IMPERVIOUS SURFACES FOR POPULATION ESTIMATE IN BRAZILIAN CITIES

Gabriel Crivellaro Gonçalves[1], Maria Isabel Sobral Escada[2], Silvana Amaral[3]

[1,2,3] National Institute for Space Research - INPE, CP 515, CEP 12227-010, São José dos Campos – SP, Brazil,
[1]gabriel.goncalves@inpe.br; [2]isabel@dpi.inpe.br; [3]silvana.amaral@inpe.br

ABSTRACT

Impervious surfaces are directly related to urban development and population growth and they can be used as indicators of urban infrastructures. This work aimed to estimate urban population from impervious surface modeling considering two Brazilian cities of similar population size and different urban context: Sinop, in Brazilian Amazon, state of Mato Grosso, and Guaratinguetá, a city in the Southeast Region, state of São Paulo. A linear spectral mixing model was used to map impermeable surfaces from Landsat images. Using population density from 2010 Census, and impervious surface intensity, a linear model was proposed to estimate population. The absence of linear relations in Sinop indicates that other approaches to relate population and urban structures should be tested for Amazonian cities. In Guaratinguetá, significant R^2 indicates that impervious surfaces is helpful to infer population density. Despite the high average relative errors for both study sites, the methodology demonstrated efficiency in the mapping of impervious surfaces.

***Key words* —** Impervious Surface, Population estimate, Landsat, Census data, modeling.

1. INTRODUCTION

Impervious surfaces (IS) correspond to structures that do not allow the penetration of water, and in urban areas are results of anthropic activities. IS are formed by materials such as concrete, asphalt, ceramic materials and so on, structured on sidewalks, roads and construction [1,2]. Urbanization and population growth are in general directly related to the increase of impermeable surfaces [3]. On the other hand, accurate and timely estimates of population distribution are essential to plan urban growth and to understand the relationships between population growth and social, economic and environmental factors [4]. As an alternative to traditional population counting methods, remote sensing provides tools and methodologies to estimate population in large areas and different dates, through historical imaging series [5].

From the 1970s, remote sensing began to be used to estimate population distribution through four main methods: (a) counting of housing units [6]; (b) relationship between population and built-up areas [7]; (c) classification of land use [8]; (d) spectral characteristics of pixels (pixel-to-pixel classifications) [9]. From the 1990s, SI began to be used as an alternative for population estimates. SI are relatively stable and their proportions within administrative units depend on the nature of land use patterns [5]. For the detection of SI in optical remote sensing data, the Linear Model of Spectral Mixture has often been used in different regions and different scales of analysis [5, 10].

In this context, this work creates a population estimation model from impermeable surfaces and assess its applicability in two cities of similar population, but different urban contexts: Sinop (MT) and Guaratinguetá (SP). From the census data and impermeable surfaces obtained by remote sensing, the adjustments of the demographic density estimates were verified from a linear relation with intensity of imperviousness for the two cities.

2. MATERIAL AND METHODS

2.2. Study area

Sinop is a city located in the middle northern portion of the state of Mato Grosso (Figure 1a). The county population was of 113,099 inhabitants in 2010 [11] and it was estimated in 139,935 inhabitants in 2018 [12]. The city was founded in 1974 by the Sinop colonizer (*Sociedade Imobiliaria do Noroeste Paranaense*), as a result of federal government incentives to occupy the northern region of Brazil. At the beginning of the 1990s, there were more than 500 timber industries (mostly sawmills) installed in the region, the main economy of the municipality. With the reduction of wood stocks, agribusiness started to lead the economy, converting open areas into pastures and later into agriculture [13]. Today Sinop is a micro regional center [11] for services, commerce, health and education.

Guaratinguetá, founded in 1630, is an important city of the Paraíba Valley in the state of São Paulo (Figure 1b). In 2010, there were 112,072 inhabitants and an estimated population of 121,073 inhabitants in 2018 [11, 12]. Its settlement began as a strategic point on the gold route, being the start point to the port of Paraty - RJ. Today, the city is a micro-region seat, and one of the sub-seats of the Metropolitan Region of the Paraíba Valley. Known as the Capital of the “Fundo do Vale”, Guaratinguetá offers several services to neighboring cities and has important touristic and industrial economies [14].

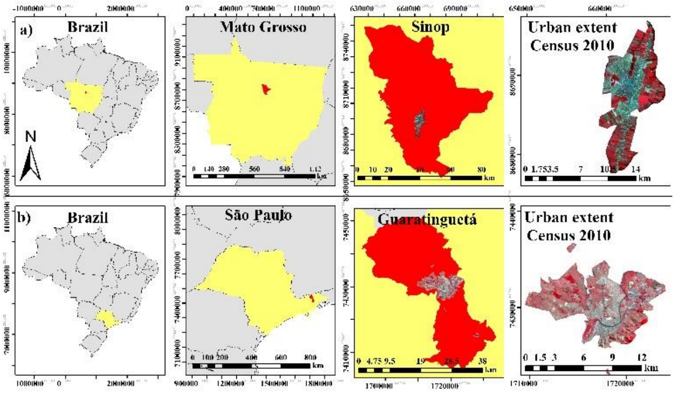

Figure 1. Study areas: a) Sinop, Mato Grosso state, and b) Guaratinguetá, São Paulo state.

2.2. Data

We used Landsat 5 TM images to estimate impervious surface areas and density. TM imagens with geometric and atmospheric correction were obtained from the EarthExplorer platform of the U.S. Geological Survey. Image dates considered the agricultural calendar and the absence of cloud cover. For Sinop (226/63), in 04/04/2010, the agricultural areas around the city were cultivated, preventing bare soil and urban areas class confusion. Because the agricultural calendar is different in the Southeast, for Guaratinguetá (218/76), the 08/01/2010 image was selected, according to the same criterion adopted for Sinop. A Landsat 5, bands 1, 2, 3, 4, 5, and 7 composition was used for image processing and identification of impervious surfaces.

For population data, we used information from the 2010 Population Census[11]. The boundaries of the urban census tracts of the two cities were accessed from the information base of the Brazilian Institute of Geography and Statistics. In addition, the variable regarding population volume within the sector, from the Universe questionnaire, was used to estimate the population density in each sector [11].

2.3. Detection of impervious surface (IS)

Due to its efficiency and simplicity, the Linear Spectral Mixture Model (LSMM) has been broadly applied to map impermeable surfaces in spatial mean resolution analyzes [2, 5]. The MLME assumes that the spectral reflectance of a pixel is a linear combination of the spectral response of the coverages within the pixel. The pure spectra of these characteristics are called endmembers [15].

For Sinop and Guaratinguetá, analyzing the LSMM endmembers for impervious surface mapping, four fractions composed the surfaces (IS): high albedo, corresponding to concrete, roofs, sidewalks and large constructions; low albedo - asphalt, water and shade; vegetation - forests, agriculture and any other vegetation type cover; and bare soil fraction [5].

Using map algebra, MLME fractions images generated binary images from thresholds empirically defined. In the high albedo fraction, mean values above 0.8 correspond to built-up areas. In the low albedo fraction, values between 0.7 and 1.05 correspond to the asphalt, and values above 1.5 correspond to water. In the fraction, all values above 0.3 correspond to the vegetation. And finally, in the soil fraction all values above 0.1 correspond to the exposed soil. The five binary images were then vectorized and joined to generate two new images: IS image, from the union of constructed areas and asphalt images; permeable surfaces (PS) image, from the union of vegetation, exposed soil and water images. Areas mapped at the same time as permeable and impermeable were subtracted from the final SI image.

2.4. Population estimation

The IS intensity in each census tract was calculated as the percentage of intersection of the IS class within each census tracts, using the landscape metric *pland* in GEODMA [16]. The IS intensity (1-100%) was used as an independent variable for the population estimation model. The dependent variable population density was calculated from the ration of the population volume by the area of the census sector, in inhabitants/km². For the linear regression (RStudio software), half of the samples (census tracts) were used to create the model and the other half for model validation.

For the model accuracy assessment, in addition to R^2, the relative error was also used to evaluate the performance of the model based on validation data (50% of the samples). Regression residues for some cases may be negative or positive. The mean of the residual errors (MRE) is usually influenced by extreme values and may not be able to evaluate the performance of the model. Therefore, the median relative error (MdRE) and standard deviation (SDRE) were also calculated.

3. RESULTS AND DISCUSSION

In Sinop, regions with highest population densities presented lower IS intensities (Figure 5). Only the streets in the central (commercial) and industrial regions are paved (asphalt). In other residential districts, there are only dirty streets, as also observed in high resolution 2010 images from Google Earth.

The lack of spatial correspondence between population density and IS intensity in Sinop (Figure 5) is also evident from the dispersion plot (Figure 6). Lu et al [5] stated that the range of 800 to 3000 hab/km² would be ideal to explore linear relations between these variables (Figure 6b). They also suggested removing rural and industrial areas or IS intensity of less than 25% and greater than 75% (Figure 6c). Even with these restrictions, there was no fit improvement for the Sinop model.

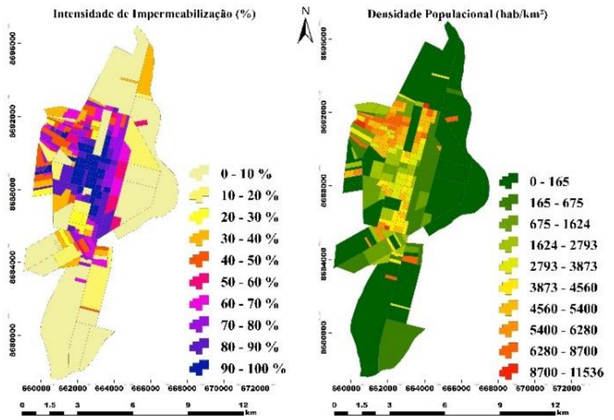

Figure 5: Impervious surface intensity (%) and population density (hab/km²) by census tracts, in Sinop - MT.

Lu et al [5] found an average relative error of 35 for Marion County - Indianapolis – USA (860454 inhabitant), while the average relative error for Sinop was 169.8 (Figure 6d). The urban context and occupation history of Amazon cities are incomparable even to cities in the south of Brazil. Both construction patterns and population concentration depend on the local context, and therefore direct comparison would be misleading.

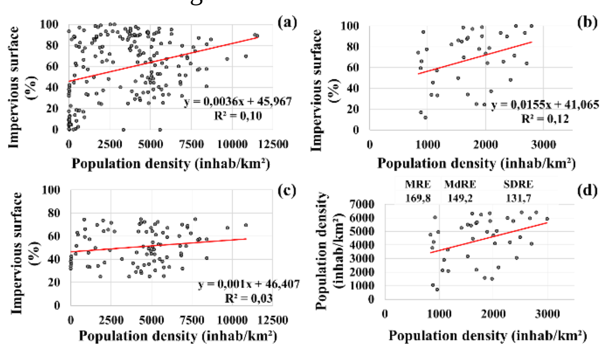

Figure 6: Linear regression between population density and impermeable surfaces for Sinop - MT.

Guaratinguetá, in the state of São Paulo, presented better correlations than those observed for Sinop. This can be explained by the more consolidated urban structure and higher IS intensities, characteristics of the cities in Paraíba Valley. Sectors with higher population densities coincide with those with higher IS intensities (Figure 7). However, there is city verticalization in the central region, and thus, the population density reaches 18000 inhab/km² in central census tracts. Vertical areas compromise population estimation based only on IS intensities [10].

The results of the population estimation model for Guaratinguetá (Figure 8) in terms of R^2 was significantly better than Sinop. Excluding the verticalized sectors, with more than 13500 inhab/km², the R^2 was from 0.36 to 0.5 (Figure 6d). Even so, the mean relative error was almost four times greater than that found for Sinop and more than ten times that found by Lu et al [5]. Applying the same data interval previously suggested, both the ideal population density range and the removal of rural and industrial areas resulted in weaker adjustments, as observed in R^2 values (Figure 8 a, b c and d).

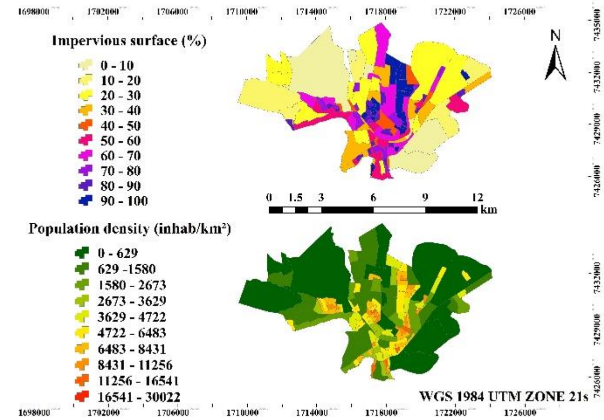

Figure 7: Impervious surface intensity (%) and population density (hab/km²) by census tracts, in Guaratinguetá - SP.

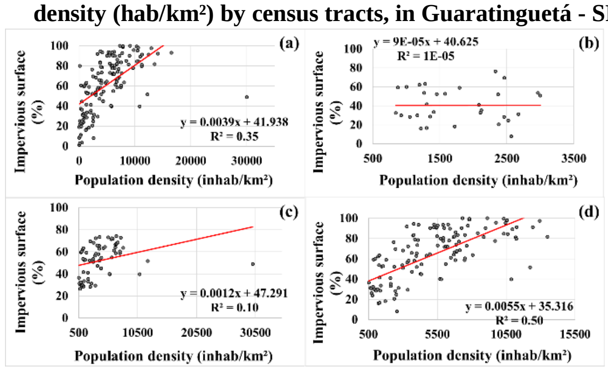

Figure 8: Linear regression between population density and impermeable surfaces for Guaratinguetá - SP.

The dispersion of observed and predicted values (Figure 9) suggest a linear relationship between them. However, the relative error demonstrates low accuracy and precision of the model. Perhaps a non-linear relationship would be more adequate to fit this model.

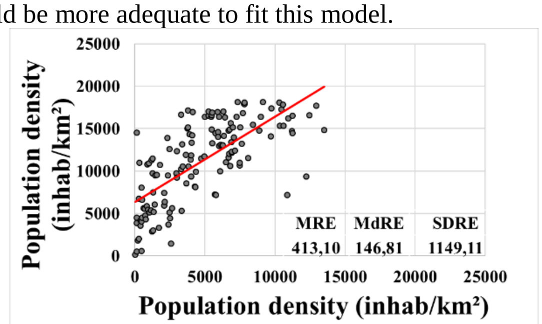

Figure 9: Dispersion between predicted population density (y) and observed (x) in Guaratinguetá.

5. CONCLUSIONS

This work highlights the usefulness of remote sensing and geoinformation to explore the relationships between population and space, considering impermeable surfaces and urban population estimates. Applying LSMM in Landsat images, we obtained the impervious surfaces for two cities - Sinop and Guaratinguetá - of similar population size, but with different urbanization history and contexts. The IS intensity information for urban census tracts and their relationship with population density were explored by linear regression modeling.

In Sinop, like most Amazonian cities, the majority of the city area in 2010 is permeable: mostly dirty streets and dwellings with little impervious areas. These conditions justify the non-existent correlation between population density and IS. These peculiarities of Amazonian cities demand the search for alternative approaches to study urban population estimates, since IS areas does not reflect population density.

In turn, the model for Guaratinguetá showed better adjustments. The higher IS intensity degree allowed, at least, to explore the relationships between IS and urban population. A limitation of this approach were the vertical areas at city center. Despite similar IS intensities, population densities higher than horizontal residential regions impairs the fit of the proposed model.

Different approaches, such as nonlinear relationships, or use of high-resolution images, or image fusions should be tested as alternatives to improve estimation models [17,18]. In addition, these analyzes should include the history and patterns of urbanization, and their implications for population distribution, so that remote sensing data and techniques would be more efficient in supporting urban population estimates.

ACKNOWLEDGEMENT

The authors thank National Institute of Space Research - INPE and MSA Project - BNDES (Number process: 1022114003005) - Sub-Project 5 - Trajectories of Patterns and Processes in the Characterization of New Dynamics of Deforestation in the Amazon for infrastructure and financial support.

6. REFERENCES

[1] Arnold, C. L., & Gibbons, C. J. (1996). Impervious surface coverage – The emergence of a key environmental indicator. *Journal of the American Planning Association*, 62(2), 243–258. 1996.

[2] Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends, *Remote Sensing of Environment*, 117:34–49. 2012.

[3] Sutton, P. C., Elvidge, C. D., Tuttle, B. T., & Ziskin, D. A 2010 mapping of the constructed surface area density for S.E. Asia – Preliminary results. In Proceedings of the 30st Asia-Pacific Advanced Network Meeting (pp. 182–190). 2010.

[4] Januzzi, P. M. *Indicadores Sociais no Brasil*. ed. 3. Alínea Editora. São Paulo. 2001.

[5] Lu, D. Weng, Q. Li, G. Residential population estimation using a remote sensing derived impervious surface approach. *International Journal of Remote Sensing*. 2007.

[6] Lo, C., Accuracy of population estimation from medium-scale aerial photography, Proceedings of the American Congress on Surveying and Mapping and American Society for Photogrammetry and Remote Sensing, Annual Convention, Washington, D.C., pp. 1–10. 1986.

[7] Lu, Z.Y., J. Im, and L. Quackenbush. A volumetric approach to population estimation using lidar remote sensing, *Photogrammetric Engineering & Remote Sensing*, 2011.

[8] Langford, M., Maguire, D.J. and Unwin, D.J. The areal interpolation problem: estimating population using remote sensing in a GIS framework. In Handing Geographical Information: Methodology and Potential Applications, L. Masser and M.Blakemore (Eds), pp. 55–77 (New York: Longman Scientific & Technical/John Wiley & Sons, Inc.). 1991.

[9] Iisaka, J., and E. Hegedus. Population estimation from Landsat imagery, *Remote Sensing of Environment*, 12(4):259–272. 1982.

[10] Wu, C.S., and A.T. Murray, 2003. Estimating impervious surface distribution by spectral mixture analysis, *Remote Sensing of Environment*, 84(4):493–505. 2003.

[11] Instituto Brasileiro de Geografia e Estatística (IBGE). Censo. Brasil, 2010.

[12] Instituto Brasileiro De Geografia e Estatística. IBGE Cidades. Rio de Janeiro, 2018.

[13] TEIXEIRA, L. A colonização no norte de mato grosso: o exemplo da gleba celeste. Presidente Prudente, 2006: Dissertação de Mestrado. Unesp, p. 117.

[14] Sampaio, H.C. Dias, R.A. Balestieri, J.A.P. Sustainable urban energy planning: The case study of a tropical city. *Applied Energy*.v.104. p.924-935. 2013.

[15] Shimabukuro, Y.E. and Smith, J.A. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. IEEE Transactions on Geoscience and Remote Sensing, 29, pp. 16–20. 1991.

[16] Körting, T. S. et al. GeoDMA: A Novel System for Spatial Data Mining. In: INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS, 1.,2008, Pisa. Proceedings. [s.l.]: Ieee, 2008. p. 975 - 978.

[17] Zhu, H. Li, Y. Liu, Z. Fu, B. Estimating The Population Distribution in a County Area in Chine Based on Impervious Surface. *Applications Paper*. p. 155- 163. 2015.

[18] Lu, D. Moran, E. Hetrick, S. Detection of Impervious Surface Change With Multitemporal Landsat Images in as Urban-Rural Frontier. *Journal of Photogrammetry and Remote Sensing*.P. 298 – 306. V.66. 2010.