A Comparison between using Ordnance Survey Landline™ and Mastermap™ Products to derive Land Cover Data for use in Socio-environmental Area Classification

N. S Walford[1], R. P. Armitage[2]

1. School of Earth Sciences and Geography,
Centre for Earth and Environmental Sciences Research,
Kingston University, Penrhyn Road, Kingston upon Thames. KT1 2EE
Tel: +44 (0)20 8547 7512
Fax: +44 (0)20 8547 7497
Email: N.Walford@kingston.ac.uk

2. School of Environment and Life Sciences,
Research Institute for the Built and Human Environment,
University of Salford, Salford. M4 5WT
Tel: +44 (0)161 295 5456
Fax: +44 (0)161 295 5015
Email: R.P.Armitage@salford.ac.uk

1. Introduction

It is 30 years since the first national area classifications were produced using statistics from the 1971 Population Census for local authorities, parishes and wards in Britain within the Planning Research Applications Group in the Centre for Environmental Studies in conjunction with the Office of Population Censuses and Surveys (Webber and Craig, 1976; Webber 1977; Webber and Craig, 1978). Successive census enumerations have seen such national classifications extended to smaller areas, enumeration districts and most recently the postcode-based Output Areas, and updated by the inclusion of more and exclusion of less relevant demographic and socio-economic statistical counts (Sleight, 1995; Wallace and Denham, 1996). Such area classifications have helped to establish geodemographics as a viable tool in the armoury of commercial organisations marketing and selling goods and services to consumers, and in the range of techniques available to those in the public sector concerned with formulating area-based policies and allocating resources targeted towards certain sectors of society. Unfortunately the term 'postcode lottery' has tarnished the image of such an approach, especially with reference to public services, although the number of companies operating in the commercial arena shows signs of increasing (Sleight, 2004).

2. Background

It is characteristic of these classification systems that, although usually based on complex multivariate statistical techniques, they encapsulate the aggregate demographic and socio-economic nature of small areas by means of a simple, relatively intuitive, thumbnail description. Paradoxically, although these descriptions often refer to both the social and environmental character of the areas in which people live by employing such terms as 'leafy suburb' or 'inner city terraced housing', environmental indicators have not been incorporated

into the analysis. The landscape of leafy suburbia is determined not so much by means of identifying the presence of housing in tree-line avenues as by recording inhabitation by 'double income families with 2.4 children'. In other words there is an implicit inference that people with certain combinations of demographic and socio-economic characteristics reside in particular types of environment.

The authors' previous research examined the effect on area classification of combining socio-economic data with environmental information, derived by using Ordnance Survey LandLine™ in conjunction with digital aerial and satellite and imagery (Walford and Armitage, 2005). The Ordnance Survey's subsequent re-engineering of its unstructured digital point, line and text topographic data into a topologically structured and polygonised database known as MasterMap™ with general land cover/use attributes provides the opportunity of making a direct comparison between these two spatial data formats. This paper presents the results of this comparison and considers the additional types of data that might enhance socio-environmental area classification. MasterMap™ has been available to commercial and public sector partners since 2004, and will come online to the UK Higher Education community through the Digimap service in 2007. The authors have been granted access to MasterMap data in advance of this release to the HE sector for Salford and a contrasting local authority (Colchester).

3. Methodology

There are two main issues to be addressed when devising a suitable methodology for classifying small areas. The first concerns the choice and format of statistical counts, including whether these are absolute or relative and the nature of any standardisation to be applied. The second relates to the types of multivariate analyses, including both data reduction (e.g. factor or principal components analyses) and/or classificatory techniques (e.g. cluster analysis), to which the input data will be subjected (Charlton, 1985). Most geodemographic classifications have been hierarchical with families, groups and clusters of areas possessing a broad, intermediate and detailed degree of similarity. In all cases such analyses may potentially suffer from commission of the ecological fallacy and from the Modifiable Areal Unit Problem, although these issues may be alleviated by judicious variable selection and careful evaluation of the classification.

The authors' previous research derived a set of land cover variables for census Output Areas in a case study local authority (Salford) within the Manchester Metropolitan area by applying image processing and spatial analytic tools with respect to aerial and satellite imagery in conjunction with OS LandLine™ data. These variables were combined with demographic and socio-economic statistics from the 2001 UK Population Census. Two forms of multivariate analysis were applied to test the impact of including environmental variables on area classification.

The first analysis replicated the National Statistics Office's methodology for classifying wards (ONS, 2004), augmenting the 43 demographic/socio-economic standardised counts with nine land cover variables, including three mixed types. This was essentially a two-stage methodology involving an iterative allocation-reallocation (K-means) clustering to produce an optimum classification from an initial set of random cluster centres. This was followed by a classical hierarchical cluster analysis using Ward's method, which was further refined to ensure each area (ward) was assigned to its correct subgroup. Second, an ordination approach was applied to explore the dimensionality of the socio-economic and land cover data.

Ordination can be achieved by either direct or indirect means, in the former case to identify the variation in a single dataset and in the latter to analyse trends in two datasets and establish the nature of the relationship between them by deriving a linear combination of one set of variables to explain the other set (canonical ordination). Redundancy Analysis (RDA) was applied with the socio-economic data entered as the "species" variables and the land cover ones as the "environmental" variables.

4. Results and Analysis

The present research has generated a similar set of land cover variables using the MasterMap™ data first using its land cover/use attributes and then with the addition of information from the ancillary aerial and satellite imagery, and repeated the previous classificatory analyses. This provides the opportunity to compare the classifications produced by both types of multivariate analysis as summarised by the analytical framework represented in Table 1.

43 census counts plus	PCA/Cluster Analysis	Redundancy Analysis
Land cover from aerial/satellite imagery and LandLine™	✓	✓
Land cover from MasterMap™	✓	✓
Land cover from MasterMap™ and aerial/satellite imagery	✓	✓

Table 1. Analytical framework

The results suggest that the MasterMap™ polygon data with its land use/cover attributes adds an environmental component to the classification. However, the relatively broad nature of these attributes is usefully supplemented by the additional information obtained from aerial imagery, for example distinguishing between concrete, tarmac and other types of hard surface material.

5. Conclusion

The paper presents results of the research as comparisons between the two multivariate analyses using the LandLine and MasterMap data as a means of delimiting and quantifying the extent of land cover types and environmental features in an urban mosaic. The results provide the opportunity to examine the potential impact of using MasterMap as opposed to LandLine data in applied academic research. The extension of the previous work to include a contrasting local authority (Colchester), which overall has a less urban character than Salford by including a rural hinterland of small towns and villages, also provides the opportunity to test out the methodologies in a different environmental context.

6. References

CHARLTON, M., OPENSHAW, S. and WYMER, C., 1985, Some new classifications of census Enumeration Districts in Britain: a poor man's ACORN, Journal of Economic and Social Measurement, 13, pp. 69-96.

CRAIG, J., 1985, A 1981 Socio-economic Classification of Local and Health Authorities of Great Britain, SMPS 48 (London: HMSO).

OPENSHAW, S., 1989, 'Making geodemographics more sophisticated', Journal of the Market Research Society, 3(1), pp. 111-132.

NATIONAL STATISTICS, 2004, National Statistics 2001 Area Classification for Statistical Wards.http://www.statistics.gov.uk/about/methodology_by_theme/area_classification/wards/default.asp

SLEIGHT, P., 1995, 'Neighbourhood watch: geodemographic and lifestyle detail the UK', Mapping Awareness, 9(6), pp.18-21.

SLEIGHT, P., 2004, New to the neighbourhood, Database Marketing, 57, pp. 35-40.

WALFORD, N. S. and ARMITAGE, R. P., 2005, Census-based socio-economic area classifications: An environmentally enhanced version, GISRUK05, Glasgow, April 2005.

WALLACE, M. and DENHAM, C., 1996, The ONS Classification of Local and Health Authorities of Great Britain, SMPS 59 (London: HMSO).

WEBBER, R., 1977, An Introduction to the National Classification of Wards and Parishes, Technical Paper 23 (Centre for Environmental Studies: London).

WEBBER, R. J., 1989, Using multiple data sources to build an area classification system: operational problems encountered by MOSAIC, Journal of the Market Research Society, 31, pp. 103-109.

WEBBER, R. and CRAIG, J., 1976, 'Which local authorities are alike?' Population Trends, 5, pp.13-19.

WEBBER, R. and CRAIG, J., 1978, A Socio-economic Classification of Local and Health Authorities, SMPS 35 (London: HMSO).

Biographies

Nigel Walford is Professor of Applied GIS at Kingston University. Research interests relate to applying GIS to the mapping and analysis of geodemographic and agri-environmental information. Recent and in press journal papers published in **Journal of Geographical Information Science, Journal of Rural Studies,** and **Population, Space and Place**.

Richard Armitage is a Lecturer in GIS at the University of Salford with research interests applying GIS and remote sensing to landscape mapping and characterisation. Publications include papers in **International Journal of Remote Sensing**, **Progress in Physical Geography** and **Vegetation Mapping: From Patch to Planet** (Alexander and Millington (editors)).