



Draft GEOGRAPHY POLICY FOR NATIONAL STATISTICS

Issue

1. This paper provides a draft statement of the Geography Policy for National Statistics. The statement is intended to provide a simple summary of the approach to be taken for reference by those producing National Statistics. It is also hoped that the approaches and standards outlined will be used more widely within the Government Statistical Service (GSS) for statistics other than National Statistics.
2. Subject to comments by GSS Heads of Profession steps will be taken early in 2007 to formalise the policy as a National Statistics standard.

Action

3. Comments on, or suggested improvements to, the policy statement are welcomed before 31st December 2006 to **alistair.calder@ons.gov.uk**

Alistair Calder
ONS Geography
30 October 2006

GEOGRAPHY POLICY FOR NATIONAL STATISTICS

- 1.1 This document provides a summary statement of the Geography Policy for National Statistics.
- 1.2 Individual policies and standards are summarised in turn below.
- 1.3 In the interests of clarity and brevity this document focuses on the principles that lie behind each of the policies and contains only very limited background information. This, together with explanation of the methodology and details of each of the policies, (will be) provided in a separate document - '**National Statistics Geography Policy – Background and Guidance**'. Where appropriate, web links to further information are also provided below.
- 1.4 The policy is intended to be flexible enough to take account of the complexity of UK geography. As ever, the real world will bring cases in which it is not appropriate to follow the standard approaches outlined below but this statement provides 'best practice' that should be adopted where possible.
- 1.5 In particular the principles outlined in bold should followed wherever possible.

GEOGRAPHIC REFERENCING OF SOURCE DATA

- **All source data should be referenced at low levels.**

Low level referencing of source data improves statistical quality and allows real flexibility in outputs. Data relating to individuals, households or businesses should ideally be recorded at address level or at least at postcode level. Data located using grid references should be stored at an appropriate level of accuracy – but at least at 1 metre resolution.

- **Where possible source data should be referenced to identifiable objects.**

When the data being referenced relates directly to a real world object (a building, street or stretch of river for example) identifiers linking to these objects should be used to reference data (perhaps in addition to the grid references). Identifiers linking to real world objects provide a better solution than simple grid referencing since this eases and promotes data matching (see below). The use of object identifiers also allow any improvements to referencing datasets – where the location of buildings, fields or roads are re-surveyed, for example – to be reflected in subsequent statistics.

- **Locational identifiers (or at least grid references) should be stored with source data for future use.**

Adding identifiers or grid references to datasets for future use provides the potential for complete flexibility in the analysis and output of statistics (subject to disclosure control). Data referenced at low levels can be re-cut using any boundary set and statistics can produced for any area - essentially future-proofing the referencing of source data.

- **Use should be made wherever possible of standard references for objects to improve data matching.**

Where standard sets of unique codes already exist for objects (e.g. addresses, streets, ports and airports) their use should be encouraged. Where there is no standard the body responsible for the primary or base datasets should be encouraged to issue codes that can be used by the wider community. Such codes should be shared via the NS website (see also Standard Names and Codes section below).

USE OF STABLE SMALL AREA BUILDING BLOCKS

- **Wherever possible stable geographies should be used as the building blocks for the creation of National Statistics.**

In England, Wales and Northern Ireland the small area building blocks should be the Output Areas (Census 2001), and the Lower and Middle Layer Super Output Areas.

In Scotland these should be the Output Areas (Census 2001), data zones and the 'intermediate geography'.

(See 'NS Geography Policy – Background and Guidance' for details on these geographies and how they are to be used.)

- **'Exact' outputs for small areas should be produced and published first for these core geographies.**

Wherever possible the OAs, SOAs, Data zones, and/or higher areas - depending upon the nature of the data - should form the core geographies at which 'exact' NS outputs will be published. Statistics should be published for these areas before (or at the same time as) other small area outputs (e.g. ward level statistics).

There is no suggestion here that outputs will not also be produced for administrative geographies such as wards and parliamentary constituencies - these will continue to be released as long as they are required. However the core NS outputs will be for the building blocks described above. Outputs for other areas will be created either direct from the source data or via best fitting or estimation from the core units. Thereafter disclosure control will have to be applied to all outputs other than those for the core units to avoid disclosure through differencing.

The key benefits of this approach are:

- Use of a common, consistent geographic base for the collection and storage of data will allow easier integration and comparison of data across topics.
- Critically, because the units are of a known and consistent size, it should be possible to publish data at lower levels while remaining confident that there is no risk of disclosure. The new approach will enable publication of a wider range of data, for smaller areas than has previously been possible. Feedback from local users is already supporting this approach.
- The national consistency of the units will allow fair comparison between different areas and better targeting of policy.
- The stability of the base will make it possible to build time series and to monitor change and the impact of policies over time.

See '**National Statistics Geography Policy – Background and Guidance**' for further details on this policy and relevant web links.

STANDARD AREA MEASUREMENTS

- **All National Statistics which incorporate the area of land (i.e. area measurements or densities) should make use of the agreed standard methodology and published figures.**

Standard Area Measurements (SAM) are sets of area measurements, created using standard methodologies, for key geographies in the UK for which statistics are produced (administrative and electoral geographies, NUTS and health geographies).

Area measurements have been produced using two approaches:

- area measurements for electoral and administrative areas have been derived from boundaries maintained by the national mapping agencies; and
- area measurements for other geographies have been derived by aggregating Output Area (OA) based area measurements.

Published SAM figures currently include 4 separate measurements for each area (e.g. each SOA, ward or district):

- area to extent of the realm;
- area to mean high water;
- area of inland water; and
- land area (area to mean high water excluding inland water).

Unless there are policy reasons for using one of the other sets of measurements the 'land area' measurement should be used for calculating densities.

SAM figures are updated annually to take account of boundary changes and any other modifications.

The methodologies are described fully in the SAM guidance notes. The notes and the area measurements are available on the NS website:

<http://www.statistics.gov.uk/geography/sam.asp>

Note: The section below describes a variation on the existing policy previously agreed by Ministers. The variation has been discussed and agreed by NSGG – this statement as drafted assumes agreement to the proposed change by ONS's Statistical Policy Committee and HoPs. The proposed change is seen by all involved as a clear improvement on the existing policy so it is hoped that the variation will be accepted.

MINIMISING THE IMPACT OF BOUNDARY CHANGE ON STATISTICS AND 'STATISTICAL WARDS'

Users of statistics have experienced difficulties in the past caused by different organisations implementing the same boundary changes for statistical purposes on different dates. To overcome these problems it has been agreed that all departments should implement electoral ward changes on the same date each year.

- **For statistical purposes ward boundary changes should become effective on 1st April – and reflect any changes operative on or before 31st December of the previous calendar year.**

For example statistics published after April 2007 (and described as for 'Wards 2007') should be for those wards that had come into statutory existence on or before 31st December 2006.

Ward changes coming into effect at any point in a given year (Jan – Dec) will be reflected in statistical outputs from 1st April of the following year.

The policy is operational in England and Wales. It has also been accepted in principle by Scotland and Northern Ireland, though the policy is largely irrelevant here because their ward boundaries change simultaneously and very infrequently and so there is more scope to control the operative dates for statistics.

Published statistics should reflect the statistical wards existing at the time of publication rather than at the date of data collection, unless there are specific requirements that dictate otherwise.

This principle is intended to ensure that National Statistics reflect as far as possible the most current geography to maximise their relevance.

See '**National Statistics Geography Policy – Background and Guidance**' for further details on the application of this policy.

NAMING AND CODING FOR UK STATISTICAL GEOGRAPHIES

- **Wherever possible standard names and codes should be used for all geographic areas for which National Statistics are published.**

The implementation plan for standard naming and coding is currently being developed based on the recommendations of a review conducted in 2005 and endorsed by NSGG.

Once implemented, the policy will:

- contain rules for naming of policy and statistical areas so that they:
 - are unique across the UK, both within the area types and between related areas;
 - are consistently constructed, generally without area type;
 - use only standard characters, with minimal punctuation and abbreviation.
- contain rules for coding of policy and statistical areas so that:
 - new coding schemes are not invented by organisations for areas owned by others;
 - codes are unique within the UK for that area;
 - disused codes are not re-used for new occurrences of the area;
 - a standard format is used;
 - a standard style is used;
 - complex coding schemes with additional information are avoided;
 - there is no implied structure, for example arrangement in alphabetical order;
 - codes do not change solely because of a change to the name.
- contain rules for maintenance so that:
 - when changes occur to the boundary of an area that affect its content, a new instance of the area is created, with a new name and code;
 - when only the name changes, for example as can occur with local authorities, the code should not change;
 - redundant codes should not be re-used.

Other aspects of the proposed policy are:

- the creation of a central metadata service containing all area unit types, with a unique 3 digit code for each, and with a central authority to administer it and provide advice to data owners and administrators;
- the adoption of a new standard coding scheme for area unit instances, with a unique numeric code (up to six digits) for each instance;
- the establishment of definitive lists (gazetteers) for each area unit type, identifying each unit instance, cross-referencing to other codes used, and recording the boundary digitally against the largest scale of mapping available;
- the gradual implementation of new codes to fit in with other developments, and used in parallel with existing codes in the interim.

The standard for naming and coding will be published on the NS website when it is available and should, thereafter, be used wherever possible.

PRESENTATION ORDER FOR STATISTICAL OUTPUTS

- **All National Statistics outputs should conform to the agreed NS standard on order of presentation.**

The presentation standard is for use in publications (paper and electronic) and also statistics disseminated on the Web. It should be adopted unless there are valid reasons for not doing so, such as ranking areas according to a significant data item.

The general principle is that whenever entities are presented in groupings (e.g. districts within county, wards within district) the order of presentation at each level of the grouping hierarchy is alphabetic. This principle applies to all levels, and all geographical breakdowns of areas, whether using standard (e.g. administrative areas) or other geographies.

Exceptions to this are the order of countries and English GORs, where a standard geographically-based order is well established, and the geographical order of the unitary authorities in Wales.

Full guidance on the agreed standard is contained in ‘**National Statistics Geography Policy – Background and Guidance**’

Further details are also available at:

www.statistics.gov.uk/geography/presentation_order_guidance.asp which provides specific guidance for common administrative, health and electoral geographies.

URBAN AND RURAL CLASSIFICATION

There is currently no UK-wide standard classification of Urban and Rural land. There are, however, agreed standards for each of the separate administrations.

- **All National Statistics outputs which include reference to urban/rural areas should make use of the standard classification for the relevant country.**

See the following links for details on the classifications and their methodologies:

England and Wales:

www.statistics.gov.uk/geography/nrudp.asp

Scotland

www.scotland.gov.uk/library5/rural/seurc-00.asp

Northern Ireland

www.nisra.gov.uk/statistics/financeandpersonnel/DMB/urban_rural.html

(No single standard classification is recommended for NI but guidance is provided on the approach to be taken.)

See ‘**National Statistics Geography Policy – Background and Guidance**’ for more details on the use of Urban/Rural classifications.

MINIMISING NEW STATISTICAL GEOGRAPHIES

- **Departments should attempt to minimise the introduction of new geographies for statistical reporting.**

The UK Statistical Geographies Database (UKSGD) maintained by ONS provides a consistent knowledge base of the range of statistical geographies which are being used across National Statistics (NS). It is intended to help to control the process of creating new statistical geographies, improve the quality of geographic information data flows, and act as a useful reference tool.

The database is re-issued on the NS website on a quarterly basis.

Departments considering the introduction of a new geography should first check that another currently used geography is not suitable.

- **Departments should ensure that any geographies upon which they report or in which they have an interest are properly recorded in the UKSGD and that any changes are reported.**

For more details on the UK Statistical Geographies Database see www.statistics.gov.uk/geography/uk_statistical_geog_database.asp

Further background information and detailed guidance on use of the individual standards outlined above (will be) contained in '**National Statistics Geography Policy – Background and Guidance**'.

Draft

Alistair Calder

Alan Taylor

ONS Geography

30 October 2006