Exploratory Analysis of Suburban Land Cover and Population Density in the U.S.A.

Francesca Pozzi [1] and Christopher Small [2]

[1] *CIESIN*
Columbia University
Palisades, NY 10964 USA
fpozzi@ciesin.columbia.edu

[2] *Lamont Doherty Earth Observatory*
Columbia University
Palisades, NY 10964 USA
small@LDEO.columbia.edu

Reprinted from:

IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas,
Paper 35
Rome, Italy
8-9 November, 2001

Exploratory Analysis of Suburban Land Cover and Population Density in the U.S.A.

Francesca Pozzi, Christopher Small

Abstract— **The objective of this study is to investigate the consistency of "suburban" population densities and land covers. We analyzed population density, extracted from the census, and vegetation abundance, derived from Landsat imagery, taking six cities in the U.S.A. as contrasting examples. Combining population density and areal vegetation abundance estimates yields univariate and bivariate distributions for the two variables. We quantify the relationship between population density and vegetation fraction in Atlanta, Chicago, Los Angeles, New York, Phoenix and Seattle. A bimodal distribution of population density in the U.S.A. suggests that it may be possible to characterize "suburban" areas on the basis of population density between 100 and 10,000 people/km^2. The maximum areal vegetation cover diminishes linearly with the Log_{10} of population density in cities with large density ranges.**

Keywords— **Suburban, Population Density, Vegetation Fraction, Land Cover**

I. INTRODUCTION

Suburban areas in the U.S.A.are often perceived as the greener "residential areas on the outskirts of a city or a large town" [1], socially and economically dependent on large cities. Suburbs have been given consideration in the past decade, given the close connections with cities and the negative consequences associated with urban sprawl. These consequences include loss of agricultural land and natural vegetation, increased traffic congestion and associated degradation of air quality.

According to the U.S.Census Bureau [2], between 1995 and 1996 more than 2 million people moved from U.S.A. cities and from non-metropolitan areas into the "suburbs". In the same report suburbs are defined as "all territory within an Metropolitan Statistical Areas (MSA) but outside of a central city". Although this definition is intuitive and easily understandable, there appears to be no consistent or formal characterization of suburban areas in terms of physical or socioeconomic characteristics.

In recent years, many authors have considered the relationship between population characteristics and environmental variables, but their interests and goals are different from the ones presented in this paper. Examples are [3], [4], [5], where the authors presented studies of integration of population and land cover for quality of life assessment or improvement of land cover classification. Studies [6] and [7] considered population and vegetation to better understand urban dynamics. Other authors analyzed land cover change related to population change and its impacts on the surrounding natural areas [8]. Approaches more focused on the correlation between socioeconomic variables, such as population count and housing density, and land cover are presented in [9] and [10].

The above mentioned studies focused mainly on the integration of population and land cover data and on single case studies. To our knowledge, no study has yet been performed to examine the demographic and land cover characteristics of suburban areas and their consistency across different physiographic environments.

The purpose of this paper is to investigate the question of whether suburban areas can be defined based on demographic and physical characteristics, specifically population density and vegetation cover. Based on the expectation that suburban areas are greener than urban centers, and that the predominant suburban land cover is vegetation, we attempt to quantify the extent to which suburban areas are vegetated in different U.S.A. cities. We consider suburban areas based on population density and on apparent spectral reflectance using Landsat data to quantify the relationship between the two variables looking at the cities of Atlanta, Chicago, Los Angeles, New York, Phoenix and Seattle.

II. DATA

Given the purpose of the study, the six cities we considered present very different characteristics, in terms of geographic location and spatial structure and dynamics. We chose cities located in a temperate climate, both in a deciduous forest biome (New York, Chicago, Atlanta) and in an evergreen forest biome (Seattle) and cities located in an arid or semi-arid climate (Los Angeles and Phoenix). We also included cities that have been among the most fast-growing of the past decades in the U.S.A. (Phoenix and Seattle), and cities that have experienced rapid growth in the past and now are characterized by large population (New York, Chicago, Los Angeles).

A. Population Density

We calculated population density from the 1990 U.S. Census Bureau population counts at the block level, the lowest in the U.S. census structural hierarchy. These data are available separately as spatial data (*Topologically Integrated Geographic Encoding and Referencing system - TIGER®*) and tabular data (*Summary Tape Files-STFs*) for each county in the U.S.A.. In this study we considered population density, expressed in persons/km^2. For each city we selected one or more counties containing the Centered Business District (CBD) and the surrounding sub-

F. Pozzi is a research associate at the Center for International Earth Science Information Network (CIESIN) of Columbia University. E-mail: fpozzi@ciesin.columbia.edu

C. Small is a researcher at the Lamont Doherty Earth Observatory of Columbia University. E-mail: small@LDEO.columbia.edu

urbs. This resulted in the selection of the following counties, for which we then created a smaller subset concordant with Landsat coverage (area reported in parenthesis):

- **Atlanta**: DeKalb and Fulton (*900 km*^{2});
- **Chicago**: Cook (*950 km*^{2});
- **Los Angeles**: Los Angeles (*3100 km*^{2});
- **New York Metropolitan Area**: Bronx, Kings, New York, Queens, Richmond, Bergen, Essex, Hudson, Passaic, Nassau, Rockland, Westchester (*2000 km*^{2});
- **Phoenix**: Maricopa (*4700 km*^{2});
- **Seattle**: King (*3200 km*^{2}).

B. Vegetation Fraction

The characteristic spatial scale and the spectral variability of urban and suburban land cover poses serious problems for traditional image classification algorithms. In areas where the reflectance spectra of the land cover vary appreciably at scales comparable to, or smaller than, the Ground Instantaneous Field Of View (GIFOV) of most satellite sensors, the spectral reflectance of a individual pixel will generally not resemble the reflectance of a single land cover class but rather a mixture of the reflectances of two or more classes present within the GIFOV. Because they are combinations of spectrally distinct land cover types, mixed pixels in urban areas are frequently misclassified as other land cover classes. Similarly, the definition of an "urban" spectral class will usually incorporate pixels of other non-urban classes.

If an urban area contains significant amounts of vegetation then the reflectance spectra measured by the sensor will be influenced by the reflectance characteristics of the vegetation. Macroscopic combinations of homogeneous "endmember" materials within the GIFOV produce a composite reflectance spectrum that can often be described as a linear combination of the spectra of the endmembers [11]. If mixing between the endmember spectra is predominantly linear and the endmembers are known *a priori*, it may be possible to "unmix" individual pixels by estimating the fraction of each endmember in the composite reflectance of a mixed pixel [12], [13].

Analysis of Landsat TM imagery suggests that the spectral reflectance of many urban areas can be described as linear mixing of three distinct spectral endmembers [14], [15]. Principal component analysis of urban reflectance consistently yield eigenvalue distributions suggesting that the majority of scene variance is contained within a two dimensional mixing plane. The triangular distribution in the mixing space defined by the principal components bears a similarity to the well known Tasseled Cap distribution discovered by [16]. The feature space distributions are similar in the sense that both contain a vegetation endmember that is distinct from a mixing continuum between high and low albedo endmembers.

The spectral endmembers determined for the areas investigated here correspond to low albedo (e.g. water, shadow, roofing), high albedo (e.g. cloud, sand, roofing) and vegetation. The strong visible absorption and infrared reflectance that is characteristic of vegetation is sufficiently distinct from the spectrally flat reflectance of the low and high albedo endmembers to allow the three components to be "unmixed" by inverting a simple three component linear mixing model [14]. The result of the unmixing is a set of fraction images showing the areal percentages, given as fractions between 0 and 1, of each endmember present within each pixel. Analysis of Landsat, Ikonos and AVIRIS imagery of several urban/suburban areas shows that a three component linear mixing model provides stable, consistent estimates of vegetation fraction for both constrained and unconstrained inversions using three different endmember selection methods [15]. Vegetation fraction estimates derived from Landsat TM data were validated with aeral vegetation fractions calculated from 2 m aerial photography and generally showed agreement to within 10% [14] The vegetation fraction estimates given here were derived from Landsat TM and validated with Ikonos MSI imagery.

III. Analysis and Results

The first step of the analysis was to quantify the distribution of population density across the entire United States to estimate whether rural, urban and suburban areas are clearly discernible based on population density. The distribution of people as a function of population density for the U.S.A. in 1990 is bimodal (Figure 1) [17]. The modes represent the spatially concentrated settlements near cities and the spatially dispersed settlements farther from cities. The larger mode has a distinct break in slope near 10,000 people/km^2, and a short, high density tail. This tail corresponds to the high density cores of large cities. For the purposes of this study we consider suburban areas to be characterized by population density between 100 and 10,000 people/km^2.

To perform the study on the six cities, spatial and tabular data from the Census were initially aggregated based on the block numeric codes for each county. The resulting vector layers were then projected to UTM coordinates,

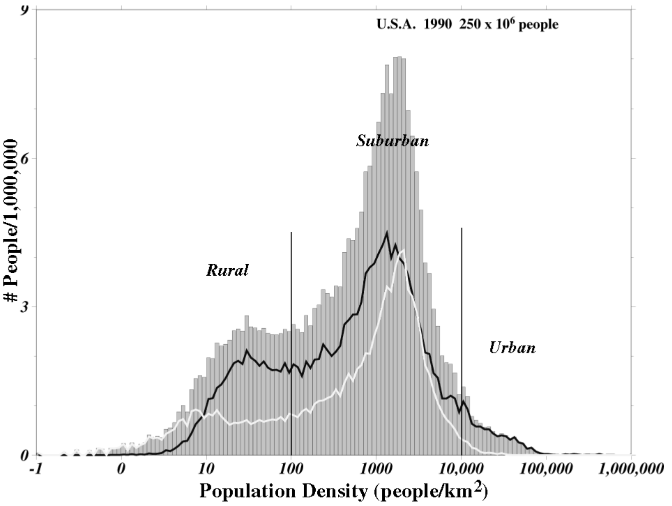

Fig. 1. Histogram of the Population Density for the U.S.A., showing also the distribution for Eastern U.S.A. (East of the 90 ° W, black line) and for Western U.S.A.(grey line).

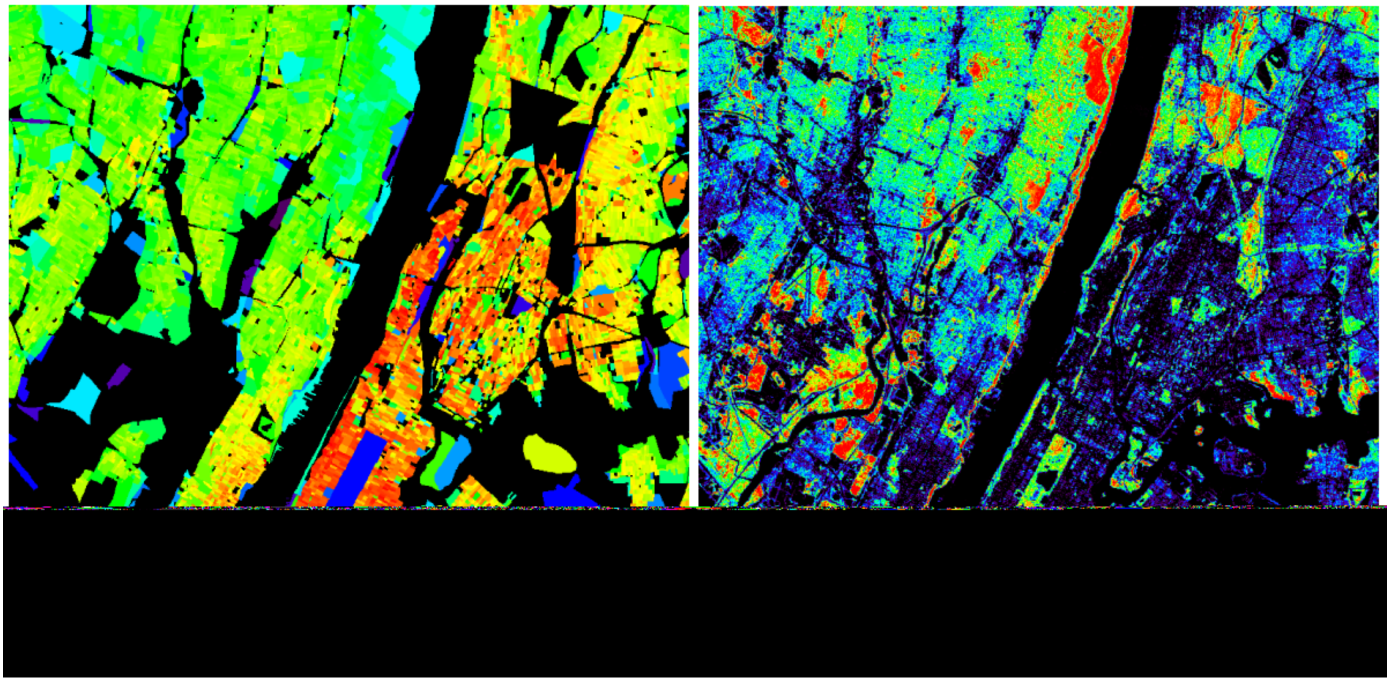

Fig. 2. Population Density and Vegetation Fraction at full resolution for part of New York Metropolitan Area. Black areas represent no data for population density and no vegetation (water) for the vegetation fraction.

rasterized to a 30 m grid and coregistered to the Landsat data. To quantify the relationship between population density and vegetation fraction, we produced bivariate distributions of people and land area as functions of population density and vegetation fraction. The bivariate population distributions are shown with population density and vegetation fraction for each city in Figure 3. We then summed the bivariate distributions to produce marginal distributions of people as functions of population density and vegetation fraction for each city (Figure 4).

IV. Discussion

In the six cities we studied there appears to be a pattern for suburban areas, both in terms of population density and vegetation fractions. The population density histograms, calculated for the counties listed in II-A, show the suburban peak characteristic of the entire U.S.A. with Atlanta and New York at the extremes (Figure 4. The vegetation fraction histograms also show a consistent pattern, with peaks varying between about 0.1 and 0.55. The cities with large urban core also present a smaller peak for vegetation fractions less than 0.01. Prominent exceptions are Atlanta, which has a symmetric distribution centered on higher values (0.5 to 0.6) and New York, which has a long tailed monotonic distribution, with a peak at less than 0.1.

The differences between the physiographic environments and the urban structures for the six cities are such that the peaks of the bivariate histograms are spread across a range of population densities and vegetation fractions. Nonetheless a consistent sub-linear relationship is seen for the largest cities (New York, Chicago and Los Angeles). These three cities have similar density distributions, with comparable peak values and with vegetation fractions linearly decreasing with Log_{10}(Population Density). Phoenix and Seattle have the most similar population density distributions, but their vegetation fraction distributions are different, due to their arid and humid climates. The bivariate distributions for these two cities are more isotropic than the larger cities, with Phoenix containing large inhabited unvegetated areas and Seattle characterized by large uninhabited and densely vegetated areas. Atlanta, on the other hand, presents a more uniform distribution, with little variations in either vegetation fraction and population density.

V. Conclusions

The objective of this study was to investigate the consistency of “suburban” settlement patterns, based on the relationship between population density and land cover among different cities in the U.S.A.. The principal conclusion of this study is that the population density distribution in the U.S.A. may provide a demographic basis for distinguishing urban, suburban and rural areas. The block level population density distribution for the entire U.S.A. shows two distinct modes corresponding to moderate density (100 to 10,000 people/km^2) settlements surrounding higher density urban cores and to low density settlements (less than 100 people/km^2) dispersed throughout the country. The high density (more than 10,000 people/km^2) cores correspond to a distinct tail delineated by a break in slope on the moderate density ”suburban” mode. This demographic classification would place 71% of Americans in suburban areas, 25% in rural areas and 3% in urban areas in 1990.

The wide variety of land use classes that character-

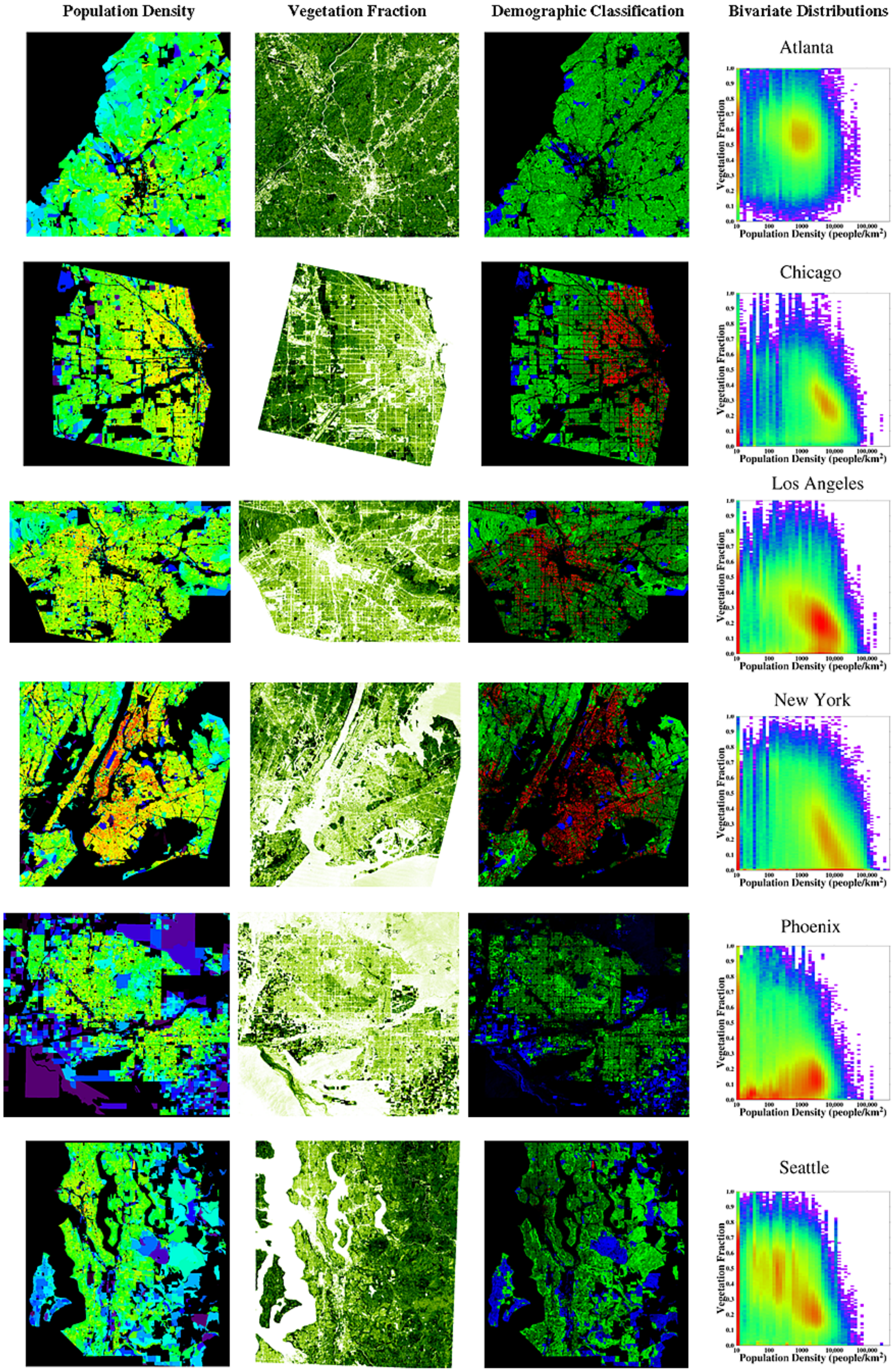

Fig. 3. Spatial Distributions of Population and Vegetation. Combining population density with vegetation fraction yields a demographic classification, where rural population densities are shown in blue, urban in red and suburban in green. Different shades of green correspond to different amounts of vegetation. Note the similarity of the peaks in the bivariate distributions for Chicago, New York and Los Angeles. Full resolution images are available at www.LDEO.columbia.edu/˜ small/Urban.html

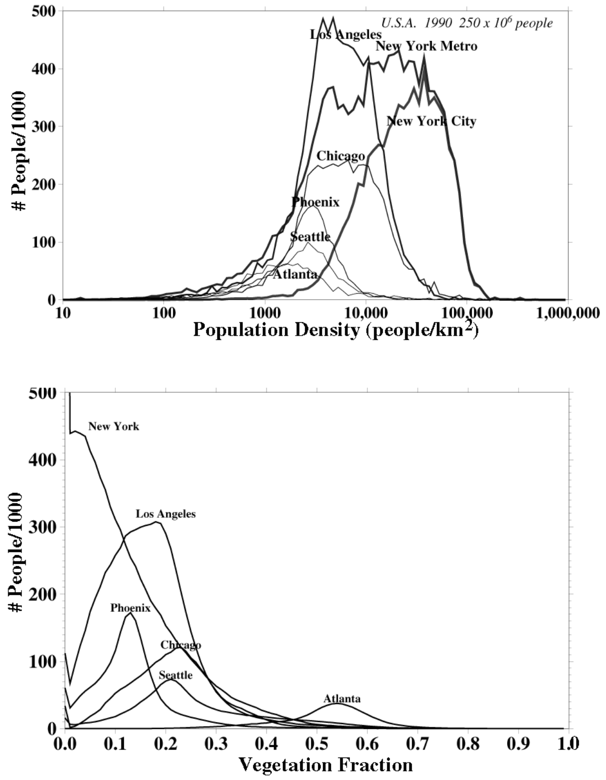

Fig. 4. Univariate Distributions of number of people as a function of Population Density (top) and Vegetation Fraction for the six cities.

ize urban and suburban areas poses serious problems for thematic classifications that rely on moderate resolution imagery. Unlike most other landcover classes, the urban/suburban mosaic is consistent only in its spectral heterogeneity at the scale of most operational satellite sensor GIFOVs [15]. The resolution difference between census tracts and the Landsat sensor does not allow for a simple spectral characterization of suburban landcover.

In the U.S.A. cities we investigated, the most consistent spectral characteristic of "demographically suburban" areas was related to the amount of vegetation cover. Maximum vegetation fraction generally diminishes with increasing population density but spectral heterogeneity still results in a wide range of vegetation fractions in demographically suburban areas. Large cities with high density urban cores do, however, show a distinct linear decrease in the modal vegetation fraction with increasing Log_{10}(population density). The different physical environments of the cities considered here result in different vegetation distributions at different settlement densities. We find no evidence for a single consistent relationship between suburban population density and vegetation abundance in the U.S.A..

Quantitative characterization of vegetation abundance in suburban areas does however provide a basis for comparison of the physical environments in which most Americans reside. Vegetation has a direct impact on solar energy flux through the environment and therefore influences the microclimate of the human habitat. Vegetation abundance and distribution also control evapotranspiration and albedo thereby influencing climate dynamics at regional and local scales. Coanalysis of settlement patterns and land cover characteristics may eventually facilitate quantitative analysis of urban sprawl, natural resource management and land use policy implications, when the relationship between these factors is understood.

AKNOWLEDGEMENTS

The authors would like to acknowledge the support and resources of CIESIN's SocioEconomic Data and Application Center (SEDAC), the Columbia Earth Institute and Lamont-Doherty Earth Observatory of Columbia University.

REFERENCES

[1] *Merriam-Webster Dictionary*, Merriam-Webster, 2001.
[2] Hansen K.A., "Geographic Mobility: March 1995 to March 1996," Tech. Rep., U.S.Census Bureau, November 1997.
[3] Lo P. and Faber B.J., "Integration of Landsat Thematic Mapper and Census Data for Quality of Life Assessment," *Remote Sensing of Environment*, vol. 2, pp. 143–157, 1997.
[4] Harris P. M. and Ventura S. J., "The Integration of Geographic Data with Remotely Sensed Imagery to Improve Classification in an Urban Area," *Photogrammetric Engineering and Remote Sensing*, vol. 61, no. 8, pp. 993–998, 1995.
[5] Vogelmann J. E. et al, "Regional Characterization of Land Cover Using Multiple Sources of Data," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 1, pp. 45–57, 1998.
[6] Emmanuel R., "Urban Vegetational Change as Indicator of Demographic Trends in Cities: the Case of Detroit," *Environment and Planning*, vol. B 24, pp. 415–426, 1997.
[7] Wagner T.W.and Ryznar R.M., "Toward Understanding Urban Processes with Remotely Sensed Data: Examples from Detroit," in *Proceeding of GeoInformatics'99 Conference*, University of Michigan, Ann Arbor, MI, 1999.
[8] Wang Y. and Zhang X., "Land Cover Change of Metropolitan Chicago Area from 1972 to 1997 and the Impact to Natural Communities in the Region," in *Proceedings of GeoInformatics'99*, University of Michigan, Ann Arbor, MI, 1999.
[9] Smith R.M. Yuan Y. and Limp W.F., "Remodeling Census Population with Spatial Information form Landsat TM Imagery," *Computers, Environment and Urban Systems*, vol. 21, no. 3/4, pp. 245–258, 1997.
[10] Radeloff V. C. et al, "Exploring the Spatial Relationship Between Census and Land-Cover Data," *Society and Natural Resources*, vol. 13, no. 6, pp. 599–609, 2000.
[11] Singer R. B. and McCord T.B., "Large Scale Mixing of Bright and Dark Surface Materials and Implications for Analysis of Spectral Reflectance," *Journal of Geophisical Research*, pp. 1835–1848, 1979.
[12] Adams J. B. et al, "Spectral Mixture Modeling: a New Analysis of Rock and Soil Types at the Viking Lander 1 Site," *Journal of Geophysical Research*, vol. 91, pp. 8098–8122, 1986.
[13] Boardman J. W., "Inversion of Imaging Spectrometry Data Using Singular Value Decomposition," in *Proceeding of the 12th Canadian Symposium on Remote Sensing*, 1989, pp. 2069–2072.
[14] Small C., "Estimation of Urban Vegetation Abundance by Spectral Mixture Analysis," *International Journal of Remote Sensing*, vol. 22, no. 7, pp. 1305–1334, 2001.
[15] Small C., "Multiresolution Analysis of Urban Reflectance," in *Proceedings of IEEE joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Rome, Italy, 8-9 November 2001.
[16] Kauth R. J. and G. S. Thomas, "The Tasseled Cap - a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat.," in *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, 1976, p. 4041.
[17] Small C., "A Global Analysis of Urban Population Distribution and the Physical Environment," in *Proceedings of the Open Meeting of the Human Dimension of Global Environmental Change Research Community*, Brazil Academy of Science, Rio de Janeiro, Brazil, 6-8 October 2001.