A New Model for Updating Origin-Destination Matrices and Path Choice Probabilities in Public Transportation Networks[1]

M. Victoria Chávez-Hernández
Universidad Autónoma de Nuevo León (UANL)
Graduate Program in Systems Engineering
Av. Universidad s/n, Cd. Universitaria
San Nicolás de los Garza, NL 66455, Mexico
E-mail: *victoria@yalma.fime.uanl.mx*

L. Héctor Juárez Valencia
Mathematics Department
Universidad Autónoma Metropolitana Unidad Iztapalapa
186 Av. San Rafael Atlixco, Mexico
E-mail: *hec@xanum.uam.mx*

Roger Z. Ríos-Mercado
Universidad Autónoma de Nuevo León (UANL)
Graduate Program in Systems Engineering
Av. Universidad s/n, Cd. Universitaria
San Nicolás de los Garza, NL 66455, Mexico
E-mail: *roger.rios@uanl.edu.mx*

Yasmín Á. Ríos-Solís
Tecnológico de Monterrey
Escuela de Ingeniería y Ciencias
Av. Eugenio Garza Sada 2501 Sur
Monterrey, NL, Mexico
E-mail: *yasmin.riossolis@tec.mx*

February 2021

[1]Preprint PISIS-RR21-01, Graduate Program in Systems Engineering, UANL, San Nicolás de los Garza, Mexico, February 2021

Abstract

The inverse problem of estimating an actual origin-destination (OD) matrix is one of the most important public transit planning steps. Most of the time, the route choice probabilities by which a user transits along the network are first calculated, and then the OD matrix is updated. In this article, these two steps are solved simultaneously by an integer linear programming model based on partial knowledge of the transit segment flow along the network. A novelty in our methodology to avoid a quadratic objective function is to measure the difference between the reference and the estimated OD matrices by defining demand deficits and excesses for the estimated OD matrix. To test our methodology, we build an instance generator based on small-world graphs to mimic real transit networks. The results are compared with an augmented Lagrangian model solved by the dual ascent and multipliers method. Our integer linear model yields high-quality estimators of the actual OD matrices, the exact flow volume segment counts, and an adequate interval for the route choice probabilities. Additionally, we test our methodology in a real-world case from the city of Monterrey, Mexico.

Keywords: Origin-destination matrix; Transit counts; Integer linear programming; Transit assignment problem.

1 Introduction

The dynamics of the city in terms of population and mobility are among the most critical challenges we face nowadays. A well-planned public transport system is an essential asset for a more livable and sustainable city. However, public transport systems often fail to offer a high-quality service that may imply long travel times. This failure may be due to the lack of information about the daily trips taking place at each period of the day, the specific trip purpose (work, school, hospital, entertainment), and how people move in the public transportation system. This information is known as the Origin-Destination matrices (OD matrices). They are usually obtained from home-based surveys every ten years [3], they are expensive, and their data processing is time-consuming (six months to one year in our case study). Thus, once the new OD matrix is available, it may be already obsolete but it may be considered as a reference matrix to be updated. For these reasons, the inverse problem of estimating the actual OD matrix based on an reference one and people flow observations at some transit segments is relevant. The basis of the planning, operation, and control of a public transport system rely on high-quality OD matrices even if demand and travel times usually follow time-dependent stochastic patterns [22, 2, 19]. OD matrices are relevant at each stage of the bus network planning process that is usually divided into two main stages [14]. The first stage is the tactical planning, where correct OD matrices are needed for the bus line design and the generation of useful timetables (or departure times of all trips). This stage focuses on offering a high-quality service for the customers: line frequency, waiting times, and short transfers, as mentioned by Ibarra-Rojas et al. [13]. The second stage is the operational planning, where the vehicle and crew scheduling problems seek to minimize the transport system operating costs. Having updated OD matrices guide the decision-makers when more trips are needed, when a driver does not show up, or when there are accidents and the network should be rapidly restored [4]. Moreover, updating OD matrices allows us to test the current system under more demanding scenarios and adapt the future demand infrastructure. Figure 1 shows an example of a transit

	0	1	2	3	4	5
0	0	**100**	80	67	96	37
1	20	0	40	44	77	81
2	34	22	0	32	153	121
3	33	15	47	0	95	64
4	43	98	11	33	0	106
5	55	15	23	39	87	0

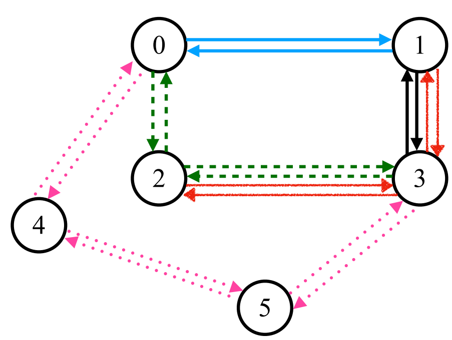


Figure 1: Reference OD matrix (left-hand side) of the transit network (right-hand side) with five lines and six stops.

network and its corresponding OD matrix. Notice that the diagonal entries of the OD matrix are 0 and that it is not a symmetric matrix (in the morning, people go downtown and few to the peripheries). The bold entry of the OD matrix means that 100 trips originate at node 0 to their final destination at node 1 in a certain period of the day. Nevertheless, this OD matrix is obsolete, so we aim to update the number of trips for every entry (we also say that the demand of the (0,1) pair must be updated). To update an obsolete OD matrix, we use flow volume observations made at some network segments obtained by fare-box and automated fare collection systems, automatic passenger counter systems, geographical positioning on cellular phones, or even surveillance videos. In this work, we estimate public transit OD matrices and consider small perturbations of the route choice probabilities to match some network flow observations. We name this OD matrix estimation as the ODA problem. Let us illustrate the ODA problem to show its importance. Consider the

lines	headway min.	transit segment (l, i, j)	travel time min.
1-blue	12	(1,0,1)	25
		(1,1,0)	25
2-green	12	(2,0,2)	7
		(2,2,0)	7
		(2,2,3)	6
		(2,3,2)	6
3-red	30	(3,2,3)	4
		(3,3,2)	4
		(3,3,1)	4
		(3,1,3)	4
4-black	6	(4,3,1)	10
		(4,1,3)	10
5-pink	30	(5,0,4)	9
		(5,4,0)	9
		(5,4,5)	12
		(5,5,4)	12
		(5,5,3)	8
		(5,3,5)	8

Table 1: The five lines (blue, green, red, black, pink) of the network of the right side of Figure 1 with with the headways and the travel times for each segment.

transit network of Figure 1 with five lines (blue, green, red, black, and pink) and six stops. Table 1 specifies the headway of the line: the difference in minutes between two different vehicles at the depot. It also indicates the travel time in minutes for each line-node-node transit segment (l, i, j). An actual but unknown OD matrix is associated with the network of Figure 1. Thus we must infer it with the observed flow volumes at some transit segments and the reference OD matrix (left-hand side of Figure 1). Suppose that we want to update the $(0, 1)$ entry of the obsolete OD matrix. Figure 2 shows only the links that a user would take to go from origin 0 to destination 1. Notice that the users paths (or strategy) do not consider taking the pink line since it would increase their travel times, as it can be validated with data from Table 1. The first and second columns of Table 2 indicate the lines and segments of the network depicted in Figure 2, respectively. The third column is the path choice probabilities that mimic how a user moves in a network and is based on a decision choice model [5]. In this work, we consider that the obsolete path choice probabilities may be updated, as shown in the fifth column of the table. The obsolete OD matrix indicates that

lines	transit segments	obsolete probabilities	updated probabilities
1 blue	(1,0,1)	0.500	0.475
2 green	(2,0,2)	0.500	0.525
	(2,2,3)	0.500	0.525
3 red	(3,2,3)	0.000	0.000
	(3,3,1)	0.080	0.090
4 black	(4,3,1)	0.420	0.435

Table 2: Transit segment probabilities associated to Figure 2.

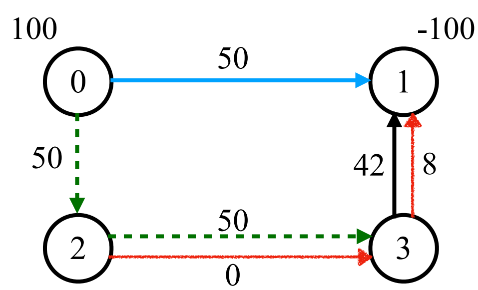


Figure 2: Transit segments that a user may take to go from origin 0 to destination 1 and the flow volume at each segment.

stop 0 generates 100 trips going to node 1 (-100 means that node 1 attracts 100 trips from node 0). Values in the segments of Figure 2 correspond to the number of passengers using each segment to go from 0 to 1 according to the obsolete OD matrix and the path choice probabilities of the lines, in the third column of Table 2. With the obsolete path choice probabilities, half of the passengers use the blue line; the other half use the (2,0,1) and the (2,2,3) segments of the green line. None use the (2,2,3) segment, while 42% of the passengers arriving at node 3 take the (4,3,1) segment. Finally, only 8% take the (3,3,1) link to arrive at node 1. Let us suppose that the number of trips originated at node 0 and destined to node 1 is no longer 100 but 200, also that the infrastructure of the network has not change. Thus, we suppose that the actual (0,1) OD matrix entry is known. Using the obsolete probabilities of Table 2 we obtain Figure 3 with the updated volumes at the

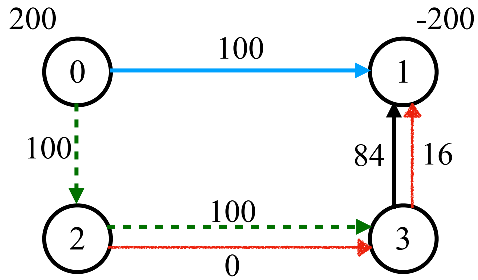


Figure 3: The segment flows are generated with the obsolete probabilities for an updated amount of 200 trips from 0 to 1.

segments. Nevertheless, this distribution of passengers along the network does not match the flow volume observations made at two segments. Through segment (2,2,3), 105 passengers are traveling, and through (3,3,1), there are 18, as shown in Figure 4. These values do not correspond to the volumes obtained with the obsolete path choice probabilities. Thus, these probabilities must also be updated as well as the number of trips of the OD matrix. Notice that the updated path choice probabilities, presented in the fourth column of Table 2, must be close to the reference ones, and the new OD matrix must be similar to the obsolete one to preserve the dynamics of the city [5]. Therefore, the ODA problem aims to update all the OD matrix trips simultaneously and the new

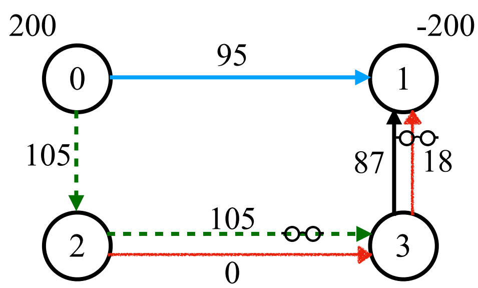


Figure 4: Updated transit segment flows for an updated amount of 200 trips from 0 to 1.

path choice probabilities by matching the flow volumes observations made at some transit segments of the network. To the best of our knowledge, we are the first to update the OD matrix and the route choice probabilities in an integrated way. In other approaches, where continuous variables are used, the obtained solutions are not integer numbers and it is necessary to make roundings to implement them in practice. Here, we formulate a mixed-integer linear program (MILP) to solve the ODA problem, which is the crucial point of our methodology since we avoid a quadratic model. Indeed, most of the models for updating OD matrices rely on a quadratic objective function that corresponds to the relative difference between the reference OD matrix and the estimated one. By having a linear model, we can solve it more efficiently, and it allows us to introduce more details in the network as the flow volume observations. Another contribution is that we consider that the number of trips and the network flow volumes are integer numbers contrary to most other approaches. We also introduce a family of valid inequalities and establish bounds on the variables to provide a tighter integer linear programming formulation. To test our methodology, we propose a random instance generator whose instances are close to real public transportation networks. We also show the performance of our methodology on a real-world case study. This paper is organized as follows. We first present the literature review in Section 2. The ODA problem is formally defined in Section 3. Then, its MILP model is presented in Section 4. Experimental results on generated networks and in a real-world case study validate our methodology in Section 5. We include a comparison of our method with a penalty-based quadratic model from the literature. Finally, in Section 6, we present the conclusions.

2 Literature review

The OD matrix estimation approaches usually combine two stages of the four-stage sequential procedure. For instance, Fisk and Boyce [10] propose a model that combines trip distribution and traffic assignment, while Fisk [9] combines the entropy maximization method with traffic assignment. Also, Yang et al. [25] extended these results to congested networks where the link choice probabilities are not constant. Some literature models update OD matrices in public transportation, but the approaches adapt to assignment procedures of vehicle traffic flows [5].

The two-stage problems have also been formulated as bi-level optimization problems where the upper level represents the OD estimation process, and the lower level represents a network equilibrium assignment [25, 11]. In Shihsien and Fricker [20], the authors propose a two-stage iterative method to estimate an OD matrix and the variation in link choices among trip makers, but inconsistencies arise when congestion effects are considered. In Yang et al. [26], the authors improved this approach by using, in the cost function, the link flows obtained from the stochastic user equilibrium traffic assignment and estimated OD flows. They use as the objective function a sum of the squares of errors and propose a successive quadratic algorithm to solve the model. Most of the shortcomings of the above studies are related to the route choice probabilities due to their computation from a separate traffic assignment model, especially in a network with congestion.

Based on the user equilibrium principle, some models succeeded in incorporating congestion effects into the estimation process, but the perception of travel costs does not vary among travelers. A more realistic approach can be considered allowing for the difference in cost perceptions and different link choice behavior among travelers using a stochastic user equilibrium assignment as in Lo and Chan [15].

Most of the mentioned approaches formulate the problem as a quadratic optimization problem in which they include observed data, such as the flow of people at some segments of the transit network, and an reference matrix obtained from surveys or projections based on the economic growth. There are relatively few authors who propose a linear model. For example, Ashok and Ben-Akiva [1] formulate a model to estimate a dynamic OD matrix. They define a state vector in terms of the departure rates from each origin to each destination. Pitombeira-Neto et al. [17] propose a linear model to estimate a dynamic OD matrix to represent the stochastic evolution of OD flows over time. Link choice probabilities are obtained through a utility model based on past link costs. They propose a Markov chain Monte Carlo algorithm to approximate the mean OD flows and the link choice model parameters.

In Chávez-Hernández et al. [7], the authors consider a penalized quadratic model to update OD matrices from observed transit flow volumes. They present an augmented Lagrangian model and its iterative solution by a dual ascent technique and the method of multipliers. In Section 5, we compare our methodology with theirs.

In this study, we introduce a model to simultaneously estimate the OD matrix and the variation of the path probabilities, representing the effects of congestion. In addition to being one of the few linear models in the literature, it has the advantage that it can be extended to model changes in the perception of the travel cost in each transit segment for each OD pair.

3 The ODA problem

In this section, we formally present the ODA problem. We are considering updating a reference OD matrix at a specific period of the day. Our methodology is based on an optimization network flow model that avoids the most often used quadratic models for this problem (e.g., Chávez-Hernández et al. [7]). Instead, we count the excess or deficit of trips at each OD pair, as shown in Section 4.

Let us consider a multimodal public transit network with a set of lines $\mathcal{L}$. The public transit system is represented by a directed multigraph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of nodes (bus or subway stops) and $\mathcal{A}$ is the multiset of transit segments (directed links) of the lines in $\mathcal{L}$. Segment or link $a \in \mathcal{A}$ is a triplet (l, i, j) indicating the line $l \in \mathcal{L}$ and the nodes i and j linked by line l, with both nodes in $\mathcal{N}$. Notice that in link (l, i, j), the line l passes first through i and then through j. We consider that all the nodes (or centroids) in $\mathcal{N}$ are an origin and a destination, thus $\mathcal{PQ} = \{(p, q) \in \mathcal{N} \times \mathcal{N} \text{ and } p \neq q\}$.

The reference OD matrix, denoted by $\widehat{\mathbf{g}} = \{\widehat{g}_{pq}\}$, corresponds to the obsolete number of trips generated in the transportation network at node p, whose final destination is node q for all $(p, q) \in \mathcal{PQ}$. The objective of the ODA problem is to determine the estimated OD matrix denoted by $\mathbf{g} = \{g_{pq}\}$, which is close to matrix $\widehat{\mathbf{g}}$ and verifies measured observation of the flow volumes at some transit segments of the network, for $(p, q) \in \mathcal{PQ}$. While the updated OD matrix $\mathbf{g} = \{g_{pq}\}$ corresponds to the variables in our methodology, the reference matrix $\widehat{\mathbf{g}} = \{\widehat{g}_{pq}\}$ values are data known a priori, for $(p, q) \in \mathcal{PQ}$.

Based on the concept of optimal strategy introduced by Spiess and Florian [22], let $S_{pq} \subseteq \mathcal{A}$ be the subset of transit segments that a traveler may take to go from p to q, with $(p, q) \in \mathcal{PQ}$. A decision choice model determines these strategies [5] together with its path choice probabilities. Indeed, an average user would not take a path between two nodes that takes twice as long. Although, this user may consider a path that takes two more minutes but without transfers. Thus, for one pair of nodes, there could be several paths that the user may take. After solving an assignment problem based on the decision choice model, each transit segment $a = (l, i, j) \in S_{pq}$ has probability π^a_{pq} of being used by a traveler going from p to q, $(p, q) \in \mathcal{PQ}$. For a segment $a' \notin S_{pq}$, this probability is $\pi^{a'}_{pq} = 0$. The obsolete path choice probabilities are used as a reference and considered a parameter in this study.

To update the OD matrix, we rely on observed flow volumes of travelers at some transit segments $\bar{v}^a$ in set $\bar{\mathcal{A}} \subset \mathcal{A}$. We use equivalent notations v^a_{pq} or $v^{(l,i,j)}_{pq}$ depending on the detail we need to express the equations, for $a = (l, i, j) \in S_{pq}$.

The ODA problem can now be formally stated: find the OD matrix $\mathbf{g}$ value that minimizes the difference between this matrix and the reference OD matrix $\widehat{\mathbf{g}}$ such that the flow volumes in the observed links $\bar{v}^a$ are verified, with $a \in \bar{\mathcal{A}}$.

4 Mixed-integer linear programming model for the ODA problem

To formulate a mixed-integer linear programming model for the ODA problem, we must determine the OD matrix variable $\mathbf{g} = \{g_{pq}\}$ with $(p, q) \in \mathcal{PQ}$. These variables are the estimated values of the OD matrix: the estimated trips from the origins to the destinations.

In this work, we use as an objective function the absolute distance between $\mathbf{g}$ and the reference OD matrix $\widehat{\mathbf{g}}$ to allow the new demand to change to reproduce the observed flow volumes at certain segments. We use two sets of variables to control the differences between the reference OD matrix and the estimated one. The excess integer variables E_{pq} with $(p, q) \in \mathcal{PQ}$ indicate that there are more trips from p to q. Thus, $\widehat{g}_{pq} < g_{pq}$ and this manner this excess is defined as $E_{pq} = \max\{g_{pq} - \widehat{g}_{pq}, 0\}$. Similarly, we introduce deficit variables D_{pq} with $(p, q) \in \mathcal{PQ}$ for the case where there are fewer trips from p to q, that is, $\widehat{g}_{pq} \geq g_{pq}$. Therefore, $D_{pq} = \max\{\widehat{g}_{pq} - g_{pq}, 0\}$. Note that when $D_{pq} > 0$ then $E_{pq} = 0$, and vice versa.

The objective function of the MILP for the ODA problem is to obtain an estimated OD matrix $\mathbf{g}$ as close as possible to the reference one $\hat{\mathbf{g}}$:

$$\min \sum_{(p,q)\in\mathcal{PQ}} \alpha D_{pq} + \beta E_{pq}. \tag{1}$$

This objective minimizes the total sum of the excess, and the deficits of the estimated OD matrix $\mathbf{g}$. Notice that linear parameters α and β allow us to give more preference to the excess or the deficits. For example, for a city that has a growing population over the years, we may expect that there will be more trips in many of its OD matrix entries, thus $\beta < \alpha$. Similarly, a rural zone may be experiencing a population decrease that should be reflected in many of the trips between origin-destination pairs ($\beta > \alpha$).

To linearly express the deficits and the excess of the estimated OD matrix, we need the following equations for each $(p, q) \in \mathcal{PQ}$:

$$D_{pq} \geq \widehat{g}_{pq} - g_{pq}, \tag{2}$$

$$E_{pq} \geq g_{pq} - \widehat{g}_{pq}. \tag{3}$$

The ODA problem updates the OD matrix and determines the volume of people traveling throughout each link $a \in S_{pq} \subset A$ that are going from p to q, $(p, q) \in \mathcal{PQ}$. Hence, we introduce integer variables v^a_{pq} to indicate the actual number of people going from p to q using segment $a = (l, i, j) \in S_{pq}$.

As mentioned before, there are some transit segments $a \in \bar{\mathcal{A}}$ where the number of flow passengers $\bar{v}^a$ is observed and counted. These observations are our most important tool to update the OD matrix. We do not know the origin nor the destination of the passengers using this segment. Thus, we have that the sum of all volumes should be equal to the observations:

$$\bar{v}^a = \sum_{(p,q)\in\mathcal{PQ}|a\in S_{pq}} v^a_{pq}, \quad \text{for } a \in \bar{\mathcal{A}}. \tag{4}$$

The usual way of modeling the passenger volumes from p to q at a link $a \in \mathcal{A}$ is to multiply the total number of trips g_{pq} by the assignment probability π^a_{pq}. Thus, $\pi^a_{pq} g_{pq} = v^a_{pq}$. Nevertheless, by using this equation, the actual passenger volumes, in some cases, might yield infeasible solutions. That implies that the assignment probabilities are not verified and may have some small deviation. Indeed, the assignment problem may establish a probability of 0.6, but in reality, it may be 0.59. Thus, we consider that these probabilities need small adjustments to reflect the actual volumes. Therefore, we propose to compute them as follows for $a \in S_{pq}$ and every $(p, q) \in \mathcal{PQ}$:

$$\lfloor \max\{(\pi^a_{pq} - \varepsilon), 0\}\ g_{pq} \rfloor \leq v^a_{pq}, \tag{5}$$

$$\lceil \min\{(\pi^a_{pq} + \varepsilon), 1\}\ g_{pq} \rceil \geq v^a_{pq}, \tag{6}$$

with $\varepsilon \geq 0$. Interval $[\max\{(\pi^a_{pq} - \varepsilon), 0\}, \min\{(\pi^a_{pq} + \varepsilon), 1\}]$ represents the allowed change in the obsolete route choice probabilities. Notice that we are not enforcing equality since we have the floor operator and positive values of ε. After solving our ODA MILP model, we obtain the updated OD matrix and the updated assignment probabilities that fit the observations of the flows in the network.

Then, we must handle the network flow constraints. The sum of the flow volumes at origin $p \in \mathcal{PQ}$ must be equal to the number of trips originated at this node, as stated by constraints (7). Similarly, with constraints (8) all flow volumes arriving at destination $q \in \mathcal{PQ}$ is equal to the total trips ending there. Flow conservation at every node is guaranteed by constraints (9): the flow entering node $k \in \mathcal{N} \setminus \{p, q\}$ must be equal to the flow leaving it.

$$\sum_{l \in \mathcal{L}} \sum_{\{i | (l,p,i) \in S_{pq}\}} v^{(l,p,i)}_{pq} = g_{pq}, \qquad (p, q) \in \mathcal{PQ}, \tag{7}$$

$$\sum_{l \in \mathcal{L}} \sum_{\{i | (l,i,q) \in S_{pq}\}} v^{(l,i,q)}_{pq} = g_{pq}, \qquad (p, q) \in \mathcal{PQ}, \tag{8}$$

$$\sum_{l \in \mathcal{L}} \sum_{\{i | (l,i,k) \in S_{pq}\}} v^{(l,i,k)}_{pq} = \sum_{l \in \mathcal{L}} \sum_{\{j | (l,k,j) \in S_{pq}\}} v^{(l,k,j)}_{pq}, \qquad k \in \mathcal{N} \setminus \{p, q\}, (p, q) \in \mathcal{PQ}. \tag{9}$$

Valid inequalities strengthen a MILP formulation since they do not cut any feasible integer solution and make the solution space polyhedron closer to the integer solutions convex hull [24, 18]. Thus, we introduce valid inequalities (10) to our MILP to decrease the computational running time without compromising the optimality of the solution since it bounds the volume of each arc by the total number of persons going from p to q:

$$v^a_{pq} \leq g^{pq}, \quad a \in S_{pq}, (p, q) \in \mathcal{PQ}. \tag{10}$$

Inequalities (10) are valid by definition. Notice that by imposing a positive integrality to the volumes, we also ensure the integrality on the estimated OD values of the matrix and the excess and deficits. Thus, these last variables may be defined as real variables but will take integer values as stated by (11)-(13).

$$v^a_{pq} \in \mathbb{Z}^+, \qquad (p, q) \in \mathcal{PQ}, a \in \mathcal{A}, \tag{11}$$

$$\delta_1 \hat{g}_{pq} \leq g_{pq} \leq \delta_2 \hat{g}_{pq} \in \mathbb{R}^+, \qquad (p, q) \in \mathcal{PQ}. \tag{12}$$

$$D_{pq}, E_{pq} \in \mathbb{R}^+, \qquad (p, q) \in \mathcal{PQ}, \tag{13}$$

where δ_1 and δ_2 are constants known by the user to bound $\mathbf{g}$ and remain close to $\widehat{\mathbf{g}}$. For example, a census or some statistical information may estimate that a particular population has grown no more than 10%.

To summarize, we denote as the ODA-MILP(ε) the MILP model of the ODA problem parametrized with a value of ε such that it minimizes objective function (1) subject to constraints (2)-(13). Our methodology consists of starting with $\varepsilon = 0$ and then increasing it by 0.02 units until a feasible solution for the ODA-MILP(ε) is reached. In this manner, we obtain an estimated OD matrix and the actual flow volumes, or equivalently, the actual assignment probabilities of the links.

5 Experimental results

The general scheme of the comparison process we use in this study to validate the ODA-MILP(ε) model is depicted in Figure 5. We start with the real matrix $\bar{\mathbf{g}}$; in Section 5.1 we explain how to generate it. This matrix is usually unknown, but we consider that we are in an ideal case where we know it to validate our approach. Then, we perturb the real matrix to obtain the reference or obsolete matrix $\hat{\mathbf{g}}$. Finally, by using the ODA-MILP(ε) we obtain $\mathbf{g}$ which estimates the real matrix. Two questions must be validated. First, we must assess how close the reference OD matrix $\hat{\mathbf{g}}$ is to the estimated one $\mathbf{g}$. That would verify the mathematical model correctness and ensure the previous knowledge on the population dynamics. Second, we must asses how close the real OD matrix $\bar{\mathbf{g}}$ is to the estimated one $\mathbf{g}$. This is the most challenging question, a fair comparison between $\bar{\mathbf{g}}$ and $\mathbf{g}$ can be drawn with this methodology, validating that the updated OD matrix is a reasonable estimate of the real population trips.

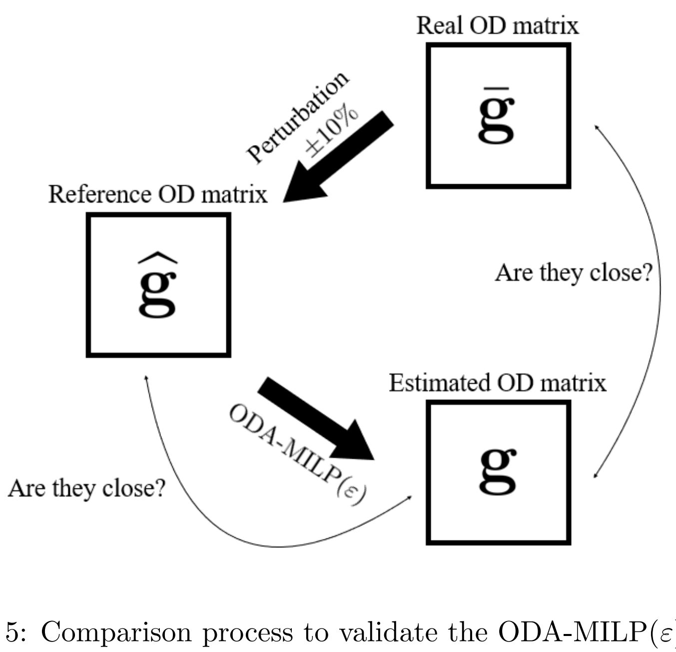


Figure 5: Comparison process to validate the ODA-MILP(ε) model.

Section 5.1 describes the randomly generated matrices $\bar{\mathbf{g}}$ and how we perturb them to obtain the reference ones. In Section 5.2, we compare the ODA-MILP(ε) with the augmented Lagrangian method introduced in Chávez-Hernández et al. [7]. Finally in Section 5.3, we present our case study based on the city of Monterrey, Mexico.

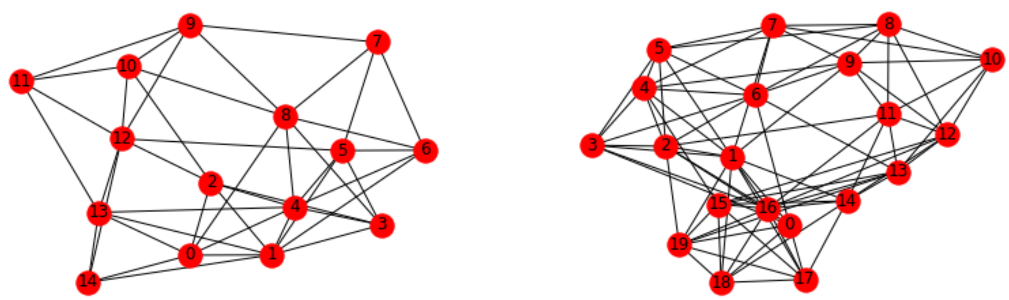
For the ODA-MILP(ε), the excess and deficit parameters of the objective function (1) are set to $\alpha = 1$ and $\beta = 1$ for all instances. Thus, no previous knowledge about the dynamics of the centroids is known in advance. The parameters in equation (12) that bound the estimated OD matrix values are set to $\delta_1 = 0.9$ and $\delta_2 = 1.1$.

The ODA-MILP(ε) was coded in Python 3.7 and solved with a branch and bound implementation by Gurobi 8.1 with the default algorithmic parameters. All experiments were executed in a computer with a processor Intel(R) Core(TM) i7 and 12 Gb of RAM.

5.1 Randomly generated instances

The generation of public transportation instances that mimic the real networks is an active research area. Public transportation networks have special properties, such as: grow in an evolutionary way, are embedded into two-dimensional space, have small-world properties, and have hierarchical organization [23, 6, 21].

Based on these studies, we generate a set of random instances that contain the matrices corresponding to the real OD matrices $\bar{\mathbf{g}}$ and the reference one $\hat{\mathbf{g}}$ and the public transit networks with the detail of the lines and the passenger volumes per link that we need for our experimental tests. All our instances and results can be found online[1].



(a) 15 nodes.

(b) 20 nodes.

Figure 6: Newman-Watts-Strogatz small-world graphs with 15 and 20 nodes.

Each instance representing a public transit system is composed by the real OD matrix $\bar{\mathbf{g}}$, the reference OD matrix $\hat{\mathbf{g}}$, its associated directed multigraph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{A}$ is the multiset of directed links between the lines in $\mathcal{L}$, and the route choice probabilities per segment in $\mathcal{A}$. The following methodology to generate the set of instances is used in this study.

1. The exact OD matrices $\bar{\mathbf{g}}$ are random integers between $[0, 500|\mathcal{N}|]$ for each pair $(p, q) \in \mathcal{PQ}$. The diagonal entries are all zeros.

2. A Newman-Watts-Strogatz small-world graph [16] is generated (with Python library Network [12]) by first creating a ring over $|\mathcal{N}|$ nodes. Each node in the ring is connected with its

[1] https://doi.org/10.6084/m9.figshare.13838819

$k = \lceil 0.3|\mathcal{N}| \rceil$ nearest neighbors (or $k-1$ neighbors if k is odd). The resulting Newman-Watts-Strogatz small-world graph has undirected edges, as shown in Figure 6 where we show two graphs with 15 and 20 nodes, respectively.

3. Now that we have a graph that resembles a public transportation network, we form the transit lines (corresponding to bus, underground, or any other transit mode) $|\mathcal{L}|$. For each OD pair of nodes $(p,q) \in \mathcal{PQ}$, we compute all the non-intersecting paths between them and select the $|\mathcal{L}|$ ones with the shortest number of segments. If there are fewer than $|\mathcal{L}|$ non-intersecting paths, we choose them all. In this manner, each path is associated with a line l and an edge (i,j) belonging to the non-intersecting shortest paths between (p,q). Notice that the lines may visit all the nodes.

4. We establish the same frequency for all the lines. The link choice probabilities π^a_{pq} for each $a \in S_{pq}$ are then evenly computed along the (p,q) OD pair paths.

5. To generate the reference matrices $\hat{\mathbf{g}}$, 15% of the OD pairs of the exact OD matrix $\bar{\mathbf{g}}$ are randomly selected and uniformly perturbed by $\pm 10\%$. The OD pairs that are not selected have the same value in the real matrix $\bar{\mathbf{g}}$ and the reference one $\hat{\mathbf{g}}$. These instances are named *Instances-ED*.

6. Using the link probabilities, we compute the segment flows v^a_{pq}, $a \in \mathcal{A}, (p,q) \in \mathcal{PQ}$, for each matrix $\bar{\mathbf{g}}$ and $\hat{\mathbf{g}}$. All the segments flow volumes in the set *Instances-ED* have been observed. Set *Instances-ED*$^{1/2}$ is composed of the same instances, but this time only half of the transit segments are observed.

7. Another set of instances is generated to test that the route choice probabilities are indeed modified. This time, the real matrix and the reference one are equal, so the reference demand entries are not perturbed. Nevertheless, 15% of the flow volumes at the network segments are perturbed by $\pm 10\%$, but all of them are still observed. The resulting instances, named *Instances-*ε, aim to show that the assignment probabilities may differ from the initial ones and must be modified together with the demand OD matrix.

In this manner, we have generated 201 instances with the number of nodes in the transit network between [4,20] and transit lines between [1,5]: 67 instances in the *Instances-ED* set, 67 instances in the *Instances-ED*$^{1/2}$ set, and 67 instances in the *Instances-*ε one.

5.2 Experimental results for the ODA problem

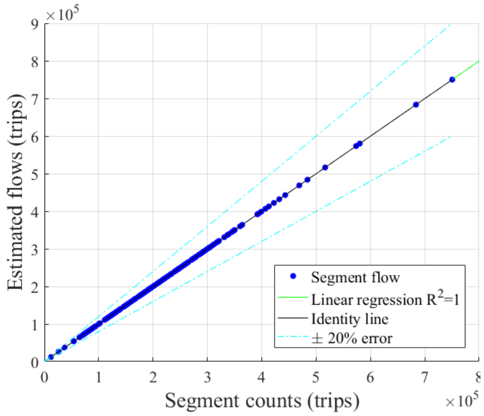
We compare the ODA-MILP(ε) performance with the augmented Lagrangian methodology of Chávez-Hernández et al. [7], which is based on an iterative dual ascent technique and the Lagrangian multipliers method. Their approach yields high-quality solutions with low CPU time.

In Table 3, we show the comparison results for the instances *Instances-ED*. The ODA-MILP(ε) is parametrized with $\varepsilon = 0$, which is sufficient for these instances to find a feasible and optimal solution (later, this parameter will be forced to change for the set *Instances-*ε). The first and second columns correspond to the number of nodes $|\mathcal{N}|$ and the number of lines $|\mathcal{L}|$ in the transit system. The third to seventh columns correspond to the results of the ODA-MILP(ε). The rest of the columns are for the augmented Lagrangian method. For both methods, rmse($\bar{\mathbf{g}}, \mathbf{g}$) is the root mean squared error (rsme) between the reference matrix $\bar{\mathbf{g}}$ and the estimated one $\mathbf{g}$, while rmse($\hat{\mathbf{g}}, \mathbf{g}$) is the root mean squared error between the exact demand $\hat{\mathbf{g}}$ and the estimated one $\mathbf{g}$. Columns labeled as rmse($\bar{v}, v$) correspond to the root mean squared error between the observed and the estimated segment flow volumes. Finally, columns "time" are the CPU time in seconds to solve the instance with each methodology. Each line of this table is an average. For example, the first line represents the average of the instances with 4 to 9 nodes but with a single line.

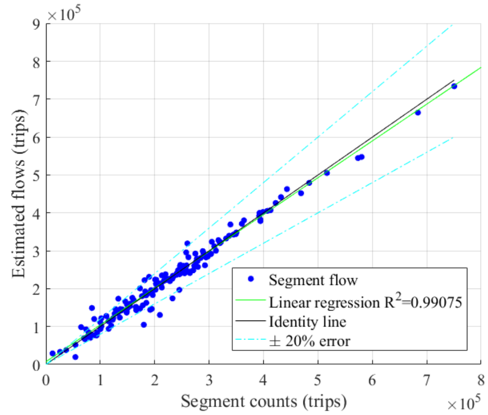
		ODA-MILP(0)				augmented Lagrangian			
$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{L}\rvert$	rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time	rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time
4-9	1	182.36	134.66	0.00	0.01	3327.32	3325.85	2782.14	0.01
	2	418.32	318.34	0.00	0.01	7053.29	7042.63	1073.07	0.00
	3	408.50	303.20	0.00	0.01	11904.13	11895.27	2789.14	0.05
10-15	1	972.13	706.25	0.00	0.02	23758.67	23754.84	9960.47	0.11
	2	1731.19	1223.90	0.00	0.02	38827.60	38769.26	10119.73	0.37
	3	2437.99	1404.72	0.00	0.03	55952.64	56148.93	8939.14	0.75
	4	3674.35	2510.42	0.00	0.04	104932.99	104909.87	7688.41	2.34
16-20	1	1399.07	746.53	0.00	0.03	33624.09	33592.08	19351.70	1.73
	2	2697.02	1919.76	0.00	0.05	74077.07	74214.09	11373.41	4.18
	3	3939.14	2303.91	0.00	0.06	109488.89	109556.34	14999.98	4.35
	4	5203.55	3212.61	0.00	0.08	129385.39	129667.00	18705.03	5.14
	5	6816.88	5045.99	0.00	0.08	181655.45	181777.15	15671.14	9.67
	Av.	2490.04	1652.52	0.00	0.04	64498.96	64554.44	10287.78	2.39

Table 3: Comparison between the ODA-MILP(ε) methodology with $\varepsilon = 0$ and the augmented Lagrangian algorithm of [7] for the *Instances-ED* set.

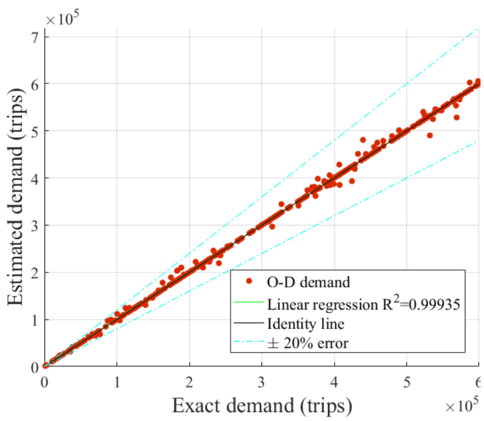
As we can observe form the table, the best results are for the ODA-MILP(ε) method. For the *Instances-ED* set with $\varepsilon = 0$, we obtain high-quality OD matrix approximations. Contrary to the augmented Lagrangian, the estimated matrices obtained with ODA-MILP(ε) method are closer to the real ones than the reference ones are. The difference of the flow volumes is equal to zero for the ODA-MILP(ε) method since the model tries to reproduce this behavior with equations (4). Notice that the augmented Lagrangian method does not remain that close to the observed flow volumes. The larger the instances, the larger the root square mean errors for both methods. Remarkably, the execution time for the ODA-MILP(ε) method is better than the augmented Lagrangian algorithm. Although the resolution time for the ODA-MILP(ε) is less than one minute, the construction of the model is very long. Indeed, the size of the number of variables is $\mathcal{O}(|\mathcal{N}|^3 + 3|\mathcal{N}|^2)$ while the size of the number of restriction is $\mathcal{O}(3|\mathcal{N}|^3 + |L||\mathcal{N}|^2 + 5|\mathcal{N}|^2)$. A research line is then about the data structures, preprocessing algorithms, and dominant solution properties to increase the size of the instances.



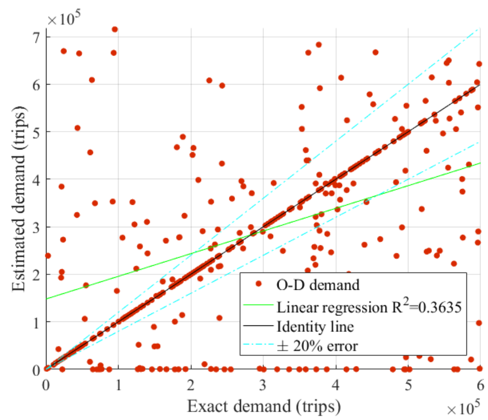
(a) ODA-MILP(ε).



(b) Augmented Lagrangian.



(c) ODA-MILP(ε).



(d) Augmented Lagrangian.

Figure 7: Scatter plots of the volumes, (a) and (b), and of the OD matrices, (c) and (d), for an instance with 20 nodes and 5 lines of the *Instances-ED* set.

Figures 7a and 7b show the scatter plots of the flow volumes at the links, and Figures 7c and 7d shows the scatter plots for the demand of the estimated OD matrix with respect to the real one for an instance with 20 nodes and 5 lines of the *Instances-ED* set. Figures 7a and 7c show the results for the ODA-MILP(ε) while Figures 7b and 7d show the results for the augmented Lagrangian approach. As can be seen, the ODA-MILP(ε) recovers almost entirely both the real OD matrices and the observed segment flow volumes. This is not the case for the augmented Lagrangian method, where we can see a relatively small scattering in the transit flow volumes but a large one in the demand OD matrices.

Most of the time, due to faulty measuring instruments, not all the flows of every transit segment can be observed. Thus, we compare the ODA-MILP(ε) performance when only half of the transit network links have been observed. These results of the *Instances-ED*$^{1/2}$ set are displayed in Table 4, which has the same structure as Table 3.

Table 4 shows that the ODA-MILP(ε) cannot exactly reproduce the flow volumes observations since the rmse($\bar{v}, v$) are no longer zero as for the *Instances-ED* set. The differences of the estimated OD matrix with the real or with the reference ones are larger, which is normal behavior since we

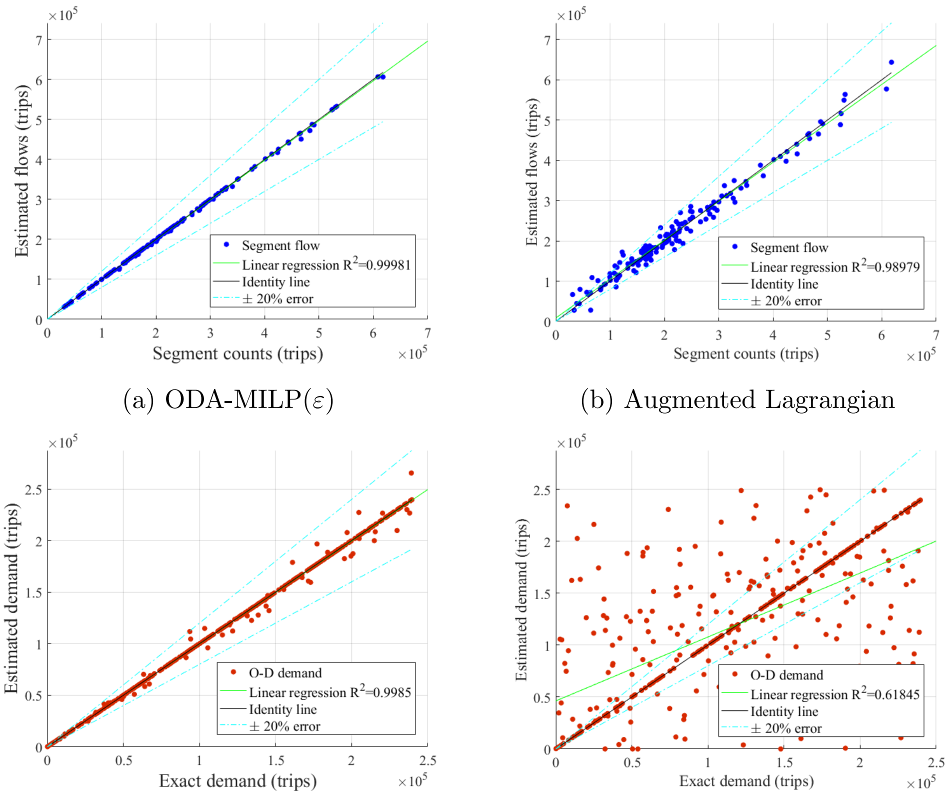
$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{L}\rvert$	ODA-MILP(0)				augmented Lagrangian			
		rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time	rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time
4-9	1	518.72	546.77	417.96	0.05	3626.40	3626.40	498.12	0.00
	2	404.27	483.43	252.36	0.15	6543.11	6543.11	1054.17	0.00
	3	336.24	464.14	191.27	0.15	11968.09	11968.09	1285.14	0.00
10-15	1	3086.89	3393.05	1941.77	0.84	17620.94	17620.94	1082.77	0.01
	2	2131.2	2533.54	1411.99	2.57	32150.13	32150.13	11373.13	0.00
	3	3318.08	3552.34	1101.28	4.19	56788.61	56788.61	6179.43	0.02
	4	4381.59	4979.35	1287.87	4.72	94491.07	94491.07	7521.56	0.15
16-20	1	3570.02	3771.38	1498.87	3.11	22171.57	22171.57	10761.14	0.02
	2	3225.94	3462.36	1738.42	6.78	70228.30	70228.30	10263.14	0.26
	3	4538.72	4807.14	1631.43	11.67	121422.94	121422.94	8709.94	0.63
	4	5421.6	6399.04	1688.64	13.9	133037.24	133037.24	17501.44	0.54
	5	6959.75	8255.14	1713.79	20.94	185924.72	185924.72	17467.71	1.27
	Av.	3157.75	3553.97	1239.64	5.76	62997.76	62997.76	7808.14	0.24

Table 4: Comparison between the ODA-MILP(ε) methodology with $\varepsilon = 0$ and the augmented Lagrangian algorithm of [7] for the *Instances-ED*$^{1/2}$ set.

have less information provided by the network. Moreover, there could be dependencies between the non-observed segments due to their geolocation. Although the computational time is still short, it takes a little longer than the time employed when considering all the observed segments. In the case of the augmented Lagrangian, we can see that the rmse($\bar{v}, v$), the rmse($\bar{\mathbf{g}}, \mathbf{g}$), and the computational time decrease with respect to the values obtained with the *Instances-ED* set.

The difference observed between the two methodologies may be explained by the assumptions made by each model. The dual ascent and Lagrange multipliers methods are accurate and efficient especially in large-scale problems with continuous variables. In small networks with integer values, these assumptions are no longer a sufficient not only for the computational efficiency but also for the result's precision. The feasible solution regions (convex hulls) of the discrete and continuous cases for the same instance have a discrepancy, thus the discrete optimum is underestimated or overestimated by the computational solution. Furthermore, in small instances the augmented Lagrangian algorithm is not as efficient as for larger instances since it is an iterative method and the desired precision cannot be reached.

Figure 8 is similar to Figure 7, but for one instance of the *Instances-ED*$^{1/2}$ set with 20 nodes and 2 lines. The scatter plots for the segment flows are depicted in Figures 8a and 8b, while Figures 8c and 8d show the scatter plots of the estimated demand. Figures 8a and 8c show the results obtained by the ODA-MILP(ε) method, while Figures 8b and 8d show the results obtained by the augmented Lagrangian. Although we do not obtain a perfect fit between the observed volumes and those calculated with the ODA-MILP(ε), the differences are small. The adjustment in both



(a) ODA-MILP(ε) (b) Augmented Lagrangian

(c) ODA-MILP(ε) (d) Augmented Lagrangian

Figure 8: Scatter plots of the volumes, (a) and (b), and of the OD matrices, (c) and (d), for an instance with 20 nodes and 2 lines of the *Instances-ED*$^{1/2}$ set.

volumes and demand of the ODA-MILP is not perfect but the dispersion between the reference values and the estimates is smaller than the dispersion obtained with the augmented Lagrangian. Furthermore, for the augmented Lagrangian we can see that although the estimated volumes remain relatively close to the observed ones, in general the estimated demand is far from the exact solution.

Frequently, the users may have small changes in how they choose their routes because of some network modifications. An example is presented in Table 2 for the network example of Figure 1. In our model, this phenomenon is modeled by constraints (5) and (6), where the new probability of traveling on each transit segment $a \in \mathcal{A}$ may change a little with respect to the probability π_{pq}^a with which the user was previously traveling from p to q, $(p, q) \in \mathcal{PQ}$.

Our previous experimental results yield an $\varepsilon = 0$, which means that the behavior of the user have not changed. That is, the path choice probabilities are equal to the obsolete ones. We now test the *Instances-*ε set where we manually change the values of ε until the problem is feasible.

These results are presented in Table 5. As for the previous tables, the first and second columns correspond to the number of nodes $|\mathcal{N}|$ and the number of lines $|\mathcal{L}|$ in the transit system. The third column is the value of the ε parameter needed to obtain a feasible estimated OD matrix. In the ODA-MILP(ε) method, we start with $\varepsilon = 0$ and then iteratively increase it by 0.02 until we obtain a feasible solution. The last column shows the time in seconds needed by the ODA-MILP(ε) method.

The values shown at each line of this table represent averages. Notice that we do not report the rmse values since our instances were constructed to force the assignment probabilities to change, such that the real and the reference matrix are the same. Moreover, the augmented Lagrangian method cannot deal with these instances since it does not modifies the assignment probabilities. Indeed, it does not converge to any solution. For example, suppose that an assignment probability at segment $a \in \mathcal{A}$ from OD pair (p, q) is $\pi_{pq}^{a} = 0.5$. If the ODA-MILP(ε) yields a value of $\varepsilon = 0.05$, then the updated probability is now in the interval [0.475,0.525], and it can be computed once we have the estimated OD matrix and its flow volumes.

		ODA-MILP(ε)	
$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{L}\rvert$	ε	time
4-9	1	0.13	0.00
	2	0.04	0.00
	3	0.09	0.00
10-15	1	0.05	0.01
	2	0.04	0.01
	3	0.04	0.02
	4	0.03	0.03
16-20	1	0.08	0.01
	2	0.03	0.03
	3	0.02	0.10
	4	0.03	0.06
	5	0.08	0.15
	Av.	0.05	0.04

Table 5: Values of ε to obtain a feasible solution and time in seconds for the ODA-MILP(ε) methodology for the *Instance-ε* set.

Table 5 shows that the ODA-MILP(ε) methodology can adjust the ε parameter to consider that path choice decisions are made differently than before. With this consideration, we can obtain an OD matrix that coincides with the real one. Moreover, the computational time does not increase, and the assignment probability variation is not radical. That is, the variations on these probabilities are relatively small.

We have validated the ODA-MILP(ε) methodology with random instances with the previous experiments. Next section, we apply it to a more realistic network.

5.3 The Monterrey transit network

Let us consider a network that represents the districts of Monterrey City, Mexico, and its surroundings consisting of 17 aggregated zones.

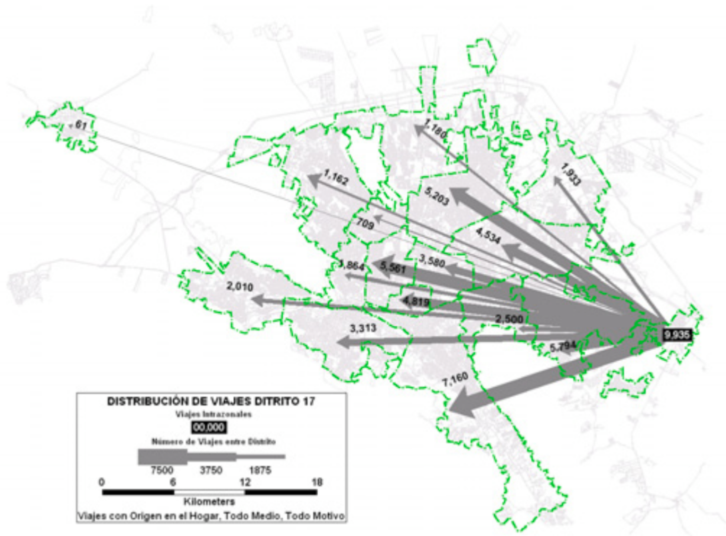


Figure 9: Trips originated at zone 17 in Monterrey to any other zone and with any purpose in 2008 (figure from [8]).

Figure 9 shows the trips from zone 17 in Monterrey to any other zone and with any purpose. The trips between all the zones in 2008 were taken from [8] and are given in the OD matrix of Table 6, which we consider as the exact demand $\bar{\mathbf{g}}$. Since we do not consider the trips within each zone, then the diagonal of the OD matrix has only zeros. The reference matrix $\hat{\mathbf{g}}$ is obtained by perturbing all the real matrix entries by a random percentage between [-10,10]%.

For this case study, since we are not introducing information about the transit network, we consider a bi-directed link (which may represent a hyper path) between each pair of zones corresponding to the network nodes. Thus, we have 272 OD pairs. Here, we made the assumption that each OD pair is connected by only one link with path probability equals to 1. So the number of trips assigned to each link can be computed as $\hat{v}_{pq}^{pq} = \bar{g}_{pq}$. Those flow volumes play the role of segment counts.

The ODA-MILP(ε) and the augmented Lagrangian methods were used to solve the instance that represent the network of Monterrey. Table 7 has the same structure as Table 3 and shows the root mean squared errors end the computing time for the ODA-MILP(ε) and the augmented Lagrangian method. Comparing the rmse values obtained for each method, we can see that the ODA-MILP(ε) estimated OD matrix remains closer to the reference one and the exact solution than the OD matrix obtained with the augmented Lagrangian; also the segment flows are estimated perfectly in a smaller computing time that the required for the augmented Lagrangian to obtain a bigger rmse($\bar{v}, v$) value. We can validate the usefulness of the ODA-MILP(ε) and its efficiency with respect to the augmented Lagrangian method. Interestingly, our methodology can recover the real matrix even if the difference with the reference one is different from zero. With more information about the growth dynamics of the city of Monterrey and some observations in the most critical transit segments of the network, we could apply our methodology to estimate the actual OD matrix in the real transportation network.

zone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	4708	2031	9653	3594	15159	8747	1808	1159	11894	1276	972	623	8872	980	58	100
2	20629	0	46083	39046	38685	51363	15097	11501	10460	40323	12123	10503	19804	11493	24976	335	987
3	18865	64948	0	37117	44471	42463	12354	7564	5669	44098	22434	25229	25721	10392	9332	270	1669
4	973	1137	520	0	2645	3906	653	211	178	3328	394	196	109	333	152	5	15
5	3420	10944	6217	15664	0	12528	2600	1170	1037	10788	3376	1408	883	1566	1039	30	103
6	7161	11795	3099	21157	9087	0	9360	2748	3631	9782	1622	1057	763	2780	1541	44	97
7	8530	6501	2021	8817	4075	2471	0	4594	3877	7117	394	795	583	3485	1198	46	80
8	35509	66705	18503	44480	27635	82193	36895	0	25605	42719	66705	18503	6678	20527	20442	816	819
9	9496	35081	5430	15030	9806	35768	13520	11548	0	12060	2343	1796	1691	4711	5028	110	183
10	27526	19382	11439	53515	22845	37734	10818	4217	2963	0	11489	6018	2966	8235	3339	148	520
11	9860	16744	21690	21839	19680	19603	5578	2795	1987	37521	0	19754	4048	4618	2795	102	649
12	19348	32579	39940	32877	26099	34886	10972	6256	4084	48654	39940	0	12005	10585	6489	277	4266
13	5426	15581	18590	8656	7878	10566	3381	2332	1570	11224	4351	6032	0	3246	3228	98	612
14	45592	14454	6559	19866	9890	30798	6559	5732	3270	22381	3650	3211	2264	0	3372	430	372
15	14955	63626	15271	22837	17180	34363	11236	12020	8188	23209	5724	5184	6759	8684	0	330	565
16	3112	3241	1416	2776	1730	4012	1643	1639	668	3490	3241	1416	606	3767	1106	0	99
17	3313	5203	4534	4819	3580	5561	1864	1162	709	7160	5203	4534	1933	2010	1180	61	0

Table 6: Origin destination exact demand for the city of Monterrey and its surroundings in 2008.

		ODA-MILP(0)				augmented Lagrangian			
$\|\mathcal{N}\|$	$\|\mathcal{L}\|$	rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time	rmse($\hat{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{\mathbf{g}}, \mathbf{g}$)	rmse($\bar{v}, v$)	time
17	272	1176.81	0.00	0.00	0.002	9337.38	9571.45	9911.50	0.016

Table 7: Comparison between the ODA-MILP(ε) methodology with $\varepsilon = 0$ and the augmented Lagrangian algorithm of [7] for the OD matrix of Monterrey.

6 Conclusions

We solve the inverse problem of estimating the actual OD matrix based on a reference one and some flow observations at some network links. Indeed, OD matrices are relevant for the bus line design and the generation of useful timetables, for adding new trips or when drivers do not show up, or when there are accidents and the network should be rapidly restored. Moreover, updating OD matrices allows us to test the current system under more demanding scenarios and adapt the future demand infrastructure.

A integer linear programming model was presented to estimate the OD matrix and simultaneously fit the path choice probabilities from a reference OD matrix and observed flow volumes in the transit segments. We compare the performance of the proposed model with the augmented Lagrangian model previously introduced by Chávez-Hernández et al. [7]. The results has shown that the ODA-MILP(ε) offers good quality solutions for small size instances. Compared to the methodologies in the literature the scatter plots of the demand and the segment flows is considerably lower than those obtained with other approaches. Moreover, the execution times are shorter with the ODA-MILP(ε). Also, we programmed one of the few instances generator that mimic transit networks to test the methodology presented in this paper.

Although our model considers only small changes both in the demand matrix and in the probabilities, most authors only consider a change in the demand. Currently, due the pandemic of COVID-19 in most cities we observe a mobility reduction that can be seen as a decrease in the number of trips represented on an OD matrix; also, in order to avoid contracting the disease, the people try to reduce their contact time with others and that modifies their path choices. This phenomena can be model by the ODA-MILP(ε) and more experiments should be carried out in scenarios where both the demand and the probabilities change. Therefore, it is an issue to handle scenarios with more substantial changes. Besides, our approach could be improved by indicating the assignment probability difference for each OD pair and each transit segment.

Our results are for relatively small networks with a solution computational cost of less than 1 minute. Nevertheless, the most time consuming is the lecture of the model before starting the branch-and-bound solver. As we mentioned before, the the model consists in $\mathcal{O}(|\mathcal{N}|^3 + 3|\mathcal{N}|^2)$ variables and $\mathcal{O}(3|\mathcal{N}|^3 + |L||\mathcal{N}|^2 + 5|\mathcal{N}|^2)$ restrictions. A research line is then about the data structures, preprocessing algorithms, and dominant solution properties to increase the size of the instances.

Finally, these results were obtained from instances generated as described in Section 5.1, this generator can be modified in such a way that the path probabilities represent an equilibrium assignment [22] for cases without congestion and consider heuristic models to represent cases with congestion and capacity limits in transport vehicles.

Acknowledgments: The research of the first and third author was supported by the Mexican Council for Science and Technology (CONACyT) under its Frontier Science Program, grant FC2016/1948. The research of the fourth author was supported by CONACyT, grant 710289.

References

[1] K. Ashok and M. E. Ben-Akiva. Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows. *Transportation Science*, 36(2):184–198, 2002.

[2] A. Balakrishnan, T. L. Magnanti, and R. T. Wong. A dual-ascent procedure for large-scale uncapacitated network design. *Operations Research*, 37(5):716–740, 1989.

[3] S. Bera and K. V. Krishna Rao. Estimation of origin-destination matrix from traffic counts: The state of the art. *European Transport / Trasporti Europei*, 49:3–23, 2011.

[4] V. Boyer, O. J. Ibarra-Rojas, and Y. Á. Ríos-Solís. Vehicle and crew scheduling for flexible bus transportation systems. *Transportation Research Part B: Methodological*, 112:216–229, 2018.

[5] E. Cascetta. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. *Transportation Research Part B: Methodological*, 18(4–5):289–299, 1984.

[6] A. Chatterjee, M. Manohar, and G. Ramadurai. Statistical analysis of bus networks in India. *PLoS ONE*, 11(12):e0168478, 2016.

[7] M. V. Chávez-Hernández, L. H. Juárez Valencia, and Y. Á. Ríos Solís. Penalization and augmented Lagrangian for O-D demand matrix estimation from transit segment counts. *Transportmetrica A: Transport Science*, 15(2):915–943, 2019.

[8] Consejo Estatal de Transporte y Vialidad, Nuevo León, México. Plan sectorial de transporte y vialidad 2008-2030, 2009. URL `http://www.cetyv.gob.mx/documentos/pstv.pdf`.

[9] C. S. Fisk. Trip matrix estimation from link traffic counts: The congested network case. *Transportation Research Part B: Methodological*, 23(5):331–336, 1989.

[10] C. S. Fisk and D. E. Boyce. A note on trip matrix estimation from link traffic count data. *Transportation Research Part B: Methodological*, 17(3):245–250, 1983.

[11] M. Florian and Y. Chen. A coordinate descent method for the bi-level O/D matrix adjustment problem. *International Transactions on Operations Research*, 2(2):165–179, 1995.

[12] A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using NetworkX. In G. Varoquaux, T. Vaught, and J. Millman, editors, *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, pages 11—15, Pasadena, August 2008.

[13] O. J. Ibarra-Rojas, R. Giesen, and Y. A. Rios-Solis. An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. *Transportation Research Part B: Methodological*, 70:35–46, 2014.

[14] O. J. Ibarra-Rojas, F. Delgado, R. Giesen, and J. C. Muñoz. Planning, operation, and control of bus transport systems: A literature review. *Transportation Research Part B: Methodological*, 77:38–75, 2015.

[15] H.-P. Lo and C.-P. Chan. Simultaneous estimation of an origin-destination matrix and link choice proportions using traffic counts. *Transportation Research Part A: Policy and Practice*, 37(9):771–788, 2003.

[16] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4):341–346, 1999.

[17] A. R. Pitombeira-Neto, C. F. G. Loureiro, and L. E. Carvalho. Bayesian inference on dynamic linear models of day-to-day origin-destination flows in transportation networks. *Urban Science*, 2(4):1–17, 2018.

[18] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley, New York, 1998.

[19] G. Shen and S. G. Aydin. Origin-destination missing data estimation for freight transportation planning: A gravity model-based regression approach. *Transportation Planning and Technology*, 37(6):505–524, 2014.

[20] L. Shihsien and J. D. Fricker. Estimation of a trip table and the θ parameter in a stochastic network. *Transportation Research Part A: Policy and Practice*, 30(4):287–305, 1996.

[21] J. Sienkiewicz and J. A. Hołyst. Statistical analysis of 22 public transport networks in Poland. *Physical Review E*, 72(4):046127, 2005.

[22] H. Spiess and M. Florian. Optimal strategies: A new assignment model for transit networks. *Transportation Research Part B: Methodological*, 23(2):83–102, 1989.

[23] C. von Ferber, T. Holovatch, Y. Holovatch, and V. Palchykov. Network harness: Metropolis public transport. *Physica A: Statistical Mechanics and its Applications*, 380:585–591, 2007.

[24] L. A. Wolsey. *Integer Programming.* Wiley, New York, 1998.

[25] H. Yang, T. Sasaki, Y. Iida, and Y. Asakura. Estimation of origin-destination matrices from link traffic counts on congested networks. *Transportation Research Part B: Methodological*, 26 (6):417–434, 1992.

[26] H. Yang, Q. Meng, and M. G. H. Bell. Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. *Transportation Science*, 35(2):107–123, 2001.