# Scottish Address Directory (SAD)

**Quality Assurance**

**November 2019**

**Table of Contents**

# 1. Introduction

This document outlines the quality assurance and quality control measures in place for the Scottish Address Directory (SAD), NRS Geography's main addressing product.

Quality assurance measures are in place to ensure that we are using appropriate methods to create products that meet the needs of Geography customers. Quality control measures are in place to verify that the results of these methods are correct and working as expected.


# 2. Background

Since 2011 NRS Geography have been working on creating a comprehensive all Scotland address list, the SAD, which could be used by a number of addressing customers both internally and externally, for example Scotland's Census 2021, Data Linkage, Scottish Government Social Surveys.

The SAD contains address information, derived from Ordnance Survey's AddressBase (AB) Premium product, classification of type and three unique identifiers; Unique Delivery Point Reference Number (UDPRN) taken from the Royal Mail Postcode Address File (PAF), Unique Property Reference Number (UPRN) and Unique Street Reference Number (USRN), both taken from the One Scotland Gazetteer (OSG).


# 3. Creation of the Scottish Address Directory (SAD)

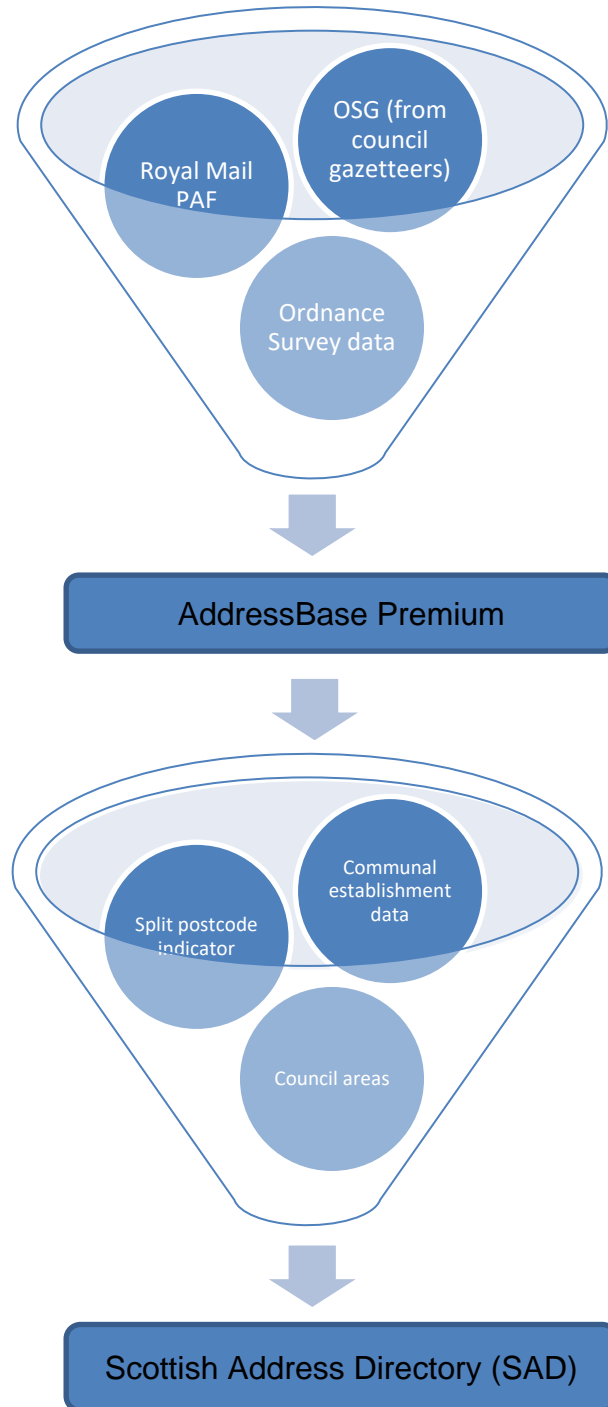SAD is an SQL database with a Microsoft Access front end.

On a quarterly basis the AB Premium product is uploaded by NRS Geography, formatted using an FME workbench, then the output is cleaned and run through an automated system to produce the SAD.  Supplementary information from Royal Mail and Geography's communal establishment register (CER) is added using the unique reference numbers as a link.

Cleaning, formatting and the addition of supplementary information is done via a .NET application and SQL queries.

By comparing the contents of the previous version of the SAD against the latest version, an updated list of historical address records and address changes is also created.

The frequency of updates is currently quarterly but will increase to 6-weekly in 2020. SAD itself is not a Census product, however it is the basis for the Census Address Register (CAR).

*Fig.1 Creation of the SAD*



OSG (from council gazetteers)

Royal Mail PAF

Ordnance Survey data

↓

**AddressBase Premium**

↓

Communal establishment data

Split postcode indicator

Council areas

↓

**Scottish Address Directory (SAD)**

# 4. Quality Assurance and Control

4.1    Quality Assurance

- Internal cleaning of the source data to remove irrelevant information such as non-relevant addresses (e.g. fields, bus shelters) via a series of SQL queries.
- Comparison of address matches contained in AddressBase Premium and those from an independent address matching of PAF and OSG done by NRS Geography.
- Cross referencing address data between a variety of sources looking for inconsistencies.
- Comparing address classifications between different source datasets and investigating discrepancies.
- Regular discussions with Office of National Statistics (ONS) and Northern Ireland Statistics and Research Agency (NISRA) to share best practice and identify common challenges in production of Census address registers.
- Regular discussions with data suppliers including Ordnance Survey (OS) and the Improvement Service (IS) to highlight issues and find solutions that can be implemented at source.
- Regular discussions with internal users of Geography address data.
- Visits to local authority data custodians to build up a picture of issues with source data and develop a timeline of change.

4.2    Quality Control

- Regular production of management information statistics, which are monitored and analysed – this includes numbers of new, changed and deleted addresses, numbers of addresses meeting certain classifications and numbers of new, changed and deleted postcodes.
- Producer/review check i.e. work produced by one member of NRS Geography shall be checked by another member of NRS Geography before being released to customers.
- Targeted mail-out exercises to identify addresses where post is undeliverable
- Field address check in areas of most concern e.g. areas where non-field address checks have a lower success rate.
- Peer review of address data by colleagues within NRS and Scottish Government.

## 5. Quality Criteria

| Quality Measure | Method | Why |
|---|---|---|
| Internal cleaning of the source data to remove irrelevant information such as non-relevant addresses (e.g. fields, bus shelters). | Identify classifications in source data as non-relevant. Remove these during SAD creation. | Ensure only building addresses are used. |
| Comparison of address matches contained in AddressBase Premium and those from an independent address matching exercise of PAF to OSG done by NRS Geography. | Clean datasets (e.g. remove leading & trailing spaces, double spaces, special characters etc.) and prepare address strings. Run a series of queries, changing linked data each time, to compare the strings in one dataset against the strings in the other dataset. | Verify accuracy of the match made in AB. |
| Cross referencing address data between a variety of sources looking for inconsistencies. | Link UPRN between datasets and pull out relevant information for investigation. | Identify issues where address information does not match (e.g. postcodes). |
| Comparing address classifications between different source datasets and investigating discrepancies. | Link UPRN between datasets and pull out relevant information for investigation. | Identify discrepancies where sources classify a record differently (e.g. school and shop). |
| Regular discussions with ONS and NISRA to share best practice and identify common challenges in production of address registers. | Attend meetings, provide information. | To promote harmonisation across the UK census offices on all address related matters. |
| Regular discussions with data suppliers including Ordnance Survey and the Improvement Service to highlight issues and find solutions that can be implemented at source. | Attend meetings, provide information, feedback issues with source data. | Resolve issues with source data. |

| Quality Measure | Method | Why |
|---|---|---|
| Regular discussions with internal users of Geography address data. | Attend meetings, provide information. | Keep colleagues informed of progress and issues which may impact use of the address data. |
| Regular production of management information statistics, which are monitored and analysed. | Compare previous versions of data and provide counts of current data. | Monitor rates of change; identify anomalies. |
| Producer/review check i.e. work produced by one member of NRS Geography shall be checked by another member of NRS Geography before being released to customers. | Run validation queries or manual checking. | Catch and correct any errors. |
| Targeted mail-out exercises to identify addresses where post is undeliverable. | Select sample addresses and mail out to occupier seeking clarification of the address details. | To quality assure the address data and help inform changes to source data. |
| Field address check in areas of most concern. | Select sample addresses and undertake a manual check to verify the address details. | To quality assure the address data and help inform changes to source data. |
| Peer review of address data by colleagues within NRS and Scottish Government. | Seek feedback from users of data, in particular where a comparison with other datasets has been made. | To quality assure the address data and help inform changes to source data. |