Remote Sensing of Environment 115 (2011) 269–280

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Spatial scaling of stable night lights

Christopher Small [a,*], Christopher D. Elvidge [b], Deborah Balk [c], Mark Montgomery [d]

[a] Lamont Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA
[b] Earth Observation Group, NOAA National Geophysical Data Center, Boulder, CO 80303, USA
[c] Baruch College, City University of New York & CUNY Inst. for Demographic Research, NY, NY 10010, USA
[d] Department of Economics, Stony Brook University, Stony Brook, NY 11794, USA

ARTICLE INFO

Article history:
Received 5 May 2010
Received in revised form 10 August 2010
Accepted 28 August 2010

Keywords:
Urban
Land cover
Night light
Landsat
Zipf
Power law
Scaling

ABSTRACT

City size distributions, defined on the basis of population, are often described by power laws. Zipf's Law states that the exponent of the power law for rank-size distributions of cities is near −1. Verification of power law scaling for city size distributions at continental and global scales is complicated by small sample sizes, inappropriate estimation techniques, inconsistent definitions of urban extent and variations in the accuracy and spatial resolution of census administrative units. We attempt to circumvent some of these complications by using a continuous spatial proxy for anthropogenic development and treat it as a spatial complement to population distribution. We quantify the linearity and exponent of the rank-size distribution of spatially contiguous patches of stable night light over a range of brightnesses corresponding to different intensities of development. Temporally stable night lights, as measured by the Defense Meteorological Satellite Program-Operational Line Scanner (DMSP-OLS), provide a unique proxy for anthropogenic development. Brightness and spatial extent of emitted light are correlated to population density (Sutton et al., 2001), built area density (Elvidge et al., 2007c) and economic activity (Doll et al., 2006; Henderson et al., 2009) at global scales and within specific countries. Using a variable brightness threshold to derive spatial extent of developed land area eliminates the complication of administrative definitions of urban extent and makes it possible to test Zipf's Law in the spatial dimension for a wide range of anthropogenic development. Higher brightness thresholds generally correspond to more intense development while lower thresholds extend the lighted area to include smaller settlements and less intensively developed peri-urban and agricultural areas. Using both Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) to estimate power law linearity and exponent of the resulting rank-size distributions across a range of upper tail cutoffs, we consistently find statistically significant exponents in the range −0.95 to −1.11 with an abrupt transition to very large, extensively connected, spatial networks of development near the low light detection limit of the sensor. This range of exponents and transition are observed at both continental and global scales. The results suggest that Zipf's Law also holds for spatial extent of anthropogenic development across a range of intensities at both continental and global scales. The implication is that the dynamics of urban growth and development may be represented as spatial phase transitions when the spatial extent and intensity of development are treated as continuous variables rather than discrete entities.

1. Introduction

Understanding of urban growth processes is challenged by the large number of variables known to influence these processes. Despite the need for large sample sizes, most comparative studies of urban growth use only a subset of the world's cities and many studies focus on a single city or small number of cities for specific time intervals. Almost all of these studies rely on administratively defined clusters of population. Understanding of the processes is further complicated by the lack of any widely accepted and consistent definition of what constitutes a city or urban area (United Nations, 2008). In spite of these challenges, some consistency has been observed in the size distribution of populations of cities. (Auerbach, 1913) observed that the size distribution for cities can often be described as a power law in which the number of cities with population greater than N is inversely proportional to N. Zipf (1949) noted that the exponent of the power law is also close to −1 for the frequency of usage of words and postulated a universal principal of least effort on the basis of these observations. The special case of a power law distribution with an exponent of −1 is generally referred to as Zipf's Law and has been tested repeatedly by economists, geographers, urbanists and physicists. Gabaix (1999) states "Zipf's law for cities is one of the most conspicuous empirical facts in economics—or in the social sciences generally. The importance of this law is that, given very strong

* Corresponding author. Tel.: +347 535 4274.
E-mail address: small@LDEO.columbia.edu (C. Small).

0034-4257/$ – see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.rse.2010.08.021

empirical support, it constitutes a minimum criterion of admissibility for any model of local growth or any model of cities".

Despite considerable research on Zipf's Law for cities, and numerous proposed explanations, there appears to be no consensus on the cause or even the generality of the law. The power law is used to describe only the upper tail of the distribution with varying criteria used for selecting a lower cutoff. While there have been challenges to the validity of Zipf's Law (e.g. Gan et al., 2006; Nitsch, 2005), numerous examples have been found for city size distributions within specific countries over the past 100 years in which the law appears to hold remarkably well (see Gabaix et al., 2004 and Pumain, 2004) for summaries). However, the exponent appears to vary from country to country (e.g. Pumain and Moriconi-Ebrard, 1997), and even within the same country over time (Guérin-Pace, 1995; Overman and Ioannides, 2001). To complicate matters further, the value of the exponent depends on the method of estimation (Rosen and Resnick, 1980; Soo, 2005). An underlying cause of many of these discrepancies is the difficulty of mapping human population. The purpose of this analysis is to explore a spatial complement to population distribution and to quantify some characteristics of the size distribution of human settlements at continental to global scales on the basis of a continuum of intensity of development (as defined below). This analysis tests the spatial manifestation of Zipf's law using an alternative measure of urbanness—nighttime lights. We use the concept of a continuum of anthropogenic development to avoid arbitrary definitions of urban land cover and to represent varying degrees of spatial connectivity among cities and other types of anthropogenic land cover.

2. Urban population and spatial extent

Much of the debate about the existence and generality of Zipf's Law seems to arise from two confounding factors related to how cities are defined and measured. The lack of a consistent definition of a city makes it difficult to compare size distributions between different countries (e.g. Pumain, 2004) because census enumerations depend on the administrative boundaries defined for each city and different countries have different, and often multiple, criteria for urban administrative boundaries.[1] The same factor confounds global analyses because inconsistency in definition precludes simple aggregation of tabular census data from different countries. A second confounding factor arises in the measurement of city size in terms of population.[2] Although administrative units are defined on the basis of spatial boundaries, the enumerations are implicitly aspatial because they discretize what are often spatially contiguous populated areas. Furthermore, cities can be defined according to different criteria—such as the population size, population density of functional criteria such as capitol cities or whether a given fraction of the labor force in engaged in non-agricultural activities. Remarkably, across time and place, the number of inhabitants has been a much more difficult criterion to measure consistently than one might imagine. As a result, there is no existing globally spatially-delineated population time-series for urban areas (Balk, 2009).

In fact, the idea of defining a city as a discrete entity with fixed boundaries may be fundamentally flawed because so many of the factors that characterize the form and function of human settlement patterns and dynamics are temporally variable over spatial gradients at some range of scales—such as, population density, housing density, road density, energy consumption, economic activity and output, emission and pollutant output. The fractal nature of urban development (Batty and Longley, 1994) combined with ubiquitous gradients in spatial density or intensity of development makes it difficult, and perhaps misleading, to define cities as discrete entities with fixed boundaries. Rather than considering the city as a single discrete entity defined in terms of inhabitants within an administrative boundary, we consider multiple spatial extents of anthropogenic land surface modification defined on the basis of varying degrees or intensities of development.

Human settlement patterns and, more generally, land surface modification can be represented as continuous spatial variation in intensity of development or degree of modification. A continuous depiction accommodates gradients while still allowing for sharp discontinuities observed in many cases (e.g. coastlines). A spatially and temporally continuous measure of the "degree of urbanness" provides a more general representation of many spatial and temporal processes while avoiding potential errors of omission and commission that inevitably result from discretization. Continuous variates also allow for the construction of different application-specific metrics from measurable quantities (e.g. those mentioned above) without the need to define a single discrete boundary separating urban from non-urban.

Numerous studies have attempted to elude the problem of urban definition by aggregating spatially adjacent population counts (see Pumain, 2006) but relatively few studies have attempted to investigate Zipf's Law in terms of area or spatial extent of anthropogenic land use—either within individual countries or aggregated to global scales. A recent study by (Rozenfeld et al., 2008) used spatially aggregated census data to demonstrate that urban agglomerations in the U.S.A. and Great Britain can be accurately described by Zipf's Law in terms of both population and area. Their results suggest that the exponent is sensitive to the scale length of the aggregation. At finer spatial scales (Fragkias and Seto, 2009) used satellite-derived urban extents of several cities in southern China to quantify temporal evolution of rank-size distributions revealing oscillations related to birth and spatial coalescence of settlements through time. Together these studies illustrate a spatial dimension to Zipf's Law that is often implied in studies using population data but rarely addressed explicitly.

As an alternative to discrete, population-based definitions of urban extent we consider a continuous, spatially explicit metric for anthropogenic land surface modification—analogous to population density—and investigate the effects of discretizing the metric at different levels or degrees of modification. Discretizing by imposing a threshold on the intensity of development results in a spatially segmented binary image with contiguous patches of developed area against an undeveloped (or less developed) background. Different thresholds result in different spatial distributions of developed area with different degrees of connectivity. Treating the size distribution of contiguous segments resulting from each threshold as a distinct depiction of anthropogenically modified or developed land area, we quantify the degree to which the spatial extent and distribution of human habitation and land surface modification can be described by a power law and how the parameters of the power law depend on the level of discretization.

3. Anthropogenic land use and night light

It is important to emphasize the difference between population counts and developed land area. Although night light brightness and population density are correlated at high levels, the relationship is less direct at the lowest brightness levels. In part, this is a result of overglow (discussed below) but also because of the variety of different sources that contribute to the lowest levels of detectable light. In this study we use satellite-derived observations of stable anthropogenic light as an indication of varying degrees of development

[1] National statistical offices not only vary in the census and administrative boundary that they collect, but also what they release, with the latter often being far more coarse than the former.

[2] In addition to variations in census accuracy and detail from country to country, most census enumerations quantify where people sleep rather than where they work and interact during the day.