

IUSSP XXVIII

IPC



2017 International Population Conference



Cape Town

29 October - 4 November 2017

Cape Town International Convention Centre

Information Bulletin 2



Improving Population Mapping and Exposure Assessment: 3-Dimensional Dasymetric Disaggregation in New York City and São Paulo, Brazil

Juliana Maantay and Andrew Maroko

City University of New York, Lehman College, Earth, Environmental, and Geospatial Sciences Department, and the Graduate School of Public Health and Health Policy

Abstract

Dasymetric mapping is a process of disaggregating spatial data from a coarser to a finer unit of analysis, using additional (or “ancillary”) data to refine the locations of population and achieve greater accuracy. Disaggregating population data reported by census tracts or other administrative/political geographic units can provide a more realistic depiction of actual population distribution and location. This is particularly important in assessing environmental exposures and impacts. Additionally, since exposures occur in 3-dimensions (for instance, air pollution is a 3-dimensional phenomenon), modeling residential population in 3-dimensions may produce more reliable estimates of exposure. Population exposure estimates are improved through dasymetric disaggregation and 3-D extrusion, using a combination of cadastral data (residential area by property tax lot), building footprint data, and building height data. Population in census units is dasymetrically disaggregated into individual buildings using residential area derived from property tax-lots and then extruded vertically based on building height. This 3-D dasymetric mapping technique is presented through a New York City-based case study, and contrasted with a case study of São Paulo, Brazil, to demonstrate the possibilities of using this technique in different settings of data availability.

Keywords: dasymetric, cadastral, population mapping, New York City, São Paulo, 3-D mapping, environmental exposures, GIS

Introduction

This study examines the importance of determining an accurate depiction of population distribution for urban areas in order to develop an improved “denominator,” allowing for more correct rates in geographic information system (GIS) analyses involving public health and urban environmental planning. Rather than using data aggregated by arbitrary administrative boundaries such as census tracts, we use dasymetric mapping, an areal interpolation method using ancillary information to delineate areas of homogeneous values. The dasymetric method has been expanded in this study to incorporate three dimensions, in order to better capture the actual population affected by three dimensional impacts, such as air pollution. In a case study of Manhattan, New York City, a comparison is made amongst several residential population exposure estimation methods, such as traditional GIS spatial selection approaches (e.g., intersection, centroid containment), 2-dimensional dasymetric disaggregation (with lot-level cadastral data as the ancillary dataset), and 3-dimensional (3-D) dasymetric disaggregation. The results of the NYC case study are contrasted with another worked example, using São Paulo, Brazil to illustrate the differences in the 3-D disaggregation technique between locations with varying degrees of data availability, and to demonstrate the possibility of 3-D dasymetric mapping improving the analysis in both types of areas. The study shows the impact that a more accurate estimation of population distribution has on current environmental and health research projects, and its potential for other GIS applications.

Environmental health and environmental justice studies require reliable estimates of exposed populations. Exposures are modeled in many ways (proximity buffers, network buffers, plumes, contaminant fate and transport modeling such as air dispersion modeling, etc.). As computer power increases and data availability is

improved, 3-D modeling has become more practical. This study aims to compare the estimated impacted population to a 3-D exposure using four methods:

- Spatial intersect (block group level)
- Centroid containment (block group level)
- 2-D dasymetric disaggregation (property lots – cadastral data - by residential area)
- 3-D dasymetric desegregation (buildings by residential volume)

Dasymetric Mapping

Dasymetric mapping refers to a process of dividing spatial data into finer units of analysis, using ancillary datasets to better locate populations or other phenomena (Eicher and Brewer, 2001; Holt et al, 2004; Mennis and Hultgren, 2006). This process seeks to create areas more closely resembling the actual “facts on the ground,” rather than geographic units based on arbitrary administrative boundaries, such as postal codes or census enumeration units. Administrative boundaries are often created arbitrarily or for other purposes and generally do not necessarily relate to the underlying data pertaining to exposures. Population totals within a given geographic unit are assumed to be distributed evenly, when in fact they are usually much more heterogeneous, especially in densely developed urban areas (Maantay and Maroko, 2008).

Two methods have been widely used to estimate populations in defined geographic districts: areal interpolation (Langford et al, 1991) and filtered areal weighting, a basic type of 2-D dasymetric mapping (Flowerdew and Green 1992; Goodchild and Lam 1980). In this study, we are using an innovative approach, 3-dimensional dasymetric mapping, building upon a previous method we designed, Cadastral-based Expert Dasymetric System (CEDS), which uses census data in conjunction with cadastral (property lot) data in order to create a more precise picture of where people actually live (Maantay et al, 2007; Maantay and Maroko, 2008; Maantay et al, 2008). The CEDS method constitutes a refinement of 2-D dasymetric disaggregation, and estimates populations better than areal interpolation and filtered areal weighting, calculating more accurate rates, and, thus, describes with more fidelity the spatial distribution and patterns of disease, risk from hazard, environmental exposures, and other issues.

Recent 3-Dimensional Dasymetric Research

In recent research by others, population estimation methods have incorporated some 3-D data elements. In a study conducted by Wang, et al (2016), population distribution was estimated by 3-D reconstruction of urban residential buildings through building detection and height retrieval with high resolution (HR) images. However, although this method utilized 3-D information to perform the estimation, it still only yielded population distribution on the ground (2-D), and not disaggregated in three dimensions. Other researchers (Biljecki et al, 2016; Lwin and Murayama, 2010, 2011; Pavia and Cantarino, 2016; Petrov et al, 2005; Xie, 2013) have undertaken similar studies, employing 3-D building information from such sources as LiDAR (Light Detection and Ranging) imagery and LiDAR-derived Digital Volume Model (DVM), building footprint data, parcel-level data, DOQQs (digital orthophoto quarter quads), and DEIMOS-2 Very-High Resolution Multispectral Imagery. These studies resulted in improved 2-D representations of population distribution, usually in the form of density surfaces, but did not actually place the population distribution in three dimensions. Sridharan and Qiu (2013) used LiDAR-derived building volumes as an ancillary variable to spatially disaggregate population, both horizontally and vertically, thus achieving the closest results to actual 3-D population distribution so far. Our new method is an advancement on this, being based on more specific data than building volume alone, and taking into account 3-D exposure assessment, in addition to population estimates.

3-Dimensional Dasymetric Method

In this current study, we are expanding upon the cadastral-based 2-D dasymetric mapping method in order to incorporate 3-dimensional modeling of population distribution, using an extrusion technique, building heights information, and other ancillary data to obtain a more accurate estimation for population distribution. This then can be useful in assessing exposure to air pollution, noise impacts, urban heat island (e.g., adverse health effects from extreme heat events), and so forth. Three-dimensional dasymetric mapping improves upon the 2-D dasymetric method by mapping population distribution in 3-dimensions, thus allowing exposures that stem from 3-dimensional impacts to be taken into consideration.

In the New York City case study, the amount of available residential area is estimated in each building by distributing the information from the tax-lot to each building footprint. Census population estimates are then disaggregated to the buildings using the ratio of residential area per building divided by total available residential area in the census unit. These estimates are then converted to 3-dimensions based on the buildings height. Exposed population is estimated by first intersecting the 3-D exposure buffer with the 3-D population, and then employing volumetric weighting to calculate the number of people exposed. For instance, if $\frac{1}{4}$ of the buildings volume is within the impact zone of exposure, then $\frac{1}{4}$ of the population estimated to reside in that particular building would be flagged as exposed. In this New York City study, we selected a six block area and compared the 3-D building-level dasymetric method with 2-D lot-level dasymetric, spatial intersect, and centroid containment methods, and the results of the comparison illustrate the benefits of 3-D dasymetric disaggregation. *[See Figures 1 – 7 and Table 1.]*

We executed a similar 3-D dasymetric disaggregation in a six-block area in the Jardim Paulista district in São Paulo, Brazil. This area has a similar total population to the 6-block area studied in New York City. However, due to lack of access to similar cadastral data, a different estimation method is used, where building volume is used as a proxy for residential area, assuming that residential area is proportional to the building volume. This was accomplished by first calculating the building volume based on footprint area and height, and then summing all the building volume per census tract. The census population was then multiplied by the ratio of building volume and the sum of all building volumes in the tract. We estimated exposed populations similarly to the methods used in the NYC example, however in this case did not include a 2-D dasymetric lot-level comparison due to the lack of cadastral data. *[See Figures 9 and 10.]*

In both the New York City and São Paulo examples, the techniques were comparable, the proportions of the populations estimated to be exposed were similar, with the intersect method being the most inclusive, then 3-D, then the centroid containment method, which tend to reflect the results typically obtained in other studies. Therefore, the intersect method gave the highest number of potentially impacted/exposed population, followed by the 2-D (in NYC only, since this we were not able to use this method in São Palo due to data constraints), then the 3-D method. The centroid containment method (showing zero exposed populations in both cases) had the lowest estimate of exposed population. *[See Table 2.]*

Conclusions

We have established the usefulness of the 3-D dasymetric method for any analyses employing population-based rates, as is commonly the case with public health and epidemiological research and hazard and risk assessment, but the 3-D dasymetric method is not limited to improving the development of rates alone. These methods will be useful in many disparate fields and serve many purposes. For instance, one can improve emergency management operations and implementation by providing more precise information about actual positions of vulnerable or susceptible populations, thereby increasing the quality of functions such as

evacuation route planning, optimal site selection for emergency shelter locations, and critical rescue and recovery prioritization for first responders.

As the morphology of cities becomes increasingly complex, the need continues to grow for immediate and well-informed decision-making, regarding both catastrophic events and chronic conditions. We anticipate that advances in dasymetric mapping, such as the 3-D dasymetric and CEDS methods, will help us to “perfect the denominator” and better our understanding of the human-urban project.

References

- Biljecki, F., Ohori, K.A., Ledoux, H., Peters, R., and Stoter, J., (2016). Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLoS One* 11(6) doi: 10.1371/journal.pone.0156808.
- Eicher, C.L., Brewer, C.A., (2001). Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science* 28(2):125–138. doi: 10.1559/152304001782173727.
- Flowerdew, R. and Green, M., (1992). Developments in areal interpolation methods and GIS. *Annals of Regional Science* 26: 67-78.
- Goodchild, M. and Lam, N.S-N., (1980). Areal interpolation: A variant of the traditional spatial problem. *Geo-Processing* 1: 297-312.
- Holt, J.B., Lo, C.P., Hodler, T.W., (2004). Dasymetric Estimation of Population Density and Areal Interpolation of Census Data. *Cartography and Geographic Information Science* 31(2):103–121. doi: 10.1559/1523040041649407.
- Langford, M., Maguire, D.J., and Unwin, D., (1991). The areal interpolation problem: Estimating population using remote sensing in a GIS framework. In: Mather, I. and Blakemore, M. (eds.), *Handling geographic information: Methodology and potential applications*. London, U.K.: Longman.
- Lwin, K. K. and Murayama, Y., (2010). Development of GIS Tool for Dasymetric Mapping. *International Journal of Geoinformatics* 6 (1):11-18.
- Lwin, K. K. and Murayama, Y., (2011). Estimation of Building Population from LiDAR Derived Digital Volume Model. In: Murayama, Y., Thapa, R.B., (eds.), *Spatial Analysis and Modeling in Geographical Transformation Process*, Springer pp. 87-9.
- Maantay, J.A., Maroko, A.R., and Herrmann, C., (2007). Mapping Population Distribution in the Urban Environment: The Cadastral-based Expert Dasymetric System (CEDS). *Cartography and Geographic Information Science* 34(2):77–102. doi: 10.1559/152304007781002190.
- Maantay, J.A., Maroko, A.R., (2008). Mapping urban risk: Flood hazards, race, & environmental justice in New York. *Applied Geography* 29 (1):111-124.
- Maantay, J.A., Maroko, A.R., Porter-Morgan, H., (2008). A New Method for Population Mapping and Understanding the Spatial Dynamics of Disease in Urban Areas. *Urban Geography* 29(7):724-738.
- Mennis, J. and Hultgren, T., (2006). Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science* 33(3):179–194. doi: 10.1559/152304006779077309.
- Pavía, J. M. and Cantarino, I. (2016). Can Dasymetric Mapping Significantly Improve Population Data Reallocation in a Dense Urban Area? *Geographical Analysis* doi:10.1111/gean.12112.
- Petrov, A., Bozheva, A., and Sugumaran, R., (2005). The effect of spatial resolution of remotely sensed data in dasymetric mapping of residential areas. *GIScience and Remote Sensing* 42(2): 113-130.
- Sridharan, H. and Qiu, F., (2013). A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes. *Geographical Analysis*, 45: 238–258. doi:10.1111/gean.12010.
- Wang, S., Tian, Y., Zhou, Y., Liu, W., and Lin, C. (2016). Fine-Scale Population Estimation by 3D Reconstruction of Urban Residential Buildings. *Sensors* 16 (10). DOI: 10.3390/s16101755.
- Xie, Z., (2013). A Framework for Interpolating the Population Surface at the Residential-Housing-Unit Level. *GIScience and Remote Sensing* 43(3): 233-251.



Figure 1

Two dimensional view of a six block study area in Manhattan, New York City. Population density shown by census block group (n=7) and building footprints shown as either having property tax-lot derived residential area data (orange) or not (grey). Fifty meter impact zone is shown along Amsterdam Avenue between 104th and 110th Streets.

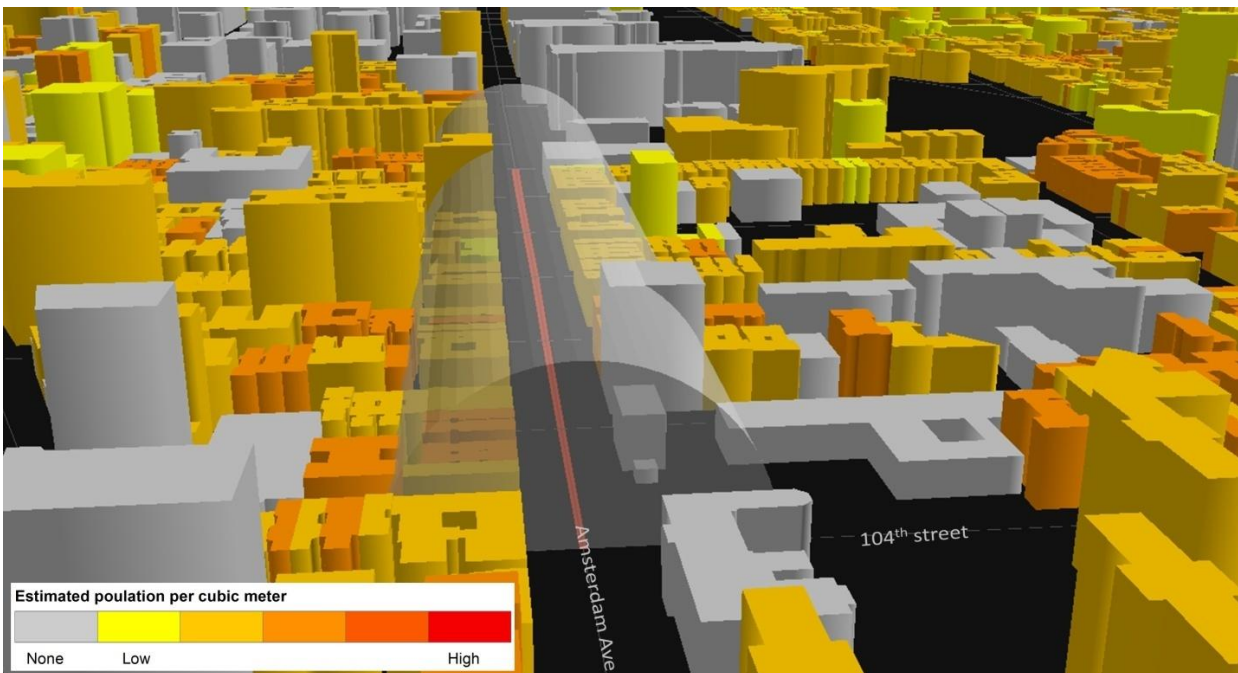


Figure 2 - 3-D depiction of population densities (by volume) and impact buffer for New York City study area
 Three-dimensional oblique view of the Amsterdam Avenue study area. Impact zone is shown as a 3-dimensional 50m buffer, and volumetric population density by building (people per cubic meter) are shown. Estimate of potentially exposed population within impact buffer = 1,971.

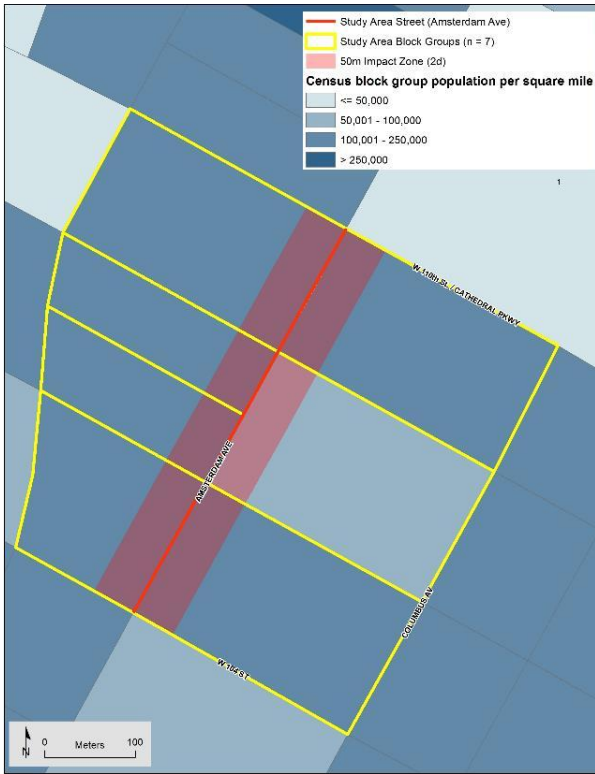
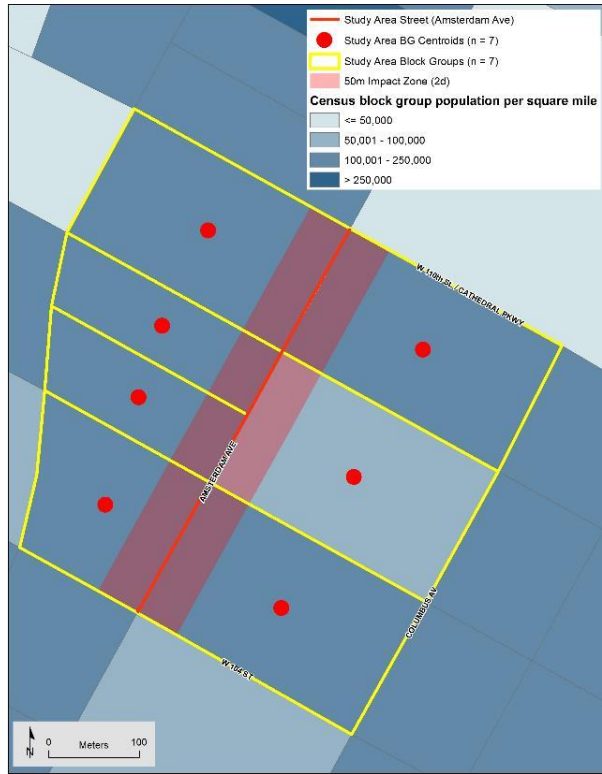


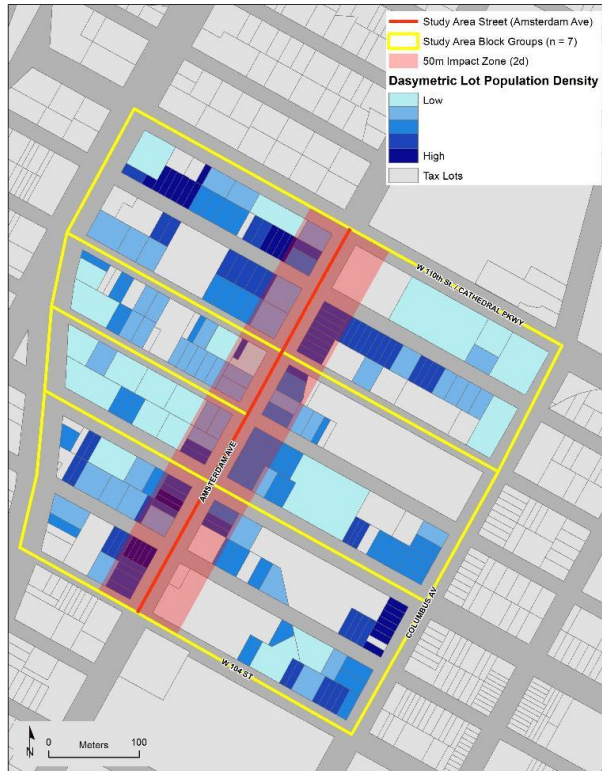
Figure 3 - Typical Spatial Selection Methods
Block group Intersect in NYC study area
 Impacted population estimate = 12,487
 Total population = 12,487



Block Group Centroid Containment in NYC study area
 Impacted population estimate = 0
 Total population = 12,487



Figure 4 - 2-D Dasymetric Method –
Tax Lots in Study Area, NYC study area



Population Estimates: Impacted Population = 3,209
Total Population = 12,487

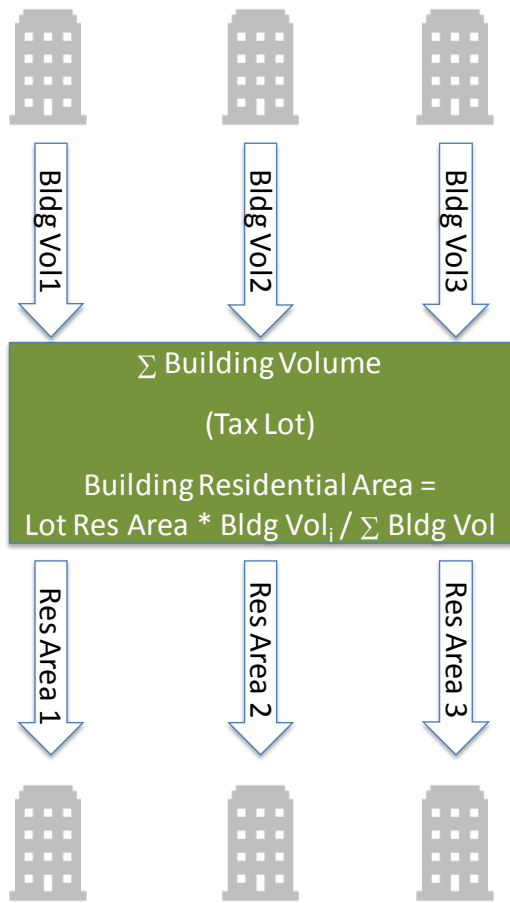


Figure 5 – 3-D Dasymetric Method - Residential Area, NYC study

Residential area is estimated in each building by distributing the information from the tax-lot to each building footprint based on building volume.

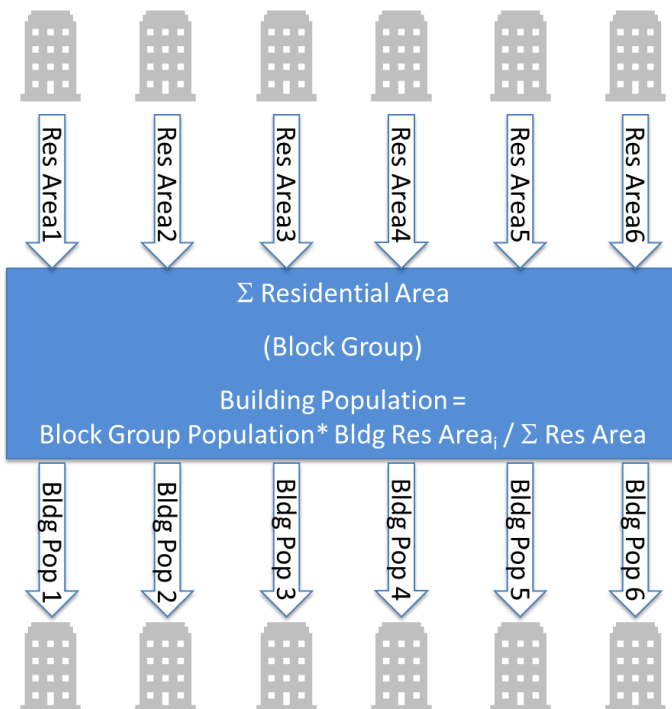


Figure 6 – 3-D Dasymetric Method – Population Estimation, NYC study: Census population data are disaggregated to the buildings using the ratio of residential area per building divided by total available residential area in the census unit.

Method	Total Population	Exposed Population	Percent Exposed
BG Intersect	12,487	12,487	100.0
BG Centroid	12,487	0	0.0
Lot-level dasymetric Intersect	12,487	3,209	25.7
3d Building Volumetric Weighting	12,487	1,971	15.8

Table 1 – New York City case study area comparison of the four methods of population estimation.
 Block Group (BG) intersect; BG centroid; 2- dasymetric (using cadastral lot level data); and 3-D dasymetric.

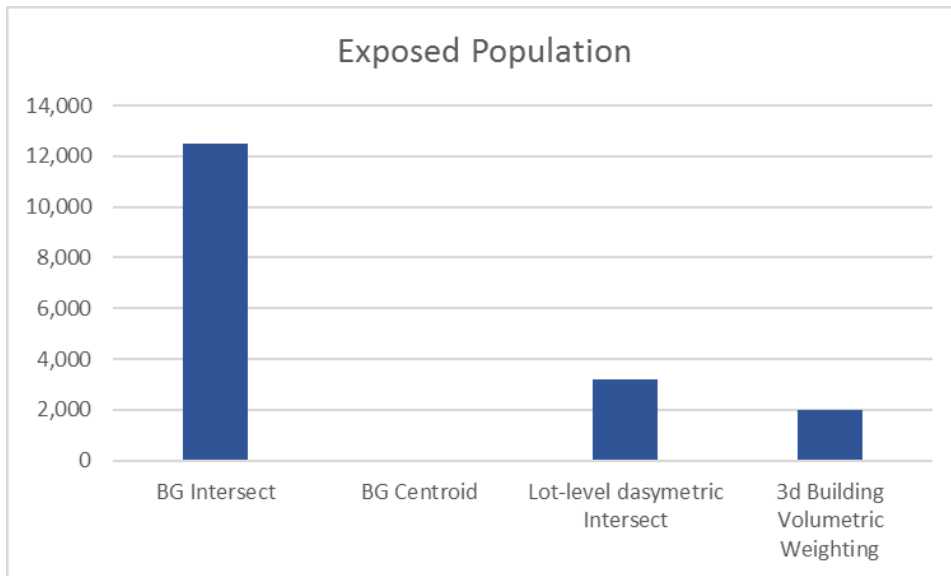


Figure 7 – Comparison of Results of 4 Methods of Population Estimation in NYC Study Area
 Block Group (BG) intersect; BG centroid; 2- dasymetric (using cadastral lot level data); and 3-D dasymetric.

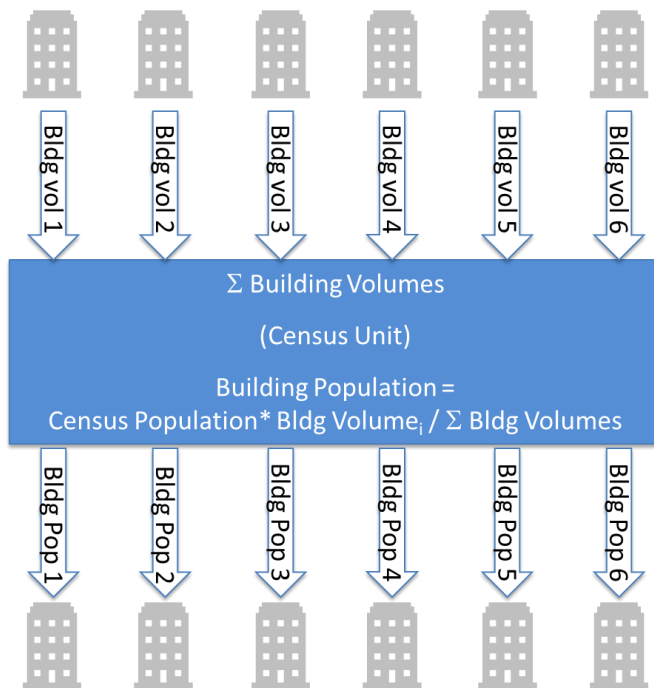


Figure 8 3-D Dasymetric Method – Population Estimation, São Paulo study: Building volume was calculated based on footprint areas and height, then summing all the building volume per census tract. The Census population was then multiplied by the ratio of building volume and the sum of all building volumes in the tract.

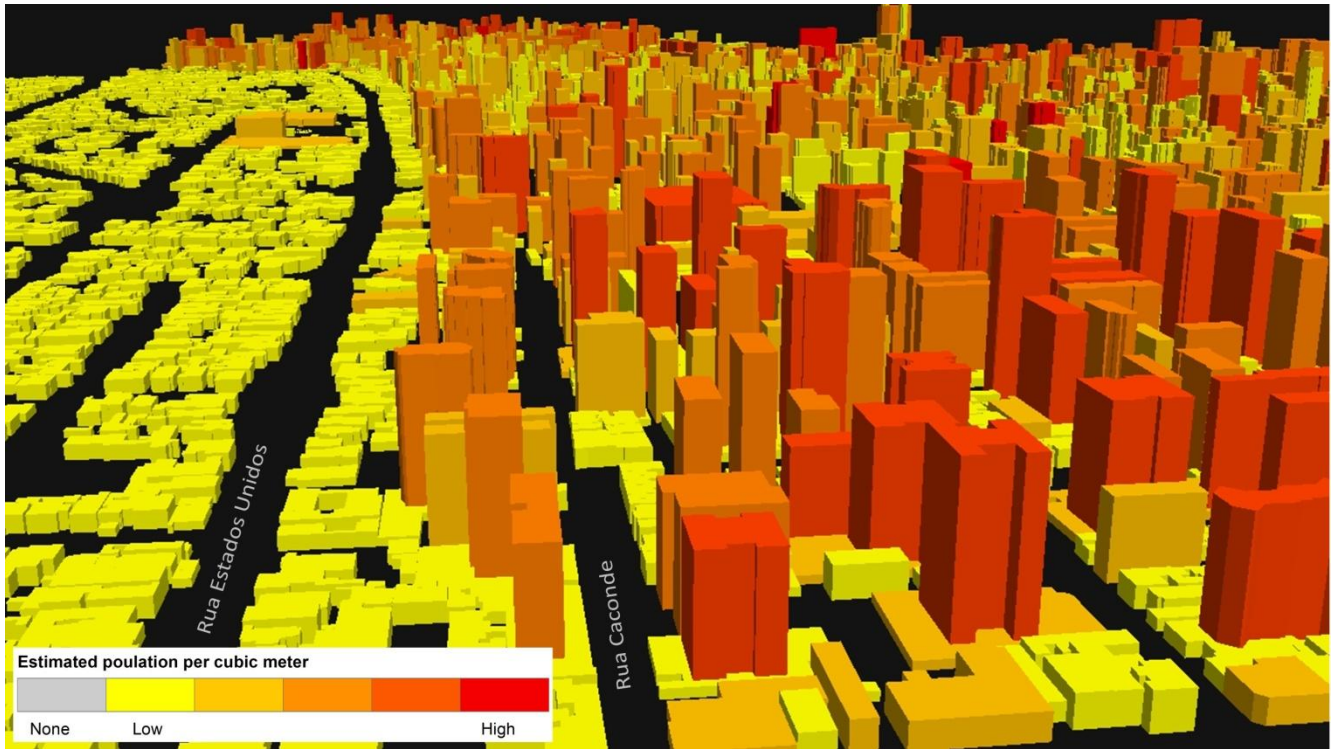


Figure 9 - The Jardim Paulista district in São Paulo, Brazil: *Estimated population per cubic meter of building volume.*

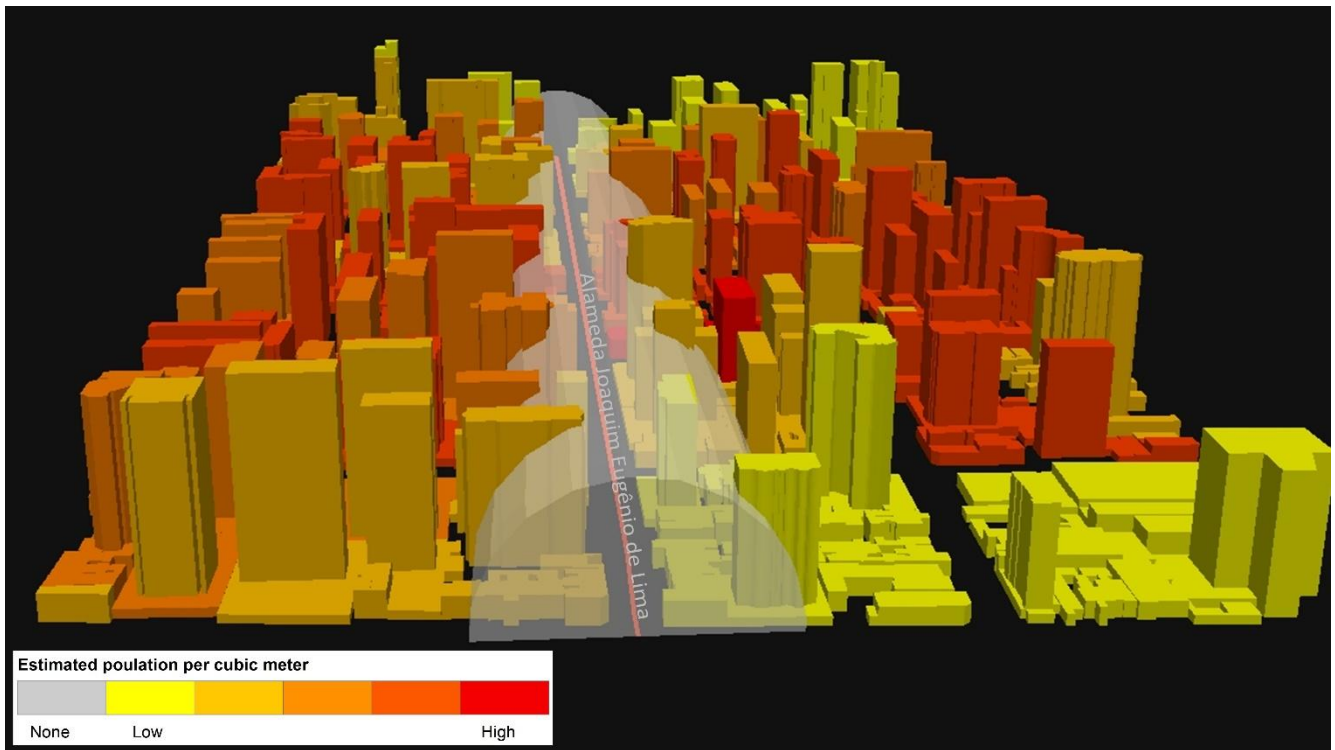


Figure 10 - 3-D depiction of population densities (by volume) and impact buffer for São Paulo study area. Three-dimensional oblique view of the Alameda Joaquim Eugênio de Lima study area. Impact zone is shown as a 3-dimensional 50m buffer, and volumetric population density by building (people per cubic meter) are shown. Estimate of potentially exposed population within impact buffer = 1,157.

Method	Population in 6-block study area		Estimated Exposed Population		Estimated Percent Exposed Population	
	NYC	SP	NYC	SP	NYC	SP
Census Intersect	12,487	11,786	12,487	9,830	100.0	83.4
Census Centroid	12,487	11,786	0	0	0.0	0.0
Lot-level dasymetric Intersect	12,487	11,786	3,209	--	25.7	--
3d Building Volumetric Weighting	12,487	11,786	1,971	1,157	15.8	9.8

Table 2 – Comparison of results for the New York City and the São Paulo case study areas

Author biosketches

Juliana Maantay is Professor of Urban Environmental Geography at City University of New York since 1998. She directs the graduate program in Geographic Information Science as well as the Urban GISc Lab, and has edited several compendia and written two widely-used textbooks and numerous other publications on the urban environment and geospatial analysis. Her main research foci are environmental justice, health disparities, and exposure assessment, specifically in urban areas. For 25 years prior to her academic career, she was an urban planner, environmental analyst, and architect. Juliana.maantay@lehman.cuny.edu

Andrew Maroko is an Associate Professor at The CUNY Graduate School of Public Health and Health Policy and Associate Director of the Urban GISc Lab at Lehman College. His research interests are in the examination of health inequities, exposures, accessibility, and environmental justice in a spatial framework.

Acknowledgements

The National Oceanic and Atmospheric Administration’s Cooperative Remote Sensing Science and Technology Center (NOAA-CREST) provided critical support for this work under NOAA grant number NA17AE162. The statements contained within this paper are not the opinions of the funding agency or the U.S. government, but reflect the authors’ opinions.

We would also like to express our thanks to our colleagues in Brazil, Reinaldo Paul Pérez Machado, Luis Bittar Venturi, and Ligia Vizeu Barrozo, of the Universidade de São Paulo, for all their help, especially in obtaining the São Paulo data we needed for this study.