Using Population Data to Address the Human Dimensions of Environmental Change

D.M. Mageean and J.G. Bartlett, University of Maine

Resource Management Requirement

In recent years researchers and policy makers have identified population-environment interactions as crucial to issues of ecology, economic development, and human welfare. It seems clear that human populations and demands on the environment are driving ecological change in such areas as global warming, ozone depletion, deforestation, biodiversity loss, land degradation, and pollution of air and water. In developing countries, the anthropogenic effects include drastic environmental deterioration in areas where population pressures exceed a particular threshold (Terborgh 1989). In developed countries, environmental problems have arisen because increasing incomes, leisure, and ease of communication have generated a stronger demand for recreation and tourism (Bayfield 1979).

Enormous gaps still exist in the scientific understanding of precisely how demographic factors—such as population size and growth rate, settlement distribution, and migration dynamics—affect resources and the environment. For example, determining what proportion of environmental impacts result from population growth versus behavior (e.g., consumption patterns) is difficult. Furthermore, because of the fragmentation of research among disciplines, achieving a holistic picture is problematic.

In 1993, Paul Stern of the National Research Council described the need for a second environmental science that would be "focused on human-environment interactions—to complement the science of environmental processors by analyzing key questions" (Stern 1993). Stern outlined three main fields of inquiry for such questions: (1) the study of human causes of environmental change, (2) the effects of environmental change on things people value, and (3) the study of respective feedback between humanity and the environment. Unfortunately, little unanimity of opinion regarding the nature of the relationship hampers efforts in the first field of inquiry. As recently as 1992, Daniel Hogan wrote that when the relationships between population growth and the physical environment are considered, demography had advanced little beyond Malthusian arithmetic (Hogan 1992). Of particular importance is the need to more fully understand the nature of dynamic interactions between population—a major driving force in global environmental change—and natural resources. Changes in land use can be affected by both population growth and changes in population distribution resulting from migration flows.

Although considerable attention has been devoted to the impact of population growth on sensitive ecosystems in the developing world, the study of human impacts in the United States may be just as important because it affords particular opportunities for modeling and increased understanding. Given the projected growth rate for the country and certain regions within it, this understanding is vital in order to project environmental consequences. Between 1990 and 1995, the U.S. population grew by 1.0% per year on average, a slight increase from the average annual growth rate of 0.9% in the 1980s. However, growth patterns varied substantially by geographic region.

The U.S. population is one of the most mobile in the industrialized world; over the past three decades it has been steadily shifting from northern Frostbelt states to southern and western Sunbelt states. Between 1990 and 1995, the most rapid growth has been in the West, particularly in the mountain states of Nevada, Idaho, Arizona, Colorado, Utah, and New Mexico, where average population growth rates of 2% or more are common. If sustained, albeit unlikely, such growth rates would double the populations of certain states in just 35 years—a pace faster than many developing countries. Generally perceived as an economic benefit, short- and long-term population growth can also present tremendous environmental challenges. Moreover, responding to such challenges is complicated by rapidly shifting patterns of population growth. Between 1993 and 2020, the U.S. population is projected to climb from 258 million to 326 million. The South and West regions—areas already under environmental stress—are expected to account for 82% of the growth during this period. Most of the growth will occur in eight states, so land use issues and the ability to model human impacts on the environment are likely to become critical.

Data sets on population are available for a range of spatial scales for most regions on the Earth and are particularly reliable in the developed world. Two key questions that must be addressed in this context follow.

- How should such data sets be exploited to advance understanding of the human dimensions of changing land use/land cover?
- How compatible are the data sets with the diverse biological and environmental data sets currently available, particularly in relation to spatial units of analysis?

According to the Human Dimensions Program "…a high priority in research on the human dimensions of global environmental change must be placed on conceptual and methodological issues since, without appropriate concepts and methodologies, research cannot be undertaken" (Jacobson and Price 1991). This case study shows how new concepts emerge from methodological and analytical advances that seek to integrate remote sensing, environmental, and demographic data that target changes in U.S. population distribution and growth. It describes collaborative work with the Biodiversity Research Consortium on links between a suite of measures of human activity obtained from the 1980 and 1990 U.S. Censuses and landscape metrics that quantify spatial patterning of landscapes in ecologically relevant ways that can be tracked using remote sensing.

Methods

The primary unit of analysis used was the 640km^2 hexagon derived from the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP; see Kiester et al. 1993). A digital grid of 12,600 hexagons was overlaid onto the conterminous United States. Hexagons were chosen as the basic sampling unit because the distance between the centroids of any two adjacent hexagons is a constant 27km. Mapped data of landscape and habitat types were available from an analysis of Advanced Very High Resolution Radiometry (AVHRR) meteorological satellite images (NOAA) (Loveland et al. 1991). The satellite's sensor resolution is 1.1km^2.

The Loveland et al. land cover classification, derived from the AVHRR data, formed the basis for the study's landscape metrics. One hundred fifty-nine land cover classes of the scheme were aggregated into 13 coarser classes of an Anderson Level II scheme, and a final urban class was added from the Digital Chart of the World (Danko 1992). The 14 land cover types were cropland/pasture, grassland/cropland, woodland/cropland, grass-dominated, shrub-dominated rangeland, mixed grass/shrub rangeland, deciduous forest, coniferous forest, mixed deciduous/coniferous forest, water bodies, coastal wetlands, barren or sparsely vegetated land, alpine tundra, and urban areas.

Land cover characteristics were calculated for each hexagon at Anderson Level II (14 land cover classes) by summarizing the distribution of each land cover class across all 1.1km^2 AVHRR-derived pixels in each hexagon. In addition, landscape pattern metrics were computed for variables such as patch size distributions for various cover classes, shape complexity and fractal dimension, types and frequency of edges between habitat types, and measures of road abundance and total length of all major riparian systems present per hexagon (O'Connor et al. 1996, Hunsaker et al. 1994). Long-term weather data for average annual precipitation, mean January and July temperatures and annual temperature variation for the same period were derived from the Historical Climate Network Database.

One goal of the case study was to determine the extent to which population density or its near-equivalents either captures most anthropogenic interaction with the environment or is merely one facet thereof. Consequently, nine variables from the 1990 county-level census data file (U.S. Bureau of the Census, 1990) deemed most likely to capture key demographic facets and be positively linked to variation in land use/land cover patterns across the conterminous U.S. were collected and examined using principal components analysis (PCA). The list of variables follows.

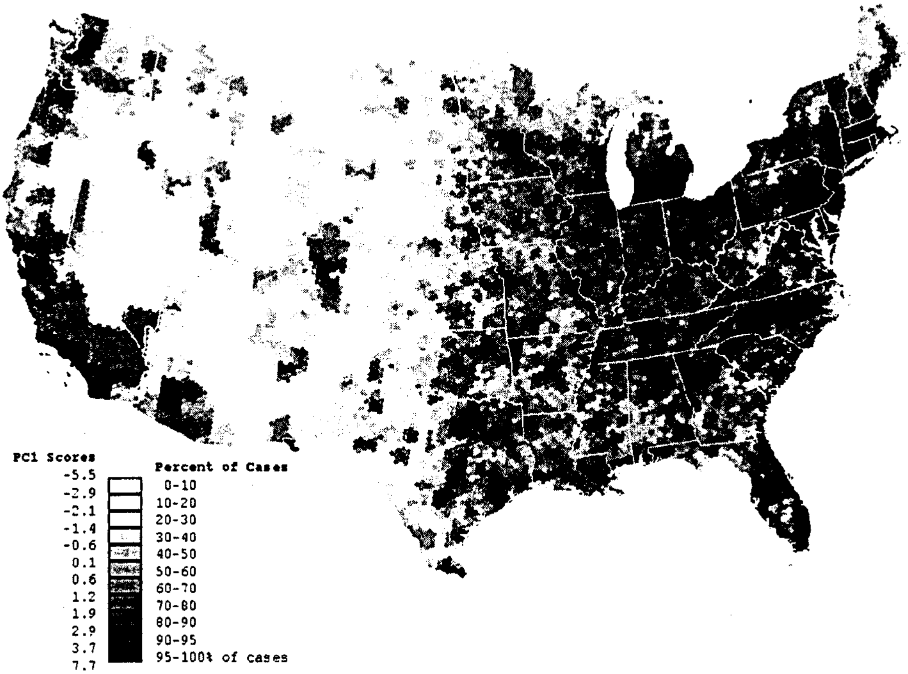
- Change in population (1980-1990)
- Mean age of structure (1990)
- Metropolitan or nonmetropolitan status (1990)
- Total number of farms (1987)
- 1980 and 1990 population
- Total acreage in farms (1987)
- Total number of housing units (1990)
- Per capita income (1989)

The variables provided measures of population density and growth, surrogate measures of date and intensity of settlement, and measures of the urban or rural nature of the area. Quantified in terms of per capita income, wealth (a measure of affluence) was included because of the potential relationship to consumption patterns.

Because landscape pattern metrics were calculated for each hexagon and census variables were at the county level, the digital county level boundary file was overlaid onto the digital EMAP hexagon grid in ARC/INFO. Weighted values for each census variable per hexagon were calculated from the intersected coverages. Area weighting was used for density measures and population density weighting was used for per capita income. All census variables were appropriately normalized prior to PCA.

Results and Discussion

PCA was used to create a set of composite indices of human effects. The analysis generated variables that minimized the total residual sum of squares after fitting linear functions in all census variables across all hexagons. This process yielded two axes of interest, which were then interpreted as an index of *human settlement* in the case of PC1 and *density independent growth and settlement* in the case of PC2 (Mageean and Bartlett 1996). The first principal component accounted for 54% of the total variance and had positive major loadings on four variables—1980 population, 1990 population, wealth index, and housing density. When mapped across the 48 conterminous states (see the next figure), the PC1 scores broadly parallel the pattern of what most demographers would call population density, but the PC axis has the virtue of using information from multiple census variables and is better described as an index of human settlement. Note that darker tones in the image identify more densely populated regions.

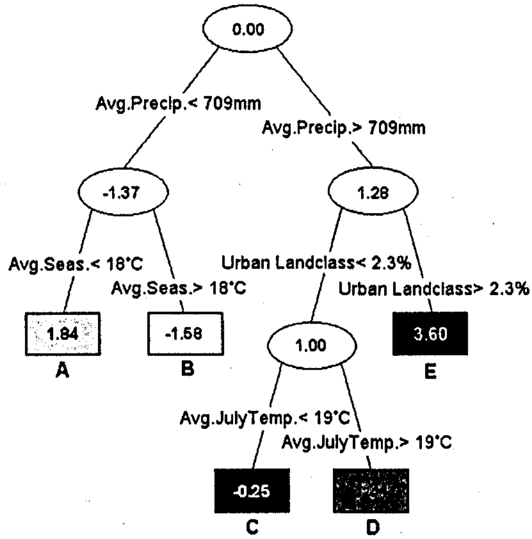


Distribution of PCA scores for human settlement index based on nine variables extracted from the 1990 U.S. Census. Black areas denote highest values (top 5%) for the index.

From the human dimensions of the environment perspective, exploring the relationship of the index to environmental factors is worthwhile. For example, climate and topography may constrain settlement, or settlement may determine subsequent land use. Because correlation between specific landscape metrics and the study's demographic indices may be modified by other cultural, political, and/or socioeconomic variables, and because specific landscape variables may have distinct effects in different parts of the country, correlation analysis and traditional linear regression modeling are inappropriate for this analysis. Consequently, an adaptive statistical technique was used to identify significant, nonlinear, regionalized relationships among land use and climate covariates. Called classification and regression tree (CART) analysis (Breiman et al. 1984), the technique recursively partitions a focal variable (e.g., human settlement index) with respect to a set of independent variables. For each independent variable, a splitting threshold is chosen to maximize differences in the response variable (maximum between-group diversity), and the data set is split into two subsets. The independent variable that best splits the response variable explains the most variation in the data; that variable is used in the tree as a splitting variable. The process is then repeated independently and recursively on each increasingly homogenous subgroup until a stopping criterion is satisfied.

CART analysis was employed to partition the variance in the human settlement index among the environmental variables considered, and recursive partitioning yielded the hierarchical model depicted in the next image. Ovals denote split points (splitting variables are listed), while rectangles denote end points. Both ovals and rectangles contain within-group mean values for the human settlement index. The human settlement index was first split on the basis of annual precipitation with a threshold of 709mm. Drier areas followed the left-hand branch, while wetter areas followed the right-hand branch. Drier areas were subsequently segregated into nodes A and B based on seasonal differences. Wetter areas were partitioned into urban areas with greater than 2.3% of the land area classified as urban (node E) and nonurban areas. Nonurban areas were further segregated into nodes C and D based on an average July temperature threshold of 19°C.

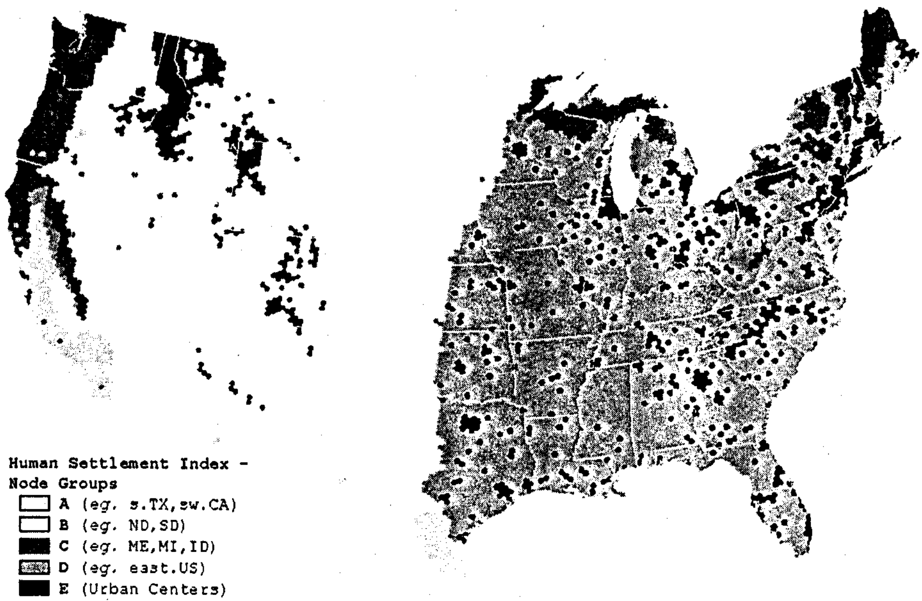
Regression tree model rules relating an index of human settlement to environmental and remotely sensed land use variables, resulting in five end nodes or sets of hexagons with shared environmental conditions of relevance to settlement index.



Regression tree model rules

Node	Environmental condition
Node A	Hexagons with less than 709mm annual precipitation and minimal difference between January and June temperatures (low seasonality).
Node B	Drier areas with high seasonality.
Node C	Wetter (i.e., with precipitation greater than 709mm), nonurban hexagons (less than 2.3% urban representation) with cooler summer temperatures (average July temperatures less than 19°C).
Node D	Wetter, nonurban hexagons with warmer July temperatures.
Node E	Wetter, urban hexagons with greater than 2.3% urban representation.

The next image shows the location of hexagons in each end node summarized in the table.



Regionalization of regression tree generated determinants for the human settlement index.

The results suggest two general conclusions. First, urban centers are primarily in the wetter East (node E, $\bar{x}$ = 3.60) and are otherwise driven by geographical factors such as proximity to rivers and the Atlantic Ocean, as well as historical factors such as area of initial settlement and penetration of the country. The latter phenomenon may have been limited by aridity in the West, an idea supported by the concentration of node E sites along the Pacific Coast (see the previous image). Second, strong interactions among climatic variables appear to influence settlement patterns; in wetter, nonurban areas summer temperatures were critical (nodes C and D), with warm summers favored (node D, $\bar{x}$ = 1.34) but in arid areas annual temperature variation was critical, with seasonably equable areas favored for settlement (node A, $\bar{x}$ = 1.84). Note that a series of sensitivity analyses revealed no collinearity among climate variables.

While in a sense these findings are well known to geographers, this analysis allows quantification of the settlement pattern's dependence on, or independence from, environmental factors. In effect, it identifes *interactions* between settlement and environment rather than *correlations.*

Of particular interest to those concerned with the environmental impact of population is the study's second principal component. This was a multivariate structure contrasting areas of high population growth accompanied by new building with areas of farming and established settlement patterns, in essence measuring the effect of population growth and redistribution. Because this differential growth is orthogonal to the first, such differential growth is independent of the general pattern of settlement. Consequently, the index is described as density independent growth and settlement (DIGS), and it measures growth over the 1980-1990 period involving new development away from land allocated to agriculture. When analyzed in a CART tree, a node that segregated the locations with the highest values of this index was obtained. The locations of this growth were selectively concentrated on coastal barrier islands and dunes, and along the edges of desert areas, all locations of scarce fragile ecosystems. This national pattern of impact, which does not appear to have been previously documented, has major conservation implications.

While the scientific understanding of environmental and demographic change is dramatically increasing when studied separately, an ability to link the two in a synthetic and holistic way has proven elusive. Furthermore, while much of the attention surrounding the population-environment issue has been directed toward population growth, there is a need to examine the influence of other demographic processes such as migration and urbanization. This project uses GIS technology to explore new methods for evaluating the spatial relationships between population and land use. The combination of digital data, methods, and conceptual analysis incorporating remotely sensed data presented here appears to offer a powerful way of detecting human impact on the environment.

Acknowledgments

We wish to thank Biodiversity Research Consortium collaborators Carolyn Hunsaker, B. Jackson, and S. Timmons (Oak Ridge National Laboratory) for provision of landscape metrics; R. Neilsen, D. Marks, J. Chaney, C. Daly, and G. Koerper (U.S. Environmental Protection Agency Environmental Research Laboratory, Corvallis, Oregon) for assistance in computing climate data; Tom Loveland (EROS Data Center, U.S. Geological Survey) for land cover class data; and Raymond J. O'Connor (Wildlife Ecology, University of Maine) and Denis White (Geosciences, Oregon State University) for assistance with spatial analysis. John Bartlett acknowledges financial support for this work from Interagency agreements DW12935631 between USEPA and USDA Forest Service, USEPA Cooperative Agreements CR818843-01-0 and CR823806-01-0, and USDA Forest Service Cooperative Agreement PNW93-0462 with the University of Maine (Raymond J. O'Connor, principal investigator).

References

Bayfield, N.G. "Recovery of four heath communities on Cairngorm, Scotland, from disturbance by trampling." *Biological Conservation* 15 (1979), 165-79.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification and Regression Trees*. Monterey, California: Wadsworth & Brooks, Inc., 1984.

Danko, D.M. "The digital chart of the world." *GeoInfo Systems* 2 (1992), 29-36.

Hogan, D.J. "The impact of population growth on the physical environment." *European Journal of Population* 8 (1992), 109-23.

Hunsaker, C.T., R.V. O'Neill, S.P. Timmins, B.L. Jackson, D.A. Levine, and D.J. Norton. "Sampling to Characterize Landscape Pattern." *Landscape Ecology* 9 (1994), 207-26.

Jacobson, H.K. and M.F. Price. *Human Dimensions Program: A Framework for Research on the Human Dimensions of Global Environmental Change.* Report No. 1. International Social Science Council, 1991.

Kiester, A.R., D. White, E.M. Preston, L.M. Master, T.R. Loveland, D.F. Bradford, B.A. Csuti, R.J. O'Connor, F.W. Davis, and D.M. Stoms. *Research Plan for Pilot Studies of the Biodiversity Research Consortium.* Washington, D.C.: U.S. Environmental Protection Agency, 1993.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown. "Development of a land-cover characteristics database for the conterminous U.S." *Photogrammetric Engineering and Remote Sensing* 57 (1991), 1453-63.

Mageean, D.M. and J.G. Bartlett. "Putting people on the map: integrating social science data with environmental data." Proceedings of the PECORA Thirteen Conference, Sioux Falls, South Dakota, August 1996. Bethesda, Maryland: American Society of Photogrammetry and Remote Sensing, 1996.

O'Connor, R.J., M.T. Jones, D. White, C. Hunsaker, T. Loveland, B. Jones, and E. Preston. "Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States." *Biodiversity Letters* 3 (1996), 97-110.

Stern, P.C. "A Second Environmental Science: Human-Environment Interactions." *Science* 260 (1993), 1897-99.

Terborgh, J. *Where have all the birds gone?* Princeton: Princeton University Press, 1989.

U.S. Bureau of the Census. Census of Population and Housing. Summary File-3C. U.S. Department of Commerce, 1990.