Do Local Information Systems Hide the Bigger Picture? An analytical approach to measuring the strength of local boundaries.

Sarah Goodwin[1] and Jason Dykes[2]

[1]Manchester Geomatics Ltd, Beehive Mill, Jersey St, Manchester M4 6JG

sarah.m.goodwin@gmail.com

[2]giCentre, School of Informatics, City University London, EC 1V 0HB

Tel. (+ 44 20 7040 8906) Fax (+44 20 7040 8845)

jad7@soi.city.ac.uk, www.gicentre.org/jad7

KEYWORDS: Local Information Systems, polygon Wombling, boundaries, output areas, geodemographics

1 Introduction

The growth of freely available statistical data at small area geographies, and associated initiatives such as e-Government and the Freedom of Information Act 2000 have encouraged local governments and partnerships to develop web-based Local Information Systems (LIS). While the capabilities and functionality of operational LIS differ widely (Foley *et al.,* 2007), they are all constrained by scale and local boundaries. Administrative boundaries such as Local Authority (LA) boundaries are not visible on the ground and are susceptible to change, however, these are often used within LIS for decision making and analysis. Constraining data at LA boundaries in LIS could potentially restrict the information and subsequent knowledge acquired for the areas on the borders of such systems and split communities with similar characteristics. If this effect exists it may be reinforced by the growing LISfocussed approach to the consideration of geography within LAs. Here we investigate whether LA boundaries follow patterns of social change or whether they divide regions of homogenous population.

2 Methods

Two analytical techniques were used to investigate this concept in respect to the study area of Manchester and its eight surrounding LAs.

The ONS Output Area Classification (OAC) data was used to create ‘homogenous areas’ (HAs) within the defined study area, at the three OAC cluster levels of Super Groups, Groups and Sub Groups. These HAs were analysed to discover whether LA boundaries divide regions considered homogenous according to the three levels of classification.

A 'Geographical Boundary Analysis' technique known as ‘Polygon Wombling^{Õ} (Jacquez and Greiling, 2003) was used to analyse multiple census variables at Output Area (OA) level to identify abrupt demographic change. Boundary analysis techniques seek to define and evaluate *'edges of homogenous areas or zones of rapid change in a variable's spatialfield'* (Jacquez *et al,* 2000:225).

2.1. Polygon Wombling

Womble (1951) developed a method for discovering boundaries in a collection of continuous variables. Polygon Wombling (or areal Wombling) has been developed much more recently (Jacquez *et al,* 2000) and aims to identify difference boundaries - shared borders between two adjacent polygons that have dramatically different observed response values (Boots, 2001). There are relatively few academic articles relating to areal Wombling (Lu and Carlin, 2005), however, the method provides an effective means of identifying the important boundaries in social data. It also offers a unique representation of areal data enabling zone boundaries to be variously emphasized according to the strength of the demographic boundary that they represent.

During the Wombling process the polygon edges are analysed to distinguish the boundaries of abrupt variable change between neighbouring areas. Each polygon edge is assigned a *boundary likelihood value* (BLV) which measures the rate of variable change as well as the spatial change between two areas through a distance metric. Edges with higher BLVs represent abrupt changes in variable values and therefore are more likely to form part of the *'Wombled Boundary'* (BoundarySeer, 2007). The polygon edges with the highest BLVs are known as *Boundary Elements* (BEs). When edge BLVs exceed an established threshold , candidate BEs become part of the *Wombled Boundary.* While most analysts use the upper 5 or 10 percent relative BLV thresholds, an absolute BLV threshold that is determined by using the full range and distribution of all the edge BLVs is recommended. In order to investigate the difference between thresholds, the top 5, 10 and 20 percent thresholds were explored in this analysis along with an absolute threshold defined specifically to relate to the BLV distribution of the study area.

Multi-variable polygon Wombling allows boundaries that represent discontinuities in multivariate data to be identified. As demographic patterns are related to many social variables such as housing, employment and education, multi-variable polygon Wombling with small area geographical units, such as OAs, enables the identification of strong and weak social barriers at the local level. In order to generate a comprehensive picture of social patterns the forty-one Census 2001 variables used to create OAC were employed in multi-variable Wombling.

3.Results

The locations of the Wombled boundaries derived using an absolute threshold of 19.2% are shown in Figure 1. The LA boundaries are more likely to be identified as Wombled boundaries than other edges in the study area as a whole for each of the three relative thresholds - a third of all LA edges are within the top 20% barriers of abrupt social differences as measured through the multi -variable Wombling (Table 1). X^2 tests give us confidence that these differences are unli kely to be random in all cases (see Table 1). The OAC analysis supports this suggestion that LA edges are associated with social difference as a small percentage of LA boundaries cut across HAs. However, both analyses show that LA boundaries can divide areas that have similar social characteristics and so the consideration of LAs as distinct units may in places "*hide the bigger picture*". Importantly, there are spatial trends to this pattern – the answer to our question is geographically variable in our study area. HAs cross LA boundaries most frequently in specific areas, particularly around Manchester City Centre, indicating that the strength of the LA boundaries varies across the region. Note also that the OAC data is based on clusters generated at a national level, whereas the raw data for the Wombling is standardised within the study area to ensure that results are locally discriminating. The multi-variable Wombled boundaries identify some edges that the OAC analysis does not and these occasionally divide OAC HAs, demonstrating that even though the two techniques use the same data they detect social discontinuity with different precision.

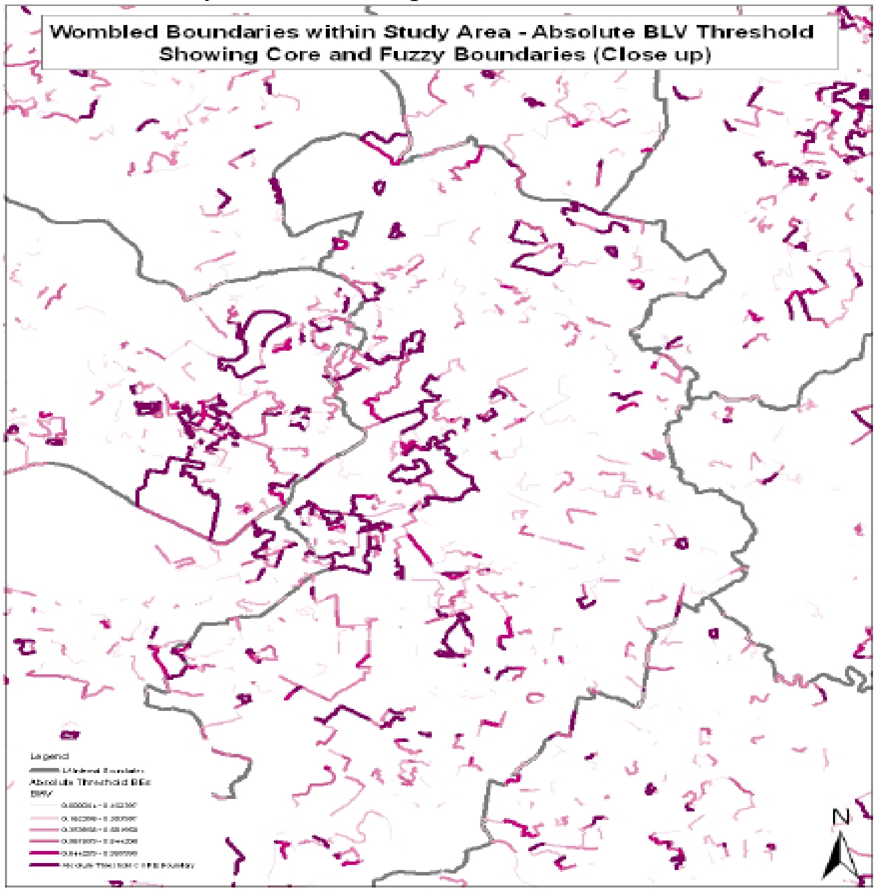

Figure 1. A close up of the Manchester City Centre area illustrating the absolute Wombled boundaries using a fuzzy representation - the darker the boundary the higher the BMV.

	threshold	expected	observed (wombled)		X^2	p	significance and trend	
All LA edges	5%	27.8	41	7.4%	6.27	**1.23%**	**at 5%**	more than expected
	10%	55.6	89	16.0%	20.07	**0.00%**	**at 1%**	more than expected
	20%	111.2	174	31.3%	35.47	**0.00%**	**at 1%**	more than expected
Manchester LA edges	5%	15.8	34	10.7%	20.79	**0.00%**	**at 1%**	more than expected
	10%	31.7	64	20.2%	32.92	**0.00%**	**at 1%**	more than expected
	20%	63.4	114	36.0%	40.39	**0.00%**	**at 1%**	more than expected
Other LA edges	5%	11.9	7	2.9%	2.05	15.22%	no	as expected (a few less)
	10%	23.9	25	10.5%	0.05	82.18%	no	as expected
	20%	47.8	60	25.1%	3.12	7.76%	no	as expected (a few more)

Table 1. Observed and expected Wombled edges for all LA edges and Manchester / non-Manchester LA edges at three thresholds. Expectations calculated under the assumption that Wombled edges will be distributed randomly across all edges. Significance evaluated with X^2 test (no Yates' correction).

Comparison of edges on the Manchester LA boundary with other LA edges helps us identify where consideration of LAs as distinct units may "*hide the bigger picture* ". Wombled edges occur much more frequently on the Manchester LA boundary than would be expected if they were randomly distributed, with 10.7% of the Manchester edges within the top 5% BLV threshold. Edges with the highest BLVs occur particularly frequently here. However this trend is not apparent when all other (non-Manchester) LA edges are considered. Indeed whilst the result is not significant, fewer Wombled edges are observed than expected in this case. The polygon Wombling provides no evidence that other LA edges constitute difference boundaries between socially distinct neighbours.

The four distributions of BLVs for a) all (OA) edges, b) all LA boundary edges, c) all Manchester LA edges and d) all non-Manchester LA edges are shown in Figure 2. The distribution of BLVs in each of the histograms shows some similarities. The outliers to the skewed distribution represent the edges of very abrupt change. This positive skewed distribution indicates that very abrupt change in social data is in fact relatively rare across the study area. Our overall findings suggest that it is more likely to occur across LA boundaries than across non-LA boundaries, but when Manchester LA boundaries and non-Manchester LA boundaries are considered separately this is only true in the former case . This trend is reflected by the less dramatic decay evident in Figure 2a than 2b. The strength of the Manchester LA boundary is apparent when comparing Figures 2c and 2d.

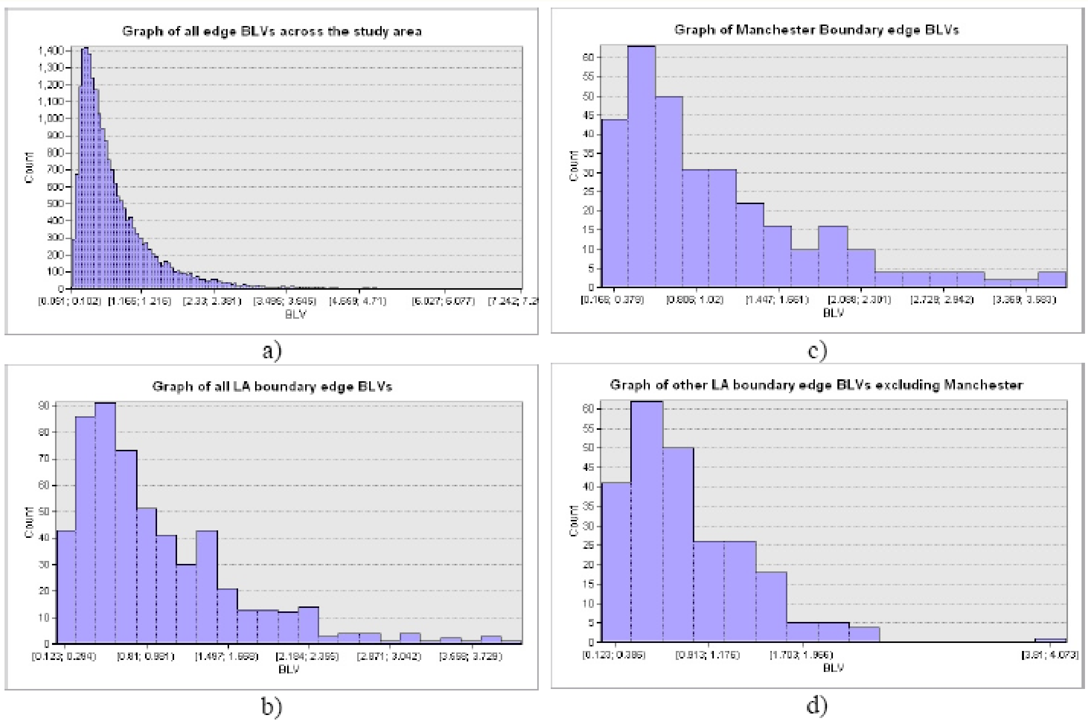

Figure 2. Histograms illustrating the distribution of BLVs for: a) all the edges in the study area, b) All LA boundary edges, c) the Manchester LA edges, d) the non-Manchester LA boundary edges

A number of specific examples of Wombled boundaries were investigated. Some occur at identifiable physical barriers, many following main street networks which are particularly well illustrated in and around the city centre of Manchester (Figure 3). *'Wombled Polygons',* whereby an area is highlighted as an island of abrupt social differences when compared to surrounding neighbours, are particularly interesting. For example, the area of Salford Quays, which was newly developed at the time of the 2001 Census through heavy investment in urban regeneration, is clearly identified as an island through Wombled boundaries.

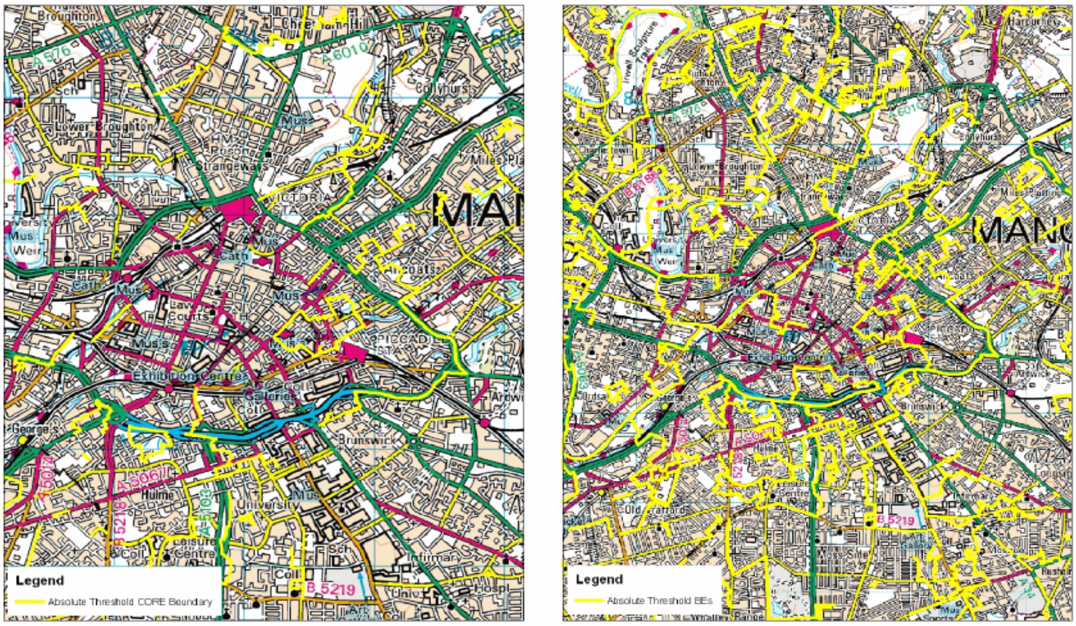

Figure 3. Absolute BLV threshold showing the core and fuzzy Wombled boundaries (in yellow) on a
1:50,000 OS raster map. © Crown Copyright/database right 2007. An Ordnance Survey/EDINA
supplied service.

4. Conclusions / Implications

Polygon Wombling allows us to measure, visualize and compare discontinuity between zones. It provides a unique perspective on area data and the opportunity to remove areal units from maps that use line thickness and colour to reflect boundary strengths (Figure 1). The degree of social difference revealed through Wombling is a particularly interesting concept for the analysis of the geography of social change in the urban environment as it enables abrupt social differences that may lead to possible tensions to be identified. The analysis of data over time could subsequently lead to a better understanding of social barriers and aid the monitoring of regeneration projects in urban environments.

In the context of LIS, we have shown that the answer to our research question is geographically variable, drawing attention to some of the potential risks associated with a discontinuous approach to data use. The high frequency of boundaries with BLVs along the Manchester LA is very unlikely to be due to random processes and the strength of these boundaries in comparison to the surrounding LA boundaries may be of interest for social analysis within the area. Whilst some of the abrupt demographic changes are due to physical barriers on the ground the frequency with which Wombled boundary edges fall on the Manchester boundary indicate that the political boundaries themselves could influence these social barriers. Where LA boundaries do not represent discontinuities in data about those who they divide, as is the case in the non-Manchester LA boundaries, then LIS may be hiding a significant 'bigger picture' by forcing administrative divisions where social divisions do not exist. Disregarding this may even contribute to social divisions in the long run. These results and the analytical techniques used in this research project are applicable to a number of areas of social research such as urban regeneration, neighbourhood renewal and local boundary definition.

5. Acknowledgements

We acknowledge Dr Robert Barr, Senior Lecturer of GIS, University of Manchester and Managing Director of Manchester Geomatics Ltd. and Dr Jo Wood, Senior Lecturer, City University London for advice and support that have contributed to this work.

References

Boots, B. (2001) Using local statistics for boundary characterisation. *GeoJournal,* 53, 339-345.

BoundarySeer (2007) Online Help - Boundary Detection - Wombling. Available online at http://www.terraseer.com/products/bsr/help/index.htm#Wombling/About_wombling.htm accessed on 2 1/12/07. TerraSeer.

Foley, P., Alfonso, X. & Wiseman, I. (2007) Local Information Systems: A review of their role, characteristics and benefits. IECRC for the Department for Communities and Local Government.

Jacquez, G. M. & Greiling, D. A. (2003) Geographic boundaries in breast, lung and colorectal cancers in relation to exposure to air toxics in Long Island, New York. *International Journal ofHealth Geographics,* 2.

Jacquez, G. M., Maruca, S. & M.-J. Fortin (2000) From fields to objects: A review of geographical boundary analysis. *Journal of Geographical Systems,* 2, 221-241.

Lu, H. & Carlin, B. P. (2005) Bayesian Areal Wombling for Geographical Boundary Analysis. *Geographical Analysis,* 37, 265-285.

Womble, W. H. (1951) Differential Systematics. Science, 114,315-322.

Biographies

Sarah Goodwin is a GeoData Consultant at Manchester Geomatics Ltd. She works closely with local government clients to ensure the best use of their spatial data. This includes the development of Local Information Systems. She recently received an MSc in Geographic Information Systems from City University London.

Jason Dykes is a Senior Lecturer in Geographic Information and member of the giCentre at City University London. He has research interests in geovisualization and is involved in work that is using and evaluating geovisualization techniques in local government.