i \in I, \forall r \in T \tag{21}$$

$$LB_{jr} \leq \beta_j \sum_{i \epsilon I} \alpha_i \, x_{ijr}^0 \leq UB_{jr} \ \forall j \in J, \forall r \in T \tag{22}$$

That is $(|I| + |J|) \times |T|$ non-linear constraints.

5.2 Variant 2

However, from a computational perspective while reducing significantly the size of the constraint set, the sets of constraints (21) and (22) still have a major computational disadvantage since they are nonlinear constraints. A suitable approximation, replacing them by linear ones, could improve the computational performance while holding the main physical characteristics of the underlying transport problem.

The formulation (20) assumes that the entry cell (*i*,*j*) of the estimated OD matrix for time period *r*, x_{ijr} is given by:

$$x_{ijr} = \alpha_i \beta_j x_{ijr}^0$$

Then the trip generation of origin *i* for time period *r* is given by:

$$O_{ir} = \sum_{j \in J} x_{ijr} = \sum_{j \in J} \alpha_i \beta_j x_{ijr}^0 = \alpha_i \sum_{j \in J} \beta_j \, x_{ijr}^0 \rightarrow \frac{O_{ir}}{\alpha_i} = \sum_{j \in J} \beta_j \, x_{ijr}^0 \tag{23}$$

Let's assume that:

- $\hat{O}_{ir}$ is an estimate of O_{ir}, for instance from a suitable historical OD matrix $\boldsymbol{X^H}$. Remark, this could be a good chance to include alternative OD sources, as for example those from KINEO's (or similar matrices from Smartphones) in this "fusion" procedure.
- If upper and lower bounds for α_i, α_i^l, and α_i^u, respectively, can be estimated such that $\alpha_i \in \left[\alpha_i^l, \alpha_i^u\right]$

Then:

$$\widehat{UB}_{ir} = \frac{\hat{O}_{ir}}{\alpha_i^l} \quad and \ \widehat{LB}_{ir} = \frac{\hat{O}_{ir}}{\alpha_i^u}$$

Would respectively be upper and lower bound estimates of (23), and then, constraints (21) could be approximated by:

$$\widehat{LB}_{ir} \leq \sum_{j \in J} \beta_j \, x_{ijr}^0 \leq \widehat{UB}_{ir}, \ \ \forall i \in I, \forall r \in T \tag{24}$$

Similarly:

$$D_{jr} = \sum_{i \in I} x_{ijr} = \sum_{i \in I} \alpha_i \beta_j x_{ijr}^0 = \beta_j \sum_{i \in I} \alpha_i \, x_{ijr}^0 \rightarrow \frac{D_{jr}}{\beta_j} = \sum_{i \in I} \alpha_i \, x_{ijr}^0 \tag{25}$$

Then, if:

- $\hat{D}_{jr}$ is an estimate of D_{jr}, for instance from a suitable historical OD matrix $\boldsymbol{X^H}$.

- And upper and lower bounds for β_j, β_j^l, and β_j^u, respectively, can be estimated such that $\beta_j \in [\beta_j^l, \beta_j^u]$

$$\widehat{UB}_{jr} = \frac{\widehat{D}_{jr}}{\beta_j^l} \quad and\ \widehat{LB}_{jr} = \frac{\widehat{D}_{jr}}{\beta_j^u}$$

Would, respectively, be upper and lower bound estimates of (6), and then, constraints (3) could be approximated by:

$$\widehat{LB}_{jr} \leq \sum_{i \in I} \alpha_i\, x_{ijr}^0 \leq \widehat{UB}_{jr}, \quad \forall j \in J, \forall r \in T \tag{26}$$

Resulting in the alternative formulation of the Assigment Free DODME:

$$\min_{\alpha_i, \beta_j} \left\{ w_1 \left[\sum_{i \in I} \sum_{j \in J} \sum_{r \in T} \left(x_{ijr}^H - \alpha_i \beta_j x_{ijr}^0 \right)^2 \right] + w_2 \sum_{t \in T} \sum_{l \in \hat{L}} \left[\hat{y}_{lt} - \sum_{i \in I} \sum_{j \in J} \sum_{r=1}^{t} \bar{a}_{ijr}^{lt} \alpha_i \beta_j x_{ijr}^0 \right]^2 \right\} \tag{27}$$

$$\widehat{LB}_{ir} \leq \sum_{j \in J} \beta_j\, x_{ijr}^0 \leq \widehat{UB}_{ir}, \quad \forall i \in I, \forall r \in T \tag{28}$$

$$\widehat{LB}_{jr} \leq \sum_{i \in I} \alpha_i\, x_{ijr}^0 \leq \widehat{UB}_{jr}, \quad \forall j \in J, \forall r \in T \tag{29}$$

$$\alpha_i \geq 0, \forall i \in I$$

$$\beta_j \geq 0, \forall j \in J$$

In which the $(|I| + |J|) \times |T|$ constraints (28) and (29) are linear.

6. Preliminary conclusions and future work

The detailed analysis of the analytical approaches to DODME has made evident the role of the assignment matrix in these formulations, what has been computationally confirmed by the numerical experiments with the Dynamic Spiess procedure. The identification of the role of the assignment matrix raised the main research question addressed in this research report of whether the empirical traffic measurements enabled by the ICT applications, could be interpreted in terms of an empirical assignment and then, from them an empirical assignment matrix could be generated. This research question has found so far a positive answer as we report, enabling a new formulation of the DODME problem as a nonlinear optimization problem. This research report provides a proof of concept of the novel approach, that is, it checks the feasibility of the algorithmic chain. The way in which it has been implemented does not allow to guarantee the quality of the results. The proof of concept looks very promising from the computational stand point but the quality of the results is strongly dependent on the quality of the available data, since it is a data driven process, and the ways in which some of the computational steps have been implemented (i.e. the computation of the route choice sets). Further computational experiments with better and more reliable data, improved route procedures, and seed matrices fusing information from other sources, are expected before drawing reliable conclusions. Improvements can also be expected from the proposed alternative approaches, as well as, from other formulation of the objective function in the optimization problem.

Acknowledgements

This work has benefited of the technical discussions with Dr. Klaus Nökel, Head of technological Innovation at PTV Group, addressing the data processing aspects as well as the route choice selection; the technical support of Dr. Arne Schnek, Chief Software Engineer at PTV Group, with respect to the Visum issues. The heuristic for the calculation of the link travel times from the waypoints is based on a private information from Frauke Petersen from PTV Group and Alessandro Attanassi from PTV Sistema and, last but not least with Professor Guido Gentile of the Dipartimento di Ingegneria Civile Edile e Ambientale, Sapienza Università di Roma on the seminal ideas about the new optimization approach.

This research was funded by Industrial-PhD-Program 2017-DI-041, TRA2016-76914-C3-1-P Spanish R+D Programs and Secretaria d'Universitats-i-Recerca-Generalitat de Catalunya-2017-SGR-1749.

References

Codina, E. and L. Montero (2006). "Approximation of the Steepest Descent Direction for the O-D Matrix Adjustment Problem." Pp.329–62 in *Annals of Operations Research*. Vol.144. Springer US.

Frederix, R., Viti F. and Tampère C. (2013). "Dynamic Origin-Destination Estimation in Congested Networks: Theoretical Findings and Implications in Practice." *Transportmetrica A: Transport Science* 9(6):494–513.

Frederix, R., F. Viti, R. Corthout, and C. Tampère (2011). "New Gradient Approximation Method for Dynamic Origin-Destination Matrix Estimation on Congested Networks." *Transportation Research Record: Journal of the Transportation Research Board* 2263(1):19–25.

Janmyr, J. and Wadell D. (2018), Analysis of vehicle route choice during incidents, Ms Thesis, University of Linköping, Department of Science and Technology, LiU-ITN-TEK-A--18/020—SE

Krishnakumari P., van Lint H., Djukic T. and Cats, O. (2019), A data driven method for OD matrix estimation, Transportation Research C, https://doi.org/10.1016/j.trc.2019.05.014

Lundgren, Jan T. and A. Peterson (2008). "A Heuristic for the Bilevel Origin-Destination-Matrix Estimation Problem." *Transportation Research Part B: Methodological* 42(4):339–54.

Nassir N., J. Ziebarth, E. Sall, and L. Zorn,(2014), "Choice Set Generation Algorithm Suitable for Measuring Route Choice Accessibility," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2430, no. 1, pp. 170–181.

Ros-Roca X., Montero L. and Barceló J. (2019), Investigating the Quality of Spiess-Like and SPSA approaches for Dynamic OD Matrix Estimation, accepted for publication in Transportmetrica.

X. Ros-Roca, L. Montero, A. Schneck, J. Barceló (2018), Investigating the performance of SPSA in Simulation-Optimization approaches to transportation problems, International Symposium of Transport Simulation (ISTS'18), Transportation Research Procedia, Volume 34, pp. 83-90. www.sciencedirect.com

Spiess, Heinz. 1990. "A Gradient Approach for the OD Matrix Adjustment Problem." *Centre for Research on Transportation, University of Montreal, Canada* 693(Publication No. 693, CRT):1–11.

Toledo, Tomer and Tanya Kolechkina. 2013. "Estimation of Dynamic Origin-Destination Matrices Using

Linear Assignment Matrix Approximations.” *IEEE Transactions on Intelligent Transportation Systems* 14(2):618–26.

Wang Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.