

Multiobjective design of sustainable public transportation systems

Diego Gabriel Rossit^{1,2}, Sergio Nesmachnow³ and Jamal Toutouh⁴

¹ Department of Engineering, Universidad Nacional del Sur, Argentina

² Instituto de Matemática de Bahía Blanca UNS-CONICET, Argentina

³ Universidad de la República, Uruguay

⁴ Massachusetts Institute of Technology, USA

E-mail: diego.rossit@uns.edu.ar, sergion@fing.edu.uy, toutouh@mit.edu

Abstract. The design of the bus network is a complex problem in modern cities, since different conflicting objectives have to be considered, from both the perspective of bus companies and the citizens. This article presents a multiobjective model for designing a sustainable public transportation network that simultaneously optimizes the covered travel demands by passengers, the total travel time, and the generated pollution. The proposed model is solved using exact weighted sum and a heuristic procedure based on the standard shortest path problem. Preliminary tests were performed in small real-world instances of Montevideo, Uruguay. Experiments allowed obtaining a set of compromising solutions that in turn allow exploring different trade-off among the optimization criteria. The proposed heuristic was competitive, being able to find a good compromising solution in short computing times.

1. Introduction

The expansion of urban population in the last century has put pressure on decision-makers to provide efficient public services to the citizens. Among these services, having an efficient public transportation system is of paramount importance in order to diminish the impact of private cars. In the recent past, this system had to fulfil two main goals: provide a good Quality of Service (QoS) to the citizens and burden the operating costs for the operators of the system [1]. However, in the last decades, sustainability has also been integrated as an important criterion in public transportation system [2]. Public transportation can be an important contributor to air pollution [3], affecting the environment and quality of life of the citizens. In this sense, several cities have implemented diverse policies for reducing the impact of this vital service. For example, in the last years, the City Hall of Montevideo has increasingly acquired a number of electric and hybrid buses [4].

Among the diverse decision-making problems involved in public transport, the Bus Transit Network Design Problem (TNDP) is a common computationally complex [5] problem that arises in transportation planning in modern cities. It includes the definition of bus lines layouts in an urban area while enhancing some desired criteria.

2. Related work

The Bus Transit Network Design Problem (TNDP) has been addressed in several works in the related literature. A recent review of the topic was performed by Ibarra et al. [1]. Among the diverse authors that used mathematical programming to deal with this problem, it is relevant the work performed by Cancela et al. [6], who revised different formulations from the literature and proposed a new one that modelled the waiting time of users and the behaviour of users when multiple lines are available. Wan and Lo [7] proposed a formulation that minimizes the cost of the operator company considering the capacity of the bus. Szeto and Jiang [8] have proposed a bi-level formulation. The upper level determines the routes path and bus frequencies with the objective of minimizing the number of transfers subjected to constraints on the maximum number of routes, fleet size, minimum frequency and enough line capacity. In the second level, the model proposed in Spiess and Florian [9] that applies strategies to improve the travel time is used. However, the improvements over the travel time in this second level are limited by the decisions taken on the first level [6]. More recently, Liang et al. [10] proposed a two-step model framework to design the bus transit network. Firstly, a column generation method is applied to identify the candidate set of path of the buses lines. Then, they used stochastic linear programming to optimize the bus frequency considering uncertain demand and travel times.

In last decade, with the rise of environmental awareness, societies have started to require public transport systems that are not only economically efficient and respond to the transport needs of citizens but also that they are environmentally sustainable [2]. In this line, Duran et al. [5] have proposed a mathematical model for the TNDP which simultaneously considers the minimization of travel time and CO² emissions. Articles from our group also studied the paradigm of sustainable mobility and specific case studies in Montevideo, addressing the characterization of recent sustainable initiatives developed for the public transportation of Montevideo, Uruguay [4]. In turn, recommendations based on a specific analysis of quantitative (e.g., coverage, accessibility, affordability) and qualitative (e.g., public finance, integration, comfort and pleasure) were proposed for the development of sustainable mobility in Parque Rodó neighborhood [11]. The design of an optimized backbone for the public transportation network (i.e., using light railway trams) considering combinatorial optimization approaches was analyzed in Risso and Nesmachnow [12]. The approach integrated real demands, travel and waiting times, and deployment costs, to provide a proper design for sustainable mobility in Montevideo, to improve quality of service. Data analysis plays a major role to characterize the mobility needs and the current reality, and it has been a subject of study by applying urban informatic approaches [13].

According to the World Health Organization, air pollution is a significant-top health hazard in the cities due to its impact on the citizens' health [14, 15]. Understanding the phenomena that have implications for the production and dissipation of air pollutants has attracted the interest of the scientific community [16]. Thus, there is an important field of research on modelling and forecasting ambient air pollution [17, 18]. Physics-based approaches have been applied to address air pollution modeling [19]. However, these methods are computationally expensive. During the last years, with the rapid development of artificial neural networks and deep learning and their successful use for forecasting applications, several researchers have proposed them to deal with air outdoor pollution modelling, prediction, and forecasting [20, 21]. Most of these methods predict the air pollution concentration giving the previous pollution concentration and other external conditions, such as the weather. It has been shown that Urban road mobility affects air quality [22, 23]. Thus, different approaches based on ANNs model the air pollution according to the current road traffic density [24].

This article introduces a new mathematical formulation that aims at defining a bus network considering the simultaneous maximization of the number of passengers and minimization of the total travel time and the pollution generated. The formulation is solved by means of an

exact solver. As far as we are concerned, these three objectives have not been simultaneously considered in an exact resolution for the TNDP. For the sake of comparison with the exact solver, we also present a fast heuristic approach. In this initial work, we predict the pollution generated considering the population density and the type of bus (electric, hybrid or diesel). In future approaches, as TNDP impacts the overall road traffic, it is expected to use the aforementioned ANNs models based on the road traffic density as we implemented in Jamal et al. [25].

3. Mathematical formulation

The TNDP is addressed with a mixed-integer programming model. This model aims at optimizing three different objectives: minimizing the travel time, minimizing the pollution generated by the buses and maximizing the number of passengers served by the buses. The network is designed as a set of paths for each bus from a starting node up to the end node. It is assumed that the bus return through the same path, a simplified idea that has been used in the related literature [7]. In this model, the capacity of buses is not considered, since it is supposed to be adjusted in a posterior decision-making process when considering the frequencies setting of buses lines [1, 26]. Thus, a bus line captures all the demand of the neighbourhoods that it connects. In this model, the neighbourhoods coincide with the administrative divisions of the city (census segments). Thus, the model considers the following sets:

- a set of ordered buses $K = (k_0 \dots k_e, k_{e+1} \dots k_{e+h}, k_{e+h+1} \dots k_{e+h+d})$, where the first e buses are electric buses, the following h are hybrid buses and the final d buses are diesel buses;
- a set of census segments (hereafter segments) $I = (i_0 \dots i_{|I|})$.

Given a set of segments, a specific notation is used for the set of edges between adjacent segments: $A(\cdot)$, such that: $A(I) = \{(i, j) | i \in I, j \in I, i \neq j, i \in N_I(j), j \in N_I(i)\}$ where $N_I(i)$ represents the subset of neighbors of segment i among the segments of set I . Additionally, superset $I^{sg} = I \cup i^s \cup i^g$ is defined, where i^s is the base from which buses start their paths and i^g is the terminal from which buses end their forward trip and start their return trip.

The parameters of the proposed model are: w_i^k , the pollution generated by the bus k on segment i ; p_{ij} , the estimated demand (number of passengers) that are willing to travel from segment i to j ; t_{ij} , the travel time going from segment i to j ; and TL^k maximum travel time of bus k . All variables are binary: x_{ij}^k indicates if bus k does the trip between arc $(i, j) \in A(I)$, y_{ij}^k indicates if bus k does the trip between nodes i and j , and f_{ij} indicates if any bus does the trip between nodes i and j . The problem proposes finding a set of layout functions $X^k = \{x_{ij}^k\}$ stating for each bus the path that simultaneously maximizes the passengers served by the network and minimizes the travel time of the used buses and the generated pollution.

$$\text{minimize } TT = \sum_{ij \in A(I^{sg})} \sum_{k \in K} \left(x_{ij}^k (t_{ij} + t_{ji}) \right) \quad (1a)$$

$$\text{maximize } D = \sum_{ij \in I^{sg}} \sum_{k \in K} (f_{ij} (p_{ij} + p_{ji})) \quad (1b)$$

$$\text{minimize } P = \sum_{ij \in A(I^{sg})} \sum_{k \in K} \left(w_j^k x_{ij}^k \right) \quad (1c)$$

$$\text{subject to } \sum_{i \in N_I(j)} x_{ij}^k = \sum_{i \in N_I(j)} x_{ji}^k, \forall k \in K, j \in I \quad (1d)$$

$$\sum_{j \in N_I(i)} x_{ij}^k = 1, \forall k \in K, i \in I^s \quad (1e)$$

$$\sum_{ij \in A(I)} t_{ij} x_{ij}^k \leq TL^k, \forall k \in K \quad (1f)$$

$$\sum_{ij \in A(I)} t_{ji} x_{ij}^k \leq TL^k, \forall k \in K \quad (1g)$$

$$\sum_{j \in N_I(i^g)} x_{ji}^k = \sum_{N_I(i^s)} x_{i^s j}^k, \forall k \in K \quad (1h)$$

$$\sum_{j \in N_{I^{sg}}(j)} x_{ij}^k \geq y_{im}^k, \forall k \in K, i, m \in I^{sg}, i \neq m \quad (1i)$$

$$\sum_{j \in N_{I^{sg}}(m)} x_{jm}^k \geq y_{im}^k, \forall k \in K, i, m \in I^{sg}, i \neq m \quad (1j)$$

$$\sum_{k \in K} y_{im}^k \geq f_{im}, \forall i, m \in I^{sg}, i \neq m \quad (1k)$$

$$x, f \in \mathbb{B} \quad 0 \leq y \leq 1$$

Equations (1a) and (1b) are the total travel time and the number of passengers served by the buses, respectively, considering the contribution of both directions of traversed arc since the bus is supposed to travel the same trip in the return path. Eq. (1c) is the estimated generated pollution by buses. Regarding constraints, Eq. (1d) establishes that if a bus enters a node, it has to leave it. Eq. (1e) sets that at most one segment can be visited once by each bus. Equations (1f) and (1g) limit the maximum travel time of each bus in the forward and return trip. Equations (1i)–(1k) establish that f_{im} is one if and only if neighborhoods i and m are visited by the same bus. Eq. (1h) enforces that if a bus has departed from the starting node, it reaches the ending point. Due to the mathematical structure of the model, the binary nature of variable y can be relaxed—i.e., to a real value in $[0,1]$ —without affecting the validity of the model.

4. Experimental evaluation

The proposed model was validated on two problem scenarios from Montevideo, Uruguay, defined over zones including one hundred census segments of the city, each. The two zones are depicted in Figure 1: the red scenario is located near downtown and has densely populated small segments; the yellow scenario is located in a peripheral zone of the city, where segments are relatively large. The red scenario considers three buses: one electric, one hybrid, and one diesel, and yellow scenarios consider five buses: two electric, one hybrid, and two diesel. Passengers demand between segments was retrieved from a public database of origin/destination of trips in Montevideo [27, 28]. The pollution generated by a bus on a segment is estimated considering two aspects: the type of bus, i.e., if it is electric, hybrid or diesel, and the population density of the segment. Travel times were measured using the method proposed by Vázquez for OSRM [29].

Ten different instances were solved for each scenario. Three instances correspond to single-objective versions of the problem presented in Section 3: SO-TT considers only the optimization problem of Eq. (1a), SO-P considers the optimization problem of Eq. (1b) and SO-D considers the optimization problem of Eq. (1c). In turn, seven instances consider the multiobjective problem, using the weighting sum approach [30], in which the sum is normalized using the single-objective results as in our previous work [31]. Seven different weight vectors of the form (w_{TT}, w_P, w_D) were used (w_{TT} , w_P , and w_D are the respective weights assigned to Eq. 1a, Eq. 1b, and Eq. 1c in the multiobjective weighted sum). The first three correspond to highly biased instances in which the weight of one of the objectives is much larger than the others. Thus, MO-TT* instance uses (0.98,0.01,0.01), MO-D* uses (0.01,0.98,0.01) and MO-P* uses (0.01,0.01,0.98). The other

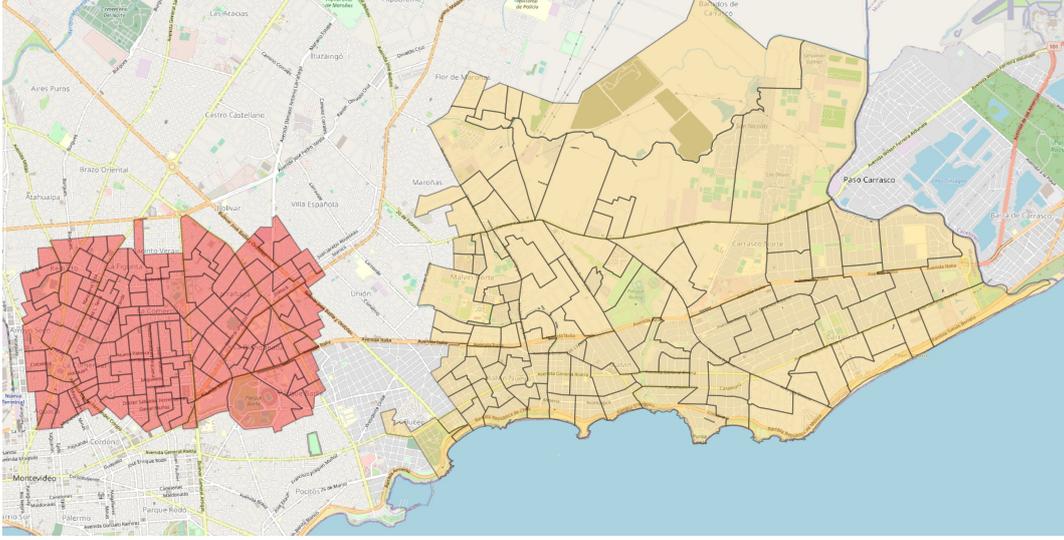


Figure 1. Test instances on Montevideo city (background map from OpenStreetMap).

four multiobjective instances aim at computing more balanced compromise solutions. MO-TT instance uses $(0.6, 0.2, 0.2)$, MO-P uses $(0.2, 0.6, 0.2)$, MO-D uses $(0.2, 0.2, 0.6)$, and MO-CS uses $(0.33, 0.33, 0.33)$.

In turn, a fast heuristic (FH) procedure is proposed as a baseline for comparing exact solutions computed by the proposed model. FH solves a simplified version of the model presented in Section 3. The schema of FH is outlined in Algorithm 1. The objective function minimizes the impedance of the used arcs, which is calculated as a random weighted sum of the normalized travel time of the arc and the pollution generated at the destination segment. The impedance function is optimized within a simplified linear programming model SP only considering constraints Eqs. (1d)-(1h).

Algorithm 1 Global procedure of the fast heuristic FH

```

1: procedure HEURISTIC( $I, K, w_{ki}, t_{ij}$ )
2:    $\max_t \leftarrow \max(t_{ij})$  ▷ Gets the maximum travel time
3:    $\min_t \leftarrow \min(t_{ij})$  ▷ Gets the minimum travel time
4:    $\max_{w_k}, \min_{w_k} \leftarrow 0$ 
5:   for  $k \leftarrow 0; k ++; k \leq K$  do ▷ For each bus
6:      $\max_w.append(\max(w_{ki}[k]))$ ; ▷ Gets the maximum pollution generation
7:      $\min_w.append(\min(w_{ki}[k]))$  ▷ Gets the minimum pollution generation
8:   Create impedance matrix  $\text{imp}$ 
9:   for each  $(i, j) \in A(I)$  do ▷ For each edge in the adjacency matrix
10:    for  $k \leftarrow 0; k ++; k \leq K$  do
11:      $\alpha \leftarrow \text{random}(0, 1)$  ▷ Build impedance matrix
12:      $\text{imp}_{ij} \leftarrow \frac{t_{ij}[i, j] - \min_t}{\max_t - \min_t} (1 - \alpha) * \frac{w_{ki}[k, j] - \min_w[k]}{\max_w[k] - \min_w[k]}$ 
13:   Solve  $SP$  with  $\text{cost} \leftarrow \sum_{(i, j) \in A(I)} \text{imp}_{ij} - \times x_{ij}^k$ 

```

Solutions were obtained using Pyomo [32] as the modelling language and Gurobi [33] as the exact solver. The experimental analysis was performed on a computer with Intel Processor i7-4790 CPU @3.60GHz and 32GB RAM, using a time limit of 2400 seconds for the solver.

Table 1 presents the experimental results. For each instance the Table reports: the value of each objective (TT , P , and D); the deviation from the ideal solution Σ , computed as the distance to the ideal solution with the L^2 norm (Equation 2, where $O = \{TT, P, D\}$ is the set of objectives and $best_o$ and $worst_o$ are the best and worst value achieved for each objective); the computing time in seconds; and the optimality gap estimated by Gurobi gap_G . For the FH heuristic gap_G is not reported, since the solution is not obtained by an exact approach.

$$\sqrt{\sum_{o \in O} \left(\frac{value - best_o}{worst_o - best_o} \cdot 100\% \right)^2} \quad (2)$$

Instance	TT	P	D	Σ	time (s)	gap_G
<i>yellow scenario</i>						
SO-TT	172.55	11.43	8968	108.20%	1	0.00%
SO-P	384.35	8.26	17956	142.73%	1	0.00%
SO-D	642.65	46.91	395029	200.00%	2400	6.78%
MO-TT*	172.55	11.43	9329	108.11%	1	0.00%
MO-P*	319.20	8.26	16587	129.22%	5	0.00%
MO-D*	555.31	31.18	342059	154.44%	2400	139.54%
MO-TT	172.55	11.34	9329	107.88%	3	0.00%
MO-P	<i>No feasible solution obtained within the specified time limit</i>					
MO-D	345.94	15.71	172667	113.76%	2400	80.40%
MO-CS	<i>No feasible solution obtained within the specified time limit</i>					
FH	286.44	10.32	27256	124.82%	3	-
<i>red scenario</i>						
SO-TT	71.256	1.616	35536	119.37%	1	0.00%
SO-P	93.366	0.426	89216	105.49%	1	0.00%
SO-D	387.726	6.526	3988966	200.00%	2400	0.06%
MO-TT*	71.256	1.346	29986	115.08%	2	0.00%
MO-P*	92.86	0.846	109286	111.71%	3	0.00%
MO-D*	238.15	4.09	205100	161.85%	2400	98.10%
MO-TT	71.25	1.34	3376	114.99%	11	0.00%
MO-P	92.36	0.84	11395	111.43%	11	0.00%
MO-D	187.47	2.67	340330	88.40%	2400	19.09%
MO-CS	87.96	1.02	36567	106.64%	512	0.00%
FH	122.00	0.84	9976	121.16%	1	-

Table 1. Objective function values for the proposed instances.

Results in Table 1 show that the proposed model was able to obtain a set of solutions with different trade-off among the optimization criteria. In the case of the yellow scenario, the instance with the smallest deviation was MO-TT* (108.11%); in the red scenario, it was MO-D (88.40%). The instance with the largest deviation was SO-D (200% percentage deviation) in both cases. Regarding the heuristic procedure (FH), it was able to compute good compromising solutions, since the overall deviation is not far from the instance with the smallest deviation: about 16.94% (compared to MO-TT*) in the yellow scenario and 32.76% (compared to MO-D) in the red scenario.

Another important outcome of the experimentation is the impact that the number of relevant integer variables of the model has on the resolution efficiency. Related to this issue, the increment in the integer variables is usually negatively correlated to the computational complexity of the problem [34]. In line with this concept, SO-TT and SO-P—in which the objective function Eq. (1c) is not involved and, thus, Eqs. (1i)–(1k) (and integer variable f) become irrelevant since they are not part of the optimization criteria—are easily solved by Gurobi. On the other hand, in all the time-consuming instances (those for which the time limit is reached without obtaining the optimal solution and those for which no solution is found) Eq. (1c) is included. In these cases, Eqs. (1i)–(1k) (and integer variable f) become relevant for the optimization process—since they participate in the objective function—and, thus, the problem has a larger number of relevant integer variables.

5. Discussion and future work

The Bus Transit Network Design Problem arises in the context of public bus transportation in modern cities and basically consists in the definition of buses lines layouts in an urban area while aiming at enhancing some desired criteria. In this article, a novel multiobjective mathematical formulation is presented for this problem which simultaneously optimizes two traditional objectives, i.e., the number of passengers that are served by the network and the total travel time of the buses, and a less conventional objective related to the pollution generation of the network. This model was solved with exact and heuristic procedures for two realistic scenarios of Montevideo city, Uruguay, obtaining different compromising solutions that allow exploring the trade-off among the objectives.

The main lines for future research include the application of more advanced exact multiobjective approaches for the exact resolution, such as augmented ε -constraint method. Another feature that will be addressed to enhance the model is to consider uncertainty in the pollution generation since this parameter can be affected by several elements (e.g., the season of the year, the weather and traffic conditions, economic activities that are performed in the segment, and buildings skyline affecting air circulation). In this sense, we are developing a model based on artificial neural networks to estimate more accurately the pollution generation of a bus that transits a certain area of the city. This network will be trained with data gathered by air-quality monitors displaced in Montevideo.

Acknowledgements

This research was partially funded by European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 799078, by the European Union H2020-ICT-2019-3 and UMA18-FEDERJA-003, and the Systems that Learn Initiative at MIT CSAIL.

References

- [1] Ibarra-Rojas O J, Delgado F, Giesen R and Muñoz J C 2015 *Transportation Research Part B: Methodological* **77** 38–75
- [2] Miller P, de Barros A G, Kattan L and Wirasinghe S 2016 *KSCE Journal of Civil Engineering* **20** 1076–1083
- [3] Reisi M, Aye L, Rajabifard A and Ngo T 2014 *Ecological Indicators* **43** 288–296
- [4] Hipogrosso S and Nesmachnow S 2019 *Communications in Computer and Information Science* **1152** 93–108
- [5] Duran J, Pradenas L and Parada V 2019 *Public Transport* **11** 189–210
- [6] Cancela H, Mauttone A and Urquhart M E 2015 *Transportation Research Part B: Methodological* **77** 17–37
- [7] Wan Q K and Lo H K 2003 *Journal of Mathematical Modelling and Algorithms* **2** 299–308
- [8] Szeto W Y and Jiang Y 2014 *Transportation Research Part B: Methodological* **67** 235–263
- [9] Spiess H and Florian M 1989 *Transportation Research Part B: Methodological* **23** 83–102
- [10] Liang J, Wu J, Gao Z, Sun H, Yang X and Lo H K 2019 *Transportation Research Part B: Methodological* **126** 115–138
- [11] Hipogrosso S and Nesmachnow S 2020 *Smart Cities* **3** 479–510
- [12] Risso C and Nesmachnow S 2020 *Smart Cities* pp 228–243
- [13] Nesmachnow S, Baña S and Massobrio R 2017 *EAI Endorsed Transactions on Smart Cities* **2** 153478

- [14] World Health Organization 2018 Ambient (outdoor) air quality and health [https://www.who.int/en/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health) Accessed: 2019-07-07
- [15] Lebrusán I and Toutouh J 2020 *Air Quality, Atmosphere & Health* **14**(3) 333–342
- [16] Jorquera H, Montoya L D and Rojas N Y 2019 *Urban Air Pollution* (Cham: Springer International Publishing) pp 137–165
- [17] Venkatram A 2015 *Lectures on air pollution modeling* (Springer)
- [18] Zannetti P 2013 *Air pollution modeling: theories, computational methods and available software* (Springer Science & Business Media)
- [19] Chuang M T, Zhang Y and Kang D 2011 *Atmospheric environment* **45** 6241–6250
- [20] Cabaneros S M, Calautit J K and Hughes B R 2019 *Environmental Modelling & Software* **119** 285–304
- [21] Toutouh J, Lebrusán I and Nasmachnow S 2020 *International Conference on Optimization and Learning* (Springer) pp 115–127
- [22] Lebrusán I and Toutouh J 2019 *Ibero-American Congress on Information Management and Big Data* (Springer) pp 9–24
- [23] Lebrusán I and Toutouh J 2020 *Smart Cities* **3** 456–478
- [24] Toutouh J 2021 *Smart Cities* ed Nasmachnow S and Hernández Callejo L (Cham: Springer International Publishing) pp 90–105 ISBN 978-3-030-69136-3
- [25] Toutouh J, Nasmachnow S and Rossit D G 2020 *1st International Workshop on Advanced Information and Computation Technologies and Systems (AICTS)* (Irkutsk, Russia)
- [26] Nasmachnow S, Muraña J and Risso C 2020 *Smart Cities (Communications in Computer and Information Science* vol 1359) (Springer International Publishing) pp 183–198
- [27] Massobrio R and Nasmachnow S 2020 *Applied Sciences* **10** 1–20
- [28] Massobrio R Origin-destination transport matrices of the city of Montevideo, Uruguay <https://www.fing.edu.uy/renzom/msc/od-matrices.md> Accessed: 2020-11-26
- [29] Vázquez Brust A 2018 *Rpubs by RStudio* <https://rpubs.com/HAVB/osrm>
- [30] Zadeh L 1963 *IEEE transactions on Automatic Control* **8** 59–60
- [31] Rossit D G, Toutouh J and Nasmachnow S 2020 *Waste Management* **105** 467–481
- [32] Hart W E, Laird C D, Watson J P, Woodruff D L, Hackebeil G A, Nicholson B L and Siirola J D 2017 *Pyomo-optimization modeling in python* 2nd ed vol 67 (Springer Science & Business Media)
- [33] Gurobi Optimization L 2021 Gurobi optimizer reference manual URL <http://www.gurobi.com>
- [34] Glover F 1984 *Journal of Information and Optimization Sciences* **5** 69–71