Geostatistics as a Tool to Map the Spatio-Temporal Evolution of Car Ownership in São Paulo Metropolitan Area

Eduardo Henrique de Moraes Takafuji [1]

Bruna Lahos de Jesus Bacic[2]

Juliana Siqueira-Gay[2]

Mariana Abrantes Giannotti[2]

Marcelo Monteiro da Rocha[1]

[1] Geosciences Institute - University of São Paulo – IGc/USP
Sala 105, Bloco I Superior, Rua do Lago, 562 Cidade Universitária 05508-080 São Paulo – SP, Brasil
eduardo.takafuji@usp.br; mmrocha@usp.br
[2]Polytechnic School of University of São Paulo – EP/USP
Av. Prof. Almeida Prado, Travessa 2, nº 83 – Cidade Universitária 05508-080 São Paulo – SP, Brasil
bruna.lahos@usp.br; mariana.giannotti@usp.br; siq.juliana@gmail.com

Abstract. The spatial distribution knowledge of the car ownership per family is essential to characterize urban mobility, accessibility, transport related social exclusion, among others. The commuting from population living at the peripheral areas to access jobs saturate the transport system and car ownership could be a proxy indicator for understanding the population's behavior and its relation with transport infrastructure. To represent better car ownership pattern of behavior by family in the São Paulo Metropolian Area (SPMA), this paper aims to map the cars owned by family using ordinary kriging. Kriging estimation is considered a Best Linear Unbiased Estimator (BLUE) technique that bring the uncertainties associated with the value estimation. This estimation is also important in order to perform multitemporal analysis of car ownership, as census limits changes overtime. So, this paper also provide a comparison between krigged maps from the last two census surveys – year 2000 and year 2010. Thus, a final map was compiled subtracting the resultant of 2010 by the 2000, in order to properly evaluate the car ownership evolution. It could be observed that the car ownership per household has increased in most of the SPMA, especially in cities where the provision of public transport is less present.

Key-words: Geostatistics, Car Ownership, São Paulo Metropolitan Area (SPMA), Kriging.
Palavras-chave: Geoestatística, Posse de automóvel, Região Metropolitana de São Paulo (RMSP), Krigagem

1. Introduction

The spatial distribution knowledge of the car ownership per family is essential to characterize urban mobility, accessibility, transport related social exclusion, among others. The increased car ownership could be explained by different factors, for instance, demographic variables, households sizes (Ritter and Vance (2013); Whelan (2007)) or economic factors, such as policies to promote the car's purchase and gasoline prices (Yagi and Managi (2016)).

For Loureiro et al. (2006) the urban travel studies are influenced by geographical attributes such as high residential density and socioeconomic activities, proximity between areas, spatial coverage of the transport network, travel impedance on the road network, etc.

The metropolitan area of São Paulo (SPMA) is extremely heterogeneous and complex given the particularities of city expansion patterns and concentrated supply of labor in some industrial and financial centers. The commuting from population living at the peripheral areas to access jobs saturate the transport system, and the price of land is, therefore, influenced by accessibility to jobs conditions (Wegener and Furst (1999)). Thus, an important portion of population tends to settle in the skirts of the city (Correa (1995)) which implies in high travel times to commute from home to work.

Aranha (2005) studied the commuting process between municipalities from SPMA. The author used 2000 census data where it could be observed relationships from migration trends and how these contribute to the structuring of the metropolitan space and noticed that more than half of the economically active population works in a municipality different from where they live, making it necessary to commute, usually by car.

Gomes et al. (2016) stresses the need to study the relationship of modal choice (e.g. bus, car, bike) and spatial locations of households, considering the influence of urban areas in decisions to use certain mode of transport.

Thus, car ownership could be a proxy indicator for understanding the population's behavior and its relation with transport infrastructure. Besides that, the rise level of motorization could cause significant consequences as increased energy consumption, atmospheric pollution and traffic congestion (Whelan (2007)) and can serve as a good parameter to inform urban and environmental planning stakeholders. The car ownership in a household level is a variable collected by the Demographic Census, conducted by the Brazilian Institute of Statistics and Demography (IBGE). The census tracts limits changes every ten years, what make it difficult to perform a multitemporal analysis. To overcome this interpolation and estimations techniques can be used as explored at this article. Kriging estimation is considered a Best Linear Unbiased Estimator (BLUE) technique, that bring the uncertainties associated with the value estimation.

To represent better car ownership pattern of behavior by family in the SPMA, this paper aims to map the cars owned by family using ordinary kriging. As well as, to compare the maps generated with the last two census surveys – year 2000 and year 2010. Thus, a final map was compiled subtracting the resultant of 2010 by the 2000, in order to properly evaluate the car ownership evolution.

2. Metodology

The research steps followed: gathering of census data and respective car ownership, centroids calculation, kriging estimation and finally map algebra operation (Figure 1).

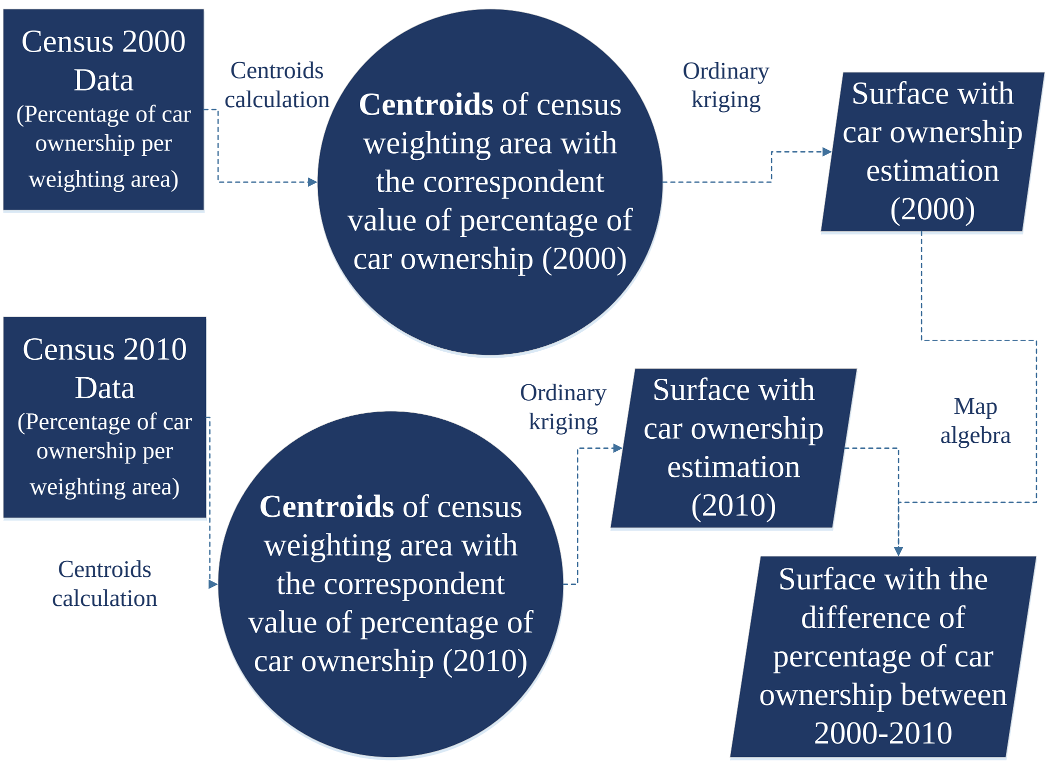

Figure 1: Methodological Scheme

2.1. Census 2000 and Census 2010

Brazilian Population Census is developed by the (IBGE) every ten years. There is two possible data source from Census, the results of the Universe and from the Sample. The car ownership is from the result of the Sample which is divided according to the weighting area. Such spatial area is defined as a geographical unit, comprising a mutually exclusive group of contiguous census tracts. The application of weights in survey data collected is a calibration procedure in order to produce estimates consistent and related to the population as a whole. The minimum size of the interviews sample done by weighting area is 400 occupied households, except for the municipalities that not reach sufficient number of respondents and in this case, the municipality itself is considered a weighting area.

2.2 Study Area

The study area is São Paulo Metropolitan Area (Figure 2), consisting of 39 municipalities approximately 20 million people in an area of 8,051 km² (IBGE, 2010). The GDP (Gross Domestic Product) in this area accounts for over 55% of São Paulo (SEADE, 2015).

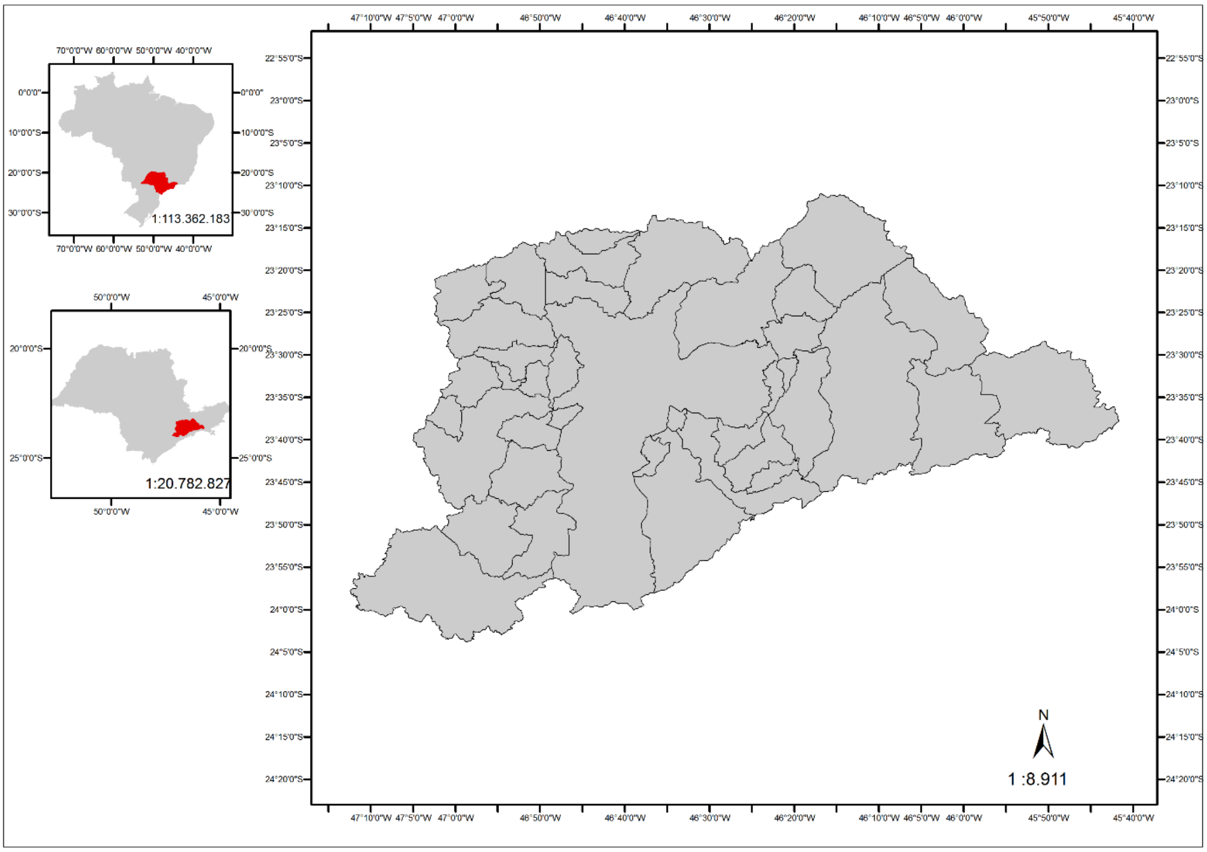

Figure 2: São Paulo Metropolitan Area

2.3 Geostatistics

Matheron (1962) define a regionalized variable as, sensu stricto, an actual function, taking a definite value in each point of space. From the point of view of physics or geology, a given number of qualitative characteristics are linked to the notion of regionalized variable, such as localization, geometrical support (volume of the sample, with its geometrical shape, its size and orientation), continuity (in its spatial variation), and anisotropy. This is made possible owing to a simple mathematical tool: the variogram. The variogram is a curve representing the degree of continuity of mineralization. Experimentally, one plots a distance h in abscissa and, in ordinate, the mean value of the square of the difference between the grades of samples picked at a distance “h” one from the other. Armstong (1998) shows the variogram function of an intrinsic

ISBN: 978-85-17-00088-1

random function - for stationary and intrinsic variables, the mean of Z(x+h) - Z(x) is zero, and so γ(h) is just the mean square difference – as:

$$\gamma(h) = 0.5\, E\, [(Z(x+h) - Z(x)^2)]$$

Deutsch (2002) explains that the experimental variogram points are not used directly in subsequente geostatistical steps such as kriging and simulation; a parametric variogram model is fitted to the experimental points. There are two main reasons: (a) the variogram function is required for all distance and directions vectors h within the search neighborhood of subsequent geostatistical calculations and (b) the variogram measure and γ(h) must have the mathematical property of "positive definiteness" for the corresponding covariance model. Moreover the most used variogram models are Spherical, Exponential and Gaussian as shown by Armstrong (1998):

$$Sph(h) = \begin{cases} (C\left(1.5\left(\frac{h}{a}\right) - 0.5\left(\frac{h}{a}\right)^3\right) & , if\ a > h \\ C & , if\ a \leq h \end{cases}$$

$$Exp(h) = C(1 - \exp\left(-\frac{h}{a}\right))$$

$$Gauss(h) = C(1 - \exp\left(-\frac{h^2}{a^2}\right))$$

Isaaks & Srivastava (1989) explains that the choice of a covariance model (or, if one prefers, a variogram model or a correlogram model) is a prerequisite for ordinary kriging. The term ordinary kriging is associated with the acronym B.L.U.E., where "linear" is because its estimates are weighted linear combinations of the available data; it is "unbiased" since it tries to have the mean residual or error, equal to 0; it is "best" because it aims at minimizing, the variance of the errors. The distinguishing feature of ordinary kriging (compared to all of the other estimation methods), therefore, is its aim of minimizing the error variance. This estimator equation is

$$Z^*_{OK}(x_0) = \sum_{i=1}^{n} \lambda_i . Z(x_i)$$

with the sum of $\lambda_i = 1$, where λ_i is the kriging weight of each sampled point Z(x_i). The set of kriging weights that minimize the error variance are retrieved from the ordinary kriging system, written in matrix notation as

$$\begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} & 1 \\ \vdots & \cdots & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

where γ_{nn} is the variance of the distance of the sampled points and γ_{n0} is the variance of the distance between the sampled point to the point being estimated and μ is the Lagrange parameter. More details about geostatistical estimations, mathematical proofs, assumptions, parameters and its applications can also be found also at: Matheron (1978); Journel & Huijbregts (1978); Journel (1989); Goovaerts (1997); Deutsch & Journel (1998); Chilès & Delfiner (1999); Olea (2003).

3. Results and Discussion

The data used are a segment of the 2000 and 2010 Brazilian census and it was provided by IBGE. The chosen variable is the car ownership and its spatial distribution is shown in Figure 3. All the geostatistical procedures were done with SGeMS (Stanford Geostatistical Modeling Software) (Remy et al. (2011)).

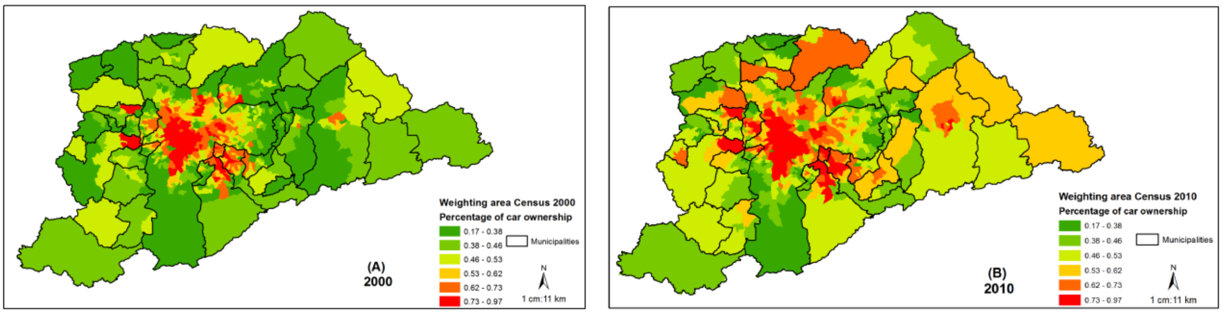

Figure 3 – Car ownership per weighting area in SPMA according Census 2000 (A) and Census 2010 (B).

The experimental variogram were computed and both models were fitted and they are shown in Figure 4 and the parameters of the fitted variogram is shown in Table 1 .

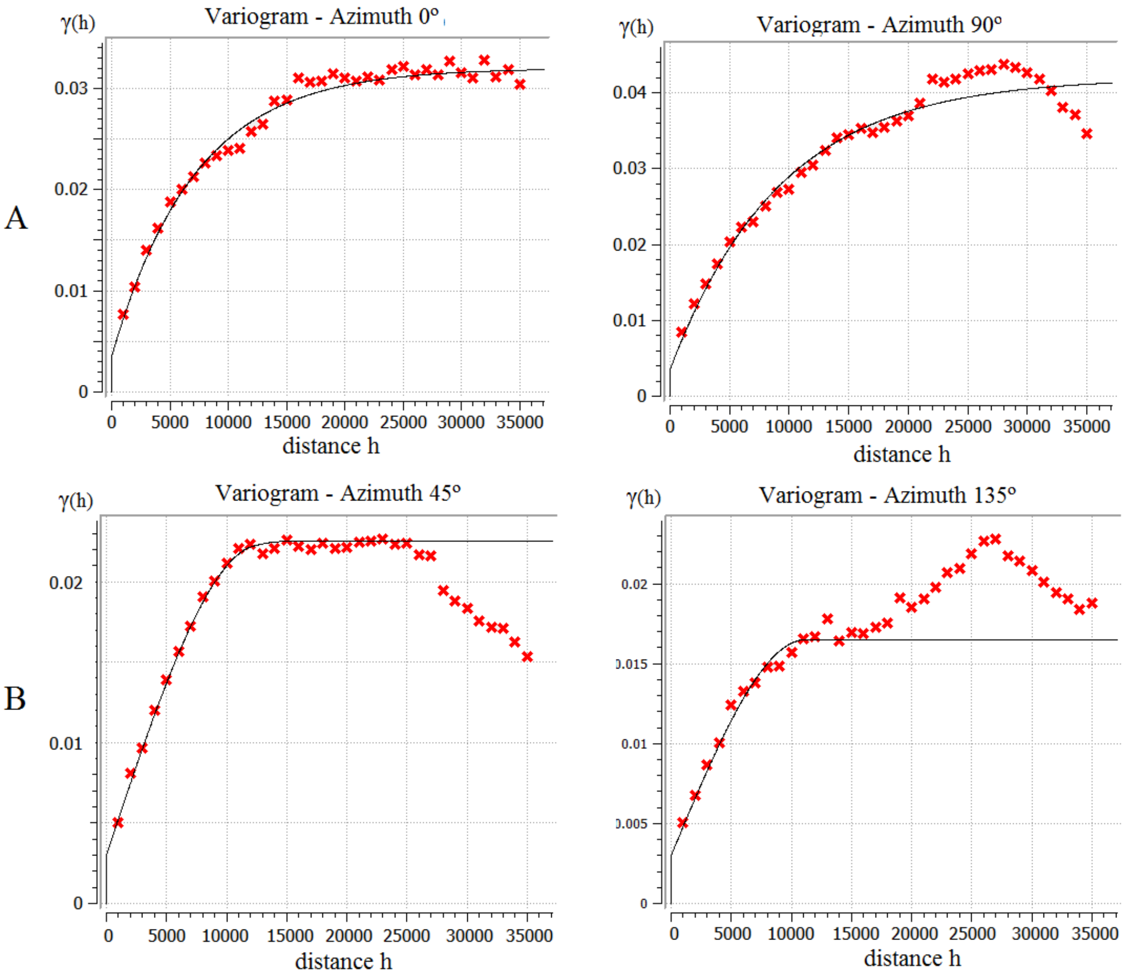

Figure 4 – Experimental variogram and its adjusted model of the car ownership per family by 2000 Census (A) and by 2010 Census (B).

ISBN: 978-85-17-00088-1

Table 1 - Parameter of model variogram fitted for the car ownership variable of Census 2000 and 2010.

	Census 2000		Census 2010	
Model	Exponential		Spherical	
Nugget Effect	0.0035		0.003	
Structures	1	2	1	2
Spatial Variance	0.0285	0.01	0.0135	0.006
Max. Range	26,950	30,000	12,250	15,000
Med. Range	21,350	∞	11,200	∞
Min. Range	0	∞	0	∞
Rotation in Z	90	90	45	45
Rotation in X	0	0	0	0
Rotation in Y	0	0	0	0

The search neighborhood was chosen with the exhaustive cross validations tests. The selected search neighborhood is by octants and it selected at least 2 point and maximum of 8 and 2 is the minimum of non-empty octants, with minimum of 1 point per octant and 2 points maximum. The search ellipse is rotated and have the same ranges of the anisotropy ellipse shown on Table 1. The best search neighborhood found by cross validation show a coefficient of correlation 0.842 for Census 2000 data and 0.805 for Census 2010 data. The maps calculated by ordinary kriging are shown on Figure 5.

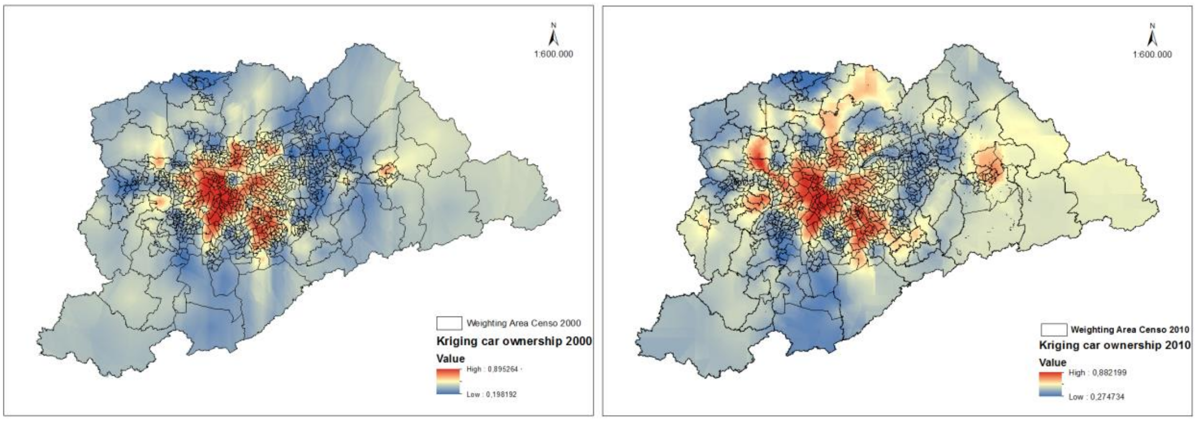

Figure 5 – Kriged car ownership 2000 (A) and 2010 (B).

In order to understand the difference of car ownership in SPMA, the year 2010 map can be subtracted from the year 2000 map and the aftermath highlights the evolution of the car ownership per family (Figure 6). Thus, the deep blue regions (negative values) shows where the car ownership percentage is lower in year 2010. The values in that car ownership remains almost the same are represented between blue and yellow and the growth is in orange and red colors. Moreover, the car ownership percentage per family increased in the adjacent towns of São Paulo (the larger municipality), it shows that in regions where public transport supply is inefficient, population tends to use individual transport as a means of locomotion.

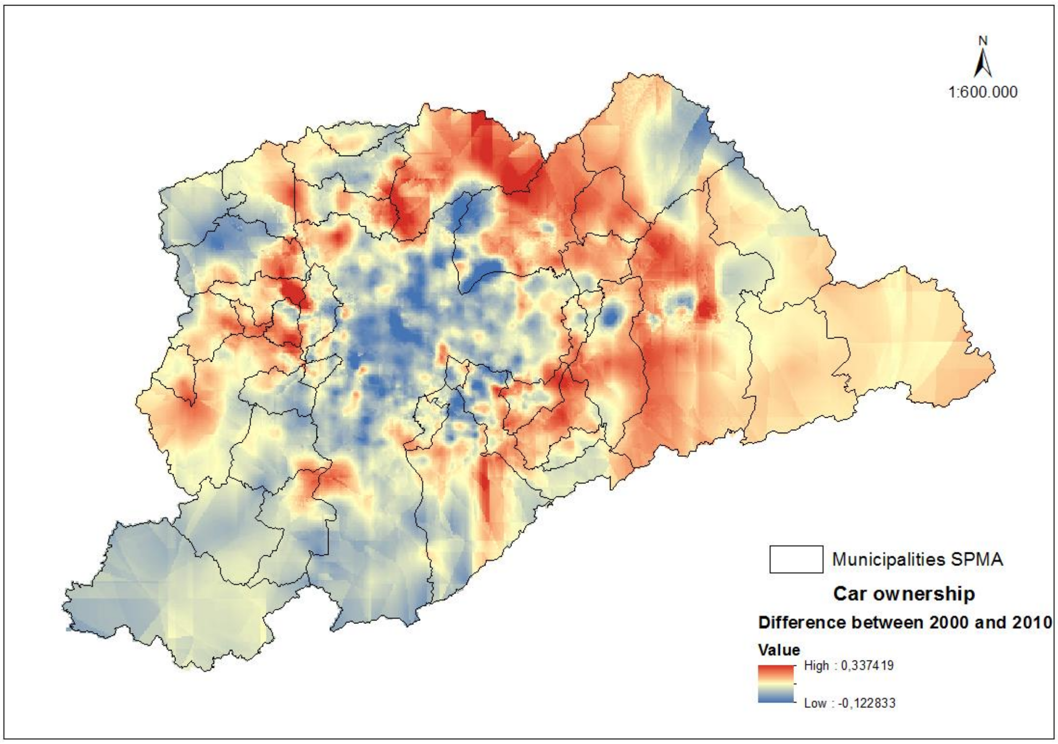

Figure 6 - The difference of car ownership between 2010 – 2000.

4. Final Thoughts

The results of Kriging were satisfactory to understand the spatial distribution of households by car ownership in the SPMA. It was observed that São Paulo neighboring cities had significant growth while São Paulo showed low values of growth and even decline. The western portion of the study area consists of the municipalities of great economic activity.

The south, east and west obtained representative increase in car ownership. These areas also received new installations of residential complexes attracting the middle class population, what could corroborate with the hypothesis raised by Correa (1996) and Aranha (2005) that people tend to live in nearby cities and commute to access jobs, which pressures the ownership level.

The NE portion also experimented the rise of car ownership, comprising the municipality of Guarulhos, which has international status airport. A small patch of decrease between São Paulo and Guarulhos is the region "Parque Ecológico do Tietê" where there is no resident population and therefore.

Finally, it can be observed that the car ownership per household has increased in most of the SPMA, especially in cities where the provision of public transport is not as good as in the state capital.

5. Acknowledgements

The authors are grateful to Geosciences Institute and the Polytechnic School of university of São Paulo for the infrastructure and support during the realization of this study. As well as, CAPES (Coordination for the Improvement of Higher Education Personnel) and CNPq (National Counsel of Technological and Scientific Development) for the authors scholarship.

References

Armstrong M. **Basic linear geostatistics.** Berlin: Springer, 1998. p.15-72

Aranha, V. Mobilidade pendular na metrópole paulista. **São paulo em perspectiva**, v. 19, n. 4, p. 96-109, out./dez. 2005

Chilès J.P.; Delfiner P. **Geostatistics: Modeling Spatial Uncertainty**. New York: Wiley-Interscience Publication, 1999

Correa, R. L. **O Espaço urbano**. Editora Ática, Série Princípios, 3a. edição, n. 174, 1995. p.1-16.

Deutsch C. V.; Journel A.G. **GSLIB: Geostatistical Software Library and User's Guide**. New York: Oxford University Press, 1998. 369 p.

Goovaerts P. **Geostatistics for National Resources Evaluation**. New York: Oxford University Press, 1997

Gomes, V.; Pitombo, C.; Rocha, S.; Salgueiro, A. Kriging Geostatistical Methods for Travel Mode Choice: A Spatial Data Analysis to Travel Demand Forecasting. **Open Journal of Statistics,** v. 6, p. 514-527, 2016. doi: 10.4236/ojs.2016.63044

IBGE. **Censo Demográfico 2010**, Instituto Brasileiro de Geografia e Estatística, 2010.

IBGE. **Censo Demográfico 2000**, Instituto Brasileiro de Geografia e Estatística, 2000.

Isaaks E. H.; Srivastava R. M. **An Introduction to Applied Geostatistics**. New York: Oxford University Press, 1989. p. 278 – 322

Journel A. G. **Fundamentals of Geostatistics in Five Lessons**. Washington D.C.: American Geophysical Union, 1989

Journel A. G.; Huijbregts Ch. J. **Mining Geostatistics**, London: Academic Press, 1978. p. 600

Loureiro, C.F.G.; Silva, H.N.; Carvalho, L.E.X. Metodologia de análise de regressão geograficamente ponderada aplicada ao fenômeno das viagens intermunicipais. Panorama Nacional de Pesquisa em Transportes, XX ANPET, v.1, p.479-491, 2006.

Matheron G., Principles of Geostatistics, Economic Geology 58. In VanLandinghan S. L., 1983, **Economic Evaluation of Mineral Property** (Benchmark papers in geology), Hutchinson Ross Publishing Company, p. 157-177, 1963.

Matheron, G., Estimer et choisir, Les cahiers du CMM de Fontainebleau, Fasc. 7, Ecole des Mines de Paris, 175 p., 1978.

Olea R. A., **Geostatistics for Engineers and Earth Scientists**. New York: Kluwer Academic Publishers, 2003. p. 141-162

Remy N., Boucher A. e Wu J., 2011, **Applied Geostatistics with SGeMS – A User's Guide**. New York: Cambridge University Press. p. 286

SEADE – Fundação Sistema Estadual de Análise de Dados. Portal de Estatística do Estado de São Paulo. São Paulo. Available: < http://www.seade.gov.br/>. Access in: 15/06/2015.

Ritter, N.; Vance, C. Do fewer people mean fewer cars? Population decline and car ownership in Germany. **Transp. Res. Part A Policy Pract**. 50, 74–85., 2013 doi:10.1016/j.tra.2013.01.035

Whelan, G., 2007. Modelling car ownership in Great Britain. **Transp. Res. Part A Policy Pract**. 41, 205–219. doi:10.1016/j.tra.2006.09.013

Wegener, M.; Fürst, F.; Land-Use Transport Interation: State or Art. Publicação 46 (Berichte aus dem Institute für Raumplanung 46). Institute für Raumplanung, Universität Dortmund, Dortmund. 1999.

Yagi, M., Managi, S., 2016. Demographic determinants of car ownership in Japan. **Transp. Policy**. 50, 37–53. doi:10.1016/j.tranpol.2016.05.011