A bilevel model for public transport demand estimation

Neila Bhouri, Jean Patrick Lebacque, Pablo A Lotito, Victoria M Orlando

23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16-18 September 2020, Paphos, Cyprus

A bilevel model for public transport demand estimation

Neila Bhouri[a], Jean-Patrick Lebacque[a], Pablo A. Lotito[b,c], Victoria M. Orlando[b,c,*]

[a] *Univ. Gustave Eiffel, IFSTTAR/COSYS/GRETTIA, 77454 Marne-la-Vallee, France*
[b] *PLADEMA, Fac. C. Exactas, Universidad Nacional del Centro de la Prov. de Buenos Aires, Pinto 399, Tandil, Argentina*
[c] *CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina*

Abstract

For the case of public transport, we consider the problem of demand estimation. Given an origin-destination matrix representing the public transport demand, the distribution of flow among different lines can be obtained assuming that it corresponds to a certain equilibrium characterized by an optimization problem. The knowledge of that matrix is expensive and sometimes unaffordable in practice. In this work, we explore its estimation through the numerical solution of a bilevel optimization problem.

Keywords: transit assignment; public-transport demand

1. Introduction

Transit assignment models have become an interesting research area because knowing the passenger behavior allows comparing different planning scenarios in terms of network performance, always assuming that the transport demand is known.

Many models for passenger behavior have been proposed. Most of them consider that when a passenger decides to travel between certain O-D pairs and is waiting for a vehicle at a stop, he must decide which transit line should he take to minimize his total expected travel time. Among the first models that considered congestion effects, we can cite (1) that work with the concept of hyperpath composed by ”strategies of attractive lines”, but failed to be realistic in cases of high demand.

De Cea and Fernandez (2) began to consider the congestion effects at bus stops and inside the bus. This model was improved in (3) formulating a transit equilibrium problem that uses effective frequencies functions that vanish if the in-vehicle flow exceeds its capacity (see 3). The main limitation of these methods is that the technical assumptions are very limiting in the first case and there no efficient algorithms to compute the solution in both cases.

Cepeda et al (4) decided to continue this idea and reformulated the equilibrium problem as the minimization of a nonconvex and nondifferentiable gap function. To solve this problem a heuristic method was proposed, using an adaptation of the Method of Successive Averages (MSA) and obtaining the lines flow vector. This method can

* Corresponding author.
E-mail address: victoria.orlando.35@gmail.com

be applied on high scale networks without computational drawbacks but can generate line flows that exceed the capacity when the demands are high. To improve this method, Codina and Rosell (5) presented an algorithm with strict capacities that find the solution of the fixed point inclusion formulation derived from the problem of variational inequality proposed by Codina (6). At each iteration an assignment problem is solved, using Lagrangian duality and the cutting-planes method.

The use of the previous models of transit assignment in any planning study requires the knowledge of the transport demand, commonly known as the origin-destination matrix. To obtain that matrix could be very expensive and sometimes unaffordable in practice. As has been made for the case of traffic assignment (see (7)), in this work we explore its estimation through some directly measurable quantities like the real frequencies of the buses. As we know how to compute, given the demand, the flows, and hence the frequencies, we pose a kind of inverse problem whose solution estimates the actual demand. As far as we know, there is no previous work about public transport demand estimation using this approach. Most of them are based on statistical or econometrical considerations, see (8; 9; 10; 11).

In the next section, we present a detailed description of the assignment model following the one presented in (4). In section 3 we pose the inverse problem used for demand estimation and in section 4 we present the numerical experiments made with the example given in (4).

2. Transit assignment model

Following the notation of previous works (1; 3; 4; 6) we consider a directed graph $G = (N, A)$ where N is the node set and A the link set, each one with cardinality N_N and N_A. The set of nodes is composed of the bus-stop nodes N_s and the line nodes N_l. The arcs are divided in the alighting and boarding arcs connecting the bus-stop nodes with the line nodes, the on-board arcs (or line segments) connecting line-nodes and the walk arcs connecting bus-stop nodes, see figure 1 for a sketch.

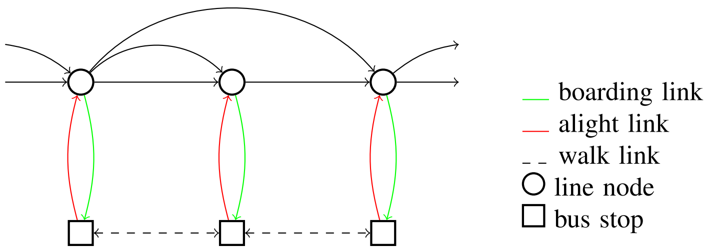


Fig. 1. Public transport network.

For some origin-destination (od) pairs $(i, d) \in W \subset N \times N$, there is a transport demand called g_i^d, and we call D the set of all nodes that are destinations of some od pair. For a node i we call A_i^+ the set of emerging links and A_i^- the incoming link set. We also define the node-link incidence matrix $A \in \mathbb{R}^{N_N \times N_A}$ where $A_{ia} = 1$ iff $a \in A_i^+$, $A_{ia} = -1$ iff $a \in A_i^-$ and otherwise zero.

We call v_a^d the flow through link a with destination d. For each destination d we define the set of feasible flows with destination d and the set of total feasible flows as

$$V^d = \left\{ v^d \in \mathbb{R}_+^{N_A} : Av^d = g^d \right\}, \quad V = \left\{ v \in \mathbb{R}_+^{N_A} : v = \sum_d v^d, v^d \in V^d, \forall d \right\}. \tag{1}$$

We call v_a^d the flow trough link a with destination d, and we call $V(g)$ the set of feasible flows for the demand g, that is the set of all $v_a^d \geq 0$ such that $v_a^d = 0$ for all $a \in A_d^+$.,

$$g_i^d + \sum_{a \in A_i^-} v_a^d = \sum_{a \in A_i^+} v_a^d, \quad \forall i \neq d. \tag{2}$$

Two functions of the full flow vector v are associated to each arc, the travel time function $t_a(v)$ and the effective frequency $f_a(v)$. Both have non negative values and the frequencies can have the constant value $+\infty$. As mentioned in

(4) the case when t_a and f_a are constants is called the *uncongested* case and the case where only the travel time t_a is fixed is called the *semicongested* case. Here we will consider a third case where the travel time function is constant but the frequencies are not. To model the impact of the bus load on the frequency the function 3 is used.

$$f_a(v) \begin{cases} \mu\left[1 - \left(\frac{v_a}{\mu c - v_{a'} + v_a}\right)^{\beta}\right], \text{ if } \; v_{a'} < \mu c, \\ 0, \qquad\qquad\qquad\qquad \text{otherwise,} \end{cases} \tag{3}$$

where $v_a = \sum_{d \in D} v_a^d$ is the total flow boarding at stop and using arc a and $v_{a'}$ is the total flow after the stop ($v_{a'} \geq v_a$). The parameter μ is the nominal frequency of the lines and c is the physical capacity of the buses, thus, $\mu c - v_{a'}$ is the residual capacity waiting at the stop.

The rationale behind the model is that each passenger at each node chooses an arc to continue its trip. The decision is based on minimizing the total travel time. Thus, at each node a Common Line Problem should be solved, where, now, the frequencies depend on the flows. In the paper (4) it is shown that the corresponding (equilibrium) flow $v \in V_0^*$ is the global minimizer of the so-called gap function G of the flow v, that we write here also as a function also of the demand g,

$$G(v, g) = \sum_{d \in D} \left[\sum_{a \in A} t_a(v) v_{a'}^d + \sum_{i \neq d} \max_{a \in A_i^+} \frac{v_a^d}{f_a(v)} - \sum_{i \neq d} g_i^d \tau_i^d(v) \right], \tag{4}$$

where t_a is the travel time, τ_j^d is the total expected travel time from j to d, A_i^+ is the set of links emerging from i, f_a models the impact of the congestion on the frequency, μ is the nominal frequency of the line and c its capacity, β is a calibrated parameter and $v_{a'}$ is the on-board flow right after the stop.

Then the transit assignment for a given demand g is obtained minimizing $G(v, g)$ over the flows in $V(g)$. It is known, also by the work (4), that the optimal value is 0.

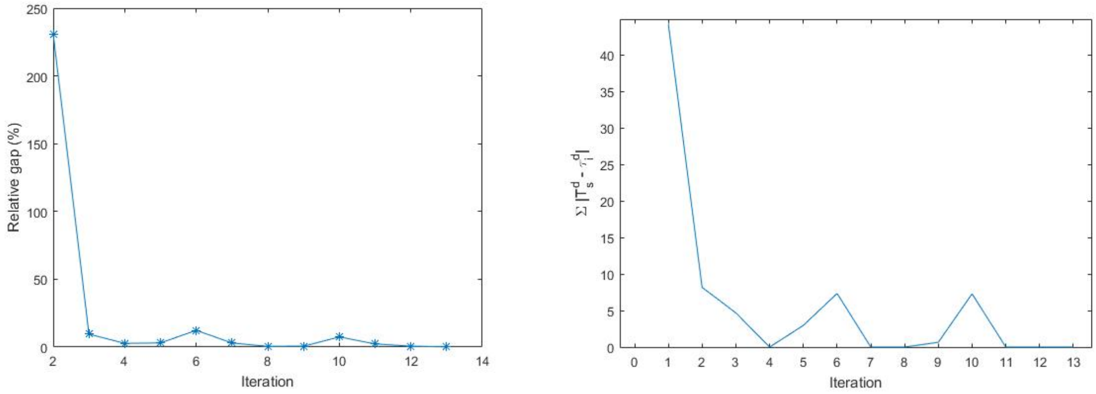
To solve the assignment problem in (3; 4) the authors propose the MSA (Mean Successive Average) method. It means that starting with an all-or-nothing assignment, at each iteration the travel times are updated and a new assignment (for fixed travel times and frequencies) is averaged with the previous one. Interestingly enough, in contrast to the traffic assignment problem, here we have a computable stopping criterium as we know that $G(v, g) = 0$ for an equilibrium. The assignment with fixed travel times and frequencies is made using the Hyperpath Dijkstra method as it was proposed in (4; 1).

For the sake of completeness we reproduce the MSA algorithm below:

Result: Flow at equilibrium
Let $\alpha_k \in (0, 1)$ such that $\alpha_k \to 0$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$;
Find $v^0 \in V_0$ and let $k = 0$;
while $G(v^k) > \epsilon G(v^0)$ **do**
  Compute $t_a = t_a(v^k)$ and $f_a = f_a(v^k)$;
  Compute the shortest hyperpath for each $d \in D$;
  Compute the induced flows $\hat{v}_a^d$;
  Update $v^{k+1} = (1 - \alpha_k)v^k + \alpha_k \hat{v}$;
  Set $k = k + 1$;
end

In order to obtain the first flow $v(0)$, an all-or-nothing assignment is made computing the shortest hyperpath for $t_a = t_a(0)$ and $f_a = f_a(0)$. If $f_a(v^0) = 0$ for some arc a, then the next iteration will be unfeasible. To avoid this situation, the effective frequency can be augmented to $\tilde{f}_a(v) = max\{f_a(v), \varepsilon\}$, for a small enough ε. In this way, even for a large flow, there always will be a feasible arc.

The Figure 2 shows the typical performance of MSA, computed for the second example described in the section 4, using the parameters defined therein.



(a) Relative gap. (b) Sum of the differences between T_s^d and τ_i^d.

Fig. 2. MSA algorithm performance for the example 2 in section 4.

3. Demand estimation problem

Assuming that the model carefully represents the real dynamic of the passengers, it is possibly to use it to detect anomalies or changes in the demand data when the observed flow or frequencies are different from the computed ones.

Here we focus on correcting the given demand to comply with the observed frequencies. That is, given a nominal demand $\bar{g}$ and observed (measured) frequencies $\bar{f}$ over some observed arcs in $A_{obs} \subset A$, we look for the demand g that minimizes

$$\min_{g,v} \sum_{a \in A_{obs}} \left(\frac{\bar{f}_a - f_a}{\bar{f}_a} \right)^2 + \gamma \sum_{a \in A} \left(\frac{\bar{g}_a - g_a}{\bar{g}_a} \right)^2 \quad (5)$$

$$s.t.$$

$$v \in V(g), \quad (6)$$

$$G(v, g) = 0. \quad (7)$$

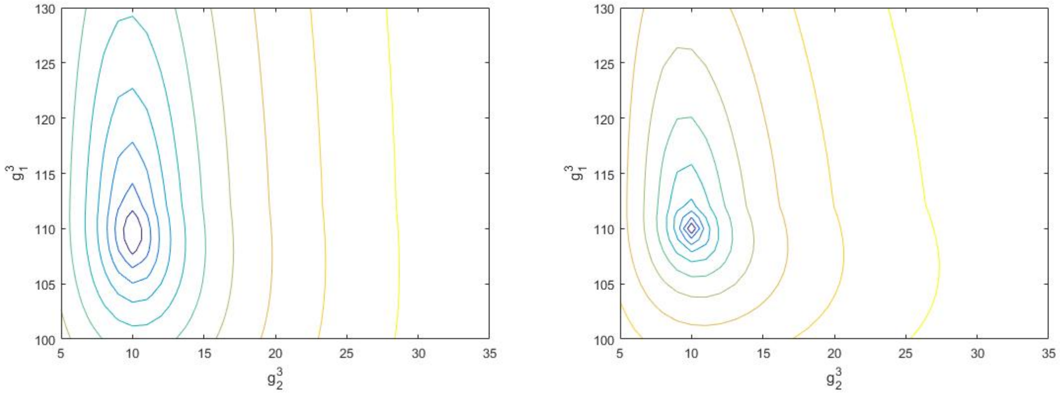
More general quadratic criteria can be considered, for example including coefficients for each arc that represent the confidence of the measures on that arc. The regularization parameter γ represents the trade-off between adjusting the observed flows and conserving the nominal demand; in figure 3 we show the level curves computed for different values of γ in the case of the first example in section 4. The regularization term has a beneficial effect on the convexity of the problem and also on the uniqueness of its solution (see again figure 1, where sublevel sets are "more convex" for γ higher), but large values of γ make the problem to ignore the observations.

Nevertheless, even for large values of γ, i.e, for a more convex problem, the numerical solution of this bilevel problem is rather involved because the flow $v(g)$ is given implicitly by $G(v, g) = 0$ and there is not an easy way to compute variations of v with respect to g.

4. Numerical experiments

For a first numerical experiment, we consider the small example that Cepeda et al. proposed in (4) (Section 4.1.1). We assume that we have the real frequency data and the objective is to estimate the O-D matrix that induces these frequencies.

To find the minimizers in 5 we use the Nelder-Mead method (see (12)). It is a derivative free method included in Matlab through the command *fminsearch* (13), and we considered a precision value of 0.01.



(a) $\gamma = 1/5$ (b) $\gamma = 1/100$

Fig. 3. Level curves (log scale)

4.1. Cepeda et al. network

Consider the network in figure 4 with three nodes and two transit lines connecting them: L_1 (local line, connecting nodes 1, 2 and 3) and L_2 (express line, connecting node 1 with node 3). Suppose that we have demands of 10 trips from node 1 to node 2, 100 trips from node 1 to node 3 and 10 trips from node 2 to node 3. Considering that the capacity of each bus is 20 passenger by bus, the dwell time at stops is 0.01 minutes and the effective frequencies are defined by 3 with $\beta = 0.2$.

Finally suppose that the frequency of lines L_1 and L_2 is 6 and 16 vehicles per hour, respectively, and travel times over each link are $t_{12} = 20.01$, $t_{23} = 20.01$ and $t_{13} = 24.01$ minutes.

In order to obtain the equilibrium assignment we applied the MSA Algorithm. It is important to note that demands g_1^2 and g_2^3 can only use the line L_1 while demand g_1^3 can choose L_1 or L_2. Taking this into account we obtained the following link volumes:

$$v_{12} = 25.7, \quad v_{23} = 25.7, \quad v_{13} = 84.3$$

where it can be seen that passengers who want to travel from node 1 to node 3 choose a strategy that considers both lines, local and express.

For this assignment the total time (travel + wait) of each strategy for each demand g_i^d satisfies the equilibrium condition $T_s^d = \tau_i^d$. In the particular case of g_1^3 the total travel time is equal to 40.02 minutes.

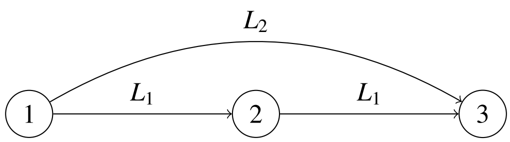


Fig. 4. Small network proposed by (4)

The effective frequencies based on these assignment are $f_{12} = 0.0265$, $f_{23} = 0.0374$ y $f_{13} = 0.0625$.

Suppose we can measure the current effective frequencies and based on this and a nominal demand we want to estimate the current O-D matrix. Consider, for example, the following observed frequencies:

$$\bar{f}_{12} = 0.0215, \quad \bar{f}_{23} = 0.0362, \quad \bar{f}_{13} = 0.0624$$

These frequencies are obtained when we perform the flow assignment with $g_1^2 = 10$, $g_1^3 = 110$ and $g_2^3 = 10$. Taking into account these frequencies and considering the nominal O-D matrix $\bar{g}_1^2 = 10$, $\bar{g}_1^3 = 100$ and $\bar{g}_2^3 = 10$ we solve the problem 5 with $\gamma = 1/5$ and obtain the estimated O-D matrix $g_1^2 = 10.05$, $g_1^3 = 109.5$ and $g_2^3 = 9.98$, which can be considered a good approach to the assumed real O-D matrix $\bar{g}_1^2 = 10$, $\bar{g}_1^3 = 110$ and $\bar{g}_2^3 = 10$. The progress of the objective function of problem 5 during the O-D matrix estimation can be seen in the Figure 6.

4.2. *Another network*

In order to reproduce the previous methodology in another network we consider a new example with four nodes and four lines serving it as shown in Figure 5. The data of each line are summarized in Table 1. Considering the demands $g_1^3 = g_1^4 = g_4^3 = 100$ the MSA Algorithm was applied and the results are exposed in Table 2. Table 3 summarizes the link flows obtained summing over all destinations and considering all demands. The effective frequencies obtained for this assignment are also shown there.

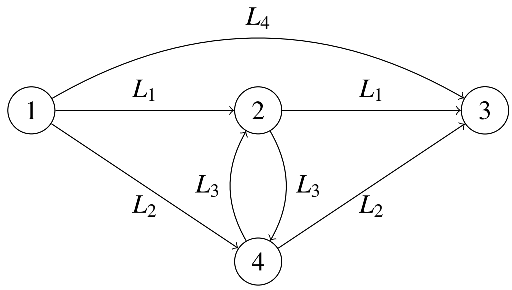


Fig. 5. Network with 4 nodes and 4 lines

Table 1. Service data

Line	Route	Travel times (min)	Frequencies (veh/h)
L_1	$1 \to 2 \to 3$	$t_{12} = t_{23} = 20.01$	8
L_2	$1 \to 4 \to 3$	$t_{14} = t_{43} = 22.01$	16
L_3	$2 \to 4 \to 2$	$t_{24} = t_{42} = 5.01$	16
L_4	$1 \to 3$	$t_{13} = 28.01$	10

Table 2. Disaggregated flows resulting for assignment in example 2.

Demand	Link flows	Lines used	Total cost
g_1^4	$v_{12}^4 = 32.54$, $v_{14}^4 = 67.46$, $v_{24}^4 = 32.54$	L_1, L_2, L_3	$T_s^4 = 39.1260$
g_4^3	$v_{23}^3 = 6.37$, $v_{43}^3 = 93.63$, $v_{42}^3 = 6.37$	L_1, L_2, L_3	$T_s^3 = 41.0354$
g_1^3	$v_{12}^3 = 21.69$, $v_{23}^3 = 21.69$, $v_{14}^3 = 39.38$ $v_{43}^3 = 39.38$, $v_{13}^3 = 38.93$, $v_{43}^3 = 39.38$, $v_{13}^3 = 38.93$	L_1, L_2, L_4	$T_s^3 = 45.1520$

Table 3. Total flows and effective frequencies resulting for assignment in example 2.

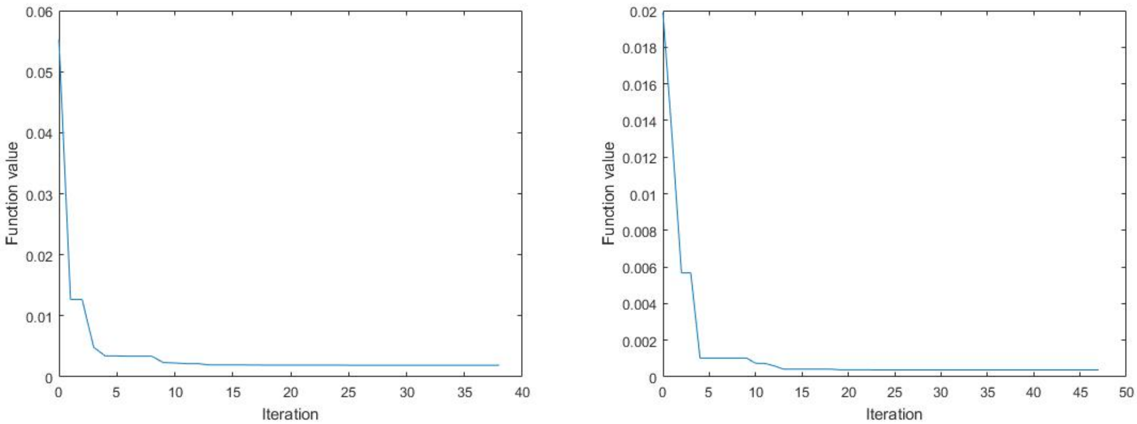
Link (i, j) / Results	(1,2)	(2,3)	(1,4)	(4,3)	(2,4)	(4,2)	(1,3)
Link flows	54.24	28.06	106.83	133.01	32.54	6.37	38.93
Effective frequencies	0.0259	0.0613	0.0525	0.0526	0.0978	0.1448	0.0465

In order to estimate the O-D matrix we have the measured frequencies:

$$f_{12} = 0.0243, \quad f_{23} = 0.0591, \quad f_{14} = 0.0489, \quad f_{43} = 0.0515$$
$$f_{24} = 0.0977, \quad f_{42} = 0.1419, \quad f_{13} = 0.0428,$$

that are obtained when an assignment is made with $g_1^3 = 120$ and $g_1^4 = g_4^3 = 100$.

Using the nominal demands $g_1^3 = g_1^4 = g_4^3 = 100$ we solved the problem (5-7) with $\gamma = 1/100$ and observed frequencies obtained for a demand of $g_1^3 = 120$ and $g_1^4 = g_4^3 = 100$, obtaining the following demand estimation $g_1^3 = 118.86$, $g_1^4 = 100.75$ and $g_4^3 = 100.18$. The progress of the objective function of problem (5-7) during the O-D matrix estimation can be seen in the Figure 6.



(a) Example 1 (b) Example 2

Fig. 6. Progress of objective function in (5).

5. Conclusions

In this work, we have proposed an approach to public transport demand estimation. Given a model of flow distribution for public transport according to its demand, we propose the solution of an inverse problem to update the demand for observed flow variations. Preliminary results show that it can be done with derivative-free optimization algorithms over small-sized networks. The numerical analysis for larger networks and the search for analytical derivation of descent directions are currently under work.

Acknowledgements

We thank the advice of two anonymous referees and the help of R. Cominetti to understand the computations in (4).

References

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

[2] J. D. Cea, E. Fernández, Transit assignment for congested public transport systems: An equilibrium model, Transportation Science 27 (2) (1993) 133–147.

[3] R. Cominetti, J. Correa, Common-lines and passenger assignment in congested transit networks, Transportation Science 35 (3) (2001) 250–267.

[4] M. Cepeda, R. Cominetti, M. Florian, A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria, Transportation Research 40 B (2006) 437–459.

[5] E. Codina, F. Rosell, A heuristic method for a congested capacitated transit assignment model with strategies, Transportation Research 000 B (2017) 1–28.

[6] E. Codina, A variational inequality reformulation of a congested transit assignment model by cominetti, correa, cepeda, and florian, Transportation Science 47 (2) (2013) 231–246.

[7] J. Walpen, E. M. Mancinelli, P. A. Lotito, A heuristic for the od matrix adjustment problem in a congested transport network, European Journal of Operational Research 242 (3) (2015) 807 – 819. doi:https://doi.org/10.1016/j.ejor.2014.11.009.
URL http://www.sciencedirect.com/science/article/pii/S0377221714009138

[8] J. d. D. Ortuzar, L. G. Willumsen, Modelling transport (3rd edn.), Wiley, New Tork, 2001.

[9] E. Cascetta, TRANSPORTATION SYSTEMS Springer Optimization and Its Applications, Vol. 10. doi:10.1016/0041-1647(76)90008-3.
URL http://www.springer.com/mathematics/applications/book/978-0-387-75856-5

[10] D. N. Dike, C. C. Ibe, E. A. Ejem, O. Erumaka, O. E. Chukwu, Estimation of inter-city travel demand for public road transport in nigeria, Journal of Sustainable Development of Transport and Logistics 3 (2018) 88–98.

[11] A. García-Ferrer, M. Bujosa, A. de Juan, P. Poncela, Demand forecast and elasticities estimation of public transport, Journal of Transport Economics and Policy (2006) 45–67.

[12] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, Convergence properties of the nelder-mead simplex method in low dimensions, SIAM Journal of Optimization 9 (1) (1998) 112 – 147.

[13] MATLAB, version 9.3.0 (R2017b), The MathWorks Inc., Natick, Massachusetts, 2017.