reviewed paper

Quantifying Town Development in Space and Time using Land Use Data

Miriam Steurer, Caroline Bayr

(Mag. Dr. Miriam Steurer MA, JOANNEUM RESEARCH – POLICIES, Leonhardstraße 59, 8010 Graz, Austria, miriam.steurer@joanneum.at)

(DI Caroline Bayr, JOANNEUM RESEARCH – POLICIES, Leonhardstraße 59, 8010 Graz, Austria, caroline.bayr@joanneum.at)

1 ABSTRACT

The most popular methods to measure urban sprawl are based on population growth and/or density measurements. The data needed for these calculations are easily available as long as the focus lies on politically drawn boundaries or jurisdictions. If however the emphasis lies on the organic de-facto town boundaries, these population data are not always readily (and cheaply) available. Here we employ and develop indicators that focus on the organically grown (functional) size of a town or village rather than its politically defined boundary. We use the structure provided by the Corine Land Cover (CLC) Project (2006) to define usage categories as well as grid size (100m x 100m).

In this paper we follow two lines. First, we identify mismatches between legal and de-facto town boundaries using land use data. Second, we employ several land-use-based urban sprawl indicators and compare their performance. Some of these indicators have been used in the literature before, others we developed ourselves. These indicators are based on the structure and/or composition of urban space.

2 INTRODUCTION

Urban sprawl is a complex phenomenon. For some the main characteristic of urban sprawl is poor accessibility, or as Al Gore defined it: "a gallon of gas can be used up just driving to get a gallon of milk". [1] For others the main components of urban sprawl are low density zoning (single family homes on large plots) or single-use zoning rules that lead to a geographical separation between residential and commercial areas. For some urban sprawl is an aesthetic criterion. Sprawl in this sense is defined as boring housing developments, interchangeable commercial developments containing malls and fast food chains. Not everyone sees urban sprawl in a negative way however, Glaeser and Kahn (2003) for example see urban sprawl as the "inexorable product of car-based living" and argue that "sprawl has been associated with significant improvements in quality of living", and that "the environmental impacts of sprawl have been offset by technological change". Brueckner (2000) also see urban sprawl as the by-product of rising incomes, growing population and a fall in car-commuting costs. His argument however takes a different note in that he argues that sprawl is associated with three different types of market failure: The failure to take into account the value of open space, the failure to take congestion and ecological consequences of urban sprawl into account, and the failure to make developers pay for the infrastructure they need. As urban sprawl means different things to different people, it is not surprising that many different definitions of urban sprawl exist in the literature. However everyone seems to agree that urban sprawl is an important phenomenon. Frenkel and Ashkenazi (2008) summed up the situation very well when they wrote: "we know that sprawl is significant, but we are not yet sure what it is exactly or how to measure it" (Frenkel and Ashkenazi, 2008).

Of course, measuring urban sprawl is important not only because experts agree on the importance of the sprawl phenomenon. Measuring urban sprawl is important because it is associated with a variety of major ecological, social, and health effects and developments. The impacts associated with urban sprawl range from the lack of scale economics (Frenkel and Ashkenazi, 2001), ecological problems such as air pollution and congestion due to increased car use associated with urban sprawl (Brueckner, 2000, Nechyba and Walsh, 2004), to social problems of increased segregation (e.g. Glaeser and Kahn, 2003) or increased isolation (Frumkin, 2002), to health problems such as obesity [2] (e.g. Ewing et al., 2003, Bray et al., 2005).

Our objective here is to compare alternative measures of urban spawl. Some have been previously used and some we have developed ourselves. Our land use data comes from the Corine Land Cover Project (European

---

[1] Quote from the speech "war on sprawl" by Al Gore during his campaign for the U.S. presidency, January 1999.

[2] Obesity and urban sprawl are statistically correlated. Most researchers argue that obesity is a negative by-product of urban sprawl, partly induced by the car dependency of the population. However some researchers argue, that the causation between urban sprawl and obesity is actually the reverse with obese people being more likely to choose to live in a sprawling environment (see e.g. Eid et al., 2008).

Environment Agency, 2010). We use the grid structure as well as the land use labelling of the CLC Project to calculate alternative measures of urban sprawl. The grid size in the Corine Land Cover Project is 100 by 100 meters. There are 42 land use classes present in the CLC Project. Cells are classified according to their dominant usage. Population is mainly found in CLC class 1 (continuous urban fabric) and 2 (discontinuous urban fabric).

In our empirical comparions we focus on Austria's second largest city, Graz, as well as two smaller regions in the same province, Gleisdorf and Leibnitz. We choose different size urban developments on purpose, as urban sprawl does not only affect large cities: "urban sprawl is a process which can affect even the smallest of villages" (Sudhira et al., 2004).

3 DENSITY MEASURES

The most popular methods to measure urban sprawl are based on population density measurements. The main input in these measures is the number of people living in a particular area (e.g. people per square kilometer). Low population density numbers suggest high degrees of urban sprawl. We consider two different ways of measuring population density.

3.1 Density Measure 1: number of people divided by total area of district/municipality

The standard version of the density measure is to divide the number of people by the total area of a district/municipality ( ). Using this method we can see large differences in the population density for major international cities. The 6,299km^2 city area of Los Angeles has a population of 14,900,000, which corresponds to a population density of 2,400 people per square kilometre according to this method. For the 2,163km^2 area of Seoul-Incheon in South Korea with a population of 22,547,000 people, the population density is 10,400 people per square kilometre. The city of Dhaka in Bangladesh has by far the highest population density in the world with 44,400 people per square kilometer (New Geography, 2013). According to the density measure of urban sprawl, if a specific area experiences positive population growth, the population density becomes higher and urban sprawl is reduced.

The formula for this density measure reads as follows:

$$D_1 = \frac{\sum_i pop_i}{\sum_i x_i}$$

where pop_i is the population in cell i for $i = 1, \ldots, n$ with n as the total number of cells in the city and $\sum_i x_i$ is the whole area of the city (sum of all cells in the town).

For the three Austrian municipalities the population densities per square kilometer are more moderate. While the population density within the official city limits of Graz is 2,234 people per km^2, for the 57 surrounding communities of Graz it is only 128 people per km^2. Within the official political city limits, the density index for the regional town of Gleisdorf is 1,214 and for the southern styrian town of Leibnitz it is 1,322 people per km^2.

3.2 Density Measure 2: number of people divided by built-up cells in district/municipality

We include a second density measure here which considers the population density in built-up areas only (. The rationale for doing this is that the population may not be spread evenly across all cell types in a municipality. Suppose for example that the district/municipality contains a large park. This acts to reduce overall population density and hence increases the level of measured urban sprawl according to density measure 1. However, it could be legitimately argued that the park should be ignored when measuring population density. This is what density measure 2 does.

We use land use data from the Corine Land Cover Project (2006) to classify individual 100m x 100m grid cells according to their dominant usage into 42 possible land use classes. Our first task is to decide which of these land use classes constitute a built-up area. This may differ across regions. We define a cell as built-up if there is residential housing in that cell. While we do not have population data at the level of individual cells, we can determine which cell types contribute population using a regression model.

The city of Graz has 285,387 inhabitants (including principal and secondary residence) in 2006. The city is divided into 259 voting districts ("Zählsprengel") for each of which we have the number and land use type of the cells lying within that area as well as the number of inhabitants. Within Graz only 12 CLC cell types are observed.

We estimate the following regression:

$$bev_{ZSP} = \beta_i \cdot x_i, \ i = 1, \dots, 12$$

where bev_{ZSP} is the population in a district, x_i is the number of observations of cell type i observed in that district. The regression is run over the 259 districts in Graz. We find that only two cell types are significant with a p-value [3] smaller than the significance level of 5 %: these are CLC class 1 (continuous urban fabric) and CLC class 2 (discontinuous urban fabric).

To find the relative distribution of population between CLC classes 1 and 2 within the city region, we run the regression above for the entire area of Graz and include only CLC class 1 and 2 as explanatory variables. For the city area of Graz we find an imputed average cell population of 101.85 people for CLC class 1 and 22.84 people for CLC class 2. This corresponds to a relative population weight of 0.82 for CLC class 1 and 0.18 for CLC class 2.

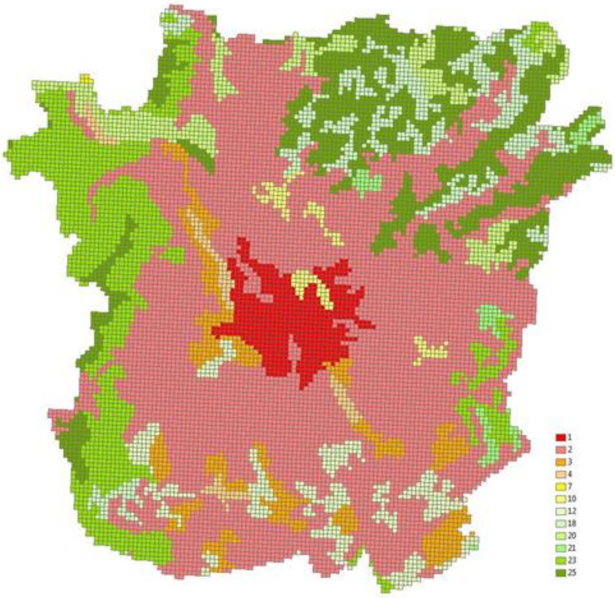

Fig. 1: Corine Land Cover classes of the city of Graz.

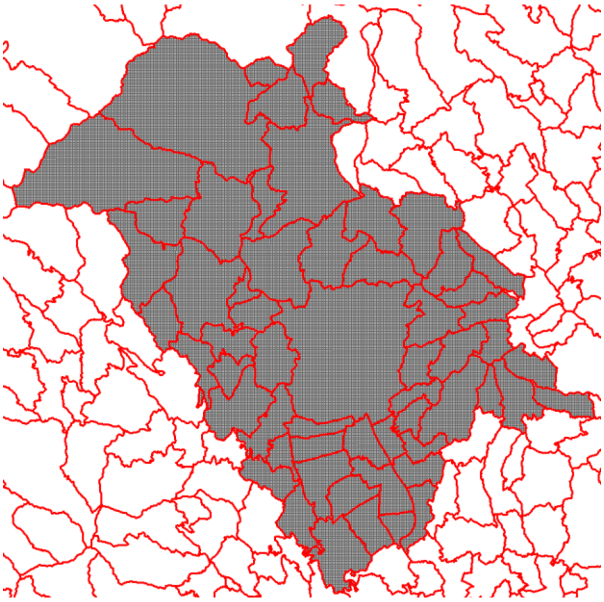

Fig. 2: The city of Graz with its surrounding 57 municipalities.

[3] In our case the null hypothesis is . The smaller the p-value, the more strongly the test rejects the null hypothesis which means that the tested variable has a significant influence on the population variable bev_{ZSP} .

ISBN: 978-3-9503110-4-4 (CD-ROM); ISBN: 978-3-9503110-5-1 (Print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI

There are 57 small municipalities surrounding Graz. Equivalent to the first regression above, we now run regressions for these 57 municipalities. The surrounding area of Graz contains 22 of the 42 possible land use classes.

This time we find that in addition to the CLC classes 1 and 2 (continuous and discontinuous urban fabric), three further CLC classes are significant for housing population, these are CLC class 18 (pastures), CLC class 20 (complex cultivation patterns) and 23 (broad-leaved forest). Even though these additional cell types are statistically significant, they house a very low number of people on average.

Population density according to our density measure is therefore calculated as follows: For each of the 259 sub-districts within the city limits of Graz we divide its population by the sum of the number of cells of class 1 and 2 within this district. Population density for each of the 57 municipalities outside of Graz is calculated by dividing its population by the sum of CLC cells of class 1, 2, 18, 20 and 23 within that area.

Therefore the formula for this density measure reads as follows:

$$D_2 = \frac{\sum_i pop_i}{\sum_j x_j}$$

where pop_i is the population in cell i for $i = 1, \dots, n$ with n as the total number of cells in the city and $\sum_j x_j$ is the built-up area of the town.

4 ENTROPY MEASURES

4.1 Measuring entropy at the district/municipality level with population data

Entropy measures of urban sprawl focus on differences in density across districts rather than the density level itself. A high entropy score implies that population density does not vary much across districts. This is interpreted as a high level of urban sprawl. A low entropy score conversely implies high variability in population density across districts and a low level of urban sprawl.

Here we use Shannon's measure of entropy. This has been used previously to measure urban sprawl by Yeh and Li (2001) in China, Sudhira, Ramachandra, and Jagadish (2004) in India and Sun, Forsythe and Waters (2007) in Canada.

Urban spawl according to Shannon's entropy measure is calculated as follows:

$$E = \sum_i p_i \cdot \log(p_i),$$

where

$$p_i = \frac{pop_i / \sum_j x_{ij}}{\sum_i \left(pop_i / \sum_j x_{ij}\right)}$$

are the weights that sum to one. The index j counts the number of land use classes, pop_i is the population, and x_i is the land area in district i. Its upper bound is $\log(n)$, where n is the number of districts. E takes the value $\log(n)$ when the population density is the same in all districts. Dividing E by $\log(n)$ provides us with an index that lies between zero and one. This normalized entropy measure makes it easier to compare entropy results between regions.

$$E_{1,2} = \frac{\sum_i p_i \cdot \log(p_i)}{\log(n)}$$

Using our two measures of density from the previous section (i.e., density measures D_1 and D_2) we obtain two different measures of Shannon's entropy $(E_1$ and $E_2)$. We calculate these entropy measures for the city of Graz, the surrounding area of Graz, and greater Graz (the latter being the sum of the city and surrounding area).

4.2 Entropy measures that do not use population data

The Shannon entropy measure can also be applied to density measures that are not based on population data. One example of such a density measure for a district is the number of built-up cells divided by the total number of cells in the district. In this case, the entropy measure takes the following form:

$$E_3 = \sum_i p_i \cdot \log(p_i)\text{ ,where}$$
$$p_i = \frac{\sum_k x_{ik}/\sum_j x_{ij}}{\sum_i(\sum_k x_{ik}/\sum_j x_{ij})}$$

are the weights summing up to one. The variable x_{ik} denotes the number of cells in district i with built-up cells k ($k = 1, 2$ for the city of Graz, and $k = 1, 2, 18, 20, 23$ for the surrounding communities of Graz) and x_{ij} is the number of cells in district i of land use class j for $j = 1, \dots, 12$ in the city of Graz, and $j = 1, \dots, 22$ in the surrounding area of the city.

5 SHAPE BASED MEASURES

Urban sprawl can also be thought of as a land use issue rather than one of population density per se. We now consider a class of methods for measuring urban sprawl that do not require any population data. This feature can be useful since obtaining population density measures can be problematic when the political boundary of a town does not coincide with its functional boundary. [4] For example, in Figure 3 we see that our town A* lies mostly within its political boundary (A). Most of its borders however lie outside its political boundaries in the political jurisdictions of B, C, and D. Data on population changes are readily available for each political jurisdiction, but these data are not geographically referenced – they do not show in which cell of each jurisdiction people live. Suppose that the town A* expands rapidly towards the north and west but not anywhere else. If we just measure population change in A, we completely miss this expansion. If on the other hand, we include all of the population growth that happened in A, B, C, and D together, we overestimate the expansion of A*, since each of these jurisdictions also has settlements that are not connected to A*. In this section we will illustrate a method to overcome this problem.

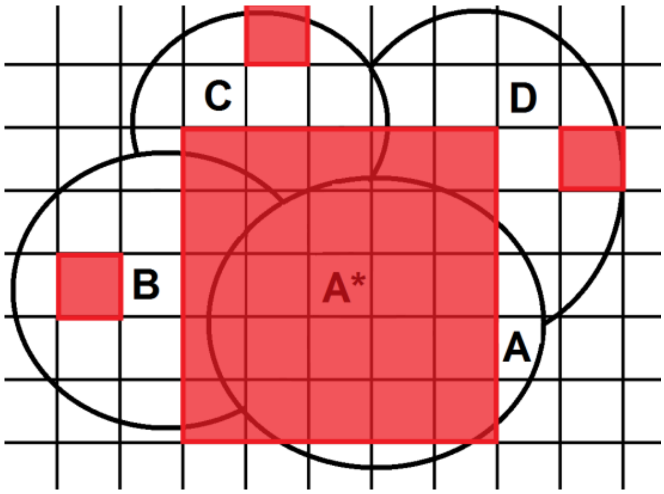

Fig. 3: Political vs. functional boundary of a town.

A class of urban sprawl measure focuses on the shape of a town's boundary. The most compact boundary is a circle. This minimizes the level of urban sprawl according to this type of measure. The more the shape departs from a circle the higher is the resulting measure of urban sprawl. Also, the wiggliness of the boundary can play a role in Fractal-type measures from this class.

The Corine Land Cover Project (European Environment Agency, 2010) provides an ideal framework for illustrating these types of methods. The landscape is divided into 100 by 100 meter cells, thus allowing the boundary to be identified at this fine level. This is a topic that we will consider in future research.

[4] The question of „where does a town end?" has various possible answers. We define the "end of town" as the end of cells with Corine Land Cover classification 1 or 2 (continuous or discontinuous urban fabric).

ISBN: 978-3-9503110-4-4 (CD-ROM); ISBN: 978-3-9503110-5-1 (Print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI

6 RESULTS

Sprawl Measures	Density Measure D_1	Density Measure D_2	Entropy measure E_1	Entropy Measure E_2	Entropy Measure E_3
Graz	2234/km²	4183/km²	0.927	0.937	0.990
Surrounding communities	128/km²	309/km²	0.917	0.932	0.979
Graz including surrounding area [5]	360/km²	735/km²	0.918	0.933	0.980
Gleisdorf	1130/km²	×	×	×	×
Leibnitz	1242/km²	×	×	×	×

Table 1: Results of Density Sprawl Measures.

7 CONCLUSION

Urban sprawl has many interpretations. Different measurement concepts capture different aspects of this complex issue. We have shown how entropy measures and land use data can be used together to capture the phenomenon of urban sprawl in new ways. Measuring sprawl at the functional settlement level rather than the political level is a big benefit for city and regional planners. Being able to do so without gathering expensive cell level population data is another benefit of our approach.

8 REFERENCES

BRAY R, VAKIL C., ELLIOTT D: Report on Public Health and Urban Sprawl in Ontario: a review of the Pertinent Literature, In Energy Technology Data Exchange World Energy Base (ETDEWEB), 2005.

BRUECKNER Jan K.: Urban Sprawl: Diagnosis and Remedies, In: International Regional Science Review, Vol. 23, Issue 2, pp. 160-171, 2000.

EWING Reid, SCHMID Tom, KILLINGSWORTH Richard, ZLOT Amy, RAUDENBUSH, Stephen: Relationship Between Urban Sprawl and Physical Activity, Obesity, and Morbidity. In: American Journal of Health Promotion, Vol. 18, No. 1, pp. 47-57, 2003.

EUROPEAN ENVIRONMENT AGENCY: Corine Land Cover Project 2006, Version 13, http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster, 2010.

FRENKEL Amnon, ASHKENAZI Maya: Measuring Urban Sprawl; How Can We Deal With It? In Envirionment and Planning B: Planning and Design, Vol. 35, Issue (1), pp. 56-79, 2008.

FRUMKIN Howard: Urban Sprawl and Public Health, In: Public Health Reports, Vol. 117, Issue 3, pp.201-217, 2002.

GLAESER Edward L., KAHN Matthew E.: Sprawl and Urban Growth, In Working Paper 9733, National Bureau of Economic Research, Cambridge, MA 02138, May 2003.

EID Jean, OVERMAN Henry G., PUGA Diego, TURNER Matthew A.: Fat City: Questioning the Relationship between Urban Sprawl and Obesity, In: Journal of Urban Economics, Vol. 63, Issue 2, pp. 385-404, 2008.

NECHYBA Thomas J., WALSH Randall P.: Urban Sprawl, In: Journal of Economic Perspectives, Vol. Volume 18, Issue 4, pp. 177–200, 2004.

NEW GEOGRAPHY: World Urban Areas Population and Density: a 2012 Update, http://www.newgeography.com/content/002808-world-urban-areas-population-and-density-a-2012-update, accessed on 27.2.2013.

SUN Heng, FORSYTHE Wayne, WATERS Nigel: Modelling Urban Land Use Change and Urban Sprawl: Calgary, Alberta, Canada. In: Networks and Spatial Economics, Vol. 7, pp. 353-376, 2007.

YEH Anthony G.O., Li Xia: Measuring and Monitoring of Urban Sprawl in a Rapidly Growing Region Using Entropy. In: Photogrammetric Engineering and Remote Sensing, Vol. 67, Issue 1, pp.83-90, 2001.

[5] The weights for the assignment are according to their relative area (relative number of cells).