Dasymetric Mapping for Environmental Health GIS

Juliana Astrud Maantay (1) (2) (3) Andrew R. Maroko (1) (2) Jun Tu (1) (2) Brian Morgan (1) Alvin Hal Strelnick (4)

South Bronx Environmental Justice Partnership

(1) Lehman College, City University of New York, Department of Environmental, Geographic, and Geological Sciences; and Geographic Information Sciences Program (2) City University of New York Graduate Center, Earth and Environmental Sciences Ph.D. Program
(3) Lehman College, City University of New York, Department of Health Sciences, Public Health Graduate Program (4) Albert Einstein College of Medicine; Montefiore Medical Center

Synopsis

This poster depicts the importance of determining an accurate depiction of total population and specific sub-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, crime, 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 mapping technique is presented through a case study of New York City, and shows the impact that a more accurate estimation of population distribution has on a current environmental justice and health disparities research project, and its potential for other GIS applications.

Background of Dasymetric Mapping

What is "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. 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 zipcodes or census tracts. Administrative boundaries are created arbitrarily or for other purposes and do not necessarily relate to the underlying data. Population totals within a given geographic unit are assumed to be distributed evenly, when in fact they are usually much more heterogeneous.

Methods and Data

Two methods are widely used to estimate populations in defined geographic districts: areal interpolation and filtered areal weighting. In this study, we designed a new method, Cadastral-based Expert Dasymetric System (CEDS), which uses census data in conjunction with cadastral (tax lot) data in order to create a more precise picture of where people actually live. This method better estimates populations to calculate more accurate rates, and, thus, describe with more fidelity the spatial distribution and patterns of disease, crime, hazard, exposure, and other issues. Figure 1 contrasts a hypothetical example comparing the three methods. Table 1 compares different aspects of these three methods.

Figure 1. Comparison of Three Disaggregation Methods

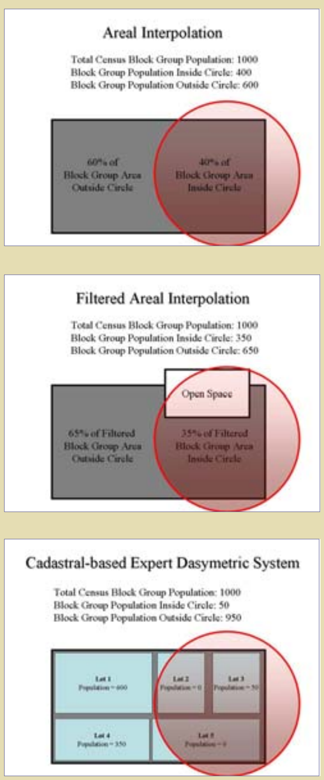
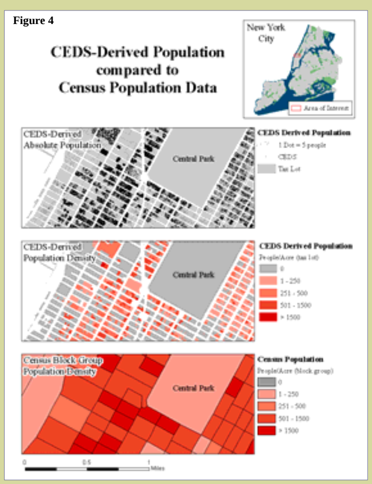


Table 1. Areal Interpolation vs. Filtered Areal Weighting vs. CEDS

	Areal Interpolation	Filtered Areal Weighting (Binary Method)	Cadastral-based Expert Dasymetric System (CEDS)
Assumption	Population is distributed homogeneously throughout each geographic unit	Population is distributed homogeneously over inhabitable areas	Population distributed heterogeneously over inhabitable areas reflected by tax lot data
Ancillary Data	No	Land use / Land Cover	Cadastral data (Tax lot properties)
Methodology	Calculate populations based on ratio of affected geographic area to baseline unit	Calculate populations based on ratio of affected area to baseline unit excluding uninhabitable areas	Use tax lot data to disaggregate population, then aggregate tax lot level values to geographic units of interest
Advantage	Simple	Relatively simple, useful when working with areas without cadastral data	High spatial resolution, high accuracy
Disadvantage	Lack of accuracy	Less accurate for highly urbanized area	Not applicable for areas without cadastral (tax lot) data



Proposed dasymetric mapping method: CEDS

General Methodology and Analysis

Objective

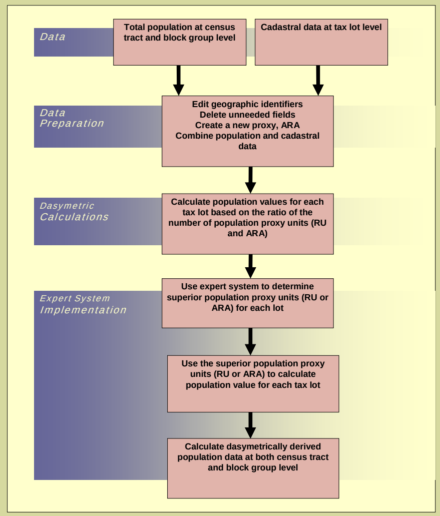
Disaggregate the total population counts from the census tract (CT) level (N = 2,217 CTs in New York City) and census block group level (N = 5,733CBGs, mean = 2.6/CT) to the tax lot level (N = 847,153 tax lots, mean = 382/CT).

Methodology

The technique uses adjusted residential area (*ARA*) and number of residential units (*RU*) as proxies for population distribution. The population in each census block group was disaggregated among the tax lots based on either *ARA* or *RU*. The proxy unit used in the disaggregation was determined by an expert system individually for each geographic unit. The results were then validated against census data and compared to commonly used disaggregation techniques to assess predictive accuracy.

The CEDS disaggregation of census populations can be compartmentalized into three fundamental steps: data preparation, dasymetric calculations, and expert system implementation, represented in the flow chart shown in Figure 2.

Figure 2 Flow chart of CEDS methodology

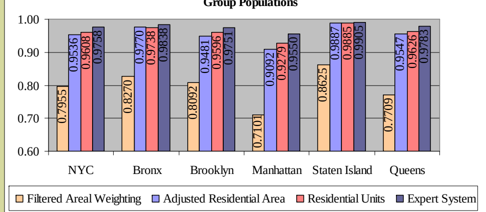


Results Evaluation

The dasymetrically derived population data produced by CEDS was compared to the population data produced by using filtered area weighting, adjusted RA, and RU alone. The estimated block group populations from the four disaggregation methods were regressed against the block group population data from the census to evaluate their relative effectiveness in New York City as a whole and separated by borough. The adjusted R^2 values for the results of CEDS ranged from 0.96 to 0.99, higher than the three other methods, indicating that CEDS produced the most accurate population data (Figure 3).

The result of the CEDS methodology is tax lot-level rather than block group-level population data, an areal unit that has approximately 150-times finer resolution. Figure 4 compares CEDS-derived population, CEDS-derived population density by tax lot, and traditional choroplethic population density by census block group.

Figure 3 Adjusted R^2 Values of Estimated Block Group Populations vs. Census Block Group Populations



Data Sources

Total population data by census tract and block group from U.S. Census 2000.
Cadastral data at tax lot level from Lotinfo, a product of LotInfo, LLC, which combines spatial data from the New York City Departments of City Planning and Finance.

Asthma and Air pollution Case Study Using CEDS

Asthma is extremely prevalent in the Bronx. The asthma hospitalization rate for children is 70 percent higher in the Bronx than in New York City as a whole, and 700 percent higher in the Bronx than for the rest of New York State (excluding New York City). The purpose of this case study is to compare filtered area weighing and CEDS methodologies by examining asthma hospitalization rates inside and outside of a 150m buffer zone along limited access highways (LAH) in the Bronx.

Data sources

Asthma hospitalization data from 1995 to 1999 provided by the Statewide Planning and Research Cooperative System; Census, roadways, TIGER landmark, and TIGER water body geographic files from the U.S. Census Bureau; Cadastral data provided by Lotinfo.

Methodology

The asthma hospitalization rate for inside the buffer was calculated by dividing the cases falling inside the buffer by the population within the buffer. The populations inside the buffer obtained by filtered area weighing method and CEDS were different, yielding different rates. The same method was used for calculating cases, populations, and rates outside of the buffer.

Filtered areal weighting used areas with no likely residential population (open spaces) as a "cookie cutter" to remove uninhabited areas from the areal weighting calculations. The estimated inside/outside buffer block group population was calculated through multiplying the original census block group population by the ratio of the buffer area (inside or outside) to filtered block group area (excluding open spaces).

The CEDS methodology was used to disaggregate the census block group population to the tax lot level. Populations inside and outside the buffer were calculated by selecting the tax lots whose centroids fall inside or outside the 150 meter LAH buffers and summing the associated populations.

Results

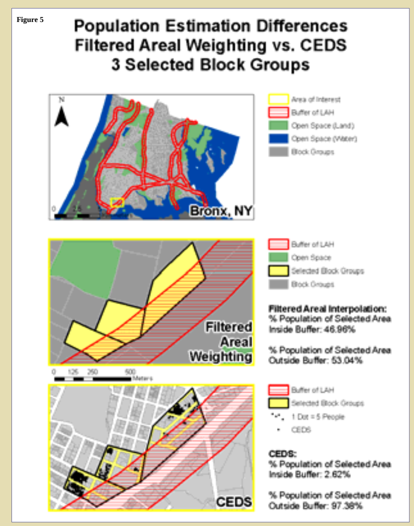
Difference in Population for 3 Selected Block Groups

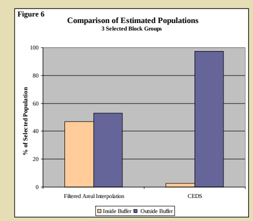
Three block groups were selected in the South Bronx to illustrate the methods and respective results more explicitly. The filtered areal weighting method estimated 1,017 people (~47%) resided within 150 meters of LAH and 1,149 people (~53%) lived outside of the buffer zone. The CEDS method estimated that only 57 people (~3%) were within the LAH buffer, and 2,109 (~97%) lived beyond the 150 meter threshold (see Figures 5 and 6).

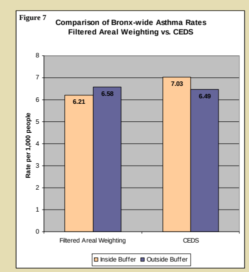
Difference in Asthma Rates Borough-wide

For the entire Bronx, the filtered areal weighting and CEDS methodologies produced differing results as well. When using the filtered areal weighting methodology, the asthma hospitalization rate inside the buffer was 6.21 per 1,000 per year, whereas outside the buffer the rate was 6.58 per 1,000 per year, greater than the inside rate. If exposure to certain outdoor air pollution increases asthma hospitalizations, and close proximity to limited access highways increases exposure to this pollution, one would expect that living in close proximity to an LAH would show an elevated rate of asthma hospitalizations when compared with those living beyond this threshold. The filtered areal weighting methodology, however, claims that the opposite is true. When using the CEDS method, rates inside the buffer were found to be 7.03 per 1,000 per year and outside the buffer were 6.49 per 1,000 per year (Figure 7).

Examining the standardized incidence ratios (SR), filtered areal weighting showed that residing within the LAH buffer was *protective*, with a 5% lower asthma hospitalization rate. However, when CEDS is used to calculate the LAH buffer denominators, the SR shows a 7.5% higher asthma hospitalization rate. These CEDS-derived results, based on the more precise population location information, support the hypothesis that exposure to pollutants released from the limited access highways in the Bronx elevate asthma hospitalization rates.







Conclusions

Based on the application of the CEDS methodology to New York City population data and the case study example of asthma hospitalization rates in the Bronx, we have demonstrated that the Cadastral-based Expert Dasymetric System can improve research and analyses that utilize population distribution information, and create more realistic models of real-world conditions.

ACKNOWLEDGEMENTS

This research was partially supported by grant number 2 R25 ES01185-05 from the National Institute of Environmental Health Sciences of the National Institutes of Health.

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

Thanks are also due to the member organizations of the South Bronx Environmental Justice Partnership, who understood the relevance of this project to environmental health justice and gave their unstinting encouragement and assistance in the effort.