

**SPATIAL GROWTH PATTERNS AND TRANSPORTATION IMPACTS OF
MAJOR TRIP GENERATORS – THE CASE OF CAMPINAS, BRAZIL**

Anna Beatriz Grigolon

abgrigolon@gmail.com

Gustavo Garcia Manzato

gmanzato@sc.usp.br

Antônio Néelson Rodrigues da Silva

anelson@sc.usp.br

University of São Paulo

São Carlos School of Engineering

Department of Transportation

Av. Trabalhador São-carlense, 400 – São Carlos, SP, Brazil

ABSTRACT

Given the evidences that the location and size of new Trip Generators (TG's) may be strongly influenced by existing TG's, the objective of this paper was to evaluate this hypothesis in a Brazilian medium-sized city. The analysis may be useful for forecasting the impacts that a large business development can produce in the transportation systems around and connected to it. A methodology was then proposed for modeling the spatial growth patterns of TG's and applied in the city of Campinas. The results can be an important reference for cities that are experiencing fast and intense changes in their spatial occupation and land use patterns.

1. INTRODUCTION

In the last decades many Brazilian cities have experienced intense growth, with impacts on the economic activities and internal and external population flows. The changes produced in the infrastructure for transportation were quite often directed to serve the private car (Scarlatto, 1989). Along with the changes in the transportation infrastructure, changes in the distribution of land uses were also observed. They also reflect the dominance of the private motorized mode. One of these changes is the concentration of business and activities in particular zones of the city apart from the traditional CBD (Central Business District). Those clusters of activities have a significant impact on transportation, because they are usually large Trip Generators (TG's).

Many efforts have been done to characterize the TG's and to develop methodologies for assessing their impacts in the urban area, even in developing countries. In the case of Brazil, for example, urban and transportation planners have access to studies about the topic since the early 1980's (CET, 1983; Goldner, 1986). More recent studies are also available, such as Portugal and Goldner (2003). Other developing countries also have studies in the area, such as Espejo (2001). The main focus of those studies, however, lies on aspects of circulation, traffic safety, and accessibility. This is justified by the fact that large TG's produce several impacts on traffic in their areas of influence, what is visible in large Brazilian cities, according to ANTP (2005).

Other aspects of TG's still require studies, however. One of them is the process of site selection for new business. That was discussed by Cheng *et al.* (2007) for the case of shopping malls. Yu *et al.* (2007) extended the previous study looking at the distribution of shopping centers that could satisfy vendors and shoppers in terms of realizing the shortest car based shopping trips. Our study is complementary to both, because we

have focused on an actual aspect that drive the spatial distribution of TG's within the urban areas, which is the evidence that the location and size of new Trip Generators (TG's) may be strongly influenced by existing TG's. Therefore, the objective of this paper is to evaluate this hypothesis in a Brazilian medium-sized city. Two branches of spatial analysis are explored in the present study while dealing with several levels of aggregation of spatial information. They are: spatial statistics and spatial modeling.

Firstly, concepts of Exploratory Spatial Data Analyses (ESDA), like Moran's I index and scatterplot, are used to characterize regions of similar behavior in terms of particular variables. ESDA techniques strongly rely on indexes, graphs and maps, and techniques such as the Moran's I index and scatterplot were already explored as the first step of a modeling effort in previous studies (Ramos and Silva, 2003a, 2003b, and 2007; Ramos *et al.*, 2004; Manzato *et al.*, 2006; Manzato and Silva, 2006). The use of the Moran's I index and scatterplot for that purpose is particularly interesting because the areas under analyses can assume four conditions only (Q1, Q2, Q3, and Q4), which are the Quadrants of the scatterplot. Those Quadrants indicate the relative position of the attribute value of a particular zone in relation to the values of all other zones in the studied area. Q1 and Q2 indicate areas where the value of the attribute is similar to the average of values found in the immediate neighbors. In Q1 both are high values (i.e., above the overall average value), while in Q2 both are low values (i.e., below the overall average value). Both conditions indicate a positive spatial autocorrelation. Conversely, Q3 and Q4 are zones with negative spatial autocorrelation. Q3 zones have attribute values lower than the overall average, while the average of the neighbor zones is higher than the overall average. Q4 zones are just the opposite (more details can be found in the literature, for example, in Ramos and Silva, 2003a and 2003b).

Secondly, spatial modeling is carried out using Cellular Automata concepts and the growth patterns observed in the consolidated areas can be used to build scenarios for new developments. When allowing the identification of patterns and the projection of trends, this kind of analysis may be useful for forecasting the impacts that a large business development can produce in the transportation systems around and connected to it. This is particularly true for Brazilian medium-sized cities, which are experiencing a very intense growth process, or even for cities of other countries that present similar characteristics. In the case studied here, the proposed methodology is applied for modeling the spatial growth patterns of TG's in the neighborhoods of some shopping malls in the city of Campinas. The spatial analysis and modeling techniques are applied to examine the development of clusters around the shopping malls. The outcomes of the models can be later used to forecast their impacts on transportation systems. The proposal is specifically applied to shopping malls due to their strong attraction power to other business. Campinas was selected to be the case study because of its status as a transition example between the conditions usually observed in Brazilian medium-sized cities (including those in the state of São Paulo) and those present only in the largest Brazilian cities.

2. METHODOLOGY

The method developed for the present study can be divided in four phases: *i)* Data acquisition and treatment; *ii)* Exploratory spatial data analysis, by means of spatial statistics; *iii)* Construction of a forecast model using Cellular Automata concepts; *iv)* Calibration and validation of the model. Before describing the phases in details, however, some of the main characteristics of the city analyzed must be highlighted. Campinas, which was founded in 1774, is nowadays a sprawled city with a strong

commercial sector. It is also a provider of specialized services, what have attracted a significant proportion of high income groups for living within and around the city.

The continuous urbanized area of Campinas has relatively high values of population density and it can be described as a region of easy circulation and high mobility, what favors the social contact of the citizens living in the metropolitan area. With around one million inhabitants (in 2003), it is the second largest city in the state of São Paulo, only smaller than the capital. The population is formed predominantly by people who moved from other regions mainly in the period comprised from 1970 until 1980, when two-thirds of the inhabitants were not born in the city. Those conditions apparently did not affect the quality of life in the city, as one can see through the average value of the Human Development Index (HDI) observed there. The HDI value of 0.852 is considered high, in the scale from zero to one (Prefeitura Municipal de Campinas, 2006).

Three regions of the city were particularly important for the case study: the CBD (Central Business District), and two major Trip Generators (the shopping malls Iguatemi and Galleria). The City Hall building was considered the central point of the CBD. Its historical function as a city landmark has driven the commercial development in the central area of Campinas for many years. In a similar fashion, both shopping malls selected play an important role in the urban area due to the intense commercial activities and flows they generate. The Shopping Iguatemi Campinas was opened to the public in 1980. With 281 stores offering different products and located in an easily accessible spot, it currently attracts an estimated number of twenty-two million visits per year. The other TG studied, the Shopping Galleria, started to operate only in 1992. It has 145 shops and it currently attracts an estimated number of five million customers

per year (Grigolon and Silva, 2006). According to the mall administration, the catchment area of the TG encompasses more than 500,000 inhabitants.

2.1. Data acquisition and treatment

The municipality provided all data needed for the application, as previously described by Grigolon and Silva (2006). The dataset contained information about the land uses for a large number of parcels, summing up 170,497 records, what represented 47 % of the total number of urban lots. The database was already georeferenced, what made possible the use of a Geographic Information System for the analyses. The lots were associated with points in a map and the records were classified in subdivisions of the three following categories of land use: residential, commercial, and industrial. The focus of our study was the commercial use, although some important TG's could also be found among the industries. However, as the industries were much more restricted by the zoning regulations than the other uses we decided to leave them out the model, at least for the present study. In order to be able to apply the ESDA concepts aforementioned, the data associated to the points were aggregated into regular cells of 100 x 100 meters. That aggregation significantly reduced the size of the dataset, from more than 170,000 points to 4,858 areas.

After organizing the data in the way described above, two variables were extracted from the dataset to be used in the following phases of the analysis: the *absolute number of records with commercial use* per cell, and the *proportion of records with commercial use* in relation to the total number of records (i.e., residential, commercial, and industrial) per cell. Those variables were separated according to the 'age' of the building. That information came from the year registered in the individual records,

which were grouped in the following five classes: before 1979, from 1980 to 1985, from 1986 to 1991, from 1992 to 1997, and from 1998 to 2003.

2.2. Exploratory spatial data analysis

The second phase of the proposed method was based on an exploratory analysis of the spatial data, by means of spatial statistics. After a preliminary examination of the spatial distribution of the two variables, the variable *proportion of records with commercial use* was selected for the next phase. The application of the ESDA concepts was done with tools available in the GIS *ArcView* (version 3.2) combined with the extension *Spacestat* (Anselin and Bao, 1997; Anselin and Smirnov, 1998), for the calculation of the spatial proximity matrix, and a spreadsheet. From the two alternatives related to the neighborhood condition for the construction of the spatial proximity matrix available in *Spacestat* (*Rook* or *Queen*), we selected the latter. That means that the cells are already considered as neighbors if they have a single common point. That condition is particularly important in the case of squared cells, as here. The results of the exploratory analysis are quadrants of the Moran's scatterplot, which can be displayed in *Box Maps*. Those maps can be built for different periods of time, allowing the visualization of the development process under a common criterion, (i.e., the four quadrants of the Moran's scatterplot).

2.3. The proposed model

This section summarizes the third and fourth phases of the methodology, which are: the construction of a forecast model using Cellular Automata concepts; and the calibration, validation, and application of the model.

As the model was built using Artificial Neural Networks, an additional treatment of the variables was performed. The two variables described in section 2.1 were combined in a single categorical variable. In order to do so, the *absolute number of records with commercial use* was transformed in five equal size classes, named as A, B, C, D, and E. A similar procedure was done with the variable *proportion of records with commercial use*. In that case, the classes were: 1, 2, 3, 4, and 5. The combination of the groups in a single categorical variable produced then twenty-five possibilities (e.g., A1, A2, ..., E4, E5), what can give a clear indication of the intensity of commercial use in the cells. That categorical variable is therefore the output of the neural net. Moreover, the original variables from the precedent periods of time are also inputs of the model, along with: the Moran's scatterplot Quadrant of each cell, the average value of the variable *proportion of records with commercial use* considering only the immediate neighbors of each cell in the calculation, and the number of neighbor cells in each of the quadrants Q1, Q2, Q3 and Q4.

The structure of the model is such that the output is always based on the transitions observed in the cells in previous periods of time. Those transitions along time are reflected in changes in the relative position of the cells in the Moran's scatterplot. That modeling approach using neural nets was proposed by Ramos and Silva (2007), as an alternative to another modeling concept based on fixed transition rules (as used by Ramos and Silva, 2003a e 2003b). Among the advantages of the neural net approach are: the possibility of using the variable under analysis as the output of the model (instead of the Quadrants), and the possibility of using parts of the dataset for validation and test.

In summary, the approach proposed here uses a *Multilayer Perceptron* (MLP) neural network with a *backpropagation* algorithm to model the behavior of the cells regarding their transitions in the period from 1979 until 2003. In the calibration phase, 50 % of the records were selected for training, 25% for validation and 25 % for testing, using input data from the following periods: before 1979, from 1980 to 1985, and from 1986 to 1991. The output in that case was the period from 1992 to 1997.

The calibrated model was subsequently validated using input data from the following periods: from 1980 to 1985, from 1986 to 1991, and from 1992 to 1997. The output in that case was the period from 1998 to 2003. Finally, as the performance of the model in the validation phase was acceptable, it was used for predictions in the next time step. The input periods were: from 1986 to 1991, from 1992 to 1997, and from 1998 to 2003; while the output was from 2004 to 2009. The results obtained are presented in the next section, along with a discussion of the application.

3. RESULTS

The application, in the city of Campinas, of the methodology described in the previous section produced the results summarized in Figures 1, 2, 3, 4, and 5. The maps displayed in those figures are *Box Maps*, i.e., maps showing the distribution of the zones belonging to each one of the four quadrants of a Moran's scatterplot. In the case of the figures 1 through 5, the variable used to build the scatterplots (and also the maps) is the share of commercial activities in the cells in five periods of time: before 1979, from 1980 to 1985, from 1986 to 1991, from 1992 to 1997, and from 1998 to 2003.

The cells in black represent the values in Quadrant 1 (Q1), which is the case of cells with high values (above the overall average value) surrounded by cells also with high

values. They are predominantly located around the CBD as well as along some of the main streets. In addition, after the points in time when the large shopping malls considered in the analysis were opened, the cells around them immediately changed to the Q1 condition. The some sort of change was observed along some important links of the urban network of streets. Not surprisingly, those changes were accompanied by a reduction in the number of Q1 cells around the CBD.

Another piece of information available in Figures 1 through 5 is the distribution of cells containing values in Quadrants 2, 3, and 4 (respectively, Q2, Q3, and Q4). Most cells of the studied area are in Q2, which is the case of cells with low values (below the overall average value) surrounded by cells also with low values. In our case, however, cells in Q2 do not play an important role in the modeling effort, unless they change to another quadrant along time. In contrast, cells in Q3 deserve special attention because of their propensity to naturally move to Q1. The last group, which is formed by the cells in Q4, has usually isolated cells. This is an indication of their strong local power, given they are cells with high values (above the overall average value) surrounded by cells also with low values. That condition, however, does not affect the model developed for the present study, which is mainly affected by the main TG's, such as large shopping malls.

The estimation values indicated a good overall performance of the model. In the calibration phase, 89.3 % of the predictions were in the right classes, while in the validation phase the figure was close to 86.0 %. That performance encouraged the use of the model for predicting values in a future time step, from 2004 to 2009. Although the predictions in the entire area can play an important role in a more comprehensive planning effort, the results of the predictions around the shopping malls Iguatemi and Galleria are particularly relevant for the present study. The values used in the following

analysis are displayed in Figures 6 and 7, which show the classes of data in time steps from 1979 until 2009, Figure 6 shows the distribution of data along time regarding the Iguatemi shopping mall, while Figure 7 shows the case of the Galleria shopping mall. The vertical bars in Figures 6 and 7 are associated with classes representing the proportion of commercial use in the map cells. The values between 80 and 100 % are the focus of the present analysis because they are likely to be the main TG's. The sites of the shopping malls were already stressed in Figures 1 through 5.

In the analysis of the shopping mall Iguatemi (Figure 6), there is a continuous growth in the number of cells in the bars representing the class 80-100 % since 1979 until 2003, with a reproduction of the current figure in the future time step (2004-2009). In the case of the shopping mall Galleria (Figure 7), there is a reduction in the actual values of the class 80-100 % after 1979. With the exception of the period 1992-1997, the value of the class 80-100 % remains stable and equal to three. In the prediction produced by the model, however, there is an increase in that figure for the near future. That is in accordance with the expectations, given that the opening date of that particular shopping mall should affect only recent (and future) years.

4. CONCLUSIONS

The case study carried out in the city of Campinas, Brazil, for testing the hypothesis that existing Trip Generators can influence the location and size of new TG's produced some interesting outcomes. Initially, the good results in the case studied has demonstrated the suitability of our model for reproducing the temporal dimension involved in the location of new TG's based only on the location and size of the TG's. Moreover, the use of the model to build future scenarios suggests an increase or at the least the stability of the number of new TG's around the main existing TG's. The

intensity of those trends, however, is quite sensitive to the time span analyzed. In any case, the confirmation of the original hypotheses can bring significant impacts for transportation planning, due to the increase of traffic that can be foreseen with the clustering of TG's. Further studies should be conducted in order to evaluate the extension and intensity of those impacts.

It is also important to notice the possibility of underestimation in the results produced by the model as a consequence of the peripheral TG's considered in the analysis. In order to improve the predictive power of the model, it could be valuable to include other input variables in the analyses, as well as to test it with different spatial settings. In such a way, not only the performance of the model would eventually be improved but also its adaptability would be tested. One of the potential elements to become an input variable in the neural net model would certainly be the distribution of transportation infrastructure (such as main roadways, arterial streets, etc.) throughout the urban area, what could possibly add the strength of the transportation-land use relationship to the model.

5. ACKNOWLEDGEMENTS

The authors would like to express their gratitude to 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), which have supported our efforts for the development of this work in different ways and periods. Finally, we thank the referees for their valuable comments.

6. REFERENCES

Anselin, L. and Bao, S. (1997) Exploratory spatial data analysis linking SpaceStat and ArcView, in Fisher, M. and Getis, A. (eds.) *Recent Developments in Spatial Analysis*, Springer-Verlag, Berlin.

Anselin, L. and Smirnov, O. (1998) The SpaceStat extension for ArcView 3.0, *Regional Research Institute*, West Virginia University, Morgantown.

ANTP (2005) *Planejando o desenvolvimento das Cidades*, Associação Nacional de Transportes Públicos, São Paulo, Available in: http://www.antp.org.br/telas/desenvolvimento_urbano/capitulo2_urbano.htm.

CET (1983) *Pólos Geradores de Tráfego*, Companhia de Engenharia de Tráfego, Boletim Informativo nº 32, Prefeitura de São Paulo, São Paulo.

Cheng, E.W.L.; Li, H. and Yu, L. (2007) A GIS approach to shopping mall location selection. *Building and Environment*, 42, pp. 884–892.

Espejo, C.P.L.(2001). *Estimación de tasas de generación de viajes para actividades comerciales en el A.M.C.: Propuesta Metodológica*, Master's Thesis, Universidad Simon Bolívar, Caracas, Venezuela. Available in: <http://redpgv.coppe.ufrj.br>.

Goldner, L.G., (1986). *A interferência dos pólos geradores de tráfego no sistema viário: análise e contribuição metodológica para Shopping Centers*, Master's Thesis, Programa de Engenharia de Transportes, COPPE/Universidade Federal do Rio de Janeiro, Brazil.

Grigolon, A.B. and Silva, A.N.R. (2006) Impactos dos padrões de ocupação do solo no entorno de Pólos Geradores de Viagens, in: *Proceedings of the XX Congresso de Pesquisa e Ensino em Transportes*, Brasília (DF), Brazil, 1, pp. 617-628.

Manzato, G.G. and Silva, A.N.R. (2006) Incorporando um indicador de oferta de infraestrutura de transportes na definição de regiões metropolitanas, in: *Proceedings of the XX Congresso de Pesquisa e Ensino em Transportes*, Brasília (DF), Brazil, 1, pp. 341-352.

Manzato, G.G.; Soares, A.J. and Silva, A.N.R. (2006) Aglomerações urbanas e oferta de transporte no estado de São Paulo, Brazil, in *CD-ROM Proceedings of the 2nd Congresso Luso Brasileiro para o Planeamento, Urbano, Regional, Integrado, Sustentável*, Braga, Portugal.

Portugal, L.S. and Goldner, L.G. (2003) *Estudo de Pólos Geradores de Tráfego e de seus impactos nos sistemas viários e de transportes*, 1st ed., Edgard Blücher, São Paulo, Brazil.

Prefeitura Municipal de Campinas (2006). Available in: www.campinas.sp.gov.br/portal_2003_sites/conheca_campinas/cc_campinas_a_cidade.htm, Accessed on: November 24th, 2006.

Ramos, R.A.R. and Silva, A.N.R. (2003a) A data-driven approach for the definition of metropolitan regions, in: *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management – Reviewed Papers*, Sendai, Japan.

Ramos, R.A.R. and Silva, A.N.R. (2003b) Um contributo para a delimitação da área metropolitana do noroeste de Portugal, *Revista Portuguesa de Estudos Regionais*, pp. 61-82.

Ramos, R.A.R. and Silva A.N.R. (2007) A spatial analysis approach for the definition of metropolitan regions – the case of Portugal, *Environment and Planning B*, **34** (1), pp. 171–185.

Ramos, R.A.R.; Silva, A.N.R. and Miranda, V.P. (2004) Comparação de metodologias de análise espacial para definição de zonas urbanas homogêneas, Available in: www.igeo.pt/IGEO/portugues/servicos/

Scarlato, F.C. (1989) *Metropolização de São Paulo e o Terceiro Mundo*, 2nd ed., Iglu, São Paulo, Brazil.

Yu, B.; Yang, Z. and Cheng, C. (2007) Optimizing the distribution of shopping centers with parallel genetic algorithm. *Engineering Applications of Artificial Intelligence*, **20**, pp. 215–223.

ILLUSTRATIONS

Figure 1: Box map representing the variable *proportion of records with commercial use* before 1979 – page 16.

Figure 2: Box map representing the variable *proportion of records with commercial use* from 1980 to 1985 – page 17.

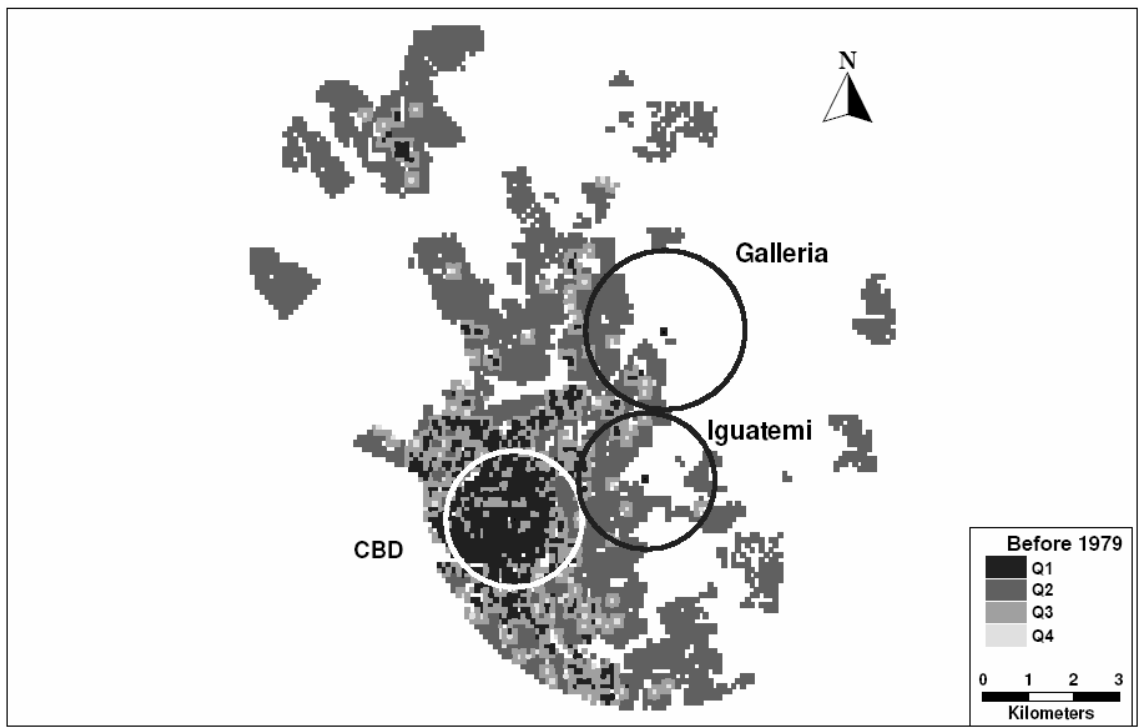
Figure 3: Box map representing the variable *proportion of records with commercial use* from 1986 to 1991 – page 18.

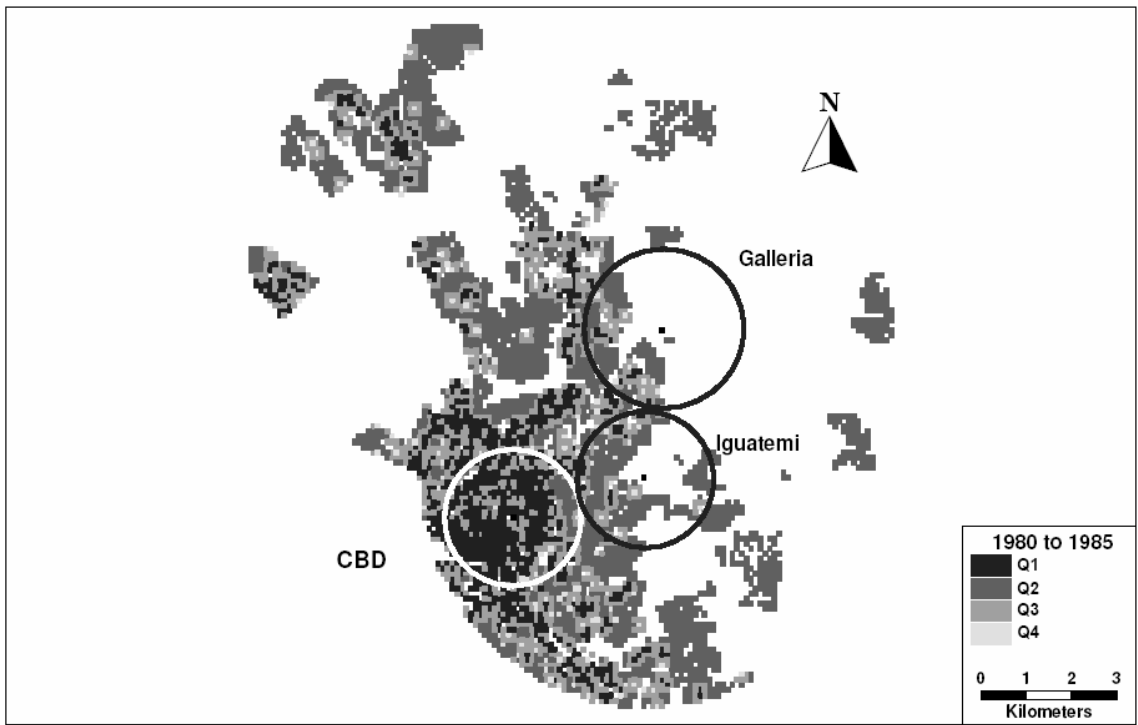
Figure 4: Box map representing the variable *proportion of records with commercial use* from 1992 to 1997 – page 19.

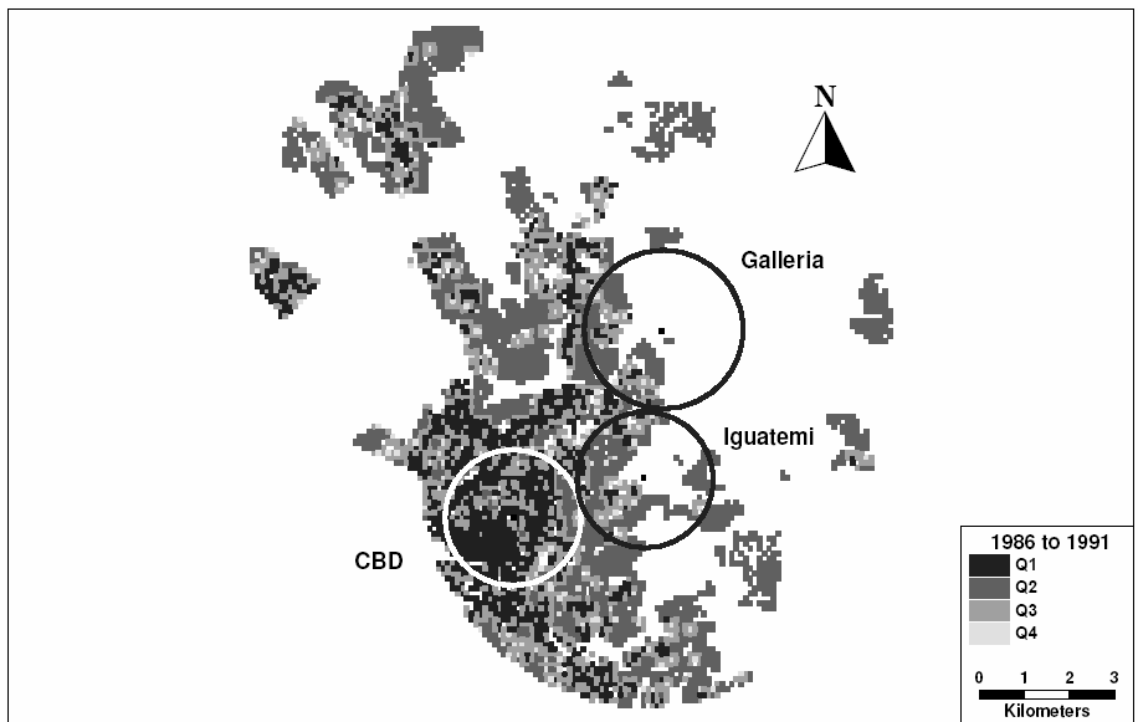
Figure 5: Box map representing the variable *proportion of records with commercial use* from 1998 to 2003 – page 20.

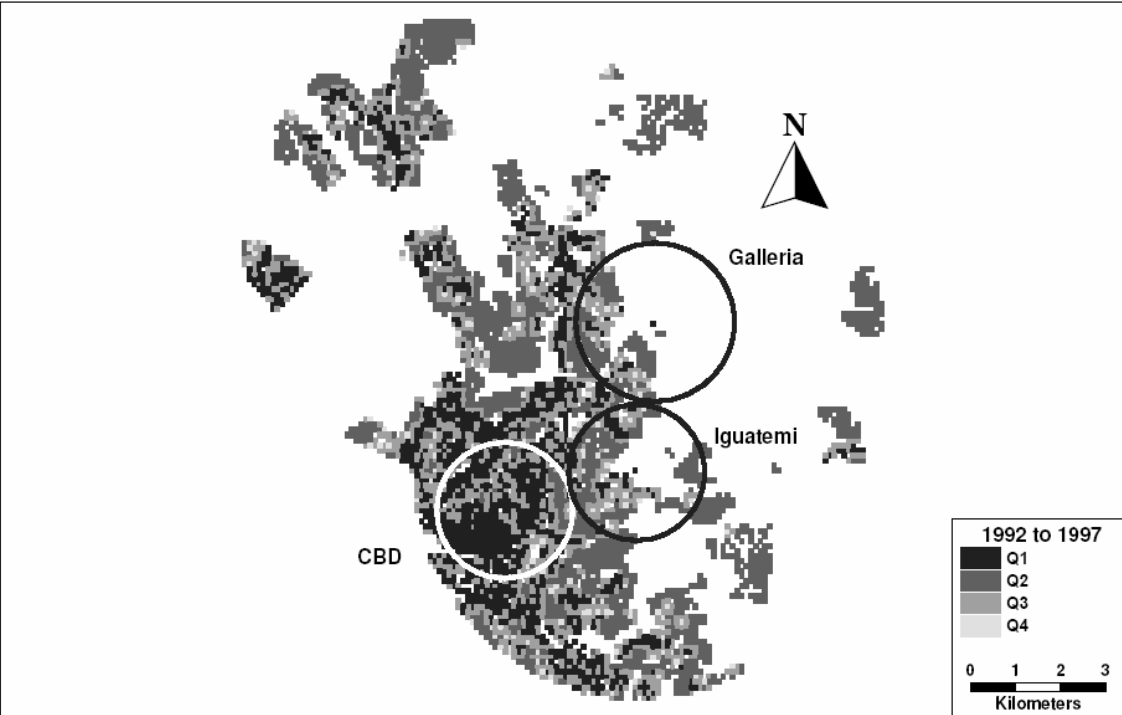
Figure 6: Graph showing the distribution of data along time around the Iguatemi shopping mall – page 21.

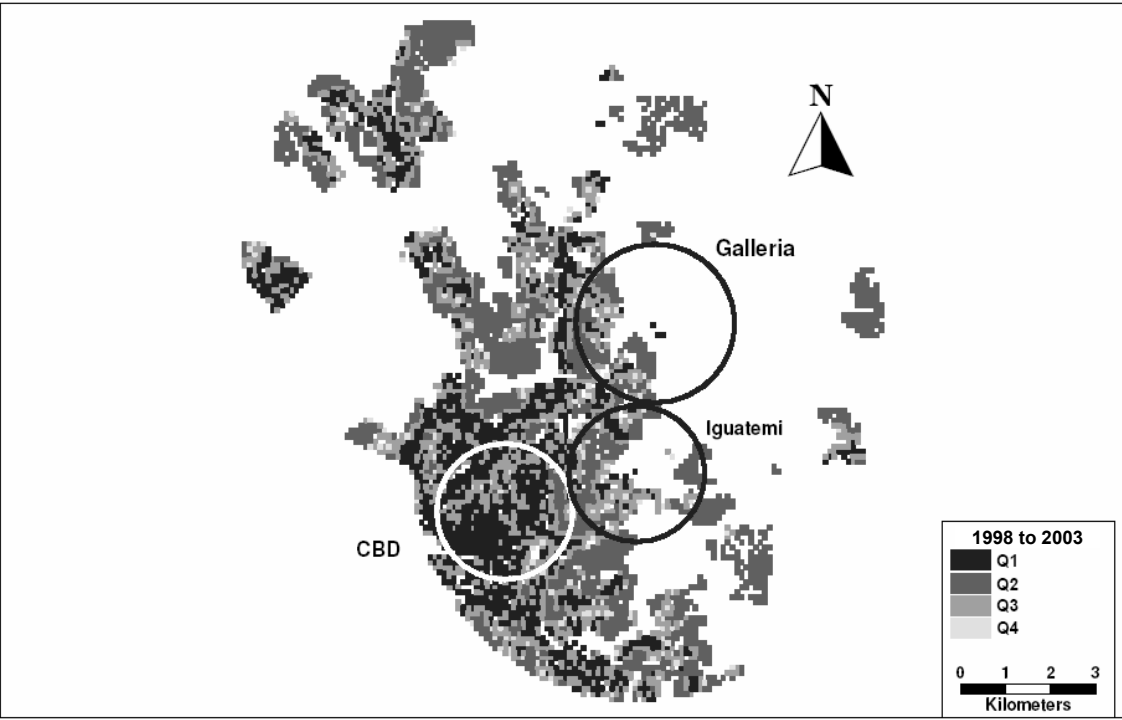
Figure 7: Graph showing the distribution of data along time around the Galleria shopping mall – page 22.











Shopping Iguatemi

