Small Area Population Estimates in the transformed population estimation system: methods development

A working paper that explores different methodological options for Small Area Population Estimates in the transformed population estimation system

Contact:
Valentina Gribanova
Small.Area.Estimation.
Modelling@ons.gov.uk

Release date:
7 December 2023

Next release:
To be announced

Table of contents

1 . Why transform the population estimation system?

Current population estimation system

In the current population estimation system for England and Wales, the most accurate estimates of the population are produced every 10 years from the census. Population estimates from a census are updated each year using the cohort component method, described in our Methods guide, to produce mid-year population estimates (MYE). The cohort component method adds on births, removes deaths, and accounts for internal, cross-border and international migration to roll the census estimates forward over time. Certain special populations (armed forces and prisoners) are also adjusted for. Whenever a new census is conducted, MYE for the previous decade are revised (rebased) to ensure consistency. For example, in November 2023, the rebased MYE for 2012 to 2020 were released in our Rebasing of mid-year population estimates following Census 2021, England and Wales bulletin to ensure that they are consistent with Census 2021. The cohort component method is used to produce MYE for geographical breakdowns down to local authority (LA) by age and sex.

For geographical breakdowns below LA level, commonly referred to as Small Area Population Estimates (SAPE), different methodologies are used (see our Methodology note on production of population estimates). Middle layer Super Output Areas (MSOAs) and Lower layer Super Output Areas (LSOAs) both use a ratio change methodology (see our Census 2021 geographies methodology). This method takes the census as a base and rolls the estimates forward each year using the change in the population recorded in administrative sources for consecutive years as an indicator of change in the true population. The method is used to produce these estimates in intercensal periods. Prior to 2013, both the NHS Patient Register (PR) and Department for Work and Pensions (DWP) Child Benefit data were used as the administrative data sources for the ratio change. From 2013, only the PR data was used though the PR was discontinued in 2020. For the 2020 to 2021 period, the NHS Personal Demographic Service (PDS) was used for the ratio change and estimates at LSOA level were published in November 2023 for mid 2021 (not accounting for Census 2021). Special populations (prisoners and armed forces) are also adjusted for in the ratio change method. For consistency, LSOA mid-year population estimates are constrained to MSOA estimates, which in turn are constrained to the LA MYE. Unlike the cohort component method used for LA estimates, no flows data (births, deaths, and migration) are used to inform the ratio change method.

Estimates at Output Area (OA) level use an apportionment method instead of a ratio change method. We acknowledge the importance of OA-level estimates, since they form the building blocks for producing other geographical breakdowns, such as National Parks. This paper focuses on MSOA and LSOA-level estimates; details on current methods for OA estimates are in our Methodology note on production of population estimates. Note that the MSOA and LSOA population total estimates were previously granted National Statistics status, with OA estimates provided as supporting information only, as explained in our SAPE: Summary of methodology review and research update.

Transformed population estimation system

We are transforming the population and migration statistics system for England and Wales, making use of the best available data sources with a focus on the use of administrative data (see our Overview of the transformation for more information). At the heart of this is the our Dynamic Population Model (DPM). The DPM uses statistical modelling techniques to bring together a range of data and demographic insights to estimate the population. Similar to the method previously described for the LA MYE, the DPM also uses a (model-based) cohort component method to update the census population with birth, death and migration flows. However, the DPM also uses admin-based measures of population stocks between censuses, primarily our Statistical Population Dataset (Statistical Population Dataset (SPD). The SPD is a linked administrative dataset, which applies a set of inclusion rules to approximate the usually resident population. The DPM incorporates information on data quality and uncertainty and includes system models to incorporate expert knowledge on demographic behaviours.

The DPM currently uses census data to adjust the administrative data stocks for coverage issues, but has a future vision to replace this with sustainable methods that make use of other sources, which could include the use of a regular coverage survey or audit surveys to validate SPD inclusion rules, or both. The DPM currently produces estimates for LA by age and sex, but is not designed to produce estimates for geographical breakdowns below LA level. This is partly because of the challenge of combining the administrative data sources, some with the required geographic granularity, others without, to get to the required level of detail and precision. In addition, there would be a high percentage of zero or small counts in the inputs to the DPM at MSOA and LSOA level split by age and sex that would prevent the model running for some small areas.

In Section 7 of our recent DPM, improvements to data sources and methodology for LAs methodology, methodological options for estimates below LA level in the transformed system were discussed. One option is to consider the same ratio change method for MSOA and LSOA estimates, as described in the previous subsection, but to use a DPM LA benchmark instead of an MYE LA benchmark. Another option would be to consider benchmarking SPD counts at MSOA or LSOA level to DPM LA benchmarks.

Here we will build on these different options and compare their performance for estimates at MSOA and LSOA level in terms of bias. The methods considered and how they will be compared are discussed in Section 2: Baseline methods for SAPE estimates in transformed population system. The emphasis of Section 2 is on existing and simpler methods (ratio change and benchmarking) that provide a baseline in terms of quality to compare other methods with. Section 3: Results provides results of the comparisons between the methods outlined in Section 2. Section 4: Geospatial methods discusses a method that is currently under investigation which makes use of geospatial data at low levels of geography. Section 5: Conclusion provides conclusions.

2 . Baseline methods for Small Area Population Estimates in transformed population system

Several different methods have been considered for producing Small Area Population Estimates (SAPE) in a transformed (admin-first) system. Here, we focus on Middle layer Super Output Area (MSOA) and Lower layer Super Output Area (LSOA) totals only using ratio change and benchmarking methods. We refer to these as baseline methods, since they have been used previously at the Office for National Statistics (ONS) and, in the case of the ratio change method, are used for official estimates. Therefore, they serve as a comparative baseline for more innovative methods. Future work could consider estimates by age and sex as well as Output Area (OA)-level estimates.

Ratio change method

The ratio change method has been used for MSOA and LSOA estimates in the current population system for many years. The method is conceptually simple and has the advantage of being familiar to users. However, there are disadvantages. The method does not have a natural mechanism for producing measures of uncertainty to accompany the estimates. However, note that we have never published uncertainty (variance) measures for population estimates below local authority (LA) level. Both our Census 2021 confidence intervals (XLSX, 2.71MB) and Mid-year estimates (MYE) confidence intervals dataset are only available at LA level (or higher). The ratio change method is also reliant on having a consistent annual supply of administrative data either from a single administrative data source or a combination of sources that are linked in an appropriate way (such as the Statistical Population Dataset (SPD)). The administrative data used for the ratio change method must also be deemed suitable for reflecting population change. Finally, the ratio change method is currently reliant on a census starting point and it uses the census data directly at MSOA or LSOA level.

In this report, we vary both the LA benchmarks and the administrative datasets used for the ratio change from those that are considered in the current population system. The idea is to perform a sensitivity analysis that makes better use of experimental data, such as LA estimates from the Dynamic Population Model (DPM), to assess the impact this has on the quality of the estimates for lower levels of geography. To measure quality, we focus on bias by comparing the ratio change MSOA and LSOA estimates for MYE 2021 (rolled forward from census MYE 2011) to the Census 2021 estimates (source of truth). At the time of writing this report, MSOA and LSOA estimates for census-based MYE 2021 were not available, so we use Census 2021 instead (which have a March rather than June reference date) and acknowledge that we are not accounting for all the bias between Census 2021 and MYE 2021.

The different variations in data inputs of the ratio change method are as follows:

- method 1a replicates the method used in the current population system with the rolled forward MYE being used as the LA benchmark (for 2012 to 2021) and the PR being used as the administrative dataset for ratio change up until 2020 at the time of writing this report, the official rolled forward MYE for 2021 at MSOA and LSOA level have not been released); in addition, the Personal Demographics Service (PDS) data for 2020 to 2021 could not be sourced for this work, so the SPD data is used to reflect the population change between 2020 and 2021

- method 1b is the same as method 1a, except a DPM LA benchmark is used and, to ensure a fair comparison, the DPM without Census 2021 as an input is used; this model is referred to as “Admin Based Population Estimates (ABPE) basic” in our DPM, improvements to data sources and methodology for LAs methodology

- method 1c uses the rolled forward MYE as the LA benchmark, but changes the administrative datasets used for the ratio change; the PR is used for 2011 to 2016, with SPDs used for all the years they are available (2016 to 2021)

- method 1d is the same as method 1c, except a DPM (without Census 2021) LA benchmark is used

Note that all PR and SPD data used has an MYE reference date (30 June) for each year.

Benchmarking method

The second baseline method discussed in this section is a benchmarking approach. In this method, 2021 SPD counts at MSOA and LSOA level are benchmarked (or calibrated) to appropriate 2021 LA benchmarks. By incorporating benchmarks of sufficient quality at LA level the estimates at lower levels should be improved, versus not including any benchmarks at LA level and using the SPD MSOA and LSOA counts directly. The benchmarking method has a similar advantage to the ratio change method in the sense that it is conceptually simple. Unlike the ratio change method, the benchmarking method is not dependent on the census for the inputs below LA level. However, both the LA benchmarks considered are dependent on 2011 Census. Like the ratio change method, the benchmarking method does not have a natural mechanism for producing measures of uncertainty to accompany the estimates. It is reliant on having a consistent annual supply of SPDs since the method uses the administrative data counts directly below LA level. Although the PR was deemed suitable to use for the ratio change method (since it does not use the PR counts directly at MSOA and LSOA level), it would not be suitable to use for the benchmarking method because of the notable net overcount seen in the PR (and its successor the PDS). While SPDs are not suitable for population estimation on their own, they are the best representation of population available using administrative data alone.

The benchmarking method has been applied before at the ONS. In June 2023, we released a calibrated LSOA population difference dataset. In this work, SPD counts at LSOA level were benchmarked to DPM LA estimates. The DPM that included Census 2021 as an input was used because the aim of that work was to produce the best estimates available. Here, we use the DPM without Census 2021 as an input to enable a fair comparison to the ratio change methods previously proposed, which are rolled forward from 2011 and do not use Census 2021. We also extend the SOA ABPE work by considering MSOA as well as LSOA level outputs for the benchmarking method (see our DPM, improvements to data sources and methodology for local authorities methodology for more information).

The different variations in data inputs of the benchmarking method are as follows:

- method 2a benchmarks the 2021 SPD counts at MSOA and LSOA level to the 2021 rolled forward MYE at LA level from the current population system
- method 2b is the same as method 2a, except the DPM (without Census 2021) has been adopted as the LA benchmark
- method 2c uses the 2021 SPD MSOA and LSOA counts without LA level benchmarks; this method is a baseline scenario and is expected to have the lowest performance of all the ratio change and benchmarking options

3 . Results

In this section, we compare both the ratio change and benchmarking methods in terms of bias based on the different scenarios outlined in Section 2: Baseline methods for Small Area Population Estimates (SAPE) in transformed population system. Results are shown for both Middle layer Super Output Area (MSOA), and Lower layer Super Output Area (LSOA), 2021 population totals separately. All results shown in this section use the 2021 local authority (LA) boundaries, but at MSOA and LSOA level, the 2011 boundaries were used because of data availability.

Middle layer Super Output Area level

Table 1 compares the different variations of the ratio change method (1a to 1d) in terms of bias at MSOA level for 2021, as well as also providing comparisons with the different variations of the benchmarking method (2a to 2c). The focus of this analysis is on absolute relative bias (ARB), which is defined as the absolute value of:

100*[(estimate-true value)/true value]

The estimate is a given MSOA estimate in 2021 and true value is the corresponding Census 2021 estimate.

Apart from Census 2021 (true values), the rest of the data used in this analysis for 2011 to 2021 relates to the mid-point of the year. Given this, we define the estimates of bias as relating to mid-2021 but acknowledge that we are not accounting for all the bias between Census 2021 and mid-2021.

Holding everything else constant (method type and MSOA data used), including a Dynamic Population Model (DPM) LA benchmark as opposed to a rolled forward MYE LA benchmark, leads to a notable improvement in median bias for both the ratio change and benchmarking methods at MSOA level. For the ratio change method, using Statistical Population Datasets (SPDs) for reflecting population change between 2016 and 2021 leads to an improvement over only using SPDs for the final year of the decade (holding everything else constant).

As we would expect, the MSOA SPD counts without LA benchmarks (method 2c) had the highest bias on average at MSOA level. The median bias for this method across all 7,201 MSOAs was 3.24%. The median bias for the other methods varies between 2.05% (method 2b) and 2.61% (method 1a).

Despite the simpler nature of the benchmarking method, the bias achieved is similar to that of the ratio change method. In fact, the best performing method in Table 1 is the benchmarking method with a DPM LA benchmark (method 2b) which has a median bias of 2.05%. However, it should be noted that the ratio change method does not have the option of using SPDs for the 2011 to 2015 period. It is possible that, if SPDs were available for the entire 2011 to -2021 period, that the ratio change method may outperform benchmarking method 2b.

For the best comparisons in Table 1, only one of the MSOA data and LA benchmark should be varied for a given method type (ratio change or benchmarking). For example, comparing methods 1a and 1b where only the LA benchmark is varied is more appropriate than comparing methods 1a and 1d, where both the MSOA data and LA benchmark are varied.

Table 1: MSOA totals, comparison of methods in terms of Absolute Relative Bias (ARB), mid 2021, England and Wales (2011 MSOA boundaries)

Method	Method type	MSOA data	LA benchmark	Minimum bias	Lower quartile	Median	Upper quartile	Maximum bias
1a	**Ratio change**	PR for 2011-2020 SPD for 2020-2021	Rolled forward MYE	0	1.23	2.61	4.90	100.07
1b	**Ratio change**	PR for 2011-2020 SPD for 2020-2021	DPM without Census 2021	0	0.99	2.18	4.10	93.57
1c	**Ratio change**	PR for 2011-2016 SPD for 2016-2021	Rolled forward MYE	0	1.18	2.49	4.65	79.33
1d	**Ratio change**	PR for 2011-2016 SPD for 2016-2021	DPM without Census 2021	0	0.95	2.08	3.88	61.16
2a	**Benchmarking**	SPD counts	Rolled forward MYE	0	1.19	2.58	4.77	97.04
2b	**Benchmarking**	SPD counts	DPM without Census 2021	0	0.92	2.05	3.93	69.07
2c	**Raw SPD counts**	SPD counts	NA	0	1.75	3.24	5.08	57.79

Source: Office for National Statistics, NHS

A comparison of the median bias at MSOA level for each of the methods is shown for all English and Welsh regions in Table 2. The median bias at MSOA level for 2021 is notably higher in London than for other regions, for all the methods except for the MSOA SPD counts (method 2c) where Wales had a notably higher bias than London.

Table 2: MSOA totals, comparison of median MSOA Absolute Relative Bias (ARB) at regional level, mid 2021, England and Wales (2011 MSOA boundaries)

Method	Median ARB Ratio change (Method 1a)	Median ARB Ratio change (Method 1b)	Median ARB Ratio change (Method 1c)	Median ARB Ratio change (Method 1d)	Median ARB Benchmarking (Method 2a)	Median ARB Benchmarking (Method 2b)	Median ARB Raw SPD counts (Method 2c)
East Midlands	2.53	1.87	2.62	1.82	2.61	1.98	3.24
East of England	2.61	2.01	2.46	1.92	2.53	1.91	3.56
London	4.45	3.80	3.98	3.86	4.12	3.50	3.75
North East	2.18	1.97	2.00	1.90	2.26	1.75	2.09
North West	2.24	1.85	2.19	1.90	2.23	1.88	2.50
South East	2.75	2.21	2.53	1.91	2.40	1.79	3.53
South West	2.07	1.91	2.04	1.79	2.10	1.85	3.35
Wales	2.45	1.89	2.34	1.86	2.80	2.17	5.14
West Midlands	2.84	2.10	2.54	2.04	2.68	2.09	2.84
Yorkshire and The Humber	2.25	2.26	2.27	2.06	2.60	1.92	2.88

Source: Office for National Statistics, NHS

The conclusions for 2021 in Table 1 and Table 2 are broadly in line with the previous findings from our SAPE 2011 evaluation report. In this work, the ratio change method was run from 2001 to 2011, starting from Census 2001 and using PR or child benefit data for the ratio change and rolled forward MYE as the LA benchmark. The ratio change estimates for 2011 at MSOA level were compared with the 2011 Census in terms of bias.

Although direct comparisons between the 2011 MSOA bias estimates and those shown for the ratio change method in Tables 1 and 2 are challenging because of differences in the administrative data sources and or LA benchmarks used, we note that the median ARB) for MSOAs in 2011 was 2.5%. This is broadly in line with the results in Table 1. Similar conclusions were also found at LSOA level.

In our evaluation report, London also had notably higher median (ARB) than all other regions in 2011 at MSOA level. The same was also true at LSOA level.

Lower layer Super Output Area level

Table 3 compares the different variations of the ratio change method (1a to 1d) in terms of bias at the LSOA level for 2021, as well as also providing comparisons with the different variations of the benchmarking method (2a to 2c). The focus of this analysis is on (ARB), with Census 2021 LSOA estimates treated as the true values.

As at MSOA level, including a DPM LA benchmark, as opposed to a rolled forward MYE LA benchmark, leads to a notable improvement in median bias for both the ratio change and benchmarking methods at LSOA level (holding everything else constant). For the ratio change method, using SPDs for dictating population change between 2016 and 2021 does not lead to an improvement over only using SPDs for the final year of the decade (holding everything else constant). This is perhaps unexpected, but the combined results in Tables 1 and 3 suggest that the SPD offers more benefit for the ratio change at MSOA level than it does at LSOA level. However, a full assessment would require SPD data available for the entire decade from 2011 to 2021.

The SPD counts at LSOA level (method 2c) perform comparatively poorly in terms of bias compared with other methods. However, they perform better than the ratio change method estimates that use rolled forward MYE LA benchmarks (methods 1a and 1c) which is perhaps unexpected. The median bias for the raw SPD counts was 3.56% compared with 3.71% and 3.89% for methods 1a and 1c, respectively. The median bias for the other methods vary between 2.66% (method 2b) and 3.52% (method 1d).

Despite the simpler nature of the benchmarking method, the bias achieved is superior to that of the ratio change method. In fact, the best performing method in Table 3 is the benchmarking method with a DPM LA benchmark (method 2b), which has a median bias of 2.66% (far lower than all other methods). The benefits of the benchmarking method over the ratio change method appear more substantial at LSOA level than at MSOA level. However, it should be noted that the ratio change method does not have the option of using SPDs for the 2011 to 2015 period.

Table 3: LSOA totals, comparison of methods in terms of Absolute Relative Bias (ARB), mid 2021, England and Wales (2011 LSOA boundaries)

Method	Method type	LSOA data	LA benchmark	Minimum bias	Lower quartile	Median	Upper quartile	Maximum bias
1a	**Ratio change**	PR for 2011-2020, SPD for 2020-2021	Rolled forward MYE	0	1.72	3.71	6.76	386.59
1b	**Ratio change**	PR for 2011-2020, SPD for 2020-2021	DPM without Census 2021	0	1.54	3.31	6.14	317.50
1c	**Ratio change**	PR for 2011-2016, SPD for 2016-2021	Rolled forward MYE	0	1.80	3.89	7.16	214.63
1d	**Ratio change**	PR for 2011-2016, SPD for 2016-2021	DPM without Census 2021	0	1.61	3.52	6.61	221.46
2a	**Benchmarking**	SPD counts	Rolled forward MYE	0	1.46	3.13	5.72	206.42
2b	**Benchmarking**	SPD counts	DPM without Census 2021	0	1.22	2.66	5.00	187.80
2c	**Raw SPD counts**	SPD counts	NA	0	1.81	3.56	5.81	178.96

Source: Office for National Statistics, NHS

A comparison of the median bias at LSOA level for each of the methods is shown for all English and Welsh regions in Table 4. Similar to MSOA level, the median bias at LSOA level for 2021 is notably higher in London than for other regions, for all of the methods except for the LSOA SPD counts (method 2c), where Wales had a higher bias than London.

Table 4: LSOA totals, comparison of median LSOA Absolute Relative Bias (ARB) at regional level, mid 2021, England and Wales (2011 LSOA boundaries)

Method	Median ARB Ratio change (Method 1a)	Median ARB Ratio change (Method 1b)	Median ARB Ratio change (Method 1c)	Median ARB Ratio change (Method 1d)	Median ARB Benchmarking (Method 2a)	Median ARB Benchmarking (Method 2b)	Median ARB Raw SPD counts (Method 2c)
East Midlands	3.70	3.10	4.16	3.47	3.29	2.55	3.52
East of England	3.53	3.00	3.70	3.18	3.01	2.52	3.81
London	5.80	5.27	5.83	5.26	4.99	4.44	4.59
North East	2.90	2.93	3.18	3.16	2.71	2.33	2.52
North West	3.52	3.15	3.67	3.47	2.82	2.63	2.92
South East	3.71	3.19	3.77	3.23	2.88	2.32	3.66
South West	3.22	2.99	3.37	3.12	2.59	2.34	3.46
Wales	3.50	3.18	3.64	3.36	3.30	2.92	5.30
West Midlands	3.74	3.15	3.81	3.39	3.17	2.61	3.28
Yorkshire and The Humber	3.24	3.04	3.65	3.40	3.00	2.45	3.19

Source: Office for National Statistics, NHS

4 . Geospatial Methods

Alternative model-based methods for small-area population estimates

We are considering alternative approaches to the ratio change and benchmarking methods to produce small area population estimates (SAPE). One option is the use of geospatial analysis techniques to produce high-resolution population and demographic data, described in the article, Advances in mapping population and demographic characteristics at small-area level.. Geospatial approaches may offer a possible avenue for providing population estimates for small areas using frequent and high-quality data sources that can inform population distributions at low-level geographies. The primary intention of geospatial methods is to complement other population estimation systems such as census, registration systems or from model-based approaches such as the Dynamic Population Model (DPM), rather than replace these. Conceptually, integrating and drawing strength from several data sources, including aggregated local authority (LA) estimates from methods such as the DPM, in conjunction with data at disaggregated levels of geography, including administrative based estimates (from various administrative sources) and other geospatial information, aims to provide high-quality SAPE, that each data source on their own, may not satisfactorily achieve.

Organisations such as WorldPop and LandScan have developed a portfolio of work providing low-level "gridded" estimates, making use of relationships between geospatial information and population estimates at these fine-grained levels of geography. Gridded methods involve dividing geographic areas into grids of equal size and can range in size. Typically, grid squares can be of any size between 100 meters squared (m2) to 1 kilometres squared (km2). These grids are not nested within traditional Office for National Statistics (ONS) administrative boundaries, but estimates may potentially be produced at small area statistical or administrative boundaries by a process of fitting to aggregate gridded estimates within a common higher-level administrative boundary. Alternatively, some methods could be used to model estimates at a desired geographical boundary (such as, OAs, LSOAs, and MSOAs). Geospatial approaches have typically been used for lower and middle-income countries that don't have recent and reliable census or registration data. However, the Disaggregating population data article from Yue Qiu and others suggests they are now being increasingly used where high-quality geocoded data are available. Within the remit of the transformation of population and migration statistics, we consider these novel methods for providing SAPE in England and Wales.

Methods

The essence of geospatial approaches is to make use of data that has a "location component", capturing various geographical, demographic, and socio-economic information that relates to the population living there. The approach capitalises on more recent advances in computing power, remote sensing and other new forms of satellite imagery and digital data about the Earth's surface, described by an article in the 'International Journal of Applied Earth Observation and Geoinformation'; examples include night-time lighting, land use, building footprints plus many others. Mapping to gridded geographies rather than the typical boundaries used in England and Wales, particularly Output Areas (OAs), Lower layer Super Output Areas (LSOAs), or Middle layer Super Output Areas (MSOAs) where boundaries are based on evenly distributed population estimates that may lack spatial and distribution detail, could better suit the properties of geospatial data. Geospatial data may allow greater discrimination between areas and potentially stronger models of their relationship to population distributions. See the Disaggregating population data article for more detail.

Geospatial methods for SAPE generally fall into two categories: "bottom-up and top-down" approaches: both methods share common goals to produce population estimates for small areas or uniform, high spatial-resolution grids. However the bottom-up methods rely on accurate counts of populations within small, defined sample areas at a target level of geography and by aggregating these high-resolution predictions, population totals can be produced for different levels of geography (see the article, Spatially disaggregated population estimates in the absence of national population and housing census data, from N.A. Wardrop and others for more information). Bottom-up methods could be suitable for different situations, for example, when countries lack complete or any census data, or when high-quality geocoded data (see the article from Yue Qiu and others is available. However, both situations require high-quality survey and ancillary data to produce estimates of sufficient quality at granular breakdowns (see the Review of geospatial methods for population estimation from Kristine Nilsen and others). Population totals for all areas are predicted by modelling relationships between the sampled populations and geospatial information (available for all areas at the target level of geography). The statistical models, the most common bottom-up method being Bayesian hierarchical modelling, are built to predict population numbers in unsampled areas, together with uncertainty intervals based on covariate and spatial relationships.

Alternatively, top-down disaggregation approaches take known population totals at a higher level of geography and spatially redistribute these to lower levels of geography. Geospatial covariates at the lower level are used to inform these spatial redistributions. In Yue Qiu and others's 2022 paper, methods for disaggregating population data provide comprehensive comparisons of dasymetric methods. Several disaggregation methods exist, though the most commonly used method is dasymetric mapping, which refers to subdividing the source zone into smaller areas that can reflect spatial changes in the population, based on ancillary data. This is widely adopted by organisations such as WorldPop. WorldPop have extensive experience of disaggregating higher-level estimates to low-level grids across the world, including England and Wales, as well as Scotland and Northern Ireland, as shown in the University of Southampton's Projection-disaggregated gridded population datasets for 189 countries.

Both top-down and bottom-up applications have their strengths and limitations. When considering the application of the bottom-up methods to ONS contexts, these methods are contingent on having a high-quality survey which is capable of producing accurate estimates at the target level of geography. ONS would require a high-quality survey or several surveys combined to produce reliable and high-quality low-level population totals. If sufficient survey data are available, the quality of estimates from the modelling procedure can be assessed at the level that modelling took place. On the other hand, top-down disaggregation methods are dependent on high-quality aggregated population totals, which could be available at national, regional and LA level, for example, our Admin-based population estimates from the DPM. Top-down methods are also dependent, to some degree, on the modelled relationships between population and covariate information being sufficiently similar at aggregated and disaggregated levels of geography. However, current applications of top-down methods, similar to ratio-change and benchmarking approaches, only allow for measuring bias. Uncertainty measures could be derived at the modelled level of geography, but currently we would not have model-derived uncertainty measures for target levels of geography (MSOAs and LSOAs). However, there are currently workstreams within the ONS exploring methods to derive uncertainty measures in administrative datasets and population estimates, that may have potential to be used in other contexts. The disaggregation methods are also more readily available for the ONS to consider. Therefore, we give more consideration in this paper to the top-down dasymetric mapping methods adopted by WorldPop, as described in the Disaggregating Census Data for Population Mapping article from Forrest R. Stevens and others.

Conceptual application

WorldPop have developed a top-down semi-automatic dasymetric model to produce estimates for several applications, including producing estimates for low-level 100m grids, as well as to statistical or administrative boundaries. These methods use random forest regression models to establish relationships between geospatial covariates and population estimates at a level of geography where the population estimates are of a high quality. For ONS contexts, this is likely to be DPM estimates at LA level. These models are used to make predictions for lower levels of geography where only geospatial covariate information is available. As mentioned, this could be down to low-level grids (for instance, 100m) but also to statistical or administrative boundaries, for example, to OA level. Inherently, the dasymetric methods WorldPop employ calibrate the lower-level predictions to the known high-quality population totals. Disaggregated estimates can then be built-up to higher level geographies. For instance, we could disaggregate to gridded or OA-level geographies, and then aggregate these estimates up to LSOA or MSOA boundaries, as well as other geographies, such as National Parks.

The dasymetric mapping method requires covariate information to be derived at both aggregated and disaggregated geographies. The geospatial covariates used in the model could include geospatial variables, such as:

- night-time lights data
- elevation and slope of the terrain
- land cover and use
- building stocks that count households and communal establishments grouped by either number of rooms or property type

Other geospatial covariates are more descriptive when modelling to very fine-grained levels of geography (for instance, 100m grid squares). Examples include points of interest (POI) data on infrastructure, such as access to amenities (hospitals, schools, post offices, GP surgeries, streetlights, and so on). We will explore the different dasymetric mapping methods to produce estimates at grid-level geographies, as described by WorldPop's report (PDF, 376KB), which may be pertinent with the current population transformation work in the ONS, where several low-level estimates would be in demand. One benefit of these top-down methods is that current implementations are available from programming languages, such as R. One potential challenge to investigate is the differences in geographical scale between model training and small-area prediction to explore conceptual drift relationships between population and geospatial information from LA to small area level.

The High resolution population distribution maps article from Andrea E. Gaughan and others shows that gridded data could be used to provide outputs for several types of geographical hierarchies: OA-LSOA-MSOA, parishes-wards-local authority districts-counties, electoral wards, and many others. Part of the research for producing gridded estimates would be to explore how flexible these approaches can be to produce high-quality population estimates at various geographical hierarchies. Furthermore, the frequency of outputs would be contingent on the frequency of the supply of population data and geospatial information. For instance, some of the geospatial information we could make use of, including measures of night-time lights, climate, terrain, air quality, as well as POI data, can be sourced at monthly intervals. If alternative data sources, particularly the LA-level population estimates, can also be sourced at more frequent intervals than annual outputs, then there is the possibility that disaggregated outputs could be produced at a similar frequency. The primary distinction between geospatial methods and the baseline methods (ratio change and benchmarking) is that geospatial methods hold the potential of providing more frequent estimates for several low-level geographies. Furthermore, the flexibility of using georeferenced data at high spatial resolution may allow for the estimation of different population definitions, not just the residential population but others such as the total population. For instance, geospatial data has been used by Population 24/7 to produce estimates of the day-time mobile population,and we are already working with it, as described in our Working paper.

Using fine-grained geospatial information about England and Wales will not only be useful for estimating population, but also characteristics of the population. For instance, previous WorldPop applications for England and Wales have produced population by age and sex breakdowns. There is also work in other countries, particularly work from the World Bank in collaboration with WorldPop, who have produced estimates of Income, wealth (PDF, 1.2MB) and poverty (PDF, 1.3MB). Different types of gridded information will be important for producing various characteristic outputs, where varied data sources will be uniquely indicative of specific characteristics of the population.

Future research

There is scope to explore several areas of research with the geospatial approaches. One work stream is to consider the types of geospatial covariates we use in the models. For example, POI and administrative housing stock datasets have been shown in the research article from Noée Szarka and Filip Biljecki to produce promising results to predict ages of resident populations in Singapore. POI datasets may reflect the characteristics of the people they serve, and this level of granularity could hold potentially provide unique information about distributions at small area level (see the POI commentary article from Achilleas Psyllidis and others for more information).

Another consideration is the use of mobile phone data that could offer additional uses for the geospatial modelling approaches. We have produced experimental daytime population estimates, shown in our Working paper, derived from anonymised and aggregated mobile phone crowd movement data. Two exploratory empirical examples of using mobile phone data outside of the ONS include the creation of a Dynamic population distribution dataset developed in Finland using grid cells and advanced dasymetric interpolation methods, and population estimation from Mobile network traffic metadata (PDF, 2.9MB) to infer Population densities and mobility in Italy. These may provide further insights into the utility of mobile phone data for population estimation theory and practice.

Advances in the acquisition and processing of satellite images (see Patrick Lehnert's and others' Proxying economic activity with daytime satellite imagery article) have led to progressive analysis of aspects relevant to the population, including commercial and residential global building stock (see Thomas Esch's and others' World Settlement Footprint 3D article). With a combination of advances in geospatial satellite imagery (such as thermal imagery HotSat1) and methods that harness POI data, there could be potentially wide applications for future research.

Further work on modelling approaches could involve directly comparing bottom-up and top-down methods to understand how these different methods perform for ONS contexts (see Kristine Nilsen's and others Review of geospatial methods for population estimation article). Exploration of the use of Census 2021 outputs as a source of "truth" could be used for evaluating bias of different geospatial modelling options and provide validation of statistical outputs. Importantly, any assumptions and dependencies should be explicitly set out and tested where possible. Alongside this, understanding what different modelling techniques could be most beneficial (this could include trialling various machine learning algorithms) will be important to ensure we are adopting the most "fit for purpose" model. Future work could also include capturing uncertainties from input data. Reducing uncertainties and improving predictions in model-based estimates will require a co-ordinated effort to incorporate the latest population enumerations, frequent high resolution satellite imagery and other geospatial data into population models. Currently, a major challenge in current top-down methods' selection is choosing the most appropriate disaggregation method and ancillary data and research that facilitates this will be invaluable. Lastly, all methods are dependent on high-quality input data and the recency of the selected geospatial information to be used for potential production of population estimates.

5 . Conclusion

This paper has discussed baseline methods (ratio change and benchmarking) to produce Small Area Population Estimates (SAPE), with a particular focus on Middle layer Super Output Area (MSOA) and Lower layer Super Output Area (LSOA) population totals. Bias estimates were used as an assessment criterion for methods performance. At MSOA level in 2021, the baseline methods produced median Absolute Relative Bias (ARB) of between 2.05% and 2.61% depending on the method of interest (excluding outputs that use only Statistical Population Dataset (SPD) MSOA counts). As expected, at LSOA level in 2021, the baseline methods produced higher median ARB than at MSOA level, ranging from 2.66% to 3.89% depending on the method of interest (excluding outputs that use only SPD LSOA counts). Using a Dynamic Population Model (DPM) (without Census 2021) local authority (LA) benchmark offered benefit over using a rolled forward mid-year estimate MYE LA benchmark at both MSOA and LSOA level.

The availability of administrative data over the 2011 to 2021 period caused notable challenges for this work, most notably for the ratio change method. The Patient Register (PR) being discontinued in 2020 and SPDs only being available for 2016 to 2021 meant that there was no consistent administrative data stock across the decade. Although the impact of including SPDs in the latter half of the decade in the ratio change method was considered, conclusions on their impact should be treated with caution. More robust results would have been reached if SPDs were available across the entire decade.

The bias seen in the ratio change method for 2021 (rolled forward from 2011) are broadly in line with our Previous report subject to the caveats discussed in Section 3: Results. We have previously concluded in our Methodology review and research update that, for the great majority of SAPE areas, the methodology used produces useable estimations of those geographies. However, at the extremes of the distribution, there are a minority of areas where the methodology does not appear to have performed so well. The results shown in Section 3: Results also support this statement.

There were some regional variations in the quality of estimates. The bias of the estimates in 2021 within London areas were notably higher (on average) than in other regions at both MSOA and LSOA level for all the baseline methods considered (excluding SPD counts). All the baseline methods considered at MSOA and LSOA level saw some notable outliers which led to substantial bias. A potential area of future research for geospatial methods is to investigate improvements into outlier bias and improving accuracy of SAPE.

Both baseline methods are conceptually simple and the ratio change method has been used for intercensal estimates at MSOA and LSOA level for many years and is familiar to users. Both the methods are reliant on a census with the ratio change method requiring census counts as a starting point at the level of estimation (MSOA and LSOA level) and the benchmarking method currently using census information at the LA level. The ratio change method requires an annual supply of suitable administrative data at MSOA and LSOA level for reflecting population change from a census starting point, whereas the benchmarking method uses the administrative data (SPDs) directly at MSOA and LSOA level. Both methods benefit from consistency in the administrative data over time. It should also be noted that none of the baseline methods are developed to produce uncertainty measures for the estimates of interest.

Acknowledging the shortcomings of these baseline methods, we are exploring alternative options that may offer an improvement for SAPE. Similar to the ratio-change and benchmarking methods, top-down geospatial methods constrain low-level population estimates to appropriate LA benchmarks. We have outlined the potential for using granular geospatial information that can inform the population distributions at fine-grained levels of geography, whether that be grid cells or small area statistical or administrative boundaries. The essence of top-down approaches is to integrate multiple data sources, using geospatial information at low levels in conjunction with existing population estimates and datasets, breaking down the high-quality aggregated population estimates to lower levels that cannot be achieved when using these data sources in isolation. There is also the potential for some of the methods, namely the bottom-up modelling approaches, to provide measures of uncertainty for small area outputs that neither the top-down geospatial methods, nor ratio-change and benchmarking methods, can currently achieve.

6 . Cite this methodology

Office for National Statistics (ONS), released 7 December 2023, ONS website, methodology, Small Area Population Estimates in the transformed population estimation system: methods development