Changing spatial perception: dasymetric mapping to improve analysis of health outcomes in a megacity

Ligia Vizeu Barrozo a, Reinaldo Paul Pérez-Machado a, Christopher Small b and William Cabral-Miranda a

a Departamento de Geografia, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, Brazil; b Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

ABSTRACT

Choropleth representation has been the most widely applied method to represent rates in disease maps due to its consistency in depicting relative data. However polygons in a choropleth map may give the erroneous notion of homogenous distribution over area in cases where the mapped quantity varies in its spatial distribution. In the case of population maps, choropleth maps suggest uniform distribution of people within large peri-urban administrative areas where population is known to be unevenly distributed within the administrative units. Dasymetric mapping can provide a more accurate and detailed distribution of population data by using ancillary information to spatially disaggregate population within administrative units. We have developed a procedure to use more detailed fiscal cadastre blocks to disaggregate census data within less detailed enumeration and sample areas. Here we explain the procedure and provide simple examples of this dasymetric representation as applied to population density, socioeconomic and health indicators. This approach may help to identify fine-scale risk patterns of infectious and chronic diseases and associated socioeconomic or environmental risk factors. It is hoped that better visualization through this approach will help specialists in planning to reduce social injustice in complex urban environments.

1. Introduction

In recent years, spatial expression of health outcomes and social conditions has become central to public health and geographic research (Matisziw, Grubesic, & Wei, 2008). Maps have been recognized as powerful tools not only because they communicate geographic facts but also because they lead to spatial reasoning. They have been widely used to identify risk areas and possible risk factors, to evaluate access to health systems and for health planning and management. Some authors have specifically emphasized the importance of maps as an effective tool for cancer control planning and evaluation (Bell, Hoskins, Pickle, & Wartenberg, 2006; Parrott, Hopfer, Ghetian, & Lengerich, 2007).

Disease maps have been used for a long time, even before the incorporation of cartographic principles for thematic representation. For instance, some maps recovered by Koch (2005) depict quarantines used during the black plague epidemic in Bari, Italy, in the end of the seventeenth century. At that time cartographic representation did not assimilate abstract elements as rates. Alongside the evolution of cartography, statistics provided methods to treat quantitative data as linear arithmetic, necessary to the development of rate maps. The first known choropleth map was designed by C. Dupin, in 1826 to represent popular education in France, using hues to indicate levels of increasing values (Cauvin, Reymond, & Serradj, 1987). Since then, choropleth representation has been widely applied to represent rates due to its consistency in depicting relative data (Richards et al., 2010a, 2010b).

Intrinsic to the definition of the administrative unit, polygons in a choropleth map may give the erroneous notion of homogenous population distribution over area. For example, this results in obvious misrepresentation when uninhabited areas such as water bodies, forested land and other uninhabited areas are not excluded from administrative units. Benjamin Semenov-Tian-Shansky first directly addressed this issue by introducing the dasymetric concept in his report to the Russian Geographic Society in 1911 (Kamenevskiy, 1930, apud Petrov, 2008). According to him, on dasymetric maps ‘(…) population density, irrespective of any administrative boundaries, is shown as it is distributed in reality. That is, by natural spots of concentration and rarefaction’ (cited in Kamenevskiy, 1930, p. 176; translated by Petrov, 2008). The prerogative for dasymetric disaggregation is the availability of more detailed ancillary data about spatial population distribution. Dasymetric mapping can be especially useful to better represent socioeconomic and health indicators. Some studies have applied the dasymetric technique in crime mapping (Bowers &
Dasymetric mapping can improve cartographic accuracy even in countries with very detailed census enumerations, like Brazil. The municipality of São Paulo has a population of 11.2 million in a territory of approximately 1,520 km² with a high degree of urbanization. The administrative districts with lower median income are generally on the periphery of the city while the wealthier districts are generally in the inner city. Analyzing the variability of demographic trends, real estate development, urban sprawl, deforestation and urban infrastructure in the metropolitan region of São Paulo, Torres, Alves, and Oliveira (2007) verified that population growth occurs due to the extension of low-income areas, predominantly in the suburbs. This pattern is contrary to the pattern typically observed in North American cities where the suburban periphery is generally associated with medium to higher income levels. In the Brazilian metropolitan areas, land market dynamics, affected by land-use regulations, as well as by public policies for transportation and housing, strongly influence the urban sprawl process and its environmental impact. São Paulo has been losing population in the same places where real estate investments are growing more significantly. Population density increases mainly where the price of land is low (Torres et al., 2007). Cartographic misrepresentation affects mainly those lower income areas, impairing urban planning and reinforcing health inequalities.

In order to improve the accuracy and detail of population density maps of São Paulo, we have developed a procedure to increase spatial resolution by linking higher spatial resolution fiscal cadastre blocks to lower resolution census data at the enumeration and sample area (SA) levels. Here we present such a procedure and provide simple examples of dasymetric representation applied to population density, socioeconomic and health indicators. We believe that this approach can be adapted to other localities to better identify demand for services, and for evaluation of strategies to address urban planning and health issues.

2. Changing spatial perception

Dasymetric mapping has a direct effect on population density calculation. More reliable statistics may better guide urban policies regarding social housing or transportation planning, thereby improving accessibility for all. Maps 1A and 1B (Main Map) depict the choropleth and dasymetric approaches for population density. The dasymetric map may change perception about the city due to its visual impact when compared to the traditional choropleth approach (Map 1A). Map 1B shows changes in population density and extent that result from the spatial constraints used in the dasymetric map. Dasymetric population density increases within districts vary from 36% to 10 times the densities represented by the choropleth approach. Thus, the dasymetric map helps to better understand what Solà-Morales Rubió (1992) calls pseudo-density. According to Sales (2012), adapted to the reality in the outskirts of São Paulo, the pseudo-density is the coexistence of the high density of the slums with the urban voids interspersed among these settlements, thus creating a lower average density for these areas. For instance, in the southern periphery, the district with the largest population in the city (the district of Grajaú with more than 360,000 inhabitants) is 5 times denser than previously calculated based on the choropleth approach. This periphery is occupied by precarious settlements and slums in irregular division of lots. High population density in this area is especially consequential because of the locations of the slums along the shores of the main reservoirs, which supply nearly one-third of the water São Paulo consumes (Leite, 2012). The lack of sewage infrastructure in these informal settlements results in adverse public health consequences within the settlements as well as contamination of the general water supply for much of São Paulo.

Another impact is related to the perception of intra-district social contrasts. Socio-spatial inequalities are easily observed. Map 2A (Main Map) depicts mean monthly income by householder. Important spatial inequality is remarkable in the west circle where the light orange class depicts the second largest slum of the municipality (Favela de Paraisópolis) within a high-income neighborhood.

From 2006 to 2009 the main causes of infant mortality in São Paulo were related to neonatal mortality. 11.5% of deaths were caused by bacterial sepsis of newborns (P36.9, according to the International Statistical Classification of Diseases and Related Health Problems 10th Revision – ICD-10) and 9.3% by respiratory distress syndrome of newborns (P22.0). Other causes were congenital malformation of heart (Q24.9) with 4.3% and bronchopneumonia (J18.0) caused 3.1% of deaths. In Brazil, neonatal mortality is related to care
during labor and delivery (Lansky et al., 2014). Then, identifying risk areas in a more accurate map may help to make a better evaluation of access to the health system. Comparing mean monthly income (Map 2A – Main Map) to relative risk of infant mortality (Map 2B – Main Map), we observe that Favela de Paraisópolis is in the same class as part of its surrounding neighborhood, possibly due to the better health service provided in the wealthier area.

3. Methods

3.1. Data and procedures

The census boundary units used in the present study are enumeration areas (EAs) and SAs. The EA is the smallest geographic entity for which the Brazilian Census tabulates decennial census data. It is roughly equivalent to the Census Block Group in the US Census. An EA is a territorial unit for data collection in census operation with definite borders in continuous areas, based on political and administrative divisions, sized in terms of territory and number of housing units that can be visited by only one enumerator. Generally this equals 250–300 households. Boundaries of units change as population grows to maintain stable population numbers within units. Average population in EAs in São Paulo in 2010 was 611 (std. dev. = 313). SAs are the geographic level defined to apply statistical procedures that allow the use of sample surveys as valid to the entire population. Thus, an SA encompasses a set of EAs.

The municipality of São Paulo comprises 18,435 EAs grouped in 310 SAs (Instituto Brasileiro de Geografia e Estatística, 2011). EAs boundaries are defined and stored as closed polygons by the Brazilian Institute of Geography and Statistics (IBGE), responsible for the demographic census in Brazil. This study refines the spatial resolution of EAs by incorporating underlying land use and fiscal cadastre blocks. Dasymetric mapping was implemented using the ArcGIS geographic information system (GIS) to combine multiple ancillary data-sets to detect uninhabited lands and nonresidential blocks in the area. Ancillary data include the digital cartographic database of the city blocks, the database of municipal fiscal cadastre and land cover information derived from Landsat satellite images with a spatial resolution of 30 m and digital aerial orthophotos with mean spatial resolution of 45 cm. All ancillary data correspond to the 2010 base year.

The cartographic database of the city blocks was provided by the Secretariat of Finance and Economic Development as the Digital Map of the City of São Paulo. This map was originally digitized from aerial photographs at a 1:7500 scale and updated with field observations in 2004 and 2005.

Current land cover information is derived from a multi-season composite of Landsat satellite imagery. In contrast to error-prone discrete thematic classification of land cover, we use a linear spectral mixture model to estimate the relative areal fractions of vegetation, water, shadow, soil and impervious substrate within each Landsat pixel (Small, Perez-Machado, Barroso, & Luchiari, 2015). Multi-season imagery is used to distinguish spectrally stable impervious substrate from spectrally similar but seasonally variable rock and soil substrates. Built-up areas are characterized by sub-pixel mixtures of spectrally invariant impervious substrate and persistent shadow while pervious soils are characterized by changing reflectance resulting from soil moisture variations. The resulting land cover fraction map is able to represent continuous gradients in land surface properties (like vegetation abundance) as well as sharp discontinuities related to contrasting land cover types. These multi-season land cover fraction images have been used to map decadal changes of land use throughout São Paulo and the surrounding area.

In the peri-urban interface, where expansion pattern corresponds to isolated built-up areas smaller than the spatial resolution of Landsat images, we complemented land-use mapping through visual classification of digital orthophotos.

Next, we overlaid EAs on urban blocks and peri-urban built-up areas to capture identity features of the census to each polygon on the map. SAs codes were linked through database procedures. Polygons where the fiscal cadastre database shows no residential land use or population was equal to zero in the EA of the census database were designated as nonresidential.

3.2. Dasymetric representation of population density, income and health outcome

Population density

In the periphery of the large metropolises in Brazil, there is a complex mixture of formality and informality due to high demand for housing. The common solution found for low-income population results in the development of informal precarious settlements and slums (favelas), increasing demographic density in the outskirts (França & Barda, 2012). Thus, to achieve a more reliable population density by district, we compared the produced population density maps based on the choropleth and dasymetric approaches (Maps 1A and 1B, on Main Map respectively). In the former, the entire area of the district was used for calculation. In the dasymetric approach, only the effective residential area was used.

Socioeconomic variable

The dasymetric base map of the territory of São Paulo was designed to allow joining socioeconomic variables

Development as the Digital Map of the City of São Paulo. This map was originally digitized from aerial photographs at a 1:7500 scale and updated with field observations in 2004 and 2005.
using SA and EA identifiers. As an example, we produced a map of mean monthly income by household by SA.

Because of the large size of the study area and the fine detail provided by the census data, we show details of the mapping with inset examples. This is illustrated in the 4 circles in more detail (~1:60,000 scale) on Map 2A (Main Map). The classification method was defined according to guidelines from Cauvin et al. (1987, pp. 76–77), applying Natural breaks (Jenks’ algorithm). Color schemes for each of the maps were selected using the ColorBrewer procedure (Brewer, MacEachren, Pickle, & Herrmann, 1997).

Health outcome
We calculated relative risk of infant mortality aggregated by SA using numbers of deaths in young children between birth and 1 year of age occurred from 2006 to 2009. Anonymized residential addresses of deceased children were provided by DATASUS, Data Processing of the Department of Brazilian Public Health System and geocoded by the São Paulo State System for Data Analysis Foundation (SEADE). Relative risks were derived from indirect standardization taking into account gender as covariate, using the software SaTScan™ v9.3 (Kulldorff, 1997). The expected number of infant deaths for each SA, adjusting for gender, was calculated using the respective internal standard of the gender-specific population in the Municipality of São Paulo. With indirect standardization, estimates of rates and relative risks have lower variance (Kulldorff, 2003), which is especially important for small areas such as SAs. The indirectly standardized covariate adjusted Relative Risk (RRadj) (Kulldorff, 2014), is:

$$RR_{adj} = \frac{\sum c_s}{\sum n_s} \cdot \frac{C_s}{N_s},$$

where $c_s$ is the observed number of cases in gender group $s$ in the SA, $n_s$ is the population in gender group $s$ in the SA, $C_s$ is the observed number of cases in gender group $s$ in the Municipality of São Paulo and $N_s$ is the population in gender group $s$ in the Municipality of São Paulo. A RR value below 1 means that it occurred less cases than expected for that location taking into account the standard of the gender-specific population for the municipality.

Population data by SA were derived from the 2010 census.

The map was produced by joining relative risks to the dasymetric cartographic base in the GIS. Classification method and color scheme were defined as in Map 2A (Main Map).

Additionally we prepared a detailed map (Figure 1) to depict the potential of the dasymetric representation to help in a better assessment of environmental health risk when socioeconomic data (e.g. mean dwellers by household) are overlaid by a pollution source (gas station).

Figure 1. Detailed dasymetric representation of socioeconomic data overlaid by pollution source locations and their corresponding Euclidian buffers, in São Paulo, Brazil.
4. Discussion and conclusions

We adapted a method to visualize morbidity, mortality and socioeconomic data using the fiscal cadastre block level for geographic analysis by EA and SA applying dasymetric concepts to the municipality of São Paulo. This has been accomplished by using a comprehensive dataset as ancillary data such as demographic census, municipal fiscal cadastre, Landsat satellite images and, digital aerial orthophotos, conferring enough reliability to support urban and health planning.

Favelas have evolved within Brazilian cities and are part of their urban morphology. In general they occupy public lower soil stability risk areas or environmentally protected areas, increasing erosion and causing pollution to water courses. As consequence of the housing problems, mobility poses an additional challenge to be resolved. According to contemporary urbanism this precariousness requires interventions to install basic infrastructure, recognizing this ‘urban plurality’ as belonging to the city, with the focus on landscape as an element for urban regeneration (França, 2012). This entire scenario requires integrated actions based on the analysis of multiple scales which cannot be addressed through the visualization of the current choropleth approach. This dasymetric map linked to the Census database in the Municipality of São Paulo is the first attempt to contribute to the urban planning tasks to address some of these issues which are usual in the current Brazilian settings.

In addition to interventions of urban design, this approach may help in health planning, by disclosing fine-scale risk patterns of infectious and chronic diseases. Although the dasymetric approach does not result in a more accurate risk analysis, the better geospatial visualization of risk in relation to the actual population location makes it possible to study the spatial clusters of morbidity and mortality within the geographic context of their occurrence. Further, this approach may help to define neighborhoods and to extract data for multi-level analysis in epidemiologic studies.

More accurate maps of population distribution may also help to improve accessibility measures, mainly in the peri-urban areas. Universal access to the health care system has been believed to contribute in the reduction of health inequalities.

It is hoped that better visualization of health and census data through the dasymetric approach will help health and urban specialists in the planning purposes to allow social inclusion and reduce social injustice in complex urban environments.

Acknowledgements

We are grateful to Empresa Paulista de Planejamento Metropolitano (EMPLASA), Secretaria Municipal de Finanças e Desenvolvimento Econômico da Prefeitura de São Paulo and SEADE (Fundação Sistema Estadual de Análise de Dados) for data availability.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the State of São Paulo Research Foundation – FAPESP [grant number 2010/12619-7], [grant number 2013/17617-0] and the National Council of Technological and Scientific Development-CNPq [grant number 303028/2013-0]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Software

The dasymetric map was produced using ESRI ArcGIS 10.1. ENVI v4.7 was used for satellite image analysis. The numeric database was linked to the cartographic base in Maptitude 2014. SPSS 17.0 was used to aggregate Census data by EA to SAs and to calculate socioeconomic indicators.

ORCID

Lígia Víceu Barrozo https://orcid.org/0000-0001-7322-6789

References

Bell, B. S., Hoskins, R. E., Pickle, L. W., & Wartenberg, D. (2006). Current practices in spatial analysis of cancer data: Mapping health statistics to inform policymakers and the public. International Journal of Health Geographics, 5(1), 49. doi:10.1186/1476-072X-5-49

Bowers, K., & Hirschfield, A. (1999). Exploring links between crime and disadvantage in north-west England: An analysis using geographical information systems. International Journal of Geographical Information Science, 13(2), 159–184. doi:10.1080/1365881992414109

Brewer, C. A., MacEachren, A. M., Pickle, L. W., & Herrmann, D. (1997). Mapping mortality: Evaluating color schemes for choropleth maps. Annals of the Association of American Geographers, 87(3), 411–438. doi:10.1111/1467-8306.00061

Cauvin, C., Reymond, H., & Serradj, A. (1987). Discrétisation et représentation cartographique: Maison de la Géographie. Montpellier: GIP RECLUS, coll. Reclus modes d’emploi.

França, E. (2012). Territory of informality or expressions of territory in the Contemporary City? In E. Franca & M. Barda (Eds.), Entre o céu e a água: O Cantinho do Céu (A. A. Puntch, Trans.) (pp. 170–177). São Paulo: Prefeitura de São Paulo, Habisp.

França, E., & Barda, M. (2012). Entre o céu e a água: O Cantinho do Céu [Sky and Water, the living in between: The Cantinho do Céu] (A. A. Puntch, Trans.). São Paulo: Prefeitura de São Paulo, Habisp.

Instituto Brasileiro de Geografia e Estatística. (2011). Base de informações do Censo Demográfico 2010: Resultados do
Universo por Setor Censitário. Rio de Janeiro: Centro de Documentação e Disseminação de Informações.
Koch, T. (2005). Cartographies of disease: Maps, mapping, and medicine (Vol. 1). Redlands, CA: ESRI Press.
Kulldorff, M. (1997). A spatial scan statistic. Communications in Statistics: Theory and Methods, 26 (6), 1481–1496. doi:10.1080/03610929708831995
Kulldorff, M. (Producer). (2003, 01 May 2015). CDC presentation: Mapping rates and proportions. Retrieved from http://www.satscan.org/presentation/
Kulldorff, M. (2014). SaTScan user guide for version 9.3 (Version 9.3). Retrieved from http://www.satscan.org/
Langford, M., & Higgs, G. (2006). Measuring potential access to primary healthcare services: The influence of alternative spatial representations of population. The Professional Geographer, 58(3), 294–306. doi:10.1111/j.1467-9272.2006.00569.x
Lansky, S., Friche, A. A. D. L., Silva, A. A. M. D., Campos, D., Bittencourt, S. D. D. A., Carvalho, M. L. D., ... Cunha, A. J. L. A. D. (2014). Pesquisa Nascer no Brasil: perfil da mortalidade neonatal e avaliação da assistência à gestante e ao recém-nascido. Cadernos de Saúde Pública, 30, S192–S207. doi:10.1590/0102-311×00133213
Leite, R. P. (2012). Democratic water. In E. França & M. Barda (Eds.), Entre o céu e a água: O Cantinho do Céu (A. A. Puntch, Trans.) (p. 166). São Paulo: Prefeitura de São Paulo, Habisp.
Linard, C., Gilbert, M., Snow, R. W., Noor, A. M., & Tatem, A. J. (2012). Population distribution, settlement patterns and accessibility across Africa in 2010. PLoS One, 7(2), e31743. doi:10.1371/journal.pone.0031743
Maantay, J. A., Maroko, A. R., & Porter-Morgan, H. (2013). Research note – A new method for mapping population and understanding the spatial dynamics of disease in Urban Areas: Asthma in the Bronx, New York. Urban Geography, 29(7), 724–738. doi:10.2747/0272-3638.29.7.724
Matisziw, T. C., Grubesic, T. H., & Wei, H. (2008). Downscaling spatial structure for the analysis of epidemiological data. Computers, Environment and Urban Systems, 32(1), 81–93. doi:10.1016/j.compenvurbsys.2007.06.002
Parrott, R., Hopfer, S., Ghetian, C., & Lengerich, E. (2007). Mapping as a visual health communication tool: Promises and dilemmas. Health Commun, 22(1), 13–24. doi:10.1080/10410230701310265
Petrov, A. N. (2008). Setting the record straight: On the Russian origins of dasymetric mapping. Cartographica: The International Journal for Geographic Information and Geovisualization, 43(2), 133–136. doi:10.3138/carto.43.2.133
Poulsen, E., & Kennedy, L. W. (2004). Using dasymetric mapping for spatially aggregated crime data. Journal of Quantitative Criminology, 20(3), 243–262. doi:0748-4518/04/0900-0243/0
Richards, T. B., Berkowitz, Z., Thomas, C. C., Foster, S. L., Gardner, A., King, J. B., ... Royalty, J. (2010a). Choropleth map design for cancer incidence, part 1. Preventing Chronic Disease Public Health Research, Practice and Policy, 7(1), 1–6.
Richards, T. B., Berkowitz, Z., Thomas, C. C., Foster, S. L., Gardner, A., King, J. B., ... Royalty, J. (2010b). Choropleth map design for cancer incidence, part 2. Preventing Chronic Disease Public Health Research, Practice and Policy, 7(1), 1–6.
Sales, M. M. L. D. (2012). Urbanization and transformation of the landscape. In E. França & M. Barda (Eds.), Entre o céu e a água: O Cantinho do Céu (A. A. Puntch, Trans.) (pp. 183–187). São Paulo: Prefeitura de São Paulo, Habisp.
Small, C., Perez-Machado, R., Barrozo, L. V., & Luchiari, A. (2015). Mapping decades of urban growth with multitemporal spectral mixture models. In INPE (Ed.), XVII Simpósio Brasileiro de Sensoriamento Remoto – SBSR (pp. 7088–7095). João Pessoa-PB: Instituto Nacional de Pesquisas Espaciais.
Solá-Morales Rubió, M. D. (1992). Projectar la perifèria. UR: urbanismo revista, (9–10), 2–4
Torres, H., Alves, H., & Oliveira, M. A. D. (2007). São Paulo peri-urban dynamics: Some social causes and environmental consequences. Environment and Urbanization, 19(1), 207–223. doi:10.1177/0956247807076784