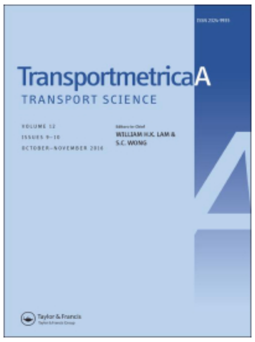

Taylor & Francis
Taylor & Francis Group

Transportmetrica A: Transport Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ttra21

A maximum entropy optimization model for origin-destination trip matrix estimation with fuzzy entropic parameters

Héctor López-Ospina, Cristián E. Cortés, Juan Pérez, Romario Peña, Juan Carlos Figueroa-García & Jorge Urrutia-Mosquera

TRANSPORTMETRICA A: TRANSPORT SCIENCE
https://doi.org/10.1080/23249935.2021.1913257

A maximum entropy optimization model for origin-destination trip matrix estimation with fuzzy entropic parameters

Héctor López-Ospina [a], Cristián E. Cortés [b], Juan Pérez [c], Romario Peña [a], Juan Carlos Figueroa-García [d] and Jorge Urrutia-Mosquera [e]

[a]Universidad del Norte, Barranquilla, Colombia; [b]Civil Engineering Department, Universidad de Chile, Instituto Sistemas Complejos de Ingeniería (ISCI), Santiago, Chile; [c]Universidad de Los Andes, Chile; [d]Universidad Distrital Francisco José de Caldas, Bogotá, Colombia; [e]Facultad de Ingeniería, Universidad del Desarrollo, Santiago, Chile

ABSTRACT

We formulate a bi-objective distribution model for urban trips constrained by origins and destinations while maximizing entropy. We develop a flexible and consistent approach in which the estimations of generated/attracted parameters are fuzzy with entropic membership functions. Based on a fuzzy-entropy approach, we measure the uncertainty associated with fuzzy variables. We solve the problem by means of compromise programming considering a weighted sum objective function. We compute and extend concepts such as accessibility, attractiveness, and generalized cost, typically obtained in transport economic analyzes. Considering that our formulation is convex, we solve the problem in one step only, maintaining the uniqueness of the the optimization problem solution. We present two numerical examples to illustrate the proposed methodology, analyzing the impact of the results based on strong mathematical and statistical arguments. Finally, we show that our approach has better prediction capabilities than traditional fuzzy models regarding aggregated indicators and structural distribution patterns.

ARTICLE HISTORY
Received 16 August 2019
Accepted 17 March 2021

KEYWORDS
Entropy optimization; Origin-destination trip matrix; Transport distribution; Fuzzy sets; Fuzzy entropy

1. Introduction

In the context of transportation planning, the four-stage modeling approach, which utilizes information regarding infrastructure provisions and public transport services as well as socioeconomic data from the population, is a well-known methodology used for the evaluation of hypothetical scenarios at a strategic level (Feng, Zhang, and Fujiwara 2009; Ortuzar and Willumsen 2011). Normally, the urban system is split into a tiling system of zones matched with different networks for each available mode; the sub-models are thus defined as static representations of the assignment of trips during one or several periods (for example, a whole day or peak hours). The four-stage model follows a sequential structure, even though in some cases, the formulation represents a simultaneous execution of

CONTACT Juan Pérez jperez@uandes.cl Facultad de Ingeniería y Ciencias Aplicadas, Universidad de Los Andes, Chile

the model's stages (ESTRAUS 1989; Siegel et al. 2006). Overall, the four-stage modeling approach has the following structure: first, the total number of trips that began and ended in each zone are estimated separately. Next, in the second stage, the distribution of the trips is computed (origin-destination (OD) matrix). Then, in the third stage, the trips distributed in a pair of zones are assigned to the available modes serving that pair. Finally, the trips are split into different transport modes or combinations and are assigned to routes or strategies to fill up the networks representing the utilization of infrastructure by cars and public transport vehicles (Park et al. 2019).

Currently, aspects other than the traditional indicators associated with traffic congestion, such as sustainability, accessibility provision, people mobility, big data, and dynamism, should be considered when planning a proper and modern urban transport system (Frederix, Viti, and Tampère 2013; Pourebrahim et al. 2019). Although these new aspects could require different modeling paradigms, most strategic decisions regarding transport infrastructure and public transport system provisions in the world are made based on static four-stage models such as the one described above.

In most applications, the number of trips generated from and attracted to a zone (namely, o_i and d_j for zones i and j, respectively) are estimated through either multiple linear regression or growth factor models. Such estimations are based on the characteristics of households and residents in each zone (habitants per household, household structure, average income, number of cars, etc.) Additionally, they are based on land use features (industrial activities, commerce, educational institutions, regulation plans, and so on). Under the above modeling structure, obtaining the o_i and d_j parameters has an associated level of uncertainty inherent to both the estimation models and the functional form chosen to represent certain phenomena within the models. As discussed above, the vectors of generation and attraction (namely, o and d) from stage 1 are the required inputs for the trip distribution stage, where an OD matrix is constructed to account for the total number of trips performed at each OD pair over a modeling period.

Thus, for each pair of zones, i and j, the model output are the trips predicted to occur between zones i and j over the modeling period (t_{ij}). The distribution can be computed with different levels of aggregation or information (Ortuzar and Willumsen 2011; López-Ospina 2013; Thomas and Tutert 2013; Delgado and Bonnel 2016; Lim and Kim 2016; Chen, Liu, and Zhou 2019; Li et al. 2019; Shen 2019). The typical models used to compute this matrix are based on the maximization of entropy, which is a measure of the uncertainty in an urban system; the maximization problem is dependent on balancing the constraints associated with the level of information provided by vectors o_i and d_j, the generalized cost and the time constraints. These constraints can be analyzed at aggregate or disaggregate levels depending on the available information from the urban system and according to the modeling approach proposed (Cascetta, Pagliara, and Papola 2007). The idea of maximizing the entropy is to search for the most probable distribution of a variable at an aggregate level built from options obtained from more disaggregated levels.

When the vagueness or ambiguity of the origin o_i and destination d_j parameters is not considered, the base model with trip entropy maximizes trip dispersion and aims to maximize the information contained in the resulting matrix. A fuzzy characterization of the origin and destination parameters in the base problem is, in fact, an alternative method for considering the ambiguity and vagueness of these parameters; these considerations become relevant in cases where methodological differences in the design of surveys for estimating

OD patterns are observed, together with the existence of possible errors in data collection, omissions in data recording, and the impossibility of measuring the complete time lapses, among other aspects. Overall, this method recognizes that these parameters are not known with certainty. The base model without fuzzy OD parameters cannot handle this issue and assumes that parameters are known and deterministic.

It is evident from the previous arguments that vectors o_i and d_j are computed through models that deal with many sources of uncertainty. In addition, calibrating parameters for generation and attraction models has always been a difficult task due to the diversity of information from different sources that has to be collected for this purpose.

1.1. Contributions and outline of the paper

Given that OD matrix estimation is still a research challenge, since OD matrices are not fully observable (Hu et al. 2020; Ros-Roca, Montero, and Barceló 2021) and as a result of the arguments presented above, it is necessary to search for a modeling approach to properly handle sources of uncertainty in the context of solving stage 2 of the four-stage model to obtain a reliable distribution matrix. We address these uncertainty sources by defining o_i and d_j as input variables of a fuzzy set framework and allowing o_i and d_j to have varying membership degrees. This is different from classical sets in which the element memberships are assigned values of one. The objective of considering o_i and d_j as fuzzy parameters is that their membership functions are well defined in predefined ranges of uncertainty, and more importantly, we can use entropic membership functions to link this approach with the distribution problem to be optimized. It can be shown that under this modeling structure, the distribution stage is equivalent to solving a fuzzy optimization problem through a bi-objective formulation without losing relevant mathematical conditions such as the concavity of the objective function in the solution process.

In this paper, we develop a mathematical formulation for a doubly constrained transport distribution model based on the maximization of entropy, assuming that the parameters o_i, d_j and c (total system costs) are fuzzy parameters with entropic membership functions. As mentioned before, this allows flexibility in the model as well as in obtaining the theoretical and practical results. We take advantage of entropic membership functions, which is equivalent to transforming an entropic function to obtain the distribution of a Bernoulli random variable, in order to solve the problem in one stage unlike the classical models for fuzzy optimization. The use of fuzzy parameters allows the modeler to realize that he(she) does not know with certainty the OD parameters, and the further consideration of fuzzy entropy reinforces the information gain in this process; these membership functions could be applied to measure the subjective value of information under the condition of uncertainty related to the parameter of interest, which follows the concepts from fuzzy entropy theory (Aggarwal 2021). With our approach, the trips' matrix estimation is more robust, and the decision process behind the design of transportation systems becomes more illustrated. Moreover, the functional form of the entropic function can be understood as an approximation of the confidence interval associated with predictions. These arguments support our decision to model uncertainty through fuzzy sets.

Following the previous arguments and theoretical reasons for using entropic membership functions in the maximization of entropy for the doubly constrained trip distribution model, we identify the four major contributions of this research:

- The theoretical aspects of our modeling contributions are associated with the information gain when a fuzzy entropy approach is used; in fact, to the best of our knowledge, this aspect is not present in the literature, and in our opinion, this is a relevant issue to consider in the design process of transportation systems. Note, we are able to address the aspects behind the estimation of OD matrices with our approach. We are able to maximize the information gain, including randomness, vagueness, and ambiguity of the information on o_i and d_j parameters, in the determination of the trips. Considering only the entropy maximization on trips allows us to maximize the information gain but does not allow us to address the nature of the OD parameters. These values are typically calculated with estimation methods (Ortuzar, 2011) that use data gathered from transportation systems, which usually contain errors, consider different measured times and reflect situations where the total trips generated and attracted do not match ($\sum_i o_i \neq \sum_j d_j$). The base model cannot handle these issues. The output does not allow us to reflect them in the trip matrix, and it is not robust in capturing the uncertainty in the aforementioned parameters, which could result in incorrect or ineffective decisions for transportation system designs.

  The use of fuzzy parameters o_i and d_j allows us to realize that the modeler does not know with certainty the OD parameters, and the further consideration of fuzzy entropy reinforces the information gain in this process. With our approach, the trips' matrix estimation is more robust, and the decision process behind the design of transportation systems becomes more illustrated.
- The proposed formulation exclusively contains equality constraints. This allows us to directly extend the well-known gravitational algorithm as the core of our trip distribution model. In the case of trapezoidal and triangular specifications for the fuzzy parameters, the inequality constraints that appear when formulating the complementary slackness conditions make the extension of the gravitational algorithm quite difficult. We also highlight the fact that for the implementation of the gravitational algorithm, we combine fixed-point methods with root computations (proven convergence), which provides good support when nonlinear optimization software is not available. It also provides good support when the objective is to solve real-sized problems; in which case, the resulting nonlinear optimization problems could become unsolvable for most commercial packages.
- Unlike models that apply trapezoidal and triangular specifications, in this proposed model, we find analytical expressions for the balance factors, which contribute to the evaluation of projects by computing accessibility and attractiveness measures, which are very useful for performing planning and policy studies. By following this same line of reasoning, we are able to properly compute the dual variables in urban transportation problems; this allows us to conduct many sensitivity analyzes, which are omitted in the previous fuzzy entropic specifications.
- With respect to the numerical experiments, we show that statistically our model performs as good as previous trapezoidal specification models on the fuzzy parameters but with better flexibility in terms of potential sensitivity analyzes of different types as well as in terms of algorithmic potentialities when solving real-size problems. Moreover, we have developed a methodology to compare the adjustment of our model to other fuzzy specification models, namely, trapezoidal and triangular specification models, in terms of predicted structural changes in the distribution patterns embedded in each estimated

matrix. To conduct this comparison, we performed a χ^2 test between trapezoidal and entropic specifications with respect to the CRISP model, which shows that the entropic functions (for the different parameters considered in the bi-objective specification) have better prediction capabilities than the traditional fuzzy models when capturing changes associated with the structural patterns in a distribution; this analysis, inspired by the work of Black (2018) allows for a much better statistical analysis of the model than focusing only on aggregated indicators. In a particular example, trapezium, triangle and CRISP generated exactly the same OD matrix, although variability in o_i and d_j were observed. That is, such models were not able to capture such variability to generate different solutions along the Pareto frontier. In contrast, with our model, we were able to find different matrices to capture the dispersion and dynamics behind such parameters in a better way.

In the next section, we perform a comprehensive literature review covering several aspects of the research, from transport planning models to the application of fuzzy logic in similar contexts. Next, in Section 3, we justify and present the model and solution algorithm. Then, in Section 4, we show a numerical application of the model and finalize the paper with conclusions and further research insights.

2. Literature review

For the sake of clarity, we have decided to split this section into two subsections. The former briefly discusses the most traditional trip distribution models with the concept of entropy, and the latter subsection relates fuzzy optimization and entropy. This covers all the relevant aspects that contribute to our work.

2.1. Trip distribution and entropy models

The first approximations to distribution models were the growth models, where an expansion factor previously set for a specific period and population is applied. Later, distribution models were developed on aggregate and dissagregate data. Examples of dissagregate models are intervening opportunities models, where trip making is not explicitly related to distance but to the relative accessibility of opportunities for satisfying the objective of the trip (Ortuzar and Willumsen 2011). They are based on the initial work by Stouffer (1940), although Schneider (1959) presented the theory utilized at present.

In the case of considering dissagregate models, data are obtained directly from individual users, unlike aggregate models, where data come from geographical divisions supported by a transportation network (de Grange, Fernández, and de Cea 2010).

The classical model by Hitchcock (1941) provides an aggregate distribution, in which for certain goods, produced in several factories, are moved to their destinations at minimum cost. The resulting model is a linear programing formulation with constraints associated with constant costs. In the past, obtaining individual data to calibrate dissagregate models has been a difficult task; therefore, many authors have based their developments on models similar to the classical Hitchock formulation. Among the most important contributions, we highlight Wilson (2011), who proposed the concept of maximization of entropy to provide a theoretical foundation for the approximation of the gravitational model. The entropy is used as a measure of the dispersion of trips between origins and destinations (Samanta and Roy 2005).

Aiming to add microeconomic bases to the entropy model, Cochrane (1975) devised a sub-model of trip generation that incorporates a type of accessibility measurement. On the other hand, some authors have defended the convenience of nonlinear regression models of Poisson type, given their adaptability to different distribution scenarios (Flowerdew and Aitkin 1982; Winkelmann and Zimmermann 1995).

Following the concepts presented by Wilson, different studies have been developed regarding entropy maximization models. One of the outstanding studies was carried out by Fang and Tsao (1995), who presented a dual approach of unrestricted convex programing to provide a solution to the problem of entropy maximization for travel distribution, focusing on models with quadratic costs. In the same way, de Grange, Fernández, and de Cea (2010) analyzed and compared various spatially aggregated trip distribution models within a common theoretical framework for formulating and solving multiple target optimization problems, including quadratic costs associated with congestion phenomena.

An interesting study that captures the variability of parameters in entropy models is the one developed by Tsekeris and Stathopoulos (2006); they formulate a double dynamic model (DDGM) that takes into account the evolution of trip demand between periods and within a period (short term or within a day).

In da Silva and de Almeida (2013) a mathematical model is proposed to estimate the OD matrix integrating entropy models with multiple objective optimization techniques (goal programing) to ensure a balance between originated and attracted trips, thus avoiding the use of balance factors to adjust these parameters.

In Section 3.1, some recent applications of the maximum entropy model on trip estimation will be presented.

2.2. Fuzzy logic and entropy models

The estimation of parameters for modeling is quite complicated within the context of transport and urban systems due to the large number of factors influencing the parameters. In many cases, it is not possible to obtain sufficiently accurate and relevant data (Samanta and Roy 2005). To overcome these inaccuracies, fuzzy logic has been included as part of classical modeling. For example, Zimmermann (1978), and Bit, Biswal, and Alam (1992) presented a solution process to the multiobjective transport problem under uncertain environments modeled with fuzzy parameters. In Chanas and Kuchta (1996) they defined, programed and solved the distribution problem including fuzzy coefficients for costs. In Kalic and Teodorovic (1997) they improved the previous results with a mixed solution combining fuzzy logic and programing of a genetic algorithm.

In Samanta and Roy (2005), the authors analyze the problem of trip distribution with an objective function of entropy maximization based on Shannon's measurement. This measure acts as a trip dispersal factor between origins and destinations. In addition, they proposed making the supply and demand constraints more flexible through fuzzy optimization using trapezoidal numbers. The solution to this problem was obtained through parametric optimization.

In Ojha et al. (2009) they consider the solid multi-objective transport problem, STP, in which inaccuracies in cost units and in-route travel time are modeled as fuzzy variables.

Additionally, the availabilities in the supplier, the demands of the clients and the capacities of each mode are modeled as fuzzy variables, having as objective to compare the performance of the model that includes maximization of the entropy opposite to the one that does not include it. A year later, the same authors present an extension of this problem by including discounts on transport costs (Ojha et al. 2010). They considered discounts on all units and incremental discounts according to the quantity of purchase. To solve this new version of the STP, given the complexity associated, a genetic algorithm of the type 'roulette-wheel selection' was used.

Later, the STP was also used by Baidya, Bera, and Maiti (2014) under an entropy approach oriented at intervals, including budget constraints and product fragility. The interval theory applied to this model allows researchers to capture the uncertainty associated, for example, with the cost of transportation. The modeling of the problem has as objectives the reduction of the total cost and the increase of the profit, however, the authors achieve a solution as a mono-objective problem by applying a weighted sum method.

3. Fundamentals of the modeling approach

This section exposes the basic concepts of the entropy maximization problem for the trip distribution problem and defines the entropic membership function in detail. These concepts are required to develop the model proposed in this article.

3.1. Entropy maximization model

In this section, we present the entropy maximization formulation for the distribution stage following the traditional Wilson (1970) approach (Williams 2019). This model and its extensions are still being studied (in theoretical advances and applications). Among the most relevant recent work we mention the following: modeling with GPS data (You and Ritchie 2019), dynamic OD estimation using smart card data (Ait-Ali and Eliasson 2019), modeling to identify the source of foodborne disease outbreaks (Schlaich et al. 2020), modeling with mobile phone data (Demissie, Phithakkitnukoon, and Kattan 2018; Caceres, Romero, and Benitez 2020), modeling evacuation trip distributions during typhoons dynamically (Hu and Ho 2016), estimating OD matrices under travel demand constraints (Sun et al. 2019), modeling interregional transportation (Velichko 2016), modeling taxi trip distributions (Tang et al. 2018), input-output analysis (Hewings and Fernandez-Vazquez 2019), modeling highway traffic flows (Hu et al. 2020), etc.

Consider a set of m origins $i = \{1, 2, 3, \dots, m\}$ and n destinations $j = \{1, 2, 3, \dots, n\}$, where o_i is the number of outgoing trips from i, and d_j is the number of incoming trips to j. The doubly constrained entropy maximization model (in origins and destinations) (León, Figueroa-García, and López-Ospina 2016) for the distribution model is:

$$\max_{t_{ij}} \quad z = -\sum_{i=1}\sum_{j=1} t_{ij} \ln t_{ij} \tag{1}$$

s.t.

$$\sum_{j=1} t_{ij} = o_i, \quad \forall\, i \in m,\ (\tau_i) \tag{2}$$

$$\sum_{i=1} t_{ij} = d_j, \quad \forall j \in n, \ (\gamma_j) \tag{3}$$

$$\sum_{i=1} \sum_{j=1} c_{ij} t_{ij} = c. \ (\beta) \tag{4}$$

The function to be maximized represents the uncertainty degree associated with the function of the urban system. That is, the objective function reflects the behavior of the entropy, which must be maximized to find the most likely state. The constraints on the total number of generated and attracted trips are the typical constraints (same as Equations (2) and (3)). The parameters o_i and d_j are usually estimated through a multiple lineal regression or other models used in the generation/attraction stage of a transportation planning model. c_{ij} is the generalized cost of an OD pair $i–j$ and c is the total cost. Most applications consider the generalized travel cost of an OD pair $i–j$ (c_{ij}) as a linear function of the specific attributes (access time, fare, travel time, etc.) of the trip. Every attribute is weighted with an importance value, which is basically a subjective weight, assigned by the agent who performs the trip. τ_i and γ_j are Lagrange multipliers for the origin and destination constraints, respectively. β is the Lagrange multiplier for the total cost constraint.

The classical solution of this problem is given by

$$t_{ij} = \exp\left(\tau_i + \gamma_j + \beta c_{ij}\right) \tag{5}$$

A notation change yields:

$$t_{ij} = a_i * o_i * b_j * d_j * \exp\left(\beta c_{ij}\right) \tag{6}$$

where $\exp(\tau_i) = a_i * o_i$ and $\exp(\gamma_j) = b_j * d_j$. In addition to this

$$a_i = \frac{1}{\sum_k b_k d_k e^{c_{ik}\beta}} \tag{7}$$

$$b_j = \frac{1}{\sum_k a_k o_k e^{c_{kj}\beta}} \tag{8}$$

Parameters $a_i, b_j \in \mathbb{R}$ are known as balance factors associated with the number of trips generated by zone i, namely, (o_i), and the number of trips with destinations in zone j, where β is the sensitivity parameter for the generalized cost of travel ($\beta < 0$). a_i can be interpreted as a measure of the potential of accessibility to zone i, while b_j can be seen as an attraction factor in zone j (Caschili, De Montis, and Trogu 2015). As noted by Martínez (1995) and Martínez and Araya (2000), the concept of potential accessibility is related to the aggregate benefits of potential trips in the whole urban area.

In particular, measures of accessibility and attractiveness for each zone i (acc_i and att_j) have economic interpretations as monetary units per trip (see Martínez 1995; Martínez and Araya 2000).

$$acc_i = -\frac{1}{\beta} \ln a_i \tag{9}$$

$$att_i = -\frac{1}{\beta} \ln b_j \tag{10}$$

Therefore, a modeler can use balancing factors from the estimated spatial interaction travel model to calculate the access measures. The information embedded in these measures of benefits is based on the interactions associated with individual trips. These measures represent the expected benefits associated with the most likely distribution of trips for a given set of o_i's and d_j's.

Finally, Equations (7) and (8), which represent balance factors (a_i, b_j), define a fixed-point system of equations, where the iteration process starting with any initial values converges to the solution (see Algorithm 1 described below).

Algorithm 1: Entropy maximization

procedure MAXIMIZATION(*Entropy*)
  $\beta \leftarrow \beta_0$
  $\epsilon \leftarrow \epsilon$ ▷ This is a tolerance value
  $\triangle_\beta \leftarrow \beta^0$
  $o_i, \forall i$ ▷ Calculated before
  $d_j, \forall j$ ▷ Calculated before
  c_{ij}, C ▷ Calculated before
WHILE $\triangle_\beta \geq \epsilon$ **DO**:
  Calculate a_i^t and b_j^t using Equations (7) and (8). (A system of fixed-point equations)

$$a_i^t = \frac{1}{\sum_k b_k^t d_k e^{c_{ik}\beta^{t-1}}}$$
$$b_j^t = \frac{1}{\sum_k a_k^t o_k e^{c_{kj}\beta^{t-1}}}$$

  Calculate β^t using Equations (4) and (6). (A root-finding algorithm) That is, solve:

$$\sum_{i=1}\sum_{j=1} c_{ij} a_i^t b_j^t o_i d_j e^{c_{ij}\beta^t} - c = 0$$

  $\beta^t = \beta^{t-1}$
  $\triangle_\beta \leftarrow |\beta^t - \beta^{t-1}|$
  ENDWHILE
  $a_i \leftarrow a_i^t$
  $b_j \leftarrow b_j^t$
  $\beta \leftarrow \beta^t$

$$t_{ij} = a_i b_j o_i d_j e^{c_{ij}\beta}$$

  return $t_{ij};\ a_i;\ b_j; \beta$

3.2. *Fuzzy versions of the entropy maximization problem*

We describe the fuzzy versions of the entropy maximization problem in order to compare them with our model. We focus on the mathematical formulation and the parametric optimization methodologies used to solve the above problems.

In the transport entropy maximization problem, we consider the total cost and the origination and destination parameters to be fuzzy. We use the trapezoidal form of a fuzzy parameter because it can also handle fuzzy parameters in the form of a triangle or CRISP.

Definition 3.1: A fuzzy set, namely, A, is characterized by a membership function $\mu_A : X \to [0,1]$ defined over a universe of discourses $x \in X$. Thus, fuzzy set A is the set of ordered pairs $x \in X$ and its membership degree, $\mu_A(x)$, i.e.

$$A = \{(x, \mu_A(x)) \mid x \in X\} \tag{11}$$

Definition 3.2: Trapezoidal fuzzy number (TrFN): The TrFN can be represented by a quadruplet $\mathbf{x} = (\underline{x}, \mathbf{x_1}, \mathbf{x_2}, \overline{x})$, and its membership function has the following form:

$$\mu_{\mathbf{x}}(x) = \begin{cases} \dfrac{x - \underline{x}}{\mathbf{x_1} - \underline{x}}, & \text{for } x \in [\underline{x}, \mathbf{x_1}) \\ 1, & \text{for } x \in [\mathbf{x_1}, \mathbf{x_2}] \\ \dfrac{\overline{x} - x}{\overline{x} - \mathbf{x_2}}, & \text{for } x \in (\mathbf{x_2}, \overline{x}] \\ 0, & \text{for } x \notin \mathbf{x} \end{cases} \tag{12}$$

where $\underline{x} \le \mathbf{x_1} \le \mathbf{x_2} \le \overline{x}$.

With the trapezoidal representation, one can form parameters as real numbers, triangles, or intervals. Let us note that:

- If $\mathbf{x_1} = \mathbf{x_2}$, then we have a triangular fuzzy parameter (Kurtulmuşoğlu, Pakdil, and Atalay 2016).
- If $\underline{x} = \mathbf{x_1}$ and $\mathbf{x_2} = \overline{x}$, then we have a CRISP number in an interval.
- If $\underline{x} = \mathbf{x_1} = \mathbf{x_2} = \overline{x}$ then $\mathbf{x}$ is a real number.

Following the ideas of Samanta and Roy (2005), we show the fuzzy version of the entropy maximization problem with trapezoidal parameters $o_i = (\underline{o_i}, \mathbf{o}_{i1}, \mathbf{o}_{i2}, \overline{o}_i)$, $d_j = (\underline{d_j}, \mathbf{d_{j1}}, \mathbf{d_{j2}}, \overline{d}_j)$ and $c = (\underline{C}, \mathbf{C_1}, \mathbf{C_2}, \overline{C})$.

$$\max_{t_{ij}} \quad z = -\sum_{i=1} \sum_{j=1} t_{ij} \ln t_{ij} \tag{13}$$

s.t.

$$\sum_{j=1} t_{ij} = (\underline{o_i}, \mathbf{o_{i1}}, \mathbf{o_{i2}}, \overline{o_i}), \quad \forall\, i \in m \tag{14}$$

$$\sum_{i=1} t_{ij} = (\underline{d_j}, \mathbf{d_{j1}}, \mathbf{d_{j2}}, \overline{d_j}), \quad \forall\, j \in n \tag{15}$$

$$\sum_{i=1} \sum_{j=1} c_{ij} t_{ij} = (\underline{C}, \mathbf{C_1}, \mathbf{C_2}, \bar{C}) \tag{16}$$

The model defined by Equations (13)–(16) is equivalent to the following bi-objective problem (Equations (17)–(23)).

$$\max_{t_{ij},\lambda} \quad \left(-\sum_{i=1}\sum_{j=1} t_{ij} \ln t_{ij}, \lambda \right) \tag{17}$$

s.t.

$$\sum_{j=1} t_{ij} \leq \lambda \mathbf{o_{i2}} + (1-\lambda)\overline{o_i}, \quad \forall\, i \in m \tag{18}$$

$$\sum_{j=1} t_{ij} \geq \lambda \mathbf{o_{i1}} + (1-\lambda)\underline{o_i}, \quad \forall\, i \in m \tag{19}$$

$$\sum_{i=1} t_{ij} \leq \lambda \mathbf{d_{j2}} + (1-\lambda)\overline{d_j}, \quad \forall\, j \in n \tag{20}$$

$$\sum_{i=1} t_{ij} \geq \lambda \mathbf{d_{j1}} + (1-\lambda)\underline{d_j}, \quad \forall\, j \in n \tag{21}$$

$$\sum_{i=1}\sum_{j=1} c_{ij}t_{ij} \leq \lambda \mathbf{C_2} + (1-\lambda)\bar{C} \tag{22}$$

$$\sum_{i=1}\sum_{j=1} c_{ij}t_{ij} \geq \lambda \mathbf{C_1} + (1-\lambda)\underline{C}$$

$$t_{ij} \geq \ 0, \quad \forall\, i, j,\ \lambda \in [0, 1] \tag{23}$$

In Equations (17)–(23), the two objectives are to maximize the entropy and λ, where λ is the minimum membership level of the origin, destination and cost constraints. Compared to a non-fuzzy model of entropy maximization, the number of constraints doubles. In addition, to the best of our knowledge, there are no accessible mathematical expressions for calculating the system, and there is no method extending the gravitational version of the non-fuzzy (classic) entropy maximization problem; this could be due to the inherent difficulties in interpreting the complementary slackness conditions for the constraints.

In particular, to solve this problem, we consider the model defined by Equations (24)–(31), in which there is a lower bound for λ, which represents the minimum membership E. According to the ϵ-constraint method, for different values of $E \in [0, 1]$, we solve this bi-objective model. Note that the trip entropy function and λ are both concave functions. In addition, the optimization problem only contains linear constraints. These two reasons guarantee that the ϵ-constraint method allows us to obtain the Pareto frontier of the previous bi-objective optimization problem (Pereyra, Saunders, and Castillo 2013; Bonnel and Collonge 2015).

$$\max_{t_{ij},\lambda} \quad -\sum_{i=1}\sum_{j=1} t_{ij} \ln t_{ij} \tag{24}$$

s.t.

$$\lambda \geq E \tag{25}$$

$$\sum_{j=1} t_{ij} \leq \lambda \mathbf{o_{i2}} + (1-\lambda)\overline{o_i}, \quad \forall\, i \in m \tag{26}$$

$$\sum_{j=1} t_{ij} \geq \lambda \mathbf{o_{i1}} + (1-\lambda)\underline{o_i}, \quad \forall\, i \in m \tag{27}$$

$$\sum_{i=1} t_{ij} \leq \lambda \mathbf{d_{j2}} + (1-\lambda)\overline{d_j}, \quad \forall\, j \in n \tag{28}$$

$$\sum_{i=1} t_{ij} \geq \lambda \mathbf{d_{j1}} + (1-\lambda)\underline{d_j}, \quad \forall\, j \in n \tag{29}$$

$$\sum_{i=1}\sum_{j=1} c_{ij}t_{ij} \leq \lambda \mathbf{C_2} + (1-\lambda)\bar{C} \tag{30}$$

$$\sum_{i=1}\sum_{j=1} c_{ij}t_{ij} \geq \lambda \mathbf{C_1} + (1-\lambda)\underline{C}$$

$$t_{ij} \geq\ 0, \quad \forall\, i,j\,\lambda \in [0,1] \tag{31}$$

One of the shortcomings of the previous models is that no analyzes have been performed on the accessibility measures or marginal effects, such as the marginal utility of travel costs. Furthermore, no extensions of the gravitational algorithm have been proposed or discussed.

3.3. Entropic membership function

In this section, we define the entropic membership function based on the concept of entropy associated with a fuzzy set. The connection between fuzzy set theory and the entropy concept is in the computations of the fuzzy information measurements.

Entropy can be understood as a degree of randomness and has been used to measure the fuzziness in a fuzzy system or set. Fuzzy entropy deals with vagueness and ambiguous uncertainties, while traditional entropy formulations address probabilistic uncertainties. Measures of 'fuzziness' can be interpreted as measures of the amount of fuzzy information (Bellman and Zadeh 1970). The fuzzy entropy value of the set A is as follows:

$$H_A = -k\sum_{i=1}^{n}(\mu_i \log(\mu_i) + (1-\mu_i)\log(1-\mu_i)) \tag{32}$$

where μ_i is a membership function and $k \in \mathbb{R}$ is constant.

Analogous to the entropy value of a Bernoulli distribution, which is defined over the interval $[0, 1]$, it is possible to define a membership function for the entropy values of a variable X enclosed in its closure/support $\boldsymbol{x}$, as follows.

Definition 3.3: Let $\mathbf{x} = \{x \in \mathbb{R} : \underline{x} \leqslant x \leqslant \overline{x}\} = [\underline{x}, \overline{x}]$ be an interval set over the reals; then, the log-natural function $f_{\mathbf{x}}$ of $\mathbf{x}$ is defined as follows:

$$f_{\mathbf{x}}(x) = \begin{cases} -q\ln(q) - p\ln(p), & \text{for } x \in \mathbf{x} \\ 0, & \text{for } x \notin \mathbf{x} \end{cases} \tag{33}$$

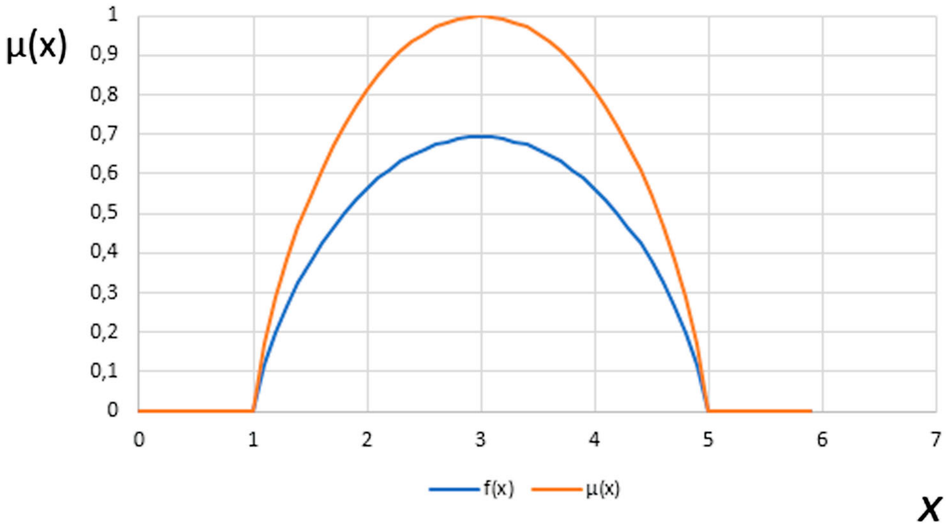

Figure 1. Membership function $\mu_{\mathbf{x}}(x)$.

where

$$q = \frac{\bar{x} - x}{\bar{x} - \underline{x}} \tag{34}$$

$$p = \frac{x - \underline{x}}{\bar{x} - \underline{x}} \tag{35}$$

$$p + q = 1 \tag{36}$$

Roughly speaking, $f_{\mathbf{x}}$ is not a normalized function. To normalize it, we define the entropic fuzzy set as the normalized log-natural function of the interval **x**, as follows:

Definition 3.4: Let $f_{\mathbf{x}}$ be the log-natural function of **x**. The entropic fuzzy set, namely, $\mu_{\mathbf{x}}$ of **x**, is as follows:

$$\mu_{\mathbf{x}}(x) = \frac{f_{\mathbf{x}}(x)}{\max_{x \in \mathbf{x}} f_{\mathbf{x}}(x)}. \tag{37}$$

As mentioned above, here, we can visualize that the fuzzy entropic membership function could be applied to measure the subjective value of information under a condition of uncertainty related to the parameter of interest.

A graphical display of $\mu_{\mathbf{x}}$ and $f_{\mathbf{x}}(x)$ with $\underline{x} = 1$ and $\bar{x} = 5$ is shown in Figure 1.

4. Problem statement and model formulation

The objective of the present work is to incorporate the effects of the uncertainty inherent in the generation and attraction of trips at each zone within a studied urban area in the distribution through the concepts of fuzzy logic and optimization by means of entropic membership functions (see Equation (37)). Such functions allow us to resolve the resulting fuzzy problem through a bi-objective formulation without losing important mathematical conditions, such as the concavity and derivability of the objective function, in the process of searching for a solution. Unlike the trapezoidal model, the constraints are of equality, which allow us to obtain closed expressions for the measures of accessibility and attractiveness, t_{ij} and the balance factors.

A fundamental assumption of the proposed model in this article is that o_i and d_j are modeled as fuzzy parameters, in which the membership functions are defined positively over intervals of the form $\mathbf{o}_i = [\underline{o}_i; \overline{o}_i]$ and $\mathbf{d}_j = [\underline{d}_j; \overline{d}_j]$. The extreme values for the interval ranges have to be provided, either by the expertize of the modeler or from an analysis of the uncertainty in the computation of the parameters during the process of the generation and attraction of trips, for example, through multiple regression models.

The next step is to assume that the membership function is entropic, as stated in Equation (37). More precisely, we use an entropy function transformation to obtain the distribution of a Bernoulli random variable, which allows us to solve the entire problem in just one stage, unlike many classical fuzzy optimization models for the distribution of urban trips (Samanta and Roy 2005; Ojha et al. 2009). Moreover, the functional form of entropy can be understood as an approximation of the confidence level of the predictions.

In addition, as mentioned in Section 3.2, some previous studies used triangular and trapezoidal membership functions in the same context as the gravitational model (Equations (24)–(31)). The entropic membership function can be interpreted as an approximate (derivable and concave) function of these types of formulations (triangular and trapezoidal membership).

Our hypothesis is that it is possible to include vagueness and ambiguity through fuzzy numbers, but the consideration of information gain can only be provided by a fuzzy-entropy approach. Our proposal is useful for performing this task because it applies fuzzy entropy on parameters and performs entropy maximization over trips.

In summary, the main contributions of considering fuzzy entropic parameters in an entropy maximization model can be clarified by stating the following:

(1) The inclusion of fuzzy entropy on OD parameters allows us to include aspects that the base model cannot handle, namely, assuring the information gain while considering the vagueness and ambiguity of the parameters.
(2) The use of fuzzy parameters allows a modeler to realize that OD parameters are known with uncertainty, and the consideration of fuzzy entropy reinforces the information gain in this process.

To include uncertainty over o_i and d_j, we define the following bi-objective model:

$$g^1 := \max_{t_{ij}} - \sum_{i=1}^{K} \sum_{j=1}^{L} t_{ij} \ln(t_{ij}) \tag{38}$$

$$g^2 := \max_{\alpha_j, \rho_i, \eta} \{ -\alpha_j \ln(\alpha_j) - (1-\alpha_j) \ln(1-\alpha_j), \quad \forall j \in n; \tag{39}$$

$$- \rho_i \ln(\rho_i) - (1-\rho_i) \ln(1-\rho_i), \ \forall i \in m; \tag{40}$$

$$- \eta \ln(\eta) - (1-\eta) \ln(1-\eta) \} \tag{41}$$

s.t.

$$\sum_{j=1}^{L} t_{ij} = \rho_i \underline{o}_i + (1-\rho_i)\bar{o}_i, \quad \forall i \in m \tag{42}$$

$$\sum_{i=1}^{K} t_{ij} = \alpha_j \underline{d}_j + (1-\alpha_j)\bar{d}_j, \quad \forall j \in n \tag{43}$$

$$\sum_{i=1}^{k} \sum_{j=1}^{L} c_{ij} t_{ij} = \eta \underline{C} + (1-\eta)\bar{C}, \tag{44}$$

$$t_{ij} \geq 0, \quad \rho_i,\ \alpha_j,\ \eta \in [\varepsilon; 1-\varepsilon], \tag{45}$$

where ε in Equation (45) is an extremely small positive number (to avoid ln(0)). Functions of the second goal are based on the entropy value of a Bernoulli random variable, and the maximum value is reached at the center of the respective interval and computed independently for α_j, ρ_i and η. That is, if $0.5 = \alpha_j = \rho_i = \eta$ (medium point of each interval: $\frac{1}{2}(\underline{d}_j + \overline{d}_j)$, $\frac{1}{2}(\underline{o}_j + \overline{o}_j)$ and $\frac{1}{2}(\underline{C} + \overline{C})$, respectively), then functions $-\alpha_j \ln(\alpha_j) - (1-\alpha_j)\ln(1-\alpha_j)$, $-\rho_i \ln(\rho_i) - (1-\rho_i)\ln(1-\rho_i)$, and $\eta\underline{C} + (1-\eta)\overline{C}$, reach their maximum value. However, in the integrated model, as described in Equations (38)–(45), we are not necessarily obtaining those maximum values in all the studied cases.

On the other hand, outside each interval $\mathbf{o}_i = [\underline{o}_i, \overline{o}_i]$, $\mathbf{d}_j = [\underline{d}_j, \overline{d}_j]$, and $[\underline{C}, \overline{C}]$, the membership degree is zero.

The physical meanings of Equations (42) and (43) are analogous. In the case of Equation (42), for each zone i, all the generated trips at that zone fall in the range $(\underline{o}_i, \overline{o}_i)$. If $\rho_i \to 1$, then the number of trips will approach $\underline{o}_i$. On the other hand, if $\rho_i \to 0$, then the number of trips will approach the lower limit $\overline{o}_i$. Finally, if $\rho_i \to \frac{1}{2}$, the number of generated trips should approach $\frac{1}{2}(\underline{o}_i + \overline{o}_i)$.

In the case of Equation (43), we refer to the trips attracted to each zone j, that is, all the attracted trips falling within $\underline{d}_j, \overline{d}_j)$. In particular, if $\alpha_j \to 0$, then the number of trips attracted in zone j will approach $\overline{d}_j$. On the other hand, if $\alpha_j \to 1$, then the number of trips will approach the lower limit $\underline{d}_j$. Finally, if $\alpha_i \to \frac{1}{2}$, the number of attracted trips should approach $\frac{1}{2}(\underline{d}_j + \overline{d}_j)$.

In the case of Equation (44), this expression guarantees that the total costs of the system fall between $[\underline{c}, \overline{c}]$. Analogous to Equations (42) and (43), the value of such costs depends on the value of η.

In this article, we propose a solution method based on a compromise programing method for the bi-objective formulation by considering a linear combination of two weighted objectives described below:

(1) Goal 1, i.e. g^1: The maximization of total system entropy weighted by a parameter θ_1.
(2) Goal 2, i.e. g^2: The maximization of the sum of the membership functions for each generation/attraction of trips per zone and the membership total cost function. This goal is weighted by a parameter θ_2.

Parameters θ_1 and θ_2 have a sum of 1 (since the objectives are normalized). To compare our model with models (24)–(31), we note that $\theta_2 = \theta$ and $\theta_1 = 1 - \theta$. That is,

θ will represent the valuation of the aggregate membership of the fuzzy parameters. Additionally, as we increase the value of θ, we expect the membership levels of the different parameters to be close to 1. Thus, such a parameter plays an analogous role to parameter E in Equation (25). This allows us to make additional comparisons of the numerical results in Section 4.

Note that we could have disaggregated values in more areas than those proposed here. For example, one for the generation of trips, another for the attraction of trips and a last one for the costs. However, this apparent simplification allows us to analyze with clarity the impact of uncertainty of such parameters in the estimation of trip distribution. Moreover, such extensions or modeling disaggregation do not add value to the conceptual phenomenon addressed in this research and therefore are left for further developments (as discussed in the conclusion of the paper) to allow for a clear presentation of the phenomenon of uncertainty introduced through this entropic approach with fuzzy parameters.

Thus, the bi-objective model can be aggregated as follows:

$$\begin{aligned}\max_{t_{ij},\alpha_j,\rho_i} \quad & (1-\theta) * \frac{-\sum_{i=1}^{K}\sum_{j=1}^{L} t_{ij}\ln(t_{ij}) - \underline{g}^1}{\bar{g}^1 - \underline{g}^1} \\ & + \theta \frac{1}{\bar{g}^2 - \underline{g}^2} * \Bigg[-\sum_j [\alpha_j \ln(\alpha_j) + (1-\alpha_j)\ln(1-\alpha_j)] \\ & - \sum_i [\rho_i \ln(\rho_i) + (1-\rho_i)\ln(1-\rho_i)] \\ & + [-\eta \ln(\eta) - (1-\eta)\ln(1-\eta)] - \underline{g}^2 \Bigg]\end{aligned} \tag{46}$$

s.t.

$$\sum_{j=1}^{L} t_{ij} = \rho_i \underline{o}_i + (1-\rho_i)\bar{o}_i, \quad \forall\, i \in m \tag{47}$$

$$\sum_{i=1}^{K} t_{ij} = \alpha_j \underline{d}_j + (1-\alpha_j)\bar{d}_j, \quad \forall\, j \in n \tag{48}$$

$$\sum_{i=1}^{k}\sum_{j=1}^{L} c_{ij} t_{ij} = \eta \underline{C} + (1-\eta)\bar{C}, \tag{49}$$

$$t_{ij} \geqslant 0,\ \rho_i,\ \alpha_j,\ \eta \in [\varepsilon; 1-\varepsilon] \tag{50}$$

where

- $\bar{g}^1$ is the solution of the problem of maximizing the entropy without considering the membership objective associated with the functions of generation and attraction or trips (namely, $\theta = 0$).

- $\overline{g}^2$ is the solution of the maximization of the sum of the membership functions for the total attraction and generation estimates and total costs, without considering the total system entropy objective ($\theta = 1$).
- $\underline{g}^1$ is the value of the trips' entropy in the optimal solution of g^2.
- $\underline{g}^2$ is the value of the sum of the membership functions for the total attraction and generation estimates and total costs in the optimal solution of g^1.
- ε in Equation (50) is an extremely small positive number (to avoid $\ln(0)$).

Note that the trip entropy functions and the sum of the membership entropic functions for the total attraction and generation estimates and total costs are strictly concave functions. In addition, the optimization problem only contains linear constraints. These two reasons guarantee that the variation in parameter θ between 0 and 1 allows us to obtain the Pareto frontier of our bi-objective optimization problem (Pereyra, Saunders, and Castillo 2013; Bonnel and Collonge 2015). In Appendix, we explain in more detail some theoretical aspects about this Pareto frontier.

This formulation allows us to generate a sensitivity analysis for the results of t_{ij} for different values of the parameter weights in the objectives of the bi-objective formulation. In the literature, we found some previous studies that used triangular and trapezoidal membership functions in the same context as the functions used in this research (Kalić and Teodorović 2003; Samanta and Roy 2005). In our work, the proposed membership functions are more generic, as they can be interpreted as an approximation that could be recovered from these types of functions. Moreover, the problem can be solved optimally with any classical convex nonlinear optimization software.

In addition, through this formulation, we are able to analyze the variability of the classical results of the gravitational trip distribution models used typically in transport economy studies, such as studies on accessibility, attractiveness and valuation of generalized travel costs.

Another very important feature of the proposed model is related to a basic conservation assumption behind entropy models for the estimation of transportation demand under deterministic scenarios. That is, for deterministic values of o_i and d_j, this assumption is stated as:

$$\sum_i o_i = \sum_j d_j \tag{51}$$

However, models used in practice to estimate both o_i and d_j differ in input information, and as a consequence, it is common that this equality does not hold in practical applications. To correct this, expansion factors are normally used. This practice can of course generate errors in the estimations, which can cause potential mistakes in transportation policies and predictions.

The proposed model does not need to use expansion factors, and we only require relaxing Equation (51) to find at least one feasible solution, as follows:

$$\left[\sum_i \underline{o}_i; \sum_i \overline{o}_i\right] \cap \left[\sum_j \underline{d}_j; \sum_j \overline{d}_j\right] \neq \emptyset \tag{52}$$

This assumption relaxes the condition inherent to its deterministic counterpart (see Equation (1)), allowing us to use an initial estimation of the attraction and generation of trips. In this way, with the Lagrange first-order conditions for this problem, we obtain the following results:

$$t_{ij} = \exp\left(g_0^1(\tau_i + \gamma_j + \beta c_{ij})\right) \tag{53}$$

$$\rho_i = \left(1 + \exp\left(\frac{\tau_i}{\theta}(\bar{g}^2 - \underline{g}^2)(\bar{o}_i - \underline{o}_i)\right)\right)^{-1} \tag{54}$$

$$\alpha_j = \left(1 + \exp\left(\frac{\gamma_j}{\theta}(\bar{g}^2 - \underline{g}^2)(\bar{d}_j - \underline{d}_j)\right)\right)^{-1} \tag{55}$$

$$\eta = \left(1 + \exp\left(\frac{\beta}{\theta}(\bar{g}^2 - \underline{g}^2)(\bar{C} - \underline{C})\right)\right)^{-1} \tag{56}$$

$$g_0^1 = \frac{\bar{g}^1 - \underline{g}^1}{1 - \theta} \tag{57}$$

where τ_i, γ_j, β are Lagrange multipliers and g_0^1 is the initial condition.

As mentioned before, $\exp(\tau_i)$ and $\exp(\gamma_j)$ are associated with measures of accessibility and attractiveness for each zone (Martínez 1995). Moreover, β can be interpreted as the marginal valuation of the generalized cost of individuals and agents. Note also that ρ_i, α_j, t_{ij}, η depend on the length of the interval associated with each parameter o_i, d_j and C, as well as on the weight assigned to each objective.

By following Ortuzar and Willumsen (2011), Equation (53) becomes

$$t_{ij} = \exp\left(g_0^1\right) * a_i * (\rho_i \underline{o}_i + (1 - \rho_i)\bar{o}_i) b_j * (\alpha_j \underline{d}_j + (1 - \alpha_j)\overline{d}_j) \exp\left(g_0^1 \beta c_{ij}\right) \tag{58}$$

where $\exp\left(g_0^1 \tau_i\right) = a_i * (\rho_i \underline{o}_i + (1 - \rho_i)\bar{o}_i)$ and $\exp\left(g_0^1 \gamma_j\right) = b_j * (\alpha_j \underline{d}_j + (1 - \alpha_j)\overline{d}_j)$. Thus, the balance factors a_i and b_j can be obtained as follows:

$$a_i = \left(\exp\left(g_0^1\right) \sum_j b_j * (\alpha_j \underline{d}_j + (1 - \alpha_j)\bar{d}_j) \exp\left(g_0^1 * \beta c_{ij}\right)\right)^{-1} \tag{59}$$

$$b_j = \left(\exp\left(g_0^1\right) \sum_i a_i * (\rho_i \underline{o}_i + (1 - \rho_i)\bar{o}_i) \exp\left(g_0^1 * \beta c_{ij}\right)\right)^{-1} \tag{60}$$

a_i can be interpreted as a measure of the potential of accessibility to zone i, while b_j can be seen as an attraction factor in zone j. In both cases, these measures depend on the membership level in each parameter; this shows the potential impact of our formulation with fuzzy parameters on project valuation subjects.

Equations (53)–(60) can be solved by integrating the constraints of the optimization problem (Equations (47)–(49)) using nonlinear fixed-point methods and root finding numerical methods.

Solution algorithm

In this section, we describe the algorithm used to solve the maximization problem proposed. First, the sensitivity to the cost value is initialized to proceed with the optimization procedure through iterations with fixed-point equations.

We start with a seed value for β^0, along with all the additional required parameters, such as the cost matrix and tolerance values.

At iteration t, using β^{t-1}, we can compute the values of $\tau_i{}^t$ and $\gamma_j{}^t$ for all i, j. Then, we replace the optimal values of ρ_i and α_j from Equations (54) and (55) in Equations (47) and (48) to obtain a system of fixed-point equations for the next iteration:

$$\tau_i^t = \frac{\theta}{(\bar{o}_i - \underline{o}_i)(\bar{g}^2 - \underline{g}^2)} \ln\left(\frac{\underline{o}_i - \bar{o}_i}{\sum_{j=1}^n \exp(g_0^1(\tau_i^t + \gamma_j^t + \beta^{t-1}c_{ij})) - \bar{o}_i} - 1\right), \quad \forall\, i \in m \quad (61)$$

$$\gamma_j^t = \frac{\theta}{(\bar{d}_j - \underline{d}_j)(\bar{g}^2 - \underline{g}^2)} \ln\left(\frac{\underline{d}_j - \bar{d}_j}{\sum_{i=1}^n \exp(g_0^1(\tau_i^t + \gamma_j^t + \beta^{t-1}c_{ij})) - \bar{d}_j} - 1\right), \quad \forall\, j \in n \quad (62)$$

With the values of the multipliers from Equations (61) and (62), we can compute t_{ij} and η for the t^th iteration, which are denoted as: t_{ij}^t, η^t.

$$t_{ij}^t = \exp((g_0^1) * (\tau_i^t + \gamma_j^t + \beta^{t-1} * c_{ij})), \quad \forall\, i \in m,\, j \in n \quad (63)$$

$$\eta^t = \left(1 + \exp\left(\frac{\beta^{t-1}}{\theta}(\overline{g}^2 - \underline{g}^2)(\overline{C} - \underline{C})\right)\right)^{-1} \quad (64)$$

Finally, β^t is computed with the following nonlinear equation (the root-finding algorithm).

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} \exp\left(g_0^1 * \left(\tau_i^t + \gamma_j^t + \beta^t * c_{ij}\right)\right) = \eta^t \underline{C} + (1 - \eta^t)\overline{C} \quad (65)$$

The algorithm stops when the following condition is satisfied: $|\beta^{t-1} - \beta^t| \leq epsilon$

A pseudocode for the solution algorithm described here is presented next in Algorithm 2.

Note that with the algorithm results, we can obtain the values of Goal 1 (g^1) and Goal 2 (g^2) and the balance factors (Equations (59) and (60)) for each variation of weight θ.

As mentioned in the Introduction section, this algorithm is useful for modeling OD matrices of large sizes, such as those found in realistic problems, in which case, most nonlinear optimization packages are unable to find a solution for these problems.

5. Numerical examples

We present two numerical examples, a small example with synthetic data with five origins and destinations and a second example with one hundred origins and destinations that represents the city of Concepción in Chile. In the first 5×5 case, we expose the attractiveness and accessibility indicators that are obtained with our model. In addition, we compare the results of our methodology against the triangular, trapezoidal, and CRISP membership functions.

5.1. First example: the synthetic scenario

To apply the model described above, we first develop the methodology for a synthetic example consisting of 5 origin and 5 destination zones. In addition, for each zone, the number of trips attracted and generated are not known with certainty. Furthermore, the total

Algorithm 2: Entropy maximization with fuzzy entropic parameters

procedure MAXIMIZATION(*Entropy*, *Membershipfunction*)
θ
$\beta \leftarrow \beta_0$
$\epsilon \leftarrow \epsilon$ ▷ This is a tolerance value
$\triangle_\beta \leftarrow \beta_0$
$\underline{g}^2 \leftarrow \underline{g}^2$ ▷ Calculated before
$\overline{g}^2 \leftarrow \overline{g}^2$ ▷ Calculated before
$\underline{g}^1 \leftarrow \underline{g}^1$ ▷ Calculated before
$\overline{g}^1 \leftarrow \overline{g}^1$ ▷ Calculated before
WHILE $\triangle_\beta \geq \epsilon$ **DO**:
Calculate τ_i^t and γ_j^t with Equations (61) and (62). (a system of fixed-point equations)
Calculate η^t with Equation (64)
Calculate β^t using Equation (65). (a root-finding algorithm)
$\triangle_\beta \leftarrow |\beta^t - \beta^{t-1}|$ **ENDWHILE**
$\tau_i \leftarrow \tau_i^t$
$\gamma_j \leftarrow \gamma_j^t$
$\beta \leftarrow \beta^t$
$\eta \leftarrow \eta^t$
$t_{ij} \leftarrow exp\left(g_0^1 * \left(\tau_i^t + \gamma_j^t + \beta^{t-1} * c_{ij}\right)\right),\ \forall\, i \in m, j \in n$
$\rho_i \leftarrow \left(1 + \exp\left(\frac{\tau_i}{\theta}(\overline{g}^2 - \underline{g}^2)(\overline{o}_i - \underline{o}_i)\right)\right)^{-1}$
$\alpha_j \leftarrow \left(1 + \exp\left(\frac{\gamma_j}{\theta}(\overline{g}^2 - \underline{g}^2)(\overline{d}_j - \underline{d}_j)\right)\right)^{-1}$
$o_i \leftarrow \rho_i \underline{o}_i + (1 - \rho_i)\overline{o}_i$
$d_j \leftarrow \alpha_j \underline{d}_j + (1 - \alpha_j)\overline{d}_j$
$g^1 : \frac{-\sum_{i=1}^{K}\sum_{j=1}^{L} t_{ij}\ln(t_{ij}) - \underline{g}^1}{\overline{g}^1 - \underline{g}^1}$
$g^2 : \frac{-\sum_j[\alpha_j \ln(\alpha_j) + (1-\alpha_j)\ln(1-\alpha_j)] - \sum_i[\rho_i \ln(\rho_i) + (1-\rho_i)\ln(1-\rho_i) - \eta\ln(\eta) - (1-\eta)ln(1-\eta)] - \underline{g}^2}{\overline{g}^2 - \underline{g}^2}$
return $t_{ij}; \eta; \tau_j; \gamma_j; \beta; \rho_i; \alpha_j; o_i; d_j; g^1; g^2$

cost of the system (C) is also not fully determined. Thus, we assume that the actual number of effective generated and attracted trips per zone during the whole modeling period are contained in a box of uncertainty bounded by a lower limit and an upper limit that define the range of potential values for the demand parameters. Similarly, there are bounds (lower and upper) for the total cost of the system. Note that, in practical terms, the actual average values of the numbers of trips are not known, although they can be estimated according to some relevant confidence level, which allows us to define parameters such as fuzzy numbers. For this small application case, ranges for the numbers of trips generated and attracted by the 5 origin and destination zones are listed in the following Table 1.

At this point, we must verify that the model is able to find at least one feasible solution so that the total number of trips for each range of values can be computed. The sum of the limits of the different uncertainty boxes for the generated trips is equal to $T_{prod} = [855, 1196]$,

Table 1. Bounds for the generated and attracted trips.

Generated and attracted trips					
Generated trips			Attracted trips		
ZO	Lower limit $\underline{o}_i$	Upper limit $\overline{o}_i$	ZD	Lower limit $\underline{d}_j$	Upper limit $\overline{d}_j$
O1	175	243	**D1**	183	249
O2	169	236	**D2**	178	251
O3	171	239	**D3**	180	244
O4	166	238	**D4**	184	245
O5	174	240	**D5**	185	252

Table 2. Distances between each pair of zones.

Zones	**D1**	**D2**	**D3**	**D4**	**D5**
O1	23	20	19	31	25
O2	12	16	20	10	13
O3	11	13	17	19	13
O4	17	20	28	15	21
O5	10	12	9	15	17

while the sum for the attracted trips is $T_{Atra} = [910, 1241]$. Note that the intersection of the intervals is not empty. In addition, $\overline{C} = 78$ and $\underline{C} = 0.1$.

On the other hand, the distances for each i–j pair are presented in the following Table 2.

The limits of the entropy function are $[\underline{g}^1, \overline{g}^1] \approx [-4207.2; -3841.5]$, while the limits of the sum of the membership functions are $[\underline{g}^2, \overline{g}^2] \approx [6.412; 7.294]$.

In this way, the optimization problem for this synthetic example is written as follows:

$$\max_{t_{ij},\alpha_j,\rho_i} = (1-\theta)\ \frac{-\sum_{i=1}^{5}\sum_{j=1}^{5} t_{ij}\ln(t_{ij}) + 4207.2}{-3841.5 + 4207.2}$$

$$\frac{\theta}{7.294 - 6.412} * \left[-\sum_{j=1}^{5}[\alpha_j \ln(\alpha_j) + (1-\alpha_j)\ln(1-\alpha_j)] \right.$$

$$\left. -\sum_{i=1}^{5}[\rho_i \ln(\rho_i) + (1-\rho_i)\ln(1-\rho_i)] - [\eta \ln(\eta) + (1-\eta)\ln(1-\eta)] - 6.412 \right]$$

s.t.

$$t_{1,1} + t_{1,2} + t_{1,3} + t_{1,4} + t_{1,5} = \rho_1 175 + (1-\rho_1)243$$

$$t_{2,1} + t_{2,2} + t_{2,3} + t_{2,4} + t_{2,5} = \rho_2 169 + (1-\rho_2)236$$

$$t_{3,1} + t_{3,2} + t_{3,3} + t_{3,4} + t_{3,5} = \rho_3 171 + (1-\rho_3)239$$

$$t_{4,1} + t_{4,2} + t_{4,3} + t_{4,4} + t_{4,5} = \rho_4 166 + (1-\rho_4)238$$

$$t_{5,1} + t_{5,2} + t_{5,3} + t_{5,4} + t_{5,5} = \rho_5 174 + (1-\rho_5)240$$

$$t_{1,1} + t_{2,1} + t_{3,1} + t_{4,1} + t_{5,1} = \alpha_1 183 + (1-\alpha_1)249$$

$$t_{1,2} + t_{2,2} + t_{3,2} + t_{4,2} + t_{5,2} = \alpha_2 178 + (1 - \alpha_2)251$$
$$t_{1,3} + t_{2,3} + t_{3,3} + t_{4,3} + t_{5,3} = \alpha_3 180 + (1 - \alpha_3)244$$
$$t_{1,4} + t_{2,4} + t_{3,4} + t_{4,4} + t_{5,4} = \alpha_4 184 + (1 - \alpha_4)245$$
$$t_{1,5} + t_{2,5} + t_{3,5} + t_{4,5} + t_{5,5} = \alpha_5 185 + (1 - \alpha_5)252$$
$$23t_{1,1} + 20t_{1,2} + 19t_{1,3} + 31t_{1,4} + 25t_{1,5} + \dots$$
$$+ 10t_{5,1} + 12t_{5,2} + 9t_{5,3} + 15t_{5,4} + 17t_{1,5} = \eta 0.1 + (1 - \eta) * 78$$
$$t_{ij} \geq 0, \ \rho_i, \ \alpha_j, \ \eta \in [\varepsilon; 1 - \varepsilon]$$

where ε is an extremely small positive number (to avoid $\ln(0)$).

First, the results of the proposed model are compared with the results of 4 optimization models with membership functions from the literature. These are two trapezoidal-type functions, one triangular function and a function with CRISP (interval) parameters. Figure 2 shows the parameters used in the membership functions for $\mathbf{o}_i$. The structures of the membership functions for $\mathbf{d}_j$ and C are equivalent to the structure of the $\mathbf{o}_i$ membership function. Some features of these membership functions are as follows:

(1) The CRISP function takes the value 1 within interval $[\underline{o_i}, \overline{o_i}]$ and takes the value zero outside this interval (Figure 2(d)).
(2) Similar to all entropic functions, triangular membership functions reach the highest value in the center of the interval $\frac{\underline{o_i}+\overline{o_i}}{2}$ (Figure 2(b)).
(3) Membership functions for trapezoidal numbers $(\underline{o_i}, \mathbf{o_{i1}}, \mathbf{o_{i2}}, \overline{o_i})$ were designed with a symmetrical structure; that is to say, the distance between $\underline{o_i}$ and $\mathbf{o_{i1}}$ is equal to the distance between $\mathbf{o_{i2}}$ and $\overline{o_i}$. In the case of trapezoidal model 1, this distance is $\frac{1}{4}(\overline{o_i} - \underline{o_i})$ (Figure 2(a)), while In the case of trapezoidal model 2, this distance is $\frac{1}{3}(\overline{o_i} - \underline{o_i})$ (Figure 2(c)).

The main objective for designing these functions, with this symmetrical structure, is the possibility of comparing different structures of entropic membership functions.

In Table 3, we display a summary of the results for trapezoidal membership functions 1 and 2, CRISP range, and our entropic function. In all cases, for $\theta \in \{0.1, \dots, 0.9\}$ or $E \in \{0.1, \dots, 0.9\}$, the mean, standard deviation, and coefficient of variation for the number of trips associated with each OD matrix obtained are reported.

According to the results in Table 3, the OD matrices of the entropic membership model are the least dispersed with respect to the values of θ. In addition, in this model, the number of trips is not as dispersed in terms of the variation in the values of θ; the number of trips ranges from 44.4 ($\theta = 0.1$) to 41.951 ($\theta = 0.9$). The most dispersed results are obtained with trapezoidal model 2 in the case of $E = 0.8$. However, in the triangular model, greater variability is noted with small values of E. It is worth stating that in some cases, feasible models cannot be achieved due to the α-minimum cut of the membership function. In our formulation, all the results are feasible due to the formulation of the bi-objective problem as a compromise optimization model.

To analyze the variability and similarity in the OD matrices obtained with the triangular model, trapezoidal model 1, trapezoidal model 2, CRISP model and entropic membership

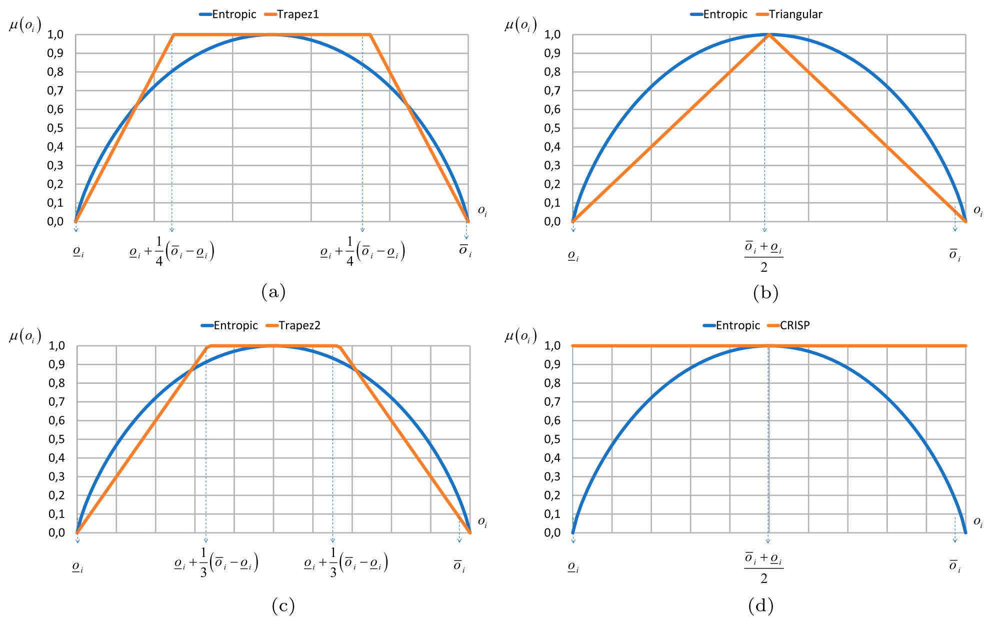

Figure 2. Membership functions for o_i. (a) Fig 2a. (b) Fig 2b. (c) Fig 2c and (d) Fig 2d.

Table 3. Summary of the membership function results.

	Entropic			Trapezoidal model 1			Trapezoidal model 2			Triangular		
θ, E	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV
0.1	44.398	0.791	1.8%	47.055	4.826	10.3%	46.750	5.030	10.8%	46.114	5.327	11.6%
0.2	43.082	1.163	2.7%	46.114	5.327	11.6%	45.469	5.615	12.3%	44.166	6.226	14.1%
0.3	42.59	1.705	4.0%	45.143	5.755	12.7%	44.166	6.226	14.1%	42.220	7.449	17.6%
0.4	42.337	2.443	5.8%	44.166	6.226	14.1%	42.862	6.955	16.2%	39.672	7.939	20.0%
0.5	42.184	3.462	8.2%	43.188	6.763	15.7%	41.527	7.906	19.0%	39.710	22.43	56.5%
0.6	42.084	4.932	11.7%	42.220	7.449	17.6%	39.672	7.939	20.0%		INFEASIBLE	
0.7	42.016	7.214	17.2%	41.072	7.824	19.%	39.489	16.456	41.7%		INFEASIBLE	
0.8	41.972	11.211	26.7%	39.672	7.939	20.0%	39.931	28.822	72.2%		INFEASIBLE	
0.9	41.951	19.977	47.6%	39.379	13.576	34.5%		INFEASIBLE			INFEASIBLE	

function model, the χ^2 statistical test is used, following the ideas of Black (2018). Generally, given two matrices T and S, the test seeks to determine whether matrix T fits matrix S with the following test statistic:

$$\chi^2(T, S) = \sum_{i=1}^{K} \sum_{j=1}^{L} \frac{(T_{ij} - S_{ij})^2}{S_{ij}} \tag{66}$$

This test statistic χ^2 is distributed with $K - 1 \times L - 1$ degrees of freedom (16 for our first example). If p−value $= P(\chi^2_{gl=K-1\times L-1} > \chi^2(T, S))$ is greater than a significance level (for example, 5%), then the hypothesis that matrix T fits matrix S cannot be rejected. Note that this test is not necessarily symmetric. In other words, if a matrix T fits S, S does not necessarily fit T. For the analysis of the results, the two tests are carried out, and thus, it is shown that the two OD matrices are statistically equal. On the other hand, if in any case the probability is less than the significance level, it can be concluded that the matrices are different. In addition, the following indicator is calculated to analyze the percent difference between the two matrices.

$$diff(T,S) = \frac{1}{K*L}\sum_{i=1}^{K}\sum_{j=1}^{L}\frac{|T_{ij}-S_{ij}|}{S_{ij}} \tag{67}$$

If $T = S$ then $diff(T,S) = 0$. This indicator measures the average percentage change in the number of trips between the OD pairs of the matrix T with respect to the matrix S. Note that S_{ij} must be greater than zero for Equations (66) and (67) to make sense.

First, the results of the model with CRISP interval parameters (T_{crisp}) are compared with the models with trapezoidal (1 and 2), triangular and entropic membership functions, which are noted (T^{E}_{trap1}; T^{E}_{trap1}; T^{E}_{triang}; T^{θ}_{entr}). The parameter E is the minimum level of membership in the trapezoidal membership functions within the model (see Equation (25)), and θ is the valuation of the goal g^2 within the optimization model described in Equations (46)–(50).

The following Table 4 shows the value of the χ^2 statistic for each case (the trapezoidal models 1 and 2 and the triangular and entropic membership functions), the p-value, and the level of difference in the obtained matrices, where T_{crisp} is set as the adjustment or reference matrix.

The results in Table 4 show that the four models adjust with the CRISP model when there are small values of θ (E) (that is, it cannot be shown that the solutions of the four models are significantly different from a CRISP solution). For example, trapezoid model 1 result, when E is less than or equal to 0.5, fits with the CRISP result; however, as E becomes larger, that fit becomes much smaller. For trapezoid model 2, when E is less than or equal to 0.4, the result can be adjusted to the CRISP model result. For the triangular model, when E is equal to 0.1 and 0.2, the result can be adjusted. In the entropic model, the results can be statistically adjusted in more cases, up to when θ is less than or equal to 0.6. In general, it is shown that all the models generate different solutions as the values of θ and E increase. This implies that the real existence of two conflicting objectives (entropic and fuzzy membership) can be affirmed, which validates the basic assumption of multi-objective optimization models (in this case, a bi-objective model). On the other hand, when analyzing the indicator *diff*, it is observed that all the models generate different solutions by varying the values of θ and E. We can highlight the following results from this indicator:

- For small values of θ (E), the OD matrix of the membership function with entropy is the least similar to that of the CRISP model.
- For high values of θ (E), the OD matrix of the trapezoid 2 membership function is the least similar to that of the CRISP model.
- There is no direct relationship between the difference indicator and the χ^2 statistic. For example, for trapezoidal model 1 with $E = 0.5$, there is a p-value of 0.126 and an average percentage difference of 11%. In trapezoidal model 2, with $E = 0.4$, there is an indicator of difference of 11.7% and a p-value of 0.062, and finally, in the entropic version, with $\theta = 0.4$, there is a difference indicator of 11.4% and a p-value of 0.403.

In Tables 5, 6 and 7, we show that there is a relationship among the solutions of the trapezoidal (1 and 2), triangular and entropic membership function models. This result permits

Table 4. Results of the chi-square test and difference indicator.

	Trapezoidal model 1			Trapezoidal model 2			Triangular			Entropic		
E, θ	χ^2	*p*-value	*diff*	χ^2	*p*-value	*diff*	χ^2	*p*-value	*diff*	χ^2	*p*-value	*diff*
0.1	1.752	1	3.1%	2.729	0.9999	3.8%	4.791	0.997	5.01%	10.177	0.857	7.80%
0.2	4.792	0.997	5%	7.459	0.963	6.2%	14.939	0.529	8.94%	14.512	0.561	9.70%
0.3	9.043	0.912	6.9%	14.939	0.529	8.9%	31.82	0.011	13.31%	16.033	0.451	10.70%
0.4	14.939	0.529	8.9%	25.468	0.062	11.7%	51.428	1.35E−05	18.74%	16.729	0.403	11.40%
0.5	22.549	0.126	11%	38.287	0.0014	14.9%	242.515	1.77E−42	38.77%	17.683	0.343	11.80%
0.6	31.82	0.011	13.3%	51.428	1.35E−05	18.7%		INFEASIBLE		20.4058	0.202497	12,30%
0.7	40.357	0.0007	15.7%	139.034	1.11E−21	29.7%		INFEASIBLE		28.743	0.02573	13.60%
0.8	51.428	1.35E−05	18.7%	394.385	5,47E−74	50.3%		INFEASIBLE		56.266	2.20E−06	19.10%
0.9	102.528	1.16E−14	25.3%		INFEASIBLE			INFEASIBLE		175.021	8.36E−29	33.10%

Table 5. Results of the chi-square test on the entropic and trapezoidal 1 membership functions.

E	θ	$\chi^2(T^E_{trap1}, T^\theta_{Entro})$	p-value	$\chi^2(T^\theta_{Entro}, T^E_{trap1})$	p-value
0.1	0.1	14.57	0.556	12.968	0.675
0.2	0.2	16.998	0.386	14.852	0.536
0.3	0.3	15.719	0.473	14.027	0.597
0.4	0.4	14.237	0.581	13.218	0.657
0.5	0.5	13.801	0.614	13.567	0.631
0.6	0.6	14.876	0.534	15.398	0.496
0.7	0.7	12.453	0.712	13.562	0.631
0.8	0.8	15.824	0.465	17.845	0.333
0.9	0.8	13.581	0.63	16.127	0.444

Table 6. Results of the chi-square test on the entropic and trapezoidal 2 membership functions.

E	θ	$\chi^2(T^E_{trap2}, T^\theta_{Entro})$	p-value	$\chi^2(T^\theta_{Entro}, T^E_{trap2})$	p-value
0.1	0.1	14.880	0.533	13.376	0.645
0.2	0.2	16.737	0.403	15.061	0.520
0.3	0.3	16.499	0.419	15.624	0.479
0.4	0.4	17.700	0.342	18.152	0.315
0.5	0.5	19.852	0.227	22.283	0.134
0.6	0.6	13.332	0.648	15.605	0.481
0.7	0.9	14.605	0.554	15.770	0.469
0.8	0.95	7.164	0.970	7.916	0.951

Table 7. Results of the chi-square test on the entropic and triangular membership functions.

E	θ	$\chi^2(T^E_{triang}, T^\theta_{Entro})$	p-value	$\chi^2(T^\theta_{Entro}, T^E_{triang})$	p-value
0.1	0.5	13.821	0.612	12.021	0.743
0.2	0.6	10.768	0.824	9.857	0.874
0.3	0.7	15.984	0.454	16.338	0.430
0.4	0.8	15.824	0.465	17.845	0.333
0.5	0.9	16.169	0.441	21.047	0.177

us to conclude that the model with the entropic membership function reproduces the solutions of the classic models from the literature. After the tables, the practical implications of these results are explained.

Table 5 shows the χ^2 test between the entropic model matrices (T^θ_{Entro}) and the trapezoidal model 1 matrices (T^E_{trap1}) (the test is performed in both directions to show the equivalence of the two matrices).

The Table 5 shows that in almost all cases, when $\theta = E$, the matrices $T^E_{trap1}, T^\theta_{Entro}$ are statistically equal. The minimum p-value was 0.386, which implies that the statistical evidence does not reject the hypothesis that they are equivalent matrices. When E equals 0.9, the best adjustment of the matrix $T^{0.9}_{trap1}$ is with $T^{0.8}_{Entro}$ ($\theta = 0.8$). As previously mentioned, these results permit us to conclude that the entropic model replicates the solutions not only for the CRISP model but also for the fuzzy optimization model with trapezoidal 1 membership functions of the type shown in Figure 2.

For trapezoidal model 2, the following results are obtained (Table 6):

According to Table 6, the solutions of the model with the entropic membership function are equivalent to those of the model with trapezoidal function 2. For the cases when E is 0.7 and 0.8, it is necessary to adjust with values of θ to 0.9 and 0.95, respectively. The minimum p-value is 0.134, which indicates that there is no statistical support to reject the statement that the matrices are the same for the cases mentioned in that table.

The Table 7 the results for the triangular model. For the feasible cases of this model ($E = 0.1, \ldots, 0.5$), values of θ where the OD matrices are equal to the results of the model with entropic membership functions are found. In this case, a fit is found when $\theta = E + 0.4$. In other words, small values of θ (less than 0.5) may be the least suitable for analyzing equivalence because the triangular model approaches the center of the intervals faster.

In numerical terms, from the examples developed in this section, the entropic membership model reproduces solutions that are statistically equivalent to those of the trapezoidal (1 and 2) and triangular models. However, the trapezoidal (triangular) models, given the inequality constraints that appear in their specifications, result in multipliers that are approximately equal to zero in most of the cases since the obtained solutions are interior points (complementary slackness conditions). This fact does not allow us to perform a direct analysis of the marginal effects of the variations in o_i and d_j on t_{ij}. This drawback also appears in the accessibility and attractiveness measures. The entropy model, with equality constraints, allows us to find values greater than zero for the multipliers in all the cases. On the other hand, the marginal utility of income becomes zero in the trapezoidal model because its multiplier is in fact zero (complementary slackness condition). In the entropy model, this value (β) is also different from zero. It is also worth mentioning that based on the solutions for t_{ij}, o_i and d_j obtained from the trapezoidal and triangular models, we effectively could obtain estimates for these valuations (accessibility and attractiveness). However, for conducting such analyzes, it would be necessary to implement either a classical entropic model or its dual, and to the best of our knowledge, those analyzes have not been reported in the previous literature related to fuzzy optimization for entropy and trip distribution.

The results of the marginal analysis and Pareto frontier of the model proposed in the present research are presented in Appendix.

5.2. Second test example: Concepción case

To illustrate the application of the model in a real setting, we use a set of zones from the city of Concepción in Chile. In this example, there are 100 zones distributed throughout the city. The data were obtained from the household survey conducted in 2002. For the 100 zones chosen in Concepción, there are records in the database to identify and characterize a household as a unit of analysis.

Before looking for a solution, it is important to analyze whether the inputs from the distribution stage could generate a solution. To check if this is true, the following procedure is carried out. As we want a distribution of trips not limited by a fixed value but falling within a range of values, the intersection of the boxes of variability must not be empty. Following this idea, the total sum of the 100 limits associated with the boxes of generated trips is equal to $T_{gene} = [185249; 186659]$. On the other hand, the sum of the boxes corresponding to the attracted trips is $T_{Atra} = [186441; 281860]$. Since the intersection of these sets is not empty, one could assume that it is possible to find at least one solution for the distribution

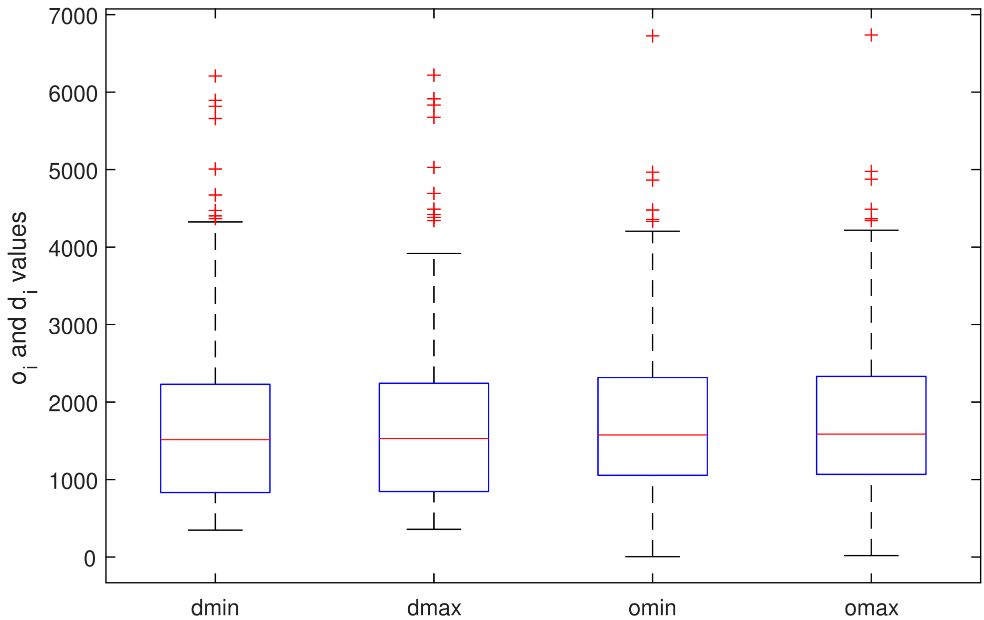

Figure 3. Box-and-whisker plots for the bounds of o_i and d_j.

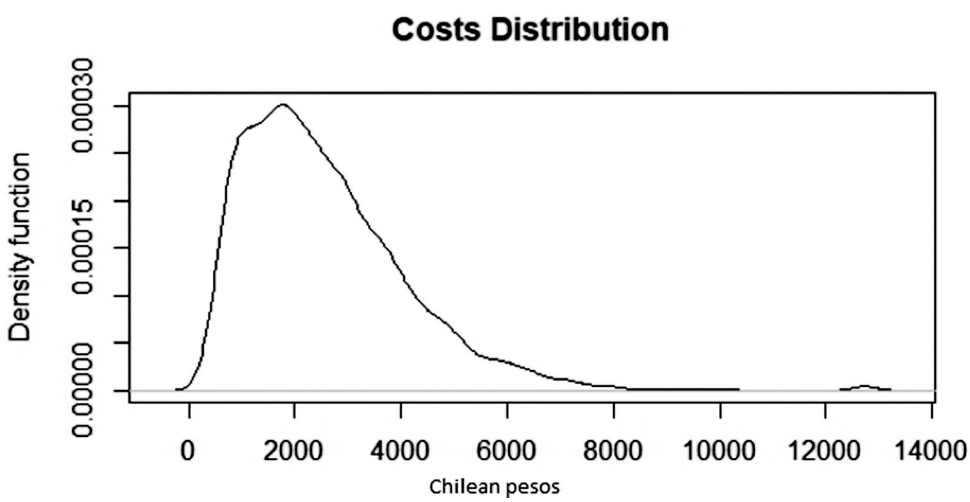

Figure 4. Density function of the generalized costs (Chilean pesos).

problem. However, the interval of attracted trips is much wider. This condition will have a relevant effect on the solutions.

Figure 3 (box-and-whisker plots) shows the distribution (dispersion and symmetry) of the bounds of the attracted and generated trips, namely, $d_{\min} = \underline{d}$, $d_{\max} = \overline{d}$, $o_{\min} = \underline{o}$ and $o_{\max} = \overline{o}$.

Figure 4 shows the density function of the generalized costs (Chilean pesos: 1 dollar is approximately 750 Chilean pesos).

Table 8. Feasibility of the models.

Model	Feasibility
CRISP	Always feasible
Trapezoidal 1	$E = 0.1, 0.2$ and 0.3
Trapezoidal 2	$E = 0.1, 0.2$
Triangular	$E = 0.1$
Entropic	$g^1(Crisp), g^2, \theta = 0.1, 0.2, 0.3, \ldots, 0.9$

Table 9. Statistical results of T_{CRISP}.

Mean	18.67
SD	19.92
CV	1.07
1st quartile	5.83
Median	12.79
1st quartile	24.48
5th percentile	0.00593

Table 10. Statistics for the absolute values of the differences between the triangular (Trapezoidal 1 and Trapezoidal 2) and CRISP models.

	$ABS(T^E_{Trap1} - T_{Crisp})$			$\mathrm{ABS}(T^E_{Trap1} - T_{Crisp})$		$ABS(T^E_{triang} - T_{Crisp})$
	$E = 0.1$	$E = 0.2$	$E = 0.3$	$E = 0.1$	$E = 0.2$	$E = 0.1$
Mean	0.012	0.023	0.035	0.016	0.031	0.023
SD	0.038	0.080	0.132	0.052	0.113	0.080
10th percentile	0.001	0.002	0.003	0.001	0.003	0.002
3th quartile	0.002	0.005	0.007	0.003	0.006	0.005
Median	0.004	0.008	0.012	0.005	0.011	0.008
1st quartile	0.007	0.014	0.020	0.009	0.018	0.014
99th percentile	0.212	0.451	0.726	0.286	0.627	0.451

In the following cases (Table 8), for the CRISP, triangular, trapezoidal 1, trapezoidal 2, and entropy models, the feasibility is obtained:

In this case, since all the obtained matrices have small travel values (much less than 1) in a considerable number of cells, it is not convenient to use the χ^2 statistic. To compare the results, the absolute values of the differences between each obtained matrix and the matrix of the CRISP model are used. That is, if there is a solution matrix S, the matrix $ABS(S - T_{Crisp})$ will be obtained, and to make these comparisons, the mean, first quartile, median and third quartile are analyzed.

First, some statistical results of the matrix T_{Crisp} are presented (Table 9).

The 5th percentile is presented to show that in at least 500 OD pairs, the number of trips is 0.

The Table 10 shows some statistical results of the absolute difference matrices between the CRISP model and the triangular membership function ($E = 0.1$), trapezoidal model 1 with $E = 0.1, 0.2$ and 0.3 and trapezoidal model 2 with $E = 0.1$ and 0.2.

The results of Table 10 show that there is no difference between the trip distribution matrix of the CRISP model and the matrices of the triangular and trapezoidal membership functions. For example, the 99th percentile has a maximum absolute difference of less than

Table 11. Statistics for the absolute values of the differences between the entropic and CRISP models.

	ABS($T^{\theta}_{entro} - T_{Crisp}$)					
	$\theta = 0.1$	$\theta = 0.3$	$\theta = 0.5$	$\theta = 0.7$	$\theta = 0.9$	$\theta = 0.99$
Mean	0.099	0.306	0.669	1.513	4.350	8.529
SD	0.195	0.572	1.299	2.967	10.560	34.837
10th percentile	0.006	0.017	0.029	0.131	0.345	0.594
1st quartile	0.018	0.057	0.123	0.333	0.891	1.445
Median	0.045	0.153	0.333	0.769	1.923	2.897
3rd quartile	0.097	0.338	0.738	1.672	4.763	7,390
99th percentile	1	2.478	5.632	12.306	35.877	77.231

1 trip (0.7626). Furthermore, the maximum value of the 3rd quartile is 0.02. In other words, the variation in the parameter (E) in the traditional models failed to reflect the diversity of the generated and attracted trips in the resulting OD matrices.

Below (Table 11), the same results (the absolute values of the differences between the OD entropic model matrices and the CRISP model matrices) are shown for some values of θ. Note that as the value of θ increases, the solutions obtained by the proposed model T^{θ}_{entrop} move away from the CRISP model solutions. That is, in this case, a bi-objective model that better captures the variability of the parameters of the attracted and generated trips and the total cost of the system is indeed found. Unlike the classic models, a range of travel variability can actually be generated between each OD pair. It is worth mentioning that starting at $\theta = 0.7$, a significant difference between the matrices T^{θ}_{entrop} and T_{crisp} is noticed.

6. Conclusions

In this research, an extension of the doubly delimited model of maximum entropy is proposed for the distribution of trips, including the uncertainty inherent in the estimation of the generated and attracted trips by zone (oi and dj, respectively) and the total cost. To incorporate this uncertainty, a fuzzy membership function is formulated using the entropy value of a Bernoulli random variable. This membership function allows us to consider a differentiable bi-objective optimization problem. For this reason, this method can be implemented easily. On the other hand, the entropic membership function can be seen as a (differentiable) smooth approach to other membership functions, such as triangular and trapezoidal ones. This fuzzy entropic membership function can be applied to measure the subjective value of information under a condition of uncertainty related to the parameters of interest. In addition, we highlight the tactical and operational gains provided by our approach regarding their impact on the process of designing transportation systems as a result of the inclusion of information gain, randomness and ambiguity in the OD parameters.

Another advantage of the proposed model is that it is not necessary to use expansion factors to fulfil the condition that $\sum_i o_i = \sum_j d_j$ because the interval estimations of $\mathbf{o}_i$ or $\mathbf{d}_j$ allow the relaxation of such conditions. In addition, the interval estimations allow us to analyze, within the different stages of decisions, the impact generated by the uncertainty or variability in the parameters obtained in the first stages of an urban transport planning methodology.

This model is resolved by means of compromise programing considering a weighted sum of objectives. From such a formulation, we are able to analyze, compute and extend the results typically found in transport economic analyzes, such as accessibility, attractiveness and generalized cost valuation. In addition, this formulation, as a weighted sum of concave functions, allows us to solve the problem in just one step, keeping the uniqueness of the solution of the associated optimization problem.

Through some illustrative numerical examples, we have shown that the model based on entropic membership functions is able to capture the variations in distribution patterns associated with changes in the weights embedded in a bi-objective formulation of the problem, where we explicitly model the trade-off between the maximization of total system entropy and the maximization of the sum of the membership functions for the generation/attraction and total cost functions. The results show that the entropic model can make a difference with respect to the CRISP version when some combinations of weights are tested in the bi-objective formulation, which is not possible with many of the alternative membership function models. In addition, from the reported experiments, we observe that the entropic model is feasible regardless of the values of the weights, which are found to be relevant features because, in several cases, the alternative models report infeasible solutions.

Although the strategic aspects indicated above are relevant, there are some tactical and operational issues worth highlighting. In this sense, we note that the proposed method generates additional valuable information from the result, such as (i) the possibility to reproduce (under certain specific conditions) the results of the fuzzy methods in particular cases of our more general Pareto frontier, and (ii) indicators, including a_i and b_j, in the proposed formulation and factors that are not defined in the fuzzy cases are used as benchmarks. Thus, we recognize some important aspects in terms of the more robust results and interpretability of the obtained parameters, such as the balance factors as accessibility and attractiveness measures. In fact, in terms of policy or project evaluation, it is feasible to generate boxes of uncertainty for the measures of accessibility and attractiveness, as well as for the total number of trips between different zones. These values are relevant to urban and transport economics and are not easy to obtain with alternative specifications of membership functions. In entropic specifications, computing the dual values from complementary slackness conditions is straightforward, as the model is written in terms of equality constraints.

For future research, we can first generalize this modeling approach to consider other methods for modeling the distribution of trips, which consider different constraints and aggregation levels (modes, route choice, land use, etc.) (Wang et al. 2018; Li et al. 2019). We can also consider a different disaggregation of the objectives (for example, one for attracted trips, another for generated trips and one that focuses on costs). Several extensions of the results can be envisaged in other contexts of mathematical modeling, such as logistic applications (Teye, Bell, and Bliemer 2017; Teye and Hensher 2021), game theory (Roy and Das 2013), queuing systems (Jain and Bhagat 2015; Jain, Sharma, and Rani 2016), portfolio optimization (Zhou et al. 2015), urban and land use systems (Briceño-Arias and Martínez 2018; Purvis, Mao, and Robinson 2019), transportation and land use interactions (Li et al. 2019), and social systems (Mishra and Ayyub 2019) among others.

Acknowledgments

The authors are grateful for the financial support provided under Projects ANID/FONDECYT/Regular 1191200, CONICYT/FONDECYT 11160320, Instituto Sistemas Complejos de Ingeniería (ANID PIA/APOYO AFB180003) and project 4.148 of Fundación para la Promoción de la Investigación y la Tecnología. Banco de la República. Colombia.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors are grateful for the financial support provided under Projects Fondo Nacional de Desarrollo Científico y Tecnológico ANID/FONDECYT/Regular 1191200, ANID/FONDECYT 11160320, Complex Engineering Systems Institute ANID PIA/APOYO AFB180003 and project 4.148 of Fundación para la Promoción de la Investigación y la Tecnología. Banco de la República. Colombia.

ORCID

Héctor López-Ospina http://orcid.org/0000-0001-8144-6394
Cristián E. Cortés http://orcid.org/0000-0001-7041-4602
Juan Pérez http://orcid.org/0000-0003-3640-2514

References

Aggarwal, M. 2021. “Redefining Fuzzy Entropy with a General Framework.” *Expert Systems with Applications* 164. doi:10.1016/j.eswa.2020.113671.

Ait-Ali, A., and J. Eliasson. 2019. “Dynamic Origin-Destination Estimation Using Smart Card Data: An Entropy Maximisation Approach.” Preprint, arXiv:1909.02826.

Baidya, A., U. Bera, and M. Maiti. 2014. “Solution of Multi-Item Interval Valued Solid Transportation Problem with Safety Measure Using Different Methods.” *Opsearch* 51 (1): 1–22.

Bellman, R. E., and L. A. Zadeh. 1970. “Decision-Making in a Fuzzy Environment.” *Management Science* 17 (4): B–141.

Bit, A., M. Biswal, and S. Alam. 1992. “Fuzzy Programming Approach to Multicriteria Decision Making Transportation Problem.” *Fuzzy Sets and Systems* 50 (2): 135–141.

Black, J. 2018. *Urban Transport Planning: Theory and Practice*. Vol. 4. London: Routledge. doi:10.4324/9781351068604.

Bonnel, H., and J. Collonge. 2015. “Optimization Over the Pareto Outcome Set Associated with a Convex Bi-Objective Optimization Problem: Theoretical Results, Deterministic Algorithm and Application to the Stochastic Case.” *Journal of Global Optimization* 62 (3): 481–505.

Bonnel, H., and J. Collonge. August 2014. “Stochastic Optimization Over a Pareto Set Associated with a Stochastic Multi-Objective Optimization Problem.” *Journal of Optimization Theory and Applications*162 (2): 405–427. doi:10.1007/s10957-013-0367-8.

Briceño-Arias, L., and F. Martínez. 2018. “Short-Term Land Use Planning and Optimal Subsidies.” *Networks and Spatial Economics* 18 (4): 973–997.

Caceres, N., L. Romero, and F. G. Benitez. 2020. “Exploring Strengths and Weaknesses of Mobility Inference From Mobile Phone Data Vs. Travel Surveys.” *Transportmetrica A: Transport Science* 16 (3): 574–601.

Cascetta, E., F. Pagliara, and A. Papola. 2007. “Alternative Approaches to Trip Distribution Modelling: A Retrospective Review and Suggestions for Combining Different Approaches.” *Papers in Regional Science* 86 (4): 597–620.

Caschili, S., A. De Montis, and D. Trogu. 2015. “Accessibility and Rurality Indicators for Regional Development.” *Computers, Environment and Urban Systems* 49: 98–114.

Chanas, S., and D. Kuchta. 1996. "A Concept of the Optimal Solution of the Transportation Problem with Fuzzy Cost Coefficients." *Fuzzy Sets and Systems* 82 (3): 299–305.
Chen, B., Y. Liu, and T. Zhou. 2019. "An Entropy Based Solid Transportation Problem in Uncertain Environment." *Journal of Ambient Intelligence and Humanized Computing* 10 (1): 357–363.
Cochrane, R. 1975. "A Possible Economic Basis for the Gravity Model." *Journal of Transport Economics and Policy* 9 (1): 34–49.
Das, I., and J. Dennis. 1997. "A Closer Look at Drawbacks of Minimizing Weighted Sums of Objectives for Pareto Set Generation in Multicriteria Optimization Problems." *Structural Optimization* 14 (1): 63–69. doi:10.1007/BF01197559.
da Silva, M. A. V., and D. Marcio de Almeida. 2013. "A Model to Estimate the Origin–Destination Matrix for Soybean Exportation in Brazil." *Journal of Transport Geography* 26: 97–107.
de Grange, L., E. Fernández, and J. de Cea. 2010. "A Consolidated Model of Trip Distribution." *Transportation Research Part E: Logistics and Transportation Review* 46 (1): 61–75.
Delgado, J. C., and P. Bonnel. 2016. "Level of Aggregation of Zoning and Temporal Transferability of the Gravity Distribution Model: The Case of Lyon." *Journal of Transport Geography* 51: 17–26.
Demissie, M. G., S. Phithakkitnukoon, and L. Kattan. 2018. "Trip Distribution Modeling Using Mobile Phone Data: Emphasis on Intra-Zonal Trips." *IEEE Transactions on Intelligent Transportation Systems*20 (7): 2605–2617.
ESTRAUS. 1989. "Estudio estratégico de transporte del gran santiago."
Fang, S.-C., and H.-S. Tsao. 1995. "Linearly-Constrained Entropy Maximization Problem with Quadratic Cost and Its Applications to Transportation Planning Problems." *Transportation Science* 29 (4): 353–365.
Feng, X., J. Zhang, and A. Fujiwara. 2009. "Analysis of Feedback to Obtain Steady-State Solutions in Four-Step Modelling." *Transportmetrica* 5 (3): 215–227.
Flowerdew, R., and M. Aitkin. 1982. "A Method of Fitting the Gravity Model Based on the Poisson Distribution." *Journal of Regional Science* 22 (2): 191–202.
Frederix, R., F. Viti, and C. M. Tampère. 2013. "Dynamic Origin–Destination Estimation in Congested Networks: Theoretical Findings and Implications in Practice." *Transportmetrica A: Transport Science* 9 (6): 494–513.
Grandinetti, L., F. Guerriero, D. Laganà, and O. Pisacane. 2010. "An Approximate ϵ-Constraint Method for the Multi-Objective Undirected Capacitated ARC Routing Problem." In *Experimental Algorithms*, edited by P. Festa, 214–225, Berlin, Heidelberg: Springer Berlin Heidelberg.
Hewings, G. J., and E. Fernandez-Vazquez. 2019. "Entropy Maximization and Input–Output Analysis." *Interdisciplinary Science Reviews* 44 (3–4): 272–285.
Hitchcock, F. L. 1941. "The Distribution of a Product From Several Sources to Numerous Localities." *Journal of Mathematics and Physics* 20 (1–4): 224–230.
Hu, T.-Y., and W.-M. Ho. 2016. "A Modified Entropy-Based Dynamic Gravity Model for the Evacuation Trip Distribution Problem During Typhoons." *Journal of the Chinese Institute of Engineers* 39 (5): 548–555.
Hu, B., Y. Ma, Y. Pei, and W. Gao. 2020. "Statistical Analysis and Predictability of Inter-urban Highway Traffic Flows: A Case Study in Heilongjiang Province, China." *Transportmetrica A: Transport Science*16 (3): 1062–1078.
Jain, M., and A. Bhagat. 2015. "Analysis of Bulk Retrial Queue Using Maximum Entropy Principle." *International Journal of Operational Research* 23 (4): 477–496.
Jain, M., G. Sharma, and V. Rani. 2016. "Maximum Entropy Approach for An Unreliable Mx/g/1 Queueing System with Bernoulli Vacation, Restricted Admission and Delayed Phase Repair Under N-Policy." *International Journal of Services and Operations Management* 23 (4): 407–442.
Kaisa, M. 1999. *Nonlinear Multiobjective Optimization*, volume 12 of *International Series in Operations Research & Management Science*. Boston: Kluwer Academic Publishers.
Kalic, M., and D. Teodorovic. 1997. "Trip Distribution Modeling Using Soft Computing Techniques." In *EURO XV/INFORMS XXXIV (Book of Abstracts, p. 74), Barcelona*.
Kalić, M., and D. Teodorović. 2003. "Trip Distribution Modelling Using Fuzzy Logic and a Genetic Algorithm." *Transportation Planning and Technology* 26 (3): 213–238.

Kurtulmuşoğlu, F. B., F. Pakdil, and K. D. Atalay. 2016. "Quality Improvement Strategies of Highway Bus Service Based on a Fuzzy Quality Function Deployment Approach." *Transportmetrica A: Transport Science* 12 (2): 175–202.

León, J. P., J. C. Figueroa-García, and H. López-Ospina. 2016. "A Model for Fuzzy-Cost Travel Distribution Problems Using Entropy Measures." In *Workshop on Engineering Applications*, 268–279. Springer.

Leyffer, S. 2009. "A Complementarity Constraint Formulation of Convex Multiobjective Optimization Problems." *INFORMS Journal on Computing* 21 (2): 257–267. doi:10.1287/ijoc.1080.0290.

Li, T., H. Sun, J. Wu, Z. Gao, Y.-e. Ge, and R. Ding. 2019. "Optimal Urban Expressway System in a Transportation and Land Use Interaction Equilibrium Framework." *Transportmetrica A: Transport Science* 15 (2): 1247–1277.

Lim, Y., and H. Kim. 2016. "A Combined Model of Trip Distribution and Route Choice Problem." *Transportmetrica A: Transport Science* 12 (8): 721–735.

López-Ospina, H. 2013. "Maximization Model of Entropy and Generalized Costs for Urban Travel Distribution." *Ingeniería y Universidad* 17 (2): 391–407.

Martínez, F. J. 1995. "Access: The Transport-Land Use Economic Link." *Transportation Research Part B: Methodological* 29 (6): 457–470.

Martínez, F., and C. Araya. 2000. "Transport and Land-Use Benefits Under Location Externalities." *Environment and Planning A* 32 (9): 1611–1624.

Messac, A., A. Ismail-Yahaya, and C. A. Mattson. July 2003. "The Normalized Normal Constraint Method for Generating the Pareto Frontier." *Structural and Multidisciplinary Optimization* 25 (2): 86–98. doi:10.1007/s00158-002-0276-1.

Mishra, S., and B. M. Ayyub. 2019. "Shannon Entropy for Quantifying Uncertainty and Risk in Economic Disparity." *Risk Analysis* 39 (10): 2160–2181.

Ojha, A., B. Das, S. Mondal, and M. Maiti. 2009. "An Entropy Based Solid Transportation Problem for General Fuzzy Costs and Time with Fuzzy Equality." *Mathematical and Computer Modelling* 50 (1–2): 166–178.

Ojha, A., B. Das, S. Mondal, and M. Maiti. 2010. "A Solid Transportation Problem for An Item with Fixed Charge, Vechicle Cost and Price Discounted Varying Charge Using Genetic Algorithm." *Applied Soft Computing* 10 (1): 100–110.

Ortuzar, J. D. D., and L. Willumsen. 2011. "Modal Split and Direct Demand Models." In *Modelling Transport*, 207–225. West Sussex, United Kingdom: John Wiley & Sons Ltd.

Park, K., S. Sabouri, T. Lyons, G. Tian, and R. Ewing. 2019. "Intrazonal Or Interzonal? Improving Intrazonal Travel Forecast in a Four-Step Travel Demand Model." *Transportation* 47: 1–22.

Pereyra, V., M. Saunders, and J. Castillo. 2013. "Equispaced Pareto Front Construction for Constrained Bi-Objective Optimization." *Mathematical and Computer Modelling* 57 (9–10): 2122–2131. https://doi.org/10.1016/j.mcm.2010.12.044. *System Dynamics in Project Management & Applied Mathematics and Computational Science and Engineering-Selected Papers of the Seventh PanAmerican Workshop – June 6–11 2010*, Venezuela.

Pourebrahim, N., S. Sultana, A. Niakanlahiji, and J.-C. Thill. 2019. "Trip Distribution Modeling with Twitter Data." *Computers, Environment and Urban Systems* 77: Article ID: 101354.

Purvis, B., Y. Mao, and D. Robinson. 2019. "Entropy and Its Application to Urban Systems." *Entropy*21 (1): 56.

Ros-Roca, X., L. Montero, and J. Barceló. 2021. "Investigating the Quality of Spiess-Like and SPSA Approaches for Dynamic OD Matrix Estimation." *Transportmetrica A: Transport Science* 17 (3): 235–257.

Roy, S. K., and C. B. Das. 2013. "Multicriteria Entropy Bimatrix Goal Game: A Fuzzy Programming Approach." *J. Uncertain Syst* 7: 108–117.

Samanta, B., and T. K. Roy. 2005. "Multiobjective Entropy Transportation Model with Trapezoidal Fuzzy Number Penalties, Sources, and Destinations." *Journal of Transportation Engineering* 131 (6): 419–428.

Schlaich, T., A. L. Horn, M. Fuhrmann, and H. Friedrich. 2020. "A Gravity-Based Food Flow Model to Identify the Source of Foodborne Disease Outbreaks." *International Journal of Environmental Research and Public Health* 17 (2): 444.

Schneider, M. 1959. "Gravity Models and Trip Distribution Theory." *Papers in Regional Science* 5 (1): 51–56.
Shen, Y. 2019. "Segregation Through Space: A Scope of the Flow-based Spatial Interaction Model." *Journal of Transport Geography* 76: 10–23.
Siegel, J. D., J. De Cea, J. E. Fernández, R. E. Rodriguez, and D. Boyce. 2006. "Comparisons of Urban Travel Forecasts Prepared with the Sequential Procedure and a Combined Model." *Networks and Spatial Economics* 6 (2): 135–148.
Stouffer, S. A. 1940. "Intervening Opportunities: A Theory Relating Mobility and Distance." *American Sociological Review* 5 (6): 845–867.
Sun, C., Y. Chang, Y. Shi, L. Cheng, and J. Ma. 2019. "Subnetwork Origin-Destination Matrix Estimation Under Travel Demand Constraints." *Networks and Spatial Economics* 19 (4): 1123–1142.
Tang, J., S. Zhang, X. Chen, F. Liu, and Y. Zou. 2018. "Taxi Trips Distribution Modeling Based on Entropy-maximizing Theory: A Case Study in Harbin City-china." *Physica A: Statistical Mechanics and its Applications* 493: 430–443.
Teye, C., M. G. Bell, and M. C. Bliemer. 2017. "Urban Intermodal Terminals: The Entropy Maximising Facility Location Problem." *Transportation Research Part B: Methodological* 100: 64–81. ISSN: 0191-2615.
Teye, C., and D. A. Hensher. 2021. "A Commodity-Based Production and Distribution Road Freight Model with Application to Urban and Regional New South Wales." *Transportmetrica A: Transport Science* 17 (4): 566–592. doi:10.1080/23249935.2020.1802361.
Thomas, T., and S. Tutert. 2013. "An Empirical Model for Trip Distribution of Commuters in the Netherlands: Transferability in Time and Space Reconsidered." *Journal of Transport Geography* 26: 158–165.
Tsekeris, T., and A. Stathopoulos. 2006. "Gravity Models for Dynamic Transport Planning: Development and Implementation in Urban Networks." *Journal of Transport Geography* 14 (2): 152–160.
Velichko, A. 2016. "Interregional Transportation Modeling for the Far East of Russia Macro-Region." In *DOOR (Supplement)*, 394–403.
Wang, J., S. Peeta, X. He, and J. Zhao. 2018. "Combined Multinomial Logit Modal Split and Paired Combinatorial Logit Traffic Assignment Model." *Transportmetrica A: Transport Science* 14 (9): 737–760.
Williams, H. 2019. "The Contribution of Sir Alan Wilson to Spatial Interaction and Transport Modelling." *Interdisciplinary Science Reviews* 44 (3–4): 232–248.
Wilson, A. G. 1970. "Inter-Regional Commodity Flows: Entropy Maximizing Approaches." *Geographical Analysis* 2 (3): 255–282.
Wilson, A. 2011. *Entropy in Urban and Regional Modelling*. Vol. 1. London: Routledge. doi:10.4324/9780203142608
Winkelmann, R., and K. F. Zimmermann. 1995. "Recent Developments in Count Data Modelling: Theory and Application." *Journal of Economic Surveys* 9 (1): 1–24.
You, S. I., and S. G. Ritchie. 2019. "Tour-Based Truck Demand Modeling with Entropy Maximization Using GPS Data." *Journal of Advanced Transportation* 2019: Article ID: 5021026.
Zhang, K. shi, W. ji Li, and W. ping Song. 2008. "Bi-Level Adaptive Weighted Sum Method for Multidisciplinary Multi-Objective Optimization." doi:10.2514/6.2008-908.
Zhou, R., Z. Yang, M. Yu, and D. A. Ralescu. 2015. "A Portfolio Optimization Model Based on Information Entropy and Fuzzy Time Series." *Fuzzy Optimization and Decision Making* 14 (4): 381–397.
Zimmermann, H.-J. 1978. "Fuzzy Programming and Linear Programming with Several Objective Functions." *Fuzzy Sets and Systems* 1 (1): 45–55.

Appendix. Additional comments on methodology and results

This section includes two interrelated topics. We first analyze the balance factors; in the second part, we provide more information and an example with regard to the Pareto frontier.

Balance factors analysis

The results of the marginal analysis of the model proposed in the present research are presented in this section. We focus on the calculation of accessibility and attractiveness measures. These results are not comparable to the previous models, because in those studies such considerations are not carried out.

Regarding the balance factors of the zones considered, a_i and b_j, the results are presented below according to three different configurations of the weights associated with the two considered objectives.

Additionally, the values of β (marginal utility of income) are presented in Table A2.

It can be observed that the results in Table A1 are similar among zones. However, it cannot be stated that the zones have the same accessibility or attractiveness ratings, due to the changes in β reported in Table A2. In this example, for a fixed value of θ almost the same attractiveness or accessibility values (using Equations (9) and (10)) for all zones are obtained. This indicates that these zones have similar characteristics for those who travel. In other words, for users, destination zones are almost equally accessible from all origins, and destinations attract similar flows from all origins. However, by varying θ, an interval where these measures vary in each zone can be obtained, allowing for the flexibility of these indicators which are relevant in models for land use and the evaluation of transportation projects. The following table shows the average results of those values (average accessibility and attractiveness) for different values of θ.

Table A4 shows the minimum, maximum and average values of ρ_i, α_j and η. In addition, this table shows the level of membership of each entropic function according to different values of θ.

Pareto frontier

It is relevant to mention that there is more than one alternative to determine Pareto frontiers for bi-objective optimization problems; each method has drawbacks and advantages in its implementation. We utilized two different approaches, the weighted sum and the inclusion of an inequality constraint, to consider one of the objective functions; this last method is known as ε-constraints. Notice that there are more alternatives in the literature for exploring the Pareto frontier.

Table A1. Balance factors a_i and b_j.

θ	0.2	0.5	0.8	θ	0.2	0.5	0.8
Zones		Values a_i		Zones		Values b_j	
O1	0.00459	0.00469	0.00476	**D1**	0.00462	0.00473	0.00484
O2	0.00471	0.00482	0.00483	**D2**	0.00465	0.00479	0.00490
O3	0.00465	0.00477	0.00477	**D3**	0.00471	0.00483	0.00493
O4	0.00470	0.00483	0.00483	**D4**	0.00465	0.00476	0.00486
O5	0.00462	0.00473	0.00472	**D5**	0.00458	0.00469	0.00479

Table A2. β (Marginal utility of income).

θ	β
0.1	−0.000480
0.2	−0.000810
0.3	−0.001139
0.4	−0.001457
0.5	−0.001760
0.6	−0.002036
0.7	−0.002267
0.8	−0.002410
0.9	−0.002343

Table A3. Average accessibility and attractiveness for values of θ.

θ	Average accessibility and attractiveness
0.9	2263.1
0.8	2213.8
0.7	2353.9
0.6	2622.1
0.5	3037.9
0.4	3672.0
0.3	4706.2
0.2	6629.9
0.1	11240,9

Table A4. Membership levels and values of ρ, α and η.

θ	$\min(\alpha_j)$	$\max(\alpha_j)$	$\overline{\alpha}$	$\mu(\overline{\alpha})$	$\min(\rho_i)$	$\max(\rho_i)$	$\overline{\rho}$	$\mu(\overline{\rho})$	η	$\mu(\eta)$
0.1	0.387	0.404	0.396	0.969	0.244	0.268	0.252	0.815	0.061	0.333
0.2	0.493	0.501	0.495	0.999	0.344	0.361	0.349	0.933	0.091	0.440
0.3	0.529	0.537	0.532	0.997	0.381	0.396	0.385	0.961	0.104	0.481
0.4	0.546	0.556	0.551	0.992	0.400	0.414	0.403	0.973	0.112	0.506
0.5	0.557	0.568	0.563	0.989	0.411	0.424	0.415	0.979	0.119	0.527
0.6	0.564	0.577	0.570	0.986	0.418	0.431	0.422	0.982	0.127	0.549
0.7	0.569	0.583	0.576	0.983	0.423	0.435	0.427	0.985	0.137	0.577
0.8	0.572	0.587	0.579	0.982	0.427	0.438	0.430	0.986	0.153	0.618
0.9	0.573	0.589	0.580	0.981	0.429	0.439	0.432	0.987	0.186	0.692

Some of the alternatives for exploring and finding Pareto frontiers are weighted sum (Das and Dennis 1997; Kaisa 1999; Zhang, Li, and Song 2008; Leyffer 2009), weighted metrics (Kaisa 1999), the method in Kaisa (1999) and Grandinetti et al. (2010), the normal boundary intersection (Das and Dennis 1997), the normalized normal constraint (Messac, Ismail-Yahaya, and Mattson 2003) and achievement scalarizing functions (Kaisa 1999). The advantages and disadvantages of these methodologies depend on the mathematical features, variables, constraints, and objective functions involved.

In our setting, one of the objective functions is a concave maximization of the total entropy; the other is also a concave maximization. It is also related to the entropy but, in this case, associated with the OD matrix's fuzzy parameters. On the other hand, all constraints are linear. This means that our formulation falls in the category of bi-objective convex optimization problems.

Therefore, in our development, we can establish the properties that apply when a bi-objective optimization problem is convex. In particular:

(1) Every local Pareto optimal solution is also a global Pareto optimal solution.
(2) For a Pareto optimal solution, there is a weighting vector $(\theta, 1-\theta)$ such that the solution belongs to the Pareto front of a weighted sum expression, as proposed in our work.
(3) However, this is valid not only for the weighted sum approach but also for the ε-constraints method. This is because for every solution to this problem, a set of optimal weights in the context of a weighted sum approach can be found. Therefore, by transitivity, through this method, it is also possible to find the Pareto frontier in the case of convex bi-objective problems.

These statements are true, because they have been proved by theorems developed in the context of the multi-objective optimization literature. Specifically, Kaisa (1999) developed points 1, 2, and 3 in Theorems 2.2.3 (a more general version for quasiconvex problems is in 2.2.4), 3.1.4, and 3.2.6, respectively. Please note that these proofs are valid not only for bi-objective but also for convex multi-objective optimization problems. For more information for interested readers on Pareto set and front optimization we recommend Pereyra, Saunders, and Castillo (2013), and Bonnel and Collonge (2014).

Table A5. Values of g^2 and g^1 for different values of θ.

θ	g^2	g^1
0.1	0.0%	100.0%
0.2	60.3%	99.9%
0.3	72.2%	96,9%
0.4	76.4%	93.4%
0.5	78.6%	90.8%
0.6	80,4%	79,8%
0.7	82.3%	62.0%
0.8	85.1%	57.9%
0.9	100.0%	0.0%

In conclusion, according to our literature review and considering the mathematical features of the problem, the proposed weighted sum and methods allow us to find the Pareto frontier for our problem. As an example, we provide an example of the construction of the Pareto frontier in the example presented in Section 5.1 of the paper.

Table A5 shows the solutions of g^1 and g^2 for different values of θ (weighting of g^2 in our bi-objective problem).

$$g^1 : \frac{-\sum_{i=1}^{5}\sum_{j=1}^{5} t_{ij}\ln(t_{ij}) + 4207.2}{-3841.5 + 4207.2}$$

$$g^2 : \frac{1}{7.294 - 6.412} * \left[-\sum_{j=1}^{5}[\alpha_j \ln(\alpha_j) + (1-\alpha_j)\ln(1-\alpha_j)] \right.$$

$$\left. -\sum_{i=1}^{5}[\rho_i \ln(\rho_i) + (1-\rho_i)\ln(1-\rho_i)] - [\eta \ln(\eta) + (1-\eta)\ln(1-\eta)] - 6.412 \right]$$

Note that values of g^1 and g^2 correspond to compliance percentage levels of each objective (entropy of trips distribution and fuzzy entropic membership) and represent a discrete subset of the Pareto frontier.

For values of θ between 0.4 up to 0.6 it is noted that high levels of compliance can be achieved for both objectives jointly. For example, for $\theta = 0.6$ an 80% compliance is achieved in the two objectives. In this way, the O-D matrix obtained with $\theta = 0.6$, noted as $T_{entro}^{\theta=0.6}$ could have a high frequency of occurrence. Note that, unlike the classic multiobjective optimization models (where the decision maker selects a value of θ to implement the mathematical model), in our research it was relevant to implement the model with different values of θ to obtain the variation of the origin destination matrix due to its dynamic nature.