Bias and Variance quality standards for 2023 recommendation

1. Background

The Beyond 2011[1] programme was established following the 2011 Census to assess the different possible approaches to producing population and housing statistics in the future through a programme of consultation and research. This work culminated in March 2014[2] with a recommendation from the National Statistician to the government for the future provision of population statistics. The National Statistician recommended a predominantly online census in 2021 supplemented by the further use of administrative and survey data. The Beyond 2011 program came to an end in early 2015. The National Statistician will be giving a further recommendation on the future of population statistics in 2023.

As part of the Beyond 2011 program, a set of quality standards[3] were developed. The standards were designed to reflect the quality achieved within the current system based on census and mid-year population estimates. Separate quality standards were developed for both population estimates (national and local authority totals and split by age and sex) and population characteristics e.g. estimates of economic activity and ethnicity at different geographical levels. The Beyond 2011 quality standards focused on Census 2011 and Mid-Year Estimates (MYE) for 2002-2010 with a particular emphasis on variance and less so bias. To inform the evidence and research for the 2023 Recommendation, here we focus on population estimates and provide updated quality standards for bias and variance based on Census 2021 and MYE for 2012-2020. This initial report focuses on national and Local Authority (LA) totals only with age and sex estimates being dealt with in future work. Quality standards for population characteristics may be considered in future work.

2. Bias and Variance standards for population estimates

Table 1 outlines the theoretical quality standards of future population estimates. These reflect the quality that is currently achieved through the population estimates system based on census and MYE.

P1 refers to achieving the same quality as the most recent census (2021) in every year and is referred to as the maximum quality. P2 is equivalent to what we have in the current population system i.e. maximum quality in the census year but declining quality as you move further away from the census with minimum quality seen in 2020. P3 refers to the average quality achieved in the current system so for the most recent decade this would be equivalent to the quality achieved at the mid-point (2016).

Table 1: Theoretical quality standards for population estimates

	Quality standard
P1 (Maximum)	Maximum quality achieved in the current system (ie. what is achieved in a census year), every year
P2 (Variable)	The quality achieved in the current system (with a peak and then a decline during the decade)
P3 (Average)	The average quality achieved in the current system, every year

We consider P1-P3 for both variance and bias separately. For more details on P1, P2 and P3 for variance, see Sections 2.1-2.3 and for bias see Table 4 in Section 4.

2.1 Maximum level of **precision** (P1)

Census 2021 had two quality aims[4] for precision (variance) which were the same as those for Census 2011:

- All 331 Local Authority (LA) population estimates have a 95% relative confidence interval width of +/-3% or better;
- The national population estimate (England and Wales combined) has a 95% relative confidence interval width of +/- 0.2% or better.

The confidence intervals produced for Census 2021[5] differ to those produced for Census 2011 in the sense that they are asymmetric rather than symmetric around the point estimate. This means we have two percentages (relative confidence interval widths) associated with each point estimate for evaluating whether the above quality aims are met, one relating to the lower bound and the other the upper bound. For example, for the LA quality aim stated above we aim not to exceed the 3% target on either side of the point estimate.

At LA level the following was achieved for Census 2021:

- **97.9% of the 331 LA population estimates had a 95% confidence interval lower bound of -3% or better relative to the point estimate.**
- **96.1% of the 331 LA population estimates had a 95% confidence interval upper bound of 3% or better relative to the point estimate.**

Most if not all LAs met the 3% quality aim (on both sides of the point estimate). We adopt what was achieved in Census 2021 (highlighted in bold above) as an indicative quantitative measure for P1, to reflect the maximum precision in the current system at LA level. If an option proposed in the future meets this standard, then it can produce population estimates that are of similar quality to those produced from the 2021 Census at LA level.

It is also important to consider the precision of those LAs whose Census 2021 population estimate was not within the standard highlighted in bold above, so an additional standard is needed.

At LA level the following was achieved for Census 2021:

- All LA population estimates (except three LAs) had a 95% confidence interval lower bound of -3.7% or better relative to the point estimate. The greatest relative difference between point estimate and lower bound was -5.20%.
- All LA population estimates (except eight LAs) had a 95% confidence interval upper bound of 3.6% or better relative to the point estimate. The greatest relative difference between upper bound and point estimate was 13.3% and the second highest was 6.99%.

The 2021 Census also had the following quality aim for the precision of the national population estimate:

- The national population estimate (England and Wales combined) has a 95% relative confidence interval width of ±0.2% or better.

This aim was exceeded in the 2021 Census:

- **The 2021 Census national population estimate (England and Wales combined) exceeded the quality aim of 0.2% (on both sides of the point estimate). The national population estimate had a 95% confidence interval lower bound of -0.16% and a 95% confidence interval upper bound of 0.07% relative to the point estimate (59,597,542 people).**

We adopt what was achieved in Census 2021 (highlighted in bold above) as a quantitative measure for P1, to reflect the maximum precision in the current system at the national level.

It is also of interest to look at England and Wales separately. No interval for England was produced for Census 2021 but it is expected to be very similar to that given above for England and Wales combined.

The population estimate for Wales had a 95% confidence interval lower bound of -0.45% and a 95% confidence interval upper bound of 0.37% relative to the point estimate (3,107,494 people). As expected**,** the precision achieved for the population estimate of Wales is less than that for England since we are estimating for a smaller population.

2.2 Variable (minimum) precision (P2)

The precision (variance) of population estimates varies in the inter-censal period. The minimum precision of population estimates within the current system can be taken as the precision of the MYE at the end of the most recent decade (2020) before the 2021 Census.

In January 2022, confidence intervals were published for all LAs in England and Wales for the 2011-2020 MYE[6].

Of particular interest for the minimum precision is the confidence intervals for MYE 2020. Excluding LAs with very small populations (Isles of Scilly and City of London), defined as outliers, there were 334 LAs in England and Wales in 2020. Results indicate that:

- **90% of LA population estimates had a 95% relative confidence interval width of +/-7.9% or better (excluding outliers)**

- **97% of LA population estimates had a 95% relative confidence interval width of +/-11.9% or better (excluding outliers)**

- **All LA population estimates had a 95% relative confidence interval width of +/-24.8% or better (excluding outliers).**

We adopt what was obtained for MYE 2020 (highlighted in bold above) as quantitative measures for P2 to reflect the minimum precision in the current system at LA level.

Note: The confidence intervals produced for MYE 2011-2020 are different to those produced for the 2002-2010 period. For the 2002-2010 period, traditional symmetric confidence intervals were calculated meaning that confidence intervals were symmetric around the point estimate. This allowed relative confidence interval widths to be calculated in the traditional way (width of interval divided by estimate). However, for the 2011-2020 period, asymmetric confidence intervals were calculated meaning that MYEs do not lie in the centre of the confidence intervals and in some cases, estimates lie outside the interval. This makes calculating relative confidence interval widths challenging. For the analysis in this paper, we take the lower and upper limits as they are and assume that the MYE lie in the middle of the interval. Given this, comparing results from this report to those previously calculated for the Beyond 2011 reports is not appropriate.

2.3 Average precision at the mid-point of the decade (P3)

The average precision (variance) of population estimates within the current system refers to the precision of the MYE at the mid-point of the decade (2016) before the 2021 Census.

Of particular interest for the average precision is the confidence intervals for MYE 2016 (mid-point of decade). Excluding LAs with very small populations (Isles of Scilly and City of London), defined as outliers, there were 346 LAs in England and Wales in 2016. Results indicate that:

- **97% of LA population estimates had a 95% relative confidence interval width of +/-6.6% or better (excluding outliers)**

- **All LA population estimates had a 95% relative confidence interval width of +/-10.9% or better (excluding outliers)**

We adopt what was achieved for MYE 2016 (highlighted in bold above) as quantitative measures for P3 to reflect the average precision in the current system at LA level.

2.4 Bias

As well as considering the variability of estimates for assessing quality, it is also important to consider the bias (accuracy) of the estimates.

2.4.1 Bias in the census

The level of bias adjustment for the 2011 Census population estimate of England and Wales was 0.5%. At the national level 0.5% bias was considered unacceptable and hence was adjusted for in Census 2011. Similarly, for Census 2021 at national level (England and Wales), an objective was to have bias less than 0.5%[4]. When evaluating options for future population estimates a bias of 0.5% or greater at the national level would not be acceptable for meeting the established quality standards of the Census.

2.4.2 Bias in non-census years

Consideration of the bias in the MYE is important as they form a key part of the current population estimates system. This will allow us to provide an indication as to the highest bias that is currently seen in population estimates.

It is difficult to evaluate the bias in MYE for years where there is no census. For this analysis we derive estimates for the census based MYE in 2012-2020 (at both national and LA level) which are referred to as proxy census based MYE under the assumption that there is a linear trend between the census based MYE in 2011 and 2021. The bias is then calculated as the difference in percentage terms (absolute relative bias) between the MYE for a given year from 2012-2020 and the corresponding true value (proxy census based MYE). Whilst this is clearly not ideal, it allows us to provide some indication of bias in the MYE across the decade. The hypothesis is that the bias will increase as estimates move further away from the census.

National level

Table 4 relates the Absolute Relative Bias (ARB) estimates at national level (England and Wales combined) to the criteria P1, P2 and P3 described in Table 1. For P1 we use the ARB as estimated in previous censuses which was 0.5%. For P2, the highest ARB was 0.88% in 2018 with the bias generally increasing as we move further away from the 2011 Census. This provides an indication of the maximum bias seen in the current population system at national level. For P3, the average ARB across the decade is taken to be the value in 2016 (0.82%).

Table 4: Absolute Relative Bias (ARB) in terms of P1-P3 criteria, national level

Criteria	Description	Absolute Relative Bias (ARB)
P1 (Maximum)	Lowest ARB achieved in the current system (what is achieved in a census year, every year)	0.5% in Census year
P2 (Variable)	The highest ARB achieved in the current system	0.88% in 2018 represents the highest bias (minimum quality)
P3 (Average)	The average ARB achieved in the current system, every year, using the value in 2016.	0.82% in 2016

LA level

At the start of the decade (2011) there were 346 LAs in England and Wales (excluding Isles of Scilly and City of London). By the end of the decade (2021) there were 329 due to LAs being merged. For this analysis on bias, we use the 329 definition across the whole time period 2011-2021. Table 5 summarises the bias at LA level for the years 2012-2020.

In 2012, 23.4% of LAs (77 out of 329) had a bias greater than 0.5%. In the middle of the decade (2016), 73.6% of LAs (242 out of 329) had a bias greater than 0.5%. However, at the end of the decade (2020), 81.5% of LAs (268 out of 329) had a bias greater than 0.5%.

In 2012, only 4.0% of LAs (13 out of 329) had a bias greater than 1%. In the middle of the decade (2016), 50.8% of LAs (167 out of 329) had a bias greater than 1%. However, at the end of the decade (2020), 65.0% of LAs (214 out of 329) had a bias greater than 1%.

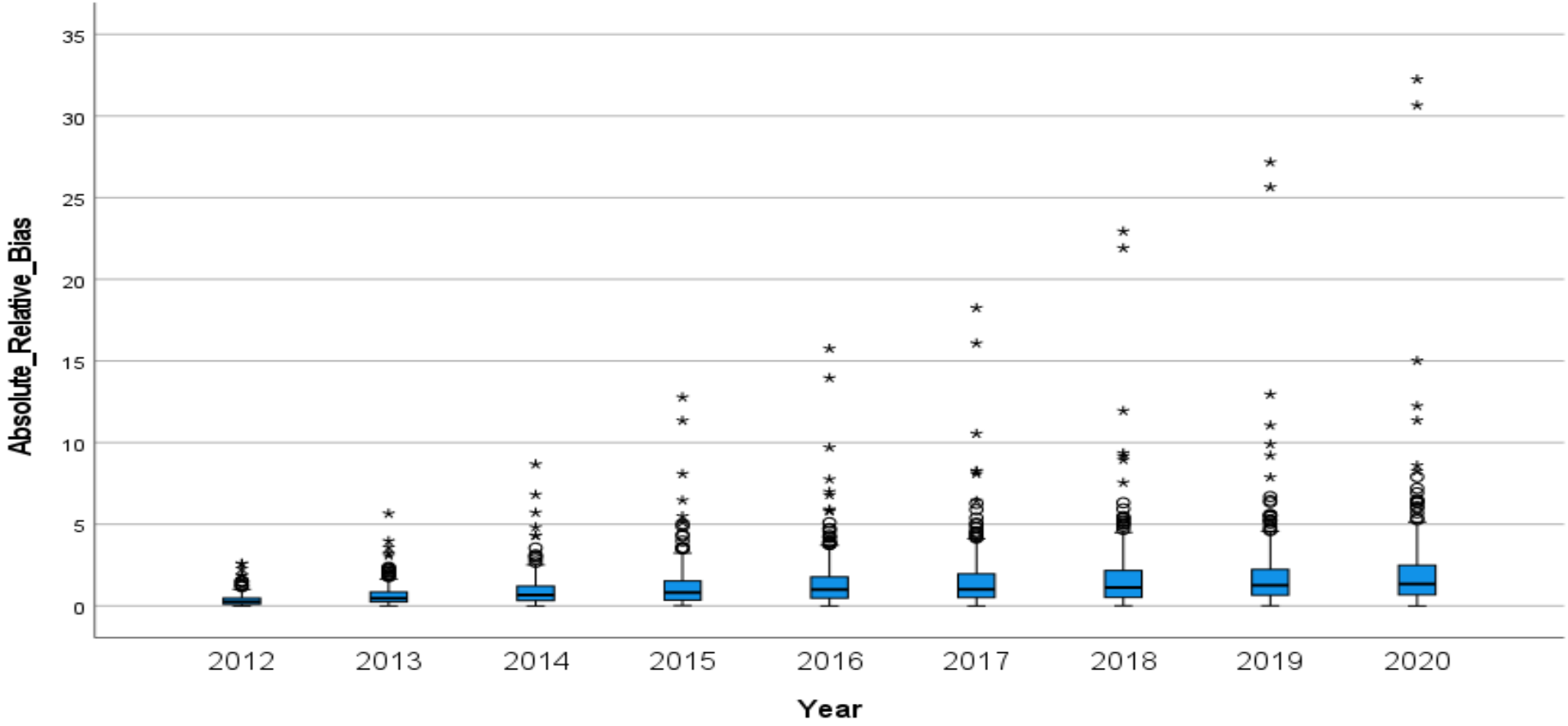
Table 5: Number of LAs with Absolute Relative Bias (ARB) greater than specified thresholds, LA level, 2012-2020

Year	No. of LAs with ARB > 0.4%	No. of LAs with ARB > 0.5%	No. of LAs with ARB > 1%	No. of LAs with ARB > 3%	No. of LAs with ARB > 5%	No. of LAs with ARB > 10%
2012	101	77	13	0	0	0
2013	189	155	67	5	1	0
2014	223	205	110	9	3	0
2015	241	228	143	23	7	2
2016	260	242	167	34	10	2
2017	262	249	162	38	11	3
2018	263	251	185	43	16	3
2019	272	265	202	57	18	4
2020	280	268	214	61	22	5

Note: Excludes Isles of Scilly and City of London. The number of LAs with negative relative bias across the decade ranged between 39.2% in 2018 and 51.4% in 2012.

Figure 1 provides a graphical representation of the ARB at LA level in the form of a box and whisker plot. **The median ARB across LAs varies between 0.26% in 2012 to 1.34% in 2020. In 2016 the median ARB across LAs was 1%.**

Figure 1: Box and whisker plot for the ARB at LA level for 2012-2020



Note: Excludes Isles of Scilly and City of London

It is important to bear in mind that the 2021 Census took place in the middle of a third COVID-19 lockdown. In this bias analysis we assume a linear trend between census based MYE 2011 and census based MYE 2021 but if the 2021 estimate is affected by where people were situated in the third lockdown then this will affect the trajectory of the population in the intervening years. It could be that biases seen for 2012-2020 for certain LAs are overestimated (or underestimated). In some cases, it might be hard to unpick whether the high biases are to do with problems in the MYE or whether they are impacted by the timing of the 2021 Census and the subsequent effect that has on deriving proxy census based MYE for 2012-2020.

3 Discussion

This report has focused on updating quality standards from the Beyond 2011 Program using more recent estimates from the current system for population estimation (Census 2021 and MYE 2012-2020).

The results presented for Census 2021 confidence intervals relate to the maximum precision achieved in the current population estimates system. For quality requirement P1 to be satisfied a future population estimates system would need to achieve the same precision as Census 2021 in every year. At the national level (England and Wales combined), Census 2021 exceeded its quality target for relative confidence interval widths of +/- 0.2% or better. At the LA level, Census 2021 met its quality target for relative confidence interval widths of +/- 3% or better in most if not all LAs.

For quality requirement P2 to be satisfied, a future population estimates system would need to achieve the same precision as Census 2021 once a decade with diminishing quality across the next decade. For P2, the MYE uncertainty at the end of the decade (2020) defines the minimum precision achieved in the current population estimates system. Uncertainty measures for MYE in 2020 were only available for LA totals and are presented in this report. For future iterations of this report, MYE uncertainty for 2020 will ideally also be available at national level and ideally also split by age and sex for both national and LA level.

For quality requirement P3 to be satisfied, a future population estimates system would need to achieve the average precision from the current system every year. For P3, the MYE uncertainty at the mid-point of the decade (2016) defines the average precision achieved in the current population estimates system. Uncertainty measures for MYE in 2016 were only available for LA totals and are presented in this report. For future iterations of this report, MYE uncertainty for 2016 will ideally also be available at national level and ideally also split by age and sex for both national and LA level.

The report also focused on bias in the current population estimates system at both national and LA level. It is challenging to evaluate the bias in the MYE for years where there is no census. In this report we assume a linear trend between census based MYE 2011 and census based MYE 2021 to estimate proxy census based MYE for 2012-2020. Whilst this is clearly not an ideal assumption, it allows us to approximate the level of bias present in the current population estimates system as we move further away from the census. At national level, the level of bias ranged between 0.5% in census years and 0.88% for years towards the end of the decade.

This report has not focused on quality standards for the Dynamic Population Model (DPM) or administrative based estimates such as those based on Statistical Population Datasets (SPDs). However, this report has provided quality standards for bias and variance in the current population estimates system which will be beneficial for future comparisons to other approaches.

3.3 Further work

There are several areas for further work:

- The MYE uncertainty estimates for 2012-2020 are only currently available for LA totals[6]. However, MYE uncertainty for national totals (England and Wales separately and combined) are currently being calculated by ONS colleagues and will ideally be available for the next iteration of this paper. These estimates will feed into quality standards P2 and P3 at national level.
- The MYE uncertainty split by age and sex for 2012-2020 is not currently available at national or LA level. These would be more time consuming for ONS colleagues to calculate but they can be done depending on work priorities. Currently age by sex estimates of uncertainty at national and LA level are only available from Census 2021.
- The bias estimates calculated so far have only been for national and LA totals. This could be extended to include age and sex in future work.
- The bias estimates at LA level have not currently been related to criteria P1-P3 but there is intention to do this in the future.
- The lowest level geography considered in this report was LA level. Further work could consider MSOA and LSOA level that were discussed in the Beyond 2011 reports.
- This report has focused on population totals only and not split by age and sex or population characteristics/attributes. There is scope to build on the work of the Beyond 2011 reports for population characteristics such as ethnicity at small area level.

References

[1]Office for National Statistics (2012). Beyond 2011 Public Consultation on User Requirements Report. Beyond 2011 Public Consultation on User Requirements - Report

[2]Office for National Statistics (2014). Beyond 2011: Final Options Report (04) Beyond 2011: Final options report (481.7 Kb PDF)

[3]Office for National Statistics (2013). Beyond 2011: Options Report 2 Beyond 2011 Options Report 2

[4]Office for National Statistics (2020). Design for Census 2021. Design for Census 2021 - Office for National Statistics (ons.gov.uk)

[5]Office for National Statistics (2022). Measures showing the quality of Census 2021 estimates. latest version of the local authority comparison tool

[6]Office for National Statistics (2022). Measures of Uncertainty in ONS local authority mid-year population estimates – all confidence intervals. Measures of uncertainty in ONS local authority mid-year population estimates - all confidence intervals - Office for National Statistics

[7]Office for National Statistics (2021). Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland - Office for National Statistics (ons.gov.uk)

[8]Office for National Statistics (2013). Difference between rolled forward mid 2011 estimates and 2011 census MYE:

differencebetweenrolledforwardmid2011estimatesand2011censusmyes