Applied Geography 77 (2016) 1–7

Contents lists available at ScienceDirect

Applied Geography

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

Combining spatial data with survey data improves predictions of boundaries between settlements

Beccy Wilebore*, David Coomes

Forest Conservation and Ecology Group, Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK

A R T I C L E I N F O

Article history:
Received 8 October 2015
Received in revised form
1 September 2016
Accepted 11 September 2016
Available online 10 October 2016

Keywords:
Circular buffers
GIS
Land-use
Thiessen polygons
Voronoi polygons
West Africa

A B S T R A C T

Studies of land-use change often require the combination of socioeconomic survey data with spatially continuous maps of land-cover change. One approach is to define maps of land ownership, assuming that all land-use change can be attributed to the owners or managers of each parcel of land. Unfortunately, records of administrative boundaries between towns and villages are commonly unavailable in developing countries and prohibitively costly or time consuming to map for individual projects. However, point locations of the settlements themselves can be obtained easily from existing maps or remotely-sensed imagery. In this paper we compare three methods – circular buffers, unweighted Voronoi polygons (sometimes referred to as Thiessen polygons) and multiplicatively weighted Voronoi polygons – for estimating boundaries between villages in an agricultural landscape in West Africa. The benefits and limitations of each approach are discussed, and their accuracy assessed using 98 independently collected GPS coordinates of village boundaries. We present a novel method for generating and optimising weights for multiplicatively weighted Voronoi polygons using survey data of village sizes from a subset of villages. By using both spatial information and survey data from villages, we show that multiplicatively weighted Voronoi polygons outperform other methods of predicting village boundaries, and increase the correlation coefficient between surveyed village area and mapped areas from 0.18 to 0.68 compared with more commonly used unweighted Voronoi polygons. Our method of weighting Voronoi polygons can be implemented with data and software commonly available to researchers and non-governmental organisations.

1. Introduction

The study of human-driven land-cover changes often requires merging spatially continuous, or pixel-level, data of land-cover with spatially discontinuous data from socioeconomic surveys (Liverman & Cuesta, 2008; Skole, 1994; Walsh & Welsh, 2003). Household-level or village-level surveys can provide detailed information about people living within a single location. On the other hand, satellite imagery and other remotely sensed products can produce spatially continuous maps of land-cover change through time, over large areas, and often at high spatial resolution. However, understanding the influence of people on the landscape by linking social data with satellite imagery is a key and persistent challenge for land-change science (Liverman & Cuesta, 2008). Whilst some types of survey data, such as population estimates, can be interpolated between points to produce information of population density at the pixel level (Balk et al., 2006; Herold, Scepan, & Clarke, 2002), others are aggregated within administrative boundaries to produce panel data, or are categorical in nature such that interpolation of the values is not meaningful (Walsh & Welsh, 2003). Where the scale of analysis or policy intervention is at the village or settlement level, an understanding of administrative boundaries between settlements is of interest to link measured socioeconomic variables to observed land-cover changes and studying the drivers of land-use change (Irwin & Geoghegan, 2001; Robinson, Holland, & Naughton-Treves, 2014). In developed countries, land tenure is often well-established and administrative boundaries between farms or towns are generally accurately mapped and publicly available. In developing countries, however, the data on private land holdings and fine-scale administrative boundaries, such as between towns and villages, are commonly unmapped (Blackman, 2013). Even where land tenure is well-

* Corresponding author. Current address: Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK.

E-mail address: beccy.wilebore@zoo.ox.ac.uk (B. Wilebore).

http://dx.doi.org/10.1016/j.apgeog.2016.09.007
0143-6228/

understood locally it can be prohibitively costly or infeasible to map these administrative boundaries for large study areas (Liverman, Moran, Rindfuss, & Stern, 1998). But, unlike administrative boundaries that might have no physical basis, the locations of the settlements themselves are often available even for remote regions, since they can be mapped quickly using hand-held GPS units or remotely with high resolution aerial imagery from satellites or aircraft (Herold et al., 2002; Lu, Tian, Zhou, & Ge, 2008). In the absence of better records, these point-locations can be used to predict the administrative boundaries between settlements.

Looking beyond land change science to other disciplines studying spatial processes, two approaches for dividing landscapes between point data have been favoured to date. The first and simplest uses only data about each point to generate a circular buffer around them, ignoring the spatial arrangement and properties of other points in the landscape (Pigot, Owens, & Orme, 2010). The size of the buffer is determined from assumptions about the influence of the generator point (in the case of settlements for example, villages might be assumed to manage land in a radius that is proportional to their population). This approach has been used in studies across numerous academic fields, including to examine access to retailers (Duncan et al., 2014), in neutral models of species ranges (Baselga, Lobo, Svenning, & Araújo, 2012; Pigot et al., 2010) and species occupancy models (Efford & Dawson, 2012). The benefits of circular buffers are in their simplicity; they require only the location of each point and some information about range. As new data are collected, they may be easily added to the study landscape, as each point is treated as independent of the others. However, a strong drawback of circular buffers is that they result in ranges that may overlap, and they have uniform shapes. Whilst this may be an acceptable for some spatial phenomenon, it is typically an unrealistic result when mapping land ownership.

The second approach, “Thiessen” or unweighted Voronoi polygons, does not account for differences in the influence of each generator point and instead uses only their spatial configuration to divide the landscape through tessellations (Okabe, Boots, Sugihara, Chiu, & Kendall, 2008). A tessellation is the complete division of a defined portion of Cartesian plane into non-overlapping regions given a finite number of generator points (Okabe, Satoh, Furuta, Suzuki, & Okano, 2008). Unweighted Voronoi polygons are constructed such that each polygon contains only one generator point, and any area within a polygon is closer to its associated point than any other. They therefore assume that, for any given location, the most likely land manager is from in the nearest village (measured by Euclidean distance to the village settlement). Unlike circular buffers, they require no information about the points other than their spatial configuration. Unweighted Voronoi polygons have been used to map bird territories (Schlicht, Valcu, & Kempenaers, 2014), determine movement of organisms across landscapes (Galpern & Manseau, 2013), map vegetation (Gooding, Rackham, Holland, & Robertson, 1997), create neighbourhoods for housing market studies (Kryvobokov, 2013) and for land-use models (Muller & Zeller, 2002). However, unweighted Voronoi polygons assume that all generator points have the same influence in the landscape. This may not be the case if, for example, settlements with larger populations are expected to have more land than those with small populations, regardless of their location. Unweighted Voronoi polygons also require a complete dataset of all villages within the bounds of the study area; any missing villages will lead to an overestimation in the size of neighbouring polygons and introduce error to the results.

This paper presents a third approach, weighted Voronoi polygons, and compares its effectiveness against circular buffers and unweighted Voronoi polygons as a method for estimating village boundaries in a region of West Africa. Weighted Voronoi polygons combine both the spatial configuration of generator points with additional information about the expected size or influence of each point. Starting with an unweighted Voronoi tessellation, polygon sizes are adjusted by multiplying the distance measures between points by a weighting term for each point (Okabe, Satoh et al., 2008). Multiplicatively weighted Voronoi polygons have been used in some research fields including in ecology to predict species growth and competition (Aakala, Fraver, D'Amato, & Palik, 2013; Du et al., 2012), generate theoretical forest landscapes (Li, Bettinger, & Weiskittel, 2010; Passolt, Fix, & Tóth, 2013) and study territory sizes and shapes (Adams, 2001; Wilkin, Garant, Gosler, & Sheldon, 2006). They have also been used in studies of urban hierarchies in the US (Mu & Wang, 2006), in logistics (Galvão, Novaes, Souza De Cursi, & Souza, 2006) and in measuring consumer preferences (Boots & South, 1997). However, to our knowledge there are no other studies that use multiplicatively weighted Voronoi polygons to determine administrative boundaries between settlements. Here we present a novel method for generating the village-level weightings using incomplete village-level survey data. Methods for generating the weighted Voronoi polygons were adapted from Dong (2008).

2. Methods

2.1. Study area

Our study area is 286,200 ha of community land outside of the Gola Rainforest National Park (GRNP) in Sierra Leone, West Africa (Fig. 1). The landscape is managed predominantly for subsistence slash-and-burn agriculture and, with the exception of isolated private estates and the National Park itself, land is almost exclusively owned and managed by 476 villages within the seven chiefdoms around the National Park. Despite the region recovering from a civil war ending in 2002, administrative boundaries between villages are well-defined by local institutions. However, the demarcations have not been officially mapped or registered with the government. For this study we assume that village lands are contiguous and contain the village that owns them, and that all land is owned by only one village with no shared land. Using local

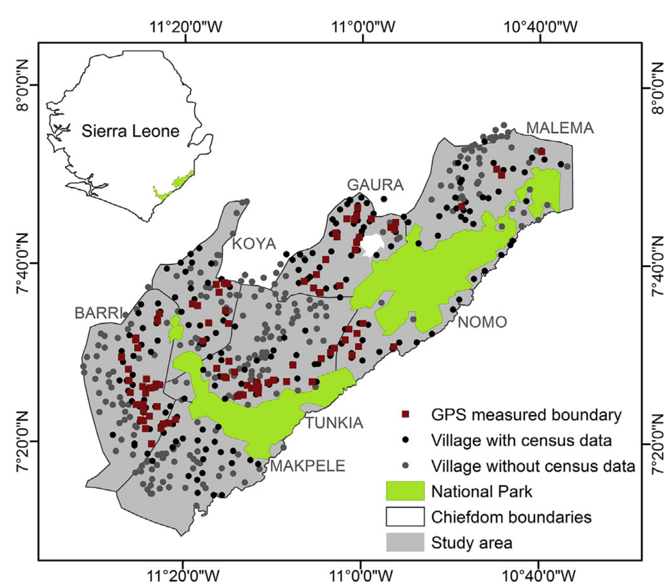

Fig. 1. Study site in Sierra Leone, West Africa. Villages surround the Gola Rainforest National Park, across seven chiefdoms. The distribution of GPS coordinates at known boundaries between villages are shown.