Land Use Change Statistics

Statement of quality assurance

This document/publication is also available on our website at www.gov.uk/dclg

If you have any enquiries regarding this document/publication, email contactus@communities.gov.uk or write to us at:

Department for Communities and Local Government
Fry Building
2 Marsham Street
London
SW1P 4DF
Telephone: 030 3444 0000

March 2016

For all our latest news and updates follow us on Twitter: https://twitter.com/CommunitiesUK

Contents

1. Purpose of document

This document provides information of the processes used to produce the land use change statistics and the supporting quality assurance procedures in place.

Specifically it will provide information on how the underlying data is collected, validated and analysed before becoming a finalised land use change data set. It is intended to help users understand the underlying products used to produce land use change statistics, the risks to data quality and the quality assurance procedures the department uses to mitigate these risks.

These statistics are provided to the Department for Communities and Local Government under agreement from Ordnance Survey Ltd.

This document has been produced by the department in collaboration with Ordnance Survey.

1.1 Land use change statistics

Land use change statistics are designed to show the estimated amount and location of land changing use in England.

They provide a way of monitoring development. Amongst other statistics, they can be used to derive information on:

- The percentage of new addresses created on previously-developed land and non-previously-developed land.
- The percentage new addresses created in flood risk areas, agricultural land and in the Green Belt.
- The density (addresses per hectare) of new housing.

Land Use Change Statistics have been published using a new methodology since August 2015:

https://www.gov.uk/government/collections/land-use-change-statistics

Details of the new methodology are available in the methodology changes guidance:

https://www.gov.uk/government/publications/land-use-change-statistics-methodology-changes-guidance

2. Overview of data sources

The underlying Ordnance Survey products used to produce the land use change statistics are listed below further details are available on the Ordnance Survey website. The products include:

MasterMap®

The topography Layer of Ordnance Survey MasterMap® provides a detailed, current and comprehensive map dataset of Great Britain.

http://www.ordnancesurvey.co.uk/business-and-government/products/topography-layer.html

AddressBase®

AddressBase® Premium is a current, comprehensive and accurate geographic dataset of addresses, properties and land areas where services are provided.

http://www.ordnancesurvey.co.uk/business-and-government/products/addressbase-premium.html

Points of Interest®

Points of Interest® is a location based directory of all public and privately owned businesses, transport, health, education and leisure services.
http://www.ordnancesurvey.co.uk/business-and-government/products/points-of-interest.html

Each of these products has its own extensive set of quality assurance procedures in place. Ordnance Survey has provided separate documentation detailing their quality assurance principles and standards which will apply to the production and update of these products. Full details of these principles were provided to the department and are available in Annex A.

Further details of the underlying products are provided in the following sections.

2.1 OS MasterMap®

A visual example of the OS MasterMap® Topography layer is provided in figure 1. All objects with the topography layer have a unique identifier for ease of referencing. There are over half a billion built and natural landscape features within the product.

Figure 1: OS MasterMap® Topography Layer

The topography layer of OS MasterMap® is a key flagship product for Ordnance Survey and it is updated and published every six weeks.

Alongside each publication documentation is provided detailing all of the areas where specific content has been improved as well as highlighting any current known issues with the topography data.

This includes instances where polygons are overlapping each other or if there are any gaps in the topography layer.

2.2 AddressBase®

AddressBase® Premium is a comprehensive and accurate geographic dataset of 40 million addresses, properties and land areas where services are provided. A visual example of AddressBase® data is shown in figure 2, which is address points overlaid onto the OS MasterMap® topography layer.

Figure 2: AddressBase® with underlying topography
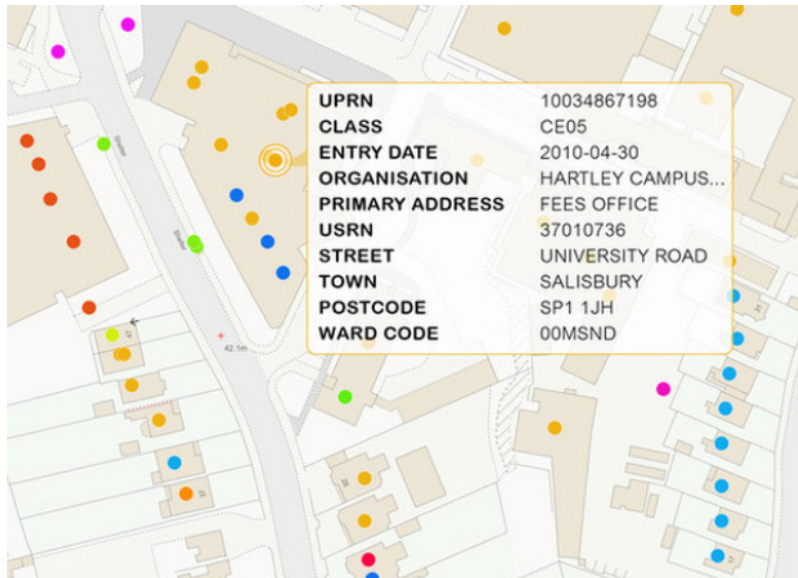
AddressBase® is the key product for identifying the residential address change data in the land use change statistics.

The information for AddressBase® comes predominantly from local authorities, who work with various sources in their organisation; council tax, electoral registration, planning and building control (amongst others) to identify and verify the existence and location of properties and their official address.

Once these checks have taken place, the property is added to their local land and property gazetteer. This gazetteer will have a number of validity and quality assurance checks to ensure the information input meets the British Standard for addressing (BS7666).

This information is then supplied on a daily basis to GeoPlace (a Limited Liability Partnership jointly owned by the Local Government Association and Ordnance Survey) who run additional quality checks against the data submitted.

Anything that fails this validation is passed back to the local authority custodian to fix before re-submitting. The address information that passes this second validation is then added to the AddressBase products and matched to other address sources, primarily the Postal Address File from Royal Mail.

Once the product is built, the information from GeoPlace is taken by Ordnance Survey and run through quality assurance tests within Ordnance Survey to ensure the product is correct to specification. Only when it has met this is the product then updated and published. The methodology then uses the closest published releases to the start and end dates of the period (1st April and 31st March).

2.3 Points of Interest®
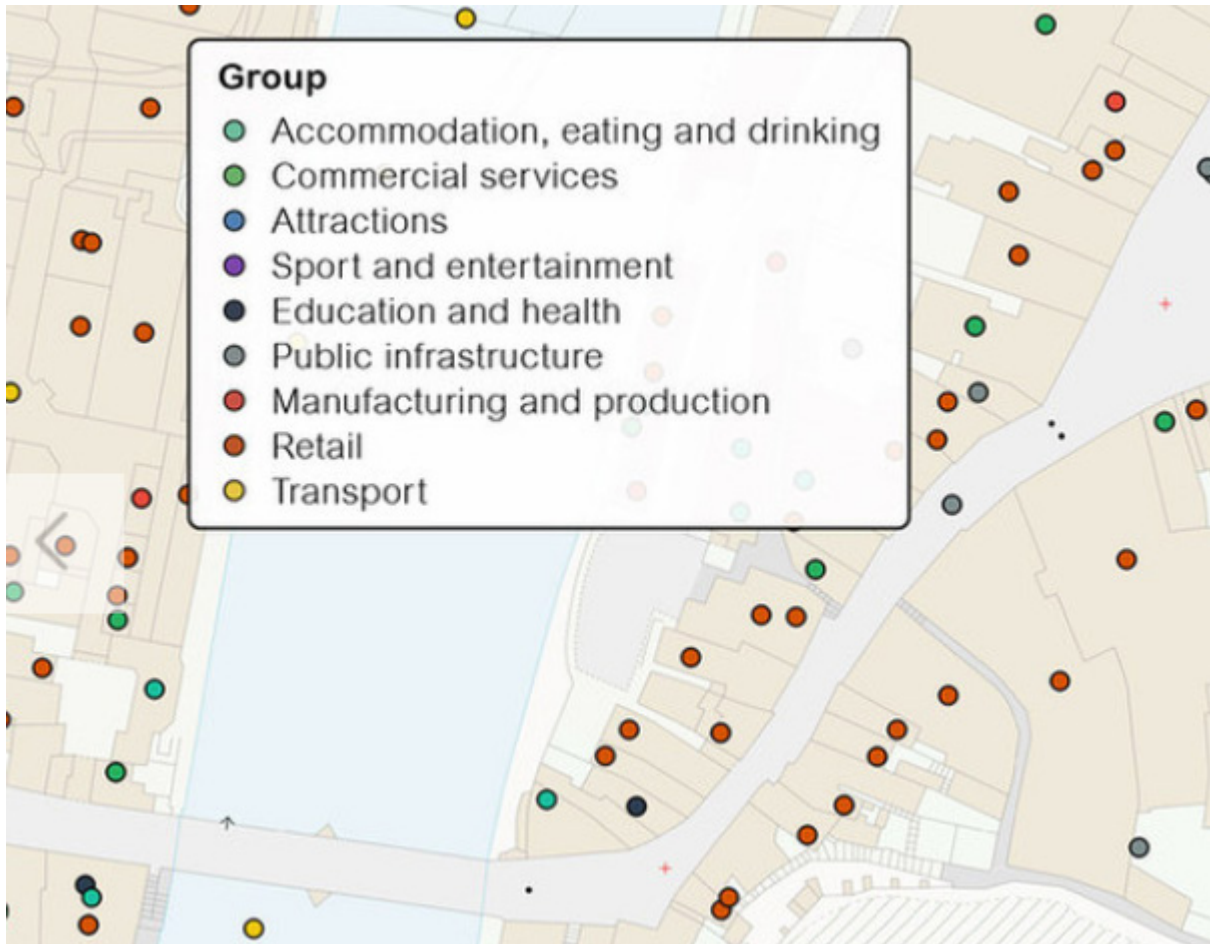
Points of Interest is a location based directory of all public and privately owned businesses, transport, health, education and leisure services in Britain. It drives and informs customer-facing decisions, plans and strategies in business and the public sector.

It contains information on over 4 million points in Britain and is updated quarterly. Records are sourced and quality-checked from over 100 leading listing suppliers.

Points of Interest are produced by PointX and distributed by the Ordnance Survey Where possible PointX match the address provided from its suppliers against the structured Ordnance Survey address data. This means that the address reported conforms to the ordering, spelling and structuring used in the address data.

3. Production of Land Use Change statistics

Once data is up to date in the underlying products, Ordnance Survey can then use them in the land use change statistics methodology. In combining these data sources they are able to infer land use change. Further details of this process are provided in the Land use change statistics methodology changes guidance

3.1 Quality Assurance checks performed by Ordnance Survey

Prior to running the LUCS data creation process, Ordnance runs several checks on the raw input datasets to ensure that they are correct. These checks include:

- The date metadata, at individual record level, within each of the input datasets is checked to ensure that all dates fall within the expected ranges (ie. each input dataset does indeed correspond to the advertised dates).
- All unique identifier fields are tested for uniqueness within each dataset.
- Geometries are checked for duplicates and overlaps.
- Attribute lookup tables are joined to each of the input tables to ensure that every attribute combination has a corresponding entry in the relevant lookup table.
- A copy of each input dataset is retained in its raw form so that it can be referenced during the intermediate checks in the process and in the final quality assurance.
- Irrelevant records (e.g. landform records within OS MasterMap Topography Layer) are removed before entering the process.

Ordnance Survey's land use change statistics team conducts additional checks at each stage of the data creation process to ensure that all records flow through the process as expected, and that no duplicate records are created at any stage.

- Record counts are checked against those of equivalent input datasets after each stage of the process to ensure that the numbers match exactly.
- Every intermediate table created is checked for duplicate geometries and/or unique identifiers to ensure that no duplication occurs during table joins.

Once the data products have been run through each step of the methodology an output data set is then generated. This output data set is itself then subject to quality assurance checks before an initial version is provided to the department. These checks are used to validate the records produced as output from the methodology and to filter out any incorrectly identified change.

These checks include
- Examining the data is at the macro scale (whole of England) in a Geographical Information Software package to identify observable geographical or class anomalies.
- The land use change data is then analysed by the number of records and total area of change by both the TO (new land usage) and FROM (previous land

usage) classes across several of the other data components to sense check at the macro level.

- All of the TO and FROM classes by local authority in stacked bar graphs. This has helped visualise a local authority's change profile and by comparison draw attention to unusual or unexpected amounts of change at the more local level. Any unusual data have then been micro inspected at the individual site level in a GIS to understand the process generating that specific polygonal change and where appropriate tested for validity against other sources such as imagery or news stories that could confirm the existence of the change in the real world.
- Ordnance survey factor in information from their teams on any specific changes to the underlying products which may have occurred during the year. This allows them to filter out change which is captured as part of data improvement programs as opposed to real world change.

In addition to this checking process, any unusual spatial patterns or the unusual classification of specific sites is recorded and can be used to further refine the methodology's algorithms to improve both the identification of classes and the overall quality assurance for subsequent iterations of the data.

Once the data has passed through his series of data quality checks an initial version of the data is provided onto the department.

3.2 Quality Assurance checks performed by the department

Following the checks on the initial data run by Ordnance Survey comprehensive further checks are run on the data when it is received in the department. This is to provide assurance on the quality of the data before proceeding to run the analysis required to produce the statistical release and tables.

The checks run by the department prior to analysis include:
- Verifying all points provided lie within the boundary of England.
- No duplicate records or overlapping areas.
- Investigation of high density figures in Local Authorities. Specifically looking at areas with unusually high standard deviations and analysing high density clusters of points.
- Identified high density points spot checked against real world changes. Checking is focused on the points with the biggest potential to influence the national and local totals.
- Checking local authority totals sum to national figures.
- Checking for unusual year on year change at both local and national level.
- Data is validated at national level as of a similar magnitude to other departmental data on housing statistics and external data sources on the

Green Belt, planning permissions and floor space of new residential developments.

Once the department is satisfied that the data provided can pass these checks and therefore is of sufficient quality then it is signed of as being finalised. It will then be the finalised data set used to produce the statistical publication.

3.3 Quality Assurance checks performed prior to publication

Once the data is finalised the analysis and production of the statistical publication and associated tables is begun. The production of the release is typically carried out by the lead statistician. They are responsible for producing and checking the tables, charts and maps and then analysing the results to produce the commentary for the statistical release.

When the statistical release is ready then the lead statistician will quality check it by running through the quality assurance checklist. This is a list of checks to ensure that the data is correct and consistent throughout the publication. Each check is signed off to say it has been completed and the finalised checklist is saved for internal assurance purposes. A second independent statistician then reviews the statistical release and using the same checklist will quality assure the final products. Once this process has been completed then the statistics are ready to publish.

4. Relationship with suppliers and users

4.1 Relationship with suppliers

There are well maintained collaborative working relationships between the department and Ordnance Survey to produce this data. During the production lifecycle regular discussions on progress are on-going with the department's lead statistician. Throughout the project there is an overarching project plan and timetable for delivery.

Meetings between the department and Ordnance Survey were regularly called when finalising the methodology. During these meeting Ordnance Survey were able to give full briefings on any technical challenges and detailed explanations of any quality issues. This has included deciding on how best to produce the new measure of density and the most accurate way to identify previously developed land.

Any issues arising are dealt with by Ordnance Survey with the department providing oversight on what the most agreeable methodological solution would be to provide data which met requirements.

There is a full contract between the department and Ordnance Survey for the provision of data. It covers the timetable of the expected data supply. Following the initial provision of data under the contract the responsible statistician is able to review the product and take final for acceptance of the data.

Details of checks run before acceptance covered under the previous section of this document. Should any issues be encountered with the data provided the department are not required to make payment until Ordnance Survey are able to satisfactorily resolve them.

This contract is currently scheduled to run until November 2015 and will deliver data for the years 2013/14 and 2014/15 with the option to extend for an additional year to cover 2015/16.

4.2 Relationship with users

Statistical data users are part of the Central Local Information Partnership Planning (CLIP) sub group meetings and are regularly briefed on the progress of the land use change statistics. A planning alert mailing list is used to flag key information to users. Any interested party can be registered on the mailing list and this is advertised in the statistical publication. All users are invited to feedback on the data and identify any issues.

Feedback on the statistics is welcomed at any time and is used to improve the statistics going forward. Anybody wishing to be provided with regular updates on the statistics can be added to the mailing list.

Please email planning.statistics@communities.gsi.gov.uk with any feedback or requests to be added to the mailing list.

5. External input on the quality assurance process

UK Statistics Authority

The UK Statistics Authority has been working closely with the department on the transition for the old to the new methodology. The UK Statistics Authority provided advice on how the changes to the methodology could be best publicised to users and helped in reviewing some of the documentation explaining these changes.

Further to this they were able to provide further advice on documenting the quality assurance processes within the new methodology. The department is grateful for their input which has allowed it to improve the quality and clarity of documentation provided to users.

Upon the initial publication of the data the department was able to produce an assessment of the new methodology against the [Quality Assurance of Administrative Data framework.](#)

[This assessment](#) was designed to provide users with initial assurance that the quality assurance in place was of sufficient standard for them to have confidence in the new methodology. This document aims to build on that assessment and provide a more narrative view of the quality assurance processes in place.

6. Overall summary of the quality assurance process

The underlying data products used to produce land use change data are well established, frequently updated and have comprehensive validation checks built into the Ordnance Surveys collection processes. After applying the land use change methodology to this underlying data significant further checks are performed by both the Ordnance Survey and the department.

There is a moderate level of interest by local authorities, planning and development organisations, government bodies and media organisations in the land use change data. The information is also used to monitor and develop the impacts of government policy. It is therefore important, for the full range of users, that the published data is accurate and reliable.

As demonstrated throughout this report, the department has extensive quality assurance in place throughout the production and analysis of land use change data. Quality assurance is applied from the point at which data is collected, through to the production of the final statistical release.

The department is responsible for regularly reviewing any risks to the data quality and improving the quality assurance processes where necessary. The department can be confident that the main risks to data quality are sufficiently reduced and managed, and that the data published is high quality and fit for purpose.

7. ANNEX A- Ordnance Survey's approach to quality

Quality for Ordnance Survey refers to (but is not limited to): data quality, volume of delivery, schedule adherence and cost. Whilst these elements may be managed separately for operational purposes, to truly meet customers' needs the quality of Ordnance Survey data must best be considered in its entirety. Specifications and requirements must be stated, agreed and effectively communicated to allow management of data supply across the organisation and to serve a complex supply chain that ultimately populates our data stores and products.

We apply relevant Standards to our processes, products and services which help us better meet the customers' requirements. For example ISO19157: *Geographic Information – Data quality* (2013), ISO19158: *Geographic information - Quality assurance of data supply* (2012) and ISO2859: *Sampling procedures for inspection by attributes* (series). In addition we both recognise that customer expectations may sometimes differ from those stated and act appropriately/accordingly and also acknowledge that the value of any quality measurement comes not from the data quality results but from the resultant identification and implementation of improvements in our capability (people, processes and tools). These principles lie at the core of our quality management approach.

Quality plans

Quality plans identify how the people, processes and tools will be managed and combined to provide the capability to meet a given customer requirement. Quality plans are the responsibility of production area managers.

Our policy is to ensure all quality plans include the following:

- Reference to a clear and controlled specification
- Reference to a training plan that all operators will need to complete before they start working on the process.
- Reference acceptable quality levels for all relevant elements of delivery. For example data quality (how complete the data is) and what rate of output is expected.
- how overall quality will be monitored and reported (Metrics)
- how the quality of individuals will be monitored and reported (individual accreditation)
- how exceptions/quality issues will be handled

- how improvements to the process will be made and targets for those improvements.

Accreditation

Through Accreditation Ordnance Survey satisfies itself that its suppliers, both external and internal, are capable of delivering data to the required quality. Quality in this instance refers specifically to data quality, schedule of delivery, volume of delivery and cost. Accreditation achieves this through an appropriate mix of timely data testing, and process review. The Accreditation process is based on our experiences of quality management systems, complex requirements and different production environments. In 2012 this internal Ordnance Survey process was published as an International Standard: ISO19158: *Geographic information – Quality assurance of data supply* (2012).

Accreditation is used both to identify the level of risk in meeting the customers' requirement and to assist in identifying appropriate controls to better meet the customers' requirements. Achieving accreditation, in itself, is not the overall aim.

Our policy on Accreditation includes the following:

- Accreditation will apply to all production processes.
- All staff working on any process will have individual accreditation (or be actively seeking to achieve it to an agreed schedule).
- Accreditation status will be used as one of the inputs used to consider approval to advance into production from pilot or trial activity.
- Full Accreditation is awarded in relation to the entire end to end supply chain. To achieve full Accreditation the supply chain must be capable of delivering to the customer requirement. A key input into an end to end assurance review is customer feedback (internal and external) on the product or service being supplied.
- All processes will use the Accreditation guidance document to implement Accreditation which is based on ISO19158: *Quality assurance of data supply* (2012).

Quality Control (QC)

The purpose of QC is to ensure that each individual and process is delivering to the required quality within a production process. QC is part of the production activity and is the responsibility of Ordnance Survey's internal and external production teams. QC directly supports the maintenance of individual accreditation. The aim of QC is to identify the level of quality achieved by an individual and to assist in identifying

opportunities for improvement to better meet the customers' requirements. Passing QC, in itself, is not the overall aim.

Our policy on QC includes the following:

- All production processes will have QC procedures.
- The level of QC (sample size) will be identified in the process Quality Plan.
- If appropriate the level of QC required will be determined using ISO 2859 – Sampling Procedures for Inspection by Attributes.
- Results of internal QC will be recorded in an appropriate store (eg Accreditation Database).
- Efficiency and effectiveness of QC will be enhanced through the use of automatic validation software as it becomes available.
- Data will not be committed to the data store until appropriate Quality Control has taken place.
- The QC results will be used to identify appropriate improvement strategies (for example learning and development of staff)

Data assurance

Data assurance testing is a check on data that has already been committed to the data store. The purpose of this is to ensure that the quality control within any production process is working effectively and that the process is delivering the required quality. The aim of data assurance testing is to ensure consistency and appropriate quality throughout our supply chains and to assist in identifying opportunities for improvement to better meet the customers' requirements. Testing, in itself, is not the overall aim.

Our data assurance policy includes the following:

- Assurance will always be performed independently of the original production process.
- Where appropriate to do so, the level of testing required will be determined using ISO 2859 – *Sampling Procedures for Inspection by Attributes*. Consideration will also be given to the cost and revenue and risk posed.
- Results of any testing will be recorded in an appropriate store (e.g. Accreditation database).
- Results will be used to identify appropriate improvement strategies (e.g. system improvements including enhancement of validation rules and the learning and development of staff).

Automated data quality validation

Validation through software provides one of the best opportunities for quality assurance. In many cases validation can be applied to entire datasets with minimal manual input, providing ourselves and the customer with prompt and substantial confidence in the quality of that dataset. Automated validation is however limited by the rules that can be applied logically and by the complexity of those logical rules. Development time, run time and false positive results are all factors taken into account when we consider the role of automated validation.

Our policy on automated data quality validation includes the following:

- Opportunities for automated data quality validation must be considered against the benefits, the risks posed by the dataset as well as the cost of any manual intervention.
- Data validation planning will be an integral part of any changed or new system.
- Data validation will be undertaken on import and export prior to and following data capture and maintenance activity.
- The type of validation to be run will vary depending on what process has been carried out on the data.
- Existing validation will be regularly reviewed for effective improvement opportunities for overall quality assurance.

Product assurance

Product assurance is a final opportunity for assurance before any product is released to customer. The aim of this activity is to identify issues and risks in customer expectations of the product in terms of accessibility and usability within the context of the product specification

Our policy on product assurance includes the following:

- Any assessment will be made against the defined product specification (and defined user requirements and user guide documentation).
- Product assurance data quality results will be used as a key input in any decision to accept new or changed product data.
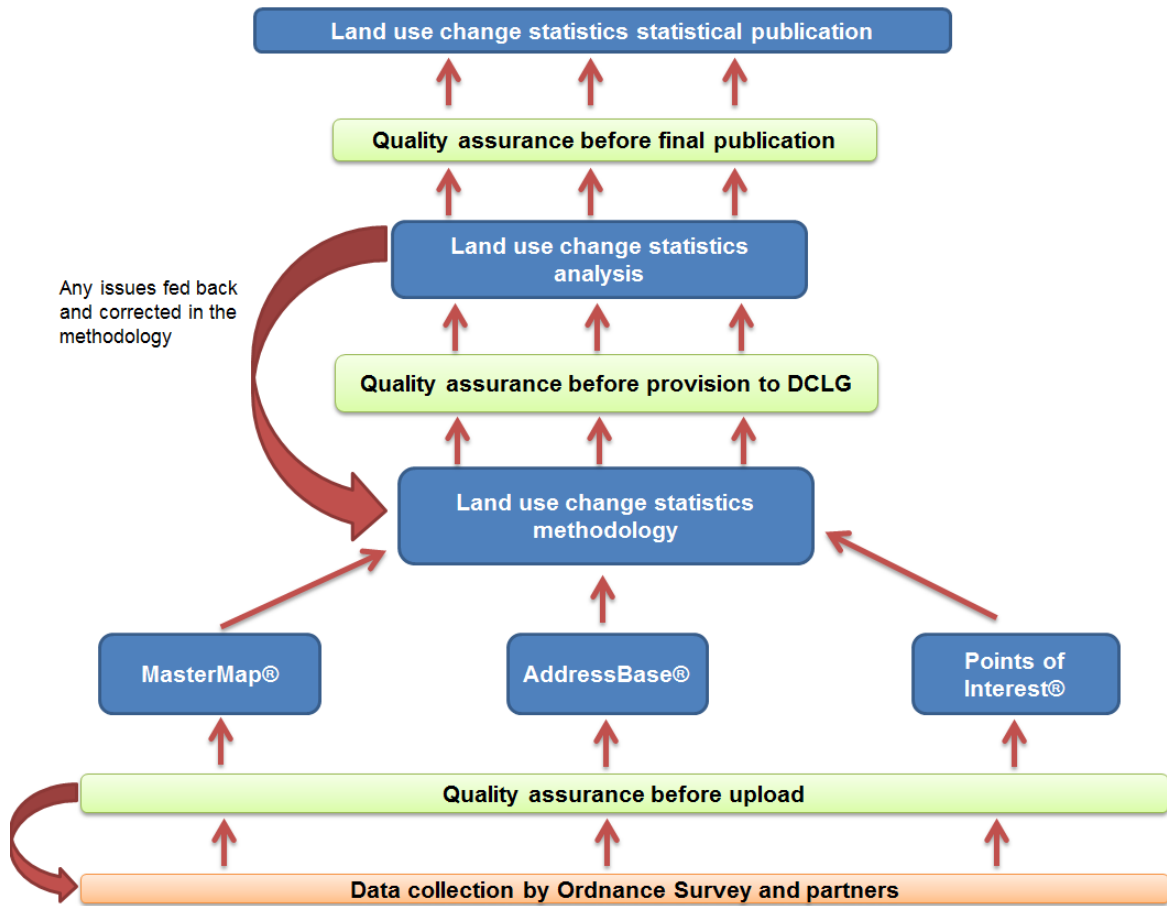
Quality measures and levels

Quality measures provide the metrics for the measurement of quality that is important to the customer to meet their requirement. Once appropriate measures are identified acceptable quality levels (AQLs) can be agreed. Quality Levels or Acceptable Quality Levels are a means of measuring the relative quality

conformance delivered against the requirement. With any manual processes we recognise that variation is likely to occur, however AQLs provide the minimal level of conformance acceptable to the customer. Increases in AQL values may have substantial impact on cost and time. These factors are always considered as flexible variables that can be adjusted to better meet the customer requirement. Allowing for some error will form part of our managed risk and is further controlled through our policy on data quality issues and errors.

Our policy on quality measures and levels includes the following:

8. ANNEX B – Overview of Quality Assurance process flow

Land use change statistics statistical publication

Quality assurance before final publication

Land use change statistics analysis

Any issues fed back and corrected in the methodology

Quality assurance before provision to DCLG

Land use change statistics methodology

MasterMap®

AddressBase®

Points of Interest®

Quality assurance before upload

Data collection by Ordnance Survey and partners

OGL

This document/publication is also available on our website at www.gov.uk/dclg

If you have any enquiries regarding this document/publication, complete the form at http://forms.communities.gov.uk/ or write to us at:

Department for Communities and Local Government

Fry Building

2 Marsham Street

London

SW1P 4DF

Telephone: 030 3444 0000

For all our latest news and updates follow us on Twitter: https://twitter.com/CommunitiesUK

March 2016