

RESEARCH NOTE—A NEW METHOD FOR MAPPING POPULATION AND UNDERSTANDING THE SPATIAL DYNAMICS OF DISEASE IN URBAN AREAS: ASTHMA IN THE BRONX, NEW YORK¹

*Juliana A. Maantay,*² *Andrew R. Maroko,* and *Holly Porter-Morgan*
Environmental, Geographic, and Geological Sciences Department
Lehman College, City University of New York

Abstract: Determining an accurate depiction of population distribution for urban areas in order to develop an improved “denominator” is important for the calculation of higher-precision rates in GIS analyses, particularly when exploring the spatial dynamics of disease. Rather than using data aggregated by arbitrary administrative boundaries such as census tracts, we developed the Cadastral-Based Expert Dasymetric System (CEDS), an interpolation method using ancillary information to delineate areas of homogeneous values. This method uses cadastral data, land-use filters, modeling by expert system routines, and validation against various census enumeration units and other data. The CEDS method is presented through a case study of asthma hospitalizations in the borough of the Bronx in New York City, in relation to proximity buffers constructed around major sources of air pollution. The analysis using CEDS shows that asthma hospitalization risk due to proximity to pollution sources is greater than previously calculated using traditional disaggregation methods. [Key words: Dasymetric, asthma, population distribution, pollution exposure, cadastral.]

THE GEOGRAPHY OF ASTHMA AND AIR POLLUTION IN THE BRONX

The objective of this research note is to demonstrate the utility of developing a more accurate way to map and estimate population distribution using a high-resolution geographic unit of analysis and an expert system, in order to understand the spatial dynamics of disease. A case study of asthma and air pollution in the borough of the Bronx in New York, illustrates this new method, the Cadastral-Based Expert Dasymetric System (CEDS).

¹This 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 paper are not the opinions of the funding agencies or the U.S. government, but reflect the authors’ opinions. This research was also supported by a Faculty Research Award from the Professional Staff Congress of the City University of New York (PSC-CUNY), Awards # 69372-0038, “Perfecting the Denominator: Developing a Cadastral-Based Expert Dasymetric System in New York City.” 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 to the effort.

²Correspondence concerning this article should be addressed to Juliana Maantay, Environmental, Geographic, and Geological Sciences Department, Lehman College, CUNY, 250 Bedford Park Boulevard West, Bronx, New York 10468; telephone: 718-960-8574; fax: 718-960-8584; e-mail: juliana.maantay@lehman.cuny.edu

The past two decades have seen a significant increase in the prevalence of asthma, and research has demonstrated that parts of New York City have among the highest rates in the United States. According to the *Asthma Facts Report*, asthma hospitalization in the Bronx for children is 70% higher than in New York City as a whole, and 700% higher than New York State, excluding the city of New York (New York City Department of Health [NYC DOH], 1999). “Overall, in recent years, the Bronx is the New York City borough with the highest rates of both asthma hospitalizations and deaths” (New York City Department of Health [NYC DOH], 2003, p. 2). Poor air quality may be one of the major factors contributing to these elevated rates. Recent studies have shown the association between asthma events and certain criteria and volatile air pollutants (e.g., ozone, particulate matter; Brunekreef and Holgate, 2002; Hoek et al., 2002; Mortimer et al., 2002; Delfino et al., 2003; Chaix et al., 2006). Additionally, researchers have reported a spatial correspondence between asthma-related illnesses and major sources of air pollution (Brauer et al., 2002; Gauderman et al., 2005, 2007; Molitor et al., 2006).

Proximity to stationary air pollution point sources, such as Toxic Release Inventory (TRI) facilities, as well as mobile sources, such as highways, has been linked to increased respiratory ailments (Peters et al., 1999; Jerrett and Finkelstein, 2005).

This relationship is especially important in the Bronx, because this borough contains multiple major sources of air pollution located within or near densely populated areas (Maantay, 2001). Understanding the magnitude and spatial distribution of this pollution and the exposure impact on susceptible populations has been problematic, because neither the pollution nor the population is distributed evenly, particularly in a hyper-heterogeneous urban environment such as the Bronx.

Previous research (Maantay, 2005/2007) examined the spatial association between major sources of air pollution and residences of individuals hospitalized for asthma during 1995–1999 in the Bronx. Four major categories of air pollution sources were included in that study: TRI facilities, other major stationary point sources (SPS) from the National Emissions Inventory, limited access highways (LAH), and major truck routes (MTR). Standard distances (i.e., proximity buffers) from each pollutant source were developed as a proxy for exposure impact (Fig. 1).

The buffers constructed for that study were based on distances established as standards by environmental agencies or used most often by other researchers as the zones of greatest potential impact from sources. One-half mile (approximately 800 m) radius buffers were constructed around TRI facilities (Chakraborty and Armstrong, 1997; Neumann et al., 1998); one quarter mile (approximately 400 m) radius buffers were established around other major stationary point sources of U.S. EPA-defined criteria pollutants that cause smog, acid rain, and other health hazards, including sulfur oxides (SO_x), nitrogen oxides (NO_x), and particulate matter (PM; New York City Mayor’s Office of Environmental Coordination, 2001); and a 150 m buffer from roadway centerlines around both limited access highways and major truck routes (Hitchins et al., 2000; Zhu et al., 2002). “The distance within which concentrations of primary vehicle traffic pollutants are raised above ambient background levels” (Venn et al., 2001, p. 2177) is 150 m from a main road. The majority of similar studies found significant associations between traffic-related emissions and respiratory symptoms within the 100–200 m range (Nitta et al., 1993; Edwards et al., 1994; Livingstone et al., 1996; Wilkinson et al., 1999).

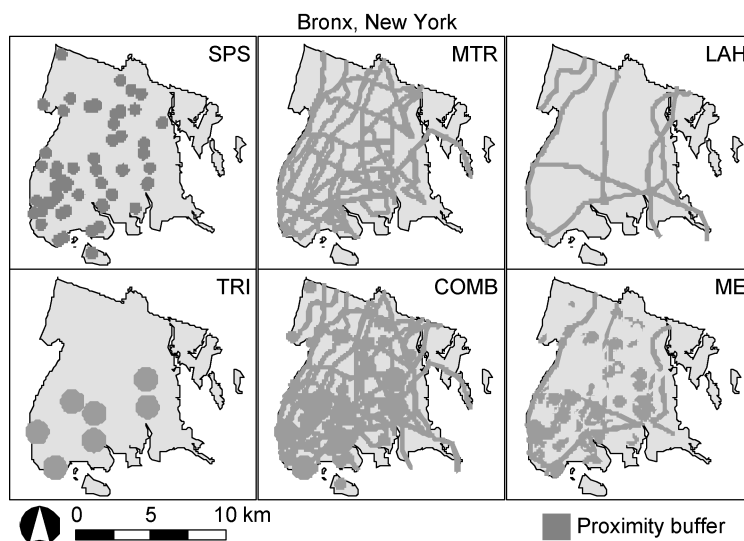


Fig. 1. Pollution proximity buffers used in the asthma hospitalization analyses. SPS = stationary point sources; TRI = Toxic Release Inventory facilities; MTR = major truck routes; LAH = limited-access highways; COMB = combined buffers; ME = multiple-exposure buffers. *Sources:* SPS: National Emission Inventory (EPA, 2006); MTR: New York City Department of Transportation (2002); U.S. Bureau of the Census TIGER files (2001b); LAH: U.S. Bureau of the Census TIGER files (2001b); TRI: U.S. EPA, Toxic Release Inventory (2002).

Because record-level asthma hospitalization data were used in this and the previous phases of our study (as opposed to data aggregated by census tract, zip code, or some other geographic unit), we were able to geocode actual patient street addresses, and thereby calculate more precisely how many cases were located within the pollution impact buffer zones, versus how many occurred outside them. The limitations of the asthma hospitalization data include the fact that these data consist only of hospital admissions, not emergency room or physician visits; nor should they be confused with asthma prevalence or incidence. Currently, hospitalization data are the only data consistently and reliably available at the patient-record level. Furthermore, there are general issues with geocoding accuracy and inclusiveness of the data that are discussed extensively in an earlier study by the senior author Maantay (2005/2007). These inaccuracies are mainly artifacts of the geocoding routines in the software plus ambiguities inherent in the spatial data to which the geocoded points are matched. Incomplete or incorrect addresses in the table to be geocoded also lead to suboptimal results. Some of the nearly 50,000 asthma hospitalization patient records for the five-year period had missing or incomplete addresses and thus could not be reliably geocoded, but approximately 85% of the street addresses were successfully geocoded, which is considered to be a very acceptable match rate.

Two common problems in spatial analysis using aggregated data are the Modifiable Area Unit Problem (MAUP) and ecological fallacy. MAUP occurs when changing the boundaries of the spatial units results in different findings (Openshaw, 1984). Depending upon where the boundaries are drawn when aggregating data, the geographic pattern

displayed (by the distribution of health events, noxious facilities, populations, etc.) can change substantially. It is generally believed that using the smallest feasible spatial unit possible will minimize the MAUP effect (Kreiger et al., 2002). Ecological fallacy is a related concept, which entails assigning characteristics of the group or area to the individual (King, 1997). This is unwarranted, because it assumes homogeneity, either spatially or by aspatial variables (e.g., income, race). This typically happens when aggregated data are misused to impute attributes of the individual. The method we outline below is designed to minimize the problems of ecological fallacy and MAUP by not having to assume uniform distribution of the population within these areas, and using the smallest feasible spatial unit with which to aggregate data.

PROXIMITY ANALYSIS AND AREAL INTERPOLATION

Rates of asthma hospitalization inside and outside the buffers were generated using census population data as the denominator. Because buffer boundaries are not typically coincident with census unit boundaries, we used an interpolation method (areal weighting) to estimate total population within partial census units. Based on these areally weighted population numbers, we were then able to calculate rates of asthma hospitalization within the buffers versus those outside the buffers.

Areal interpolation is a common method for calculating disaggregated population values. It is defined as “the transfer of data from one set (source units) to a second set (target units) of overlapping, nonhierarchical, areal units” (Langford et al., 1991, p. 56). A simple type of areal interpolation is to weight the variable’s values by a ratio derived from the relative areal measurements of the two types of zones (source and target; Goodchild and Lam, 1980). Areal weighting is based on the assumption that population (or another variable) is distributed homogeneously throughout the “source” zone (the original unit of data aggregation). The quantity of population estimated to be in the intersecting zone (or “target” zone) is assumed to be proportional to the quantity of area in the source zone versus the target zone. The ratio of area of source zone to target zone is then applied to population in the source zone to yield the population total in the target zone. In our study, for instance, if 25% of the census unit’s area was within a pollution impact buffer, then it was assumed that 25% of that unit’s population was also within that buffer.

The main limitation of the areal weighting method is that it assumes that population is homogeneously distributed throughout the census unit, which is rarely the case (Goodchild et al., 1993). This assumption is especially faulty in a hyperheterogeneous urban environment, where even within very small census units, population is not distributed uniformly, due to extremely diverse mixes of land uses. Therefore, rates developed from these denominators may be over- or underestimated.

Applying odds ratios to the asthma hospitalization rates show that it is 30% more likely for people living within the combined buffers to be hospitalized for asthma than people residing outside them. Within some of the individual buffers, such as TRI and major stationary point sources, it is 60% and 66% more likely, respectively, to be hospitalized for asthma than if living outside the buffers ($p < .01$; Maantay, 2005/2007). In hierarchical regression analysis, even after controlling for potential confounding factors, such as race/ethnicity and poverty status, the correlation between asthma hospitalization and proximity to air pollution sources remains significant. For example, in examining the

multiple exposure buffers, although race/ethnicity and poverty status account for most of the variance in the model, proximity to multiple sources of pollution remains significant ($R^2 = .429$; $p < .001$). Proximity to any major pollution source (residence within the combined buffers) yields similar results ($R^2 = .452$; $p < .05$; Fletcher, 2006).

Although the analysis conducted with areal weighting found that people within the buffers were much more likely to be hospitalized for asthma than those residing outside the buffers, the risks vary depending on the type of buffer. Living within TRI and major stationary point source buffers poses a higher risk than living within the limited access highway and major truck route buffers, according to the proximity and odds ratio analyses. People within the highway and truck route buffers generally do not appear to have an increased risk of asthma hospitalization. This contradicted the findings of other studies (Peters et al., 1999; Jerrett and Finkelstein, 2005) and did not conform to our expectation based on long-term anecdotal evidence, so we wondered what could have caused these counterintuitive results. We speculated that these neutral findings for the truck routes and highways may be an artifact of how the population numbers within the buffers were calculated. The areal weighting algorithm used to estimate population within the buffered areas assumed population was spread evenly throughout the census block group. However, these highway buffer areas may, in fact, be less densely populated than the remainder of the block group, for various reasons including building clearances at the time the highways were constructed. If the population next to the highways is actually smaller than that estimated by the areal weighting script, then the denominator used to calculate rates would be too high, making the asthma hospitalization rates lower than they actually are within these buffers. These limitations of areal weighting inspired us to think about a better way to estimate population in spatial units with noncoincident boundaries, such as partial census units.

DASYMETRIC MAPPING

A refinement of the areal weighting method is filtered areal weighting (FAW), where an ancillary dataset, usually of land use or land cover, is used to filter out or mask areas having few if any inhabitants. Areas containing parks, water bodies, and other normally uninhabited spaces are thus “removed” from census areal units, and the populations are redistributed to the remaining areas, based on areal proportion. This is a type of dasymetric mapping, which refers to a process of disaggregating spatial data to a finer unit of analysis, using additional (or “ancillary”) data to help refine locations of population or other phenomena (Mennis, 2003). However, even recalculating population distribution with FAW did not yield sufficiently accurate results in an area containing such hyperheterogeneous land uses as the Bronx. The need to develop a method to obtain a more accurate denominator with which to generate disease rates was critical in order to reflect the reality of asthma hospitalization’s connection to air pollution.

Although dasymetric mapping has been used in many studies to map population, existing methods generally did not result in a sufficiently high spatial resolution to be suitable for our needs. Methods such as the Three-Class and Limiting Variable methods (Eicher and Brewer, 2001), the Image Texture method (Liu et al., 2006), Heuristic Sampling method (Mennis, 2003), Kernel Density Surface from Population-Weighted Census Centroids (Bracken and Martin, 1989; Martin et al., 2000; Martin, 2006),

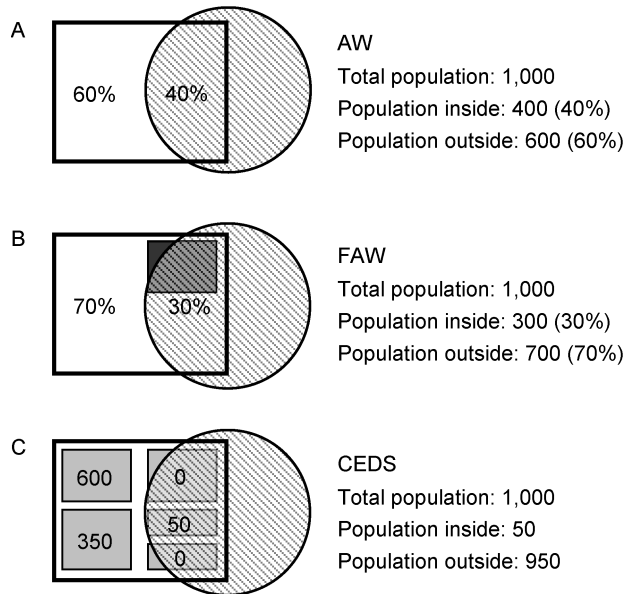


Fig. 2. Diagrammatic comparison of population disaggregation methods. (A) Areal Weighting (AW): block group intersected by an impact buffer. (B) Filtered Areal Weighting (FAW): block group intersected by an impact buffer, and showing an uninhabited area (dark rectangle). (C) CEDS: Block group showing tax lot boundaries.

and Street-Weighted Interpolation (Reibel and Bufalino, 2005; Weichselbaum et al., 2005), were evaluated as to applicability in the asthma and air pollution study; nonetheless, all were deemed to be not effective enough in providing an accurate denominator due to either comparatively lower spatial resolution of the ancillary dataset and/or attribute data inaccuracy.³ Therefore, we decided to develop our own method of dasymetric mapping, using the very high resolution New York City property tax lot data as our ancillary dataset to disaggregate census unit population by tax lot (Fig. 2).

METHODOLOGY FOR DERIVING POPULATION ESTIMATES USING THE CADASTRAL-BASED EXPERT DASYMETRIC SYSTEM (CEDS) VERSUS FILTERED AREAL WEIGHTING (FAW)

This section will describe and compare the methodologies for developing FAW and CEDS. The goal of the data disaggregation is to estimate the population residing inside versus outside the proximity buffers, in order to obtain a better denominator. The reason FAW was selected to use as a comparison to CEDS is that areal weighting interpolation is one of the most frequently used methods of disaggregating data, and filtered areal

³A full review of these methods is provided in Maantay et al. (2007).

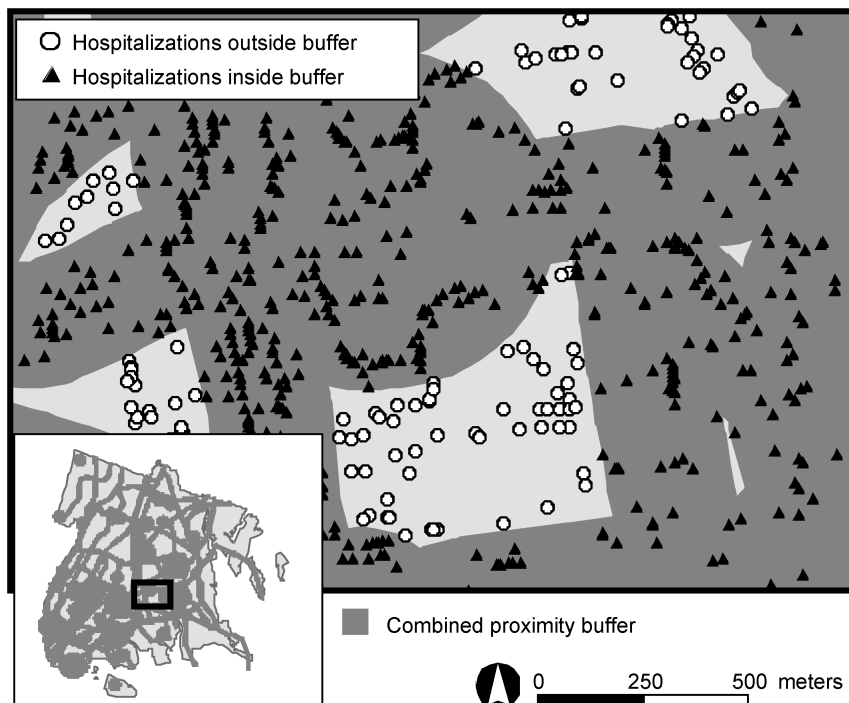


Fig. 3. Asthma hospitalizations in a portion of the Bronx, showing cases in and out of the pollution buffers. Due to patient confidentiality requirements, only hypothetical locations of asthma hospitalization cases are shown here. The actual address locations, however, were used in the spatial analyses to derive the in- and out-of-buffer rates.

weighting is a further refinement of AW, which yields good results with minimal extra expenditure of time and effort. We considered that FAW was basically the current state-of-the-art for many spatial analyses, and therefore the comparison between CEDS and the FAW method would be the most compelling.

In this analysis, the four types of impact buffers (plus the combined buffers and multiple exposure buffers) in the Bronx were created in ArcGIS using geoprocessing tools. The distances from each source, as determined by the standards cited in the literature that were detailed in the previous section, range from 150 m around the linear sources to an 800 m radius around the point sources. The residential locations for five years (1995–1999) of asthma hospitalization cases were then selected as either being within a given buffer or not, providing the numerator for asthma hospitalization rate calculations (Fig. 3). By using individual patient-level data geocoded to street addresses (obtained from the New York Statewide Planning and Research Cooperative System–SPARCS), we avoided the pitfalls inherent in using aggregated health data, and thus were able to conduct a fine-grained buffer analysis by pinpointing actual patient location with respect to the buffer boundaries.

For our comparison of methods, the denominator was calculated in two ways: (1) using filtered areal weighting (FAW); and (2) using the cadastral-based expert dasymetric system (CEDs). FAW population estimation is a relatively straightforward process. Known unpopulated locations (e.g., parks, water bodies, cemeteries, other open spaces) were used to “mask” the land area of the Bronx, similar to using a cookie cutter. The remaining land was then overlain with the proximity buffers. New area measurements were made for portions of census block groups falling inside and outside the buffers. Finally, the census population was redistributed based on the ratio of the block group–filtered areas that fall inside or outside the buffers with the unbuffered filtered census block group areas (equation 1).

$$POP_{FAW} = POP_{BG} * AREA_{BG_BUFF}/AREA_{BG} \quad (1)$$

where POP_{FAW} is the estimated portion of the block group population inside the buffer using filtered areal weighting; POP_{BG} is the census block group population; $AREA_{BG_BUFF}$ is the filtered census block group area that is also inside the buffer; and $AREA_{BG}$ is the filtered census block group area (unpopulated areas excluded). FAW population estimates for areas outside of the proximity buffers can be calculated by simply replacing the $AREA_{BG_BUFF}$ term of equation 1 with the filtered area within the block group that does not intersect the buffer.

In the present study, we compared the FAW results to those obtained with CEDs. This dasymetric system does not assume homogeneity across census units, and instead redistributes the population based on cadastral information (either the number of residential units [RU] or the residential area [RA]) at the tax lot level. The field “Residential Units” designates the number of dwelling units on the tax lot, while the field “Residential Area” consists of the square footage of the building that is used for residential purposes. Both of these serve as proxies for the population in each tax lot, neither one being an exact estimator of residential population. Because we do not know from the tax lot data how many people live within each tax lot as a whole or within each residential unit on the lot, or how many square feet of residential area there are for each person, we can only estimate the population by disaggregating from the census data, using RU or RA as proxies (equation 2).

$$RX_POP_{LOT} = POP_{BG} * U_{LOT}/U_{BG} \quad (2)$$

where RX_POP_{LOT} is the dasymetrically derived lot-level population using either RU or RA; POP_{BG} is the census block group population; U_{LOT} is the number of proxy units at the tax lot level (RU or RA); and U_{BG} is the number of proxy units at the census block group level (RU or RA).

In New York City, there are on average 150 tax lots per census block group. This fine spatial resolution allows for the detection of the large variation in population density, which is dependent upon the characteristics of the tax lot (i.e., high-density residential buildings, single family homes, commercial/industrial structures). This heterogeneity in land use and tax lot characteristics appears even at the relatively small census block group level. Block Group 1002 is an example of the heterogeneous nature of land use in the Bronx (Fig. 4). Note that much of the land use adjoining the highway is industrial or commercial, with most of the residential buildings concentrated in the northern corner of

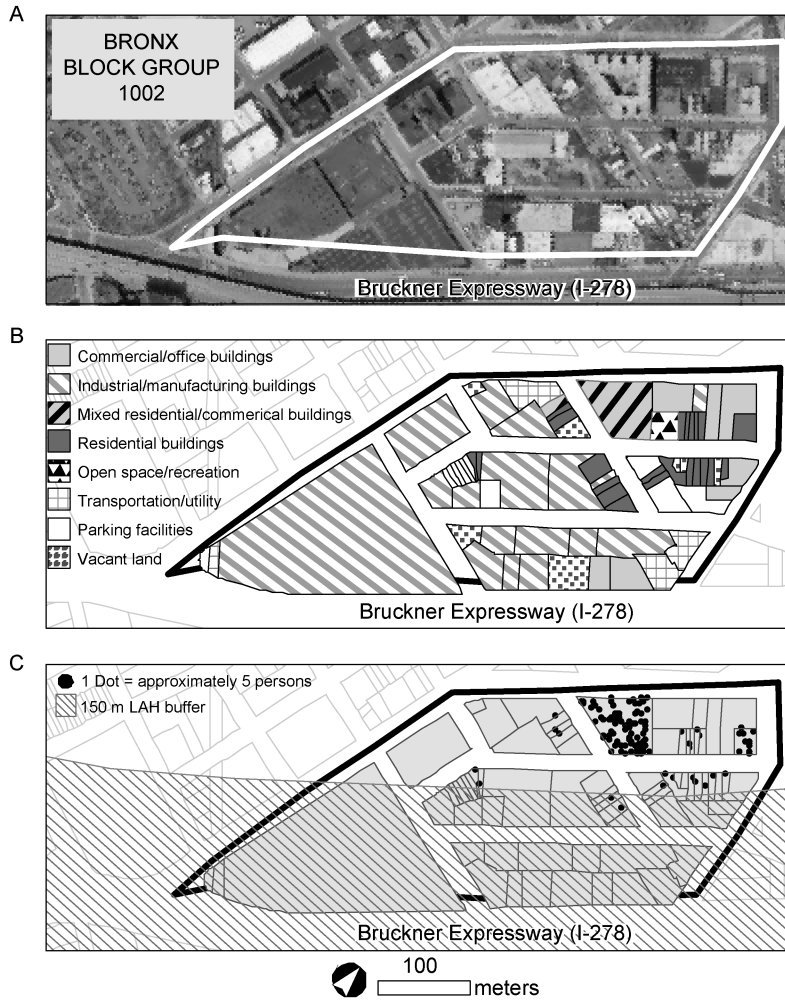


Fig. 4. Sample heterogeneous block group (A) aerial photo, (B) land use map, and (C) population estimate and LAH pollution buffer. *Sources:* U.S. Bureau of the Census (2001a), Department of Information Technology and Telecommunications (2003), LotInfo (2003).

the block group. Figure 4A is an aerial photo of Block Group 1002; Figure 4B is a land use map; Figure 4C shows the estimated population within a Limited Access Highway (LAH) pollution buffer, illustrating the comparison between FAW and CEDS. FAW-derived population is calculated by estimating the percentage of area within the buffers (in this case, 56.5%) and assigning the same ratio to the block group population. Therefore, 56.5% of the population is assumed to reside within the pollution buffer, which equals 329 people. This is an overestimation of population, when viewed against the CEDS-derived data.

An expert system is used to determine which cadastral metric (RU or RA) is most accurate for each block group. The expert system is a computerized decision-making

program that has been instructed to “decide,” based on heuristic rules and the input of “expert judgment,” which among several variables in the tax lot dataset to use for disaggregating the census data to calculate the optimally accurate tax lot-level population. Expert systems often utilize “if – then – else” rules, as in a hierarchical decision tree. The disaggregation is accomplished by estimating the census block group populations (i.e., treating block group populations as unknown) by redistributing the census tract population using CEDS. The difference in the performance of each proxy unit is assessed, and the one that most accurately predicts the census-supplied block group population is used to disaggregate all the block groups within that given census tract to the tax lot level (equation 3).

$$\begin{aligned} \text{IF } RU_POP_{DIFF} \leq RA_POP_{DIFF}, \text{ THEN } POP_{LOT} = POP_{RU_BG} \\ \text{ELSE } POP_{LOT} = POP_{RA_BG} \end{aligned} \quad (3)$$

where RU_POP_{DIFF} is the the absolute difference between the census block group population and the estimated block group population based upon number of residential units; RA_POP_{DIFF} is the absolute difference between the census block group population and the estimated block group population based upon residential area; POP_{LOT} is the final estimated tax lot population dasymetrically derived from the census block group population (not the census tract); POP_{RU_BG} is the estimated tax lot population dasymetrically derived from the census block group population (not the census tract) based on number of residential units; and POP_{RA_BG} is the estimated tax lot population dasymetrically derived from the census block group population (not the census tract) based on the residential area.

Based on the results of the expert system (whether the variable RA or RU was the most accurate predictor for a given census tract), the data are then disaggregated in a second step from the census block group data. The tax lot-level population estimates are then designated as either inside or outside the proximity buffers by using the centroid method, a simple spatial selection technique. In densely populated heterogeneous areas such as the Bronx, CEDS has been shown to perform better than FAW when estimating residential population, as validated against census data (Maantay et al., 2007). The difference between the two population estimation techniques, and the resulting potential disparity in accuracy when dealing with complex urban environments, is shown in Figure 4C.

In order to determine the possible relationship between proximity to major sources of air pollution and asthma hospitalization rates in the Bronx, exposure was classified in the following ways: (1) spatial coincidence with a proximity buffer (LAH, MTR, TRI, or SPS individually); (2) spatial coincidence with any proximity buffer (all buffers combined); (3) spatial coincidence with more than one proximity buffer (multiple exposure); and (4) no spatial coincidence with any proximity buffers (unexposed population). Asthma hospitalization rates were then calculated by dividing the number of hospitalizations by the population estimates (FAW and CEDS) for each universe of exposed populations.

RESULTS

The difference between asthma hospitalization rates when using FAW- or CEDS-derived populations as the denominators is easily seen when categorized by exposure

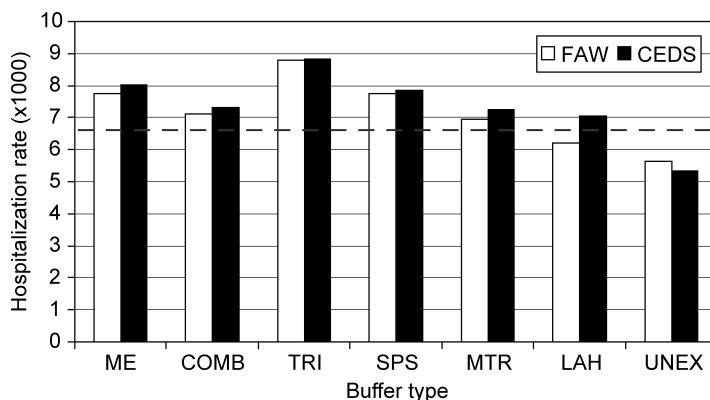


Fig. 5. Asthma hospitalizations (five-year averages, 1995–1999) in the Bronx by proximity buffer type: FAW vs. CEDS. The dotted line represents the asthma hospitalization rate for the entire Bronx. SPS = stationary point sources; TRI = Toxic Release Inventory facilities; MTR = major truck routes; LAH = limited-access highways; COMB = combined buffers; ME = multiple-exposure buffers; UNEX = unexposed (not within any buffer).

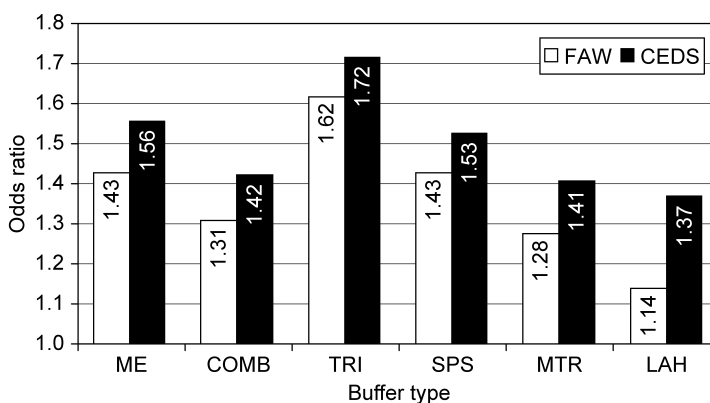


Fig. 6. Odds ratios for asthma rates (five-year averages, 1995–1999) by buffer type vs. unexposed population: FAW vs. CEDS.

type in the Bronx (Fig. 5). We used odds ratios to compare the asthma hospitalization rate (five-year average from 1995 to 1999) of the exposed populations versus that of the “unexposed” population (i.e., population which does not spatially coincide with any proximity buffer). As can be seen in Figure 6, both population estimation methods found a statistically significant increased risk (95% Confidence Interval) of asthma hospitalizations for populations residing in any single buffer, all buffers, or multiple-exposure buffers, when compared with those who do not reside in any buffer. Moreover, it should be noted that the CEDS-derived odds ratios show an elevated risk when compared to the traditional FAW technique, particularly with regard to proximity to limited access highways (OR of 1.14 and 1.37, respectively, for FAW and CEDS).

DISCUSSION AND CONCLUSION

The conventional wisdom that proximity to air pollution sources is correlated with asthma hospitalizations in the Bronx is supported by both the FAW and CEDS methods. However, CEDS provides a more realistic model for population distribution and exposure, and therefore estimates rates with increased accuracy. In this case study, FAW, assuming homogeneity across census units, tends to overestimate the population that resides close to the pollution sources when compared to CEDS, resulting in higher denominators and therefore lower rates inside the proximity buffers. Essentially, this suggests that CEDS is a better reflection of reality.

We recognize that there are certain limitations to this analysis. For example, the use of fixed buffers as a proxy for exposure is a binary model, which is a simplified reflection of reality (e.g., a case is characterized as being either “exposed” or “unexposed” based on its location vis-à-vis the buffer boundary). It should be noted that this constraint applies to any analysis using discrete boundaries rather than a continuous surface. Both FAW and CEDS are equally affected by this limitation.

Another limitation associated with fixed impact buffers is the fact that they do not necessarily represent the actual fate and transport of pollutants, as potentially affected by meteorological conditions, topography, physical properties of the pollutants, and the built environment. Air dispersion modeling, resulting in pollution plume buffers rather than fixed-distance buffers, may be able to provide a more realistic depiction of the extent of pollutant impact.

Lastly, our population numbers for the denominators rely on census data, which are known to often be inaccurate, having evinced serious problems with undercounting certain populations, especially in large urban areas. Additionally, census data only take into account the residential location of population, and do not include other locations (e.g., workplace, school) that may be at some distance from the residential address. Since exposure to air pollution occurs during the day, when people may or may not be at home, as well as in the evening when people are more likely to be home, this analysis, by using only residential locations, potentially assigns more weight than is realistic to exposure resulting from proximity of home to pollution sources.

Even after considering these caveats, it is clear that improving understanding of population distribution will enable the emergence of a more accurate picture of the spatial dynamics of disease. Additionally, fields such as crime pattern analysis, hazard and risk assessment, urban planning, environmental analysis, economic development, transportation planning, and public health provision could benefit from the CEDS method. Having more precise insights as to where people live, particularly in hyperheterogeneous urban environments, is critical when looking at any spatio-human issue.

REFERENCES

- Bracken, I. and Martin, D., 1989, The generation of spatial population distributions from Census centroid data. *Environment and Planning A*, Vol. 21, 537–543.
- Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P. H., Wijga, A., Koopman, L. P., Neijens, H. J., Gerritsen, J., Kerkhof, M., Heinrich, J., Bellander, T., and Brunekreef, B., 2002, Air pollution from traffic and the development of respiratory

- infections and asthmatic and allergic symptoms in children. *American Journal of Respiratory and Critical Care Medicine*, Vol. 166, 1092–1098.
- Brunekreef, B. and Holgate, S., 2002, Air pollution and health. *Lancet*, Vol. 360, 1233–1242.
- Chaix, B., Gustafsson, S., Jerrett, M., Kristersson, H., Lethman, T., Boalt, Å., and Merlo, J., 2006, Children's exposure to nitrogen dioxide in Sweden: Investigating environmental injustice in an egalitarian country. *Journal of Epidemiology and Community Health*, Vol. 60, No. 3, 234–241.
- Chakraborty, J. and Armstrong, M. P., 1997, Exploring the use of buffer analysis for the identification of impacted areas in environmental equity assessment. *Cartography and Geographic Information Systems*, Vol. 24, No. 3, 145–157.
- Delfino, R., Gong, H., Linn, W. S., Hu, Y., and Pellizzari, E. D., 2003, Respiratory symptoms and peak expiratory flow in children with asthma in relation to volatile organic compounds in exhaled breath and ambient air. *Journal of Exposure Analysis and Environmental Epidemiology*, Vol. 13, 348–363.
- Department of Information Technology and Telecommunications, 2003, *NYCMap*, 2002. New York, NY: Department of Information Technology and Telecommunications.
- Edwards, J., Walters, S., and Griffiths, R. C., 1994, Hospital admissions for asthma pre-school children: Relationship to major roads in Birmingham UK. *Archives of Environmental Health*, Vol. 49, 223–227.
- Eicher, C. and Brewer, C., 2001, Dasymeric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, Vol. 28, 125–138.
- Fletcher, J., 2006, *Report on the Regression Analysis of Asthma Hospitalization Rates and Proximity to Major Air Pollution Sources*. Bronx, NY: Albert Einstein College of Medicine.
- Gauderman, W. J., Avol, E., Lurmann, F., Kuenzli, N., Gilliland, F., Peters, J., and McConnell, R., 2005, Childhood asthma and exposure to traffic and nitrogen dioxide. *Epidemiology*, Vol. 16, No. 6, 737–743.
- Gauderman, W. J., Vora, H., McConnell, R., Berhane, K., Gilliland, F., Thomas, D., Lurmann, F., Avol, E., Kunzli, N., Jerrett, M., and Peters, J., 2007, Effect of exposure to traffic on lung development from 10 to 18 years of age: A cohort study. *Lancet*, Vol. 369, No. 9561, 1203–1209.
- Goodchild, M., Anselin, L., and Deichmann, U., 1993, A framework for the areal interpolation of socioeconomic data. *Environment and Planning A*, Vol. 25, 383–397.
- Goodchild, M. and Lam, N. S.-N., 1980, Areal interpolation: A variant of the traditional spatial problem. *Geo-Processing*, Vol. 1, 297–312.
- Hitchins, J., Morawsaka, L., Wolff, R., and Gilbert, D., 2000, Concentrations of sub-micrometer particles from vehicle emissions near a major road. *Atmospheric Environment*, Vol. 34, 51–59.
- Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., and van den Brandt, P. A., 2002, Associations between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *Lancet*, Vol. 360, 1203–1209.
- Janssen, N., van Vliet, P., Aarts, F., Harssema, H., and Brunekreef, B., 2001, Assessment of exposure to traffic related air pollution of children attending schools near motorways. *Atmospheric Environment*, Vol. 35, 3875–3884.

- Jerrett, M. and Finkelstein, M., 2005, Geographies of risk in studies linking chronic air pollution exposure to health outcomes. *Journal of Toxicology and Environmental Health*, Vol. 68, No. 13–14, 1207–1242.
- King, G., 1997, *A Solution to the Ecological Inference Problem*. Princeton, NJ: Princeton University Press.
- Krieger, N., Chen, J. T., Waterman, P. D., Soobader, M., Subramanian, S. V., and Carson, R., 2002, Geocoding and monitoring of U.S. Socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? *American Journal of Epidemiology*, Vol. 156, 471–482.
- Langford, M., Maguire, D. J., and Unwin, D., 1991, The areal interpolation problem: Estimating population using remote sensing in a GIS framework. In I. Masser and M. Blakemore, editors, *Handling Geographic Information: Methodology and Potential Applications*. London, UK: Longman, 55–77.
- Liu, X., Clarke, K., and Herold, M., 2006, Population density and image texture: A comparison study. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 2, 187–196.
- Livingstone, A. E., Shaddick, G., Grundy, C., and Elliot, P., 1996, Do people living near inner city main roads have more asthma needing treatment? Case control study. *British Medical Journal*, Vol. 312, 676–677.
- LotInfo, LLC, 2003, *LotInfo, 2003*. New York, NY: SpaceTrack, Inc. Available from 304 Park Ave, 11th Floor New York, NY 10010.
- Maantay, J. A., 2001, Zoning, equity, and public health. *American Journal of Public Health*, Vol. 91, No. 7, 1033–1041.
- Maantay, J., 2005/2007, Asthma and air pollution in the Bronx: Methodological and data considerations in using GIS for environmental justice and health research. *Health and Place*, Vol. 13, 32–56.
- Maantay, J. A., Maroko, A. R., and Herrman, C., 2007, Mapping population distribution in the urban environment: The cadastral-based expert dasymetric system (CEDS). *Cartography and Geographic Information Science*, Vol. 34, No. 2, 77–102.
- Martin, D., 2006, An assessment of surface and zonal models of population. *International Journal of Geographical Information Systems*, Vol. 10, No. 8, 973–989.
- Martin, D., Langford, M., and Tate, N. J., 2000, Refining population surface models: Experiments with Northern Ireland census data. *Transactions in GIS*, Vol. 4, No. 4, 343–360.
- Mennis, J., 2003, Generating surface models of population using dasymetric mapping. *The Professional Geographer*, Vol. 55, No. 1, 31–42.
- Molitor, J., Molitor, N.-T., Jerrett, M., McConnell, R., Gauderman, J., Berhane, K., and Thomas, D., 2006, Bayesian modeling of air pollution health effects with missing exposure data. *American Journal of Epidemiology*, Vol. 164, No. 1, 69–76.
- Mortimer, K. M., Neas, L. M., Dockery, D. W., Redline, S., and Tager, I. B., 2002, The effect of air pollution on inner-city children with asthma. *European Respiratory Journal*, Vol. 19, 699–705.
- Neumann, C. M., Forman, D. L., and Rothlein, J. E., 1998, Hazard screening of chemical releases and environmental equity analysis of populations proximate to toxic release inventory facilities in Oregon. *Environmental Health Perspectives*, Vol. 106, No. 4, 217–226.

- New York City Department of Health (NYC DOH), 1999, *Asthma Facts*. New York, NY: New York City Childhood Asthma Initiative.
- New York City Department of Health (NYC DOH), 2003, *Asthma Facts* (2nd ed.). New York, NY. Retrieved from the New York City Department of Health Web site at <http://www.nyc.gov/html/doh/pdf/asthma/facts.pdf>
- New York City Mayor's Office of Environmental Coordination, 2001, *City Environmental Quality Review (CEQR) Technical Manual*. New York, NY: NY City Mayor's Office.
- Nitta, H., Sato, T., Nakai, S., Maeda, K., Aoko, S., and Oho, M., 1993, Respiratory health associated with exposure to automobile exhaust. I. Results of cross-sectional studies in 1979, 1982, and 1983. *Archives of Environmental Health*, Vol. 48, 53–58.
- Openshaw, S., 1984, *The Modifiable Areal Unit Problem, Concepts and Techniques in Modern Geography*, Vol. 38. Norwich, UK: GeoBooks.
- Peters, J. M., Avol, E., Gauderman, J., Linn, W. S., Navidi, W., London, S. J., Margolis, H., Rappaport, E., Vora, H., Gong, H., and Thomas, D., 1999, A study of twelve Southern California communities with differing levels and types of air pollution. II. Effects on pulmonary function. *American Journal of Respiratory and Critical Care Medicine*, Vol. 159, No. 3, 768–775.
- Reibel, M. and Bufalino, M. E., 2005, Street-weighted interpolation techniques for demographic count estimation in incompatible zone systems. *Environment and Planning A*, Vol. 37, 127–139.
- U.S. Bureau of the Census, 2001a, *Census 2000 Summary File 1 and 3, New York State*. Washington, DC: U.S. Bureau of the Census. Available from www.census.gov
- U.S. Bureau of the Census, 2001b, *Topologically Integrated Geographic Encoding and Referencing system (TIGER)*. Washington, DC: U.S. Bureau of the Census, Geography Division, Cartographic Products Management Branch. Available from www.census.gov
- U.S. Environmental Protection Agency, Emission Inventory and Analysis Group, 2006, *National Emission Inventory (NEI), 2002*. Washington, DC: U.S. EPA. Available from <http://www.epa.gov/ttn/chief/net/>
- U.S. Environmental Protection Agency, Office of Environmental Information, 2002, *Toxic Release Inventory Data (TRI), 2000*. Washington, DC: U.S. EPA. Available from www.epa.gov/tri/tridata/tri00/
- Venn, A. J., Lewis, S. A., Cooper, M., Hubbard, R., and Britton, J., 2001, Living near a main road and the risk of wheezing illness in children. *American Journal of Respiratory and Critical Care Medicine*, Vol. 164, 2177–2180.
- Weichselbaum, J., Petrini-Monteferrri, F., Papatoma, M., Wagner, W., and Hackner, N., 2005, Sharpening census information in GIS to meet real-world conditions: The case for Earth observation. *WIT Transactions on Ecology and the Environment, Sustainable Development and Planning II*, Vol. 1, 143–152.
- Wilkinson, P., Elliott, P., Grundy, C., Shaddick, G., Thakrar, B., Walls, P., and Falconer, S., 1999, Case-control study of hospital admission with asthma in children aged 5–14 years: Relation with road traffic in northwest London. *Thorax*, Vol. 54, No. 12, 1070–1074.
- Zhu, Y. F., Hinds, W. C., and Kim, S., 2002, Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment*, Vol. 36, 4323–4335.