An entropy maximization model to estimate the origin-destination matrix of a city train with an impedance function based on trip times

Joel Paredes, Bach.[1], Edynzzon Palacios, Bach.[1] and Fernando Castro, D.Sc.[1]
[1]*Faculty of Civil Engineering, Peruvian University of Applied Sciences, Peru,*
u201524229@upcedu.pe, *u201510076@upc.edu.pe*, *pccifcas@upc.edu.pe*

Resume— A problem in urban rail public transport is the growing demand from users along their route at their boarding points. The analysis of the distribution of trips would serve for a better forecast of the future demand of the system to adapt the potential offer in terms of route design, location of stations, dimensioning of the frequency, size and quantity of wagons, establishment of the ticket value, among other components before the commissioning of urban trains, after which few are monitoring the unbridled growth in demand and urban train improvements. Given this, in order to reduce the uncertainty in estimating demand, the entropy maximization model was applied due to its advantages in relation to other methods that do not take into account variations in land use. In the context of uncertainty theory, the transport model based on uncertain entropy is transformed into its deterministic equivalent, which can be solved with general optimization methods. Finally, a numerical example is given to demonstrate the purpose.

Keywords: transport demand, trip distribution, origin-destination matrix, entropy maximization model.

I. INTRODUCTION

This research proposes the application of a forecasting model with a higher percentage of efficiency with respect to the classic models that are in force in the institutions in charge of the management, control and improvement of the main lines of public transport in urban train in the world. The conflict of over-demand and an inefficient forecast of future demand in public transport by urban train, can be affirmed as poor management in the maintenance of its activities. In the majority of urban trains they present and will present a variation of the initial demand of passengers, since the growth of the population (schoolchildren, university students and workers) is exponential year after year.

For this case, the application of the entropy maximization method is proposed in the second stage of the classical transport model, trip distribution. For this reason, previous research works on the entropy maximization model and other models with the same purpose were consulted. Solid transport based on the entropy model is an uncertain environment, but it aims to minimize the transport penalty and ensure a uniform distribution of origins and destinations [1]. This model can be derived to obtain a trip probability centered on the constraint (transportation system and cost) [2]. Another way to study trips is by comparing it with the Levenshtein distance, to create the normalized Levenshtein measurement system for origin-destination matrices (NLOD) [3].

II. METHODOLOGY

The research is a prediction of the conformation of the distribution of public transport trips by urban train based on the entropy maximization model, this proposal involves the analysis of generated trips (VG) and attracted trips (VA) with the model of entropy maximization. For a better understanding, the classic 4-stage transport model (Generation - Distribution - Modal distribution - Allocation) will be followed, but this research will have as its objective and limit the second stage (trip distribution) since it is where the forecast will be developed. future of the demand of the public transport system in urban train. The configuration of the research focuses on the review of the statistical information that will be collected from the public entities that direct and manage the main public transport mechanisms of urban trains. In addition, a data collection process through surveys. This research begins with the identification and formulation of the problem, it is through the visual identification of the conflicts that can be observed in the daily lives of citizens who travel long distances in many developed cities of the world and how they feel the experience in using it. of public transport by urban train. Once the problem is identified, we will continue with the data collection, how many enter and leave each station (generated trips - VG and attracted trips - VA) that includes the entire route of urban trains.

After the organization of the database, the construction, calibration and validation of the generation/attraction models for the forecast of future VG and VA in each of the urban train stations begins. Along with this, the future VG and VA by known route between each of the urban train stations. This activity refers to the first level of classical transport modeling (trip generation).

With the generation/attraction model carried out, the conformation of the model, calibration and validation of the trip distribution will begin using the Entropy Maximization Model (MME). Finally, we conclude with the validation of the research comparing it with another method. This activity refers to the second stage of classical transport planning model (Distribution).

Digital Object Identifier (DOI):
http://dx.doi.org/10.18687/LACCEI2021.1.1.525
ISBN: 978-958-52071-8-9 ISSN: 2414-6390

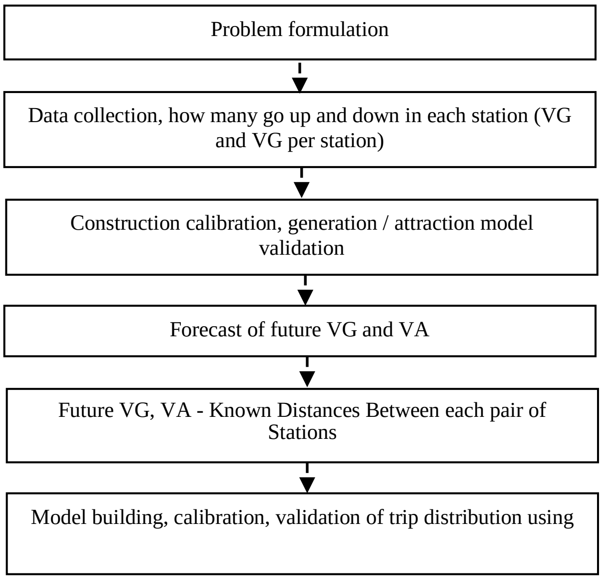

Fig. 1. Sequence of steps of the Methodology of the present study.

III. TOOLS

A. *Classic Transport Model:*

The classic model is developed as a sequence of four stages or sub-models: trip generation, trip distribution, modal split and allocation. It is now recognized that trip decisions are not actually made in this kind of sequence; a contemporary interpretation is that the positioning of each sub-model depends on the shape of the utility function to explain all these trip options [4].

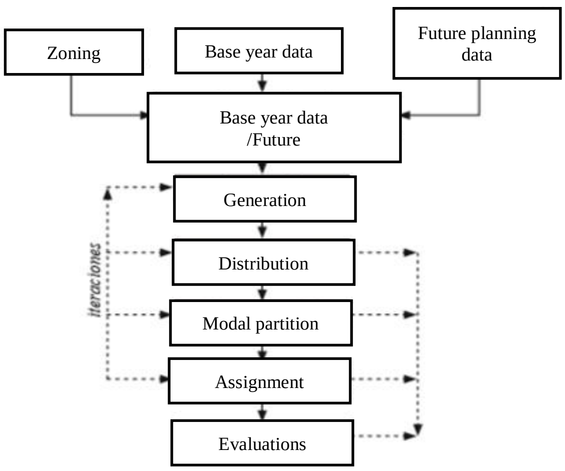

Fig. 2. The classical four-stage transport model. Adapted from Ortúzar, 2000.

B. *Trip generation:*

In the transport model, the trip generation phase aims to predict the total number of trips generated (VG) and attracted (VA) for each zone (station and / or area of influence) of the case study. Usually, the trip generation model requires a linear regression study at both the area and the household level [5].

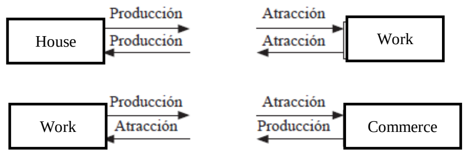

Fig. 3: Production and attraction of trips. Adapted from [5]

C. *Trip distribution:*

The zonal distribution model must predict a matrix that represents the trip structure in a study area (stations and/or area of influence).

TABLE I.
General structure of a two-dimensional trip matrix, also called an origin-destination (O/D) matrix

Origins	**Destinations**					$\sum_j T_{ij}$
	1	**2**	**3**	**... j**	**... z**	
1	T_{11}	T_{12}	T_{13}	... T_{1j}	... T_{1z}	VG_1
2	T_{21}	T_{22}	T_{23}	... T_{1j}	... T_{2z}	VG_2
...						...
I	T_{i1}	T_{i2}	T_{i3}	... T_{ij}	... T_{iz}	VG_i
...						...
z	T_{i1}	T_{i2}	T_{i3}	... T_{ij}	... T_{iz}	VG_z
$\sum_i T_{ij}$	VA_1	VA_2	VA_3	VA_J	... VA_z	

Source: [5]

The cells in each row i include the trips originating in zone i and destination in the zones of the corresponding columns. The main diagonal corresponds to intrazonal trips, while Tij represents the number of trips between origin i and destination j; VGi is the total number of trips originating from zone i, and VAj is the total number of trips attracted by zone j. [5]

D. *Entropy function:*

One of the fundamental bases of the second law of thermodynamics. It is the function called entropy that is used to calculate the degree of disorder within a process and allows us to distinguish useful energy, which is that which is converted entirely into work, from useless energy, which is lost in the environment.

E. *Impedance function (Z):*

The impedance function arises from the opposition that a circuit presents to a current when a tension. Impedance extends the concept of endurance to the circuits of alternating current (CA), and has both magnitude and phase, unlike resistance, which only has magnitude. Its formula is:

$$Z = V/I \qquad (1)$$

Where:
Z = impedance
V = voltage factor
I = intensity factor

For this research, the impedance function is an important factor in the development of the entropy maximization model because it opposes a proportional distribution, preventing the case from being unreal.

F. Entropy maximization model (MME):

The basic principles of maximum entropy, from the probabilistic point of view, assuming discrete and finite distributions.

IV. OBTAINING THE TRIP GENERATION/ATTRACTION VECTORS

The initial stage of the classical model of urban transport planning is the one that determines the generation and attraction of trips. This stage supports the data feed to the second stage, which must forecast the distribution of trips.

For the beginning of the conformation of the trip generation and attraction matrix, it is required to have a database for the assembly of the origin-destination (O/D) base matrix. This can be achieved in two ways:

A. Personal survey:

The number of surveys that must be carried out to have a representative sample to model public transport by urban train is governed by the following equation:

$$n = \frac{Z_a^2 \cdot p \cdot q \cdot N}{e^2 \cdot (N-1) + Z^2 \cdot p \cdot q} \qquad (2)$$

n = Sample size
N = Size of the universe
Z = Confidence level
p = Probability in favor
q = Probability against
e = Sampling error

Equation (2) helps us to know the minimum number of surveys to be carried out based on demand at the public transport rush hour.

B. Database:

This way of obtaining information about the trips attracted and generated in public transport is true, because these databases arise from the entities that have the function of controlling and managing public transport. The data from these are real with respect to the total demand of urban transport means, mainly in urban trains since to use these services it is necessary to have automatic service cards.

The information has been found and the origin-destination (O/D) matrix is formed. It will be necessary to have information on the population growth of the study area (city) to know the annual growth trend of users. In addition to this information, it will be necessary to have trip attraction factors, these factors can be educational centers, recreation areas, work centers, among others.

Once all the required information has been obtained (origin-destination matrix, annual population growth, attraction factors), a regression analysis will continue with the matrices obtained. To be valid, this regression must be close to 1, so that the information to be worked on is correct or has inconsistencies between them.

Along with the regression analysis is the equation of generated trips (VG) and attracted trips (VA), which facilitates the forecast of the matrix for a certain number of years in the future. To find the target cells of the trip distribution matrix, as can be seen in Table II.

TABLE II.
O/D Matrix structure with VA and VG

O / D	Di	Di + 1	Di + 2	….	Di + n	VG
Oi						**F**
Oi + 1						**g**
Oi + 2						**h**
:						**I**
Oi + n						**k**
VA	**a**	**b**	**c**	**d**	**e**	

Source: self made

V. TRIP DISTRIBUTION ENTROPY MAXIMIZATION MODEL

The entropy maximization model was defined in subtitle III, section F. This model is based on the second law of thermodynamics, since we initially assume discrete and finite data distributions (trip information).

The modeling of trip demand in its trip distribution stage is suggested to the rest of the stages that compose it since it can be projected as a simultaneous iterative process that is carried out because the general cost of traveling between zones is given by the utility of trip in a certain mode of transport and also depends on the allocation of the existing traffic and the road network (times and distances), but by considering these variables the distribution of trips can be modified and therefore the other stages as well, repeating the cycle based on the results found in the modal distribution and the traffic allocation, returning to the distribution of trips and repeating the process until finding the balance in the trips between zones.

The goal of maximum entropy is to generate probability distributions with the given information. More precisely, among all the probability distributions that are feasible in a finite set of states under a certain type of information, we want to find the one that supports the highest level of uncertainty [6].
For the case of the equilibrium of the O/D matrices, the micro-states can be interpreted as the possible origin-destination pairs that use a certain arc of the network, and in this model, each origin-destination pair is equal to the likely one as another. Therefore, the entropy of a system is given by the following formula:

$$T_{ij} = a_i * VG_i * b_j * VA_j * \exp(-\beta * c_{ij}) \qquad (3)$$

Where:

Tij = trips from origin i to destination j
ai = Factor of generated trips
bj = Factor of attracted trips
VGi = Trips generated in i
VAj = Trips attracted in j
Exp (-β cij) = Cost function for MME model
β = Calibration parameter of MME
Cij = Cost of the trip from i to j

The MME equation (3) is used to calculate the trips from origin i to destination j, based on the trips generated at origin i, the trips attracted at destination j and an impedance that depends on an exponential function of the cost of performing the trip between origin i and destination j, and also two factor vectors ai and bj that depend on the origin and destination zones of the trip between each pair of zones i, j. It is an iterative calculation where it can be seen that there is a double recurrence as seen in (3) and (4), so it must be based on arbitrary values for one of the parameter vectors in the first iteration.

$$a_i = \frac{1}{\sum_j b_j * VA_j * \exp(-\beta c_{ij})} \quad (4)$$

$$b_i = \frac{1}{\sum_j a_i * VG_i * \exp(-\beta c_{ij})} \quad (5)$$

The process of modeling the distribution of trips, that is, the process that defines how the distribution of trips is from each origin to each destination according to the impedance that exists between these areas, can be carried out in several ways: by periods of the day and by trip purposes, by socioeconomic categories or combinations of them, among others. This depends on the need for precision in the case study and its objective.

To develop the entropy maximization modeling of the distribution of trips in a city without more information than that available for the base year in the origin/destination survey, the following steps must be carried out, having previously obtained the information from the survey already refined , expanded and organized [7].

- ✓ Classification and ordering of information
- ✓ Model calibration
- ✓ Model evaluation

A. Calibration of the entropy maximization model

Once the information has been classified for the base year, the calibration of the model begins using the following procedure.

- ✓ Using the trip matrices, the average costs (mean times) of the base year and the possible impedance function, the necessary iterations are carried out to obtain the calibration coefficients of the impedance function. In many cases the model does not converge and, therefore, it is mandatory to use the discrete functions, which have a faster convergence. With this, the cost intervals (times) are found and a frequency of trips is obtained for each of them.
- ✓ The model can converge for two or more impedance costs, so it must be chosen based on the evaluation of the model discussed below. When impedance couldn´t be based on costs because all the trips have the same ticket value. Distance was not a god option because it is the same at every time of the day. Only could vary access time in peak hours and could be a little different with other hours. So trip times were the alternative to consider in impedance function setting. In order to get this information, some samples were collected in vehicle with a mobile app traveling all the train itinerary.

B. Model performance evaluation

Once the model is calibrated, it can be estimated under present and/or future conditions, for which the following steps are required:

- ✓ Carry out the distribution of the trips with the equation of the entropy maximization model to obtain the matrix of the trips using the impedance function (continuous or discrete) obtained from the calibration.
- ✓ The impedance function as it is based on trip times, intrazonal trips are not observed in the data, as distances and trip times are zero. It is logical because nobody get into the rail system paying a ticket for no distance trips (origin and destination are the same station).
- ✓ Only interzonal trips are calculated with the formulation obtained from the iterative process.
- ✓ Statistical parameters such as R^2 are obtained to see if the modeling is significant by comparing the observed and modeled trips between zones and correlating them. In this type of model, very high R^2s are not reached, that is, a suitable value could generally be greater than 0.5.

VI. CASE STUDY: LIMA TRAIN LINE 1

For the practical case, the research proposes a more efficient forecast with respect to other forecasting models, for this it is necessary to apply the entropy maximization model of the distribution stage in a real case study. The case to be developed is Line 1 of the Lima and Callao Train, of the AATE (Autonomous Authority of the Electric Mass Transportation System of Lima and Callao).

The process of forecasting the distribution of trips in the classic transport model begins at the data collection stage, the origin/destination matrix of one-business-day trips of Line 1 of the Lima and Callao electric train. This line has 26 stations along its trajectory, these stations being: Villa el Salvador (VES), Parque Industrial (PIN), Pumacahua (PUM), Villa María (VMA), María Auxiliadora (MAU), San Juan (SJU), Atocongo (ATO), Jorge Chávez (JCH), Ayacucho (AYA), Cabitos (CAB), Angamos (ANG), San Borja Sur (SBS), La Cultura (CUL), Arriola (NAR), Gamarra (GAM), Miguel Grau (MIG), El Ángel (ELA), Presbítero Maestro (PRE), Caja de Agua (CAA), Pyramid of the Sun (PIR), Los Jardines (JAR), Los Postes (POS), San Carlos (SCA), San Martín (SMA), Santa Rosa (SRO) and Bayóvar (BAY).

The matrix in Table III represents the distribution of trips on a business day in 2019 at the pm peak hour for Line 1 of the Lima and Callao electric train. This distribution has 32,338 trips made throughout its 26 stations.

TABLE III. Trip distribution matrix in p.m. peak hour on a business day on Line 1 of the Urban Train.

		Destination																										
		VES	**PIN**	**PUM**	**VMA**	**MAU**	**SJU**	**ATO**	**JCH**	**AYA**	**CAB**	**ANG**	**SBS**	**CUL**	**NAR**	**GAM**	**MIG**	**ELA**	**PRE**	**CAA**	**PIR**	**JAR**	**POS**	**SCA**	**SMA**	**SRO**	**BAY**	**VG**
Origin	**VES**	0	6	8	16	35	37	64	99	185	333	347	201	712	167	142	515	8	19	80	63	47	7	39	35	9	22	**3196**
	PIN	8	0	5	10	16	10	23	47	73	155	138	65	296	72	58	166	1	6	42	26	11	4	9	7	6	5	**1259**
	PUM	9	1	0	10	4	10	37	64	92	127	152	79	290	79	60	214	4	1	47	30	14	2	18	17	3	8	**1372**
	VMA	21	7	3	0	2	8	47	72	89	166	229	108	373	84	80	260	1	6	42	27	12	8	16	13	5	9	**1688**
	MAU	23	19	8	3	0	4	25	49	69	150	171	95	409	113	98	331	4	12	55	33	12	6	21	18	6	19	**1753**
	SJU	37	8	16	1	6	0	12	19	56	72	90	95	287	70	76	247	0	7	61	39	17	6	21	16	6	8	**1273**
	ATO	38	27	20	17	5	7	0	4	5	29	42	62	229	57	66	210	4	12	41	32	23	5	29	27	11	23	**1025**
	JCH	30	19	13	17	26	8	0	0	1	4	6	30	168	31	24	98	1	3	25	25	25	7	16	14	5	5	**601**
	AYA	44	23	9	13	32	7	7	0	0	1	13	13	100	24	13	50	6	9	20	17	14	3	11	11	6	20	**466**
	CAB	25	21	17	16	19	9	5	2	0	0	2	8	53	14	17	40	4	10	31	24	18	7	18	12	13	21	**406**
	ANG	44	29	17	31	29	22	8	4	1	2	0	2	21	14	25	83	5	10	41	38	36	10	27	27	16	34	**576**
	SBS	25	10	7	12	8	4	7	1	5	3	1	0	3	3	5	39	0	2	22	19	16	3	8	6	12	8	**229**
	CUL	79	38	27	43	55	45	34	39	33	37	14	5	0	1	13	43	1	3	41	42	44	17	46	34	31	51	**816**
	NAR	38	12	4	19	21	18	16	23	17	22	16	6	3	0	8	12	2	4	27	23	19	13	45	34	15	34	**451**
	GAM	23	9	5	7	6	3	15	13	14	18	24	4	10	5	0	1	2	0	21	22	23	11	31	21	11	23	**322**
	MIG	93	54	28	31	53	43	45	51	39	76	114	64	101	15	5	0	0	4	10	17	25	10	47	58	26	46	**1055**
	ELA	4	2	0	2	2	0	4	3	7	9	10	6	13	3	0	1	0	0	4	6	8	3	12	9	4	7	**119**
	PRE	15	6	5	5	1	2	10	9	18	33	48	24	44	5	2	2	0	0	19	15	11	9	16	21	10	12	**342**
	CAA	75	38	10	23	38	31	47	72	63	187	261	130	356	86	101	71	2	9	0	3	6	7	19	20	20	32	**1707**
	PIR	57	26	9	19	26	25	55	63	65	188	265	118	347	112	134	105	1	8	3	0	3	3	11	13	10	23	**1689**
	JAR	39	15	8	16	14	20	64	55	68	189	269	106	339	138	167	140	0	7	6	3	0	0	3	7	1	15	**1689**
	POS	16	14	5	9	9	14	46	46	57	115	224	78	221	123	137	128	7	13	10	6	2	0	1	3	0	2	**1286**
	SCA	34	19	9	8	14	13	60	41	49	187	227	90	401	125	160	149	4	11	14	8	2	0	0	0	0	2	**1627**
	SMA	21	15	3	7	18	12	46	30	27	135	215	76	306	116	119	185	9	3	45	29	14	2	2	0	0	0	**1435**
	SRO	19	12	4	11	16	22	58	35	52	149	279	104	389	145	176	218	8	14	61	39	18	1	1	1	0	2	**1834**
	BAY	67	27	10	23	36	35	144	75	144	398	574	215	777	357	407	529	8	30	121	82	44	7	10	2	0	0	**4122**
	VA	**884**	**457**	**250**	**369**	**491**	**409**	**879**	**916**	**1229**	**2785**	**3731**	**1784**	**6248**	**1959**	**2093**	**3837**	**82**	**203**	**889**	**668**	**464**	**151**	**477**	**426**	**226**	**431**	**32338**

TABLE IV.
Trip parameters on Line 1 of the Lima and Callao Electric Train taken in the field.

FROM NORTH TO SOUTH

Departure Time	STATION	DISTANCE (Km)	TIME	TIME (min)	V half (km/h)	V max (km/h)
07:27:03	BAY					
	↓	1.15	00:01:48	1.80	36.3	68.6
07:28:59	SRO		00:00:08	0.13		
	↓	1.04	00:01:40	1.67	35.4	85.4
07:30:50	SMA		00:00:11	0.18		
	↓	1.42	00:01:54	1.90	44.7	70.3
07:32:54	SCA		00:00:10	0.17		
	↓	1.43	00:01:55	1.92	44.7	70
07:35:00	POS		00:00:11	0.18		
	↓	1.31	00:01:53	1.88	41.6	79.2
07:37:09	JAR		00:00:16	0.27		
	↓	1.4	00:02:10	2.17	37	62.5
07:39:30	PIR		00:00:11	0.18		
	↓	1.46	00:02:11	2.18	39.9	69.1
07:41:53	CAA		00:00:12	0.20		
	↓	1.64	00:02:18	2.30	40.2	66.2
07:44:26	PRE		00:00:15	0.25		
	↓	0.55	00:01:16	1.27	26	51.9
07:45:52	ELA		00:00:10	0.17		
	↓	0.997	00:02:02	2.03	29.2	57.5
07:48:04	MIG		00:00:10	0.17		
	↓	1.35	00:02:03	2.05	19.4	72.9
07:50:15	GAM		00:00:08	0.13		
	↓	1.01	00:01:39	1.65	36.6	67.5
07:52:01	NAR		00:00:07	0.12		
	↓	1.66	00:02:49	2.82	35.4	62.7
07:54:59	CUL		00:00:09	0.15		
	↓	1.56	00:02:19	2.32	40.3	72.1
07:57:25	SBS		00:00:07	0.12		
	↓	1.12	00:01:48	1.80	37	63.6
07:59:21	ANG		00:00:08	0.13		
	↓	1.95	00:02:39	2.65	42.8	74.5
08:02:08	CAB		00:00:08	0.13		
	↓	0.973	00:01:50	1.83	30.8	51.1
08:04:08	AYA		00:00:10	0.17		
	↓	1.09	00:01:49	1.82	35.9	66.6
08:06:04	JCH		00:00:07	0.12		
	↓	1.59	00:02:00	2.00	47.5	80
08:08:11	ATO		00:00:07	0.12		
	↓	1.59	00:02:58	2.97	32.2	69.5
08:11:20	SJU		00:00:11	0.18		
	↓	1.2	00:01:47	1.78	40.4	69.1
08:13:18	MAU		00:00:11	0.18		
	↓	1.08	00:01:37	1.62	40	74.1
08:15:10	VMA		00:00:15	0.25		
	↓	1.43	00:02:11	2.18	39	65.3
08:17:29	PUM		00:00:08	0.13		
	↓	1.84	00:02:18	2.30	47.8	78.2
08:19:54	PIN		00:00:07	0.12		
	↓	1.39	00:02:03	2.05	40.6	78.6
	VES					
		total travel time	00:54:54	54.90		

FROM SOUTH TO NORTH

Departure Time	STATION	DISTANCE (Km)	TIME	TIME (min)	V half (km/h)	V max (km/h)
06:30:53	VES					
	↓	1.41	00:02:13	2.22	35.9	69.6
06:33:16	PIN		00:00:10	0.17		
	↓	1.86	00:02:25	2.42	46	72.6
06:35:51	PUM		00:00:10	0.17		
	↓	1.44	00:02:08	2.13	40	67.7
06:38:06	VMA		00:00:07	0.12		
	↓	1.08	00:01:47	1.78	36.1	75.9
06:40:00	MAU		00:00:07	0.12		
	↓	1.24	00:01:55	1.92	37.8	71
06:42:04	SJU		00:00:09	0.15		
	↓	1.61	00:02:10	2.17	44.3	79.3
06:44:23	ATO		00:00:09	0.15		
	↓	1.65	00:01:55	1.92	49.8	81.2
06:46:28	JCH		00:00:10	0.17		
	↓	1.11	00:01:38	1.63	38.8	81.7
06:48:14	AYA		00:00:08	0.13		
	↓	1	00:01:47	1.78	32.5	76
06:50:08	CAB		00:00:07	0.12		
	↓	1.97	00:02:20	2.33	48.5	78.9
06:52:37	ANG		00:00:09	0.15		
	↓	1.12	00:01:41	1.68	37.9	75
06:54:29	SBS		00:00:11	0.18		
	↓	1.51	00:02:09	2.15	40.6	78.9
06:56:46	CUL		00:00:08	0.13		
	↓	1.69	00:02:30	2.50	40.1	68.8
06:59:23	NAR		00:00:07	0.12		
	↓	1	00:01:37	1.62	37	78.1
07:01:07	GAM		00:00:07	0.12		
	↓	1.34	00:01:57	1.95	41.3	71.3
07:03:11	MIG		00:00:07	0.12		
	↓	1.01	00:01:47	1.78	33	58.4
07:05:05	ELA		00:00:07	0.12		
	↓	0.532	00:01:08	1.13	26.7	54.3
07:06:24	PRE		00:00:11	0.18		
	↓	1.63	00:02:17	2.28	42.7	66.8
07:08:50	CAA		00:00:09	0.15		
	↓	1.42	00:01:59	1.98	42.7	81.2
07:11:01	PIR		00:00:12	0.20		
	↓	1.37	00:01:48	1.80	44	80.3
07:12:58	JAR		00:00:09	0.15		
	↓	1.31	00:01:41	1.68	43.6	81.3
07:14:48	POS		00:00:09	0.15		
	↓	1.43	00:01:51	1.85	44	79.9
07:16:50	SCA		00:00:11	0.18		
	↓	1.42	00:01:46	1.77	46.2	79.7
07:18:44	SMA		00:00:08	0.13		
	↓	1.05	00:01:30	1.50	41.4	78.1
07:20:21	SRO		00:00:07	0.12		
	↓	1.15	00:01:43	1.72	40.45	74.42
	BAY					
		total travel time	00:51:11	51.18		

It is also necessary to know the population (for trip generation model) and current educational vacancies (for trip attraction model), and their growth factors for 2025 as shown in tables V and VI.

TABLE V.
Data on the 2019 population and vacancies in the education system in the areas of influence of each station

WHEREABOUTS	POPULATION	VACANT SYST. EDUCATIONAL
VES	55227	12476
PIN	48323	10916
PUM	23461	5726
VMA	20109	4908
MAU	25846	6290
SJU	36777	8921
ATO	31260	7583
JCH	13661	3740
AYA	15179	4156
CAB	22263	6095
ANG	22988	4465
SBS	28735	5582
CUL	34482	6698
NAR	28776	7415
GAM	33854	8723
MIG	24287	12145
ELA	29755	5819
PRE	26986	13494
CAA	48796	10944
PIR	91348	20488
JAR	74697	16753
POS	53652	12033
SCA	84294	18906
SMA	97129	21784
SRO	42205	9466
BAY	98286	22044

The growth percentages were found by conducting a study with past data. Projecting the population growth of 1995, 2000, 2005, 2010, 2015 and 2020 of the Technical Directorate of Demography and Social Studies. These percentages are shown in Table VI, which serves to suggest the expected increase in the trips generated and attracted by each station.

TABLE VI.
Matrix of percentage growth of the Population to 2025 and the educational vacancies in the areas of influence of each station

WHEREABOUTS	POPULATION	VACANT SYST. EDUCATIONAL
VES	10.27%	13.35%
PIN	10.27%	13.35%
PUM	7.97%	11.16%
VMA	7.97%	11.16%
MAU	9.60%	9.88%
SJU	9.60%	8.50%
ATO	9.60%	8.50%
JCH	7.19%	-0.98%
AYA	7.19%	-0.98%
CAB	7.19%	-0.98%
ANG	9.25%	-7.38%
SBS	9.25%	-7.38%
CUL	9.25%	-7.38%
NAR	-1.50%	8.97%
GAM	-1.50%	8.97%
MIG	-1.88%	0.40%
ELA	3.24%	14.35%
PRE	-1.88%	0.40%
CAA	12.20%	14.30%
PIR	12.20%	14.30%
JAR	12.20%	14.30%
POS	12.20%	14.30%
SCA	12.20%	14.30%
SMA	12.20%	14.30%
SRO	12.20%	14.30%
BAY	12.20%	14.30%

Using the population as an explanatory variable of the trips generated and the vacancies in the educational system as the explanatory variable of the trips attracted, a regression was applied to the data to obtain the equations of generated trips (VG) and attracted trips (VA) , obtaining equations (6) and (7) with a reliability of 78% of the adjusted R^2.

$$VGi = 0.03 * Población\ i \quad (6)$$

$$VAj = 0.08 * Vacant.Sist.Educ.j \quad (7)$$

With these equations we find the generated and attracted trips projected to 2025 as shown in tables VII and VIII.

TABLE VII.
Future generated trips in each station of Line 1 of the Lima and Callao electric train by 2025

	VG	POPULATION	VG of the model	VG futures
VES	3196	55227	1510	1666
PIN	1259	48323	1322	1457
PUM	1372	23461	642	693
VMA	1688	20109	550	594
MAU	1753	25846	707	775
SJU	1273	36777	1006	1102
ATO	1025	31260	855	937
JCH	601	13661	374	400
AYA	466	15179	415	445
CAB	406	22263	609	653
ANG	576	22988	629	687
SBS	229	28735	786	859
CUL	816	34482	943	1030
NAR	451	28776	787	775
GAM	322	33854	926	912
MIG	1055	24287	664	652
ELA	119	29755	814	840
PRE	342	26986	738	724
CAA	1707	48796	1334	1497
PIR	1689	91348	2498	2803
JAR	1689	74697	2043	2292
POS	1286	53652	1467	1646
SCA	1627	84294	2305	2586
SMA	1435	97129	2656	2980
SRO	1834	42205	1154	1295
BAY	4122	98286	2688	3016
	32338		30421.5917	33316

TABLE VIII.
Trips attracted to the future in each station of Line 1 of the Lima and Callao electric train by 2025

	GOES	VACANCIES IN THE EDUCATIONAL SYSTEM	Model VA	VA futures	VA ADJUSTED futures
VES	884	12476	959	1087	1603
PIN	457	10916	839	951	1402
PUM	250	5726	440	489	722
VMA	369	4908	377	420	619
MAU	491	6290	484	531	783
SJU	409	8921	686	744	1097
ATO	879	7583	583	633	933
JCH	916	3740	288	285	420
AYA	1229	4156	320	316	466
CAB	2785	6095	469	464	684
ANG	3731	4465	343	318	469
SBS	1784	5582	429	398	587
CUL	6248	6698	515	477	703
NAR	1959	7415	570	621	916
GAM	2093	8723	671	731	1078
MIG	3837	12145	934	938	1383
ELA	82	5819	447	512	755
PRE	203	13494	1038	1042	1536
CAA	889	10944	842	962	1418
PIR	668	20488	1575	1801	2655
JAR	464	16753	1288	1472	2170
POS	151	12033	925	1058	1560
SCA	477	18906	1454	1662	2450
SMA	426	21784	1675	1915	2824
SRO	226	9466	728	832	1227
BAY	431	22044	1695	1937	2856
	32338		20574	22596	33316

With the trip vectors generated and future attracted projected to 2025, we proceeded to estimate the distribution of trips between stations. This matrix has the same format that Table III.

CONCLUSIONS

The entropy maximization model is a model that, through its impedance function, makes it a much more realistic model in cases of travel distribution, since it considers the opposition of unrealistic circumstances.

In this research, trip time was used as the cost of the trip, which is a good approach to what the user use to compare between different modes. Monetary costs were not a good option for impedance measure because of flat rate, and distance neither because is the same value in peak hours and in other periods of the day.

For the application of MME to the classic trip transport model, it is necessary to know socioeconomic factors, since they affect the mobility probability of the users of some urban train.

The calibration of β parameter in the MME for the forecasting of trip distribution implies the necessity of finding it through the comparison of historical data in such a way that there is little dispersion between the results.

In the case study of Line 1 of the Lima and Callao Electric Train, the growth in demand was demonstrated for 2025 and in which stations there was an increase in demand, such as the San Martin station (SMA), with 5804 users in one business day.

REFERENCES

[1] SBAI, A., TOUL, A. and GHADI, F. (2018). Dynamic extension of the entropy-based gravity model for travel distribution purposes. Journal of Theoretical and Applied Information Technology, 96 (24), 8183-8193.

[2] FURLONGE, J. (2019). A probability Bias Model of Trip Distribution. Journal of transport and development, 258-268. DOI: 10.1061 / 9780784482575.025

[3] BEHARA, K., BHASKAR, A. and CHUNG, E. (2020). A novel approach to the structural comparison of origin-destination matrices: Levenshtein distance. Transportation Research Part C: Emerging Technologies, 111, 513-530. DOI: 10.1016 / j.trc.2020.01.005

[4] WILLIAMS, X. (1996). Introduction to quantitative models for administration. Grupo Editorial Iberoamérica SA Mexico

[5] ORTÚZAR, J, WILLUMSEN, L. (2000). Transport demand models. 2nd Edition. Mexico. Editions Universidad Católica de Chile, Alfaomega.

[6] LÓPEZ, H. (2013). Generalized interval cost and entropy maximization model for urban trip distribution. University of Bogotá, Colombia.

[7] CÁRDENAS, D, COLOMER, J. (2003) Model of geographical distribution of trips, based on the relative weight of tractors. Memories VI Symposium on Traffic and Transportation Engineering. University of Cauca. Popayan. Colombia.