

A DUE based bilevel optimization approach for the estimation of time sliced OD matrices

M.Bullejos^{a 1}, J.Barceló^{a 1}, L.Montero^{a 1}

^aDepartment of Statistic and Operations Research, Technical University of Catalonia (UPC) – Barcelona TECH, Jordi Girona 1-3, 08034 Barcelona, Spain

Abstract

Traffic management applications are supported by dynamic models fed with realistic real-time demand matrices in order to find efficient solutions in the short-term horizon. OD matrices have to be estimated from historic or real-time data collection and prior matrices. In this work, we present a simulation-optimization Bilevel-DUE approach for Dynamic OD matrix estimation suitable for off-line application, in which the lower level consists on a dynamic user equilibrium problem solved by a mesoscopic simulator. Improvements as the consideration of travel times and the use of a conjugate gradient combined with trust region scheme are introduced and analyzed. Time-sliced OD matrices generated off-line with the former procedure from historic data are suitable initializations for on-line and real-time dynamic OD estimation methods based on Kalman filtering. Some of those methods were developed by the authors in previous works and they provide realistic matrices to cope with day-to-day demand variability and real-time traffic conditions.

Keywords: OD matrices; dynamic user equilibrium, conjugate gradient, trust region, traffic modelling

1. Introduction

Origin-Destination trip matrices are the main input to dynamic traffic simulation models, and among them those based on either microscopic, or mesoscopic approaches are particularly sensitive to the quality of the input matrix. The quality of the results provided by those simulation models is determined by the accuracy of the model's calibration and validation, consequently a substantial research effort has been spent in developing methodological frameworks to conduct the necessary model calibration. MULTITUDE COST Action TU903 on Calibration and Validation of Traffic Simulation Models is a good example, [Antoniou et al., 2014]. However, a quite common drawback on practical applications focuses the calibration efforts on the supply and behavioral parameters of the models, assuming implicitly that the demand inputs are the right ones and have been calibrated elsewhere. The most correct way of addressing this problem is through the joint calibration of Supply and Demand, [Balakrishna et al., 2007]. However, the complexity and computational burden stemming from these procedures has led researchers to look for other approaches. One of them, successfully explored as part of WP3 of the MULTITUDE project, [Antoniou et al., 2014b], has consisted of using an off-line estimation and adjustment of the OD matrices. These matrices will be used as input to the microsimulation model. An innovative approach proposed in [Cipriani et al., 2011], formulates the problem in terms of a bilevel nonlinear optimization model combining a Simultaneous Perturbation Stochastic Approximation (SPSA) and a Dynamic User Equilibrium (DUE) model with the ability to capture the main time dependencies of traffic dynamics. The outer level is a merit function, that updates the current time-sliced OD matrix and the inner level at outer iteration k consists on the DUE whose input is the updated matrix.

This paper aims to enhance the computational procedures for the offline dynamic OD demand estimation problem using a trust-region algorithm to generate perturbed matrices at the outer level improving the computation of the descent direction.

As detailed in [Spall, 1998] the choice of the gain sequences (a_k and c_k) and the simultaneous perturbation vector (SPV) is critical to the performance of SPSA. We use an Asymmetrical Design (AD) to halve the number of objective function evaluations required in any simultaneous perturbation. At iteration k , “ m ” perturbations of

¹ Manuel Bullejos (Tel.: +34 93 401 1019), Lidia Montero (Tel.: +34 93 401 1038), Jaume Barceló (Tel.: +34 93 401 7033).
mail address: (manuel.bullejos, lidia.montero, jaume.barcelo)@upc.edu

the initial matrix are created to compute “m” gradient approximations. The dimension of the gradient is related to the statistical goodness of fit measure used in the outer merit function. Our computational experiment with the family of merit functions proposed in MULTITUDE, [Antoniou et al., 2014b], led us to conclude that the Normalized Mean Error, NME, is the one providing the more consistent results. The average gradient is used for updating the current time-sliced OD at outer iteration k , with which the next iteration begins. Reducing the number of replications is a key factor since one replication is a very time-consuming job and the larger the network size, the greater the computational time.

In a previous research, [Barceló et al., 2013], we have found that the quality of the detection layout is a key design factor. That paper provides a methodology to define a suitable Bluetooth sensor layout allowing the identification of the most likely used paths between sensors and therefore the measurements of the associated travel times. Consequently, we need that the inner level DUE problem conducted with Aimsun Meso generates not only the simulated flows and speeds at traditional detection stations, as in [Cipriani et al., 2011], but also the simulated travel time estimates from Bluetooth antennas along the corresponding paths, as detailed in [Barceló et al., 2014].

The computational results obtained in the former implementations of the Bilevel-DUE procedure were very high, therefore a potential research line is to search for more efficient implementations. In [Osorio et al., 2010] we can find a simulation-based optimization (SO) methods for complex problems with very tight computational budgets for which they report good computational results using a trust region. The main idea is to build, at each iteration, a search range which “trusts” in a neighborhood of the current solution. We have adapted it to generate the perturbations used in the SPSA scheme to approximate a gradient for the outer level function. The trust region ensures that the matrices generated by the SPSA algorithm remain in the right direction towards the real OD matrix. Avoiding replications of matrices outside of the trust region is essential for reducing the computational effort.

The paper will report on the benefits of the new procedure through computational experiments with various real-life networks.

2. Bilevel optimization approach for the estimation of time sliced OD matrices

An innovative framework proposed in [Cipriani et al., 2011], formulates the considered problem in terms of a bilevel nonlinear optimization model. Their scheme combines a Simultaneous Perturbation Stochastic Approximation (SPSA) and a Dynamic User Equilibrium (DUE) model. The outer level is a merit function, that updates the current matrix and the inner level at outer iteration k consists on the DUE whose input is the updated matrix.

2.1. Design and implementation of Bilevel-DUE

The Simultaneous Perturbation Stochastic Approximation (SPSA, [Spall, 1998]) path search optimization method is used as the solution algorithm: it works with a gradient approximation based on a perturbation of each component. Hence, all elements of the current matrix (at iteration k , x_k) are randomly perturbed together to obtain a new matrix for which the objective function is evaluated. In Fig. 1 we can see the SPSA algorithm applied to the dynamic O-D estimation problem:

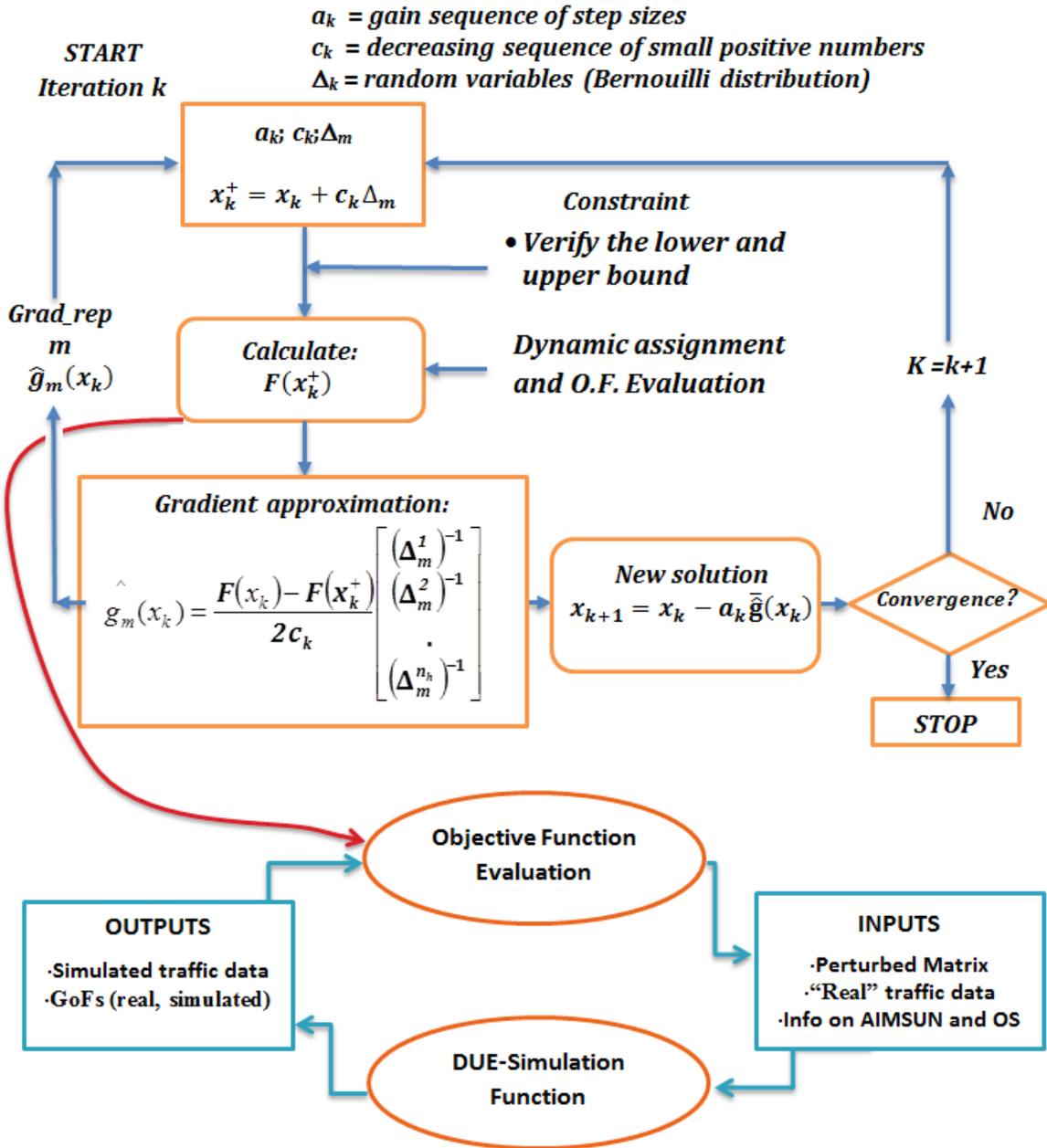


Fig. 1: SPSA algorithm applied to the dynamic O-D estimation problem

An extension and adaptation of the basic version proposed in [Cipriani et al., 2011] has been studied assuming the availability of travel times between Bluetooth sensors along the main paths connecting them in the network. The previous research reported in [Barceló et al., 2013] has proved that a suitable Bluetooth sensor layout allows the identification of the paths between sensors and therefore the measurement of the associated travel times. Consequently, to implement the proposed approach, we need that the inner level DUE conducted with Aimsun Meso generates not only the simulated flows and speeds at traditional detection stations, but also the simulated travel times estimates from Bluetooth antennas along the corresponding paths.

The computational experience reported in [Antoniou et al., 2014b], on the suggested procedure for off-line simultaneous adjustment of temporal OD matrices without introducing assignment matrices and adopting a simulation approach, finds the solution minimizing the following objective function (O.F.). This O.F. reports the “distance” between estimated and *a priori* OD demand flows and the error between actual and estimated traffic measurements: flows and speeds observations on links, and travel times between pairs of Bluetooth detectors along identified paths. Furthermore, in our proposal, the O.F. is enhanced with additional measures, such as densities and/or occupancies.

Let a network be $G = [N, A]$, where N are nodes and A are links; the period of analysis T , is divided into n_h intervals, and a subset of links $S = \{1 \dots n_s\} \in A$ equipped with sensors. The problem is finding an optimal demand vector $(d_1^* \dots d_{n_h}^*)$ such that:

$$(d_1^* \dots d_{n_h}^*) = \arg \min_{(x_1 \dots x_{n_h}) \geq 0} [f_1(x_1 \dots x_{n_h}, \hat{d}_1 \dots \hat{d}_{n_h}) + f_2(v_1 \dots v_{n_h}, \hat{v}_1 \dots \hat{v}_{n_h}) \\ + f_3(s_1 \dots s_{n_h}, \hat{s}_1 \dots \hat{s}_{n_h}) + f_4(t_1 \dots t_{n_h}, \hat{t}_1 \dots \hat{t}_{n_h})]$$

Where:

x_i estimated matrix for departing time interval i , $i = 1 \dots n_h$

v_i simulated volumes on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

s_i simulated speeds on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

t_i simulated travel times between Bluetooth sensors for departing time interval i , $i = 1 \dots n_h$

\hat{d}_i seed matrix for departing time interval i , $i = 1 \dots n_h$

\hat{v}_i traffic volumes on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

\hat{s}_i measured speeds on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

\hat{t}_i measured travel times between Bluetooth sensors for departing time interval i , $i = 1 \dots n_h$

About the term related to the distance with respect to a *a priori* OD demand matrix, preliminary outcomes have highlighted that the usual seed matrix term may too strongly constrain the search of the solution within a neighborhood of the initial matrix. Furthermore it was the dominating term and to make significant the contribution of other terms they had to be normalized. The computational experience also showed that the term of distances between x_i and \hat{d}_i had a twofold negative influence avoiding the adjusted matrix to move away from the target matrix on one hand, at a high computational cost on the other hand. However, removing this term had the opposite undesired effect and the adjusted matrix could move far away from the target one.

The finding with the new approach is that this term can be removed without the undesired effects if search is constrained within a trust region (which will be discussed in greater detail later, see section 3.2.2). Respect to different terms of the O.F., speed term showed a behavior worse than density or occupancy terms in earlier experiments. For this reason speeds have been replaced by occupancies. Therefore, assuming that $t_i = t_i(x_i)$, this means that travel times along paths are an implicit function of OD flows, our proposed O.F. is:

$$(d_1^* \dots d_{n_h}^*) = \arg \min_{(x_1 \dots x_{n_h}) \geq 0} [f_1(t_1 \dots t_{n_h}, \hat{t}_1 \dots \hat{t}_{n_h}) + f_2(v_1 \dots v_{n_h}, \hat{v}_1 \dots \hat{v}_{n_h}) \\ + f_3(o_1 \dots o_{n_h}, \hat{o}_1 \dots \hat{o}_{n_h})]$$

Where:

t_i simulated travel times between Bluetooth sensors for departing time interval i , $i = 1 \dots n_h$

v_i simulated volumes on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

o_i simulated occupancies on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

\hat{t}_i measured travel times between Bluetooth sensors for departing time interval i , $i = 1 \dots n_h$

\hat{v}_i traffic volumes on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

\hat{o}_i measured occupancies on links $\in S$ for departing time interval i , $i = 1 \dots n_h$

Each term in the O.F., represents the “distance” between observed and simulated data. To evaluate the O.F. for each new perturbed matrix, a DUE mesoscopic simulation must be performed in order to obtain the simulated data caused by this demand.

A number of goodness-of-fit (GoFs) measures can be used to evaluate the overall performance of OD estimation, as detailed in [Antoniou et al., 2014b]. After analyzing the trend of each goodness of fit measure between the considered seed matrix and the real one, it has been found that the NME (Normalized Mean Error) best represents the descendent direction through the desired solution (the real matrix). The O.F. is composed of three terms: flows, occupancies and travel times. Since we are using a normalized indicator, NME in each component (between simulated and observed data) returns a very similar value, and the weight of each part is considered equal to one.

2.2. Travel times

For the travel times term in the O.F., the sensor layout and the most likely used paths between them have been calculated on basis of the procedures already used in [Barceló et al., 2013]. Figure 2 depicts the optimized layout used with the model of Barcelona’s Eixample, highlighting the paths for which the travel times measured by Bluetooth antennas (blue circles) are available.

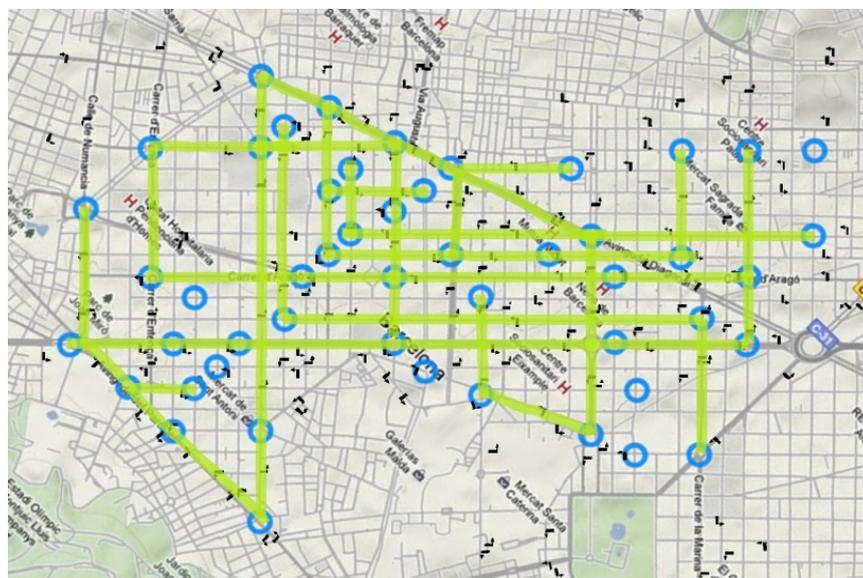


Fig. 2: Bluetooth layout and available paths

Consequently, to implement the proposed method, the Barcelona’s Eixample microscopic model has been modified to include the sensors, emulate the Bluetooth detection and retrieve the partial path travel times as depicted in Figure 3.

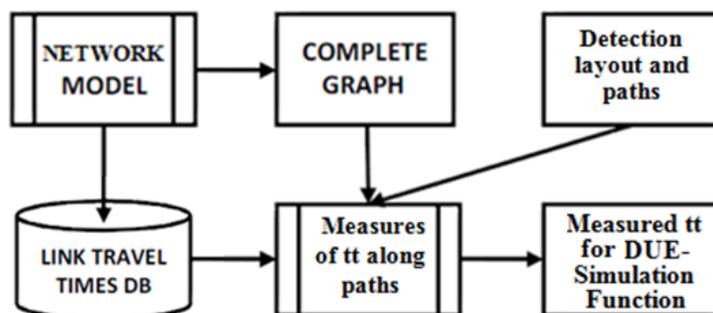


Fig. 3: Additional intermediate step in the bilevel procedure to estimate path travel times between pairs of Bluetooth antennas.

From the microscopic model we can extract the complete underlying graph of the urban network including all turnings and their associated penalties. Running the model for an estimated OD matrix $\hat{\mathbf{g}}(t_i)$ for a given time interval t_i a database with the estimated link travel times for that time interval can be generated. The link travel times, the graph of the urban network, the detection layout and the defined paths between pairs of Bluetooth antennas are the data to compute the measured path travel times \mathbf{tt} . These travel times will be one of the input data sets to the DUE-Simulation function which needs “real” traffic data (see Fig.1) in order to calculate the “distance” between observed and simulated data.

The logical diagram of the modified testing process in Figure 1 is depicted in Figure 4.

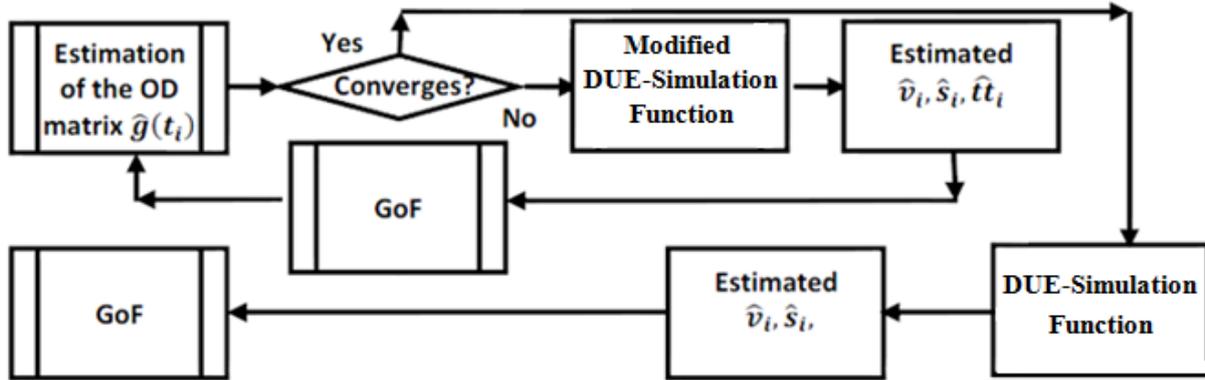


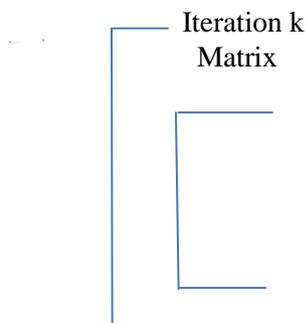
Fig. 4: Modified testing process

2.3. Asymmetrical Design

As detailed in [Spall, 1998] the choice of the gain sequences (a_k and c_k) and the simultaneous perturbation vector (SPV) is critical to the performance of SPSA. We use an Asymmetrical Design (AD) to reduce by half the number of objective function evaluations required in any simultaneous perturbation. In order to compute “m” gradient approximations at iteration k, “m” perturbations of the current matrix must be produced. Firstly, we run a simulation (Due-Simulation function call) using the current matrix at iteration k and the value of the gradient is obtained as follows:

$$gradient = \frac{objectiveplus - objective}{c_k} \cdot SPV \quad (1)$$

The *objective* value remains unchanged throughout the iteration k, but *objectiveplus* value varies in each of the “m” gradient approximations, since there is an inner loop where “m” perturbations of the matrix are created. The procedure consists schematically of the following two loops, an inner and an outer loop:



3. Computational testing

3.1. Vitoria

The first computational testing of the proposed approach was conducted using a scenario from the common evaluation and benchmarking platform developed within the framework of the EU COST Action MULTITUDE TU903. The main goal of this platform, detailed in [Antoniou et al., 2014b], was to provide a testbed in which a number of algorithms could be implemented and tested under the same conditions and our algorithm was one of them.

The selected scenario was a microscopic Aimsun simulation model of the network of Vitoria, in the Basque Country, made available by TSS to MULTITUDE project. Vitoria's network has 57 centroids and 2800 intersections (see Figure 5). Traffic data were collected from 389 loop detectors and 50 ICT (Information and Communication Technologies) detectors located using the layout models in [Barceló et al., 2012]. Almost 90% of the trips were collected twice at least in the peak-afternoon demand scenario of 1 hour of length, which accounted for 95% of the number of OD pairs and 86% of the most likely used paths identified in a DUE assignment with the 'true' historic OD matrix.

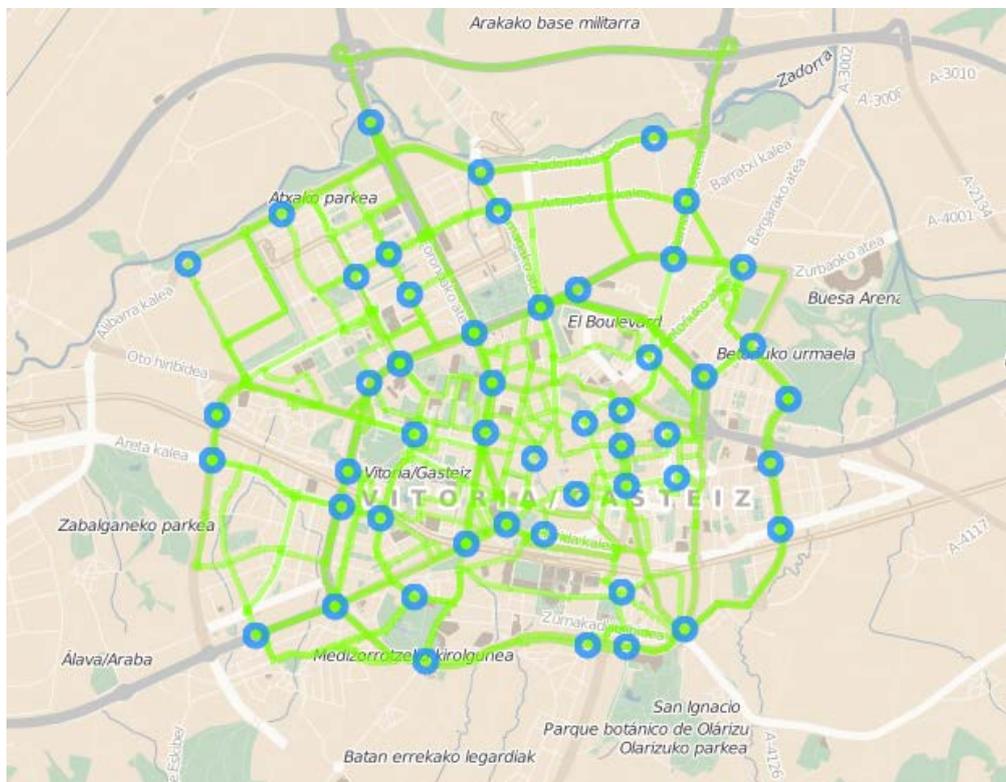


Fig. 5: Vitoria's network and subnetwork covered by ICT sensor layout.

In Vitoria network we chose 50 pairs of antennas and we calculated the most likely routes between them. The procedure **returns simulated partial path travel times** on these predefined and stored routes. The main goal of the experiments in Vitoria's network was to evaluate the importance of travel times term in the O.F. and in the matrix estimation process. Then, two versions of the Bilevel-DUE were considered: one without travel times and another one incorporating them.

3.1.1 Results for Bilevel-DUE in Vitoria's network

Since MULTITUDE platform was using Aimsun, the DUE-Function uses Aimsun Meso [TSS-Transport Simulation Systems, 2013], to perform the DUE assignment.

In Figure 6, the total OD flows for the 1h peak-afternoon period are shown, real versus estimated, for both versions of the Bilevel-DUE scheme for off-line matrix estimation, the original one without including travel times and the modification proposed by the authors. Coefficient of determination is almost 95% when times are included, higher than the 91% obtained without the travel times term.

These results confirmed that the solution is improved when travel times are taken into account in the objective function.

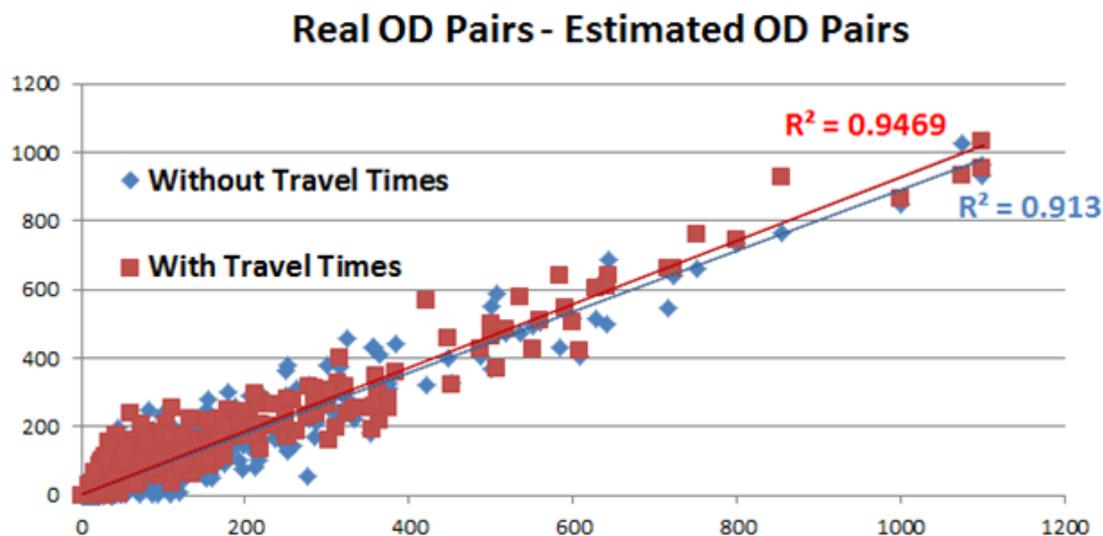


Fig. 6: Vitoria network: Bilevel-DUE real vs estimated total OD flows (1h period).

3.2. Eixample

Since the experience with the off-line Bilevel-DUE modified with the times term gave interesting results in Vitoria network, we decided to test our method in the network of Barcelona's Central Business District (CBD), Eixample, consisting of 2111 sections, 1227 nodes (grouped in intersections), 120 generation centroids, 130 destination centroids and a total of 877 non-zero OD pairs. Figure 7 depicts the Aimsun microscopic simulation model used in the computational experiments.

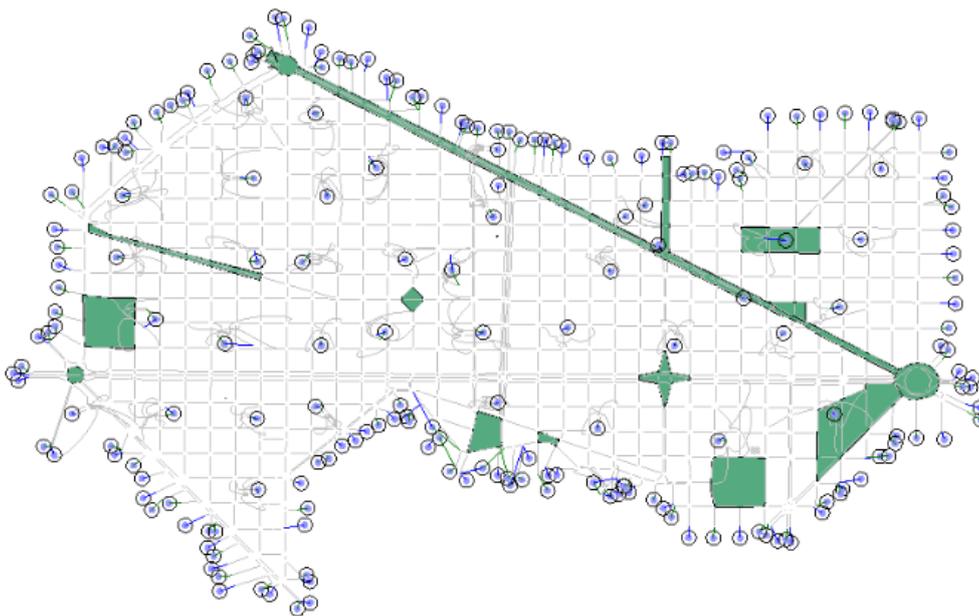


Fig. 7: Aimsun model Barcelona's Eixample

Previously to run the experiments, a swept between the seed and real matrices is performed. The goal of this is to understand the evolution (or sensitivity) of the objective function's value as the input ranges from the "seed" matrix to the "real" matrix. After studying the results and the descent direction, a suitable parameter was set for the gain sequence c_k (Eq. 1), included for the gradient update.

As detailed in [Cipriani et al., 2013], Figure 8 reports an unidimensional swept of the O.F., using a convex combination between the seed ($\alpha = 0$) and the real matrix ($\alpha = 1$) applied to the model of Barcelona's Eixample:

$$\text{newmatrix} = \alpha(\text{real matrix}) + (1 - \alpha)(\text{seed matrix})$$

with $\alpha \in [0,1]$

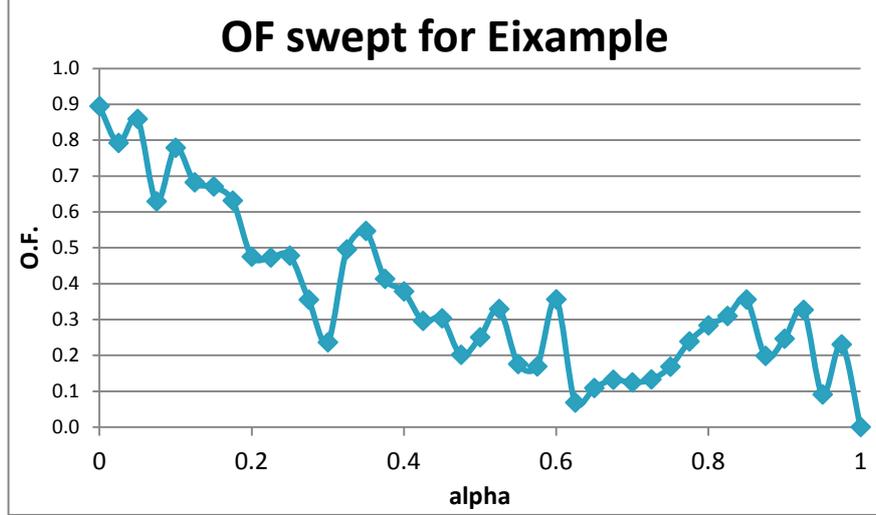


Fig. 8: O.F. unidimensional swept

3.2.1 Conjugate Gradient

The value of the gain sequence a_k (Eq. 2), which is used for the solution update, was chosen depending on the average of the calculated gradient approximations. The dimension of the gradient is related to the statistical goodness of fit measure used in the O.F. (in this case NME). The average gradient is used for updating the current time-sliced OD at outer iteration k , with which the next iteration begins. Reducing the number of replications is a key factor, since replications are time-consuming for large networks.

In [Cantelmo et al., 2014] is suggested that the computation of the approximated average gradient could be enhanced using a conjugate gradient strategy. It is known that the generated conjugate directions permit to reach faster the solution than using the basic gradient method. When we calculate the new solution (see Fig.1), the formulation can be replaced by:

$$x_{k+1} = x_k - a_k(\bar{g}(x_k) + \beta\bar{g}(x_{k-1})) \quad (2)$$

With

$$\beta = \frac{\|\bar{g}(x_k)\|^2}{\|\bar{g}(x_{k-1})\|^2}$$

This modification does not require additional computational times per iteration, since previous average approximated gradients can be stored and their values will be available to calculate β .

3.2.2 Trust Region

The computational costs required by the former implementations of the Bilevel-DUE procedure were very high, therefore a potential research line aims to develop more efficient solutions. In [Osorio et al., 2010], we can find simulation-based optimization (SO) methods for complex problems with very tight computational budgets for which they report a very promising computational performance. The efficiency and suitability of the method is due to its integration within a trust region scheme. The main idea of trust region is to build, at each iteration, a search range which “trusts” in a neighborhood of the current solution. Following this approach, we have adapted it to generate the perturbations of the current matrix involved in the computation of the descent direction. Hence, firstly we analyzed the size and the number of trips of the seed matrix. Based on this analysis, a search range is established for each time slice in particular and for the number of trips of the entire matrix in general. Let be n_s the number of time slices. Then, the trust region can be formulated as follows:

$$(1 - \gamma_1) * TS_j(\hat{d}) < TS_j(X^+) < (1 + \gamma_1) * TS_j(\hat{d}) \quad \gamma_1 \in (0.01, 0.1)$$

$$(1 - \gamma_2) * \hat{d} < X^+ < (1 + \gamma_2) * \hat{d} \quad 0 < \gamma_2 \leq \gamma_1$$

Where:

$TS_j(\hat{d})$ number of trips in the time slice j of the seed matrix, with $j = 1 \dots n_s$

$TS_j(X^+)$ number of trips in the time slice j of the perturbed matrix, with $j = 1 \dots n_s$

\hat{d} number of trips of the seed matrix

X^+ number of trips of the perturbed matrix

It is important to remember that for each perturbed matrix a DUE has to be performed. The trust region ensures in this way that the matrices generated by the SPSA algorithm remain in the right direction towards the real OD matrix. Avoiding replications of matrices outside of the trust region is essential to reduce the computational burden.

In Figure 9.a and 9.b, fifty perturbed matrices from the seed matrix used in our experiments are represented. Any of them could have been chosen by our method to evaluate the O.F. and update the solution. However, with the trust region we ensure that the chosen matrix is near to the previous one. If we do not move far away from the current matrix, a more reliable convergence process to the real matrix is achieved. The x axis of the Figure 9 shows the number of trips of the matrix and the y axis the number of trips of the first time slice (our matrix has five time slices whose components are simultaneously perturbed).

If the starting point is an obsolete matrix obtained using census and sample information, the trust region parameters can be set to find perturbed matrices with a higher number of trips than the preexisting origin-destination matrix (see Figure 9.b). Otherwise, if no information is available or for testing purposes of our algorithm with different demands, the trust region will consider matrices either with less or greater number of trips than the seed matrix (see Figure 9.a).

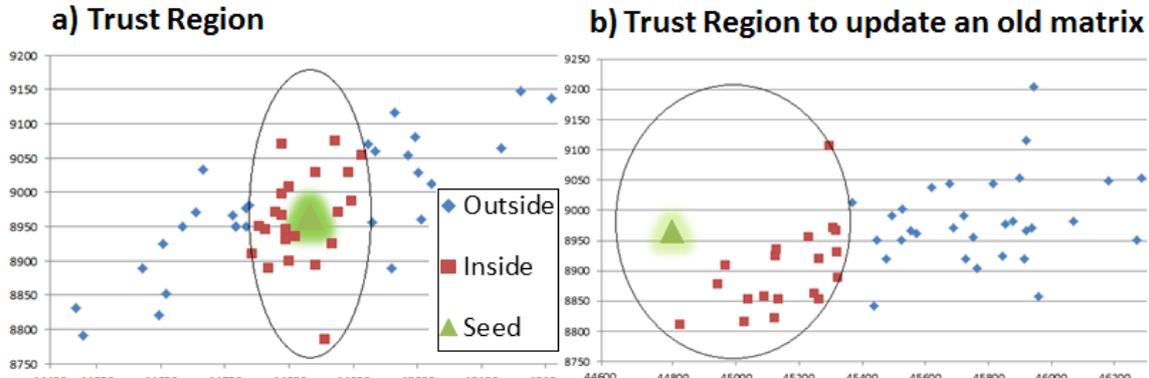


Fig. 9 a) Trust region b) Trust region to update an old matrix

4. Experiments with different demand levels

We conducted a set of experiments in the network of Barcelona's Eixample where OD pattern is common for all 5 slices, but intensities are slice-varying with a total number of trips of 59774. The 'real' OD matrix has five different time slices, accounting for 20%, 28%, 20%, 12% and 20% of the total flow and corresponds to a synthetic time-sliced matrix of 1h 15min obtained to represent a point of raised to congestion according to network's MFD (Macro Fundamental Diagram) detailed in [Daganzo et al., 2008]. There are 116 loop detectors and 50 ICT were located in intersections (see Figure 2) following indications given in Barceló et al (2012).

In the experiments presented below, the starting time-sliced OD matrix accounts for a percentage of the real OD matrix demand in all slices, and the OD Pattern is perturbed for the most important origin centroids in the network. The aim of these experiments is to test the improvement obtained using the trust region and the conjugate gradient discussed above.

Figure 10 depicts the results of the first twenty iterations of the Bilevel-DUE algorithm in Barcelona's Eixample. The algorithm works as follows: at iteration k , we start from the matrix x_k and finish with the new updated matrix x_{k+1} . Then, the O.F. is evaluated at x_{k+1} and if its value is lower than the previous one, we proceed to the next iteration $k+1$. Otherwise, if the value of the O.F. is greater than the previous one, the matrix x_{k+1} is deleted, repeating the process until a better solution x_{k+1} is found. Therefore, we remain at iteration k until we obtain a matrix which reduces the value of the objective function. Iteration k to $k+1$ can actually lead to perform several "attempts" of iterations that are thrown away, and the computational time that it entails is lost.

4.1. Demand levels

The level of demand is a key element affecting the performance of OD estimations. It is well known that the problem becomes harder under congestion, as detailed in [Antonioni et al., 2014b]. This experimental design considers two levels of demand. In constructing the two "true" demand levels, each entry in the demand matrix (defined by an OD period and departure interval) was perturbed randomly according to the following patterns, which capture various conditions around the base demand D for each network (i.e. both higher and lower):

1. Low demand (denoted as D7 in the results previously presented): $D*[0.7+0.3*\text{rand}(\)]$. The mean value of this demand is 75% of the base demand, with a range of +/- 15%.
2. High demand (D9): $D*[0.9+0.3*\text{rand}(\)]$: The mean value of this demand is 115% of base demand, with a range of +/- 15%

4.1.1 Low demand (D7)

The following Figure 10 shows the results obtained when the low demand (D7) is considered as the seed matrix. In the first twenty iterations, the O.F. (as a whole, but also each individual component) shows approximately a 95% reduction. Without the trust region and the conjugate gradient the algorithm needs 20 iterations to converge (which were actually 63 as explained in the previous paragraph). At each iteration the algorithm requires on average 3.15 attempts to reduce the objective function. As we can see in Figure 10, the value obtained after 20 iterations is worse than the obtained with the algorithm that includes the trust region and the conjugate gradient. The latter only needs 10 iterations to converge (which were actually 28, i.e. on average 2.8 attempts per iteration). The use of the trust region is effective since the algorithm in each of the first five iterations requires only one attempt to reduce the objective function. Without the trust region, the number of attempts rises to three or four at the first iterations with the consequent loss of time. The conjugate gradient also improves the performance because it converges using fewer iterations and the solution is reached faster.

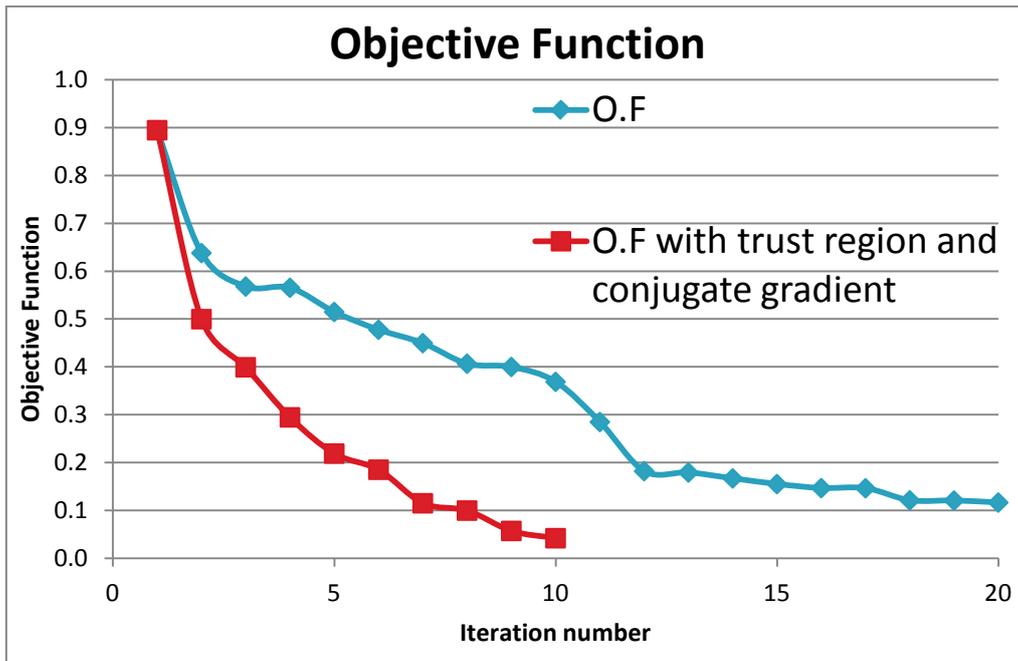


Fig. 10: Objective Function in Barcelona's Eixample

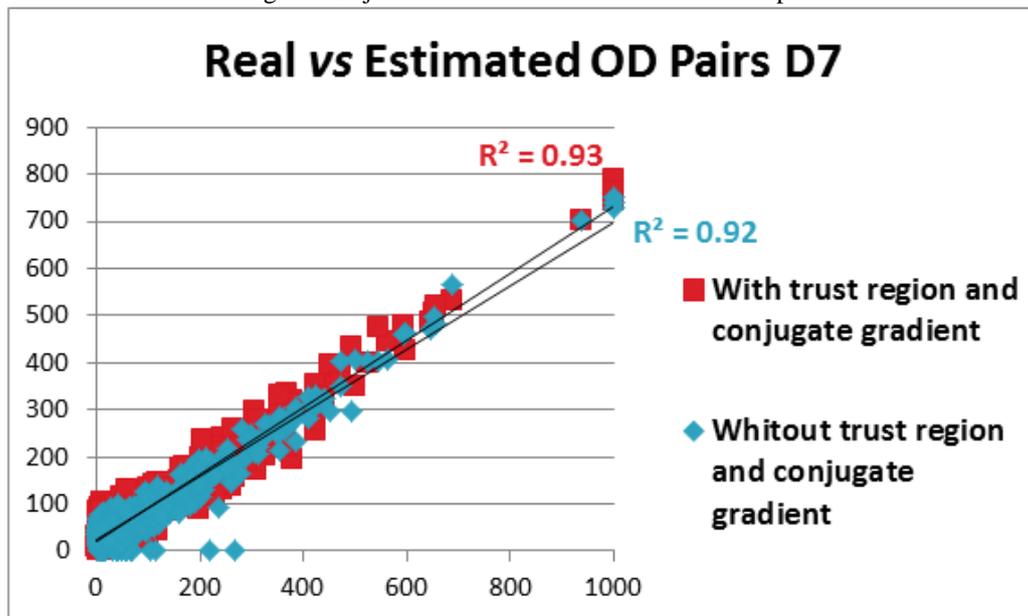


Fig. 11: Real vs Estimated total OD flows (1h15 min) with D7 demand

The regression between real and estimated total OD flows for the period of study (1h15min) shows a coefficient of determination of 93% when trust region and conjugate gradient are included (92% when they are not considered, see Figure 11). Consequently, results obtained using jointly the trust region and conjugate gradient are improved since a significant reduction in the computational time is achieved.

4.1.2 High demand (D9)

Next step in our experiments is to test the performance of the algorithms when the demand is high (D9). As we can see in the following Figure 12, the main difference between the algorithm with the trust region (TR) and the conjugate gradient (CG) and the algorithm without both of them (referred as base algorithm) is again the number of attempts to find a new minimum in the O.F. at each iteration. This number of attempts grows

exponentially when the algorithm doesn't use the trust region and the conjugate gradient, leading to 9 iterations to converge (that were actually 46, needing 18 to move from the eighth to ninth iteration). However, if the algorithm includes the trust region and the conjugate gradient, it only needs 5 iterations to converge (which were actually 11) and the computational time is 4.18 times lower.

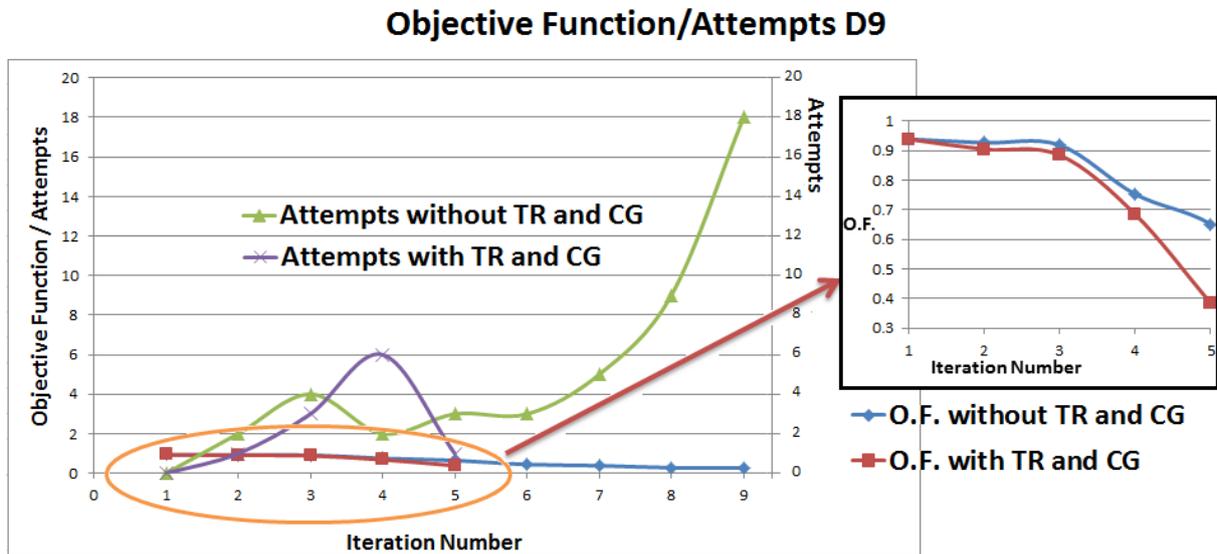


Fig. 12: Attempts per iteration and O.F. in Barcelona's Eixample with D9 demand.

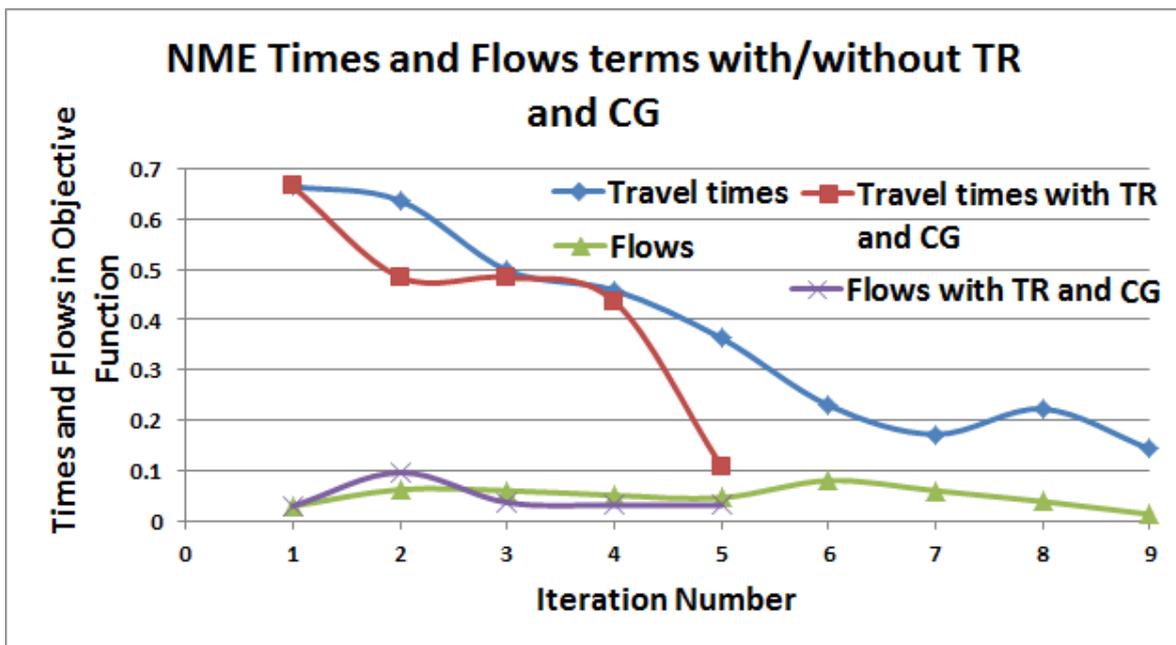


Fig. 13: Travel times and Flow terms progress along Bilevel-DUE iterates without and with TR and CG

In the above Figure 13, the progress of travel times and flow terms in the O.F. along Bilevel-DUE iterates is shown. We can see that with high demand (D9), the travel time term in O.F. is greater than flow and occupancy terms. Again, with the trust region and the conjugate gradient, the values obtained for these terms are similar or even lower than in the situation without both variants. Moreover, in this case the regression between real and estimated total OD flows for the period of study (1h15min) shows a coefficient of determination of 92%, slightly better than the 91% obtained by the base algorithm (see Figure 14). And above all, it reduces four times the computational time.

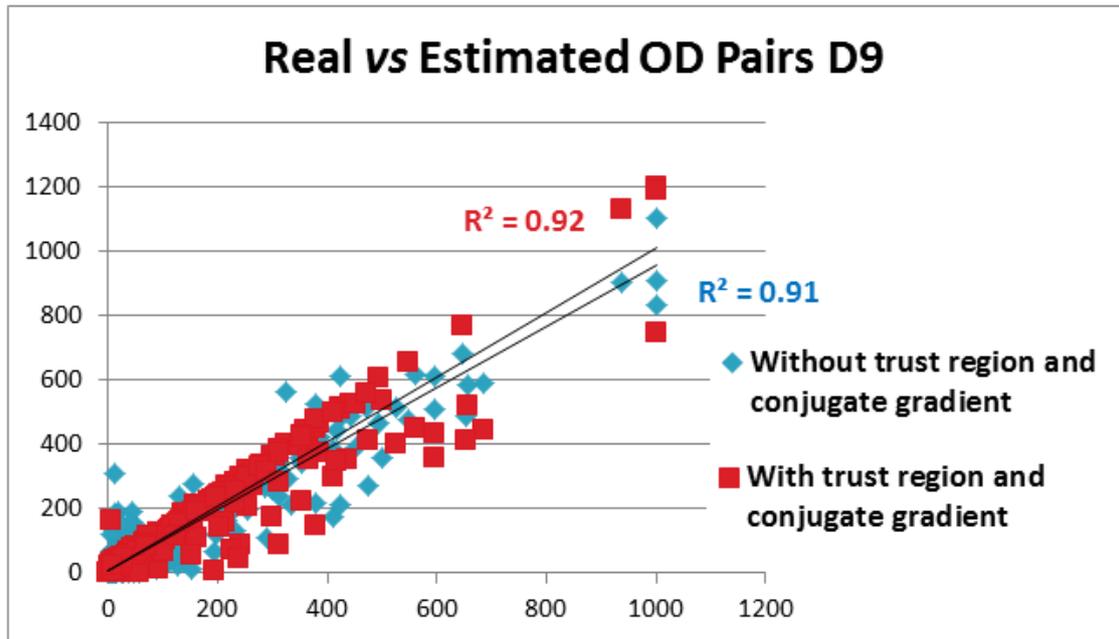


Fig. 14: Real vs Estimated total OD flows (1h15 min) with D9 demand

4.2. Changes in the structure of the matrix

Since the experience with the former demand levels provided promising results, we propose to test our method with a seed matrix where demand is not obtained randomly as previously. Starting from the base demand D , we select the three heavier origins (which inject greater number of trips on the network). For each origin, the number of trips is equally distributed between all its destinations, i.e., all OD pairs with the same origin will have the same number of trips. Therefore we are completely changing the structure of the matrix. This demand will be named H3.

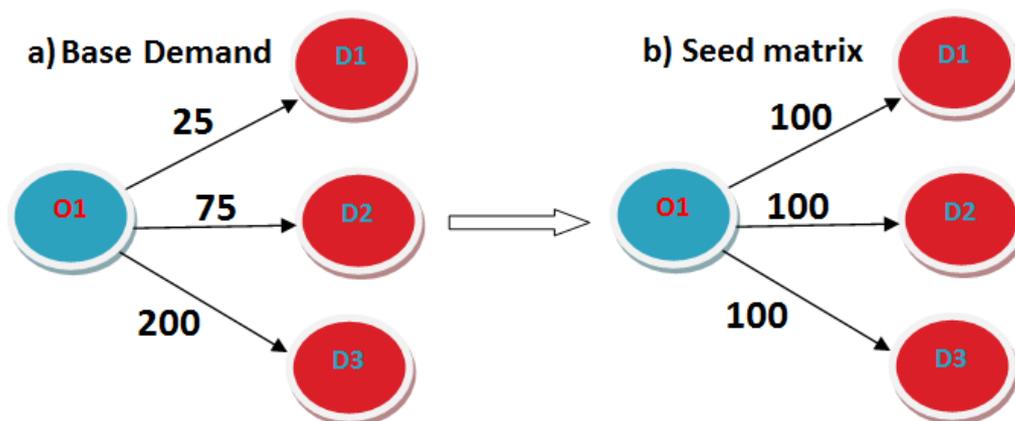


Fig. 15: Seed matrix structure for H3 demand (O1: origin 1, D_i : destination i)

In this case, as depicted in Figure 16, the base algorithm requires 7 iterations to converge (that were actually 83, needing 55 to move from the sixth to seventh iteration). However, if the algorithm includes the trust region and the conjugate gradient, then it needs 6 iterations to converge (which were actually 40). The computational time is halved obtaining a lower value of the objective function. The regression between real and estimated total OD flows for the period of study (1h15min) shows a coefficient of determination of 86%, greater than the 80% obtained with the base algorithm.

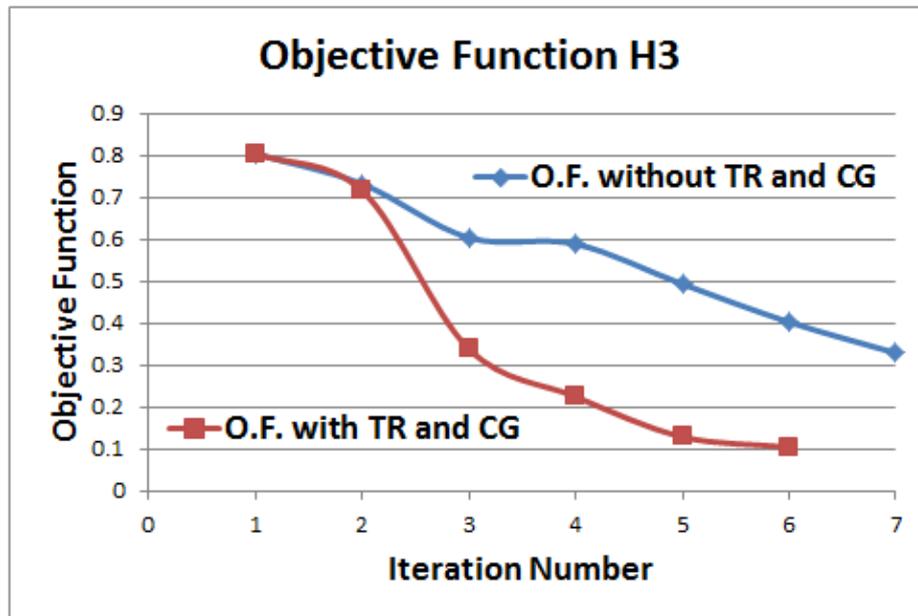


Fig. 16: Objective Function in Barcelona's Eixample with H3 demand

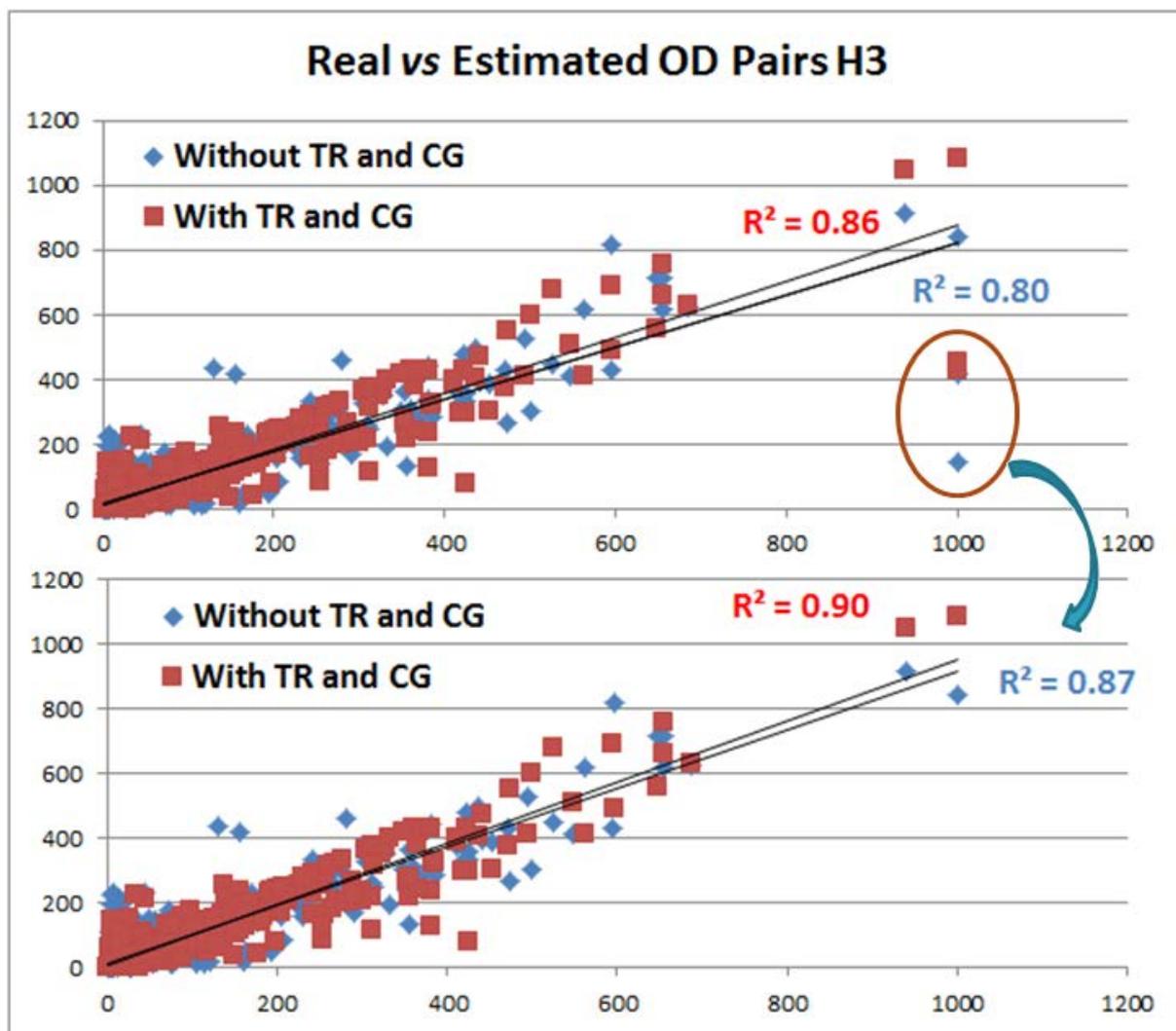


Fig. 17: a) Real vs Estimated total OD flows (1h15 min) with H3 demand

In the above Figure 17, some OD pairs are highlighted inside a circle figure. They were firstly identified as outliers and the coefficient of determination would be improved deleting these observations.. According to the x axis these OD pairs have one thousand trips in the “real” matrix. However, when the structure of the matrix is modified and the number of trips of one origin is divided between all its destinations, these OD pairs have only 150 trips in the seed matrix (the difference is 850 trips). The value of the gain sequences a_k (Eq.2) and c_k (Eq.1) and the trust region cannot raise to catch the number of trips of these particular cells. In fact, their values grow up to 400 trips, but they don't reach the one thousand trips existing in the real matrix (actually it could be achieved but affecting negatively the rest of OD pairs). A task that requires further research is to consider specific gain sequences and trust region for each OD pair.

4.2.1 H-all

Like in the previous experiment, again the structure of the matrix will be completely changed. Now we select all the origins, not only the three most important. As before, for each origin the number of trips is equally distributed between all its destinations. This new demand will be named H-all.

The algorithm with the trust region and the conjugate gradient exhibits the best performance: it gets a lower value in the O.F., needing fewer attempts to find it at each iteration (see the Figure 19). Also, the regression between real and estimated total OD flows shows a coefficient of determination of 65% (with the base algorithm is 50%, see Figure 20).

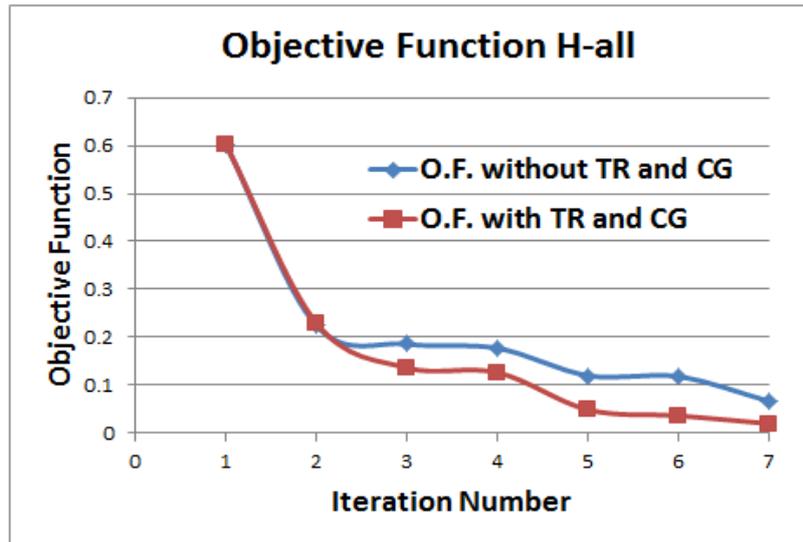


Fig. 18: Objective Function in Barcelona's Eixample with H-all demand

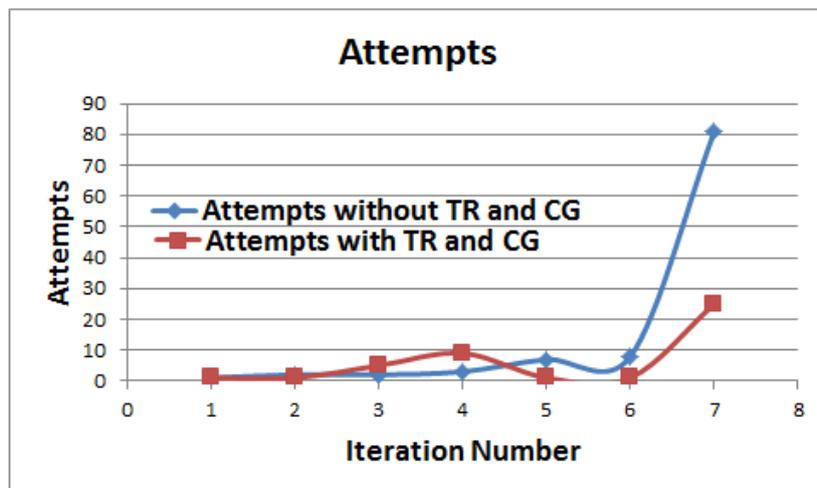


Fig. 19: Attempts per iteration with/without TR and CG

Although this coefficient of determination (65%) may seem inadequate, we have to keep in mind that the seed matrix with its completely changed structure returns a coefficient of determination of 27%. Therefore, the algorithm shows its efficiency by improving on 38 points of the coefficient of determination, as depicted in Figure 21.

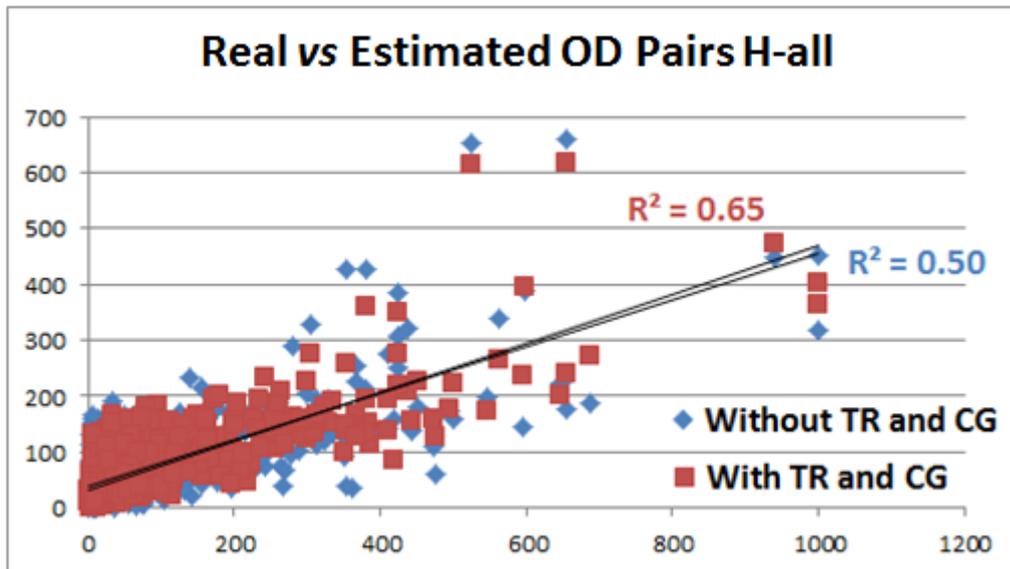


Fig. 20: Real vs Estimated total OD flows (1h15 min) with H-all demand

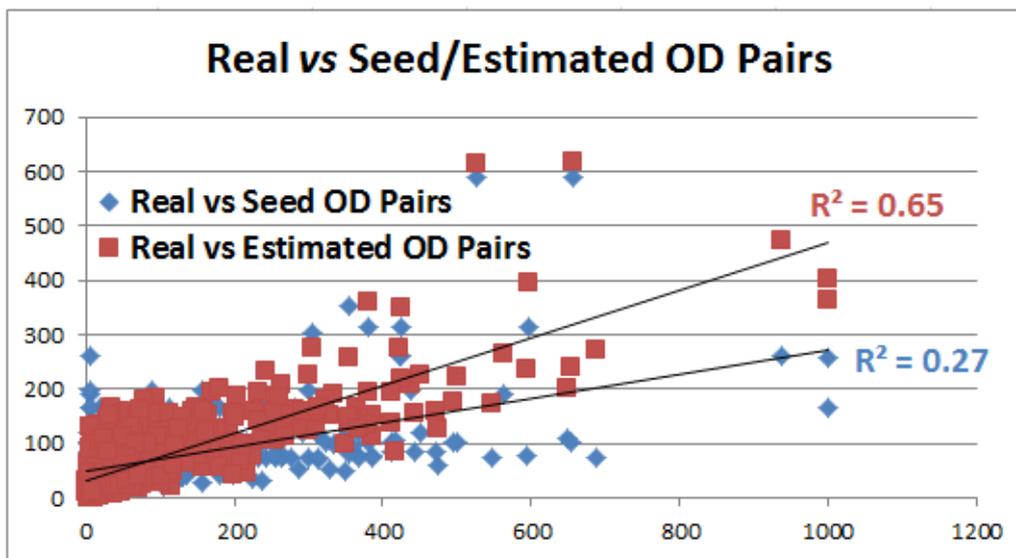


Fig. 21: Real vs Seed and Estimated total OD flows (1h15 min) with H-all demand

5. Conclusions

The modified bilevel, solving heuristically the inner level optimization by a DUE using a mesoscopic simulation for the Dynamic Network Loading (DNL), enables the exploitation of ICT measures and provides quite good time-sliced OD suitable for the initialization of the on-line and real-time dynamic OD estimation methods based on Kalman filtering. The computational times make it useful and advisable for off-line applications.

Regarding to the two main changes proposed in this paper, the conjugate gradient and the trust region, their contribution to improve the convergence and efficiency of the algorithm has been proven. The use of the trust region is effective since the algorithm requires fewer attempts at each iteration to reduce the objective function.

The conjugate gradient also enhances the algorithm performance because it converges faster using fewer iterations.

The incorporation of the conjugate gradient and the trust region also improves the quality of the results. Minimum values found in the objective function are lower with them than without them. Likewise, with them the regression between real and estimated total OD flows for the period of study (1h15min) shows a higher coefficient of determination in all cases. Overall, the number of iterations is halved and the computational time is reduced up to four times with the trust region and the conjugate gradient.

The computational experiments presented in this paper prove the robustness and quality of the O-D estimates of the proposed Bilevel-DUE approach exploiting ICT measurements. Robustness has been tested in relation to the quality of the initial O-D information. The experiments were conducted on a significantly sized, realistic network: Barcelona's CBD. The results show that it is possible to apply a Bilevel-DUE approach for estimating dynamic O-D matrices in traffic management systems.

References

- Antoniou,C., Barceló,J., Brackstone,M., Celikoglu,H., Biaggio,C., Punzo,V., Sykes,P., Toledo,T., Vortisch,P., Wagner,P., (2014). Traffic Simulation: Case for guidelines, Edts. V. Punzo and M. Brackstone, Publications Office of the European Union, JRC88526, available at: <http://publications.jrc.ec.europa.eu/repository/handle/11111111/30680>.
- Antoniou,C., Biaggio,C., Montero,L., Casas,J., Barceló,J., Cipriani,E., Djukic,T., Marzano,V., Nigro,M., Bullejos,M., Perarnau,J., Breen,M., Punzo,V., Toledo,T., (2014b). A framework for the benchmarking of OD estimation and prediction algorithms, Paper 14-2793 presented at the 93rd Annual Meeting of the Transportation Research.
- Balakrishna,R., Ben-Akiva,M., Koutsopoulos,HN., (2007). Off-line Calibration of Dynamic Traffic Assignment: Simultaneous Demand-Supply Estimation. Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington DC.
- Barceló, J., Gilliéron, F., Linares, M.P., Serch, O., Montero, L., (2012). Exploring Link Covering and Node Covering Formulations of Detection Layout Problem. Transportation Research Records 2308, pp.17-26.
- Barceló,J., Montero,L., Bullejos,M., Linares,M.P., Serch,O., (2013). Robustness and computational efficiency of a Kalman Filter estimator of time dependent OD matrices exploiting ICT traffic measurements. TRR Transportation Research Records: Journal of the Transportation Research Board, No.2344, pp.31-39.
- Barceló,J., Montero,L., Bullejos,M., Linares,M.P., (2014). A practical proposal for using origin-destination matrices in the analysis, modeling and simulation for traffic management, Paper 14-3793, presented at the 93rd TRB Annual Meeting, included in the CD-Compendium of Papers
- Cantelmo, G., Cipriani, E., Gemma, A., Nigro, M., (2014). An adaptive bi-level gradient procedure for the estimation of Dynamic Traffic Demand. IEEE Transactions on Intelligent Transportation Systems 1524-9050.
- Cipriani,E., Florian,M., Mahut,M., Nigro,M., (2011). A gradient approximation approach for adjusting temporal origin–destination matrices. Transportation Research C, 19(3),270-282
- Cipriani, E., Gemma, A., Nigro, M.,(2013). A bi-level gradient approximation method for dynamic traffic demand estimation: sensitivity analysis and adaptive approach. 16th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013). The Hague, The Netherlands, October 6-9, 2013.
- Daganzo, C.F. and Geroliminis, N. (2008) An analytical approximation for the macroscopic fundamental diagram of urban traffic. Trans. Res.B 42, 2008.
- Osorio,C., M.Bierlaire, (2010). A simulation-based optimization approach to perform urban traffic control". In Proceedings of the Triennial Symposium on Transportation Analysis (TRISTAN). Tromsø, Norway

Spall, J.C. (1998) An overview of the simultaneous perturbation method for efficient optimization. Johns Hopkins APL Technical Digest, 19 (4) (1998), pp.482–492.

TSS-Transport Simulation Systems (2013). Aimsun 7 Dynamic Simulator User's Manual, TSS, Barcelona, Spain.