SCOTTISH NEIGHBOURHOOD STATISTICS

NEIGHBOURHOOD DEFINITIONS PROJECT

Phase 2, Stage C

Robin Flowerdew, Zhiqiang Feng

School of Geography and Geosciences
University of St Andrews
St Andrews, Fife KY16 9AL

30 September 2002
Report on Stage C

Introduction

In Stage B, a methodology was established to construct a system of data zones to be implemented across Scotland. This methodology was based on the resulting zones having a population between 500 and 1000 at the time of the 1991 census, being composed of complete Output Areas as used in the 1991 Census, and nesting within local authority boundaries. Other factors involved in definition of these data zones included shape, social homogeneity as measured by the Townsend index, and accordance with natural and environmental boundaries. Non-denominational primary school catchment areas were used as a starting point in zone definition; they were approximated by output area boundaries, but were liable to be split or amalgamated according to whether their populations were above 1000 or below 500. No account was otherwise taken of perceived community boundaries.

Procedure

The task of defining data zones for the whole of Scotland was necessarily a large and complex one. It required maps of school catchment areas, which were supplied by local authorities via the Scottish Executive. Some of these were provided quickly while others took some weeks to arrive. In addition, six authorities did not have catchment area boundaries available in digital form. Orkney Islands provided a map of primary school catchment area boundaries. Highlands produced digitised boundaries, subregion by subregion, while the zone definition work was in progress.

If digitised catchment area boundaries were not available, a methodology based on Voronoi polygons was used. The point locations of non-denominational primary schools were obtained based on their postcodes, and polygons were constructed around the schools such that all points were assigned to their nearest school. These polygons were then approximated by output areas and the procedure was followed as before.

Although the procedure was operationalised in the GIS package ArcInfo, there was still a need to make subjective decisions based upon a view of the relative importance of the various factors considered. Attempts were made to speed up the process by writing and applying short programs intended to help in the selection of data zones which met the criteria relating to size and shape. It remained the case, however, that there were many ways in which the output areas in a local authority could be combined into data zones, and the method by no means guarantees that the best solution will be reached.

The work was done one authority at a time. In general, local authorities were processed in approximately the order in which their digitised primary school catchment areas were received. Most of the work was done during August, although the last ones were not completed until early September.
SCOTTISH NEIGHBOURHOOD STATISTICS

NEIGHBOURHOOD DEFINITIONS PROJECT

Phase 2, Stage D

Robin Flowerdew, Zhiqiang Feng, James Pearce

School of Geography and Geosciences
University of St Andrews
St Andrews, Fife KY16 9AL

11 November 2002
Report on Stage D

Scottish Neighbourhood Statistics: Summary of Responses

Following the completion of definition of data zones for the Scottish districts, the data zones and their boundaries were sent out to the various local authorities for their comments. This document is intended to draw out the major themes from the responses that we have received from the 32 regional councils.

Stage D was intended to be the first round of consultation, involving local authorities. First, a letter was sent to Chief Executives of all local authorities within Scotland, telling them about the exercise and asking them to expect a mailing from St Andrews giving the data zones proposed for their areas. Shortly thereafter, sets of data zones were sent out on a CD to the Chief Executives, with some GIS software to help read the data for those without appropriate software already available. There was a covering letter, which asked for comments within approximately three weeks, and a description of the methodology used in defining the zones. The maps sent out consisted of the output area boundaries and the data zone boundaries superimposed on top of them. Representations of the boundaries in .jpg format were also available on request.

Because production of the data zones was a process extending over a few weeks, the data were sent out in three batches of approximately equal size. The dates at which the zones for particular local authorities were sent out were related to how quickly they had sent in primary school catchment area boundaries. Authorities in the first set were asked to make comments by 20 September, the second set by 27 September, and the third set by 4 October. In several cases, local authorities encountered problems in reading the data, which may have delayed responses; we were able to help them resolve these problems.

In addition to the mailings to Chief Executives, copies of the data zone maps were also sent to members of the Scottish research and intelligence network.

Responses

Some of these authorities had contacted us to discuss the format in which we had sent out the definitions of data zones. In some cases, authorities asked for more time in which to submit their responses; in all cases, this was granted, and in most such cases a response was subsequently received. There was, however, concern (8, 14) that substantially more time was needed for consultation with individual authorities. No reminder letters were sent.

In addition to the local authorities, we also received responses from Communities Scotland and from the Glasgow and Clyde Valley Joint Structure Planning Committee. In collaboration with Fife Council, we displayed posters describing our work at the Fife Research Fair on 18 September 2002; this generated verbal comments from several people attending the event, mainly public sector employees.
Summary of responses

We have tried to pull out themes from the 21 written responses that we have had. It should be noted that in addition to these themes there are some comments and criticisms that are exclusive to a particular council. These comments are specific points that usually relate to errors that have occurred due to anomalies in the OA boundary data set or in the digital council area boundaries. In most cases these are points that can be easily solved by making small adjustments to the zones.

The depth of the responses varies greatly but usually the greater the depth the more the criticism of the methods. However, a number of councils provided short and positive responses (e.g. 13, 15).

Use of 1991 Output Areas

A common criticism made by many councils (1, 2, 3, 5, 6, 7, 8, 9, 12, 14, 16, 19) was that 1991 OAs had been used, instead of waiting for the 2001 areas and population data to become available. It was pointed out that it was hard to comment on the methodology and results when the zones reflected 1991 statistics. Many councils suggested that the building of new estates since 1991 had made many zones inappropriate, not least because the population in such zones may no longer be less than 1,000.

Compatibility with other boundaries

A further issue that raised concern was that the zones were not designed to be compatible with many other geographical boundaries for which councils disseminate information. Examples would be community areas, and political boundaries such as electoral ward, SIP and BNSF boundaries. (1, 2, 3, 6, 7, 8, 10, 11, 12, 16, 18, 19).

These councils differed in their opinion on how this issue should be addressed. Some feel that the process should be started again with constraints built in to force the zones to nest within other geographies whilst others request that they retain the discretion to adjust some boundaries to allow them to reflect local geographical units.

Geographical features

A number of councils suggested that some zones were inappropriate because they did not recognise features in the landscape such as motorways, other major roads, rivers and railway lines or known community boundaries (8, 10, 11, 16, 18, 19). In some cases, our methodology did not allow this, because Output Area boundaries did not respect these features. In some cases, however, it was clearly possible to adjust boundaries to create areas that had a better fit to these landscape features. Similarly, some councils gave examples of zones which were socially diverse, where boundaries could easily be adjusted to give more homogeneous units (14).

Special Output Areas

Some councils suggested that ‘special’ OAs should be treated differently as they are currently ignored (1, 8, 14). Some councils felt that special OAs were important
components of a zone. Special OAs may have sufficient population that omitting them in data zone populations would mean that some areas deemed too small in population would actually meet the 500 population criterion, and that other areas apparently between 500 and 1000 would actually exceed the maximum population of 1000. Furthermore, we were reminded that there are no special OAs in the 2001 census.

Quantity of zones

There were some diverse opinions on whether the number of zones (and hence the population thresholds) were appropriate.

Some councils felt that the size of the threshold was too big and therefore there were not enough zones to reflect important localised heterogeneity (4, 11, 17) although some felt this was only true in the rural areas (19). Furthermore, other councils felt that the range was too big (4) or was arbitrary (12) or that seeking approximate population equality was inappropriate (17).

One council (10) felt that there were too many zones on the grounds that the time and resources necessary to produce data for these zones would be too great. However, another council felt that the number of zones and the threshold was appropriate, although they did suggest that it would be useful on occasion to have smaller units for specific areas. The concerns about the extra workload were echoed by one other rural council (2).

Most councils did not comment on this important point which would suggest that they are happy with the threshold and the number of zones in their area.

Naming of zones

There was no strong consensus on how to name the zones although most suggestions were anonymous numbering systems such as the one used by the OS (6). It was recognised as a potentially problematic aspect of the exercise (14).

School catchment areas

Many councils had opinions on whether school catchments were a suitable starting point for defining the data zones. Some felt that school catchment areas were either an appropriate starting point or at least the best available starting point (3, 4, 7, 13). Other councils felt that Primary School catchment areas were an inappropriate starting point for defining data zones because they were too transient (8, 12, 19).

Islands

It was suggested that it was inappropriate to link islands with the mainland in one zone and that it was preferable not to join islands due to the heterogeneity between them (4). Communities Scotland, making a similar point, argues that islands should be amalgamated with other islands rather than with the mainland.
Shape

Only one council (10) commented that the zones were irregular and inappropriate in shape. In most cases, such problems arise from unusual shapes possessed by the Output Areas, and care was taken in constructing the data zones to make their shapes as compact as was compatible with the other criteria.

Other issues

Some of the comments have been concerned, not with the data zones themselves, but with the use to which they are to be put. One common question was whether the zones would be put into operation in their current form (or with changes suggested by local authorities in the consultations), or whether they would be revised on the basis of the 2001 census. Also, how often, and using what criteria, should the data zones be revised and updated?

Other questions include some which reflect problems in the boundary data sets available to us, for example where some areas have been omitted or incorrectly included. This is usually because the local authority boundaries have changed since 1991, and our data did not follow the updated boundaries. Communities Scotland also drew attention to problems arising from changes in unit postcode centroids, suggesting postcodes were geocoded using data for a specified base year.

Concern was expressed that it would be difficult to ensure that there is consistent local input to the process of defining neighbourhoods across the country (14). Lastly, the documentation we supplied was criticised for not being sufficiently explicit about the method, in particular the relative importance of shape, social homogeneity, and natural boundaries (14).

Conclusion

The process of consultation provided useful feedback on the zone design exercise, and seems to have been approached with care and insight by those authorities that responded. The number who did not respond was a little disappointing. It may have been due to the short amount of time made available for responses, a particular concern since it coincided with the end of the holiday season. Opinions ranged from support to fairly strong opposition, with the key issues probably being the size of data zones, the frequent failure of data zones to represent local communities, and the failure of data zones to nest within other units used by local authorities, such as wards.

Regarding the first of these issues, the problem stems from the geography of Scotland and the very different settlement patterns in different parts of the country. It is hard not to sympathise with predominantly rural areas like Scottish Borders, where data zones are often too big to identify settlements with a strong local identity, or with large urban areas like Glasgow, where the size of the city leads to a large number of separate data zones which are much smaller than the communities local people recognise and identify with. The alternative is to have data zones of different sizes in different parts of the country, but this seems problematic in that it would lead to
certain kinds of sensitive data being available for large urban data zones but not for small rural ones.

It was decided earlier in the project that it was unrealistic for the research team to identify local communities across Scotland, and the concept of data zone was adopted in recognition of this. However, some authorities criticise the data zones produced on the grounds that they do not reflect local communities, either by grouping disparate areas together or by ignoring important local boundaries. Clearly data zones would be more useful if they came as close as possible to reflecting local communities of an appropriate size. However, the time scale and resources available to the project made this impracticable except insofar as 1991 census data and Ordnance Survey maps could be used to infer the existence of communities. Ideally, there should be a mechanism for including local knowledge in the zone design process, and some thought needs to be given to how this can best be done.

One can also have sympathy with those authorities that objected to data zones that cut across the zones they currently use for dividing up their areas for planning purposes. It is particularly negative for those authorities that have already devoted considerable time and effort to devising a system of neighbourhoods for their own purposes. There is also reason to sympathise with those who felt that data zones should fit in with existing political boundaries, such as wards. This was not attempted on the grounds that ward boundaries were themselves likely to change in the near future, but as it seems that data zone boundaries will themselves be subject to change every three years or so, this argument is less convincing.

There are some relatively minor technical points on which our zones were justly criticised. The most universal criticism was our use of 1991 Output Areas and census data, and many authorities were able to point to specific examples where new building and demolitions meant that our suggested data zones were clearly inadequate. A related criticism, our exclusion of Special Output Areas, also seems well founded. The occasional criticisms of data zone shape may also be answered if the 2001 Output Areas are more compact than their forerunners in 1991.

Overall we propose the following recommendations:

1. A system of data zones should be devised on the basis of 2001 Output Areas and census data for presentation of neighbourhood statistics.
2. Data zones should normally have populations of between 500 and 1000, assuming that this size band is appropriate given the confidentiality criteria to be imposed.
3. Data zones should be made up of complete 2001 Output Areas and should nest within local authority boundaries.
4. Consideration should be given to whether data zones should nest within ward boundaries.
5. Definition of data zones should be carried out centrally for the whole of Scotland using essentially the same methodology as used in Stage C of the current project, except perhaps for point 4 above.
6. Draft data zones should be circulated to local authorities, which would have a period of three months in which to suggest modifications, which would have to accord with points 2 and 3 above.
The data zone boundaries should be subject to revision approximately every three years, but in general they should remain the same unless population has grown well over 1,000 or declined to under 500, in which case they should be split or amalgamated.

A more thorough revision should take place when the 2011 census results are available.