Landscape and Urban Planning 92 (2009) 47–52

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Spatial extent and habitat context influence the nature and strength of relationships between urbanization measures

Brian G. Tavernia*, J. Michael Reed

Biology Department, Tufts University, 163 Packard Avenue, Medford, MA 02155, USA

ARTICLE INFO

Article history:
Received 20 October 2008
Received in revised form 20 January 2009
Accepted 9 February 2009
Available online 14 March 2009

Keywords:
Landscape metrics
Roads
Road ecology
Suburban sprawl
Exurban
Wetland

ABSTRACT

A wide variety of metrics is used to quantify features of urbanization in ecological studies. Selecting statistically independent measures of urbanization depends upon the nature (multivariate collinearity) and strength (correlation coefficient) of correlations between urban metrics. We evaluated the influence of landscape extent and habitat context, factors that commonly differ between studies, on correlations between urban metrics. We examined the nature and strength of relationships between urban metrics at 1105 sites within Massachusetts, USA, including: population, agriculture cover, forest cover, wetland cover, dense residential cover, impervious surface cover, road length and greenspace cover. At each site, values were measured at five extents: 100 m, 250 m, 500 m, 1 km, and 2 km radii buffers. We also investigated the influence of habitat context on correlations by measuring values with a 1 km radius buffer for 100 sites within each of three habitat contexts (salt marsh, forest, and freshwater marsh). Principal component analysis showed that spatial scale did not affect the nature of relationships, but habitat context did. The average strength of bivariate correlations significantly increased at larger extents, and was significantly lower in salt marsh habitat context. Our results indicate that no single set of urbanization metrics is universally applicable and underline the importance of using a suite of statistical techniques to characterize independent aspects of urban environments.

1. Introduction

Urbanization can be described as an increase in human population density coupled with increased energy use and extensive alteration of the landscape (McDonnell and Pickett, 1990). The variety of measures used to quantify urban environments can be divided into three categories: demographic (e.g. population density), physical (e.g. road density), and landscape (e.g. mean patch size) (Hahs and McDonnell, 2006). Studies quantifying the effects of urbanization on species and ecological processes differ in urbanization measures selected (e.g. McIntyre, 2000; Marzluff et al., 2001; Theobald, 2004). If urbanization metrics are poorly or inconsistently correlated multiple measures would be required to sufficiently explore the relationships between urbanization and biodiversity, and integrating results across studies of ecological impacts of urbanization could be difficult. Landscape ecologists have used multivariate statistics to group landscape metrics into statistically independent subsets (e.g. Riitters et al., 1995; Cain et al., 1997); researchers have used these groups to describe the impact of urbanization on landscape structure and composition and have tried to link these patterns to socioeconomic and ecological processes (Luck and Wu, 2002; Seto and Fragkias, 2005; Yu and Ng, 2007). Recently, Hahs and McDonnell (2006) used principal components analysis (PCA) to identify independent groups of urban measures. From each group, they selected the single variable that had the highest principal component loading, considering each to be representative of their respective group, and used these measures to describe the urban environment of Melbourne, Australia. However, they cautioned that the applicability of these results to other landscapes may be limited and that investigations of factors influencing the relationships between urbanization measures were needed.

Different methods of aggregating spatial data are known to influence the output and correlations between landscape metrics (Turner et al., 1989; Jelinski and Wu, 1996; Cain et al., 1997; Wu et al., 2002). A particular set of potential problems arising from these observations is known as the Ecological Fallacy Problem (EFP) (Robinson, 1950; Openshaw, 1984; Cao and Lam, 1997). The EFP encompass three issues that can lead to an inability to draw valid inferences across scales and study regions: individualistic, ecological, and cross-level (hereafter, cross-region) fallacies. Individualistic fallacies arise from the extrapolation of patterns from smaller to larger spatial scales, while ecological fallacies are the inverse, interpolating to smaller spatial scales from larger ones. An example of an interpolation error is assuming that human population density,

* Corresponding author. Tel.: +1 617 627 5346; fax: +1 617 627 3805.
E-mail addresses: brian.tavernia@tufts.edu, btavernia@gmail.com (B.G. Tavernia), michael.reed@tufts.edu (J.M. Reed).

0169-2046/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.landurbplan.2009.02.003

which can correlate positively with endangered species richness at large spatial scales (Dobson et al., 2001), will show the same pattern at smaller spatial scales. This is a fallacy because anthropogenic drivers, including urbanization, are causing species loss at small spatial scales (McKinney, 2002). The cross-region fallacy states that relationships developed from one study region may not necessarily apply to different regions. For example, Cain et al. (1997) found that correlations between landscape metrics can differ between geographic regions.

Our overall aim was to examine the effect of changing spatial extent and habitat context on correlations between urbanization measures. Studies of the influence of landscape features on ecological processes frequently assess landscape features in a buffer around a point or site (e.g. Hagan and Meehan, 2002) or in a single cell of a predefined grid (e.g. Naugle et al., 2001); the size of these buffers or cells is the spatial extent of individual landscapes (Turner et al., 1989). Efforts to assess the impacts of urbanization often range greatly in the extent considered (e.g. White and Greer, 2006; Clark et al., 2007). If correlations between urbanization measures change with scale as suggested by the previously outlined fallacies, studies conducted at different extents may need to use different suites of metrics to characterize the urban environment. Alberti et al. (2001) examined bivariate correlations between land-cover and land-use pattern metrics at two extents and found stronger average correlations at the larger extent (5 km^2 vs. 1 km^2 cells). We extend their work by looking at a larger range of spatial extents and using multivariate statistics to identify independent subsets of predictor variables at each extent.

A second potential cause of inconsistency in correlations between urbanization measures is the effect of habitat context on urban development. Studies of the ecological implications of urbanization have been conducted in multiple habitat types: deserts (Hostetler and McIntyre, 2001), grasslands (Collinge et al., 2003), and salt marshes (DeLuca et al., 2004). Following from the cross-region fallacy and given that previous studies have found covariation between metrics depends on landscape context (Cain et al., 1997), researchers working in different habitat contexts may need different sets of urbanization metrics to characterize landscapes.

The specific goals of our study were to evaluate (1) how the nature and strength of correlations between urbanization measures changed with spatial scale and (2) the effects of habitat context on the nature and strength of these relationships. We distinguish “nature” (multivariate collinearity) and “strength” (bivariate correlation coefficient) (cf. Openshaw, 1984) of correlations to evaluate two related questions: (1) What subsets of interrelated variables are identifiable at different extents and in different habitats (nature)? and (2) Are variables within these subsets correlated to a sufficient degree to be used interchangeably (strength)? We conducted this research in Massachusetts (U.S.A.), a state whose eastern portion is dominated by the Boston metropolitan area. Urbanization is an ongoing process in this region, as evidenced by a 21.7% increase in the area developed from 1982 to 1992 (Noss and Peters, 1995; see also Porter and Hill, 1998).

2. Materials and methods

2.1. Data collection

We generated a regular grid of points across Massachusetts, extending from 41°21′ to 42°51′ N and from 70°1′ to 73°28′ W. Grid points were spaced every 4 km using Hawth’s tools, a downloadable extension for GIS ArcMap 9.2 (www.spatialecology.com). For each point, we created circular buffers at five extents: 100 m, 250 m, 500 m, 1000 m, and 2000 m-radii; because of the grid spacing, buffers from adjacent points did not overlap. We kept for analysis only those points for which all buffers were completely contained within Massachusetts ($n = 1105$). Within each buffer, we determined the value of eight urbanization measures: (1) population density (number of people/ha), (2) agriculture cover (ha; area of cropland and pasture land), (3) forest cover (ha), (4) wetland cover (ha; area of freshwater wetlands), (5) dense residential cover (ha; total area of high density residences, including multifamily and ≤0.2 ha plots), (6) impervious surface cover (ha; area of paved roads plus commercial and industrial cover), (7) road length (km; total length of secondary, primary, and highway roads), and (8) greenspace cover (ha; combined areas of open land, cropland, urban open land, pasture, forest, and woody perennial). These variables cover the spectrum of urbanization measures outlined by Hahs and McDonnell (2006): demographic (population density), physical (road density), and landscape metrics (coverage of various land use types).

We determined the value for each metric at each extent using GIS data layers provided by the Massachusetts Executive Office of Energy and Environmental Affairs (EOEEA) (http://www.mass.gov/mgis/). The census information used to calculate population density was derived by the EOEEA from TIGER data sets made available by the U.S. Census Bureau. This information is available at a coarse scale in census blocks. To determine the population density of a buffer, we first calculated the population density of each census block and the area of the block that intersected the buffer. We then multiplied the intersected area of each census block by the population density to determine the number of people contributed by the census block to the buffer. Finally, the number of people was summed across census blocks and divided by buffer area. A land use layer provided by the EOEEA was used to determine the area of the land use types; this layer is based on photo-interpretation of 1:25,000 aerial photographs. Road length was determined using a GIS layer representing the major and minor roads present in Massachusetts, which is maintained jointly by the EOEEA and the Executive Office of Transportation. Road surface area, used in calculating impervious surface area, was determined by multiplying road length by average width: local and secondary roads 6.7 m wide, primary roads 11.6 m, and highway 23.8 m (Massachusetts Highway Design Manual, 1997).

The influence of habitat context on correlations between urbanization metrics was determined by isolating polygons of salt marsh, forest, and freshwater wetland areas from the land use layer. Using Hawth’s tools, we randomly generated 100 points within each habitat type, separated by at least 2 km. For the forest and freshwater wetland categories, we retained only those points greater than 1 km from the state border, generated 1 km-radius buffers, and calculated the urban measures listed above. Given the natural proximity to coastline, this procedure was not possible for salt marsh points. For this habitat context, we generated a 1 km-radius buffer around salt marsh points and determined urbanization measures on a per land area basis (e.g. percent greenspace, road density).

2.2. Data analysis

We analyzed data using SAS v 9.1.3 (SAS Institute, 2004). First, we evaluated the multivariate nature of the relationships among urbanization measures using PCA (Princomp procedure). This was done to identify subsets of related urbanization metrics with each subset describing a different aspect of the surrounding landscape. For this analysis we were interested in the number of significant principal components (eigenvalue ≥ 1.0), and in the variables that contributed significantly to each principal component (|eigenvector| ≥ 0.3) (Hahs and McDonnell, 2006). A significant component revealed groups of measures that were correlated in