Researching long-run trends in South East England 1841-2011 for the European Union and Greater London Authority

Paula Aucott[*1] and Humphrey Southall[†1]

[1]Department of Geography, University of Portsmouth

March 13, 2015

Summary

This paper describes the sources, methods and preliminary results of two related projects on historical census data funded by government bodies for policy purposes. Both required data for diverse historical reporting areas to be redistricted to a single set of modern units. All redistricting is done by a simple vector overlap method, but this requires boundary data for both the modern and the historical units; and, as far as possible, that the historical units be more detailed than the modern ones. Even for recent periods, locating boundary maps is often much harder than locating statistics.

KEYWORDS: historical GIS, policy relevance, redistricting, census

1. Introduction

Predicting long-term trends decades ahead requires data stretching decades back. Through its censuses the UK gathers just such data today (2011), but also goes right back to 1801. However, the analytic potential is poorly exploited: "modern" census research generally looks just one or two censuses back, while "historical" census research focuses on the period 1851-1911 when reporting geographies were consistent.

This paper presents two highly applied census projects. The first, funded by the European Union, is creating time series 1951 to 2011 for the total populations of the 8,941 Wards of Great Britain as used by the 2011 census. The second, funded by the Greater London Authority (GLA), is redistricting a diverse data from censuses 1801-1961, creating consistent data for the current London Boroughs, for London's Central Activity Zone (CAZ) as defined by the GLA, and for the overall GLA area; but we focus here on constructing consistent data on London's industrial structure, converting diverse historical classifications to Standard Industrial Classification 2007 (SIC).

The research described here is ongoing but contractually must be completed by the GISRUK conference, so we concentrate here on sources and methods with findings presented at the conference.

2. Estimating historical populations of 2011 Wards

Earlier work for *Vision of Britain* (www.VisionofBritain.org.uk) included redistricting to modern units, but only to Britain's 408 districts. Redistricting to a geography twenty times more detailed requires more detailed input data, and different sources are used for each census year.

2011 and 2001 are not our concern, data being available from the Office for National Statistics. For

[*] Paula.Aucott@port.ac.uk
[†] Humphrey.Southall@port.ac.uk

1991 and 1981 population counts and vector boundaries are available for 103,419 (1991) and 105,598 (1981) English Enumeration Districts, plus Welsh and Scottish EDs, so redistricting to 8,941 British Wards is unproblematic. The historical statistics are associated with the boundaries of their contemporary administrative units and converted to the required output projection (where necessary). The historical polygons are spatially joined to the boundaries of the modern polygons, the population figures are interpolated and finally the boundaries are dissolved to give modern units with weighted statistics.

1971 is more complex. After significant investigation no digital mapping of real ward or enumeration district boundaries for this year has been found to exist. Instead we use the closest alternative, Thiessen polygons for Wards which were created by aggregating Census Enumeration Districts which themselves were generated from a centroid point dataset. The same processing method as above was employed, although the artificial boundaries meant there was some mismatching to manually correct.

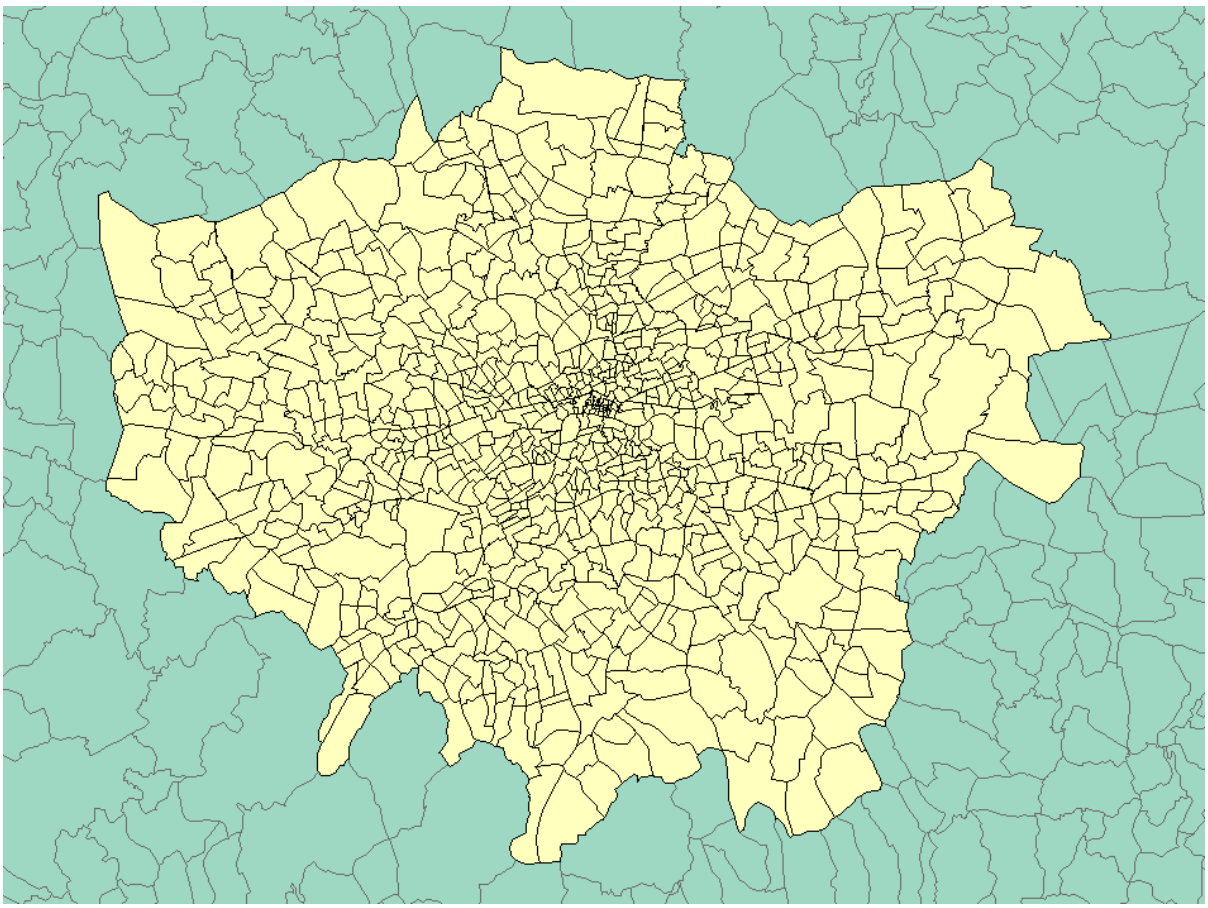

Figure 1 The Wards of Greater London in 1961, plus Civil Parishes in adjacent areas

1961 was the first census year that many urban areas were "unparished", meaning that unless we map Ward boundaries, cities as large as Birmingham, with over a million people, are single polygons. We addressed this issue by seeking ward maps from the modern councils, archives and local studies libraries. The GLA provided a map covering all London Boroughs and the new coverage above demonstrates the much higher density of polygons for wards in central London in comparison to the larger civil parish polygons in adjacent counties.

For the largest towns where we cannot obtain boundary maps, we are locating Wards as points then building Thiessen polygons. Overall, this is the most detailed mapping of the 1961 census ever done, as 1960s analyses either mapped Wards as points (Ministry of Housing and Local Government, 1966) or dealt only with parishes (e.g. Osborne, 1966). Although the current analysis covers only total

populations, this GIS could map the 1961 census microdata being restored by an Essex University project.

Figure 2 is the earliest presentation of our final results, combining the redistricted population counts for all four dates to identify, for each 2011 ward, the inter-censal period of greatest growth. We are working to add 1951, 2001 and 2011, but this is already a map which could not have been created by any other method, and brings out the complex geography of Greater London.

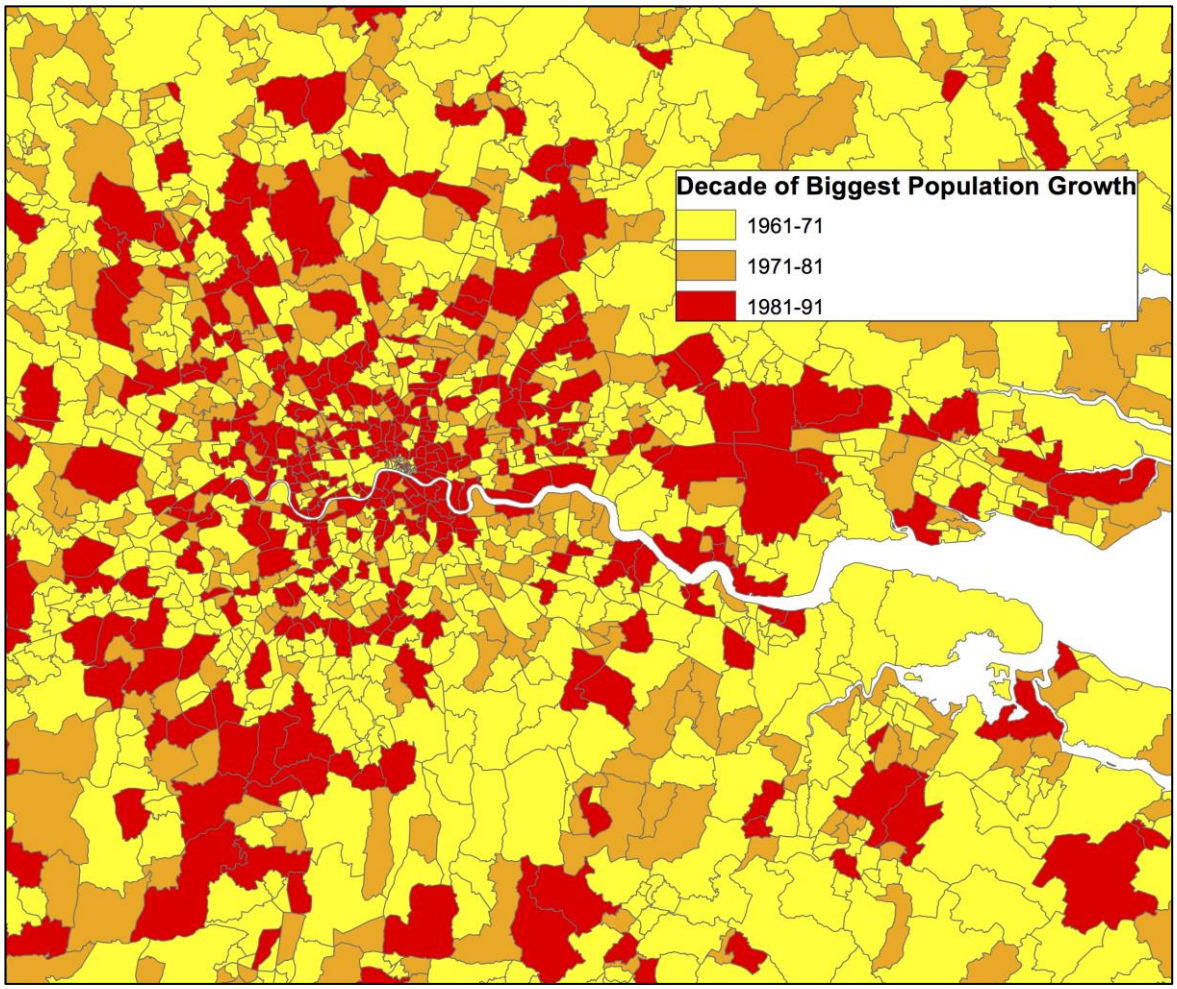

Figure 2 2011 wards for London area, showing decade of greatest population growth 1961-71-81-91

3. Longer-run trends in Greater London

A wide range of census variables are available via the Small Area Statistics from 1971 onwards, and similar data are becoming available from transcriptions of individual level census returns via Essex's Integrated Census Microdata system (http://icem.data-archive.ac.uk). However the latter is legally limited to data over a hundred years old, so long run overall perspectives must draw on the published census reports. The Greater London Authority need data on industrial structure, so our work draws on four censuses prior to born digital data.

Firstly 1841, the data existing only for Ancient Counties and thirteen specific areas within London, so results will be limited to the GLA area and the CAZ. However, the occupational classification is very detailed so mapping to the SIC is straightforward if tedious.

For 1881 data come from microdata and are held by parish and sub-district; and we have a GIS coverage for the latter for London.

The above data are for occupations and by place of residence. The first census to record place of work

and separately tabulate "industry", i.e. employer's business, was 1921, but unfortunately sub-county tabulations are not available within London, so we use 1931 data for industry, tabulated by residence. For 1951 we use an industry classification tabulated by place of work, but must still redistrict and map to SIC 2007.

4. Conclusion

The Great Britain Historical GIS project has been working for over twenty years with historical census data. Our experience shows it is quite possible for historical GIS research to achieve significant non-academic "impact", and to draw on a wider range of funding, but significant adaptations in approach are necessary. Firstly, research must come up to the present, and achieving long runs of data requires new skills: disinterring obscure statistical datasets; locating even more obscure boundary maps; even negotiating copyright.

Secondly, there is a far greater requirement among both policy makers and the general public for local time series than for maps or other purely cross-sectional presentations and analyses. Further, policy makers almost always require time series for modern units, even though it is generally easier to redistrict modern small area data to less detailed historical units. This means that once the above ingredients are assembled GIS techniques must be used to redistrict these diverse data sets to a single constant geography.

There has been much research in geographical information science developing complex redistricting algorithms, but we use simple vector overlay – "cookie cutter" – methods for two reasons (Simpson, 2002). Firstly, the improvement in accuracy from more complex methods is limited, and finding more detailed historical datasets used our time better. Secondly, these simpler techniques are much easier to describe and justify to a non-technical audience. Our focus is on real world use, both the public bodies funding the work described here and the general public accessing our web site: *Vision of Britain* had 1.65m. "unique visitors" in 2014 and the most accessible statistical content is redistricted data for modern districts.

5. Acknowledgements

This research is funded by the European Union and the Greater London Authority. We are also grateful to the many local authorities who have provided us with 1961 ward boundary maps.

6. Biography

Paula Aucott is a Senior Research Associate and doctoral candidate in the Department of Geography, University of Portsmouth; she has Masters degrees in English Local History and in GIS, and is the project manager for the Great Britain Historical GIS. Humphrey Southall is Professor of Historical Geography at the University of Portsmouth and director of the Great Britain Historical GIS.

References

Ministry of Housing and Local Government (1966) *Population change 1951-1961 by wards and civil parishes: compiled from the 1961 census. Scale 1:625,000.* Ordnance Survey, Southampton.

Osborne R H (1966) *Atlas of Population Change in the East Midland Counties 1951-1961.* Department of Geography, University of Nottingham.

Simpson L (2002). Geography conversion tables: a framework for conversion of data between geographical units. *International Journal of Population Geography*, 8(1), 69-82.