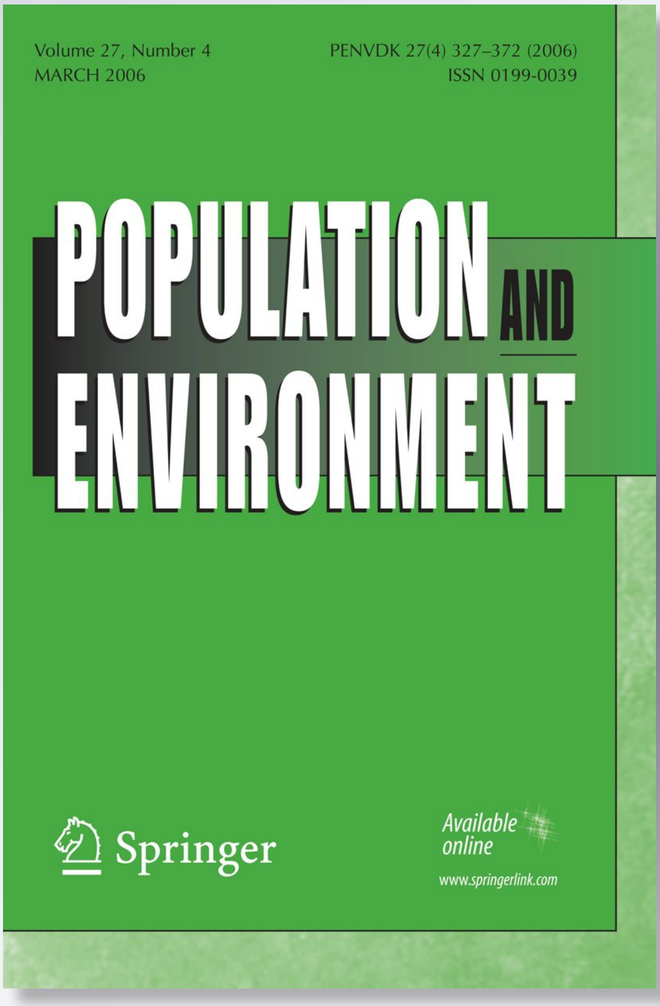
Using remote sensing and census tract data to improve representation of population spatial distribution: case studies in the Brazilian Amazon

Silvana Amaral, André Augusto Gavlak, Maria Isabel Sobral Escada & Antônio Miguel Vieira Monteiro

Population and Environment
A Journal of Interdisciplinary Studies

ISSN 0199-0039

Popul Environ
DOI 10.1007/s11111-012-0168-2



Springer

Popul Environ
DOI 10.1007/s11111-012-0168-2

ORIGINAL PAPER

Using remote sensing and census tract data to improve representation of population spatial distribution: case studies in the Brazilian Amazon

Silvana Amaral · André Augusto Gavlak · Maria Isabel Sobral Escada · Antônio Miguel Vieira Monteiro

Abstract This work proposes a methodological approach to redistribute population data obtained from polygonal census tracts into population density surfaces (grids) based on a cell space database. The methodology was first developed for the municipality of Marabá, Pará state, in the Brazilian Amazon. We used a dasymetric method to eliminate areas of environmental restriction to human presence; then integrated environmental data indicative of human presence to generate a potential surface of population occurrence; and finally, census population count data were redistributed into cells. The methodology was subsequently adapted for 13 municipalities of the Sustainable Forests District (SFD) of BR-163, generating population distribution surfaces for 2000 and 2007. The evolution of the resident population over the SFD-BR163 showed spatial patterns compatible with the occupation process described in the literature and verified by fieldwork. To be applied over other areas, the proposed methodology must be adapted with local parameters but in this way, population density surfaces can be useful as an additional data source to study population and environment relationships.

We are pleased to have this work included in the memorial issue to Professor Daniel Hogan. INPE, the Brazilian National Institute for Space Research had been running a Remote Sensing and Geoinformation graduate program since 1972. It was the year of 1999 when for the first time the late Professor Daniel Hogan approached our research group at INPE. In particular, INPE had been heading national environmental monitoring programs and projects on GIS and Satellite Image Processing open technologies. Professor Daniel, as we liked to call him, came to know INPE's headquarters at São José dos Campos, São Paulo, and talk about the possibilities of engaging ourselves in a cross-disciplinary conversation with demographers. As a start point, he right away invited us as institutional speakers for a scientific session he would organize for the ABEP 2000 meeting. In September of 2000, the MR 6 session—sustainability indicators was held and headed by him and we were there giving a talk on: *computational and mathematical challenges involved in the production and representation of spatially aware sustainability indicators: patterns, process, and territory* for a very mixed audience.

S. Amaral (✉) · A. A. Gavlak · M. I. S. Escada · A. M. V. Monteiro
Brazilian National Institute for Space Research (INPE), P.O. Box 515,
São José dos Campos, SP 12227-010, Brazil
e-mail: silvana@dpi.inpe.br

Published online: 31 March 2012

Keywords Remote sensing · Population distribution · Population density surface · Amazon · Fuzzy inference · Sustainable Forest District of BR163 · Spatial analyses · GIS

Introduction

The pressure of a growing population on natural resources has been a central issue in demographic and environmental studies (Ehrlich 1968; McMichael 1993). Instead of discussing demographic patterns and processes as a consequence of environmental conditions, human presence, and activities are typically viewed as responsible for environmental change (Hogan 1989). Recently, the predominant approach to population–environment-development research has become more moderate; demographic pressure is no longer viewed as the principal determinant of environmental problems, but instead as a factor interacting with social, economic, and political realities (Hogan 2000, 2001). Hogan (2007) emphasized the importance of including environmental issues in population research, further stating that the relationship is reciprocal. The work presented here has been motivated by this call to better integrate demographic and environmental factors within environmental social science through advancement in integration of remotely sensed data.

Remote sensing has contributed significantly to studies integrating human dimensions (demographic and social data) and biophysical parameters in the Amazon region (Frohn et al. 1996; Wood and Skole 1998; Liverman et al. 1988; Walsh 2010). In particular, analyses of changes in land use and land cover have enabled the integration of remote sensing and socio-demographic research, combining image-processing techniques with social science analytical methods (McCracken et al. 2002).

For land cover and landscape research, the unit of observation is the pixel, while the resolution, scale, and information available from spectral bands are functions of the remote sensors selected. On the other hand, social science data typically come from field surveys and/or official censuses, in which the observation unit defines the research subject (e.g., events, individuals, households, social groups, and

Every since we have been in all ABEP meetings and have developed a close link with the field of demography, in particular with the studies dealing with population, space, and environment. Also, at the time of his first visit, Silvana Amaral, this paper's first author, was engaged in her Ph.D. research at the São Paulo State University (USP), starting working with a night light satellite sensor. The meeting with Professor Daniel opened a research avenue of possibilities in exploring remote sensing integrated with GIS and spatial analytical tools for advancing new methods and methodologies for mapping and explore population spatial distribution models and assess its patterns. Her thesis: Geoinformation for Demographic Studies: Spatial Representation of Population Data in the Brazilian Amazônia (Geoinformação para Estudos Demográficos: Representação Espacial de Dados de População na Amazônia Brasileira), represented the first time that someone from INPE's staff would take a subject linked to the population studies. When she finished, in 2003, Professor Daniel was there as member of her Ph.D. committee. This paper actually is a fine proof of the intense and permanent influence that Professor Daniel has impressed in our personal and institutional lives. Since our very first meeting with Professor Daniel, the population studies has gain its own roots at INPE and in its institutional projects and its graduate programs. INPE has been involved with population studies, in particular exploring new possibilities and methodologies that deal with remote sensing data and spatial–temporal analysis that can be of use for enhancing our understanding of the complex relationships that hold the population and environment research agenda.

communities). Therefore, when choosing the geographical scale for an integrated analysis, one should consider the following: who are the social actors of interest, and what is the spatial dimension in which these actors should be considered? In addition, studies that aim to capture the implications of land use change in agricultural frontier areas should also focus on changes in the demographic composition, a comparison that requires compatible scales. These themes assume particular relevance in the Amazon region due to the close relationships among urban development dynamics, land concentration processes, and rural depopulation.

The historical process of colonization in the Brazilian Amazon as related to demographic changes and patterns of change in land use and cover has been examined primarily based on household units (Moran et al. 1994, 2003; Moran and Brondizio 1998; McCracken et al. 1999). Cohort, age, and period effects have been analyzed to interpret landscape changes, mainly deforestation rates and secondary succession. Despite their undeniable contribution to the literature, such detailed studies describe local processes; yet because of the regional heterogeneity of the Amazon, are unsuitable for generalization.

For human population studies on analytical scales larger than the family unit, population data from official censuses are often used. In Brazil, the information for decennial censuses and population counts are collected by taking the residence as the sampling unit (IBGE 2010); however, population data have also been published using the census tract[1] as the spatial unit (i.e., the information about inhabitants collected by residence is spatially aggregated in census tracts). Still, the physical borders of Amazon municipalities have changed over the past decades, leading to changes in the borders of the census tracts. Thus, comparisons between spatial analysis and census/counting population data are not straightforward. Moreover, as census tracts spatially delimit geographical areas, they can be represented as polygons in a planar subdivision (layers) in Geographical Information Systems. This representation can be easily superposed over remote sensing images for complementary visualization of physical environment and population distribution. As census tract areas contain aggregate information, one cannot easily attribute population data to an image pixel, especially for such heterogeneous areas as rural census tracts in the Amazon.

As human activities alter the landscape, remote sensing and ancillary geographical data can be used to analytically redistribute population inside a census tract (Gallego and Peedell 2001; Linard et al. 2011). This work aims to contribute to that effort. Here, we present a methodological approach to redistributing census tract population data in a cellular space on a geographical database. Cell spaces enable researchers to represent population density in an intermediary spatial unit between pixels and census tract polygons.

First, to create a population density surface from census tract population counts, we developed a methodology for Marabá, a municipality in the state of Pará, Brazil.

[1] Census tract is the territorial unit for census operations, defined by IBGE (Instituto Brasileiro de Geografia e Estatística), with physical limits identified in contiguous areas and respecting the political and administrative division of Brazil.

Then, we presented population density surfaces generated for the BR-163 Sustainable Forest District (SFD-BR163) for 2000 and 2007. This political division comprises 13 municipalities in the west of Pará, where human activities and environmental changes demand continuous studies and monitoring. We discuss the population density surfaces in comparison to fieldwork observations, as well as the evolution of the population distribution in SFD-BR163 based on population density surfaces obtained for 2000 and 2007.

Census data representations

Demographic data collected in decennial censuses are usually modeled as statistical surfaces (DeMers 1999) and commonly represented using choropleth maps (Harvey 2008). The main disadvantage of using choropleth maps is that data aggregated by census tracts assume that the population is distributed homogeneously throughout the unit, which is never the case (Tobler 1979). Choropleth models and some other surface interpolation approaches result in allocation of a non-zero population density value to every location. A way to improve the spatial detail of choropleth-based population maps is to use more detailed maps representing the distribution of human-built objects and activities (Bajat et al. 2011).

One prominent interpolation method for population data is dasymetric mapping, defined generally as the use of an ancillary data set to disaggregate coarse resolution population data to a finer resolution (Eicher and Brewer 2001). Basically, the dasymetric method aims to use any available spatial information that can provide further insight into the probable structure of source zones and thus can be informative for redistributing population counts (Langford 2003).

The increasing availability of remotely sensed imagery has driven much recent research on population interpolation using the dasymetric mapping method. Recent research suggests that dasymetric mapping can offer more accurate population estimates than many areal interpolation techniques that do not use ancillary data (Mrozinski and Cromley 1999; Gregory 2002; Mennis 2003). Other researchers also find that dasymetric mapping gives the best estimated result among all other popular methods tested (Fisher and Langford 1995, Mrozinski and Cromley 1999; Mennis and Hultgren 2006).

Classified remotely sensed imagery has been commonly used as the source for ancillary data. The land cover map is another useful GIS layer; it is crucial for the disaggregation of population data (Gallego and Peedell 2001; Linard et al. 2011). Dobson et al. (2000) say that the land cover map may be the best single indicator of population density. Dasymetric modeling methods based on land use data require the definition of relative weights associated with land use classes (Hay et al. 2005). These weights are first calculated for regions where high-resolution census data are available and then applied to other similar regions (Gallego 2010). Langford (2007) uses cartographic materials over multi-spectral satellite imagery for dasymetric-based population interpolation. Reibel and Bufalino (2005) use road networks as the ancillary predictor to downscale demographic distribution.

Gridded population distribution data are increasingly being used for resource allocation, disease burden estimation, and climate change impact assessment (among other applications) at global, national, and local scales (Linard et al. 2011). Detailed and spatially disaggregated population data are essential resources in assessing the number of impacted people when making decisions related to developmental or health issues (Bhaduri et al. 2002; Dobson et al. 2000; Hay et al. 2005). Furthermore, gridded population distribution data have applications in analyzing the impacts of climate change (McGranahan et al. 2007; Nicholls et al. 2005). The vulnerability of people to natural disasters has also been quantified (Balk et al. 2005; Maynard-Ford et al. 2008).

The methodology proposed in this paper, which is based on work by Amaral (2003), presents a refined model of spatial distribution of population that is specific to the Amazonian region. The model considers how spatial variables influence the spatial distribution of population and how environmental factors may exclude settlement. Another positive aspect is the use of cells to represent and aggregate data, which allows temporal analysis independent of possible changes in political and administrative boundaries (dismemberment of municipalities, for example) and enables integration with other demographic, social, and environment data.

Study area

We first developed this methodology for Marabá, a single municipality in the state of Pará. We then considered regional features and adapted this methodology to apply it over a broader area. This larger area is also a geopolitical unit; SFD-BR163 includes 13 municipalities in Pará (PA) state (Fig. 1a).

Marabá occupies 15,111.26 km^2 and is located in the southeast of PA state. It is a regional capital whose urban center is on the confluence of the Tocantins and Itacaiunas rivers and the PA-150 and Transamazônica roads. Marabá experienced intense migratory flux from the state of Maranhão in the 60s and from southeast states in the 70s (De Reynal et al. 1995). Population mobility has slowed recently and has become essentially rural-to-urban or rural-to-rural migration (Oliveira et al. 2001).

The Sustainable Forest District of the BR-163 (SFD-BR163), along the federal road linking Cuiabá (Mato Grosso state) to Santarém, in western PA state (Fig. 1b), was created in 2006 as the first SFD established in Brazil. An SFD is a geo-economic and social complex to promote integrated local development based on forestry activities. Public policies from different government sectors have been proposed to promote forestry activity on a sustainable basis; such policies include land policy, infrastructure, industrial development, public areas management, technical assistance, and education (MMA 2006).

The SFD-BR163 is 190,000 km^2 in size and is comprised of the municipalities of Altamira, Aveiro, Belterra, Itaituba, Jacareacanga, Juruti, Novo Progresso, Óbidos, Placas, Prainha, Rurópolis, Santarém and Trairão, among which only Trairão, Rurópolis, and Belterra are completely encompassed by its boundaries. This SFD includes a wide variety of environments and occupations. Some regions that have been occupied for more than 300 years (Coudreau 1977), while others are still in the process of consolidation or agricultural frontier expansion. The municipality of

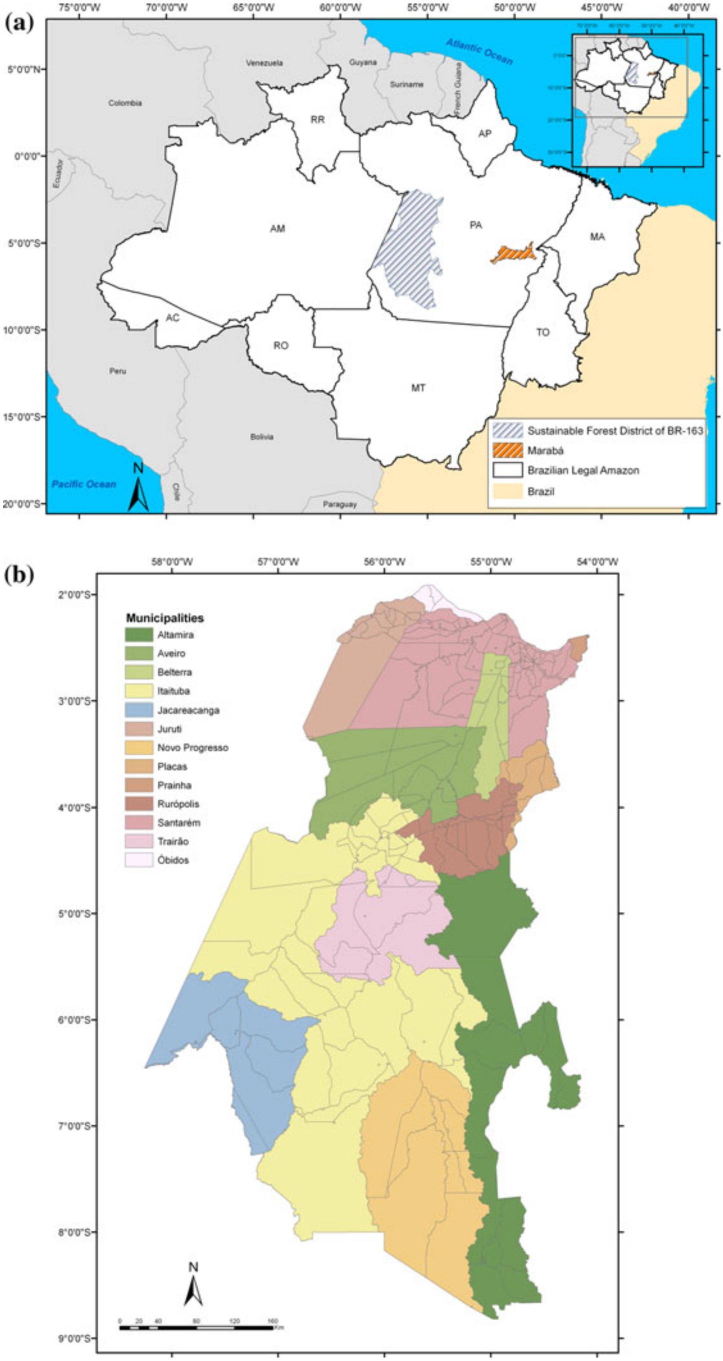


Fig. 1 Study sites: **a** the municipality of Marabá and the Sustainable Forest District/BR-163 in PA state, **b** municipalities and census tracts of the SFD-BR163. *Source* MMA (2006)

Novo Progresso is one of the latter and has shown high rates of deforestation. The proportion of deforested areas in this municipality rose from 4% (1,691 km^2) in 2000 to 14% (5,264 km^2) in 2009 (INPE 2009); in Belterra, close to Santarém municipality, this proportion increased from 15% (671 km^2) to 18% (797 km^2) over the same period.

In recent decades, the population of the municipalities in SFD-BR163 has significantly increased. At the same time, there has been a process of dismemberment and the creation of new municipalities. Figure 1b shows the current political division of the municipalities that constitute the SFD-BR163 and its census tracts (IBGE 2010). In 1980, the SFD-BR163 region was composed of the municipalities of Altamira, Aveiro, Itaituba, Santarém, Prainha, Óbidos and Juruti. In the 1991 Census, IBGE registered Rurópolis as a new municipality, and in the 2000 census, the municipalities of Belterra, Jacareacanga, Novo Progresso, Placas, and Trairão were registered. Half of the municipalities in this SFD were created during the 1990s. In contrast to the overall trend, the population of Novo Progresso decreased in its population in that decade; in general, population growth and deforestation rates are directly proportional.

Marabá was first chosen as study site because it was one focal area of GEOMA Network (DOU 2004), which supported fieldwork for the first methodological portion of the study. With the creation of SFD-BR163 in PA state, the opportunity arose to study the influence of public policies in a geo-economic and social complex. We then adapted and applied the methodology to this wider study area as part of the Cenários (Luizão 2008) and LUA/IAM (Camara 2009) Projects, which also address the spatial representation of population density and its temporal evolution.

Methodology

A model to disaggregate population data inside the census tracts

In this work, population is represented by the resident population count provided by the Brazilian Institute for Geography and Statistics (IBGE—Instituto Brasileiro de Geografia e Estatística) census. Rather than representing the population distribution based on census tracts limits (polygons) that contain information about the entire population in these geographical areas, the population distribution will be represented by a continuous surface (cell space) in which the population value will be attributed to cells. We generated a population density surface for the municipality of Marabá using a dasymetric method and a model based on environmental data indicative of human presence. Adapting this methodology to regional features, two population density surfaces were created for each of the 13 municipalities of SFD-BR163, corresponding to the 2000 and 2007 populations.

Disaggregating population data from census tract (polygons) to cell space (surface) requires the construction of a subjacent surface with modeling describing the factors that determine population distribution (Goodchild et al. 1993). We

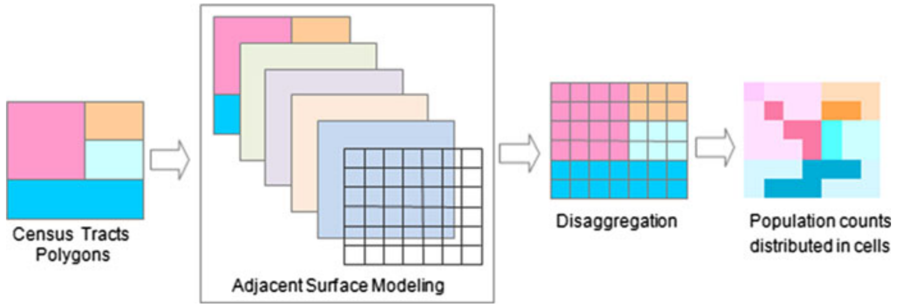


Fig. 2 General procedure to disaggregate population counts within census tracts. Adapted from Amaral (2003)

assumed that there are spatial variables related to the absence or presence of human settlement that could be used to indicate how the population is distributed (Fig. 2).

The method consists of three basic steps: (1) a dasymetric method (Mennis 2003; Sleeter 2004) to eliminate the areas of environmental exclusion of human presence (cells), using a map of land use cover classifications as a reference; (2) a multivariate interpolation method to generate a potential surface of settlement occurrence, using fuzzy inference over environmental data indicative of human presence (Zadeh 1988; Meirelles 1997); and (3) the redistribution of population count values to each cell, proportionate to a potential occurrence of population defined from (2).

For the Amazonian region, there are extensive areas of water and forest land cover where settlement is unlikely. We propose the use of ordinary digital classification of remote sensing images as thresholds settings to identify water bodies and forest land cover. The dasymetric method, applied next, consisted of removing cells where water bodies and/or forest occupied at least 95% of the population density surface in Marabá.

The multivariate interpolation method proposed to generate a potential surface of population occurrence can be summarized into five general steps (Fig. 3):

1. Selecting spatially explicit variables (environmental data indicative of human presence) related to population distribution (indicator variables);
2. Identifying the relationship between indicator variables and population distribution: this relation is quantified based on observed/previous population data and indicator variable frequency for the study site;
3. Creating a geographical database with indicator variable layers in cellular spaces: the study area is divided by regular cells containing a single value for each indicator variable;
4. Standardizing indicator variables based on fuzzy inference: the original value of every indicator variable in a cell is simplified considering fuzzy inference and are then represented in the interval of zero to one values, enabling continuous classification;

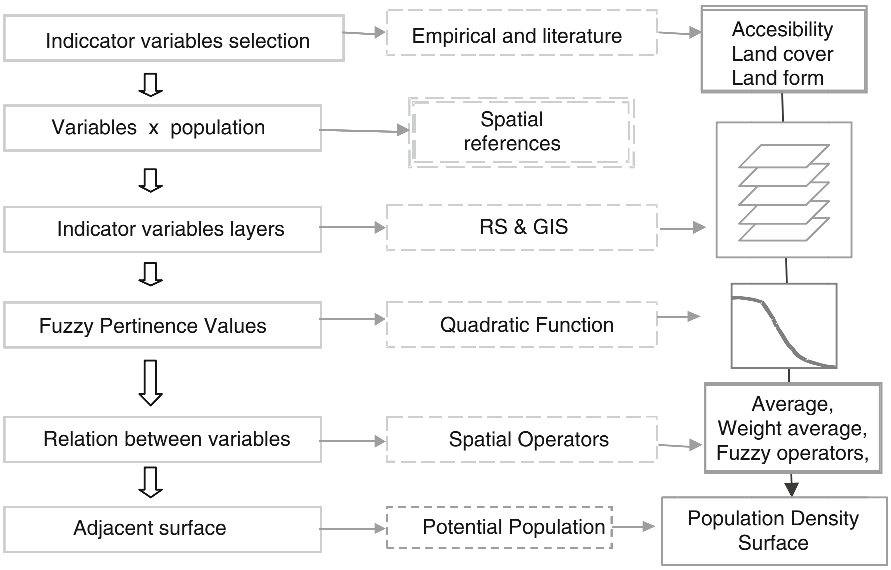


Fig. 3 The multivariate interpolation method to generate the potential surface of population occurrence. *Source* Adapted from Amaral (2003)

5. Defining operators between indicator variables: the variables are combined based on operators (averages, minimum values, etc.) to generate an adjacent model that enables the distribution of a proportional population value for every cell belonging to a certain census tract.

These steps are presented in detail below.

Selecting environmental variables related to human population presence

Different factors may determine the presence of human populations in a specific region, including historical processes, accessibility, availability of natural resources, presence of urban facilities and infrastructure and local physical characteristics, among others. The relative importance of each factor is also a fundamental variable that may vary according to local conditions. As an example, the global population distribution model proposed by Landscan used the following as indicator variables: land cover classes, distance to roads, slope classes, and the presence of night lights from the DMSP/OLS sensor (Bhaduri et al. 2002).

Access to the Amazon region (and ease of transportation within the region) has historically been a major factor associated with human presence, as described by Machado (1999). Until the 50s, occupation in the Amazon region was limited to the coastal zone and riverside areas along the main navigable rivers and a few "terra firme" areas (Costa 1997). The economy was based on extractive activities, especially on rubber extraction. In the recent Amazon colonization process, the first roads, along with the construction of the new Brazilian capital, the city of Brasília, under the Juscelino Kubischeck government (1955–1960) signaled the beginning of state intervention in the region with the National Development Plan (PDN). Migratory flow and farmers had already been established for 10 years along Belém–Brasília road (1960) when the Amazon Operation (1966) and the National Integration Plan (PIN, in 1970) were implemented. Infrastructure such as roads, an electricity power network and even natural resources inventories (RADAMBRASIL) were provided from public funds in the 70s to stimulate migration and capital flow for the new Amazon frontiers. Lands up to 100 km distant from federal roads were allocated to small farm colonization settlement projects (Costa 1997). The urbanization process also intensified following the regional colonization projects and infrastructure investments, which brought migrants from the southern and northeastern regions and changed the spatial occupation pattern. The riverine settlements were overlapped and marginalized by the new circulation axis that emerged from "terra firme" roads and villages (Godfrey and Browder 1996). From 1991 to 1996, new municipalities were created, and the population became concentrated in urban nuclei of about 20,000 inhabitants. As a result, urban nuclei were concentrated along rivers and roads axes. Becker (1998) more fully discusses the Amazon colonization process.

The presence of roads is, at certain levels, also related to deforestation in the Amazon (Skole and Tucker 1993; Alves 1999, 2002; Dale et al. 1994; Laurance et al. 2002; Fearnside 2005). Alves (2002) revealed that most of the deforestation detected from 1991 to 1996 (75%) occurred 50 km from the roads. Recently, Leite et al. (2011) reconstructed a geographically historical database of land use in Amazonia for the period of 1940–1995, through a fusion of historical census data and a contemporary land use classification. They emphasize that the spatial pattern of land use in Amazon region in this period was greatly influenced by roads and pioneer occupation areas. Even though deforestation rates are not directly related to total population counts or estimates (Geist and Lambin 2001), this type of land cover change activity indicates human presence in rural areas (Wood and Skole 1998). It is important to emphasize that deforestation in Amazon is a complex multifactor process (Camara et al. 2005).

Considering these historical factors and data availability, five variables were initially selected as indicators of human presence to disaggregate the population of Marabá municipality: distance from roads, distance from rivers, distance from urban nuclei, percentage of forest cover and slope. The three former items are related to accessibility and infrastructure. The percentage of forest cover is related to human activities, and slope (least important) is related to the general preference of human settlements for flat terrains. The LandScan project (Dobson et al. 2000) observed that most human settlements occur on soft slopes and flat land; in mountainous regions, slope values are inversely related to population density.

Identifying the relation between indicator variables and population distribution

Once the indicator variables were selected, it was necessary to determine the relationship between human presence and these indicator variables, identifying thresholds to further allocate population in a disaggregated spatial unit. For this purpose, we studied the relationships between indicator variables and the location of districts seats (for the Marabá study site) and communities (for SFD-BR163), which were considered evidence of human presence.

Marabá study site

To transform each selected variable into a population indicator, the occurrence of the districts[2] in the municipalities was assumed to be evidence of human presence related to population distribution. Each variable was studied individually to explore its relationship with the distribution of all district seats in PA State. From the frequency analysis of distance between district seats and rivers, it was observed that 90% of the districts are located up to 17 km from rivers; 50% are <3.5 km away from a river, and the average distance is 6.81 km. (Fig. 4a). Regarding distance from roads, 90% of the districts seats are <127 km far from roads and 50% are <27.5 km away (Fig. 4b). PA State is mostly flat, with slope values ranging from 0 to 7.3%. It was observed that 90% of the district seats had an average slope of <2, 50% of the districts had average slopes of <0.27% (Fig. 4c).

Regarding distance from urban centers, a nearest neighbor distance analysis over the district seats indicated that on average, such centers are 24.5 km distant from each other. Between district seats, the shortest distance is 1.5 km and the longest distance is 24.5 km.

Regarding the "forest percentage" variable in PA State, we assumed that areas with more than 95% forest cover are not likely to contain human settlements. Areas with <5% forest cover are strongly associated with human presence, and at 30% of forest cover, the likelihood of population occurrence and absence is equivalent. These forest percentage and human population occurrence are empirical values and have to be locally adjusted for each study case.

SFD-BR163 study site

The frequency distribution of distance to rivers, which is one of the relationships between the presence of settlements and an indicator variable (Fig. 5), indicated that 90% of the communities are up to 30 km from the rivers, 50% are <7.68 km away, and the average distance between the communities and rivers is 9.7 km (Fig. 5a). Concerning to the variable "distance to roads", 90% of the communities are <29.9 km from roads, and 50% are <9.7 km from roads (Fig. 5b).

Descriptive statistics indicated that SFD-BR163 communities tend to be located away from hillsides (Fig. 5c). About 90% of the communities are located among more than 1,000 m of hillsides, and no settlement <200 m from a hillside. The areas distant from the hillsides can be wetlands or plateaus, and they were mapped based on Shuttle Radar Topography Mission (SRTM) data (Farr et al. 2007), according to the vertical distance in relation to the drainage network, which was provided by the HAND algorithm (Rennó et al. 2008).

[2] According to IBGE (2000), districts in Brazil are administrative units of municipalities. Apart from the municipal seat, every district seat has the status of village ("vila").

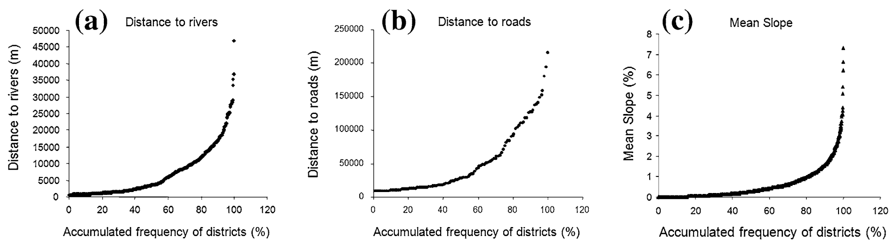


Fig. 4 Accumulated frequency of districts seats in PA state relating to their distance to rivers (**a**), distance to roads (**b**), and average slope (**c**)

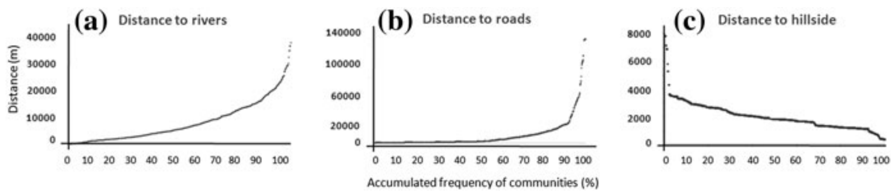


Fig. 5 Average distance to rivers (**a**), roads (**b**), and hillside (**c**) for the influence areas of the communities (accumulated frequency)

To define how the distance to the nearest communities influences populations, the distance to the nearest neighbors was analyzed, considering the locations of all communities of the SFD-BR163 municipalities. The shortest distance between communities was 2 km, and the greatest distance was 100 km. On average, the communities are 30 km distant from the nearest community.

The relationship between forest percentage and presence of population was set empirically, as settlements and population nuclei are not commonly found in regions of dense forest cover. Therefore, it was considered that over 99% forest cover, there is no possibility of population occurrence; conversely, regions with <30% forest cover are highly likely to contain settlements. The threshold of 50% of forest cover marks regions where the potential of occurrence and non-occurrence of settlement would be equal. In contrast to the Marabá procedure, several census tracts were smaller than the cell resolution (2×2 km) for the SFD-BR163. Instead of using the same threshold applied to Marabá (95% forest cover), which would exclude small census tracts, a threshold of 99% forest cover was defined for the SFD-BR163.

The cellular space on geographical database

According to the Brazilian Census (IBGE 2000), the municipality of Marabá includes 168,020 inhabitants, distributed throughout 171 census tracts: 134,373 residents in urban areas (127 census tracts) and 33,647 people at the rural areas (44 census tracts). In the urban census tracts, there was an average of 1,058.1 inhabitants (standard deviation, 324.8; median, 1,007), varying from 319 up to a maximum of

2,024 residents. In the rural census tracts, an average of 766.6 inhabitants (standard deviation of 533.0, and median equals to 721), varying from a minimum of zero up to 2,105 residents. Marabá's census tracts have an average area of 387.47 km^2, varying from 1.17 to 1,955.21 km^2. Rural census tracts in the east are dominated by agricultural activities, mainly pastures. Census tracts in the west are dominated by forest cover because of the presence of conservation units (national forests and biological reserves). To represent the heterogeneity of Marabá's census tracts, the population density surface was generated using cells of 1 km $\times$ 1 km. Only cells completely inside the Marabá boundary limits were included in population counts at the density surface.

All of the indicator variables and census tract population data formed a geographical database at TerraView GIS System (Terraview 2010), taking cells as units of analysis to generate the surface density. The ideas of cellular worlds (Couclelis 1985, 1991, 1997) and a cellular geography (Tobler 1979) support the theoretical debate in geography on representational perspectives for geographic spaces.

The classification of ETM+/Landsat Images (WRS 224/64 from 2002/08/22 and WRS 224/65 from 2002/08/13) mapped the classes "water" and "forest" for Marabá, with 30 m of spatial resolution. When images were co-registered to census tract limits, a positioning error of about one pixel (30 m) was found and projected to the UTM/SAD69 geographical reference. A simple threshold algorithm over ETM+ spectral band 4 (0.750–0.900 μm, near-infrared) classified the water bodies by area. Forest classification relied on a threshold algorithm over the ETM+ normalized vegetation index (NDVI) (Rouse et al. 1974), in which spectral information from near-infrared (band 4) and visible (band 3: 0.630–0.690 μm) light were combined $[(band4 - band3)/(band4 + band3)]$.

The general descriptive statistics related to the population in the census tracts of SFD-BR163 are presented in Table 1. It comprises the rural census tracts of the municipalities of Altamira, Aveiro, Belterra, Itaituba, Jacareacanga, Juruti, Novo Progresso, Óbidos, Placas, Prainha, Rurópolis, Santarém, and Trairão.

As SFD-BR163 is a wider area, indicator variables were obtained from different data sources than those used for Marabá (Table 2). The road network vectors available in the Ecological Economic Macrozoning database (MMA/SDS 2002) were used as a reference for the distance-to-roads variable, as computed for a regular grid at a 500 m spatial resolution.

Table 1 General population data information for SFD-BR163

	2000	2007
Number of census tracts	252	292
Minimum	0	0
Maximum	176,486	191,487
Sum	508,379	566,566
Mean	2,017.37	1,940.29
SD	11,777.78	12,205.85

Table 2 Data Sources for indicator variables at SFD-BR163

Data	Source	Reference	Analyzed years
Deforestation	Prodes (TM/Landsat 5)/INPE	INPE (2009)	2000 and 2007
Communities	Brazilian Institute of Environment and Renewable Natural Resources—IBAMA and field work survey	IBAMA (2010)	2008, 2009 and 2010
Roads	Brazilian Institute of Geography and Statistics—IBGE	IBGE (2007)	2007
Rives	Brazilian National Agency of Water—ANA	ANA (2007)	2007
Geomorphology	NASA/SRTM	Farr et al. (2007)	2000
Population	Brazilian Institute of Geography and Statistics—IBGE	IBGE (2000 and 2007)	2000 and 2007

River limits provided by National Agency for Electrical Energy (ANEEL) were used in calculating the distance to rivers. The location (points) of the districts seats (IBGE 2000) were used to analyze the distance to urban centers. The grid containing slope values (percentage) was calculated directly from the altimetry provided from SRTM data.

The variables "distance to roads", "distance to rivers", "distance to urban centers", "forest cover", and "distance to hillsides" were selected as indicator variables to generate a potential surface of population occurrence according to previous research on the occupation of this region (Alves et al. 2010; Amaral et al. 2005; Becker 2004; Furtado 2004; Pandolfo 1994). To evaluate the relationship between each of the selected indicator variables and population values, all the communities from the SFD-BR163 were studied, with 2×2 km cells of as the units of analysis.

Standardizing indicator variables based on fuzzy inference

As each indicator variable has a different scale and range, it is necessary to standardize the values to enable operations between the variables. As proposed by Turner and Openshaw (2001), fuzzy pertinence functions (Zadeh 1988; An et al. 1991) can be useful in transforming environmental data (indicator variables) into standardized variables expressing relationships to the occurrence of settlements.

The use of fuzzy sets for characterization of spatial classes is indicated when dealing with ambiguity, abstraction, and ambivalence in mathematical or conceptual models of empirical phenomena (Burrough and Mcdonnell 1998). In the concept of the pertinence function, given the value "z", the function determines whether the element evaluated belongs to a given set of analyses or not. Thus, fuzzy pertinence functions were built from maximum, minimum, and average values of each variable related to the presence of settlements. As a first approach, we proposed applying quadratic functions for all variables. Taking the distance to roads (z) as an example, the quadratic pertinence function was obtained as follows:

$$f(x) = \begin{cases} 0\,\text{if}\,4,000\,\text{m} \\ 1/(1+\alpha(z-\beta)^2) \\ 1\,\text{if}\,z \leq 1,000\,\text{m} \end{cases} \tag{1}$$

The beta (β) value corresponds to the value of the variable when the possibility of having associated population is maximum (in the case of the fuzzy value, is equal to "1"). The value of alpha (α) is obtained from the value of the variable where the occurrence or non-occurrence of the population would have the same chance of happening. In other words, alpha corresponds to the variable value where the fuzzy pertinence function is equal to 0.5 and is given by the equation:

$$\alpha = \frac{1}{(z-\beta)^2} \tag{2}$$

where z is the value of the variable when $f\,(z) = 0.5$.

From the relationships between district seats and indicator variables (for Marabá) and the communities (for SFD-BR163), fuzzy pertinence functions were obtained (Table 3).

Table 3 Fuzzy inference values for the indicator variables identified for Marabá and SFD-BR163 study sites

Indicator variable	Marabá values	SFD-BR163 values	$f\,(z)$	Marabá alfa	SFD-BR163 alfa	Marabá beta	SFD-BR163 beta
Distance to roads (m)							
≤	1,000	900	1	1.48E − 09	1.98E − 08	1,000	900
=	27,000	9,702	0.5				
>	40,000	29,900	0				
Distance to rivers (m)							
≤	1,000	900	1	2.96E − 08	5.95E − 08	1,000	900
=	6,810	7,686	0.5				
>	17,000	30,300	0				
Distance (m) to districts—Marabá; communities—SFD							
≤	1,500	2,000	1	1.89E − 09	1.28E − 09	1,500	2,000
=	24,500	30,000	0.5				
>	140,000	100,000	0				
Forest cover (%)							
≤	5	3	1	16	2.50E + 01	0.05	3.00E − 01
=	30	5	0.5				
>	99	99	0				
Average slope (%) Marabá distance to hillside (m) SFD-BR163							
≤	0.27	1,000	1	10.4,058	4.00E − 06	0.27	1,000
=	0.58	500	0.5				
>	3.5	200	0				

Combining indicator variables

After the fuzzy pertinence function was obtained for each indicator variable, it was necessary to the relationships among the variables. Establishing these relationships is a fundamental step in modeling the adjacent surface, which represents population occurrence. In the absence of a robust conceptual model, or of a standard surface that could be used to infer the relationship between variables, we proposed to apply the following: fuzzy operators (minimum, maximum, and gamma), simple average, and weighted average. In contrast to the computation of a simple average, where each variable has the same importance, in computing the weighted average, a weight has to be assigned determining the relative importance of each variable. In this work, we used the hierarchical analysis procedures (Saaty 1978) to determine the variable weights based on a paired comparison of all indicator variables.

These operators will generate a final value related to the potential of population occurrence provided by the indicator variables interactions, for each grid cell, composing the adjacent surface model.

Finally, to disaggregate population from census tracts to cells, the total population count had to be redistributed taking only the valid cells into account. Each grid cell had a potential of population occurrence assigned from the operators between indicator variables. As census tracts are represented by several cells, the population count for each cell was distributed as follows:

$$P_{\mathrm{grid}_i} = P_{\mathrm{CT}_I} \times \left(\frac{F_{\mathrm{grid}_i}}{F_{\mathrm{grid}_I}} \right) \tag{3}$$

where P_{grid_i} population count to be attributed to a grid cell i, P_{CT_I} population count for the census tract I, to which the grid cell i belongs, F_{grid_i} value resulted from the operation over the spatial indicator variables for the grid cell i (as weight average or fuzzy operators over fuzzy indicator variables), F_{grid_I} sum of all F_{grid_i} where i is a valid grid cell for the census tract I.

At the end of the procedure, population density initially depicted by the limits of census tracts (polygonal) is presented in regular cells, according to defined relationships between indicator variables and population presence.

Results and discussion

Population density surfaces

Five different population density surfaces were produced as result of the proposed methodology for the municipality of Marabá, according to the operator used to integrate indicator variables (simple average, weight average, fuzzy minimum, fuzzy maximum, and fuzzy gamma operators).

To analyze the surface results, in the absence of population data distributed in a more detailed spatial unity as census tract (2000), we took resident population data from the National Institute of Colonization and Agrarian Reform (INCRA—Instituto

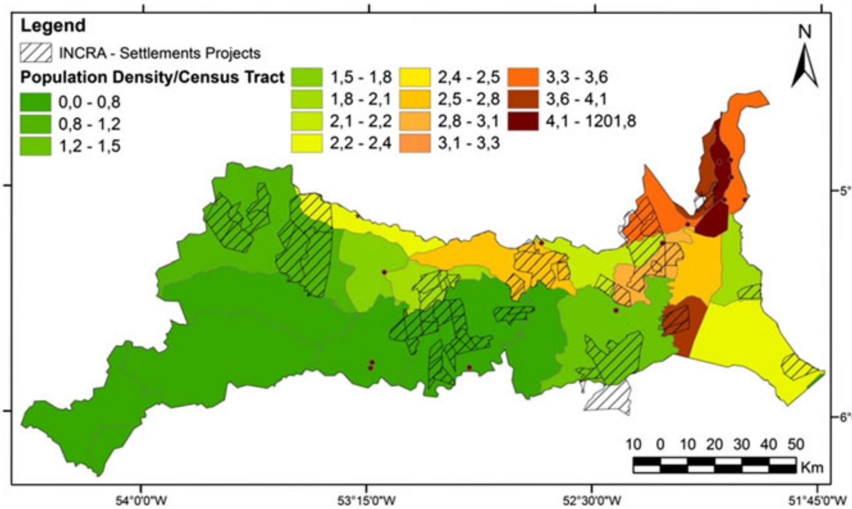


Fig. 6 Original density population from IBGE 2000 census tracts (IBGE 2000) and INCRA settlements projects localization

Nacional de Colonização e Reforma Agrária) for the official settlement projects in 2003 (Projetos de Assentamentos—PA) (MDA 2003) as a reference. Figure 6 presents the PAs limits over the original census tract representation.

When a settlement project (PA) is created, INCRA officially registers its geographical limits and the number of families that were settled in the area. In this paper, the population density for each official settlement project in Marabá was calculated by dividing the total number of residents by the area of the settlement. Because population data from PA (2003) and census tracts (2000) differ in temporal reference and spatial limits, they cannot be used to directly compare the population density values; however, it was useful to compare the surface results. In comparing these results, it was necessary to take the following into account: the population value for a PA is the total population of inhabitants for that area; population density in 2003 in the PAs is a result of the population density in 2000; we are considering population density as uniform inside a PA. There was 48 PAs, with population counts varying from 20 up to 373 inhabitants, and population density values from 0.28 up to 8.5 inhabitant/km^2.

A confusion matrix was obtained from the intersection of PA population density choropleth map with population density surfaces, considering intervals of population density. A value of global accuracy was calculated by dividing the total number of grid cells showing population density at the same population density interval in PA population density by the total number of grid cells related to PA areas. This global accuracy, given by the percentage of area correctly classified for each surface, was used only as a reference to compare the population density surfaces. We considered that the density population value inside a PA is uniformly distributed, while each PA is composed of several cells that presented heterogeneities in population density surfaces.

Table 4 Global accuracy (%) for the comparison between the population density surfaces and the resident population data from INCRA settlement projects

Population density surface	Global accuracy (%)
Simple average	14.3
Weighted average	10.4
Minimum fuzzy	10.4
Maximal fuzzy	9.5
Gamma fuzzy	18.8
Census tracts	11.8

From the global accuracy (Table 4), the fuzzy gamma operator yielded better performance (18.8%) than the other operators; however, this result was related to coincidence from areas of extreme population density values. A population density surface from the gamma operator gives a good representation of the lowest and highest population density areas, but not of the gradient of population density based on spatial heterogeneity.

The global accuracy, obtained from the simple average operator (14.3%), was related to intermediate population density ranges and was superior as a measure to the global accuracy for population density from the original census tract representation (Fig. 7a). From the visual analysis, the simple average operator also provided the best population distribution for the entire Marabá municipality, presenting more heterogeneity than weighted average surface (Fig. 7b) and the other fuzzy operators. The weighted average operator did not accurately consider the importance of forest cover. The minimum and gamma fuzzy operators were sensitive to the presence of zeros, while the maximum fuzzy operator incorporated little variability into the census tracts, generating density surface that slightly differed from the original census tract polygonal representation.

From this first result, the simple average operator was the better approach to generating population density surfaces and representing the heterogeneity of population density in Marabá. We adapted the methodology and used it to infer population distribution surfaces for a wider region, enabling the analysis of temporal evolution of the population distribution along SFD-BR163.

The resulting population density surfaces, which show the evolution of the spatial distribution of population for SFD-BR163 for 2000 and 2007, are presented in Fig. 8. Fieldwork was conducted in October 2010 to assess the accuracy of the population distribution surface, verifying geographical coordinates and population values for 98 communities in the study area (Fig. 8b). Population data for 19 communities were collected (Table 5) by interviewing residents, community leaders, health agents, and workers at the local education authority. Using field work data as the standard, the population density surface generated by the weighted average operator yielded better results than the surfaces obtained from maximum fuzzy, minimum fuzzy, gamma fuzzy, and simple average operators. The total difference between the predicted and declared population was 8%. In general, greater differences were found in estimates of small communities (Table 5). These results can be considered a good approximation, taking into account that the surface was produced using 2007 Population Counting data and that the fieldwork

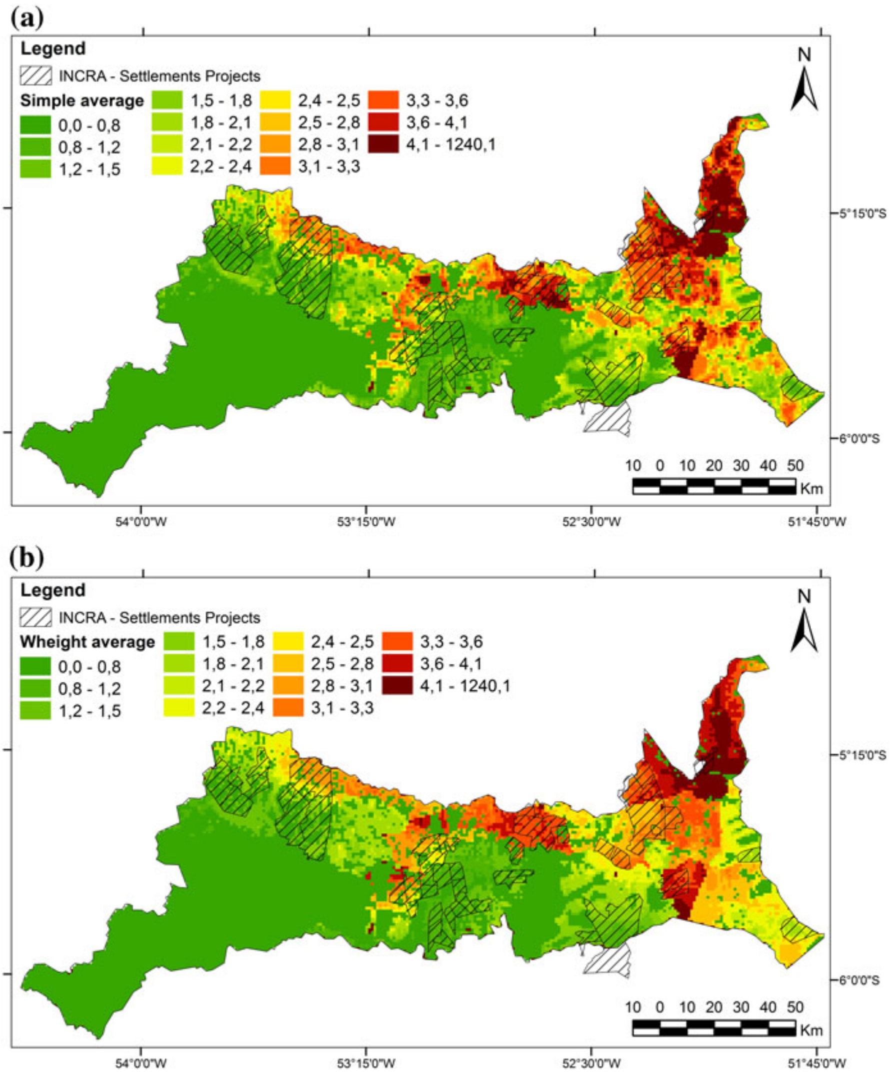


Fig. 7 Population density surface obtained for Marabá from weight average operator (**a**), and simple average operator (**b**)

population values were obtained from key informants and not estimated based on systematic survey data.

Considering the proposed methodology for disaggregating population values from census tracts into a cell space database, our observations from the Marabá and SDF-BR163 study sites make it clear that certain considerations are especially important. First, when selecting the indicator variables is necessary to take into account the particular process by which occupation grows in the region. As Marabá and SDF-BR163 are both in the Para state, and as both are involved in the process of

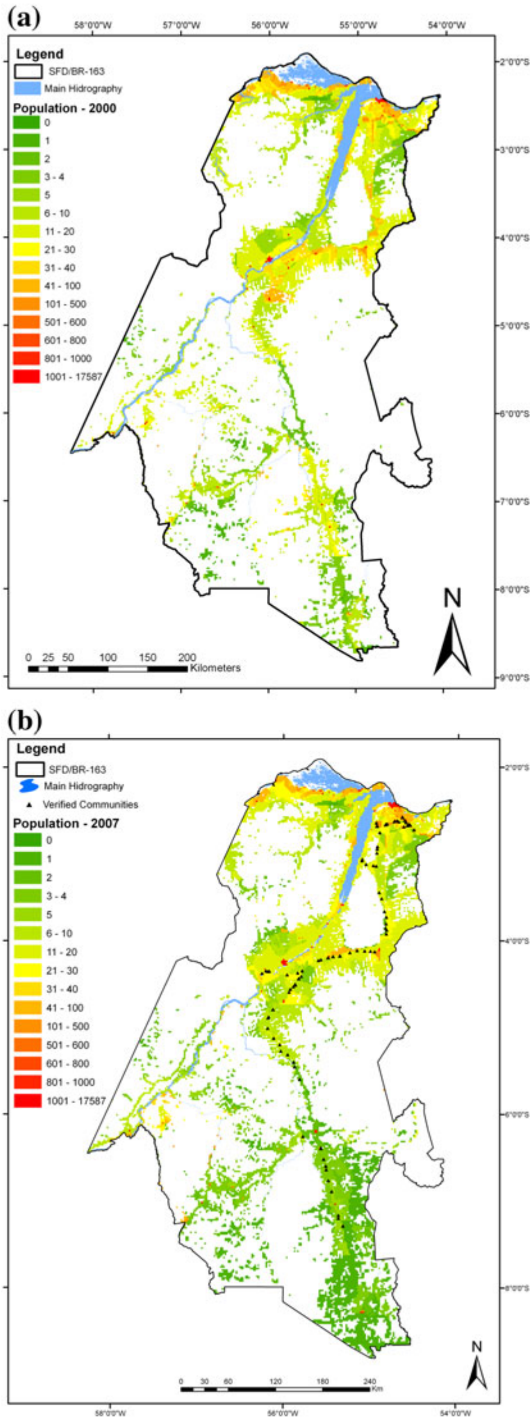


Fig. 8 Spatial distribution of population on SFD-BR163 for 2000 (**a**) and 2007 (**b**) with the location of communities verified during fieldwork (*black points*)

Table 5 Population estimates obtained from key informants in the field (2010) and from the population density surface (2007) for communities visited during the fieldwork

Community	Fieldwork population—2010	Population density surface population—2007	Differences (%)
129 do Bode	413	342	71 (17.2)
São Jorge	3,000	2,111	889 (29.6)
Galiléia	200	124	76 (38.0)
Divinópolis	3,000	2,464	536 (17.9)
Itapacuru	50	27	23 (46.0)
Itacimpasa	800	733	67 (8.4)
Nova Canaã	225	255	−30 (−13.3)
Nova Esperança	800	936	−136 (−17.0)
Bela Vista do Caracol	9,000	8,897	103 (1.1)
Jamanxim	3,500	2,990	510 (14.6)
Moraes Almeida	3,000	2,989	11 (0.4)
Alvorada	5,000	4,852	148 (3.0)
Água Azul	800	832	−32 (−4.0)
Santa Júlia	800	640	16 (20)
Três Bueiros	750	697	53 (7.1)
Riozinho	600	521	79 (13.2)
Santa Luzia	240	198	42 (17.5)
Aruri	200	163	37 (18.5)
Tucunaré	70	45	25 (35.7)

frontier expansion in the Amazon, the same indicator variables could be used for both. Even in the Amazon region, if the methodology were to be applied to the state of Amazonas, for instance, rivers navigation condition should probably be added as indicator variable.

Second, the evidence of human population used to estimate the relationship between population distribution and the indicator variable must also be defined for each area of interest. We used district distribution for Marabá and the presence of communities for SFD-BR163, with different implications for fuzzy inference (Table 3). The better the data used as evidence of population distribution, the more appropriate population density surfaces will be obtained. Obviously, the data used will also depend on data availability; we used districts for Marabá surfaces because community data were not available for the 2000 census.

Third, the quality of environmental data used as indicator variable will impact the quality of the final population density surface. As for any modeling process, the output is directly dependent on the quality of input data. In this context, remote sensing data can be a useful data source when working on large areas, such as “distance to hillside” from SRTM/NASA or deforestation mapping from Landsat/TM images from Prodes Project (Table 2), which was used in this work.

Fourth, we tested five different operators with the indicator variables. For Marabá, the simple average operator provided the best estimate of population density, whereas for SFD-BR163, the weight average operator was the best fit. Considering that simple average is an operation that gives equal weight to every indicator variable, we can infer that average operators performed better than fuzzy operators. The simple average was first chosen also because it is easier to implement, understand, and interpret.

Finally, it is always important to have field information or another data source to validate the population density surfaces. For Marabá, data from INCRA settlements projects were used as a general reference for operator comparison. In the SFD-BR163, we managed to collect field information about the resident population, enabling direct comparison between population values from fieldwork and population density surface. The proposed methodology can also be used to optimize fieldwork effort: rather than conduct a wide survey for an entire low area, population density surface can be used to stratify the area, reducing the number of points to be visit in the field.

Ultimately, we are not proposing a method to estimate new population values but a consistent criterion for disaggregating population counts from census tracts, represented by polygons delimiting areas of different sizes, into cell spaces that are smaller and present regular spatial resolution. Obviously, several other methodologies could be applied (REFS); however, our proposal seeks to use simple spatial analysis and geographical information tools to represent knowledge about how population is distributed inside census tracts.

The representation of population data by population density surfaces allows, in addition to other modeling purposes, the analysis of temporal evolution of population distribution and the study of population–environmental relations, even if the limits of census tracts differ between census surveys.

Ultimately, we are not proposing a method to estimate new population values, rather just a consistent criterion to disaggregate population counts from census tracts, represented by polygons delimiting areas of different sizes, into cells space that are smaller and present regular spatial resolution. Obviously, other methodological approach could be applied; however, our proposal seeks to use simple spatial analysis and geographical information tools to represent the knowledge about how population is distributed inside census tracts. The representation of population data by population density surfaces allows, in addition to other modeling purposes, the analysis of temporal evolution of population distribution and the study of population–environmental relations, even if the limits of census tracts differ between census surveys.

Evolution of population density in the SFD-BR163

The recent demographic dynamics in agricultural frontier areas in SFD-BR163 municipalities is treated as evidence of the occupation processes, demonstrating, among other things, their ability to attract and retain population. Table 6 shows the evolution of the total population in the DFS-BR163 between 2000 and 2007 as well as population growth, considering exclusively population values for cells within the limits of SFD-BR163. In 2000, the population of the SFD-BR163 municipalities was estimated at 476,656 inhabitants, and it reached 532,457 inhabitants in 2007, a

Table 6 Total resident population for municipalities of SFD-BR163 for 2000 (IBGE Demographic Census 2000), and 2007 (IBGE—population count 2007), and the results from density surfaces for those cells of municipalities contained in SFD-BR163 physical limits

Locality	Municipality 2000	Municipality 2007	Cells inside SFD-BR163 2000	Cells inside SFD-BR163 2007	Cells inside SFD-BR163 2007–2000	%
Brazil	169,799,170	183,987,291				8.36
PA state	6,192,307	7,065,573				14.10
Altamira	77,439	92,105	3,286	5,548	2,263	68.87
Aveiro	15,518	1,883	11,954	17,238	5,283	44.19
Belterra	14,594	12,707	14,573	12,707	−1,866	−12.80
Itaituba	94,750	118,194	95,653	117,450	21,797	22.79
Jacareacanga	24,024	37,073	12,919	19,515	6,596	51.06
Juruti	31,198	33,775	28,980	33,909	4,928	17.00
Novo Progresso	24,948	21,598	24,666	21,583	−3,083	−12.50
Óbidos	46,490	46,793	2,919	1,291	−1,628	39.28
Placas	13,394	17,898	5,170	7,200	2,031	88.11
Prainha	27,301	26,436	1,404	2,640	1,237	26.68
Rurópolis	24,660	32,950	26,011	32,950	6,939	4.00
Santarém	262,538	274,285	233,057	242,380	9,322	14.04
Trairão	14,042	16,097	14,064	16,039	1,975	−55.77
SFD-BR163 total	565,907	641,737	476,656	532,457	55,802	11.71

cumulative increase rate of 11.7% for the analyzed period. Although the environmental conditions represented by the indicator variables influenced population distribution in the same way in 2000 as in 2007, the pattern of population distribution has changed. Instead of the intensification of population density in previously occupied areas, the population is scattered over the territory of SFD-BR163 (Fig. 8).

There was some concentration of inhabitants in the northern region, along the Amazonas River and in the vicinity of Santarém; however, the concentration of population along the axis formed by Rurópolis–Itaituba–Trairão, following Transamazônica and BR-163 highways observed in 2000 was substituted by larger areas with low population density. This change may be associated with the population decrease observed in Trairão.

Table 6 shows the evolution of the total population in the municipalities near the BR-163 highway between 2000 and 2007, as well as population growth that considers exclusively the population values for the cells inside the limits of SFD-BR163. In 2000, the population of the municipalities (considering only the population count for cells inside the borders of SFD-BR163) was 476,656 inhabitants, and it reached 532,457 inhabitants in 2007. This change represents an accumulated increase rate of 11.7% at the period.

In Jacareacanga municipality, there was a high population increment taking the period 2000–2007 (Table 6). However, the latest published demographic census summary (IBGE 2010) have pointed out that the Jacareacanga's population was 14,000 residents, instead of the 24,000 counted by the population counting of 2007 (IBGE 2007). Because of these discrepant values, a consultation was carried out with the IBGE population office and it was reported that there were problems with the 2007 Population Counting in Jacareacanga. IBGE field team had to be replaced by a team from Itaituba. Because they didn't know well the region, there were double counting, resulting in overestimation of population in 2007 for this municipality.

The effect of conservation units in the SFD-BR163 becomes evident when population surfaces are compared. In the southern SFD, in the municipality of Altamira, the population started to occupy the east side of BR-163 highway, while in the west side, the population in Novo Progresso decreased. Most areas without settlement correspond to a conservation unit of restricted use such as National Parks or Indian Land. The areas where the population has spread out or concentrated during the analyzed period corresponded to unused public land or conservation units that allow sustainable use.

The resident population generally increased in the municipalities, with exception of Aveiro, Belterra, and Novo Progresso. Aveiro is almost entirely contained by conservation units. In the field, the local government explained that because they do not have the legal land tenure for rural or for urban areas, the most common economic activities, such as agriculture or cattle ranching, are restricted. They are not even allowed to build a hospital or any other public urban infrastructure necessary for settlement.

In Belterra, land ownership became concentrated when soybean producers came to the Santareno Plateau. Small producers sold their properties for grain production and migrated to urban areas (Coelho 2008). In the population density maps, this transition is expressed by the growing number of residents in the city of Belterra and the disappearance of denser areas closer to this city.

According to information collected from the fieldwork interviews, Novo Progresso received an intense immigration flux in the 2000s (Fig. 9). Men at working age (20–40 years old) went to Novo Progresso to work in the numerous sawmills in the city. With the intensification of the battle against deforestation and illegal timber practices and the creation of several conservation units in 2006, the population of this city faced a reduction of about 3,000 inhabitants from 2000 to

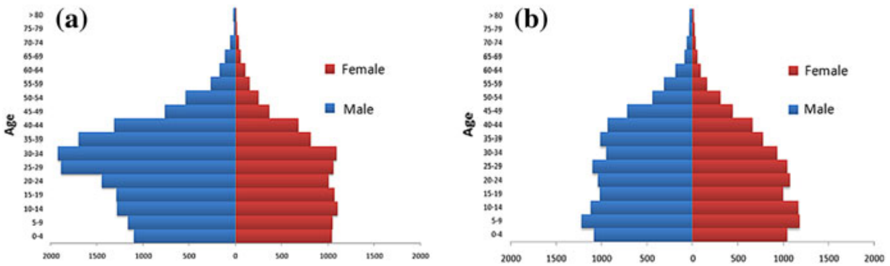


Fig. 9 Age pyramids for Novo Progresso in 2000 (**a**) and 2007 (**b**). *Source* IBGE (2000) and (2007)

2007. These individuals were mostly men, as indicated by the demographic pyramid.

Conclusions

This paper proposed a methodology to disaggregate population data provided in census tracts into smaller spatial units based on ancillary environmental data and geoinformation techniques. The results suggest it is possible to recover the heterogeneity of census tracts when the relations between the indicator variables and population occurrence are defined with criterion and local particularities are taken into account. The methodology developed for the municipality of Marabá was adapted to the Sustainable Forest District of BR-163 municipalities. As the area of interest was expanded, the cell size was enlarged, and the pattern of population distribution was obtained from the presence of communities. Data from fieldwork indicated an adequate fit between the population count predicted from the population surface and the total population for the communities along BR-163 highway.

The population density surfaces enabled the interpretation of the distribution of human presence in terms of the territory to be potentially occupied. The model allocates no population in areas where there is no possibility of human presence, such as in rivers, dense forests cover, sand islands, etc. Moreover, representing population in cell spaces enables monitoring of the population over the time. Even if the limits of municipalities or census tracts change, what is very common in such dynamic regions as Amazon, the distribution can be represented and compared in a cell space.

The evolution of the resident population over the DFS/BR-163 territory from 2000 to 2007 showed spatial patterns comparable to the occupation process described in the literature and reported in the field. Therefore, since the proposed methodology can be adapted to represent the population distribution of other areas, population density surfaces can be useful as additional data source to study population and territory dynamics.

The proposed methodology can be improved using knowledge about the spatial indicator variables and human presence relationships. With population data from the 2010 census, we will be able to represent population density evolution over a 10-year period and better monitor the impacts of the creation of a sustainable forest district on regional population distribution. Such methodological advancements, we hope, better pave the way toward the more integrated population and environment scholarship as envisioned by, and in honor of, Professor Daniel Hogan.

Acknowledgments This work was partially support by INPE—National Institute for Space Research, Cenários Project (Cenários para a Amazônia: uso da terra, biodiversidade e clima), and LUA/IAM Project—Land Use Change in Amazonia: Institutional Analysis and Modeling at multiple temporal and spatial scales.

References

Alves, D. (1999). An analysis of the geographical patterns of deforestation in Brazilian Amazônia in the 1991–1996 period. In *Paper presented at the 48th annual conference of the center for Latin American studies—Patterns and processes of land use and forest change in the Amazon*, University of Florida, Gainesville.

Alves, D. S. (2002). Space time dynamics of deforestation in Brazilian Amazonia. *International Journal of Remote Sensing, 23*, 2903–2908.

Alves, P. A., Amaral, S., Escada, M. I. S., & Monteiro, A. M. V. (2010). Explorando as relações entre a dinâmica demográfica, estrutura econômica e no uso e cobertura da terra no sul do Pará: lições para o Distrito Florestal Sustentável da BR-163. *Geografia, 35*(1), 165–182.

Amaral, S. (2003). *Geoinformação para estudos demográficos*: *Representação Espacial de Dados de População na Amazônia Brasileira.* Ph.D. thesis, Escola Politécnica da USP, Universidade de São Paulo, São Paulo.

Amaral, S., Câmara, G., Monteiro, A. M. V., Quintanilha, J. A., & Elvidge, C. D. (2005). Estimating population and energy consumption in Brazilian Amazônia using DMSP night-time satellite data. *Computers, Environment and Urban Systems, 29*(2), 179–195.

An, P., Moon, W. M., & Rencz, A. (1991). Application of fuzzy set theory for integration of geological, geophysical and remote sensing data. *Canadian Journal of Exploration Geophysics, 27*, 1–11.

ANA (Brazilian National Agency of Water). (2007). *Hidrography spatial data.* From http://hidroweb.ana.gov.br/.

Bajat, B., Hengl, T., Kilibarda, M., & Krunic, N. (2011). Mapping population change index in Southern Serbia (1961–2027) as a function of environmental factors. *Computers, Environment and Urban Systems, 35*, 35–44.

Balk, D., Gorokhovich, Y., & Levy, M. (2005). *Estimation of coastal populations exposed to the 26 December 2004 tsunami* (Vol. Working Paper). Center for International Earth Science Information Network. New York: Columbia University.

Becker, B. K. (1998). *A Especificidade do Urbano na Amazônia*: *Desafios para políticas Públicas Consequentes. Estudo elaborado para a Secretaria de Coordenação dos Assuntos da Amazônia Legal*. Rio de Janeiro: Ministério do Meio Ambiente.

Becker, B. K. (2004). *Amazônia—Geopolítica na Virada do III Milênio* (Vol. 1). Rio de Janeiro: Editora Garamond.

Bhaduri, B., Bright, E., Coleman, P., & Dobson, J. E. L. (2002). Locating people is what matters. *Geoinformatics, 5*(2), 34–37.

Burrough, P. A., & McDonnell, R. A. (1998). *Principles of geographic information systems*. Oxford: University Press.

Camara, G. (2009). *Land use change in Amazonia*: *Institutional analysis and modelling at multiple temporal and spatial scales* (*LUA/IAM*). FAPESP Research Program on Global Climate Change (FRPGCC).

Camara, G., Aguiar, A. P. D., Escada, M. I. S., Amaral, S., Carneiro, T., Monteiro, A. M. V., et al. (2005). Amazonian deforestation models. *Science, 307*(18), 1043–1044.

Coelho, A. (2008). *Modelagem de dinâmica do uso da terra e cobertura vegetal em área de expansão de grãos na região oeste do Pará*. Belém, PA: Universidade Federal do Pará.

Costa, W. M. D. (1997). *O Estado e as Políticas Territoriais no Brasil* (7a ed.). Ed. Contexto.

Couclelis, H. (1985). Cellular worlds: A framework for modelling micro-macro dynamics. *Environment and Planning A, 17*(1), 585–596.

Couclelis, H. (1991). Requirements for planning-relevant GIS: A spatial perspective. *Papers in Regional Science, 70*(1), 9–19.

Couclelis, H. (1997). From cellular automata to urban models: New principles for model development and implementation. *Environment and Planning B, 24*(1), 165–174.

Coudreau, H. (1977). *Viagem ao Tapajós*. São Paulo, SP: Editora da Universidade de São Paulo.

Dale, V. H., O'Neill, R. V., Southworth, F., & Pedlowski, M. (1994). Modeling effects of land management in the Brazilian Amazonian settlement of Rondonia. *Conservation Biology, 8*(1), 196–206.

De Reynal, V., Hebette, J., Muchagata, M. G., & Topall, O. (1995). *Agriculturas familiares e desenvolvimento em frente pioneira amazônica*. LASAT-CAT/GRET/UAG.

DeMers, M. N. (1999). *Fundamentals of geographic information systems* (2nd ed.). New York: Wiley.

Dobson, J. E., Bright, E. A., Coleman, P. R., Duree, R. C., & Worley, B. A. (2000). LandScan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing, 66*(7), 849–857.

DOU (Diário Oficial da União). (2004). *PORTARIA N° 316, DE 30 DE JUNHO DE 2004* (ISSN: 1677-7042, Seção 1, p. 14).

Ehrlich, P. R. (1968). *The population bomb*. New York: Ballantine Books.

Eicher, C., & Brewer, C. (2001). Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science, 28*, 125–138.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics, 45*(2).

Fearnside, P. (2005). Deforestation in Brazilian Amazonia: History, rates and consequences. *Conservation Biology, 19*(3), 680–688.

Fisher, P. F., & Langford, M. (1995). Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation. *Environment and Planning A, 27*, 211–224.

Frohn, R. C., McGwire, K. C., Dales, V. H., & Estes, J. E. (1996). Using satellite remote sensing to evaluate a socio-economic and ecological model of deforestation in Rondonia, Brazil. *International Journal of Remote Sensing, 17*, 3233–3255.

Furtado, C. (2004). *Formação econômica do Brasil*. São Paulo, SP: Companhia das Letras.

Gallego, F. J. (2010). A population density grid of the European Union. *Population and Environment, 31*, 460–473.

Gallego, J., & Peedell, S. (2001). Using CORINE land cover to map population density. In Eurostat, DG Agriculture, DG Environment, Joint Research Center, European Environment Agency (Ed.), *Towards agri-environmental indicators: Integrating statistical and administrative data with land cover information, Topic report 6/2001* (pp. 94–105). Copenhagen: European Environment Agency.

Geist, H. J., & Lambin, E. F. (2001). *What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence*. LUCC report series no. 4. Louvain-la-Neuve, Belgium: LUCC International Project Office.

Godfrey, B. J., & Browder, J. O. (1996). Disarticulated urbanization in the Brazilian Amazon. *The Geographical Review, 85*(3), 441–445.

Goodchild, M. F., Anselin, L., & Deichmann, U. (1993). A framework for the areal interpolation of socioeconomic data. *Environment and Planning A, 25*, 383–397.

Gregory, I. N. (2002). The accuracy of areal interpolation techniques: Standardizing nineteenth and twentieth century census data to allow long-term comparisons. *Computers, Environment and Urban Systems, 26*, 293–314.

Harvey, F. (2008). *A primer of GIS: Fundamental geographic and cartographic concepts*. New York: The Guilford Press.

Hay, S. I., Noor, A. M., Nelson, A., & Tatem, A. J. (2005). The accuracy of human population maps for public health application. *Tropical Medicine and International Health, 10*(10), 1073–1086.

Hogan, D. J. (1989). População e Meio Ambiente. *Textos NEPO, 16,* 86. Campinas: Núcleo de Estudos Populacionais—NEPO/UNICAMP.

Hogan, D. J. (2000). A relação entre população e ambiente: desafios para a demografia. In H. Torres & H. Costa (Eds.), *População e Meio Ambiente. Debates e Desafios* (pp. 21–52). São Paulo: Editora SENAC.

Hogan, D. J. (2001). Demographic dynamics and environmental change in Brazil. *Ambiente e Sociedade, Ano IV*(9), 1–30.

Hogan, D. J. (2007). População e Meio Ambiente: A emergência de um novo campo de estudos. In D. J. Hogan (Ed.), *Dinâmica populacional e mudança ambiental: cenários para o desenvolvimento brasileiro* (pp. 13–58). Campinas, SP, BR: Núcleo de Estudos de População-NEPO/Unicamp.

IBAMA (Brazilian Institute of Environment and Renewable Natural Resources). (2010). *Pará state communities—Spatial data*. From http://siscom.ibama.gov.br/shapes.

IBGE (Brazilian Institute of Geography and Statistics). (2007). *Spatial data*. From http://www.ibge.gov.br.

IBGE (Instituto Brasileiro de Geografia e Estatística). (2000). *Censo Demográfico—2000.* Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística.

IBGE (Instituto Brasileiro de Geografia e Estatística). (2010). *Censo Demográfico 2010*. From http://www.ibge.gov.br/home/estatistica/populacao/censo2010/default_sinopse.shtm.

INPE (Instituto Nacional de Pesquisas Espaciais). (2009). *Monitoramento da floresta amazônica por satélite, Projeto PRODES* 2008–2010. From http://www.obt.inpe.br/prodes/.
Langford, M. (2003). Refining methods for dasymetric mapping. In V. Mesev (Ed.), *Remotely sensed cities* (pp. 137–156). London, UK: Taylor & Francis.
Langford, M. (2007). Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems, 31*, 19–32.
Laurance, W., Albernaz, A. K. M., Schroth, G., Fearnside, P. M., Bergen, S., Venticinque, E. M., et al. (2002). Predictors of deforestation in the Brazilian Amazon. *Journal of Biogeography, 29*, 737–748.
Leite, C., Costa, M. H., Lima, C. A. A. S., & Ribeiro, G. C. S. (2011). Historical reconstruction of land use in the Brazilian Amazon (1940–1995). *Journal of Land Use Science, 6*(1), 32–52.
Linard, C., Gilbert, M., & Tatem, A. J. (2011). Assessing the use of global land cover data for guiding large area population distribution modelling. *Geo Journal, 76*(5), 525–538. doi:10.1007/s10708-010-9364-8.
Liverman, D., Moran, E. F., Rindfuss, R., & Stern, P. C. (1988). *People and pixel: Linking remote sensing and social science*. Washington, DC: National Academy Press.
Luizão, F. J. (2008). *Cenários para a Amazônia*: *Clima, Biodiversidade e Uso da Terra*. FINEP process N°: 2166/07.
Machado, L. O. (1999). Urbanização e Mercado de trabalho na Amazônia Brasileira. *Cadernos IPPUR, 13*(1), 109–138.
Maynard-Ford, M. C., Phillips, E. C., & Chirico, P. G. (2008). *Mapping vulnerability to disasters in Latin America and the Caribbean, 1900–2007*. Reston, VA: US Geological Survey.
McCracken, S., Brondizio, E., Nelson, D., Moran, E. F., Siqueira, A., & Rodriguez-Pedraza, C. (1999). Remote sensing and GIS at farm property level: Demography and deforestation in the Brazilian Amazon. *Photogrammetric Engineering and Remote Sensing, 65*(11), 1311–1320.
McCracken, S., Siqueira, A. D., Moran, E. F., & Brondizio, E. S. (2002). Land use patterns on an agricultural frontier in Brazil: Insights and examples from a demographic perspective. In C. H. Wood & R. Porro (Eds.), *Deforestation and land use in the Amazon* (pp. 162–217). Gainsville, FL: University Press of Florida.
McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization, 19*(1), 17–37.
McMichael, A. J. (1993). *Planetary overload: Global environmental change and the health of the human species*. Cambridge: Cambridge University Press.
MDA. (2003). *Projetos de Assentamentos, Gerência Operacional de Sistemas (GSO)—Web e SIPRA4.0*. Marabá: Ministério do Desenvolvimento Agrário-MDA, Instituto Nacional de Colonização e Reforma Agrária-INCRA, Superintendência Regional do Sul do Pará SR(27).
Meirelles, M. S. P. M. (1997). *Análise integrada do ambiente através de geoprocessamento: uma proposta metodológica para elaboração de zoneamentos*. Rio de Janeiro: UFRJ.
Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *Professional Geographer, 55*(1), 31–42.
Mennis, J., & Hultgren, T. (2006). Intelligent dasymetric mapping and its application to areal interpolation. *Cartography and Geographic Information Science, 33*(3), 179–194.
MMA. (2006). *Plano de Ação 2006–2007: Grupo de trabalho interinstitucional do Distrito Florestal da BR-163* (p. 27). Brasília: Ministério do Meio Ambiente.
MMA/SDS (Ministério do Meio Ambiente/Secretaria de Desenvolvimento Sustentável. Sistematização e Atualização de Informações). (2002). *Projeto Cenários para a Amazônia Legal*: *Sistema de Consulta*. Brasília, DF: Consórcio ZEE-Brasil—CDROM.
Moran, E. F., & Brondizio, E. (1998). Land-use change after deforestation in Amazonia. In D. Liverman, E. F. Moran, R. Rindfuss, & P. C. Stern (Eds.), *People and pixel: Linking remote sensing and social science* (pp. 94–120). Washington, DC: National Academy Press.
Moran, E. F., Brondizio, E., Mausel, P., & Wu, Y. (1994). Integrating Amazonian vegetation, land use, and satellite data. *BioScience, 44*, 329–338.
Moran, E. F., Siqueira, A., & Brondizio, E. (2003). Household demographic structure and its relationship to the Amazon Basin. In J. Fox, V. Mishra, R. Rindfuss, & S. Walsh (Eds.), *People and environment: Approaches to linking household and community surveys to remote sensing and GIS* (pp. 1–30). Boston: Kluwer.
Mrozinski, R., & Cromley, R. (1999). Singly- and doubly-constrained methods of areal interpolation for vector-based GIS. *Transactions in GIS, 3*, 285–301.

Nicholls, R. J., Tol, R. S. J., & Vafeidis, A. T. (2005). Global estimates of the impact of a collapse of the West Antarctic ice sheet. *Climatic Change, 91*(1–2), 171–191.

Oliveira, M. C. C., Silva, W. R., & Santos, W. A. (2001). *Estudo sobre o processo de migração de agricultores familiares na área rural da região de Marabá* (p. 46). Brazil: LASAT—Laboratório Sócio-Agronômico do Tocantins.

Pandolfo, C. (1994). *Amazônia brasileira: ocupação, desenvolvimento e perspectivas atuais e futuras*. Belém, PA: CEJUP.

Reibel, M., & Bufalino, M. (2005). Street-weighted interpolation techniques for demographic count estimation in incompatible zone systems. *Environment and Planning A, 37*(1), 127–139.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., et al. (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sensing of Environment, 112*, 3469–3481.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In *Paper presented at the third earth resources technology satellite-1 symposium*, Greenbelt.

Saaty, T. L. (1978). Exploring the interface between hierarchies, multiple objectives and fuzzy sets. *Fuzzy Sets and Systems, 1*, 57–68.

Skole, D., & Tucker, C. (1993). Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science, 260*, 1905–1910.

Sleeter, R. (2004). Dasymetric mapping techniques for the San Francisco Bay region, CA: Urban and regional information systems association. In *Paper presented at the urban and regional information systems association*, Reno, USA.

Terraview. (2010). *Terraview*. São José dos Campos, SP: INPE.

Tobler, W. R. (1979). Cellular geography. In S. Gale & G. Olsson (Eds.), *Philosophy in geography* (pp. 379–386). Dordrecht, Holland: Reidel.

Turner, A., & Openshaw, S. (2001). *Disaggregative spatial interpolation*. GISRUK, 2002. From http://www.geog.leeds.ac.uk/people/a.turner/publications/OpenshawTurner2001.html.

Walsh, S. J. (2010). *Beyond people & pixels*: *Integrating people and environment in LULC studies*. Panel contribution to the population-environment research network cyberseminar, “what are the remote sensing data needs of the population-environment research community?” from http://www.populationenvironmentresearch.org/seminars.jsp.

Wood, C. H., & Skole, D. (1998). Linking satellite, census, and survey data to study deforestation in the Brazilian Amazon. In D. Liverman, E. F. Moran, R. Rindfuss, & P. C. Stern (Eds.), *People and pixel: Linking remote sensing and social science* (pp. 70–93). Washington, DC: National Academy Press.

Zadeh, L. A. (1988). Fuzzy logic. *Computer, 21*, 83–92.