INTRAZONAL TRIP DISTANCES: AN ESTIMATION APPROACH AND APPLICATION TO A CASE STUDY

Conrado Vidotte Plaza[1]
Antônio Nélson Rodrigues da Silva[2]
UNIVERSITY OF SÃO PAULO
São Carlos School of Engineering
Av. Trabalhador São-carlense, 400, 13566-590, São Carlos - SP - BRAZIL
e-mail (1): conradoplaza@gmail.com
e-mail (2): anelson@sc.usp.br

2015 Transportation Research Board Annual Meeting

Word count: 4,596 words + 5 table and 4 figures (250 words each) = 6,848 words

ABSTRACT

This study aimed at the development of procedures for the estimation of intrazonal trip distances based on data of origin-destination (O/D) surveys. In the application discussed here, the analysis was based on an O/D survey conducted in the city of São Carlos, Brazil. The urban area was subdivided several times so as to produce 471 traffic analysis zones (TAZ). The geometrical characteristics of the zones that might influence the distances of intrazonal trips were identified. Numerical indicators of the selected geometrical patterns were compared with the average values of intrazonal trip distances to search for evidences of correlation between these variables. Two analytical approaches were explored: i) continuous and ii) discontinuous. In the first case, all trips were considered a single dataset, whereas in the second the dataset was split into two subsets of homogeneous characteristics. The datasets with higher correlation were used to build global and individual models for the estimation of intrazonal trip distances. In the global models, walking, cycling, and auto trips were considered part of a common dataset. The three transport modes were separately taken into account in the stratified models. The values obtained with these models were then compared with the estimations of other models found in the literature. The discontinuous models calibrated in this study clearly outperformed the traditional ones, which may be an indication they can be used to replace the traditional models for the estimation of average intrazonal trip distances, at least in Brazilian medium-sized cities.

1. INTRODUCTION

The information required for the description of urban journeys is associated with data sometimes not easily accessed. Moreover, it is frequently imprecise and, in many cases, unavailable. The average trip distances, for example, depend on characteristics of the street network and the trip distribution patterns. As a consequence, they can be treated as a variable influenced by the environment, the user profile, and the trip characteristics. Several procedures for the estimation of trip distances between zones can be found in the transportation planning literature. In the case of intrazonal trip distances, however, the situation is not the same. Most methods found in the literature are either outdated or based on empirical studies that refer to specific contexts.

This study aimed at the development of procedures for the estimation of intrazonal trip distances based on data of origin-destination (O/D) surveys. Such procedures can be quite useful to the transportation planning process, particularly if they are straightforward and based on readily available data.

This manuscript is organized as follows: Section 2 contains a brief theoretical discussion on intrazonal trips and reviews models developed for the estimation of trip distances, particularly those that can be applied to intrazonal trips; Section 3 describes the methodology, with a strong emphasis on the approach used for the acquisition of data; Section 4 is devoted to the main results achieved by the models developed and a comparison of the results provided by other models found in the literature review; Sections 5 and 6 address the conclusions and the references, respectively.

2. THEORETICAL BACKGROUND

Mobility is an essential condition for the development of cities. According to the U.S. Department of Commerce [1], urban trip patterns are essentially influenced by the transportation infrastructure available, the activity patterns in the area, and the socioeconomic characteristics of the population. In the specific case of intrazonal trips, they are also a direct consequence of the trip generation rates (of both, trip productions and trip attractions) within the zones. Several studies ([2], [3], [4], [5], [6], [7], and [8]) have focused on the influence of the environment on the trip behavior, but they are not conclusive regarding the conditions of intrazonal trips.

The physical characteristics of the zones are essential for the choice of the transportation mode used for intrazonal trips. According to Greenwald [9], the modal choice is influenced by the urban form, which also affects the regional distribution of trips. For Bhatta and Larsen [10], intrazonal trips often rely on non-motorized modes (walking and cycling) because distances within zones are usually short.

In the creation of origin-destination (O/D) matrices, the intrazonal trips are sometimes ignored, because the data required for their estimation are not available or do not exist. For Ortúzar and Willumsen [11], unless the intrazonal trips can be estimated with simple approaches, they should be removed from the modeling process. Bhatta and Larsen [10], in contrast, state that the intrazonal trips cannot be ignored, due to the impacts they have on important aspects of transportation, such as congestion and pollution, at a local level. When these trips are not considered in the estimation models, the output is a reduced sample that may not represent all trips. Therefore, the estimation of the parameters can be biased particularly if the trips are not concentrated around the zone centroid.

In one of the first methods to estimate intrazonal trip distances, which was developed by the U.S. Department of Commerce [12], the estimated distance would be the average distance between the centroid of a zone and the centroids of adjacent zones divided by two. In a similar approach, Venigalla et al. [13] suggested that the average intrazonal trip distance of any zone would be half the distance between the centroid of the zone and the centroid of the closest zone. In the approach proposed by Smeed [14], which was based on hypothetical and actual grid-shaped street networks, the average intrazonal trip distance ($\overline{D}$) depends only on the area (*A*) of the zone (Equation 1). Other models found in the literature were proposed by Batty [15] (Equations 2 and 3), and Fotheringham [16] (Equation 4). In Equation 4, the radius of the zone (*r*) and the distance between the centroids of neighboring zones (*z*) replace the area of the zone. In the case of intrazonal trips, however, when the value of *z* is zero, the equation can also be written as a function of *A* (Equation 5).

$$\overline{D} = 0{,}81 * A^{1/2} \tag{1}$$

$$\overline{D} = \frac{r}{\sqrt{2}} \tag{2}$$

$$\overline{D} = \sqrt{\frac{A}{2 * \pi}} \tag{3}$$

$$\overline{D} = 0{,}846 * (1{,}693)^{z/r} * r \tag{4}$$

$$\overline{D} = 0{,}846 * \sqrt{A/\pi} \tag{5}$$

Kordi et al. [17] developed a method for estimating the average trip lengths of intrazonal flows by scattering the origins and destinations of the flows within their zones. The origins and destinations of the flows were distributed in two ways: randomly and based on an available spatial density distribution. The average trip length was then calculated for all possible trip configurations. The scattered-based models and some existing assumption-based models were applied to a Swiss journey-to-work dataset and the results were compared. The density-based scattering models showed a better model fit and smaller errors.

In general, the methods for the estimation of intrazonal trip distances are based on empirical studies that should not be applied without adjustments. The present study makes a contribution to research on this field by the development and test of a new approach for the estimation of intrazonal trip distances.

3. METHOD

The procedure for the estimation of intrazonal trip distances in urban areas developed in this study involved four steps: i) identification of the variables to be considered and data collection, ii) development of scenarios and calculation of values of interest for each scenario, iii) selection of strategies for data analysis and modeling, and iv) validation of the proposed models.

3.1. Data requirements

Some of the factors that can affect the actual trip distances are related to the road infrastructure and the urban form and distribution of land uses. The following factors were selected from the literature to be applied as variables in the present study:

- Area of each study area (AREA);
- Perimeter of each study area (PERIMETER);
- SHAPE FACTOR of each study area, as proposed by [4];
- CIRCULAR SHAPE FACTOR of each study area, as proposed by [5];
- STUDY AREA/CIRCLE AREA RATIO, or the area of each study area divided by the area of the smallest circle that circumscribes it;
- STUDY AREA/RECTANGLE AREA RATIO or the area of each study area divided by the area of the smallest rectangle that best fits it;
- NS/EW RATIO, or the larger measurement in the East-West direction (E-W) divided by the larger measurement in the North-South direction (N-S) of each study area;
- DENSITY OF NETWORK CONNECTIONS, or number of endpoints in the network per km^2 of each study area;
- STREET DENSITY (length of the streets per km^2 of each study area).

The data required for the calculation of the elements aforementioned (Table 1) were acquired by an analysis of the scenarios, as described in the sequence. The databases applied, which shall be preferably georeferenced in a GIS environment, must contain the following elements of the city under analysis:

- Street network;
- Locations of trip origins and destinations, as usually found in most GIS-based O/D surveys;
- Set of homogeneous zones, such as TAZ, or Traffic Analysis Zones.

3.2. Development of scenarios

The scenarios were created in this phase to replicate different shapes and sizes of the zones considered in an urban setting. These urban zones (or study areas) that form the scenarios are used for obtaining the data required for the next steps. In general, a large number of scenarios is interesting, because it will result in larger datasets for the calibration of the models.

Two procedures can be used to build the scenarios: one manual and one supported by computer software. The manual procedure starts with the creation of a large area that encompasses the entire city. The following steps consist in a progressive subdivision of the initial area until the resulting zones have matched the zones considered references (e.g., TAZ). This subdivision process can follow the historical evolution of the city while considering the existing physical barriers. The computer-aided procedure can rely on GIS programs. In the case of this study, we used the *Regional Partitioning* procedure available in TransCAD 5.0 [20]. It enables the creation of contiguous, compact and balanced areas by the aggregation of smaller zones, which is the opposite of the first procedure. The balancing process can be based on either information on the zones (area and population, for example) or the outcomes of the procedure (such as number of study areas that constitute the new zones).

In the analysis of the scenarios, each new study area is taken into account in the calculation of the factors that may bear a relationship with the average intrazonal trip distances (listed in Table 1), according to Equation (6). A comparison of the values associated with the different factors enables the identification of trends and relationships between the factors and the average trip distances within the study areas.

TABLE 1 Data required for the study of the influence of factors related to urban form and road infrastructures on intrazonal trip distances

Data	Unit	Definition
Area of a zone	km²	Total land area covered by the study area
Perimeter of a zone	km	Measurement of outline (or boundary) of the land area covered by the study area
Diameter of the circumscribed circle	km	Diameter of the smallest circle that circumscribes the study area
Area of the circumscribed circle	km²	Area of the smallest circle that circumscribes the study area (given by $A = \pi * D^2/4$)
E-W Length	km	Larger measurement in the east-west direction (E-W)
N-S Length	km	Larger measurement in the north-south direction (N-S)
Area of the smallest rectangle around the zone	km²	Area of the smallest rectangle that best fits the study area (given by E-W Length multiplied by N-S Length)
Number of trips	number	Number of trips inside the study area
Travel distance in each trip	km	Distance traveled in each journey
Network connections	number	Number of endpoints (intersections or link's endpoints) in the network of the study area
Street extension	km	Linear extension of the streets inserted in the study area

$$\overline{D}_{obs} = \frac{\sum D}{n} \qquad (6)$$

where:

$\overline{D}_{obs}$ is the average trip distance within a study area, D is the distance of each trip and n is the number of trips in the same study area.

3.3. Strategies for data analysis and modeling

A first strategy of analysis can be based simply on a visual evaluation of scatterplots in which the values of each investigated factor are plotted against the average intrazonal trip distances. In general, trends can be easily identified with these visual comparisons. If the points in the graph follow a clear pattern, different mathematical functions (e.g., linear, exponential, or logarithmic) can be used to represent such trends. In this case, the adjustment of functions to represent the points in the graph can be used as models for the estimation of average intrazonal trip distances as a function of the studied variables. The selection of the models that better represent the actual data trends is based on the value of the coefficient of determination (R^2).

Two scales were considered in the analysis of the intrazonal trips. The first is an aggregate scale (for a global analysis, as it is called here), on which all the trips are considered together and the transport mode of each trip is not identified. In the second case, the trips are stratified according to the mode (i.e. auto trips, cycling trips, and walking trips) for the analysis. Furthermore, additional models can be obtained in the case of the global models for the estimation of trips per mode.

3.4. Validation of the models

The deviations found in the comparison of the values estimated with the models and the actual distances observed in the O/D survey can be used for the evaluation of the models in the calibration phase. The smaller the deviations, the better the model for transportation planning purposes. In this study, the average values and standard deviation values were used in the analysis of the performance of the proposed models. As usual, the models were validated with a different dataset. Their performance was further tested through a comparison with the estimation errors of other models found in the literature ([12], [13], [14], [15], and [16]).

4. RESULTS AND DISCUSSIONS

As the results of this study were obtained in São Carlos, Brazil, this section starts by briefly describing the city. In the sequence, we describe how the scenarios were built in that particular case and used to generate the data for the analyses presented in the following subsections.

4.1. General characteristics of the studied city

Located in São Paulo state, São Carlos has approximately 220,000 inhabitants, according to the most recent official census [18]. Host of two important public universities and several high-technology companies, the city is known as an important scientific and technological hub in Brazil. However, the fast urban growth process in the latest decades has generated a strong pressure for infrastructure, particularly in the transportation sector. As a result, many essential urban infrastructures have been implemented without a proper planning.

In the case of transport planning, for example, only in 2007 the city conducted its first O/D survey, which was not an initiative of the municipality, but of a research group of the University of São Paulo at São Carlos [19]. Approximately 4,000 households have been visited for the interviews and more than 19,000 trips have been registered. However, only some of these trips (3,356 walking trips, 2,631 car trips, and 340 bicycle trips) had both the origin and the destination georeferenced in a Geographic Information System.

The city has been divided into 41 TAZs, as shown in Figure 1. Its streets follow a regular grid pattern in most of the urban area, even in some parts where the terrain is not flat. These hilly segments may be important for studies like this one, because of their influence on the use of non-motorized modes (i.e. walking and cycling).

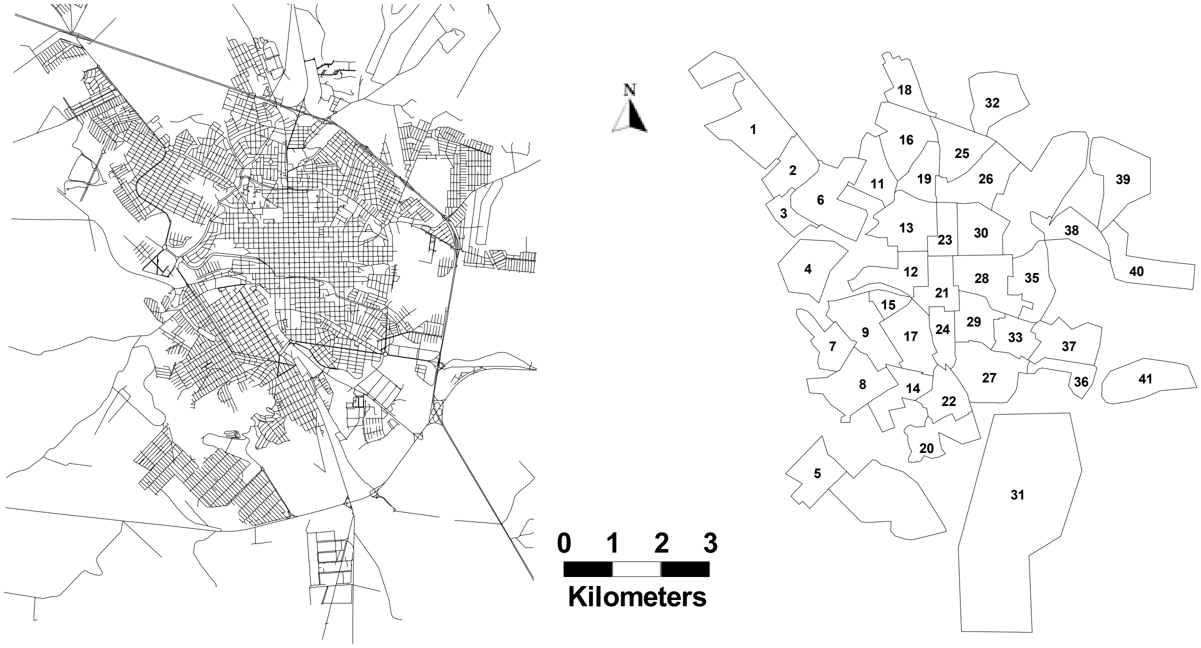


FIGURE 1 Street network and Traffic Analysis Zones of São Carlos, Brazil

4.2. Development of scenarios

The Traffic Analysis Zones used in the application discussed in this paper (Figure 1) have been used in São Carlos for at least a decade. The creation of scenarios involved the two procedures described in subsection 3.2, i.e. manual and supported by computer software. The manual procedure was essentially based on the actual areas occupied by the developments implemented in the city throughout the years. The computer-aided procedure was based on the data of population and area of the study areas and number of study areas that form new zones. The 163 scenarios built in this step of the process have generated 471 study areas for analysis.

The study areas were then organized into two groups. The first comprises zones with no intrazonal trips (or NoIT) and was formed by the study areas that had either none or only one intrazonal trip. Conversely, the second group comprehends intrazonal trips (IT) and was formed by the study areas that had two or more intrazonal trips. These groups were formed for each transport mode under analysis.

In the next step, the intrazonal trips of group IT were randomly split into two subsets: 70% were selected for analysis and calibration of the models, whereas 30% were kept aside for the validation of the models. The resulting groups were called $IT_{70\%}$ and $IT_{30\%}$, respectively. The study areas with no trips or with only one trip were not considered in the analysis because our focus is not on the reasons that influence the modal choice.

TABLE 2 Distribution of the 471 study areas into groups

Group		Number of intrazonal trips per mode		
		Walking	Bicycle	Automobile
NoIT		15	72	35
IT	$IT_{70\%}$	319	279	305
	$IT_{30\%}$	137	120	131

4.3. Data analysis and modeling

The analysis of data started with group $IT_{70\%}$ and followed two different strategies. In the first, all data in the group were used for the calibration of continuous models and in the second, the dataset with all data was filtered so as to create subgroups, which were subsequently used for the building of discontinuous models.

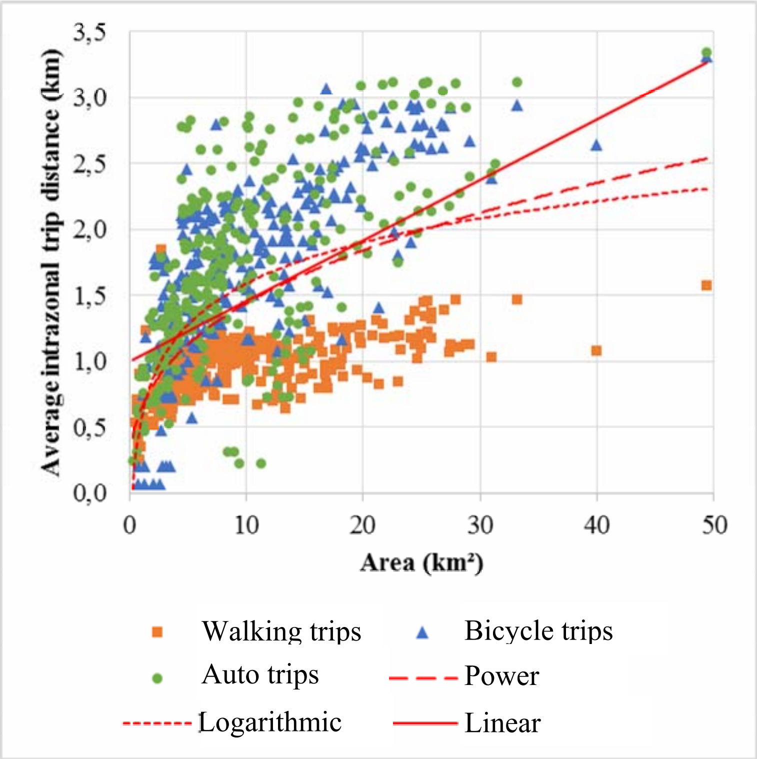
Initially, we tried to identify any strong or moderate correlations between the average intrazonal trip distances and the variables related to urban form and road infrastructures selected for investigation. These variables are listed in Table 3, along with the correlation coefficients found. The variables of highest correlation coefficients were AREA and PERIMETER, but in this study we analyzed specifically the results related to AREA.

TABLE 3 Correlation coefficient values (R) between the average intrazonal trip distances and selected variables

Variables	Datasets			
	All	Walking trips	Cycling trips	Auto trips
AREA	**0.524**	**0.622**	**0.746**	**0.652**
PERIMETER	**0.535**	**0.602**	**0.735**	**0.697**
SHAPE FACTOR	0.486	0.516	0.643	0.653
CIRCULAR SHAPE FACTOR	0.486	0.516	0.643	0.653
STUDY AREA/CIRCLE AREA RATIO	-0.142	-0.157	-0.155	-0.237
STUDY AREA/RECTANGLE AREA RATIO	-0.186	-0.182	-0.226	-0.270
NS/EW RATIO	0.079	0.130	0.149	0.033
DENSITY OF NETWORK CONNECTIONS	0.073	0.047	-0.105	0.240
STREET DENSITY	0.106	0.090	-0.086	0.286

4.3.1. Continuous models

The first procedure conducted was the building of a scatterplot with the values of average intrazonal trip distance and area per zone. The linear, power, and logarithmic functions were then adjusted to represent the data trends. A first attempt was made with the entire dataset (Figure 2), followed by an analysis stratified by transport mode (Figure 3). A power function provided the best fit in the case of the global analysis and walking trips. In the cases of cycling trips and auto trips, a logarithmic function produced the best adjustment.



$D = 0.046*A + 0.9950$

$R^2 = 0.2744$

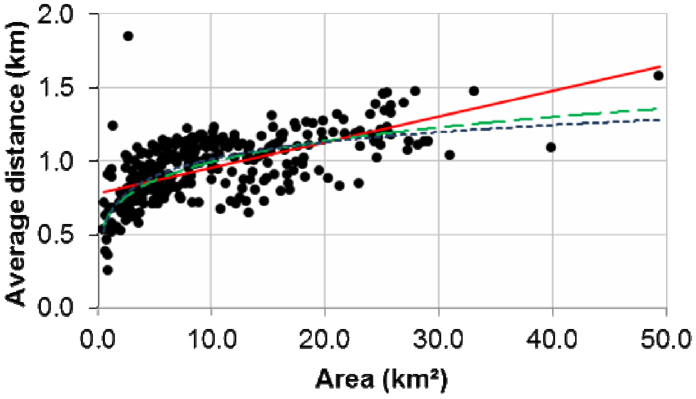
$D = 0.6378*A^{0.3541}$

$R^2 = 0.3241$

$D = 0.4503*\ln(A) + 0.5536$

$R^2 = 0.3112$

FIGURE 2 Scatterplot of *area x average intrazonal trip distance* and theoretical functions adjusted to the entire dataset (global models)


Walking trips

Linear

$D = 0.0174*A + 0.7835$

$R^2 = 0.3871$

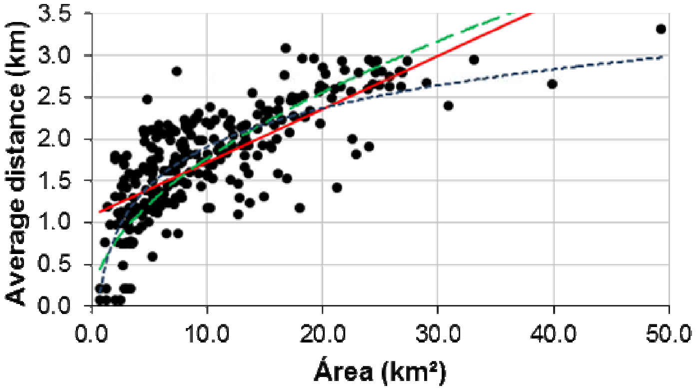
Power

$D = 0.6334*A^{0.1947}$

$R^2 = 0.4736$

Logarithmic

$D = 0.1674*\ln(A) + 0.6264$
$R^2 = 0.4516$


Cycling trips

Linear

$D = 0.0641*A + 1.0857$

$R^2 = 0.5569$

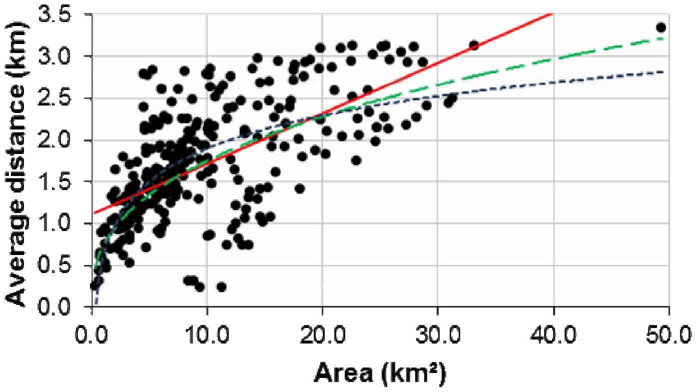
Power

$D = 0.5219*A^{0.5305}$

$R^2 = 0.4849$

Logarithmic

$D = 0.6699*\ln(A) + 0.3628$
$R^2 = 0.6217$


Auto trips

Linear

$D = 0.0605*A + 1.1179$

$R^2 = 0.4250$

Power

$D = 0.7297*A^{0.3803}$

$R^2 = 0.4299$

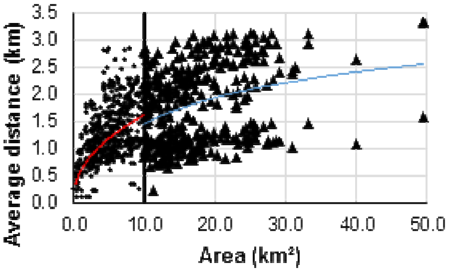
Logarithmic

$D = 0.5712*\ln(A) + 0.5823$
$R^2 = 0.4687$

FIGURE 3 Scatterplot of *area x average intrazonal trip distance* and theoretical functions adjusted to datasets stratified by transport mode

4.3.2. Discontinuous models

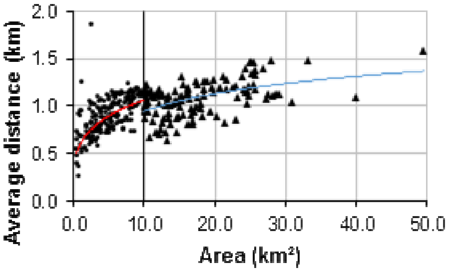
The data of group $IT_{70\%}$ were analyzed to search for the points in which the data trends, regarding the area of the zones, had changed. Three values of area were tested as transition points: 5, 10 and 15 km^2. The best fit was found with a transition point equal to 10 km^2, as shown in Figure 4.



All trips

A < 10 km²: $D = 0.5648 \cdot A^{0.4558}$ ($R^2 = 0.2989$)

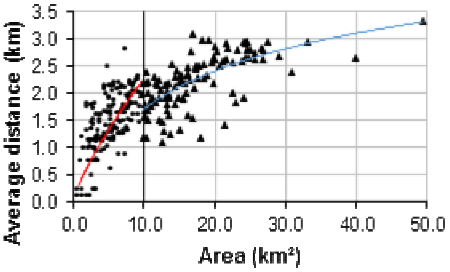
A ≥ 10 km²: $D = 0.6882 \cdot \ln(A) - 0.1319$ ($R^2 = 0.1045$)



Walking trips

A < 10 km²: $D = 0.5970 \cdot A^{0.2494}$ ($R^2 = 0.4318$)

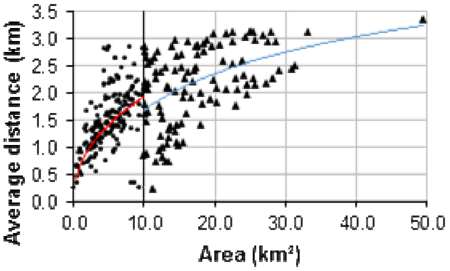
A ≥ 10 km²: $D = 0.2651 \cdot \ln(A) + 0.3334$ ($R^2 = 0.2493$)



Cycling trips

A < 10 km²: $D = 0.3641 \cdot A^{0.7950}$ ($R^2 = 0.4022$)

A ≥ 10 km²: $D = 0.9998 \cdot \ln(A) - 0.5927$ ($R^2 = 0.4605$)



Auto trips

A < 10 km²: $D = 0.6715 \cdot A^{0.4575}$ ($R^2 = 0.3987$)

A ≥ 10 km²: $D = 0.9788 \cdot \ln(A) - 0.5860$ ($R^2 = 0.2368$)

FIGURE 4 Theoretical functions adjusted as discontinuous models to the scatterplots of *area x average intrazonal trip distance*

4.4. Validation of the models

The results obtained in the previous steps of the analyses have confirmed the strength of the relationship between the average intrazonal trip distances and the areas of the zones, as observed in earlier studies. The models were validated with the data of group $IT_{30\%.}$ The performance of the models was then tested through a comparison with the estimation errors of the models proposed by other authors (Table 4).

The average and standard deviation values of the deviations found in the comparison between the values estimated with the models and the actual distances observed in the O/D survey were used in the analyses of the performance of the proposed models (Table 4). A first remark can be made about the linear models. Due to the presence of a constant term in the equation, the linear models shown in Table 4 always have an estimated value for the trip distances, even when the area of the zone is zero. The logarithmic models have a similar drawback, given that negative distance values can be found in very small areas.

Two of the models proposed by other authors ([12] and [13]) have a particularity that may be a problem for our analysis. Their calculation of the intrazonal trip distances is based on the distances between the centroid of the zone under analysis and those of the neighboring zones, therefore, they are not applicable to isolated areas. Given these negative points and the other results shown in Table 4, the discontinuous models showed the best performances in all cases considered. Although the R^2 value was not very high for the model calibrated for distances over 10 km^2, the average and standard deviation values of the residuals found with the combined models (i.e. over and under 10 km^2) were good. Furthermore, the discontinuous models clearly outperformed the models found in the literature.

5. CONCLUSIONS

A procedure for the estimation of intrazonal travel distances based on data usually available in O/D surveys has been developed. The analyses conducted in this study were based on O/D data collected in the city of São Carlos, Brazil, in the years of 2007 and 2008 [19]. The GIS files containing the street network and the traffic analysis zones (TAZ) adopted by local planners for at least one decade were used as additional information. The analysis method consisted in a search of geometrical characteristics of the TAZ that might affect the distances of intrazonal trips. Numerical values of the selected geometrical characteristics were then compared with the intrazonal travel distances so that strong or even moderate correlations could be obtained. The characteristics with the highest correlation values were used as variables in the construction of models, which were subsequently compared with five models selected from the literature.

The AVERAGE INTRAZONAL DISTANCES variable was not even moderately correlated with the following geometrical characteristics: SHAPE FACTOR, CIRCULAR SHAPE FACTOR, STUDY AREA/CIRCLE AREA RATIO, STUDY AREA/RECTANGLE AREA RATIO, NS/EW RATIO, DENSITY OF NETWORK CONNECTIONS, and STREET DENSITY. On the other hand, we have found correlations between the AVERAGE INTRAZONAL DISTANCES and the PERIMETER and the AREA of the studied areas, which could be seen as equivalent to the TAZ. Further analyses showed different patterns of the results, depending on the area values, which may be associated to distinct user's behaviors. As a consequence, we analyzed the results separately and adjusted regression models to each of the groups identified. Two models were calibrated and validated in the sequence.

TABLE 4 Performance of the models tested for the estimation of intrazonal trip distances

Models	Equations (for area A in km^2)	R²	Residuals (in km)	
			Stand. Dev.	**Average**
All trips				
Linear	$0.046*A + 0.9950$	0.2744	0.5667	0.4556
Power	$0.6378*A^{0.3541}$	0.3241	0.5528	0.4443
Logarithmic	$0.4503*\ln(A) + 0.5536$	0.3112	0.5558	0.4534
Discontinuous <10 km²	$0.5648*A^{0.4558}$	0.2989	**0.5992**	**0.5052**
≥10 km²	$0.6882*\ln(A) - 0.1319$	0.1045		
Smeed [14]	$0.81*A^{1/2}$	-	0.7393	0.9056
Batty [15]	$\sqrt{(^{A}/_{2*\pi})}$	-	0.5540	0.4973
Fotheringham [16]	$0.846*\sqrt{(^{A}/_{\pi})}$	-	0.5664	0.4436
U.S. D.C. [12]	Average distance/2	-	0.7165	0.5806
Venigalla et al. [13]	Minimum distance /2	-	0.7360	0.5758
Walking trips				
Linear	$0.0174*A + 0.7835$	0.3871	0.1639	0.1394
Power	$0.6334*A^{0.1947}$	0.4736	0.1440	0.1273
Logarithmic	$0.1674*\ln(A) + 0.6264$	0.4516	0.1429	0.1219
Discontinuous <10 km²	$0.5970*A^{0.2494}$	0.5663	**0.1412**	**0.1231**
≥10 km²	$0.2651*\ln(A) + 0.3334$	0.2493		
Smeed [14]	$0.81*A^{1/2}$	-	0.5481	0.9979
Batty [15]	$\sqrt{(^{A}/_{2*\pi})}$	-	0.2501	0.1901
Fotheringham [16]	$0.846*\sqrt{(^{A}/_{\pi})}$	-	0.3023	0.2763
U.S. D.C. [12]	Average distance/2	-	0.3932	0.5057
Venigalla et al. [13]	Minimum distance/2	-	0.3695	0.2925
Cycling trips				
Linear	$0.0641*A + 1.0857$	0.5569	0.4585	0.3804
Power	$0.5219*A^{0.5305}$	0.4849	0.4905	0.3957
Logarithmic	$0.6699*\ln(A) + 0.3628$	0.6217	0.4669	0.3880
Discontinuous <10 km²	$0.3641*A^{0.7950}$	0.4022	**0.4827**	**0.3942**
≥10 km²	$0.9998*\ln(A) - 0.5927$	0.4605		
Smeed [14]	$0.81*A^{1/2}$	-	0.6112	0.7807
Batty [15]	$\sqrt{(^{A}/_{2*\pi})}$	-	0.4561	0.6044
Fotheringham [16]	$0.846*\sqrt{(^{A}/_{\pi})}$	-	0.4592	0.4410
U.S. D.C. [12]	Average distance/2	-	0.4974	0.4209
Venigalla et al. [13]	Minimum distance/2	-	0.4223	0.5304
Auto trips				
Linear	$0.0605*A + 1.1179$	0.4250	0.5681	0.4507
Power	$0.7297*A^{0.3803}$	0.4299	0.5291	0.4193
Logarithmic	$0.5712*\ln(A) + 0.5823$	0.4687	0.5332	0.4220
Discontinuous <10 km²	$0.6715*A^{0.4575}$	0.3987	**0.5342**	**0.4186**
≥10 km²	$0.9788*\ln(A) - 0.5860$	0.2368		
Smeed [14]	$0.81*A^{1/2}$	-	0.6451	0.6392
Batty [15]	$\sqrt{(^{A}/_{2*\pi})}$	-	0.5446	0.6596
Fotheringham [16]	$0.846*\sqrt{(^{A}/_{\pi})}$	-	0.5444	0.5308
U.S. D.C. [12]	Average distance/2	-	0.8528	0.6947
Venigalla et al. [13]	Minimum distance/2	-	0.9183	0.7729

The discontinuous models summarized in Table 5 showed the best performances for the estimation of the average intrazonal distances in all cases studied and outperformed the models found in the literature, as shown in Table 4.

All trips	A <10 km²: $\overline{D}_{\text{AllTrips}} = 0.5648*A^{0.4558}$
	A ≥10 km²: $\overline{D}_{\text{AllTrips}} = 0.6882*\ln(A) - 0.1319$
Walking trips	A < 10 km²: $\overline{D}_{\text{WalkingTrips}} = 0.5970*A^{0.2494}$
	A ≥ 10 km²: $\overline{D}_{\text{WalkingTrips}} = 0.2651*\ln(A) + 0.6264$
Cycling trips	A < 10 km²: $\overline{D}_{\text{CyclingTrips}} = 0.3641*A^{0.7950}$
	A ≥ 10 km²: $\overline{D}_{\text{CyclingTrips}} = 0.9998*\ln(A) - 0.5927$
Auto trips	A < 10 km²: $\overline{D}_{\text{AutoTrips}} = 0.6715*A^{0.4575}$
	A ≥ 10 km²: $\overline{D}_{\text{AutoTrips}} = 0.9788*\ln(A) - 0.5860$

TABLE 5 Models selected for the estimation of average intrazonal trip distances

It is important to bear in mind a possible bias in these results. São Carlos is not located in a flat region, therefore cyclists and pedestrians often search for alternative paths that are not hilly, even if they are longer. This condition may certainly have affected the extension of the routes selected in the model, which were always based on a shortest path that minimizes length without considering additional sources of impedance. The city also has a few stretches of cycle paths. However, as they are short, disconnected and often located along the boundaries between zones, they probably did not affect the results of the analysis conducted for the evaluation of intrazonal trip distances. Finally, the strategy used to define the analysis zones is essential in studies like this one. It is always more difficult to interpret the outcomes of the proposed model if only geometrical characteristics of the zones are considered. The study would certainly benefit if the models take into account information about the street network, socioeconomic data of the population and infrastructure available for non-motorized trips. We believe that the association of these factors should be investigated in future studies.

Although the results of the case study may be not directly applicable to a particular context, they can provide reasonable first estimates of the average intrazonal travel distances. Moreover, the approach can be replicated elsewhere for the production of appropriate estimates.

ACKNOWLEDGEMENTS

The authors thank the Brazilian agencies CAPES (Post-Graduate Federal Agency), FAPESP (Foundation for the Promotion of Science of the State of São Paulo), and CNPq (Brazilian National Council for Scientific and Technological Development) for supporting the development of this research.

REFERENCES

1. U.S. Department of Commerce. *Traffic Assignment and Distribution for Small Urban Areas.* U.S. Government Printing Office, Washington, D.C., 1965.
2. Amâncio, M. A. *The Relationship between Urban Form and Walking Trips* (in Portuguese). Master's thesis, Federal University of São Carlos, São Carlos, Brazil, 2005.

3. Angel, S., J. Parent, and D. L. Civco. Ten Compactness Properties of Circles: Measuring Shape in Geography. *The Canadian Geographer / Le Géographe Canadien*, Vol. 54, No. 4, 2010, pp. 441-461.
4. Rodrigues da Silva, A. N., G. C. F. Costa, and N. C. M. Brondino. Urban sprawl and energy use for transportation in the largest Brazilian cities. *Energy for Sustainable Development*, Vol. 11, No. 3, 2007, pp. 44-50.
5. Costa, P. T., J. F. G. Mendes, and A. N. Rodrigues da Silva. *An Analysis of Energy Consumption for Transportation in Portuguese cities using Artificial Neural Networks*. Proceedings of the 10th World Conference on Transport Research (WCTR'04), Istanbul, Turkey, 2004.
6. Daganzo, C. F. The Length of Tours in Zones of Different Shapes. *Transportation Research Board*, Vol. 18B, No. 25, 1984, pp. 135-145.
7. Khan, M., and K. M. Kockelman. *Models for Anticipating Non-motorized Travel Choices and the Role of the Built Environment*. Proceedings of the Annual Meeting of the International Association for Travel Behavior Research, Toronto, Canada, 2012.
8. Sanches, S. P., and M. A. G. Ferreira. Comparative Analysis of the Urban Form of Brazilian Medium-sized Cities (in Portuguese). *Minerva*, Vol. 5, No. 2, 2005, pp. 177-185.
9. Greenwald, M. J. The Relationship between Land Use and Intrazonal Trip Making Behaviors: Evidence and implications. *Transportation Research Part D: Transport and Environment*, Vol. 11, No. 5, 2006, pp. 432-446.
10. Bhatta, B., and O. I. Larsen. Are Intrazonal Trips Ignorable? *Transport Policy,* Vol. 18, No. 1, 2011, pp. 13-22.
11. Ortúzar, J. D., and L. G. Willumsen. *Modelling Transport.* 4. ed. John Wiley & Sons Inc., Chichester, 2011.
12. U.S. Department of Commerce. *Calibrating and Testing a Gravity Model for Any Size Urban Area*. U.S. Government Printing Office, Washington, D.C., 1965.
13. Venigalla, M. M., A. Chatterjee, and M. S. Bronzini. A Specialized Equilibrium Assignment Algorithm for Air Quality Modeling. *Transportation Research Part D: Transport and Environment*, Vol. 4, No. 1, 1999, pp. 29-44.
14. Smeed, R. J. The Road Capacity of City Centers. *Highway Research Record*, Vol. 169, 1967, pp. 22-29.
15. Batty, M. *Urban Modeling: Algorithms, Calibrations, Predictions.* Cambridge University Press, London, 1976.
16. Fotheringham, A. S. *Market Share Analysis Techniques: a Review and Illustration of Current U.S. Practice.* In: N. Wrigley (ed.) Store Choice, Store Location and Market Analysis. Routledge, London, 1988, pp. 120-159.
17. Kordi, M., C. Kaiser, and A. S. Fotheringham. A Possible Solution for the Centroid-to-centroid and Intra-zonal Trip Length Problems. Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012.
18. IBGE. *Cidades.* Instituto Brasileiro de Geografia e Estatística, 2010. http://www.ibge.gov.br/cidadesat/topwindow.htm?1. Accessed April 30, 2013.
19. Rodrigues da Silva, A. N. *Origin-Destination Survey in São Carlos.* FAPESP Project No. 04/15843-4. Final Report. São Carlos School of Engineering, University of São Paulo, São Carlos, 2008.
20. CALIPER. *TransCAD 5.0 - Transportation GIS Software.* Caliper Corporation, Newton, MA, USA. 2008.